Innovations in Monetary Policy

Doctoral Thesis in Economics
by
Jack Meaning

School of Economics
University of Kent

This thesis is dedicated to my wife, Alice, without who's support, patience and love nothing I have done would have been possible. It is also for my newborn son, Patrick Maynard, who even in the short time I have known him has inspired me in unparalleled ways.

Acknowledgements

Many people require thanks for their part in this thesis. First and foremost is Jagjit, who first convinced me to take on a Ph.D. and who has guided me through from start to finish with an unerring mix of pushing me to stretch myself whilst acting as a safety net when things were tough. I am eternally in your debt. Also Casey Otsu for his comments at crucial moments and his calm and clear explanation of points I should often have been quicker to pick up. Of course, I am grateful to the School of Economics at Kent and all the staff there. Kent was my academic home for almost a decade and I owe much to all of you for the knowledge you have imparted to me over the years.

James Warren and Alex Waters have been my partners in macroeconomic crime for years, and the work in this thesis is heavily influenced by conversations had with them over endless cups of tea. I have also had the good fortune to find employment in a number of nurturing and exciting places. I must thank Phillip Turner, Feng Zhu and my other colleagues at the Bank for International Settlements for their worthwhile discussion on high level policy. For the last too years I have been working at the National Institute for Economic and Social Research. This has been an invaluable experience and has been both professionally enriching and incredibly enjoyable. Special thanks must go to Simon Kirby, Ian Hurst, Iana Liadze, Oriol Carreras and Rebecca Piggott who have been the best colleagues I could have hoped for.

There is also a myriad of others who have advised and aided me, be they conference attendees, referees or just interesting academics I have met. There is not space to thank you all, but I hope to continue to learn from you in the years to come.

I am indebted to the ESRC for their financial contribution to my Ph.D., which funded my young family through 3 years of study and made the whole thing possible.

And last, but by no means least, my family; Mum, Dad and Joe. You have all been incredibly supportive, not just throughout this thesis, but for as long as I can remember. It is from you all that I have garnered the drive and curiosity which has lead me to this point. Thank you. To my extended family, and family by law, you have all played your part and I hope now it is over I can spend more time enjoying your company as it deserves.

Declarations

This thesis builds on work I have carried out over the last four and a half years. Unless otherwise stated below, it is entirely my own work, as are all errors.

A number of the pieces of work which have lead to this thesis have been published or presented in one guise or another. Empirical work in the third chapter has been published to varying degrees in Caglar et al (2010) and Meaning and Zhu (2011, 2012). It was also presented at the Bank of England, the Bank for International Settlements and HM Treasury as well as to the School of Economics at the University of Kent.

The model in the fourth chapter has been published as a Bank for International Settlements Working Paper and is currently under a revise and resubmit for the Journal of Macroeconomics. It has been presented widely at the Bank for International Settlements, the Bank of England, the Money, Macro and Finance Annual Conference and the Royal Economics Society Annual Conference as well as a range of academic departments including, the University of Kent, Brunel University and Tor Vergata in Rome.

I give heartfelt thanks to all those who have allowed this work to be disseminated and given comments along the way.

Acknowledgement of coauthors

This thesis draws on a number of pieces of work which I have had the good fortune to coauthor with various colleagues. It is therefore appropriate that I both acknowledge their input and clarify my own contribution.

Chapters 1 and 2

The first and second chapters, concerning the theoretical basis of quantitative easing and the particulars of recent central bank policy, are entirely my own work, although they almost certainly draw on the multitude of discussions I have had with friends and colleagues on these topics in the last 4 or 5 years and for which I am grateful.

Chapter 3

The third chapter, which uses a range of econometric techniques to asses quantitative easing programmes in the US and UK, draws on work carried out alongside

Evren Caglar, Jagjit Chadha, Alex Waters and James Warren at the University of Kent (Caglar et al (2010)) and Feng Zhu whilst I was working at the Bank for International Settlements (Meaning and Zhu (2011, 2012)).

I conducted event studies on asset purchase programmes whilst working with both sets of coauthors, and in each case the work was my own, though the final decisions on specific details (one versus two day windows etc.) were taken jointly. The event study presented in this thesis is a completely new piece of work, although given the limited scope for variation, the results are very similar and the choices I made were informed by my work with those colleagues.

The econometric estimations based on the methodology of D'Amico and King (2010) seen in this chapter were originally carried out as part of Meaning and Zhu (2011). I had help from Jackub Demski of the BIS who sourced much of the data, but the data manipulation and actual estimations are all my own work. The results were shared with my coauthor who I thank for his insightful comments and suggestions on interpretation.

The time series analysis also formed part of a paper coauthored with Feng Zhu (Meaning and Zhu (2012)). Again, data was sourced by Jackub Demski, but I actually collated the average maturity series included in the regression myself from back issues of US Treasury publications. The regressions themselves were the result of continuous dialogue with Feng concerning appropriate controls and specifications, but I carried out the physical act of running the estimations myself, before taking the results back for further discussion with Feng and other colleagues at the BIS.

The text of the third chapter is entirely my own.

Chapter 4

The fourth chapter is an openly coauthored paper with Jagjit Chadha and Luisa Corrado. Jagjit and Luisa had previously worked on a variant of the Goodfriend and McCallum model on which this paper is based. They therefore provided me the code from their earlier variant, which formed the base of the model presented in this chapter.

My contributions can be seen in the extension of both the model and the Matlab code to include open market operations and the additional mechanisms discussed in the paper. I also carried out a lot of the data work which informs the early part of the paper and the subsequent calibration of the model. The balance sheet depiction of asset purchases was not my work, but rather that of Professor Chadha. However, I did derive the welfare loss function myself and calculate the relative gain criteria which feature in Section 7. Lastly I calibrated and ran the shocks presented, including the crisis shock. The interpretation of the results was a joint effort by all authors, as was the drafting of the text.

References

[1] Caglar, E., Chadha, J.S., Meaning, J., Warren, J. and Waters, A (2011). 'Central Bank Balance Sheet Policies: Three views from the DSGE literature', Chapter in 'Interest Rates, Prices and Liquidity' Cambridge University Press, 2011

[2] D'Amico, S and T King (2010): "Flow and stock effects of large-scale treasury purchases", Federal Reserve Board Finance and Economics Discussion Series, 2010–52

[3] Meaning, J and Zhu, F (2011). 'The Impact of Recent Central Bank Asset Purchase Programmes', BIS Quarterly Review, December 2011

[4] Meaning, J and Zhu, F (2012). 'The Impact of Federal Reserve Asset Purchase Programmes: Another Twist', BIS Quarterly Review, March 2012

Introduction

The last 10 years represent a fascinating time for monetary policy. It also coincides with my own period of study on the subject. Starting my undergraduate degree in 2005, I am incredibly fortunate to have been witness to the asking of some fundamental questions around monetary policy.

First the financial crisis of 2008 forced central banks to evaluate their relationship with the financial system. Their role as lender of last resort and the conflict it may present when banks get too big to fail was one of the big themes of the early part of the crisis. It has lead to a wider discussion on the risks and linkages within the financial sector and resulted in many central banks being given an expanded remit to consider macroprudential policy. Whilst not the focus of this thesis, these themes continue to interest me and figure in my research.

Since 2009, faced with the zero lower bound constraining their favoured policy instrument, policymakers have had to (re)find alternative instruments to spur demand stimulate their economies. This has lead them to expand their balance sheets and under take purchases of assets from the private sector on an unprecedented scale. This has raised questions about how such policies might affect the economy, and a range of empirical work aimed at helping policymakers to calibrate their interventions.

Concurrent with these policy developments, the 2008/09 crisis has served to spark a deep and fundamental debate about how economists model and view the world. The limitations of the benchmark New Keynesian model that prevailed prior to the crisis has resulted in a renewed enthusiasm for models which try to break some of the key assumptions that lay behind it.

These issues have been the focus of my own research throughout my doctorate. In the pursuit of a deeper understanding of them I have employed a range of modelling and econometric techniques, data analysis and discussion with both academics and practical policymakers. The resultant thesis draws on a number of articles and papers I have worked on in the last four years, distilling them down to what I consider some of the more interesting concepts and results. Like all good academic work, I hope it raises as many questions as it answers and represents at least a small step in the ongoing process that is the refinement of monetary policy. I know that the process of writing it has at any rate greatly advanced my own understanding

The first two chapters are intended as an introduction to these innovations, first from a theoretical and then a practical standpoint. They provide context for the work that follows and should be viewed as a springboard for further reading and research. The third chapter turns to the data and tries to take an econometric approach to understanding the largest innovation in monetary policy in recent years - large scale asset purchases. It forms part of a growing literature which finds a significant role for asset purchases by central banks in the manipulation of interest rates. The fourth chapter then investigates some of these features in a dynamic stochastic general equilibrium setting.

Through this combination of data analysis, econometrics, macroeconomic modelling and theoretical understanding, I hope that I have presented a relatively complete view of how monetary policy has evolved along many dimensions. I am sure that it will continue to do so, and I look forward to both following it, and trying to make what contribution I can in the future.

Theory for unconventional central bankers: a framework for asset purchases

Jack Meaning

Abstract

Since 2008, central banks in the major advanced economies have purchased unprecedented quantities of assets from the private sector, expanding their balance sheets and dramatically changing the characteristics of their portfolios. While there is a growing consensus in the empirical literature that these policies have succeed in lowering interest rates, agreement on a theoretical explanation as to how and why this might be the case is far less unanimous. The favoured framework of the last decade or so, the New Keynesian model, has very little to say about asset purchases, other than they may act as a signal for the path of interest rates, or a commitment to a given path of policy. This paper looks to highlight why this result is a product of the model's assumptions, before presenting a framework to think about alternative ways in which asset purchases can be motivated in a theoretical context.

1 Introduction

The widespread use of asset purchase programmes by central banks in recent years has posed a challenge for economic theory. The consensus model of the last decade or so struggles to provide any direct mechanism by which such purchases can have any traction on interest rates, or the wider economy. And yet, there is a growing body of empirical work which supports the hypothesis that large-scale asset purchases have succeeded in reducing interest rates in advanced economies, and may have stimulated demand.

This paper is intended as an overview of some of the issues at the heart of the theoretical debate around asset purchases. To that end, it develops a simple, stylised model within which the discussion can be framed. This model is a generalised version of the canonical three equation log-linear New Keynesian model, which can be returned as a special case under a specific parameterisation.

The paper will first describe the baseline New Keynesian model, its underlying assumptions and its implications for monetary policy. In this version of the world, monetary policy is conducted solely by the setting of the short-term nominal interest rate. There exists no role for asset purchases, other than as a signal for, or commitment to a future path of that policy rate. However, this result is predicated on a number of assumptions, most notably a representative agent with rational expectations who operates in unconstrained financial markets. The rest of the paper seeks to investigate how relaxing these assumptions may give rise to theoretical channels through which asset purchases can stimulate the economy.

First, money is explicitly introduced to the baseline model via a simple, stylised money demand function. This is an uncontroversial innovation as a money demand curve of some form is often assumed to lie implicitly behind the three equation model. Its omission is normally justified by the lack of additional insight it provides, as money moves passively to clear the money market at the central bank's target rate. In the context of asset purchases though, the inclusion of money and the acceptance that the supply of an asset can have implications for its price is an important initial step.

We also include two additional assets which are priced in as general a form as possible, but which under certain parameterisations allow changes in supply to have a significant impact on interest rates. This framework enables us to deviate from the baseline New Keynesian model in a number of interesting ways. We begin by analysing a world in which money demand is less than perfectly elastic and Wallace neutrality fails to hold for money. Under these conditions the central bank can alter interest rates in the economy by manipulating the money supply. In fact, in order to clear the money market, for a given money demand the central bank must vary the money supply in order to elicit an interest rate response. However, it makes no difference in this parameterisation of the model which assets are bought and sold in order to manipulate the money supply, or even whether any assets are bought at all.

The paper then turns to discussing under what conditions this may not be the case and changes in supply may have consequences for the prices of assets other than money. Drawing on the growing literature in this area a number of candidate explanations can be found, but crucially all must in some way lead to a less than perfectly elastic demand curve for assets and a failure of Wallace neutrality. The implications of such a result are then discussed in the context of the stylised model of this paper. It is shown that, should such supply effects exist, however motivated, then the central bank will be able to stabilise the economy more effectively as asset purchases will not only act to lower interest rates on wider assets indirectly through their influence on the short-term nominal interest rate, but also directly by changing the supply of the asset itself. The magnitude of this channel depends crucially on the degree of inelasticity in the asset's demand curve and the extent to which assets are imperfect substitutes for one another, both parameters which can be easily manipulated in our stylised framework.

The paper then considers the impact of the lower bound on nominal interest rates within these various versions of the world. With the short-term nominal interest rate unable to fall any lower, the New Keynesian model has only one policy prescription which is to lower agents expectations of future policy rates by committing to hold rates "lower for longer". Given forward-looking and rational agents this acts to lower long rates and stimulate activity today. With the money supply expanded such that the money market is clearing at zero, money demand is effectively satiated and so further expansion cannot reduce interest rates further. Thus at the lower bound the second parameterisation of the model collapses to the first, with the same implication for policy. Importantly though, if supply effects exist then even with the short-term nominal interest rate constrained asset purchases can directly lower interest rates by reducing the premium associated with supply.

The paper then discusses some potential mechanisms through which asset purchases can stimulate the economy other than through their direct influence on interest rates including stimulating bank lending, reducing collateral constraints, relaxing constraints on the fiscal authority and a monetarist view of monetary expansion.

2 The New Keynesian model

The New Keynesian models which came to dominate the economic debate in central banks around the turn of the century were the culmination of a number of significant developments in economic thought.

The first crucial element is that of rational expectations, which developed through the 1960s and 1970s thanks to work by Muth (1961) and Lucas (1972, 1976). This way of describing the formation of expectations states that agents within the model are aware of its structure and base their expectations on what that structure implies for the future paths of variables, given all the information available to them at that time.

From this grew the real business cycle (RBC) agenda pioneered by Kydland and Prescott (1982). RBC models brought rational expectations to the earlier neoclassical growth models. They consisted of forward-looking representative agents who optimised an objective function consistent with rational expectations. These optimisation problems were built up from microfoundations to ensure they were as theoretically consistent as possible.

Fundamental to the core RBC framework was the assumption of perfectly competitive markets in which all agents were price-takers. This left no significant role for monetary policy as perfectly flexible prices could adjust instantly to clear markets. This ran counter to the traditional Keynesian view that nominal rigidities caused price inertia which monetary policy could exploit, Keynes (1936). However, Keynesian models failed to motivate the existence of such rigidities through micro-founded optimising behaviour. What is more, rather than forward-looking rational expectations, traditional Keynesian models most commonly employed adaptive expectations, where agents views of the future depend on the past. This left them vulnerable to the Lucas Critique, Lucas (1970).

In response to these criticisms, and concurrent with the development of the RBC literature, a body of work emerged looking to introduce nominal frictions into models with rational expectations. First steps involved the use of long-term nominal contracts to generate a slow adjustment of prices, for instance Fisher (1977), Phelps and Taylor (1977) and Taylor (1979). Then, in 1983, Guillermo Calvo introduced staggered pricing contracts under which any given firm could only change the price they charged in the next period with a probability of less than one, Calvo (1983). This meant that even assuming rational and optimising agents, there was a degree of price inertia at the aggregate level following a shock as not all firms would be able to adjust instantly.

New Keynesian models as they are generally understood today began to emerge in the early 1990s. The seminal volumes by Mankiw and Romer (1991) brought together numerous contributors to lay out the foundations of the New Keynesian framework. They sought to reconcile the monopolisticly competitive firms and price and wage rigidities of traditional Keynesian models with the microfoundations and rational optimising behaviour of the RBC literature. This amalgamation of features has lead to New Keynesianism sometimes being referred to as the New Neoclassical Synthesis.

A number of expositions of the basic model can be found, for example Woodford (2003), Clarida, Gali and Gertler (1999) or Kerr and King (1996), but in its most basic form the model can be represented by three loglinearised equations; an expectational IS curve, a New Keynesian Phillips Curve and a rule for determining the nominal interest rate.

The IS curve is derived from the optimising decision of the household, based on an Euler relationship. It relates output today inversely to the real rate of interest and can be written[1]

$$y_t = E_t y_{t+1} - \sigma \left[i_t - E_t \pi_{t+1} \right] + g_t \tag{1}$$

where y is real output, i is the one period nominal interest rate and π is the one-period inflation rate. g represents an exogenous demand shock.

The supply side of the economy is represented by the NK Phillips curve. This is a log-linear approximation to the first order condition for the firm seeking to set prices optimally under Calvo-style staggered price contracts. κ is know as the Calvo parameter and represents the speed of adjustment of prices. u is an exogenous supply shock.

$$\pi_t = \kappa y_t + \beta E_t \pi_{t+1} + u_t \tag{2}$$

Lastly the model is closed out by a rule which determines the short-term nominal interest rate. Following Taylor (1993, 1999), this usually takes the form of a feedback rule on deviations in output and inflation from target

$$i_t = \rho i_{t-1} + \phi^y y_t + \phi^\pi \pi_t \tag{3}$$

A key finding of Taylor was that, in order to bring stability to the system of equations, the parameter ϕ^π must be greater than one. The reasoning for this can be intuitively seen, as only when $\phi^\pi > 1$ does the central bank definitely move the nominal interest rate sufficiently that the real interest rate moves in the

[1] All variables are written in deviations from steady state.

stabilising direction.

2.1 Implication of the NK model for monetary policy

In this representation of the world, monetary policy operates purely through manipulating the short-term nominal interest rate which passes through the economy via the traditional Keynesian interest rate channel. If the central bank cuts the nominal interest rate, this causes a change in the real interest rate as prices are slow to adjust. This in turn boosts current demand through the IS curve (1) which raises output today and subsequently generates inflationary pressure through the New Keynesian Phillips Curve (2).

Importantly, the forward looking nature of the New Keynesian model means that it is not just the contemporary interest rate which matters, but rather the expectation of the future path of rates as well. To see this, we can iterate equation 1 forward one period at a time and, under the rational expectations assumption of the model, we can substitute the resulting equation for y_{t+1} back in for $E_t y_{t+1}$. Equation 1 can therefore be written

$$y_t = \sum_{t=0}^{t=\infty} \left[\sigma_t \left[i_t - E_t \pi_{t+1}\right] + g_t\right] \tag{4}$$

and output becomes a function of the full expected future path of interest rates.

In the three equation representation of the New Keynesian model presented here, quantities have no direct impact on the economy. However, this result is a construct of the assumptions which underpin the model, most clearly the existence of complete and perfectly functioning financial markets and a representative agent.

Perfect and complete financial markets give rise to the no-arbitrage assumption which states that any deviation of interest rates from the expectations hypothesis would open up an arbitrage opportunity. Rational and unconstrained investors would then instantly exploit this opportunity so that it was immediately exhausted and prices returned to their previous level. Interest rates on all assets are thus determined purely by their expected payoffs in different states of the world, and not by their supply, Cox, Ingersoll and Ross (1981).

Even if the expectations hypothesis did not hold, it can still be argued that central bank asset purchases are irrelevant for determining interest rates in the New Keynesian framework. Assuming an infinitely lived and rational representative agent, a sale which moves assets between the private and public sectors should be ineffective. This is because the asset the private sector agent has given up is now held by the public sector, of which the agent is the ultimate owner. Therefore they are still exposed to the same payoffs as before in all states of the world, albeit that some will now be realised through taxation or reduced government spending as opposed to the direct impact of holding the asset. In order to counteract this increased exposure through the State, the private sector agent will reduce their demand for the asset sold, hedging and exactly offsetting the change in supply and leaving the price unchanged, figure 1. This irrelevance proposition, akin in spirit to Ricardian Equivalence, is known as Wallace Neutrality. It was first posited for standard open market operations by Wallace (1981), but has since been famously reiterated in the New Keynesian framework and applied to asset purchases more generally by Eggertson and Woodford (2003).

These core assumptions are vital to the policy prescription of the New Keynesian model but, as we will discuss in later sections, they are not without criticism or alternative. It should be noted here that Wallace himself saw his irrelevance proposition as an unrealistic base from which to build a more realistic representation of the world, and even Eggertson and Woodford (2003) admits it is unlikely to hold for money away from the lower bound.

3 Extending the baseline model

3.1 Introducing money

The three equation model abstracts completely from money. However, it may be a more accurate representation of the New Keynesian position to explicitly include a fourth equation which many consider to be implicitly behind the simplified three equation variant; a money demand function. Demand for money can be motivated a number of ways, but the most common is that money provides a transaction service that no other asset can replicate. This may be because it is the most widely accepted and efficient medium of exchange, or because it is the only asset which can be used to settle accounts with the State. Money is also an extremely liquid and secure store of value. These store of value, liquidity and transaction services provide a utility to the holder for which they are willing to pay.

Theoretically a number of models have looked to incorporate this feature into the New Keynesian paradigm and provide a role for money. Some look to include real money balances into the household's utility function, Walsh (2003).

These money in utility models are effectively a short-cut, as money does not in of itself provide utility, rather it is a proxy for the underlying service that money provides. A similar approach has been the cash-in-advance models, such as the one presented in the fourth chapter of this thesis and others in which agents are constrained in their consumption and spending decisions today by the cash balances they have brought over from the period before. In this way money is a determining factor in their ability to consume and so affects welfare.

Our first innovation is therefore to introduce a standard money demand function of the form

$$m_t = y_t - \alpha i_t \tag{5}$$

The inclusion of output relates to the increased demand for transaction purposes and the second term captures the (opportunity) cost of (holding) obtaining money.

This innovation ties the quantity of money to its price, the short-term nominal interest rate, in an important way. In order for the money market to clear for a given money demand, any change in the interest rate must be met with a change in the money supply, and vice versa. The central bank can approach this problem in one of two ways. First, it could set the interest rate it is targeting and supply money perfectly elastically at that rate. Money demand then pins down the quantity of money, figure 2. This is known as fixed-rate full-allotment provision and is how most practitioners characterised policy in Great Moderation. After all, the majority of central banks have used the interest rate as their principle policy instrument since at least the early 90s. The quantity of money in this case simply follows the interest rate passively to clear the market. As discussed in Ireland (2005), this particular endogenous view of money adds little to the basic New Keynesian framework and the benign nature of the quantity of money in this setup has lead many economists to consider the New Keynesian model's abstraction from money as acceptable, even if they do not believe that strict Wallace Neutrality holds.[2]

The alternative manner in which the central bank could set policy is by determining the quantity of money and leaving the market to clear, resulting in i^*. This is in essence substituting equation 3 into 5 and then re-writing them as

[2] This assertion is rejected by McCallum (2012), who demonstrates that the absence of money in the three equation New Keynesian model is only valid if the transaction cost which money serves to eliminate is additively separable in consumption and money, an assumption for which there is little to no evidence.

$$m_t = \alpha \left[\rho i_{t-1} + \phi^y y_t + \phi^\pi \pi_t\right] \tag{6}$$

and

$$i_t = \frac{y_t - m_t}{\alpha} \tag{7}$$

These two approaches are broadly analogous to each other, equating on paper to the same levels of interest and quantities of money. However, in practice the latter requires the monetary authority to know the money demand function, a non-trivial assumption given the recognised instability of money demand through time. In a seminal paper, Poole (1970) demonstrated that, although in the case of absolute certainty the two ways of setting policy are equivalent, in an economy dominated by shocks to the IS curve (g) then it is better to set policy via the quantity of money, whilst when financial or money market shocks dominate, welfare is enhanced by setting the price of money.[3] His ultimate conclusion was that in an economy subject to both real and financial shocks, the best course of action for policymakers was to use a combination of both interest rate and quantity setting, a point picked up on and elaborated by Chadha, Corrado and Holly (2014).

Our innovation here is far from controversial. Policymakers have accepted it, either explicitly or implicitly for decades and it lies at the heart of practical policy action in central banks around the world.

3.2 Adding imperfectly substitutable assets

Having introduced one asset, money, we now extend the model further with the addition of two new assets; A and B. We keep the nature of these assets as general as possible, imposing only that they are priced in such a way that the rates of interest they pay can be written as

$$i_t^A = \sum_{t=0}^{MA} i_t + \beta^A \left[S_t^{PA} + \delta S_t^{PB}\right] \tag{8}$$

$$i_t^B = \sum_{t=0}^{MB} i_t + \beta^B \left[S_t^{PA} + \delta S_t^{PB}\right] \tag{9}$$

where MA and MB are the respective maturities of assets A and B, S_t^{PA} is

[3] Poole's analysis was conducted in the IS-LM framework, so by financial market shocks, it is meant any exogenouns disturbance to the LM curve.

the publicly available supply of A and S_t^{PB} the publicly available supply of B. It is important to note here that what is meant by publicly available supply is net of assets held by the central bank, ie. total supply less that held by the monetary authority. Thus S_t^{PA} and S_t^{PB} can be written

$$S_t^{PA} = S_t^A - S_t^{CA} \tag{10}$$

$$S_t^{PB} = S_t^B - S_t^{CB} \tag{11}$$

where S_t^A and S_t^B are the total supplies of asset A and asset B and S_t^{CA} and $S_t^{CB}$are the quantities of each asset held by the central bank.

β^A and β^B can be thought of as the interest elasticity of demand, or the slope of the demand curve for each asset. δ is the extent to which assets A and B are viewed as substitutes for each other.[4]

The economic interpretation of these equations is that assets A and B are priced in accordance with the expectations hypothesis, which is captured in the first term in equations 8 and 9, but depending on the parameterisation have the potential for an additional premium which is a function of the supply of the asset and its substitutes. This is similar in spirit to Harrison (2012) who develops a model in which long rates are subject to a premium which is a function of the relative supplies of long and short bonds. The approach here, though more stylised aims to be more general in its assertion of the characteristics which distinguish m, A and B. It may, as in the case of Harrison (2012) be term, or equally, it may be liquidity, risk or any other feature which can vary between assets and for which investors may have a preference.

In order to incorporate these two assets into our wider framework, we must also adjust the IS curve so that it is dependent not only on the risk-free rate, but rather a weighted-average of the three interest rates which now prevail in the economy; i_t, i_t^A and i_t^B.

$$y_t = E_t y_{t+1} - \sigma \left[\lambda^i i_t + \lambda^A i_t^A + \lambda^B i_t^B - E_t \pi_{t+1}\right] + g_t \tag{12}$$

where

$$\lambda^i + \lambda^A + \lambda^B = 1 \tag{13}$$

This is again similar to the result derived by Harrison (2012) from an

[4]We will elaborate more on this in the next section.

optimising household with preferences for liquidity. He shows that the IS curve becomes a function of the weighted average of both short and long rates. Conceptually this idea harks back to Tobin (1969) and even Keynes who believed the monetary tightness or otherwise of an economy could not be accurately represented by a single interest rate. It is an acceptance that agents store value for the future via a range of assets, each of which pay a different return and which may or may not be perfectly connected through the expectations hypothesis. For our analysis it is assumed λ^i, λ^A and λ^B are exogenously determined and equal to $\frac{1}{3}$ in each instance.

3.3 Open market operations: a link between money and other assets

To close out the model, we need to describe how the central bank's control of the quantity of money relates to other assets in the economy. To do this, we impose that the central bank must match the liabilities on its balance sheet (money) with an equal quantity of assets, which it has purchased from the private sector. Although the mechanism is stylised, this is how the money supply is controlled in practice, both for traditional open market operations and also the large-scale asset purchases experienced in the last decade.

In our three asset world this means that S_t^{CA} and S_t^{CB} must equal the quantity of money

$$m_t = S_t^{CA} + S_t^{CB} \tag{14}$$

and in order to increase the money supply, the central bank must increase either S_t^{CA} or S_t^{CB}. All else equal, this would reduce S_t^{PA} or S_t^{PB} while simultaneously increasing m_t.

3.4 The final model

Our model can therefore be characterised by 10 core equations. First is our modified IS curve

$$y_t = E_t y_{t+1} - \sigma \left[\lambda^i i_t + \lambda^A i_t^A + \lambda^B i_t^B - E_t \pi_{t+1}\right] + g_t \tag{15}$$

Then the standard NK Phillips curve

$$\pi_t = \kappa y_t + \beta E_t \pi_{t+1} + u_t \tag{16}$$

Then we have our monetary policy rule which determines the quantity of money, and the market clearing interest rate, given that quantity of money

$$m_t = \alpha \left[\rho i_{t-1} + \phi^y y_t + \phi^\pi \pi_t \right] \tag{17}$$

$$i_t = \frac{y_t - m_t}{\alpha} \tag{18}$$

For the analysis that follows we assume that assets A and B are one period assets which differ from each other and money due to idiosyncratic features. Therefore the rates of interest they play can be written

$$i_t^A = i_t + \beta^A \left[S_t^{PA} + \delta S_t^{PB} \right] \tag{19}$$

$$i_t^B = i_t + \beta^B \left[S_t^{PA} + \delta S_t^{PB} \right] \tag{20}$$

And lastly we assume that the total supply of assets A and B is exogenously fixed so that movements in the publicly available supply of each asset are determined solely by movements in the quantity of each held by the central bank

$$S_t^{PA} = S_t^A - S_t^{CA} \tag{21}$$

$$S_t^{PB} = S_t^B - S_t^{CB} \tag{22}$$

In the simulation examples that follow we make a further simplification that the central bank controls the money supply exclusively through varying its holdings of asset A, and so

$$S_t^{CA} = m_t \tag{23}$$

and S_t^{CB} is exogenously fixed, meaning that S_t^{PB} is also exogenously fixed.

Although it appears quite simple, the general nature of this model means that it is extremely flexible and as such is a useful backdrop on which to frame and clarify one's thoughts on central bank asset purchases and how they might work. This is what we will do in the next section.

4 Central bank asset purchases as a policy tool

The first thing to note is that the standard New Keynesian model is a special case of our more general model. By setting $\alpha = \infty$ money demand becomes perfectly interest rate elastic and the quantity of money supplied does not affect the price of money, i.[5] Similarly, if assets A and B are perfect substitutes, $\delta = 1$, and there are no frictions which cause prices to deviate from the pure expectations hypothesis then demand for both assets is also perfectly elastic and so $\beta^A = \beta^B = 0$. This returns the New Keynesian result that the only manner by which monetary policy can influence the economy is through a change in i, or the future path of i, which lowers the return on assets A and B through the term structure.

4.1 A world where money matters

Let us now consider a world in which the money demand curve is less than perfectly elastic and Wallace Neutrality fails to hold for money. In our framework, this equates to the case where $-\infty < \alpha < 0$. This means we have a downward sloping demand curve for money of the type inferred by diminishing returns to transaction and liquidity services discussed earlier. For now we continue to assume Wallace Neutrality holds for wider financial markets and demand for A and B remains perfectly elastic so $\beta^A = \beta^B = 0$. Under this parameterisation an open market operation that exchanges money for asset A expands the money supply. For a given money demand, this expansion lowers i_t as agents require further inducement to hold the increased cash balances given the diminishing marginal utility of money. This feeds through the term structure and lowers $\sum_{t=0}^{MA} i_t$ which in turn lowers i_t^A. It also lowers i_t^B by the same mechanism. In fact, if $MB < MA$ then the interest rate on asset B will fall by more than that on asset A, despite B not having been involved in the operation at all.

This transmission mechanism can be seen in figure 4 which plots the impulse responses of the model following a positive shock to the money supply effected through a reduction in the supply of asset A. The expansion of m causes i to fall. As all assets in the coded version of the model are one period, all three interest rates move identically and this stimulates both output and inflation.

A key result of this parameter setting is that, under these assumptions the nature of the asset purchased is wholly unimportant. The impact on interest rates would be the same were the open market operation to exchange money for asset A or for asset B. In fact, there is no need for the central bank to make any

[5] In fact, the quantity of money is not only unimportant, it becomes indeterminate.

kind of purchase at all. The effect on the economy would be exactly the same if the central bank purely printed money and distributed it in the economy, asking for nothing in exchange. A policy of this kind has become known as a helicopter drop. The fact that all that matters is the change in the central bank's liabilities aligns this view of the world with a monetarist outlook, although the comparison is in some ways imperfect as in this model the impact on the economy from an expansion of the money supply still relies on its ability to move the short term interest rate.

4.2 Wider supply effects

So far we have looked at a world in which money is a special case, providing utility beyond its percuniary returns due to its specific characteristics. However, prices of the other assets in the economy continue to be determined purely by the expectations hypothesis, unaffected by their own supply. This is due to the perfectly elastic demand curves which are a result of the assumptions of complete and frictionless financial markets, and the Wallace Neutrality inherent in the representative agent paradigm. In the words of James Tobin

> "all non-monetary assets and debts are to be taken to be perfect substitutes at a common interest rate plus or minus an exogenous interest differential" Tobin (1982)

This feature of the New Keynesian model is contradicted by a growing empirical literature on the existence of significant supply effects. D'Amico and King (2010) and Meaning and Zhu (2011) both use individual bond level data to find significant changes in bond prices as a result of changes in the supply of government securities in recent central bank asset purchase programmes in the US and UK. Joyce et al (2010) find similar evidence for the UK, as do Breedon et al (2012) and Banerjee et al (2012). The challenge has been developing a rigorous theoretical framework to explain such results, with ex Federal Reserve Board Chairman Ben Bernanke describing asset purchases as policies which "work in practice, but not in theory" Bernanke (2014)

At its heart, it requires the breaking of some of the core assumptions of the New Keynesian model, most crucially those of a representative agent and perfect and complete financial markets.

The idea that there might be imperfect substitutability between assets which could be exploited for monetary policy purposes has its roots in the work of Tobin (1963, 1969) and Culbertson (1957) who showed, relatively intuitively that if there was imperfect substitutability between assets then changes in supply would induce movements in rates of interest. Modigliani and Sutch (1966) began to develop a more rigorous theoretical explanation as to why imperfect substitutability may exist based on the idea that investors had a preference for certain assets or areas of the term structure. For instance, they may wish to hold assets with the same maturity profile as their liabilities, or in an extreme case, be mandated by law to hold assets with certain characteristics.[6] This preference means that they value certain assets beyond their financial payoff, in much the same way as the transaction service gave an non-percuniary value to money. Investors can be tempted from their preferred mix of assets, or habitat, but require an increased payoff as inducement to compensate them for the switch. This preferred habitat theory sat between the two extremes of segmented market theory, which assumed no substitution between assets, and the expectations hypothesis which assumed perfect substitutability. More recently, Andres, Lopez-Salido and Nelson (2004) looked to build imperfect substitutability into a New Keynesian-style DSGE model. They introduce both long and short bonds which are imperfect substitutes for two reasons; long bonds incur an additional transaction cost and are also perceived as more risky and less liquid so households hold more cash to hedge against this risk which creates a second additional cost to holding long bonds. This is not however sufficient to generate a role for supply effects as households could always choose to by-pass the long bond market entirely and avoid any premia which occur. To counter this, Andres et al (2004) restrict a fraction of households to only being able to operate in long bond markets, effectively breaking the representative agent assumption and imposing the strictest degree of asset preference for one agent. The result is that changes in the relative quantities of short term bonds, characterised as money, and long bonds move the long-term interest rate and this cannot be fully arbitraged away. This then feeds into the IS curve and affects agents' behaviour.

Vayanos and Vila (2009) create a similar framework to introduce imperfect substitutability between assets. In their model there exist agents who are restricted to one market segment or another, and a limited number of agents who can arbitrage and trade across markets. When an arbitrage opportunity arises these unrestricted investors can exploit it, but in doing so become increasingly exposed to risk and so demand a higher return to induce them to take on the carry trade. This leads to an increase in rates and as such a channel by which

[6] Clear examples here are the UK pensions market which are legally required to hold long-term government securities, or the need to hold liquid assets under the latest Basel Liquidity Coverage Ratio.

changes in supply can affect interest rates.

The risk-aversion mechanism can be couched in terms of another strand of models. In the financial accelerator models of Bernanke, Gertler and Gilchrist (1999) and Kiyotaki and Moore (1997) agents issuing debt have an increasing incentive to default the more leveraged they become. For a given level of underlying collateral therefore increasing the supply of debt issued increases the risk of default. To incentivise the private sector to take on this riskier supply, the interest rate must rise accordingly.

The existence of credit constraints can also directly limit the extent of arbitrage. A central tenant of the no-arbitrage condition is that any opportunities can be exploited in unlimited quantity until the arbitrage is eliminated. If investors do not have the funds to take advantage of the arbitrage situation then complete credit markets will afford them the means. In reality credit constraints may limit their ability to do this, especially if the change in supply required is extremely large.

The same principle applies to transactions costs. If there are large costs associated with changing one's portfolio then large changes to take advantage of arbitrage would inflict significant costs for which the investor would need to be recompensed, or which might prove prohibitive to exploiting the arbitrage opportunity itself.

For asset purchases by the central bank to be a viable policy tool though, it is not sufficient that the overall supply of an asset has consequences for its price. It must also be true that investors differentiate between assets they hold themselves in their own portfolio, and those held on their behalf by the public sector. In other words, it requires that Wallace Neutrality fails. This can be achieved theoretically in a variety of ways, most of which rely on the introduction of heterogenous agents. Without a representative agent, any change in the future fiscal position resulting from taking assets on to the public sector balance sheet need not be distributed amongst private sector agents in an identical manner to the payoffs of the asset when it was held by the original private sector investor. This means agents are unlikely to adjust their demand for assets to perfectly offset the impact on prices. A clear example would be the differing payoffs to a pension fund if it holds a government bond on its balance sheet compared to if it sells that bond to the public sector and experiences a negligible increase in economy-wide inflation. Benigno and Nistico (2015) demonstrate a similar non-neutrality in the context of open market operations stemming from the removal of a fiscal indemnity which results in the central bank having to increase inflation when faced with losses from asset purchases.

Alternatively, Wallace Neutrality could break due to myopic, or only finitely-lived agents who do not take account of future tax increases when making their investment decisions. A similar argument is often heard in the discussion of Ricardian Equivalence.

The key is that agents actually perceive a change in their personal payoffs in future states of nature resulting from the change in the portfolio of assets they hold directly.

Crucially, our general and stylised model does not seek to critique and evaluate such underlying mechanisms. Rather it takes the result presented by this strand of the literature, however arrived at, as one potential version of the world and investigates its implications for monetary policy.

4.3 Supply effects in our general model

However they are motivated, we can use our general model to think about a Tobin-esque world in which the supply of assets matters for their price and return. As with our discussion of money in section 4.1, such a world implies that the demand curve for assets A and B are less than perfectly elastic, and therefore β^A and β^B are both greater than zero.

Let us reconsider our earlier thought experiment; an open market operation exchanging asset A for newly created money. The first point of note is that the existence of supply effects does not preclude the transmission mechanisms we have previously outlined. The expansion in the money supply works in exactly the same way as detailed above; i_t falls to clear the money market and this lowers both i_t^A and i_t^B through the term structure. However, we now get an additional effect on i_t^A caused by the reduction in the quantity of A available to the public. Whatever the underlying friction, the change in supply leads to a movement in the price and yield of A which cannot be fully arbitraged away.

This direct supply effect may also be accompanied by a more broadly "local" supply effect if asset B is an imperfect substitute for asset A, in other words if $\delta > 0$. Intuitively δ is the extent to which B operates as a successful substitute for A. Imagine the two assets are very close substitutes and supply matters for determining price. In this case the relevant variable is the joint supply of both A and B, as they are interchangeable, so $\delta \simeq 1$. As $\delta \to 0$ the less of an effective substitute B is for A and the less the supply of B matters in determining the price of A, and vice versa. Therefore, if $0 < \delta < 1$, as exchanging asset A for money increases the price of asset A, at least some investors will be induced to switch out of holding A and in to holding asset B. This increase in demand for asset B will push up on its price and down on its yield, lowering i_t^B by an amount in excess of the pure expectations hypothesis. This transmission mechanism has become known as the portfolio rebalancing channel.

The values of β^A, β^B and δ are therefore integral to the transmission and efficacy of asset purchases as a monetary policy tool. To see why, let us again imagine our A for money open market operation, but assuming demand for asset A is almost entirely inelastic. As β^A tends to infinity, δ tends to zero. This would mean that our open market operation would have an increasingly large impact on i^A but the transmission to i^B would diminish to nothing. In other words, there would be a large direct effect, but no portfolio rebalancing. All that the open market operation is achieving is to disconnect asset A from the yield curve. Contrast this with the case where demand for assets is perfectly elastic and so β^A and β^B tend to zero, but δ tends to one. Now the assets are perfect substitutes, so a change in the supply of one is equivalent to the change in the supply of the other, but the perfectly elastic demand means that a change in supply of either has no impact on i^A or i^B. This can be seen in figure 5 which applies the same shock to the supply of money under different parameterisations of these parameters. When $\beta^A = 0$ and $\delta = 1$ then the movement of all variables is the same as the initial shock presented in figure 4. However, as the demand curve for asset A becomes less elastic, the impact of the operation on i^A increases. Interestingly, the impact on i reduces, as the heightened sensitivity of other assets to the operation takes some of the burden of stabilising the economy and so relaxes the amount of work required by the policy rate. The impact on i^B is determined by the interaction of β^B and δ. A higher degree of pass-through from increased substitutability can be offset by a more elastic demand curve.

This has implications for the design of policy. In a world where β^A was large but δ low, asset purchases would represent a useful tool for manipulating a targeted market, or market segment. Such a scenario is likely to occur when market functioning is impaired, meaning this may be a practically relevant result. Putting it in to a real world context, it suggests that in this version of the world there was a role for targeted asset purchases in impaired markets at the inception of the 2008 crisis, as the beta values were likely to be high with strain in specific markets. However, such a characterisation would also mean that asset purchases were less effective as a tool for macroeconomic stabilisation, as there is little to no transmission to other rates in the economy. For asset purchases to be effective as a tool for wider macroeconomic and monetary stabilisation, they require a world between these two extremes in which there is less than perfect arbitrage, but assets are still to some degree substitutable. Moving along the spectrum between the two poles will alter how asset purchases work and transmit through the economy.

So, what is likely to determine the values of these parameters? The stronger the frictions that limit arbitrage opportunities, be they credit constraints, risk aversion of arbitrageurs or valuing assets for characteristics beyond their pecuniary return are at their largest the greater we would expect β^A to be. This suggests that asset purchases may be especially effective at manipulating rates in times of financial turmoil, as this is exactly when frictions are at their strongest and markets become most segmented. This is not to discount the importance of this supply channel in more normal times, as it is highly probable that frictions exist to one extent or another even when markets are functioning relative well. In fact, D'Amico et al (2012) find significant supply effects in a pre-crisis sample for the US economy.

It is also important to note that β^A and δ are unlikely to be independent of one another. The less there exists close and viable substitutes for asset A, the more inelastic demand for A will be. Therefore, β^A and δ are likely to be inversely correlated. This is a significant statement as it highlights a trade-off at the heart of the debate around supply effects and the use of asset purchases as a policy tool.

5 A demand shock

Figure 6 shows the response of the model to a demand shock under two parameterisations. In the first there are no frictions which give rise to supply effects for any asset other than money. In the second β^A, β^B and δ are set so that there are significant supply effects on assets A and B. As can be seen, the monetary authority has to reduce the policy rate by less to stabilise the economy in the latter case as the reduction in the premium on A moves i^A by more. This is an important result as it means that the correct policy setting, even in normal times must be aware the extent of supply effect. The existence of supply effects does appear to lessen the falls in output and inflation, though under this setting the magnitude is not great.

6 Policies at the zero lower bound

A dominant feature of the recent economic landscape has been the binding of the lower bound on short-term nominal interest rates. In fact, it has been this development above all others which has reawakened interest in asset purchases as an explicit policy tool and spurred the debate surrounding their theoretical basis. Our discussion thus far has taken place in the context of unconstrained rates, so all of our insights are valid away from the lower bound. However, applying our general model we can think about how these insights might change when the traditional policy channels are constrained.

The cashless version of the New Keynesian model is little affected by the lower bound. Unable to lower i_t however, the only mechanism left by which monetary policymakers can relax the stance of policy and stimulate demand is to reduce $E_t \sum i_t$. Shaping expectations in this way has become known as forward guidance, and has been adopted as part of the policy strategy of almost all major central banks at some point in recent years.[7] To change agents' expectations the central bank has to convince them it will hold the policy rate low for longer than they currently expect, transmitting along the yield curve and lowering rates at longer term. Within the New Keynesian model, and often in wider thinking this is framed as a commitment to hold rates lower than the designated policy rule would deem appropriate for a period of time, and as such equates to a commitment to higher future inflation, Eggertson and Woodford (2003). This raises issues of credibility and time-inconsistency as an inflation targeting central bank would have an incentive to renege on such a commitment once inflation was restored to the target level. This argument has lead some to suggest a subtle mechanism through which asset purchases may aid policy implementation. By holding lots of financial assets on an expanded balance sheet, the central bank is increasingly exposed to interest rate risk. As such, raising rates before these assets have been returned to the market would inflict losses on the central bank. Assuming the central bank has an aversion to losses which features in its objective function, this may stay its hand in raising rates, adding credibility to its low for longer commitment and thus making it more effective at shaping expectations and lowering rates. This "skin in the game" channel of asset purchases seems unlikely to be significant though as there is no evidence that losses from asset purchase programmes feature in the objective function of central banks, especially considering that many programmes are subject to an indemnity

[7] See chapter 2 of this thesis.

against losses from the fiscal authority. What is more, to the extent that losses are a consideration in central bank decision making, they are likely to be dwarfed by the mandated macrostabilisation objectives.

Besides this, there are a number of other reasons why a commitment to holding rates lower for longer may not be time-inconsistent in practice. Under the full information, rational expectations assumptions of the model, agents perfectly know the central bank's reaction function and all agents have the same view of the future evolution of variables because they know the nature of the underlying model structure. This means that agents' expectations of future interest rates, by design, arrive at the time-consistent outcome.

In reality though this need not be the case. In fact, it is highly unlikely to be so. This means that communication that rates will be lower for longer than markets expect need not be a deviation from the central bank's time-consistent policy rule, but rather a consequence of markets misunderstanding the central bank's reaction function, and so not accurately factoring in the true time-consistent path for policy. This may be because they have always misunderstood the reaction function, or that the reaction function itself has changed. Given the novelty of central banks finding themselves at the lower bound, it is plausible that neither they, nor markets were sure of their reaction function in this environment.

Similarly, it may be that the central bank has a different view on the outlook for the economy and is signalling that they expect the future paths of key aggregates to be weaker than markets envision. As such, in this context the signal that rates will be lower for longer than market expectations suggest moves those expectations towards a time-consistent path for policy, not away from it.

At the lower bound, the variant of our general model which introduced money collapses to the standard New Keynesian model. This is because the lower bound represents the point at which the transaction service of money has been exhausted and money demand has become perfectly elastic. Asset purchases which expand the monetary base beyond this point have no effect and the economy has entered a liquidity trap. The policy prescription then becomes equivalent to that of the standard New Keynesian baseline.

As with our analysis away from the lower bound, the final Tobin-esque incarnation of our model allows monetary policy to gain direct traction through manipulating the supplies of assets. Like the first two variants, the traditional interest rate channel is impaired as changes in the supply of money no longer lower i_t, although anything which shapes $E_t \sum i_t$ will continue to be effective. However, changes in supply of assets A or B will serve to reduce i_t^A and i_t^B. In theory, the limit of such a policy is when the supply of assets A and B have been reduced sufficiently that i_t^A and i_t^B both reach their own lower bounds. This is a theoretical proposition which seems to have begun to be tested by the Bank of Japan who have reduced premia on some assets by so much that they have had to expand their purchases to a wider range of assets in search of rates of interest above their lower bound.

7 Alternative channels for asset purchases

Our exposition so far has focussed on the ability of asset purchases to influence interest rates, and from that point has followed the traditional Keynesian interest rate channel through the IS curve to the household's intertemporal decision. However, there are a number of alternative candidate channels through which asset purchases may act to stimulate demand.

The Bank of Japan's original quantitative easing programme in the early 2000s placed much importance on the bank lending channel. The core of this idea is that asset purchases expand the level of deposits and reserves in the banking system and this leads banks to increase credit creation to the private sector, which in turn buoys activity. However, the idea that this process works through a simple function whereby banks extend loans as some fixed multiple of reserves is at best outdated. Not only is this proposition contradicted heavily by the data, but also by a growing academic literature, the main points of which are excellently distilled by McLeay, Radia and Thomas (2014). Instead, models such as Kashyap and Stein (1995) and Chadha, Corrado and Meaning (this volume) look to how the increase of reserves and deposits affects banks' funding costs and as such the optimal quantity of loans they extend. In general, by increasing deposits and reserves, central banks are providing a cheaper form of funding to banks than if they went to market and this should incentivise them to use this funding to finance loans. This has been empirically tested by authors such as Butt et al (2014) and Joyce and Spaltro (2014), in the UK and Bowman et al (2011) for Japan and found to be weak at best. A separate strand of literature on the bank lending channel looks at asset purchases as a way to mend banks balance sheets. If banks find themselves credit constrained, or limited in their ability to lend due to damaged balance sheets and risk, then an injection of safe, liquid assets in the form of reserves may alleviate such constraints and spur them to lend. Disyatat (2010) explains this mechanism clearly and develops a model in which the health of banks' balance sheets directly affects the quantity of lending.

Similarly, Gertler and Karadi (2009) extend a DSGE model with a banking sector in which the central bank can add capital to commercial banks' balance sheets via asset purchases and relax constraints on lending, boosting demand.

The principle of improving balance sheets that Disyatat (2010) applies to banks may be broadly applied to households and investors as well. If agents are credit constrained by the collateral they hold, as in the seminal models by Bernanke, Gertler and Gilchrist (1999) and Kiyotaki and Moore (1997), then providing the private sector with an increased quantity of high quality assets that they can use as collateral against which to borrow will relax their credit constraint and allow them to increase their investment or spending.

A more controversial channel through which central bank asset purchases may help to stimulate the economy is through their interaction with the fiscal authority. A large part of the assets purchased by central banks in recent years has been government securities, with the consequence of substantially reducing the publicly available stock of such securities and also the cost to governments of issuing debt. Thinking in terms of a fiscal budget constraint, this creates fiscal space, which might allow the government to increase spending or reduce taxation, boosting demand. The controversy arises when this is framed as monetary financing of fiscal spending, a subject which is taboo and off limits in the eyes of many economists. Jordi Gali (2014) shows however how such a policy can be welfare enhancing, even if the government spending which results is less than fully efficient. This channel is likely to be especially effective when the fiscal multiplier is large, as it is often suggested it is when monetary policy is constrained at the lower bound.

Last is the monetarist view, in which expansion of the money supply is stimulative in and of itself. Under this view, famously put forward by Milton Freidman, increases in the quantity of money leads to more money chasing the same quantity of goods and services in the economy. This bids up asset prices and prices more broadly and eventually leads to a pick-up in nominal spending. Bridges and Thomas (2012) provide a framework both for thinking about, and beginning to quantify such a channel.

8 Conclusions

The standard New Keynesian model contains no direct channel by which asset purchases can affect the economy. This is because, under its core assumptions of a representative agent with rational expectations and complete and perfect financial markets any arbitrage opportunities are instantly exploited and any exchange of assets between the public and private sectors is inconsequential for prices as the ultimate exposure of the household is left unchanged regardless of whether they hold the asset directly, or through the government.

The only way in which monetary policy operates is through setting the short-term nominal interest rate, or manipulating agents' expectations of its future path. Asset purchases are therefore only effective to the extent that they act to signal that path, or add credibility to the path committed to by the central bank.

By explicitly introducing a money demand function to the baseline New Keynesian model we show that if that demand is less than perfectly elastic, the central bank will have to vary the quantity of money in the economy to change the rate of interest which prevails. Such a downward sloping demand curve can be motivated theoretically by the transaction and liquidity services of money. It has also been the bedrock of practical monetary policy for some time as central banks have used open market operations to ensure the money market clears at their target interest rate. Augmenting the model in this way provides a role for asset purchases which expand the monetary base insofar as they lower the short-term nominal interest rate. However, because the New Keynesian model's assumptions still hold for all financial assets other than money, the nature of the assets for which the newly created money is exchanged is irrelevant. In fact, the same effect could be achieved by a helicopter drop which expanded the money supply without buying any assets at all. What is more, once the money supply has been expanded to the point that the transaction and liquidity services of money are exhausted then money demand becomes satiated and perfectly elastic. At this point further money creation by the central bank cannot push rates lower, the lower bound has been reached and the augmented model collapses back to the standard New Keynesian framework with the same policy prescription.

Further extension of our model allows us to relax the assumption that wider financial markets are complete without frictions. Such a setting has been motivated theoretically by the existence of imperfect substitution between assets, preferred habitat investors and risk-averse or credit constrained arbitrageurs. It could also stem more broadly from the existence of heterogenous agents who face the benefits and losses from asset purchases differently, or the failing of the New Keynesian model's strong assumption on complete financial markets. Whatever the underlying cause, if its supply has significant consequences for the price of an asset then asset purchases by the central bank will be able to exploit these and directly manipulate a whole range of interest rates The efficacy with which it can do this will be dependent on the extent to which frictions limit arbitrage and assets are imperfect substitutes, captured in our model by the parameters β^A, β^B and δ. This suggests that asset purchases may be especially effective in times of financial market turmoil when frictions are more pronounced and markets generally more segmented. Importantly such supply effects, if they exist, are little affected by the lower bound on the short-term nominal interest rate and thus provide a way for monetary policy to continue to ease policy when confronted with a constraint on its traditional policy instrument.

Having presented and discussed the main theoretical views on asset purchases we hope to show that each result rests on the assumptions underlying it. Ultimately, the only way to validate one over the other is through empirical testing, something this author attempts to do later in this thesis, and which is at the heart of a burgeoning quantitative literature. Our results, along with those of many other authors, increasingly appear to validate the existence of supply effects and imperfect substitution between assets. More often than not, these are found alongside weak expectational effects which imply an element of the signalling channel highlighted by the baseline New Keynesian model. Away from the lower bound, the money demand curve appears to be inelastic to varying degrees, with a monetary expansion leading to lower money market rates and so it is likely that all three elements exist simultaneously to one degree or another supporting the final parameterisation of the general model presented in this paper.

The challenge remains to build convincing and robust microfoundations which can explain this behaviour in a comprehensive macroeconomic framework. This is part of an ongoing and vital research agenda. For now what seems certain is that it will require deviations from some of the core assumptions behind the previous paradigm, developing a world of heterogeneous agents and imperfect financial markets.

References

[1] Andre´s, J., Lopez-Salido, J.D. and Nelson, E. (2004). 'Tobin's imperfect asset substitution in optimizing general equilibrium', Journal of Money, Credit and Banking, vol. 36(4), pp. 665–90.

[2] Banerjee, R & McLaren, N, (2012). "Using changes in auction maturity sectors to help identify the impact of QE on gilt yields," Bank of England Quarterly Bulletin, Bank of England, vol. 52(2), pages 129-137.

[3] Benigno, P and Nisticò, S (2015) "Non-Neutrality of Open-Market Operations" CEPR Discussion Paper DP10594

[4] Bernanke, B, Gertler, M and Gilchrist, S (1997) "The financial accelerator in a quantitative business cycle framework," Handbook of Macroeconomics, in: J. B. Taylor & M. Woodford (ed.), Handbook of Macroeconomics, edition 1, volume 1, chapter 21, pages 1341-1393 Elsevier.

[5] Bowman, D., F. Cai, S. Davies, and S. Kamin (2011): "Quantitative Easing and bank lending: evidence from Japan," Discussion Paper 1018, Board of Governors of the Federal Reserve System (U.S.).

[6] Breedon, F., Chadha, J.S.and Waters, A (2012). "The financial market impact of UK quantitative easing," Oxford Review of Economic Policy, Oxford University Press, vol. 28(4), pages 702-728, WITER.

[7] Bridges, J and Thomas, R (2012), 'The impact of QE on the UK economy — some supportive monetarist arithmetic', Bank of England Working Paper No. 442.

[8] Butt, N, Churm, R, MacMahon, M, Morotz, A and Schanz, J (2014),'QE and the bank lending channel in the United Kingdom', Bank of England Working Paper 511

[9] Calvo, G,A. (1983) "Staggered prices in a utility-maximizing framework" Journal of Monetary Economics, Elsevier, vol. 12(3), pages 383-398, September.

[10] Chadha, J.S, Corrado,L and Holly, S (2014). "A note on money and the conduct of monetary policy" Macroeconomic Dynamics, 18, pp 1854-1883

[11] Cox, J, Ingersoll, J and Ross, S (1981) "The relation between forward prices and futures prices" Journal of Financial Economics, 1981, vol. 9, issue 4, pages 321-346

[12] Clarida, R., Gali, J and Gertler, M (1999) "The Science of Monetary Policy: A New Keynesian Perspective." Journal of Economic Literature, 37(4): 1661-1707.

[13] Culbertson, J. M. (1957). "The Term Structure of Interest Rates". Quarterly Journal of Economics 71, 485-517

[14] D'Amico, S and T King (2010): "Flow and stock effects of large-scale treasury purchases", Federal Reserve Board Finance and Economics Discussion Series, 2010–52

[15] D'Amico, S, English, W, López-Salido, D and Nelson, E (2012). "The Federal Reserve's Large-scale Asset Purchase Programmes: Rationale and Effects," Economic Journal, Royal Economic Society, vol. 122(564), pages F415-F446, November.

[16] Disyatat, P (2010) "The bank lending channel revisited," BIS Working Papers 297, Bank for International Settlements.

[17] Eggertson, G and M, Woodford (2003)'The Zero Bound on Interest Rates and Optimal Monetary Policy' , Brookings Papers on Economic Activity, (1), pp. 139–233

[18] Fisher, S (1977) "Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule" The Journal of Political Economy, Vol. 85, No. 1 (Feb., 1977), pp. 191-205

[19] Gali, J (2014) "The Effects of a Money-Financed Fiscal Stimulus" CREI

[20] Gertler, M and Karadi, P (2009)"A model of unconventional monetary policy," Journal of Monetary Economics, Elsevier, vol. 58(1), pages 17-34, January.

[21] Gromb, D and Vayanos, D (2015) "The Dynamics of Financially Constrained Arbitrage," NBER Working Papers 20968, National Bureau of Economic Research, Inc.

[22] Harrison, R (2012) "Asset purchase policy at the effective lower bound for interest rates," Bank of England working papers 444

[23] Ireland, P (2005) "The Monetary Transmission Mechanism" Federal Reserve Bank of Boston Working Paper No. 06-1

[24] Joyce, M. A. S., A. Lasaosa, I. Stevens, and M. Tong (2011): "The financial market impact of Quantitative Easing in the United Kingdom," International Journa of Central Banking, pp. 113–161.

[25] Joyce, M., and M. Spaltro (2014): "Quantitative easing and bank lending: a panel data approach," Bank of England working papers 504, Bank of England.

[26] Joyce, M, Tong, M and Woods R (2011) "The United Kingdom's quantitative easing policy: design, operation and impact" Bank of England Quarterly Bulletin Q3 2011

[27] Kashyap, A. K., and J. C. Stein (1995): "The impact of monetary policy on bank balance sheets," Carnegie-Rochester Conference Series on Public Policy, 42, 151–195.

[28] Kerr, W and King, R,G (1996) "Limits on interest rate rules in the IS model," Economic Quarterly, Federal Reserve Bank of Richmond, issue Spr, pages 47-75.

[29] Keynes, JM (1936) "The General Theory of Employment, Interest and Money"

[30] Kiyotaki, N and Moore, J (1997) "Credit Cycles," Journal of Political Economy, University of Chicago Press, vol. 105(2), pages 211-48, April.

[31] Kydland, F and Prescott, E (1982) "Time to Build and Aggregate Fluctuations". Econometrica 50 (6): 1345–1370

[32] Lucas, R (1970) "Capacity, Overtime, and Empirical Production Functions," American Economic Review, American Economic Association, vol. 60(2), pages 23-27, May.

[33] Lucas, R (1972) "Expectations and the neutrality of money," Journal of Economic Theory, Elsevier, vol. 4(2), pages 103-124, April

[34] Lucas, R (1976) "Econometric policy evaluation: A critique," Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 1(1), pages 19-46, January.

[35] Mankiw, N,G and Romer, D (ed.) (1991) "New Keynesian Economics - Vol. 1: Imperfect Competition and Sticky Prices," MIT Press Books, The MIT Press, edition 1, volume 1

[36] Mankiw, N,G and Romer, D (ed.) (1991) "New Keynesian Economics - Vol. 2: Coordination Failures and Real Rigidities," MIT Press Books, The MIT Press, edition 1, volume 2

[37] McCallum, BT (2012) "The Role of Money in New-Keynesian Models" Banco Central de Reserve del Peru Discussion Paper

[38] McLeay, M., A. Radia, and R. Thomas (2014): "Money creation in the modern economy," Bank of England Quarterly Bulletin, 2014 Q1, 54, 14–27.

[39] Meaning, J and Zhu, F (2011). 'The Impact of Recent Central Bank Asset Purchase Programmes', BIS Quarterly Review, December 2011

[40] Modigliani, F. and Sutch, R.C. (1966). "Innovations in interest rate policy", American Economic Review, Papers and Proceedings, vol. 56(2), pp. 178–97.

[41] Muth, J (1961) "Rational Expectations and the Theory of Price Movements", Econometrica 29, pp. 315–335.

[42] Phelps, E.S. and Taylor, J.B. (1977). "Stabilizing Powers of Monetary Policy under Rational Expectations". Journal of Political Economy 85, 163-190

[43] Poole, W. (1970). "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model". Quarterly Journal of Economics 84, 197-216.

[44] Taylor, J.B (1979) "Staggered Wage Setting in a Macro Model," American Economic Review, American Economic Association, vol. 69(2), pages 108-13, May.

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

[46] Tobin, J (1969) "A General Equilibrium Approach to Monetary Theory," Journal of Money, Credit and Banking, 1, February, 15-29.

[47] Tobin, J (1982) "Money and Finance in the Macroeconomic Process" Journal of Money, Credit and Banking, 14:2

[48] Vayanos, D and Villa, J (2009) "A Preferred-Habitat Model of the Term Structure of Interest Rates," CEPR Discussion Papers 7547, C.E.P.R. Discussion Papers.

[49] Wallace, N (1981). "A Modigliani-Miller Theorem for Open-Market Operations," American Economic Review, American Economic Association, vol. 71(3), pages 267-74, June.

[50] Walsh, C (2003) "Monetary Theory and Policy", MIT Press

[51] Woodford, M (2003)."Interest and prices: Foundations of a theory of monetary policy". Princeton, New Jersey: Princeton University Press.

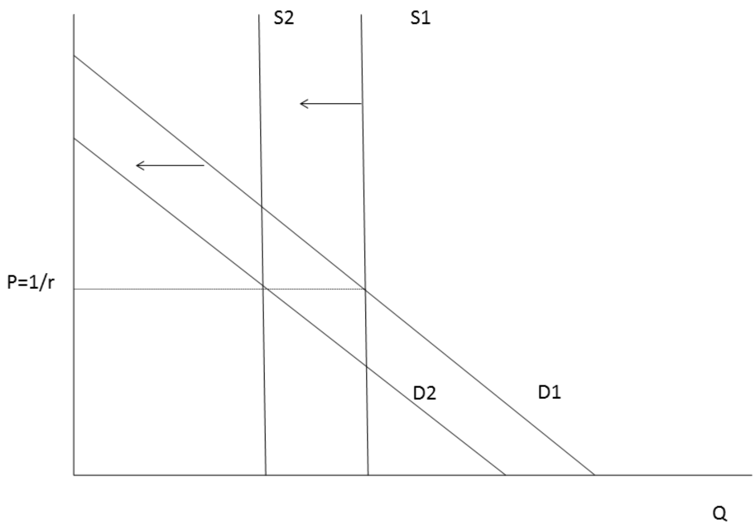

Figure 1: A graphical representation of Wallace Neutrality

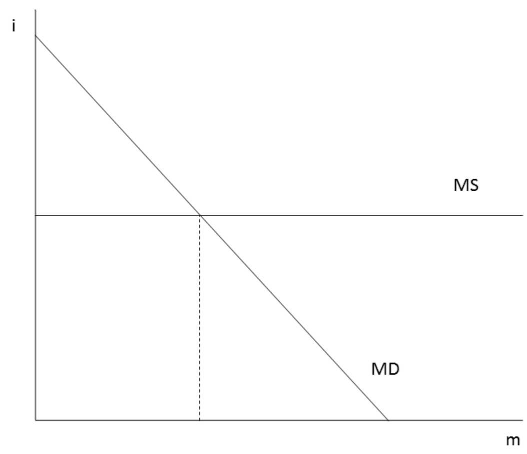

Figure 2: Setting policy via the interest rate - inelastic money demand with perfectly elastic money supply

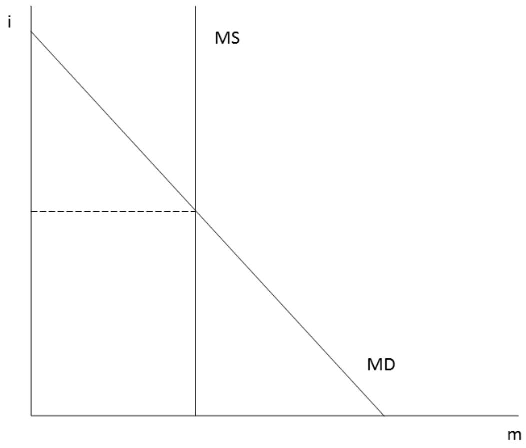

Figure 3: Setting policy via quantity of money - inelastic money demand with perfectly inelastic money supply

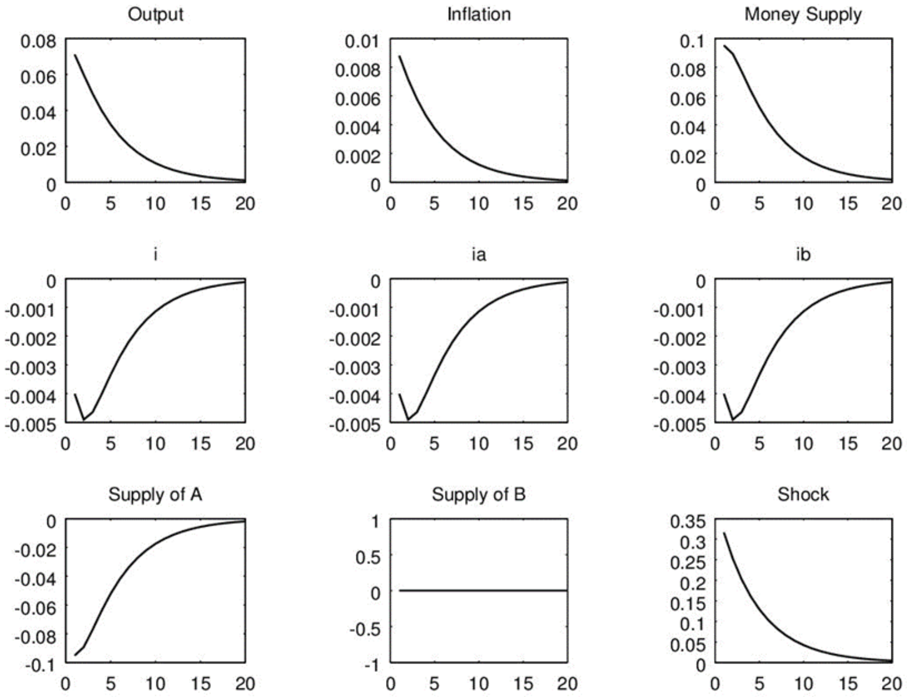

Figure 4: An open market operation which increases the money supply by purchasing asset A

$\beta^A = 0$, $\beta^B = 0$, $\delta = 1$

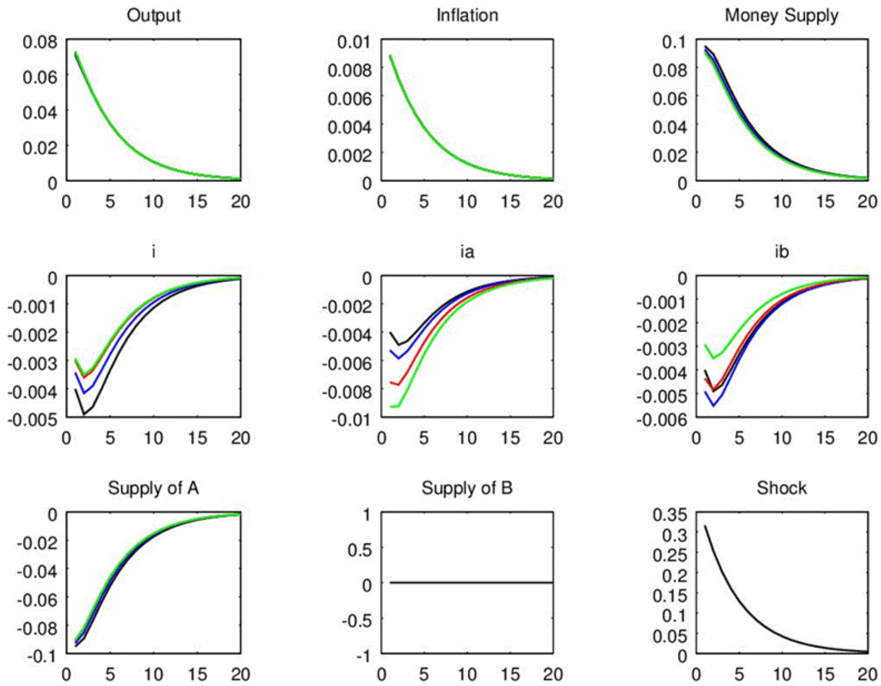

Figure 5: An open market operation swapping money for asset A under varying degrees of imperfect substitutability and inelastic demand curves

Parameter values

	Perfect substitutability		$\longrightarrow$	No substitutability
	Black line	Blue line	Red line	Green line
β^A	0	0.02	0.05	0.07
β^B	0	0.02	0.05	0.07
δ	1	0.8	0.3	0

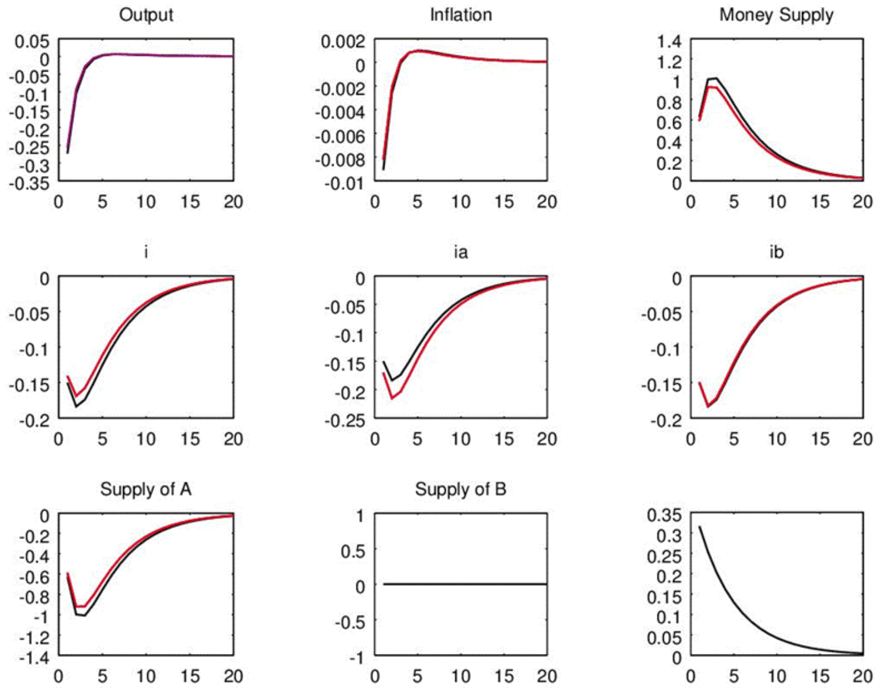

Figure 6: A demand shock with and without supply effects

Parameter values

	No supply effects (Black line)	Supply effects (Red line)
β^A	0	0.05
β^B	0	0.05
δ	1	0.3

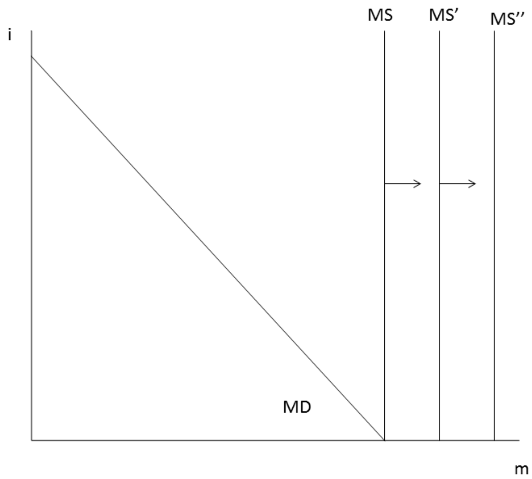

Figure 7: Monetary expansion beyond the lower bound

A recent history of monetary policy

Jack Meaning

Abstract

This paper tells the story of four central banks - the Federal Reserve, the Bank of England, the European Central Bank and the Bank of Japan - and how the way in which they conduct policy has evolved dramatically over the last decade. It moves from the Great Moderation, in which monetary policy had become thought of as a science and the short-term nominal interest rate was considered sufficient to stabilise the economy, to the present day in which the size and the composition of central bank balance sheets has taken centre stage. It details each bank's response to the crisis of 2007/8 and how their balance sheets were first used to act as lenders of last resort and provide liquidity to an ailing financial system, before the lower bound constraint forced them to innovate in order to stimulate demand more broadly.

Introduction

In 2000 Mervyn King, then Deputy Governor of the Bank of England claimed that a successful central banker should be boring, King (2000). He was speaking at a time when it looked as if the practical questions of how monetary policy could stabilise the economy had been emphatically answered. Central bankers could seemingly encourage or dampen inflationary pressure in a reliable manner by setting the short-term nominal interest rate. Consistent with the seminal analysis of Tinbergen (1952), this single instrument was sufficient to pursue the single, simplified objective of inflation targetting, which was becoming ever more widely adopted. This policy framework represented a change from the 1970s and 1980s when the focus of policy had been the explicit control of the money supply.

We are now 15 years after King's speech, and monetary policy has been anything but boring. In fact, it has undergone a period of immense change which has fundamentally altered how it is conducted. In response to a number of challenges, central banks have moved from relying solely on the short-term interest rate to employing their balance sheets in a variety of ways in the pursuit of an expanded set of objectives. The size and composition of central bank balance sheets has become (once more) a focal part of monetary policy in practice.

This paper seeks to provide the narrative of how this evolution occurred through the prism of four of the most prominent central banks in the advanced world. It is intended as a primer, to give some historical context to the more academic discussions in the rest of this thesis. Although every effort has been made to be as thorough as possible the sheer volume of policy innovations over the period considered means that this, indeed any paper, is hard pushed to be completely comprehensive. Therefore I would suggest a selection of wider reading for those in search of deeper, or more technical understanding of the policies described.

For its contribution, this paper details the consensus framework for conducting policy throughout the Great Moderation in which most central banks targeted macroeconomic aggregates by means of the short-term nominal interest rate. The early exception was the Bank of Japan who initiated a quantitative easing programme in the early part of the last decade. The paper then goes on to describe the policy responses of each central bank to the financial crisis of 2007/08. These can be broadly categorised into three eras. First a period of liquidity provision in which all four central banks fulfilled their lender of last resort function to support a financial sector in free-fall. As this subsided, the focus turned to stabilising the macroeconomy and stimulating demand. With the traditional instrument of choice constrained by its lower bound, each of the four central banks looked to other instruments by which policy could gain traction. This lead to a range of asset purchase programmes and lending facilities which caused central bank balance sheets to expand dramatically, with some now reaching 60 per cent of GDP.

Concurrent with the evolution of how central banks conduct monetary policy, many have found their remits expanded to include new objectives such as macroprudential regulation and banking supervision. Whilst of great importance and interest, these innovations are not the focus of this paper. Similarly, the paper does little to evaluate the effectiveness of the policies discussed. This is left to an extensive literature, to which the author himself has contributed.

Policy in the Great Moderation

The standard pre-crisis paradigm

The decade prior to 2007 was characterised by a largely stable macroeconomic environment. Most central banks had been granted at least operational independence subject to mandated targets for key aggregates. In the case of the Bank of England and the European Central Bank these targets were for levels of consumer price inflation,[1] whilst the Federal Reserve's remit was, and remains the broader and slightly more vague objective of "maximum employment, stable prices and moderate long-term interest rates".

As can be seen in figures 1-3, the advanced economies did indeed evolve broadly in line with these mandates throughout the late 1990s and early 2000s and this period of stability became known as the Great Moderation.

The implementation of monetary policy had been considered distilled to a science. Following the seminal analysis by Poole (1970), central banks had moved away from the quantity of money as the predominant instrument, and rather focused on the price of money; the short-term nominal interest rate. Work by John Taylor had shown that a simple policy rule involving inflation and output was sufficient to stabilise the economy as long as nominal interest rates moved more than one-for-one with deviations in inflation, as this would then ensure that real rates would move appropriately to bring the economy back to equilibrium,

[1] In 1997 the UK government set the Bank of England a symmetric target of 2.5% for the 12 month change in the Retail Price Index before updating the mandate to 2% annual growth in the Consumer Prices Index. The ECB's target is symmetric and aims for "close to but below 2%" annual change in the Euro Area Harmonised Consumer Price Index.

Taylor (1993).

At a technical level, central banks still varied the quantity of money in order to ensure that the money market cleared at their target interest rate, but this was considered of second-order importance for policy and often nothing more than a practical detail. The central bank set the aggregate quantity of reserves in the banking system, and the price at which it would be willing to provide additional funds via open market operations. By declaring this rate, central banks effectively controlled rates in money markets as no bank in need of reserves would pay significantly more than this to borrow from another bank if they could just go to the central bank and pay the policy rate. Likewise, unless the central bank had grossly oversupplied reserves to the system, no one bank would find itself so flush with excess reserves that it would accept a price much below the policy rate to lend them on. In order to lower the base interest rate in the economy, the central bank need only loosen the price at which it offered central bank money by cutting the interest rate, and then provide any additional reserves demanded on aggregate by the banking system at the new rate. If the aggregate quantity of reserves was set appropriately with respect to aggregate demand for them from the financial system, then the central bank's targeted policy rate should prevail and interbank markets should attenuate asymmetric demand and supply shocks between individual agents.

In practice, the supply of reserves was controlled through open market operations, in which the central bank provided newly created money to a counterparty in exchange for an asset in a repurchase agreement, or the pledge of an asset as collateral on a loan. These operations were normally of very short terms and relatively small amounts.

The Bank of Japan: the early exception

The first major central bank to deviate from this paradigm was the Bank of Japan.

Having grown rapidly through the 1980s, the Japanese economy began to slow significantly in the early 1990s. The Bank of Japan responded by cutting its primary policy instrument, the uncollateralised overnight call rate. Over a four year period between 1991 and 1995 the overnight call rate fell from 8% to $\frac{1}{2}$%, which although large was carried out at such a slow pace that it did little to strengthen confidence and bolster demand. Kuttner (2014) explains that the Bank of Japan's hesitancy to cut rates decisively at this juncture was even more pronounced when adjusting for the path of inflation. He posits that this was a defining difference between the Bank of Japan in the 1990s and the Federal Reserve in the 2008 crisis and was a fundamental reason for the worse outcome for the Japanese economy. The 1995 banking crisis solidified Japan's now persistently weak economic performance and on 12th February 1999 the Bank of Japan made its final cut in the overnight call rate to "virtually zero", which in practise meant that they oversupplied reserves to the system in order to "encourage the uncollateralised call rate to move as low as possible", figure 12.

Despite this, deflation intensified which meant that with the nominal rate bounded by zero, real rates began to increase, exacerbating the deflationary cycle. At this point the Bank of Japan moved its focus from the price of money to the quantity, officially changing its operating target from the overnight call rate to the level of current account balances (CABs) held with it by private banks on the 19th March 2001. The policy became known as quantitative easing, due to the prominence it afforded the quantity of central bank money. In many ways it harked back to the monetary policy of the 1970s and 1980s where targeting monetary aggregates was pervasive. The theoretical arguments put forward then by Freidman (1987), Tobin (1969) and others about why quantities of money and assets were important in and of themselves had fallen out of favour in the New Keynesian revolution of the 1990s, but they were once again pertinent in the context of the lower bound on interest rates.

From a mechanistic point of view the expansion of current account balances (CABs) differed little from traditional open market operations. The Bank of Japan would purchase assets from the banking sector and in exchange provide the counterparty with a credit in their current account. This expanded the quantity of central bank money in the system, and it was hoped that this would incentivise the banks to extend more loans to the wider economy, further boosting the monetary expansion.

Aside from the shift in operating target from the interest rate to the size of the monetary base, the other main difference between the Bank of Japan's quantitative easing policy and traditional OMOs was the sheer size of the programme. At inception, the target for CABs was ¥5trn, but this was raised a number of times as economic conditions failed to pick up. By 2004 it was ¥35trn, almost ten times the required reserves level of 2001. To achieve this the Bank of Japan bought a huge number of Japanese government bonds (JGBs), initially at a pace of ¥400bn a month, but eventually reaching ¥1.2trn a month. By the height of the policy in December 2005, the Bank of Japan's balance sheet reached ¥155trn and their assets consisted of around ¥100trn of Japanese government bonds, figure 8.

By 2006 annual consumer price inflation was reaching 1 per cent and looked to be consistent with the Bank of Japan's aim of eliminating deflation. The targets for CAB were dropped and the overnight call rate once again became the main operating tool for policy, though it was set at zero per cent. The pace of government bond purchases was scaled back significantly and CABs reduced quickly.

The onset of a crisis

By 2007 it began to become apparent that a number of large financial institutions were heavily exposed to the now souring US mortgage market. In July 2007, Bear Stearns was forced to liquidate two of its hedge funds and a month later, BNP Paribas stopped withdrawals from three of its investment funds, both due to losses sustained on sub-prime mortgage investments. In September the British bank, Northern Rock sought liquidity support from the Bank of England as it became the first UK bank in 150 years to face a bank run. These events were dwarfed in September 2008 when American investment bank Lehman Brothers was allowed to go bankrupt.

This financial turmoil and the uncertainty over who was exposed to what caused spreads in money markets to increase and the amount of interbank and wholesale lending to contract heavily. The day of the Lehman Brothers bankruptcy, LIBOR-OIS spreads, a measure of perceived risk in the banking sector, spiked at over 350 basis points. This was significant because over the Great Moderation, commercial banks had become increasingly reliant on short-term funding obtained in wholesale interbank markets. By 2007, 50% of UK banks' funding came from wholesale depositors. As part of their maturity transformation role, banks borrow short and lend long, and as they found it harder to roll-over their short-term funding they risked becoming insolvent through illiquidity.

Lender of last resort and liquidity provision

In 1873, Walter Bagehot wrote the now world-renowned book Lombard Street, Bagehot (1873). In it he described the mechanics of a financial crisis which was eerily prescient of the world economy in 2008. His remedy was for the central bank to be ready to lend freely and in an unlimited supply to all solvent institutions who could post good quality collateral, but at a penalty rate, so as to avoid inherent moral hazard. In doing so, the central bank would quell any panic in the market and may ultimately not have to advance any loans at all. This policy prescription became known as the lender of last resort role of the central bank, and as credit markets tightened in 2007 and 2008, the major central banks all undertook a range of policies aimed at fulfilling this role in one way or another.

The ECB supplied liquidity to its banking system merely by modifying the way it implemented its traditional open market operations. On the 15th October 2008 they announced that the main refinancing operations (MROs) and long-term refinancing operations (LTROs) would no longer be carried out by auctioning a fixed quantity of funds. Instead, they would employ a fixed rate, full allotment system under which they would supply funds in an unlimited quantity perfectly elastically at the designated policy rate. The quantity of loans was then determined by the demand of the banking sector, and limited only by the amount of eligible collateral each borrower had. This saw the ECB's balance sheet expand by roughly 500bn euros, to just over 2 trillion euros. This applied Bagehot's idea of lending freely, and against good collateral, but neglected to charge a penalty rate of interest.

The Bank of Japan also provided funds perfectly elastically to its banking system. In December 2008 it introduced its own version of full allotment, the Special Funds Supplying Operations. Providing counterparties had eligible collateral they could borrow for 3 month terms at the Bank of Japan's uncollateralised overnight call rate.

On the 12th December 2007 the Federal Reserve launched its Term Auction Facility (TAF) which offered loans to any depository institution which was financially sound. These loans were offered at a longer term than traditional borrowing from the Federal Reserve and could be secured against a wider range of collateral.[2] By requiring that borrowers be financially sound and charging a penalty above the market rate, TAF was close in spirit to Bagehot's principles. The policy proved effective and at its maximum size in March 2009 $500bn of funds were extended under TAF. When the crisis intensified with the collapse of Lehman Brothers, the Fed supplemented TAF with the Term Asset-Backed Securities Lending Facility (TALF). ABS markets are one of the key drivers of funding to the wider economy, supplying credit for all manner of activity to consumers and businesses. With this in mind, on the 25th November 2008 the Federal Reserve Bank of New York announced that, in an attempt to support

[2]At first they were loans up to 28 days, but this was extended further in August 2008 to 84 days.

the issuance of ABS, borrowers would be able to request non-recourse loans of a 3 or 5 year duration against AAA rated ABS. Initially the facility was granted permission to extend $200bn of loans, but less than anticipated demand meant that only $70bn was actually lent. Borrowers were eager to rid themselves of TALF financing as, like TAF, it came at a penalty, so as conditions in credit markets improved, many paid back TALF loans early, securing funds privately. The non-recourse nature of these loans meant that if the borrower cannot repay the loan, the collateral behind it, which can range from student loans and credit cards to small business loans or loans on commercial property, can be claimed by the Federal Reserve and sold. This had important implications for the risk faced by the central bank, helping to mitigate a lot of the risk which could potentially occur through fulfilling this lender of last resort role.

As the strain in financial markets subsided, banks could once again borrow from their regular sources and both TAF and TALF unwound naturally. TAF was closed officially in March 2010 and the last repayment of funds to TALF was made in October 2014.

In the UK, banks had been increasing their reserve targets since mid-2007. This had the effect of increasing the overall quantity of reserves in the system as the Bank of England met this heightened demand with increased supply consistent with maintaining its target for Bank Rate. However, concerned that banks still were not holding sufficient liquidity, towards the end of 2007 the Bank of England began to over supply reserves. In order to incentivise banks to hold these increased levels of reserves without the policy rate falling below the target for Bank Rate, the Bank of England started remunerating a larger range of excess reserve balances. This placed a floor on the rate of interest in money markets as commercial banks could always deposit their excess reserves at the Bank of England and receive Bank Rate, rather than lend them on the interbank market.

In December 2007 the Bank of England also offered term loans to banks against a broad range of collateral in the form of 3 month repos. These extended collateral long-term repos (eLTRs) were initially offered at a penalty rate (one higher than Bank Rate) in line with Bagehot's principles, but as financial conditions worsened this penalty rate was dropped. At first eLTRs were offered in relatively small amounts, auctions of £10bn. As the crisis deepened they were expanded significantly, as was the range of collateral against which they could be secured. At their height in January 2009 they amounted to approximately £180bn.

The Bank of England also made use of a number of off-balance sheet operations to improve the liquidity position of UK banks. The Special Liquidity Scheme (SLS) was launched on the 21st April 2008. Under SLS banks could swap high-quality assets such as mortgage-backed securities (MBS) for UK Treasury bills. The idea was that, although high-quality, market uncertainty was rendering MBS and other such assets temporarily illiquid, even though the asset itself was sound. To avoid moral hazard, only assets held by banks prior to the announcement of the policy were allowed to be used in SLS transactions, and the maximum term for which the swap could be effected was 3 years. At its height, the SLS reached £185bn, but as it was a swap of one asset for another for which a fee was charged, it never appeared on the Bank of England's balance sheet. What is more, no central bank money was created to fund SLS, so the monetary base was unaffected. The programme was financed by sales of T-bills and although these might be considered a close substitute for money at low interest rates, they do not appear in the monetary base.

In October 2008, the SLS was supplemented with another off-balance sheet measure, the Discount Window Facility (DWF). The DWF was similar to the SLS in that counterparties could borrow gilts from the Bank of England in exchange for posting collateral. A fee was due, but unlike the SLS, this fee varied with the quality of the collateral offered. This allowed the Bank of England to justify lending against a much wider range of collateral than before, as it was being compensated for the additional risk. The term of lending was initially 30 days, but in January 2009 this was extended to 364 days, conditional on an increased fee being paid. It was feared that borrowing from the DWF may send an unwanted signal to the market and banks would be hesitant to use it due to the implied stigma. To overcome this, data on the DWF was published with a considerable lag, allowing banks who needed it to borrow with a degree of anonymity.

The international nature of financial markets presented an additional complication for central banks over this period. The ability of a central bank to act as an effective lender of last resort originates from the fact that it can create money, so is not constrained in how much it can lend. However, this is only true in its own currency. The international nature of many financial institutions however means that they may be required to settle certain contracts in a currency which is not their own, and so need access to liquidity in a range of currencies. This problem was particularly acute with dollar funding in economies outside of the US. To combat this, in December 2007 the Federal Reserve opened up dollar liquidity swap lines with 14 other central banks around the world, including the Bank of England, the Bank of Japan and the ECB. These were designed to allow foreign central banks to lend to institutions in their economy, but in dollars.[3] The foreign central bank sold a quantity of its own currency to the Federal Reserve in exchange for dollars with the agreement that it would buy back that currency at a later date with an additional payment for interest. These dollar swaps were designed to leave the Federal Reserve as little exposed to risk as possible. As their agreement was with the foreign central bank they make no direct interaction with the ultimate borrower, so face little to no default risk. What is more, swap contracts stipulated that the second leg of the transaction was to be carried out at the same exchange rate as the first, therefore insulating the Federal Reserve from exchange rate risk.

These dollar liquidity lines were drawn on by all central banks that had them made available and peaked in December 2008 at $580bn. When the Euro Area crisis intensified in 2011 a number of these lines were reopened, but they were only drawn on to a maximum of $100bn.

As can be seen in figures 4-11 all of these liquidity provision policies acted to expand the balance sheets of all four central banks. In this way they could be characterised as a form of quantitative easing. However, the expansion of the quantity of money was not their primary objective, much like the change in the money supply was only a secondary concern for traditional open market operations throughout the Great Moderation, and so they are normally not described in this way.

Asset purchases to improve market functioning

As well as providing liquidity to their respective financial systems through loans, many central banks also made outright purchases of assets. The purpose of these purchases differed from the broad liquidity provision of the various lending facilities. They were designed to support the functioning of specific markets which were deemed to be particularly impaired but of importance to the functioning of the economy and transmission of monetary policy.

The Fed buys commercial paper and MBS

On the 7th October 2008, the Federal Reserve began buying high quality commercial paper via the Commercial Paper Funding Facility (CPFF). The purchases were of newly issued, non-interest paying commercial paper of 3 month maturities which were held for the full term. Agents who were dependent on

[3] As a precaution the Federal Reserve also opened currency swap lines in which it could borrow funding in other currencies, however it never drew on these.

commercial paper funding had seen interest rates in these markets rise over the financial crisis, and the term they could borrow at fall. The CPFF was designed, not only to act as an immediate source of funding for such agents, but also to instill confidence back into markets so that more private sector lenders would participate and rates would come down. The Federal Reserve charged a fee for making purchases, which placed the cost at lower than the prevailing market rate, but above the rate that would prevail under ordinary market conditions. This meant that, as with the lending facilities, the programme unwound naturally when it was no longer required as, under normal market functioning, issuing commercial paper to the CPFF was less attractive than going to the market. At its largest, in January 2009, the facility held around $350bn of commercial paper, approximately two-thirds of which was unsecured.

On the 25th November 2008 the Federal Reserve also announced it would buy $500bn of mortgage-backed securities guaranteed by the three large government sponsored enterprises (GSEs) Fannie Mae, Freddie Mac and Ginnie Mae.[4] This was designed to "reduce the cost and availability of credit for the purchase of houses which in turn should support housing markets and foster improved conditions in financial markets more generally." FOMC 2008. The purchases were funded through the creation of new reserves and so caused an expansion of the Federal Reserve's balance sheet.

The Bank of Japan buys commercial paper and corporate bonds

In March 2009 the Bank of Japan supplemented its ongoing monthly purchases of government bonds with a series of purchases designed to alleviate strains in particular markets, specifically ¥3 trillion of commercial paper and ¥1 trillion of corporate bonds. In relation to the Bank of Japan's balance sheet, and even to their annual purchases of JGB, these interventions were tiny.

The Bank of England sets up the Asset Purchase Facility

The Bank of England also looked to buy commercial paper and corporate bonds in early 2009. On the 19th January, following discussion with the Bank of England, the UK Treasury established a special purpose vehicle, the Asset Purchase Facility (APF), which would buy private sector assets. These assets were to be funded by the issuance of debt by the Treasury, much like the DWF, and so would effectively be sterilised and would not affect the size of the Bank of England's balance sheet.

[4] The Fed would also purchase $100bn of debt owed by the 3 GSEs.

At this stage, the APF was tasked with buying private sector assets to maintain market functioning as and how it saw strains develop, but up to a maximum of £50bn.

Policy rates hit the lower bound

Throughout this first stage of the crisis, the lender of last resort role of central banks had few implications for how they pursued their more permanent macroeconomic stabilisation objectives. As economic activity weakened and inflation fell below target, central banks began to cut their policy rates in order to stimulate demand. However, by early 2009, policy rates had been cut to such an extent that they were considered at, or approaching their effective lower bounds, figure 13. Despite these historically low rates of interest, many economies continued to contract and price growth failed to pick up. With their traditional instrument constrained, policymakers now looked to their balance sheets not only for liquidity provision, but more innovatively as a way to stimulate activity and generate inflationary pressure.

Asset purchases to boost demand

Quantitative Easing in the UK and US

The Bank of England and the Federal Reserve took very similar approaches in employing their balance sheets to stimulate demand.

In February 2009, the then Governor of the Bank of England, Mervyn King, requested permission from the Chancellor of the Exchequer to use the APF "for monetary policy purposes". This was duly granted and on the 5th March 2009 the Bank of England announced it would buy £75bn of UK government securities with a remaining maturity of between 5 and 25 years. These purchases would be funded by the creation of new central bank money, expanding the Bank of England's balance sheet and the monetary base. The initial focus of the policy was its impact on the money supply, and it was framed by King as "similar to the current implementation of monetary policy, except the instrument of policy would shift towards the quantity of money provided rather than its price" King (2009). In this respect it was similar to the policy change made by the Bank of Japan in 2001. An important difference was that unlike the Bank of Japan, who had bought its assets almost exclusively from the banking sector, the APF would buy gilts from non-bank financials. This mattered because one of the chief criticisms of the original Japanese policy was that banks used the additional funds to mend their balance sheets and hoard liquidity and in doing so failed to expand lending. This limited the increase in broad money, and therefore the impact on wider economic activity through the monetarist channel was severely dampened. By buying assets directly from the non-bank private sector the APF would expand the broad money supply and the cash holdings of economic agents directly, almost regardless of the decisions of the banking sector.

The policy was extended a further 3 times; first to £125bn in May 2009, then £175bn in August, and lastly to £200bn in November. The August extension was accompanied by a widening of the gilts which were eligible to be sold to the Bank to any gilts with a remaining maturity of greater than 3 years. The evolution of the policy can be seen in figures 11 and 12. On the asset side, the APF is included in the Other Assets section, as officially what was held on the Bank of England's balance sheet was the loan made to the APF, while the APF itself held the bonds which had been purchased. However, it is still possible to see the expansion throughout 2009 which more than offset the drawdown on the liquidity provision programmes. The corresponding liability can be seen on figure 11 as the quantity of reserves increases dramatically.

The Federal Reserve had to all intents and purposes begun a quantitative easing programme with its purchases of MBS, but its announcement on the 18th March 2009 that it would increase these by \$750bn and supplement them with \$300bn of purchases of US Treasuries of medium maturity signified an important change. Unlike the MBS purchases, which were designed to work in a specific market, the Treasury purchases were aimed at bringing rates down more generally in the economy. Unlike the Bank of Japan's QE though the focus was not on the expansion of the money supply and the central bank's liabilities, but more on the asset side of the balance sheet, building on the imperfect substitution and portfolio rebalancing theories of Tobin (1969) and Modigliani and Sutch (1966) to motivate a relationship between changes in the relative supplies of assets and their prices.

Adjusted for the rate of exchange, \$300bn was equivalent to the Bank of England's gilt purchases, but as a share of the available market they were drawing from it was considerably smaller. £200bn represented 29% of the free-float of UK gilts, while \$300bn was just a little under 5% of the outstanding supply of US Treasuries. Combining the three asset classes purchased in the Federal Reserve's first large-scale purchase programme they bought 14.5% of the total market of

US Treasuries, MBS and agency debt.

Interestingly, despite the creation of new reserves to fund the large-scale asset purchases, the Federal Reserve's balance sheet failed to expand noticeably in 2009. This was because, as with the Bank of England, the introduction of quantitative easing coincided with the drawdown of the liquidity programmes as the first stage of the crisis passed. The possibility exists that the introduction of cheap funding through large scale asset purchases hastened the withdrawal from liquidity programmes as agents substituted one form of liquidity for another.

Extensions

QE2

In the US, the recovery which had been seen at the start of 2010 had waned as the year went on and by the middle of the year there were worries about its sustainability. In his speech to the Jackson Hole conference on 27th August 2010 FOMC chairman Ben Bernanke laid the groundwork for expanding the Federal Reserve's asset purchase programme, Bernanke (2010), and on the 3rd November 2010 it was announced that the Federal Reserve would purchase a further $600bn of longer-dated US Treasuries over a period of 8 months. As with the original LSAP these would be funded by the creation of new reserves and so represented a large increase in the Federal Reserve's balance sheet.

The Maturity Extension Programme

The first half of 2011 saw the US economy pick up again, but as the mid-year slump struck again the Federal Reserve looked to loosen policy once more. Against a political backdrop of bitter rows over US Federal borrowing and the remit of the Fed itself, the FOMC was hesitant to expand its balance sheet further so instead in September 2011 they announced they would buy $400bn of longer dated Treasury securities funded by equivalent sales of short-term securities with the aim of increasing the average maturity of its securities portfolio from 75 months to over 100 months. These sterilised purchases left the size of the balance sheet unchanged and acted as pure debt management, affecting the composition of the Federal Reserve's and therefore the private sector's debt portfolio. By lowering the supply of longer dated debt in the publicly available market, especially relative to shorter-term debt, the Maturity Extension Programme (MEP) was expected to lower rates at the long end, flattening the yield curve. The MEP was as much a throwback as a new innovation in policy as it harked back to the Operation Twist policy carried out by the Federal Reserve and the US Treasury in the 1960s.

Open-ended asset purchases

In 2012 the US economy experienced its now seemingly annual mid-year slump and on the 13th September 2012 the Federal Reserve announced a return to balance sheet expansion. This time however, rather than announce an ultimate quantity of purchases, they determined to buy assets at a set pace until the economy no longer warranted monetary expansion, or specifically until conditions in the labour market improved "substantially". This was reminiscent of the Bank of Japan's original QE programme. The Federal Reserve's open-ended QE3 programme would buy $85bn of assets each month, of which $45bn would be US Treasury securities and $40bn MBS.

Tapering

In one of his last acts as Chairman of the FOMC, on 18th December 2013 Ben Bernanke announced that the pace of asset purchases would be reduced by $10bn in January 2014 to $75bn. He also announced that this reduction was expected to continue by a further $10bn a month each month until the purchases had tapered away to zero. This path for policy was honoured by the incoming Chairwoman Janet Yellen and by October 2014, no more purchases were being made under QE3. It is worthy of note here that even under the taper the Federal Reserve's balance sheet was expanding, just at a slower rate than previously. In fact, over the taper period, the Federal Reserve bought $400bn of assets. In this way, although it was considered the first step towards the Federal Reserve exiting QE, it was really a continuation of monetary easing.

The Bank of England expands QE

In the UK the Bank of England chose to restart its asset purchase programme in October 2011, despite inflation being around 3 percentage points above the 2% target. The rapid price growth was attributed to a number of temporary factors, such as the high oil price, which could not be controlled by the Monetary Policy Committee (MPC) and would drop out before the 2 year horizon with which the MPC were concerned. Meanwhile the underlying performance of the UK economy was weak. It was feared that this weak demand would weigh down on inflation in the medium term, especially once the temporary factors had dropped out of the calculation. Coupled with this, the coalition government elected in May 2010 had begun a programme of fiscal contraction which was expected to weigh down on growth further. In light of these headwinds the Bank of England looked to loosen policy.

The second round of asset purchases followed the blueprint laid out by the first. The APF purchased UK gilts from the secondary market financed by the creation of new reserves. The original announcement was for an additional £75bn of purchases, but this was subsequently extended by £50bn in February 2012 and a further £50bn in July, bringing the total size of the APF, accounting for QE1 and QE2 to £375bn.

Japan returns to quantitative easing

Comprehensive monetary easing

Faced with falling prices, on the 5th October 2010 the Bank of Japan unveiled a multifaceted comprehensive monetary easing (CME). This involved a number of elements, each designed to stoke inflationary pressure. The first was a largely symbolic cut in the uncollateralised overnight call rate target from 0.1% to 0-0.1%. More significantly, this was accompanied by an increasingly explicit clarification of the conditions required for the Bank of Japan to raise interest rates from virtually zero.[5]

The final element of the CME was an asset purchase programme (APP) through which the Bank of Japan would buy assets from a much wider range of asset classes than had been done by any central bank to-date. These included Japanese government bonds, commercial paper, corporate bonds, exchange traded funds (ETFs) and Japanese real estate investment securities (J-REITs). The APP was initially set at ¥35trn, but ¥30trn of this had already been announced under other programmes. Of the additional ¥5trn, ¥3.5trn was JGBs with the rest made up of eclectic private sector assets. The APP was only supposed to be a temporary measure, however it was extended 8 times between October 2010 and the end of 2012, from the original ¥35trn to ¥101trn. Almost all of this further expansion came through purchases of JGBs.

The Abe government and QQE

In December 2012 Japan elected Shinzo Abe as Prime Minister. The Abe government quickly set about an extensive set of policies to stimulate the Japanese

[5]See later section on forward guidance.

economy, a central tenant of which was more aggressive monetary activism. In January 2013, the Bank of Japan was given an explicit target to achieve 2% inflation within 2 years. By March 2013 the Bank of Japan had a new Governor, Haruhiko Kuroda, who had been previously outspoken about the need for additional monetary stimulus in Japan, and a month after being appointed he unveiled a massive Quantitative and Qualitative Easing programme (QQE). Under QQE the Bank of Japan officially changed its primary policy objective from the overnight call rate to the quantity of money in the system. By making ¥60-70trn worth of asset purchases a year QQE would double the monetary base in two years. The majority of the assets purchased, ¥50trn a year, would be Japanese government bonds. Importantly the majority of these bonds would be considerably longer-dated than under the Bank of Japan's previous programmes and the average maturity of their government securities portfolio would rise from around 3 years to 7 years, in line with the average maturity of the entire JGB market. This was intended to lower longer yields more by coupling the supply effects of the previous programmes with a complimentary maturity effect (see Meaning and Zhu (2012)). The remainder of the monetary expansion would be accomplished through purchases of ETFs, J-REITs and loans to banks. Although the pace of purchases was laid out for two years in line with the horizon for hitting the new inflation target, crucially QQE was open-ended, meaning that the Bank of Japan committed to carry on expanding the monetary base until inflation was stable around 2%.

QQE - a second boost

The start of 2014 saw inflation begin to pick up in Japan. Alongside monetary activism, Prime Minister Abe introduced a VAT increase in April 2014 to generate inflationary pressure. This had the effect of bringing forward households' consumption plans as they sought to spend before the tax rise took effect. This front loading spurred demand and fuelled a return to positive inflation. An artifice of the tax rise itself was that inflation jumped to 3.4% in April 2014 compared with 1.6% a month earlier in March. However, higher prices with no significant pick up in wage growth, meant that households' spending power deteriorated, and as 2014 wore on and the fall back from the front-loaded consumption took effect demand looked to be weakening. Inflation remained elevated as a result of the tax increase but it was feared that once this fell out of the calculation, inflation would fall back to zero.

So, on the 31st October 2014 the Bank of Japan moved preemptively to counter this, expanding the pace of the QQE programme to over ¥80trn a year. Almost all of this additional monetary expansion would be through increased purchases of JGBs and the Bank of Japan increased the target maturity for its portfolio further from 7 years to 7-10 years. Purchases of private sector assets increased slightly, but still made up just ¥3trn of the overall ¥80trn.

The ECB's road to quantitative easing

The European Central Bank has had a conflicted relationship with quantitative easing over the last 6 or 7 years. Vocal opposition from a number of member states, but most notably Germany to monetary expansion has at times appeared to tie the hands of the ECB's governing council. This has meant that the design and emphasis of the ECB's intervention has been quite different to the other central banks discussed above.

Corporate bond buying programme

July 2009 saw the ECB undertake its first programme of outright asset purchases. The Covered Bond Purchase Programme (CBPP) was more like the CPFF than the large scale asset purchase programmes of the Federal Reserve and Bank of England in that it was aimed at helping banks to secure longer-term funding in an impaired credit market. Between the 2nd July 2009 and 10th June 2010, the CBPP purchase 60bn euros of covered bonds. Although these purchases were not sterilised, the ECB was keen to stress that the CBPP was not expected to expand the balance sheet in a material manner as banks would use covered bonds to substitute for other forms of central bank finance such as LTROs. As such there was no quantitative easing element to the CBPP and it was instead credit easing, changing the composition of balance sheets rather than their size.

The Securities Market Programme

In 2010, as the other major economies were coming out of the worst of the financial turmoil, the Euro Area was faced with a new challenge. Over the Great Moderation large fiscal imbalances had built up in the Euro Area with periphery economies persistently running large deficits whilst Germany and other northern European economies ran surpluses. This was exacerbated by the banking crisis and recession of 2008/2009 which necessitated large increases in borrowing to support the financial system and fund the automatic fiscal stabilisers.

Markets began to question the sustainability of large debt and deficit ratios, which caused yields on Euro Area periphery debt to increase, intensifying the problem. The financial crisis had evolved into a sovereign debt crisis.

The ECB responded by launching the Securities Market Programme (SMP) on the 10th May 2010. Under the SMP the ECB would buy sovereign debt in the secondary market, much like the quantitative easing programmes of the other three central banks. The SMP differed in two crucial aspects. First, purchases were made on an ad hoc basis as deemed necessary, rather than the pre-announced purchase schedules employed by other central banks. Second, the purchases were sterilised, meaning that the extra liquidity provided by SMP purchases was offset by the issuance of short-term (one week) interest-bearing deposit certificates by the ECB. As these fine-tuning operations only lasted one week though they had to be continually renewed. The result of this was that the quantity of liquidity in the system and the size of the ECB's balance sheet were left unaffected. This was because the stated aim of the SMP was not to expand the monetary base, in fact it was explicitly not a loosening of the monetary policy stance, rather it was intended to alleviate strains in particular markets, in this case sovereign debt, which might hamper the usual transmission of monetary policy.

Initially purchases were made of Greek, Irish and Portuguese government bonds, but as the sovereign debt crisis escalated this was expanded in August 2011 to Italian and Spanish government debt. In September 2012, when the policy was at its largest, the SMP held 218bn euros of Euro Area sovereign bonds, approximately half of which were Italian. Although no more purchases were made after this point, the assets which had been bought were held to maturity and at the time of writing still total around 140bn euros.

Balance sheet expansion through LTROs

Although the SMP and the ECB's other asset purchase programmes had been designed to be balance sheet neutral, by the start of 2012 the ECB's balance sheet had expanded from roughly 1trn euros in 2006 to over 3trn euros. This tripling was almost entirely a result of growth in long-term refinancing operations, figure 6. These loans to banks had long been part of the ordinary conduct of monetary policy, but crucially from 2008 onwards their quantity was demand determined. As long as banks could post eligible collateral and were willing to pay the refinancing rate then the ECB would provide them funds. As credit conditions in the market tightened and uncertainty rose, demand for secure funding surged, and with it so did the central bank's balance sheet. What is more, the ECB repeatedly extended the assets it would accept as collateral and the term it would lend at, inviting more demand. In essence LTROs were a form of monetary expansion conducted through lending rather than outright purchases of assets.

Outright Monetary Transactions and "whatever it takes"

The ECB continued to focus its outright asset purchases on promoting good market functioning. In November 2011 it reopened its corporate bond buying programme (CBPP2) this time buying just 16.4bn euros of Euro Area corporate bonds.

However, by the summer of 2012 the sovereign debt crisis had intensified. Spreads on Euro Area periphery bonds had rocketed and there was a growing expectation that such debt servicing costs, and perhaps even the single currency itself were unsustainable. The ECB's answer was a programme of Outright Monetary Transactions (OMTs). OMTs would replace the SMP and were designed to "eliminate the unwarranted and self-fulfilling fears of a Euro Area break-up" Coeure (2013).

Like the SMP, OMTs would consist of purchases of Euro Area government bonds in the secondary market. Also like the SMP, these purchases would be sterilised. However, unlike the SMP, OMT purchases were subject to conditionality. The state whose bonds were being purchased had to be in receipt of aid from one of the Euro Area's bailout funds. These bailouts came with their own conditions on structural reform and to qualify for inclusion in the OMTs, the state in question had to adhere to their bailout terms. The economy must also prove it was capable of raising funds in private markets by successfully issuing at least a 10 year government bond. If all of these conditions were met, and the ECB deemed it necessary, then purchases could be made of the state's sovereign debt with a residual maturity of 1-3 years.

An important design feature of OMTs was that the ECB ranked pari-passu with other creditors in the event of a default. This meant they received no special treatment and if the sovereign defaulted they would lose alongside the private sector. This was designed to instill confidence in the market that the ECB would not allow a sovereign to default on its debt.

Perhaps the most crucial aspect of OMTs though was that there was no ex-ante quantitative limit. When the policy was announced Mario Draghi, the ECB Governor now famously said the ECB would do "whatever it takes to save the euro". This commitment in of itself was a powerful signal to the market and was undoubtedly effective as yields on periphery debt fell significantly and stayed down for much of the next 12 months despite not a single government bond ever being purchased under the OMT programme. By committing to step in an unlimited capacity should the need arise the ECB removed the need to intervene at all.

Targeted Long-term Refinancing Operations

By 2014 the worst of the sovereign debt crisis looked to have been weathered. The SMP and Draghi's "whatever it takes" commitment seemed to have reassured markets that there would be no break up of the currency area and yields on Southern European debt remained low. As the financial turmoil subsided, so did banks' demand for funding from the ECB and the quantity of LTROs began to reduce steadily until by the start of 2014 the central bank's balance sheet was around 2trn euros.

As the year progressed however there still remained significant weakness in credit availability, especially in the Southern European economies that needed it most. To counter this asymmetry and get credit flowing, in June 2014 the ECB announced it would offer targeted long-term refinancing operations (TLTROs). These differed from their ordinary LTROs by coming with conditions linking their cost and availability to the amount of lending banks made to the real economy. In this respect it was very like the Funding for Lending Scheme (FLS) set up in the UK.[6] There were to be two initial operations in September and December 2014 where Euro Area banks could borrow funds up to 7% of their stock of loans to the non-financial private sector, measured on 30th April 2014. In aggregate this amounted to a total potential size of around 400bn euros. Then in 2015 and 2016 the banks could borrow funds each quarter equivalent to 3 times their net lending to the private non-financials (PNFCs). This meant that the more banks lent to the real economy, the more they were rewarded with cheap term funding. It would seem that cheap funding was not what was constraining Euro Area banks as demand for the first two TLTROs was weak at just 82.6bn euros in September and 130bn euros in December. This meant that only just over half of the funds available through TLTROs were drawn upon. What is more, net lending to Euro Area PNFCs has failed to increase substantially.

Also at its June 2014 press conference, the ECB announced it would no longer continue to sterilise the assets held under the SMP. This meant it would halt the

[6]For a more detailled discussion of the FLS, see http://www.bankofengland.co.uk/markets/pages/fls/default.aspx

weekly fine-tuning operations which absorbed the funds created by the SMP and thus increase the amount of cash held by the private sector by the size of the programme.

Asset-backed Securities Purchase Programme and further corporate bond purchases

Inflation in Europe failed to pick up. In fact it deteriorated further, and by September 2014, consumer price inflation had fallen to 0.3% having been below 1% for 12 consecutive months. With inflation so far from 2% the ECB took action to stimulate the economy. Hesitant to embark on purchases of government debt, the ECB introduced two new sets of private sector asset purchases. The first, launched on 20th October 2014 was a second extension of the Corporate Bond Purchasing Programme (CBPP3). The second, unveiled a month later was the Asset-Backed Securities Purchase Programme (ABSPP) in which the ECB will buy ABS from Euro Area banks. The hope is that by taking ABS off the balance sheets of banks and giving them cash, the banks will be incentivised to use that cash to extend new credit to firms which the banks can once again securitise, rebalancing their portfolios. Both CBPP3 and ABSPP are intended to run at least until September 2016.

Quantitative Easing - The Public Sector Purchase Programme

The lack of inflationary pressure in Europe at the end of 2014 was exacerbated further by an unexpected fall in the world oil price of over 50% between November and January 2015. As this pushed the Euro Area into deflation there were growing signs that inflation expectations at the 2 year horizon were beginning to decouple from the level consistent with the ECB's mandate. With almost no option but to respond the ECB announced on the 22nd January 2015 that it would finally begin a programme of large-scale purchases of Euro Area member states' sovereign debt. The purchases would be carried out by the national central banks of member states, but would be funded by the printing of new central bank money by the ECB. What is more, the programme was open-ended, but expected to run until at least September 2016. The purchases themselves began in March 2015, targeting a pace of 60bn euros a month. The purchases would be spread across member states, divided so as not to take too much supply from any one particular sovereign market. Importantly, Greek government bonds were not eligible to be purchased under the scheme as the ECB had significant concerns about the solvency of the Greek government.

The particular institutional arrangements of the Euro Area have thrown up some unique questions for Euro Area QE. The appropriation of risk is clearly a concern. By keeping the majority of risk with the national central bank, the ECB is avoiding exposing taxpayers in Northern Europe to the default risk associated with Southern Europe. However, by eliminating such risk sharing they are perhaps hobbling a significant channel through which the programme could lower yields and stimulate demand.

Forward Guidance

As well as measures which affect the balance sheet, all four central banks considered in this paper have made use of another policy tool - forward guidance. In essence this is the shaping of agents' expectations about the future path of policy through communication. Its theoretical foundations can be found in the canonical New Keynesian model as described in Woodford (2003) which stipulates that current demand can be represented as a function not just of the contemporaneous interest rate, but also of agents' expectations of the entire path of future interest rates. Therefore, by altering these expectations, policy makers can either loosen or tighten policy without the need for interest rates to change today. As should be clear, this is a particularly attractive result when short-term interest rates are constrained by the lower bound and so cannot be moved today. The policy prescription of the New Keynesian model in this case is to commit to hold rates lower for longer, a position explained famously by Eggertson and Woodford (2003).

Forward guidance can also be used to clarify the central bank's reaction function. This has been especially useful in recent years when the degree of uncertainty has been high. By using forward guidance in this way, policymakers are not seeking to commit to anything inconsistent with their reaction function, but just to pull agents' expectations of the future path of the economy in to line with their own.

As with quantitative easing, the Bank of Japan was a forbearer for forward guidance, using it alongside the original asset purchase programme in the early 2000s. It applied what is known as state-contingent forward guidance, conditioning the moment it would raise rates on a particular outcome. In this case it was that it would maintain the monetary expansion until "the core CPI register(ed) stably zero percent, or an increase year on year".

It then reawakened the policy in 2010 when it initiated the Comprehensive

Monetary Easing. This time however it was vague, stating only that it would maintain the zero-interest rate policy "until price stability is in sight on the basis of the understanding of medium- to long-term price stability". As this proved ineffective in generating a significant shift in agents' expectations, the guidance was hardened in February 2012 when an explicit target for price growth of 1 % was introduced. This was strengthened further with the introduction of QQE. The state-contingent target was raised from 1 % to 2 % in the hope of lifting inflation expectations, with the Bank of Japan promising to continue QQE "as long as it is necessary for maintaining that target in a stable manner". This was accompanied by a partial time-contigent element of guidance though as the policy statement also noted "The Bank will achieve the price stability target of 2 percent . . . at the earliest possible time, with a time horizon of about two years".

The Federal Reserve's forward guidance began in 2008 with the loosely time-contingent comment that rates would be close to zero for "some time". This language was beefed up marginally in March 2009 when rates would be low for "an extended period". In August 2011 though it signalled an explicit time-contigent target that rates would be low at least until the middle of 2013. This was subsequently pushed back a number of times until a state-contingent form of forward guidance was introduced in December 2012.This stated that the FOMC would not begin to raise rates until the unemployment rate had fallen below $6\frac{1}{2}$ %, conditioned on inflation staying below $2\frac{1}{2}$ % and inflation expectations remaining anchored. At the time of writing the US unemployment rate stands at $5\frac{1}{2}$ % and the Fed Funds rate is yet to be increased. The FOMC has supplemented the guidance it gives by publishing what has become known as the "dot plot", a figure which shows where each of the committee's members feel rates should be over the near and medium term.

The Bank of England adopted forward guidance officially in mid-2013 with the arrival of Governor Carney, Carney (2013). Seemingly learning from the Federal Reserve and Bank of Japan, the Monetary Policy Committee opted for state-contingent guidance from the beginning. The MPC stated that it would not even consider raising Bank Rate until the unemployment rate was below 7 %. Importantly it stressed that this was not a trigger at which points rates would automatically rise, but very much a threshold after which they would reconsider their options. They also included a series of "knock-outs" which would cause them to raise rates before unemployment of 7 %. The blurred line between state- and time-contingent forward guidance was perhaps at its clearest here, as the Bank of England produced a forecast for unemployment which, when combined with the 7 % threshold could be used to intimate when the Bank expected to begin tightening policy. What transpired showed the difficulty of such an approach as the UK labour market performed unexpectedly well and the 7 % threshold was crossed just 7 months later, not the 18 months or so that the Bank of England had forecast at its inception.

Following this the Bank of England reverted to a more general form of forward guidance based on a broad range of indicators relating to slack and the labour market. In practice they have returned to a form of guidance not dissimilar to the style of communication they used prior to 2013 with individual members talking about their view of the appropriate stance of policy and the balance of risks.

The ECB has tried to avoid forward guidance in any overt manner, seemingly for fear of constraining itself in future periods. Instead Governing Council members have spoken about their personal views on the state of the Euro Area economy and the likely policy decision one meeting at a time. However, in July 2013 Mario Draghi ventured to give the vague reassurance that rates would be low for "an extended period of time", just as the FOMC had done in 2009.[7]

Exit

As economic growth has strengthened, discussion in the US and the UK has turned from further easing to the timing and nature of exit. The current market forecast is that the Federal Reserve will move first, with forward curves pricing in a rate rise in the final quarter of 2015. The Bank of England is expected to follow shortly after, with an increase in Bank Rate around the turn of the year. In both cases, policymakers have sought to avoid surprising markets by managing expectations through regular communication in the hope that this will limit the volatility around the turn in the monetary policy cycle.

Exit presents a number of challenges to policymakers. First is the coordination of the various policy instruments. Both the Bank of England and the Federal Reserve have stated that they expect to raise rates a number of times before they turn to reducing the size of their balance sheets. However, with both policy tools eventually operating simultaneously, it will be hard to judge the actual stance of policy.

[7] For further detail on forward guidance the interested reader should refer to Bernanke (2013) for the Federal Reserve, Dale and Talbot (2013) for the Bank of England, Cœuré (2013) for the ECB and Shirai (2013) for the Bank of Japan. Filardo and Hoffman (2014) also provide an informative insight into the rationale and effectiveness of forward guidance as a policy.

Then there is the question of how exactly does one unwind such a large balance sheet expansion when the time comes. There is the option to sell assets back to the market, or alternatively one could merely hold the assets to maturity and then no longer reinvest the proceeds of the redemption. As discussed in Chadha and Meaning (2012), Kirby and Meaning (2015) and McLaren and Smith (2013) the specifics of this process will not only have implications for monetary policy, but also for the fiscal authority.

Unwinding quantitative easing will also require some skilful reserve management to drain the huge amounts of liquidity the programmes have pumped into the system with losing control of the headline policy rate. As presciently discussed in Goodfriend (2002), this may require manipulation of the rate of interest paid on reserve balances.

But even before the question of how central banks should shrink their balance sheets, one must decide if this is the correct aim. Governor Carney has repeatedly stated that he expects the Bank of England's balance sheet to remain permanently bigger than it was prior to 2007. His rationale is that there has been a structural change in the demand for liquidity and reserves which central banks will have to meet. This opens up the discussion for whether or not asset purchases could be used in normal times as a direct way of targeting movements in longer-term interest rates.

Whatever form exit eventually takes, it looks certain that balance sheets will continue to be an important feature of monetary policy for the foreseeable future.

Conclusions

The last decade has seen significant change in the manner in which monetary policy is conducted. Central banks in the advanced economies have moved from relying almost exclusively on the short-term nominal interest rate to achieve price stability to manipulating the size and composition of their balance sheets in pursuit of an expanded set of objectives. At first central banks used their balance sheets to fulfill their lender of last resort role and provide liquidity to support the financial system after the 2007/08 crisis. Then, as the lower bound began to bind, their focus shifted to quantitative and credit easing policies designed to stimulate demand and act as a replacement for the traditional policy instrument. The result has been that central banks have seen their balance sheets grow vastly compared to where they were a decade ago. Not only this but they now consist of a much more varied mix of assets.

Each central bank approached this problem in its own unique way. The Federal Reserve, where the housing market collapse had been most acute, bought both mortgage-backed securities and government debt in large quantities, lowering rates in those markets whilst simultaneously increasing both liquidity and the money supply. The Bank of England focused almost solely on UK government securities, buying up to 1/3 of the market. The Bank of Japan has expanded its balance sheet so extensively that it has diversified and bought a wide range of assets, though as with the Federal Reserve and Bank of England, a high degree of importance has been given to sovereign debt. The ECB meanwhile, more constrained by political pressures and conflicting interest groups refrained for a long time from undertaking explicit quantitative easing, preferring instead to extend loans to the financial sector and sterilising any outright purchases.

We have now reached a period of divergence in monetary policy cycles. Both the Bank of England and the Federal Reserve are on the cusp of tightening policy for the first time in over six years. This turning point will no doubt be accompanied by a degree of volatility as markets adjust, but if the central banks can communicate their reaction functions clearly then this should be limited. The ECB and the Bank of Japan on the other hand are continuing to loosen monetary conditions and expand their balance sheets through large-scale asset purchase programmes. In all cases, a return to balance sheets of the magnitude and composition to which we were accustomed to before the crisis is a long way off, and perhaps not likely to be seen at all. It is highly likely that the conduct of policy on a practical level has been changed at a deep level by the developments of the last decade.

Similarly this period of flux for practical policy has provided a whole new set of questions for academic economists, and reawakened some old, half-forgotten ones at the same time. Some of these will be the subject of the remainder of this thesis.

References

[1] Bagehot, W (1873) "Lombard Street: A Description of the Money Market" (1 ed.). New York: Scribner, Armstong & Co.

[2] Chadha, J.S and Meaning, J (2012)

[3] Eggertson, G and M, Woodford (2003)'The Zero Bound on Interest Rates and Optimal Monetary Policy' , Brookings Papers on Economic Activity,

(1), pp. 139–233

[4] Goodfriend, M (2002) "Interest on reserves and monetary policy," Economic Policy Review, Federal Reserve Bank of New York, issue May, pages 77-84.

[5] King, M (2000) "Monetary Policy: Theory in Practice" Speech 7th January 2000

[6] Kirby, S and Meaning, J (2015) "The Impacts of the Bank of England's Asset Purchases on the Public Finances" National Institute Economic Review May 2015 vol. 232 no. 1 F73-F78

[7] Kuttner, K (2014) "Monetary Policy during Japan's Great Recession:From Self-Induced Paralysis to Rooseveltian Resolve"

[8] McLaren, N and Smith, T (2013) "The profile of cash transfers between the Asset Purchase Facility and Her Majesty's Treasury," Bank of England Quarterly Bulletin, Bank of England, vol. 53(1), pages 29-37.

[9] Meaning, J and Zhu, F (2012) 'The Impact of Federal Reserve Asset Purchase Programmes: Another Twist', BIS Quarterly Review, March 2012

[10] Modigliani, F. and Sutch, R.C. (1966). "Innovations in interest rate policy", American Economic Review, Papers and Proceedings, vol. 56(2), pp. 178–97.

[11] Poole, W. (1970). "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model". Quarterly Journal of Economics 84, 197-216.

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

[13] Tinbergen J.(1952). "On the Theory of Economic Policy" Amsterdam: North Holland, 2nd edition

[14] Tobin, J (1969) "A General Equilibrium Approach to Monetary Theory," Journal of Money, Credit and Banking, 1, February, 15-29

[15] Woodford, M (2003)."Interest and prices: Foundations of a theory of monetary policy". Princeton, New Jersey: Princeton University Press.

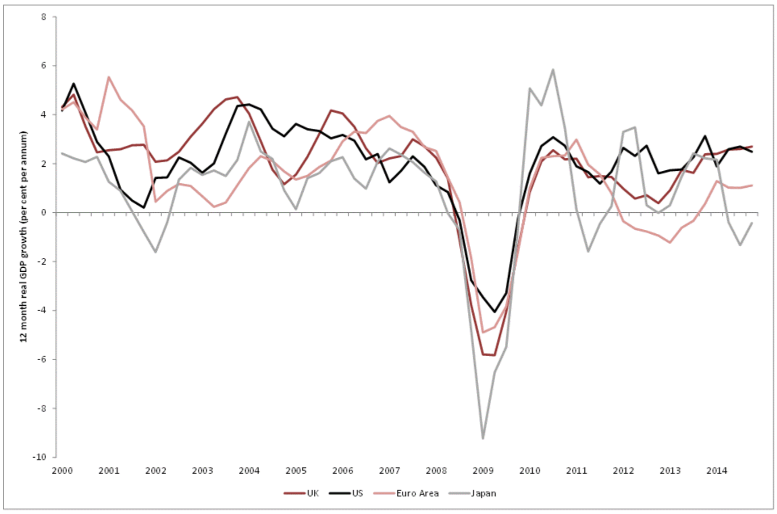

Figure 1: Real GDP growth 2000 to 2015

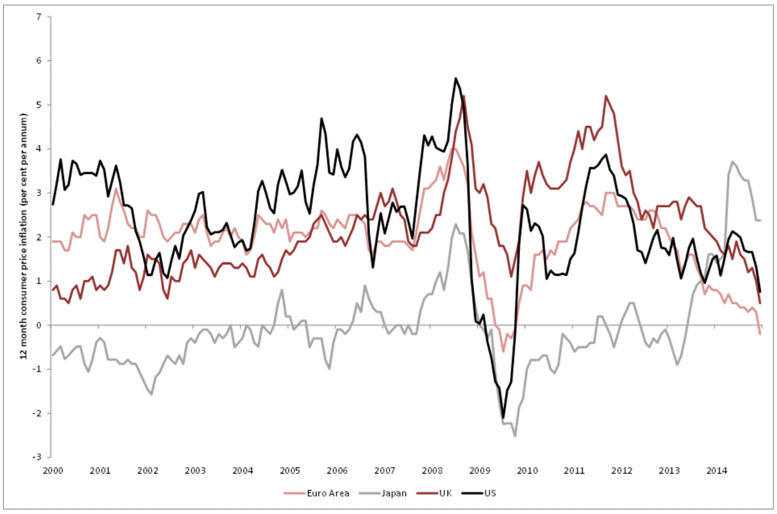

Figure 2: 12-month Consumer Price Inflation 2000 to 2015

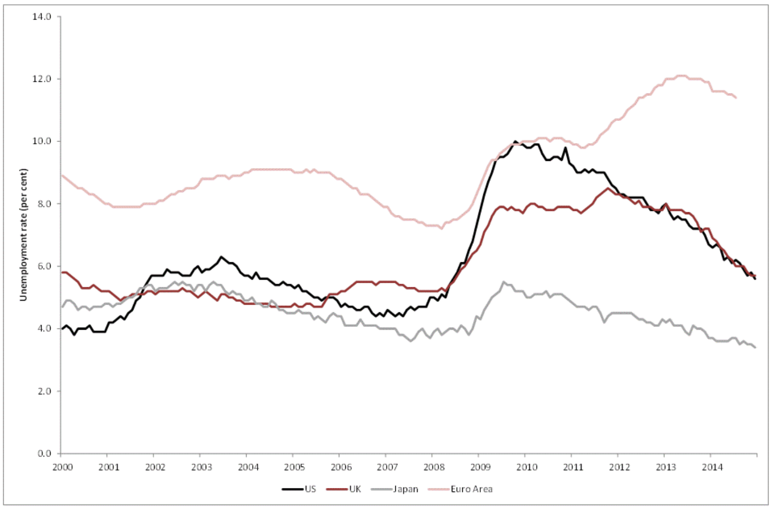

Figure 3: Unemployment Rates 2000 to 2015

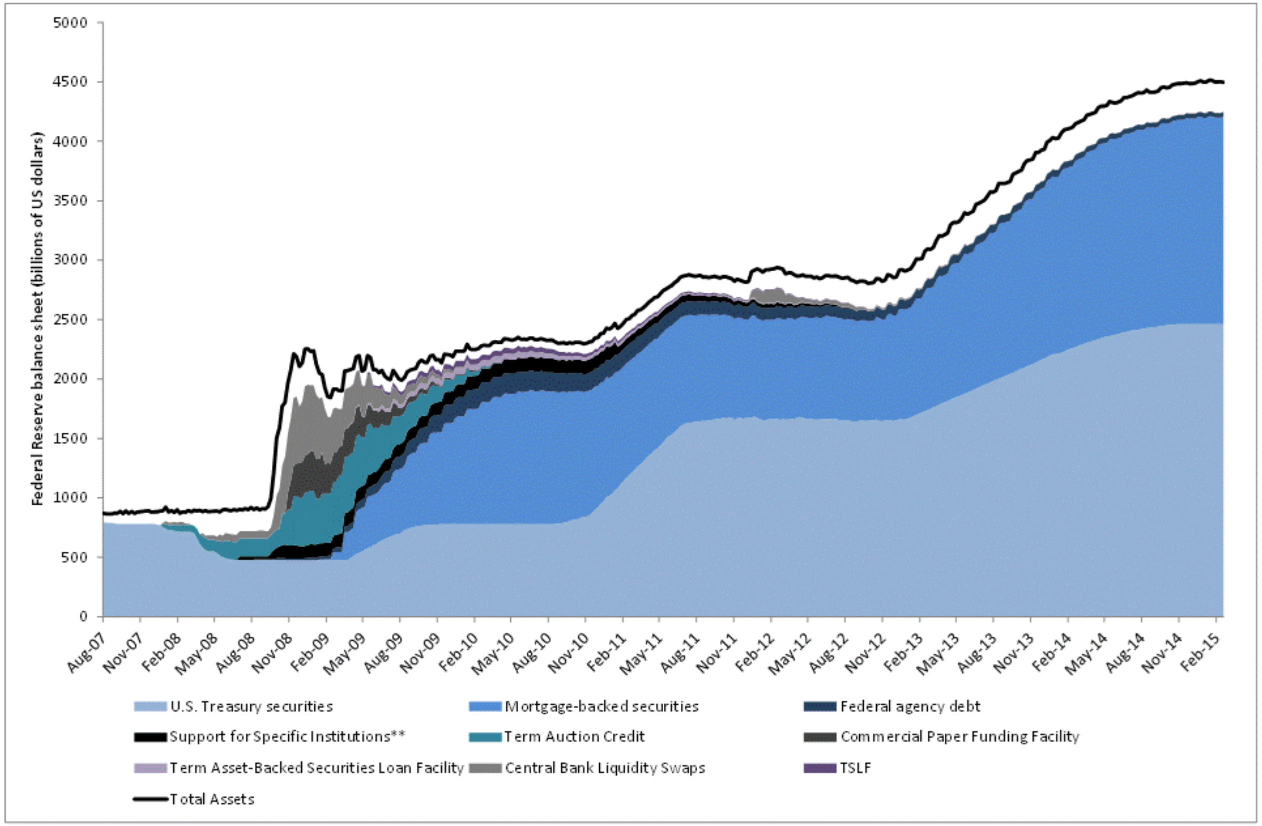

Figure 4: Federal Reserve Balance Sheet - Assets

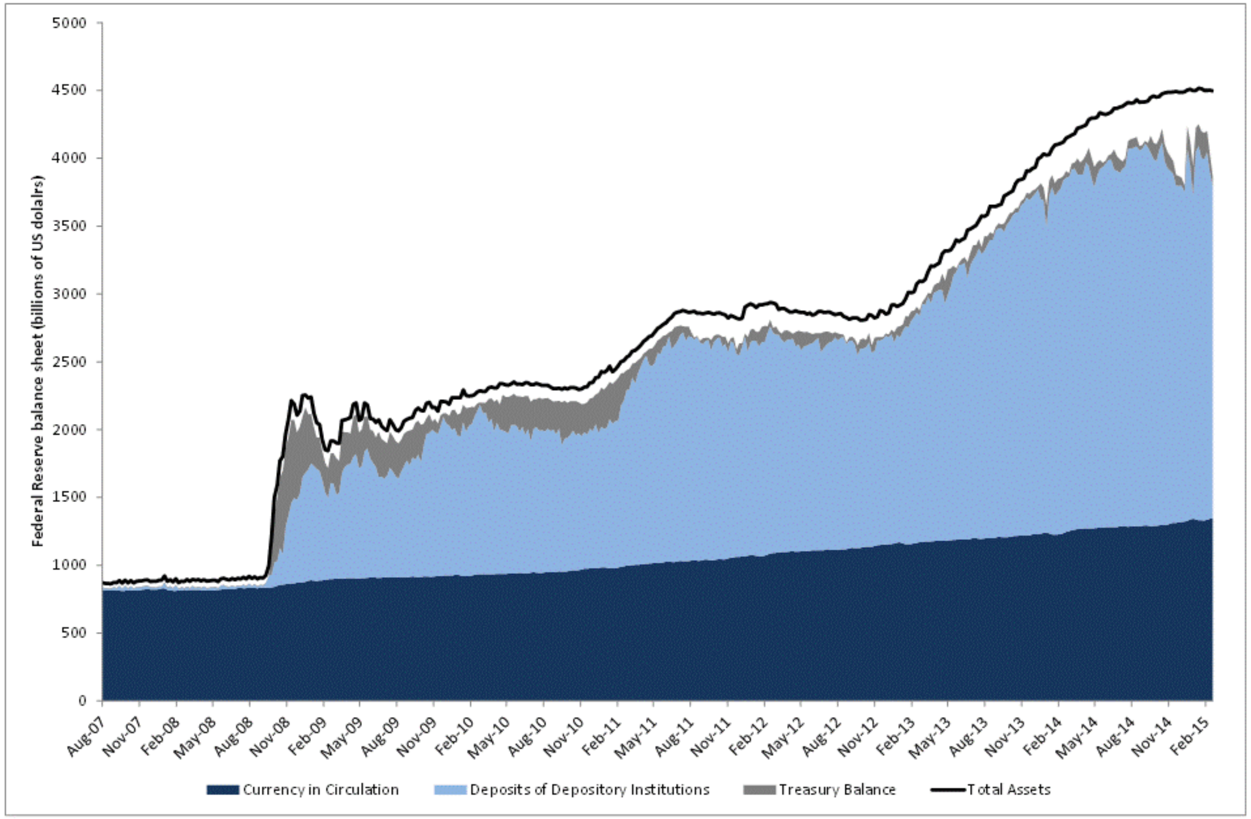

Figure 5: Federal Reserve Balance Sheet - Liabilities

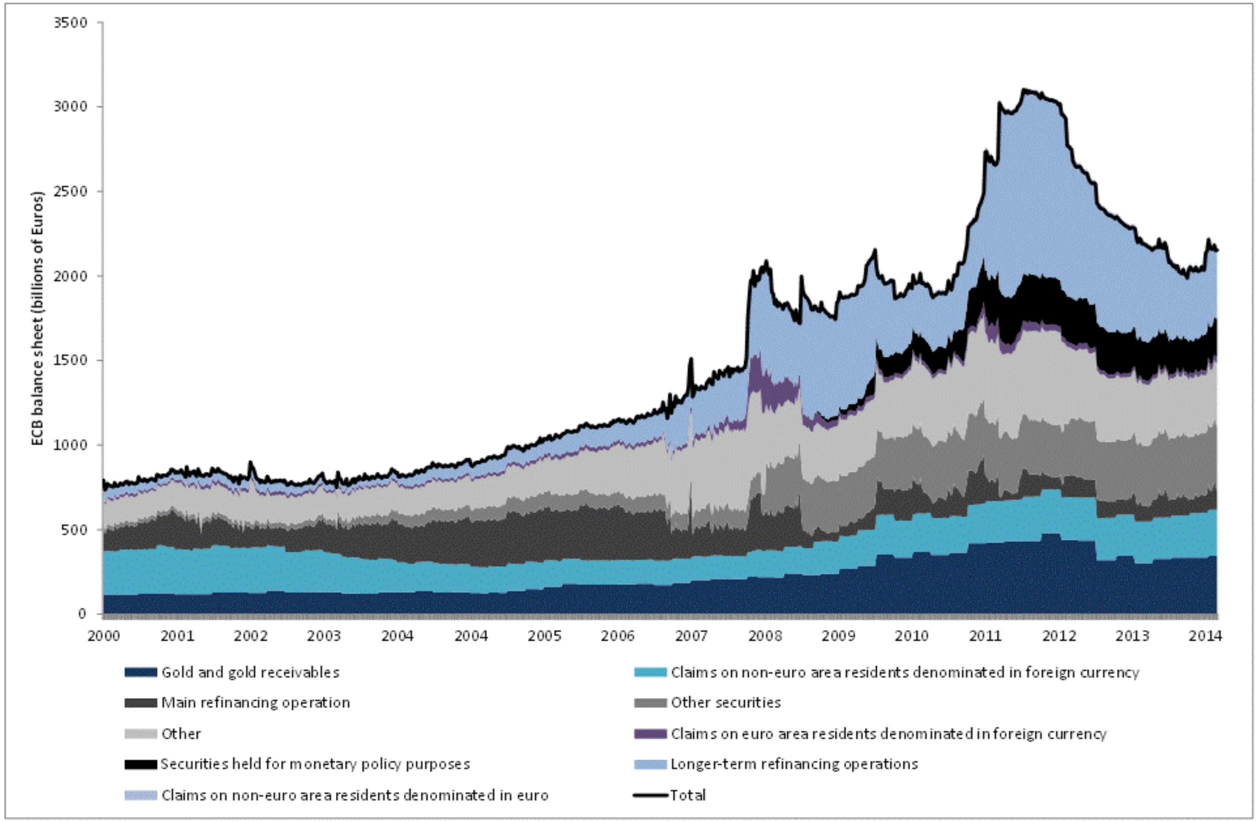

Figure 6: ECB Balance Sheet - Assets

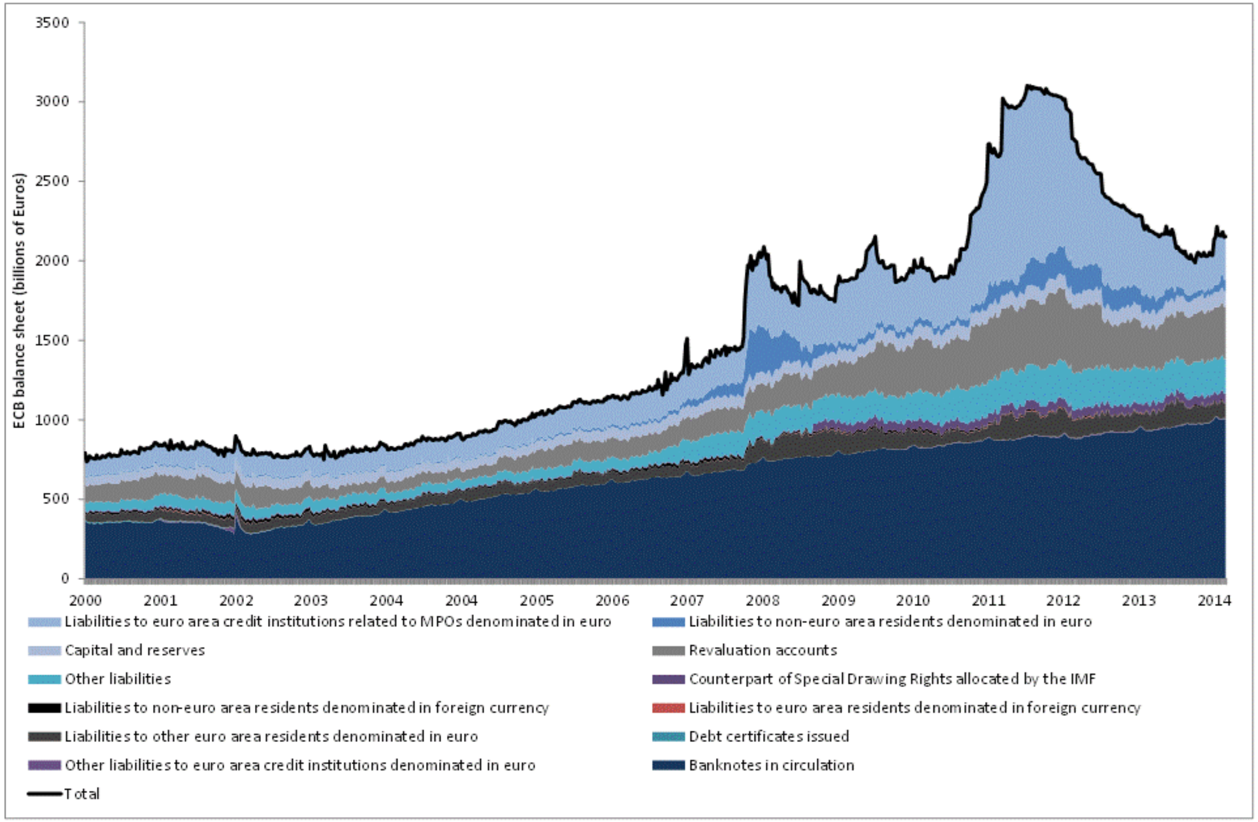

Figure 7: ECB Balance Sheet - Liabilities

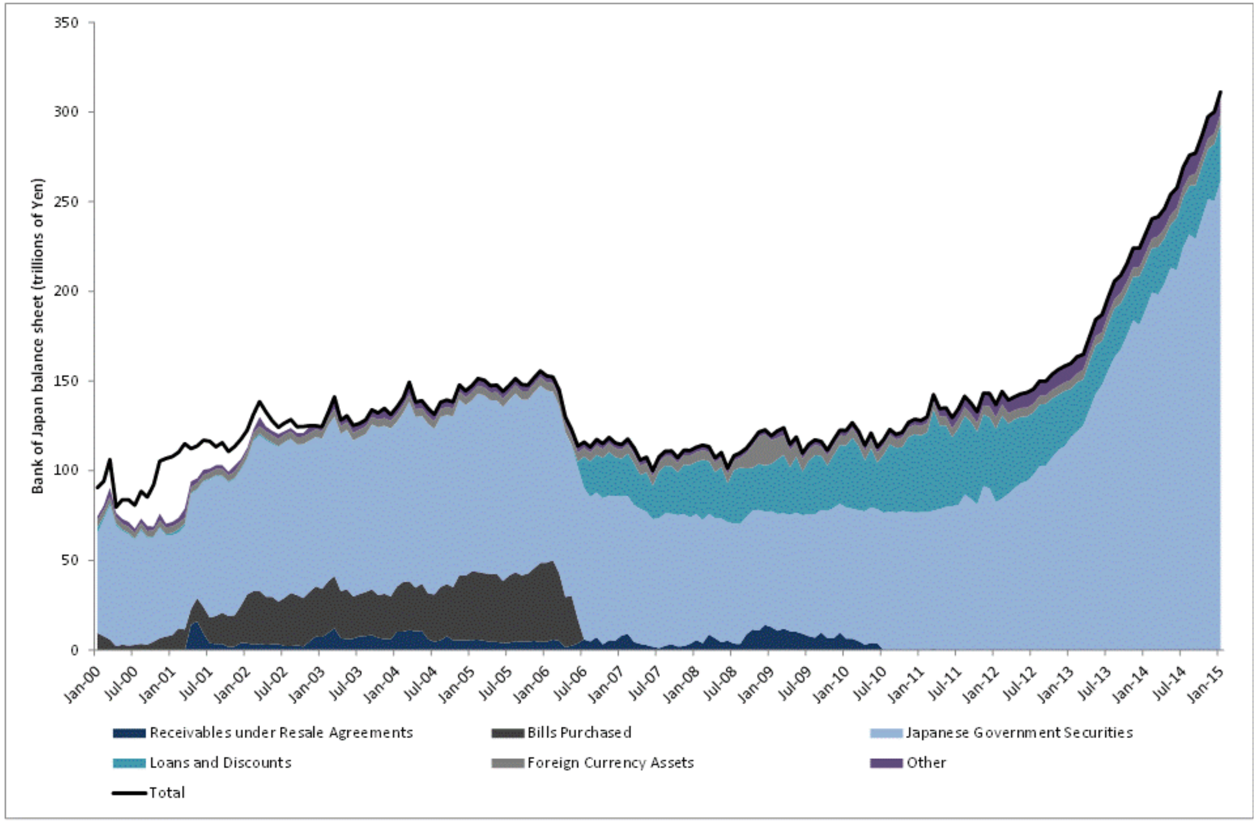

Figure 8: Bank of Japan Balance Sheet - Assets

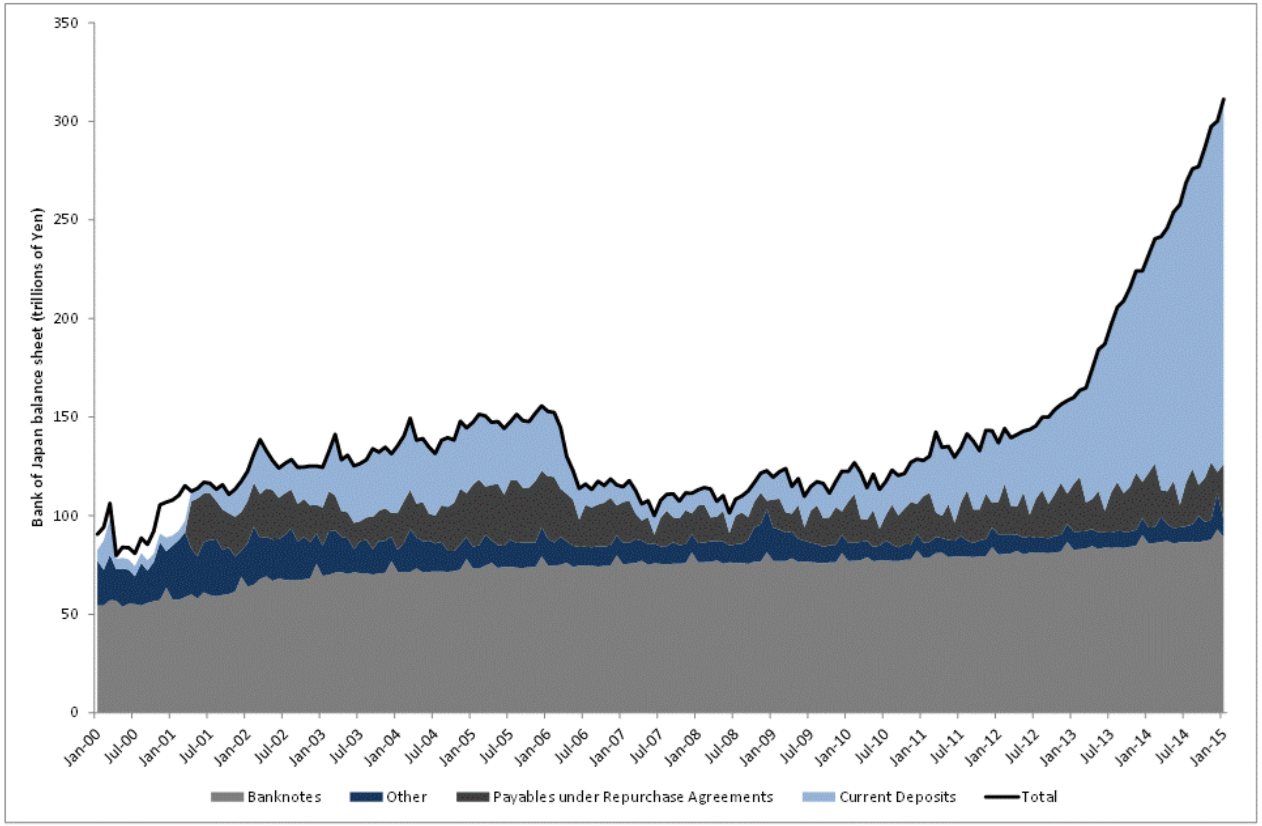

Figure 9: Bank of Japan Balance Sheet - Liabilities

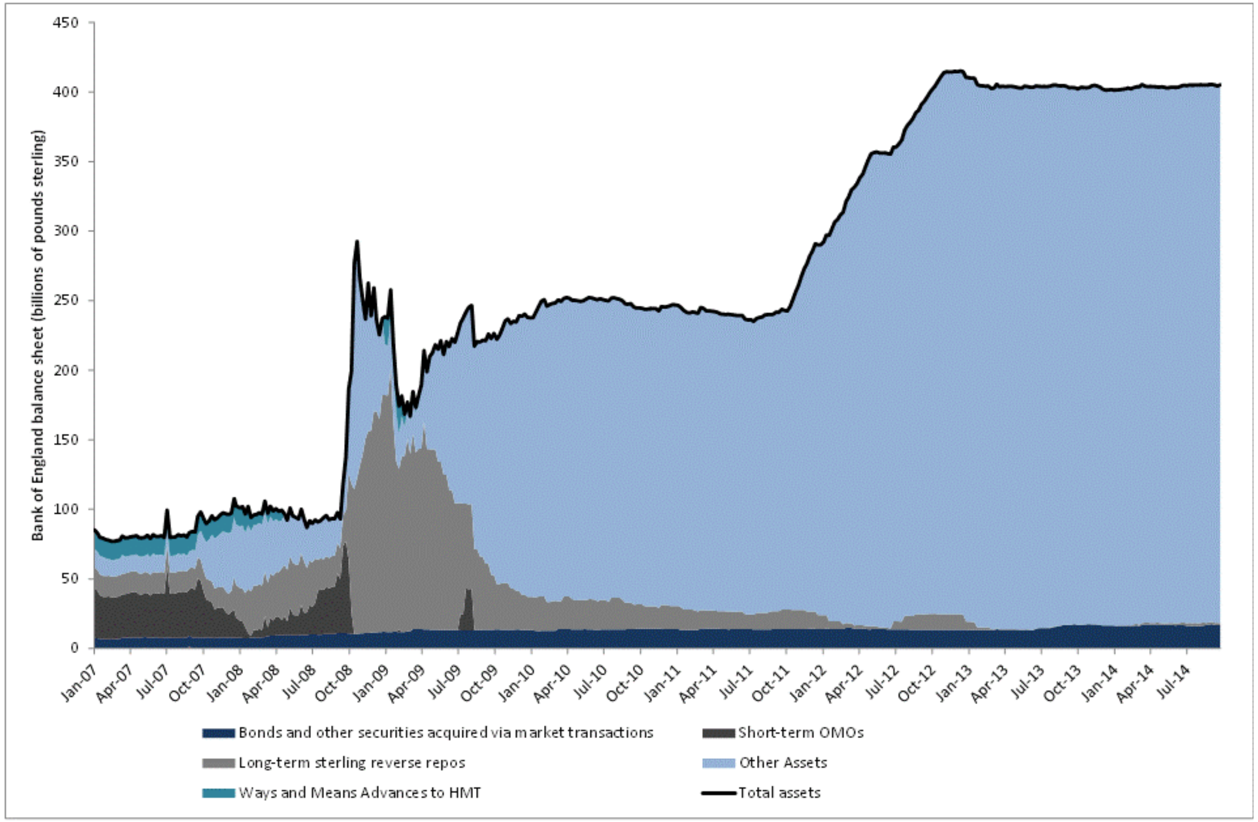

Figure 10: Bank of England Balance Sheet - Assets

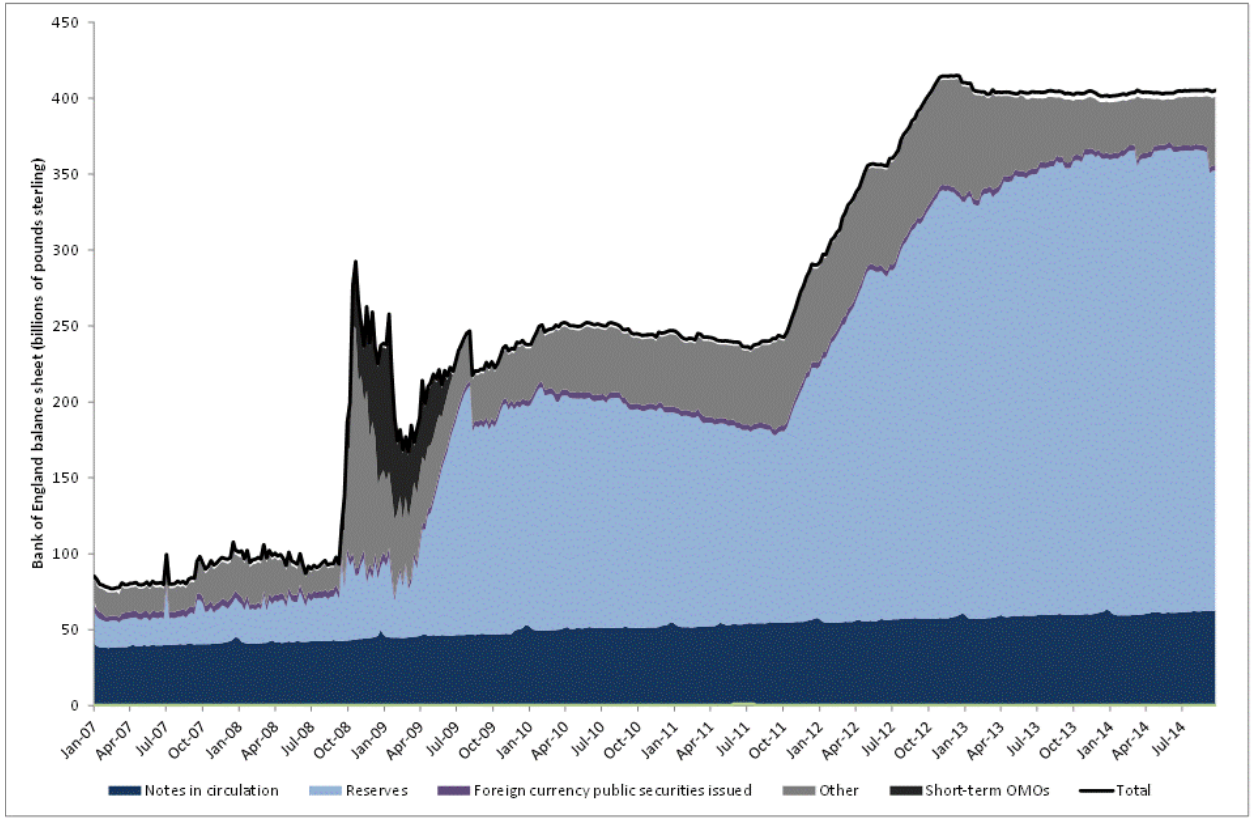

Figure 11: Bank of England Balance Sheet - Liabilities

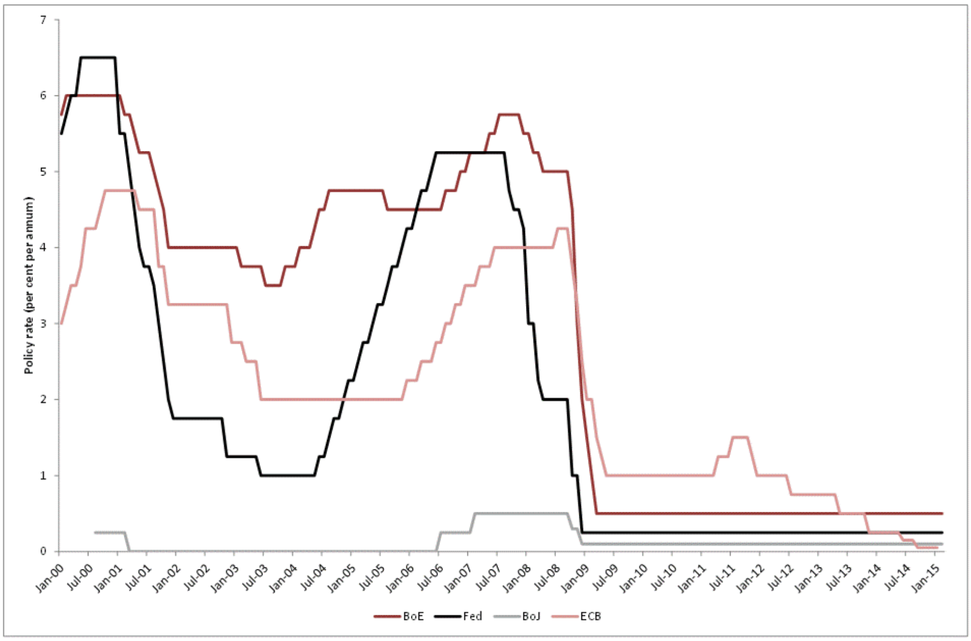

Figure 12: Central Bank Policy Rates

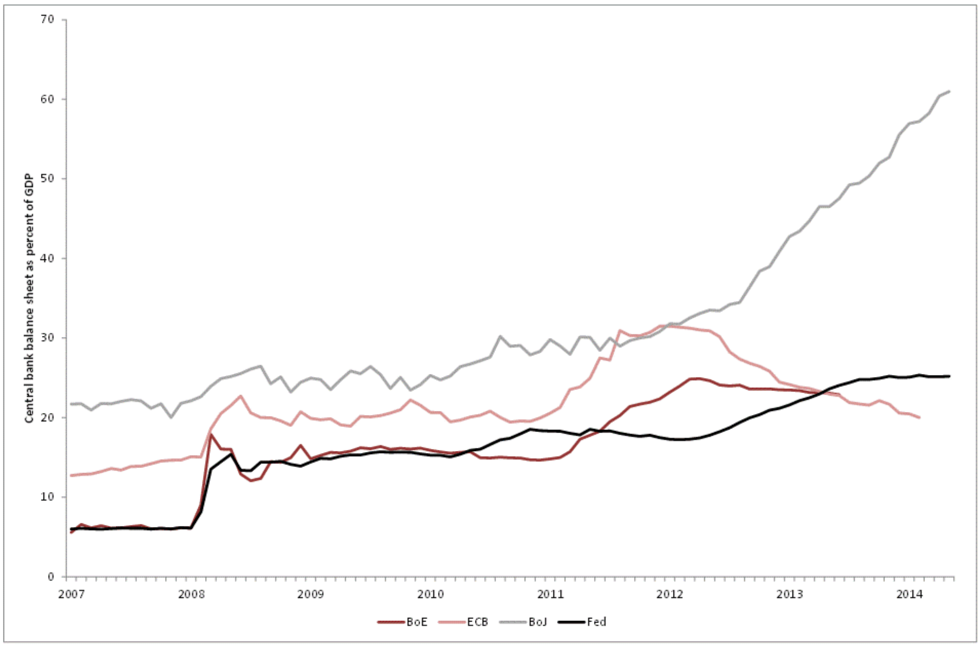

Figure 13: Central Bank Balance Sheets Relative to GDP

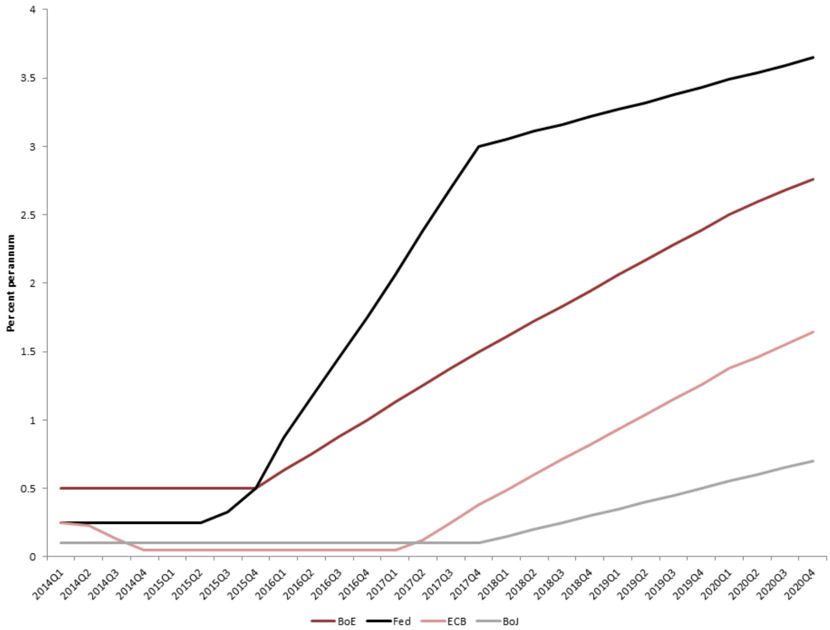

Figure 14: Expected Future Path of Policy Rates
(based on market rates and NIESR forecast: August 2015)

Quantifying Quantitative Easing

Jack Meaning

Abstract

The financial crisis of 2008 and the subsequent Great Recession have required monetary policymakers to implement a raft of new, "unconventional" policies. The most prominent of these has been the large scale purchasing of government securities known as quantitative easing. However, without reliable empirical estimates of the impact these purchases have had on key financial and macroeconomic variables it is impossible to accurately calibrate their use going forwards, or evaluate their success with hindsight. To that end, this paper employs 4 different econometric techniques to quantify this impact for policies carried out by the Federal Reserve and the Bank of England; an event study, a panel regression, a cross-sectional regression and time-series analysis. The key result, that yields on government securities have been significantly lowered by central bank asset purchases, is consistent across methodologies and can be quantified at between 20 and 150 basis points depending on the exact programme and the characteristics of the security in question. These estimates can not only be used to inform future bouts of quantitative easing, but also provide a metric to calibrate the growing array of structural models which incorporate channels for asset purchases.

1 Introduction

For any policy tool to be of practical use it must be to some degree quantifiable. Without that it is impossible to accurately calibrate or subsequently evaluate the impact of the policy.

This was one of the big challenges facing policymakers in 2009 as they moved to design and implement large scale asset purchase programmes. Unlike traditional monetary policy which had a wealth of empirical work quantifying its transmission and impact on financial markets and the real economy, prior to 2009 there existed little to no research which put firm estimates on the effects of large scale asset purchases by central banks.

Without this knowledge, policymakers cannot make fully informed decisions, for instance on the quantity of assets to buy in order to achieve a particular objective. This meant that in 2009 they were implementing these unprecedented policy measures blind, or at least partially sighted, using a combination of educated guesswork and trial and error to find their way in setting quantitative easing to stabilise their ailing economies. They were learning as they were going.

In order to make asset purchases an effective policy tool it is therefore vital that we build up a comprehensive understanding of their effects based on robust empirical evidence. This paper looks to contribute to this process by applying a number of different frameworks to provide estimates of the impact of asset purchase programmes undertaken by the Federal Reserve and the Bank of England.[1] These range from a methodologically simple event study analysis to more complex panel and cross-sectional regressions of data at the level of individual securities and finally, time series methods.

By employing a range of techniques, it is the aim of the paper to draw on common results and provide some robust, quantified statements on the impact of quantitative easing. For instance, across all frameworks we find a significant lowering of yields in government securities that were purchased by the Federal Reserve and the Bank of England. This can be quantified at between 20 and 150 basis points depending on the specific programme in question and the characteristics of the bond being purchased.

Our empirical work also sheds some light on the transmission mechanisms of asset purchases, which is intended to inform the concurrent theoretical debate.[2]

[1] While both the Federal Reserve and the Bank of England have bought a range of assets, this paper focusses on the purchases of nominal government securities which have made up the majority of the quantitative expansion of their balance sheets.

[2] See first chapter in this thesis.

We find evidence that some policy announcements reduced long-term yields by lowering expectations of the policy rate in future periods, consistent with the expectations hypothesis and the implied result of the canonical New Keynesian model. However, in line with a number of other studies, we find evidence of imperfect substitutability between assets and a significant impact on yields emanating from changes in supply. This supports theoretical positions such as the preferred habitat theory of Modigliani and Sutch (1966) and the failure of Wallace Neutrality, Wallace (1981).

In the final section we also present empirical evidence that the Federal Reserve has lowered longer yields in the US not only by reducing supply, but also by withdrawing maturity from the publicly available supply of Treasuries, whilst the US Treasury itself has acted to add maturity to the same market, a result which calls for further debate on the optimal coordination between monetary policymakers and the fiscal authority concerning asset purchases which equate to debt management in sovereign debt markets.

The paper will first discuss the results of the event study and its limitations before applying the analytical framework laid out by D'Amico and King (2010) to policies in the US and UK in order to more robustly test the hypothesis of imperfect substitution between assets. This is followed by time series analysis for the US of both supply and maturity effects in the form of an error-correction model. The paper is concluded with thoughts on what insight the range of analyses can provide.

2 Event Studies

Many of the initial studies evaluating asset purchases were based on the event study framework. This had the benefit that it was relatively simple to implement and did not require a long data set. By its nature, it could also give a result almost immediately after the policies were announced, which made it extremely useful for central bankers in need of a metric by which to measure the strength of their actions, and thus calibrate further policy.

The premise is simple; identify moments where information on the policy is released to the economy, or 'events'. Next, construct an 'event window' around each event from a moment preceding it, when the information is assumed to be unknown by agents, to a moment after the event when agents are assumed to have internalised the new information. This could be a significant speech by a policymaker, or a formal policy announcement, for instance. The change in variables over this event window is attributed to the event itself. Repeating this for all events pertaining to the policy would give its cumulative impact. The idea is to isolate moments when the market has taken on information about the policy in question from all other moments in which movements in variables are being driven by other factors.

2.1 Event studies in the literature

Gagnon et al (2011) apply this methodology to the initial US asset purchase programme looking at the change of seven key interest rates in response to all FOMC announcements and speeches which referred to asset purchases between 25th November 2008 and 17th February 2010. They find that yields on 10 year treasuries fell by 60-90 basis points over the event windows despite a trend of increasing rates over the period as a whole.[3] The reductions in rates on 10 year agency debt and mortgage-backed securities (MBS) are even greater, between 110 and 160 basis points. Joyce et al (2010) find similar results when studying how UK gilt yields across the entire yield curve reacted to key announcements surrounding the Bank of England's initial QE programme. Their event study suggests a reduction of between 55 and 120 basis points, depending on the exact maturity in question. By comparing this to the change in rates of OIS contracts of the corresponding duration they aim to separate out the extent to which this change is a consequence of changes in agents' expectations of the path of future Bank Rate (proxied by the change in OIS rates) or other factors, which they attribute to portfolio rebalancing. In their analysis, the biggest driver of the fall in interest rates is portfolio rebalancing. They also carry out a useful robustness check on the sensitivity of their results to variation in the size of the event window and find it has little material impact on the headline results.

These core findings for the Bank of England's first purchase programme are supported by the event study in Caglar et al (2011) which finds broadly equivalent falls in UK gilt yields, along with a significant flattening of the UK government yield curve.[4] They extend the narrow focus of earlier event studies to look at a wider range of variables including interest rate forwards, corporate bond rates, exchange rates and equity indices. The results suggest that QE in the UK had an impact on the economy beyond the direct effect in the government bond market. In a more recent paper Fic (2013) applies the event study methodology to

[3] They also calculated the change over just a core of 8 events and found that the reduction in yields was marginally larger.

[4] Meier (2009) also offers support to the significant downward impact of asset purchases on UK interest rates, but doesn't cover even the entirety of the first purchase programme.

policies by the Bank of Japan and the European Central Bank, as well as to those of the Federal Reserve and the Bank of England. In line with the other studies mentioned, she finds evidence of a significant reduction in yields in all cases but then goes further, feeding these estimates into the National Institute of Economic and Social Research's global econometric model (NiGEM) to investigate the macroeconomic impact on the developed economies implementing them, and the developing economies which experience policy spillovers.

2.2 Our event study

The event study presented here draws on work by the author in both Caglar et al (2011) and Meaning and Zhu (2011). We first identify a series of events relevant to each quantitative easing programme. These events are selected on the basis of when new information became publicly available to markets so comprise of significant policy announcements and speeches by the Federal Reserve and Bank of England relating to their purchases of government securities.[5] We then measure the change in variables over a two day window around each event, from the close of business the day before the event to close of business the day after it. This is consistent with Joyce et al (2010) and other studies and is preferred by this author as it allows for some slight lag in impact as markets internalise the information, or in case there exist frictions in the process of settling transactions. Evidence of such frictions is provided in the later flow analysis.[6]

We then analyse movements in a range of variables over these identified windows and compare the changes in these periods to those over the average two day window in the full sample period concerning quantitative easing.

Figure 1 shows the cumulative impact of each programme on the yields of government securities of different maturities and ten-year corporate bonds of different grades. This is defined as the sum of changes on days associated with QE events. It is clear that each asset purchase programme had a significant negative impact on yields at all maturities. For the UK this got stronger the further along the yield curve one looks, lowering yields at the long end by over 100bps. Both US programmes on the other hand exhibit a peak around 5 year rates of over 30bps. This different pattern can be explained both by the shorter average maturity of the US government debt portfolio and also the difference in the maturity of the purchases carried out by the Federal Reserve and the Bank of England. The second US programme had a stronger impact on ten-year rates

[5] A full list of events and their rationale for inclusion can be found in Table 1.

[6] The results prove unchanged in any significant way when a one day window is applied.

than the first. This is to be expected as the Federal Reserve began buying more longer-dated securities under US QE2 so as not to withdraw too much of any one area of the market and to take advantage of duration effects.

Table 2 presents a breakdown of the cumulative impact by event. In the UK the first two announcements are clearly dominant, contributing much of the final value. This is not surprising as they can be assumed to have had much more informational content than later announcements as they outlined not only the quantity of assets to be bought, but also their type and the relevant mechanisms by which this unprecedented policy for the UK economy was to be applied. They also reflect the larger size of purchases announced in March 2009, £75bn compared to later extensions of £50bn and eventually £25bn. Looking at the announcement on 6th August, one can also clearly see the market reaction to the Bank of England's shift towards purchasing longer-dated securities reflected in a larger fall in 20 year rates.

Corporate yields fell over the event windows by a comparable amount to government securities. This implies spillover effects as news of central bank purchases of government bonds encouraged demand for their corporate counterparts. This may suggest a degree of substitutability between the two asset classes, supporting the portfolio rebalancing theory of transmission, or just that the asset purchase programmes increased confidence more widely and that moved investors into asset classes which had been seen as more risky, such as corporate bonds.

Aside from the limited size of the event window, designed to keep out non-policy related factors, we also look to confirm that the fall in yields is not just due to a downward trend in the data. A limit of the event study methodology is that such a trend would cause a fall across event windows and non-event days similarly. To combat this we construct a primitive counterfactual. Tables 2-4 also show the average two-day change for yields over the whole sample period.[7] In all cases they are less than one basis point in absolute terms, and for government securities in excess of one year, they were actually increasing making the falls observed over the event windows even starker.

The impact on equity markets is less clear. On average, the FTSE100 rose by 0.12% over event windows associated with QE in the UK, Table 2. This is a negligible rise, especially when taking into account that over the full sample period the average two-day change was an increase of 0.22%. The implied volatility also rose by an average of 0.65% for events, whilst falling by 0.18% on average for

[7] The full sample average is calculated by taking the average two day change from the first event window to the last.

the full sample, suggesting that, to the small extent QE did influence UK equity markets, it only increased volatility. US asset purchases had a much stronger effect on equities, at least for the first programme where associated two day event windows experienced an average change of over 1% compared to just 0.2% for the full sample. As with corporate bonds, it seems that US QE served to bolster investment sentiment in other assets. It also reduced implied volatility by an average of 2.6%.

Initial asset purchases resulted in significant currency depreciations for both the US and UK. QE in the UK reduced the sterling nominal effective exchange rate by an average of 0.68% per event. In the US this was slightly higher, at 0.76%. However, taking into account the counterfactual of the full sample where the UK experienced a very slight appreciation and the US a small depreciation, the exchange rate implications of the two programmes are almost identical. This result may be thought to correspond to a monetary view of the exchange rate where increases in the money supply devalue one currency against another, or it may be that markets envisaged asset purchases being successful at lowering interest rates, both long and real, and adjusted the exchange rate in a manner consistent with interest rate parity.

Whilst this paper is not directly concerned with the transmission mechanisms behind asset purchases, it is interesting to note that one year OIS rates fell significantly, especially in the US. As it is based on a geometric average of the overnight rates over the period of the contract, the OIS rate is often considered a good proxy for market expectations of the policy rate. This fall then lends support to the transmission mechanism put forward by Eggertson and Woodford (2003) and others that unconventional policies work by manipulating agents' expectations of the future path of policy rates. Combined with our other results this suggests that both portfolio rebalancing and the expectations channels have a role to play.

These results must be viewed with some caveats in mind which derive from the limitations of event study methodology itself. Firstly, events may be 'muddy' to varying degrees as it is impossible in some instances to completely isolate when information reaches markets concerning purchases from the release of other information. A clear example is the 5th March 2009 announcement by the Bank of England where they simultaneously announced the asset purchase programme and a cut in Bank Rate. The March event window will inevitably capture the market absorbing both these pieces of information, although the cut in Bank Rate was widely anticipated so to a large extent would have already been built in to agents' behaviour.

This highlights another limitation of event studies. If the announcement is anticipated then it will be priced in by markets prior to an event window which straddles the announcement. This was not as much of a problem at the inception of QE as its novelty meant that people had very limited knowledge of how it would be implemented and, crucially, of the central bank's reaction function concerning asset purchases. However, as time has progressed the reaction function has arguably become clearer, to the extent that it should be almost explicit in the case of forward guidance, and thus the power of event studies is weaker. This would lead to an underestimation of the impact of the policy and is a potential explanation for at least part of the diminished impact observed in later announcements in Tables 2-4.

More generally, an effective event study relies on correctly identifying all the moments where data was internalised by the market. This is hard to do as agents may change their behaviour in anticipation of a policy, when it is announced, or perhaps not until it is completed. It could also be anywhere in between. Failure to capture a relevant event may bias the estimated impact of the policy.

3 The D'Amico and King Framework

A more complete methodology is that of D'Amico and King (2010). They propose two estimations. The first is a panel regression on daily data at the individual security level designed to capture the flow effects associated with the physical act of making purchases. The second is a cross-sectional regression, again at the individual security level, but this time looking at the change in variables from the day prior to the inception of the policy, to the day of the last physical purchase has been made. This cross-sectional approach aims to elucidate on the more permanent "stock" effect that results from reducing the supply of a security in the publicly available market, relative to the existing demand for it.

We estimate both effects for asset purchase programmes in the US and the UK.

3.1 Data

Underlying the analysis of both flow and stock effects is the same database. It contains daily data on individual government securities defined by their unique CUSIP identifier in the US and ISIN code in the UK. As in D'Amico and King, all inflation linked securities are removed, as are any securities with a remaining maturity of 90 days or less to avoid contamination by the differing behaviour of investors as a security approaches its redemption date. The sample consists of 239 securities for US QE2 and 40 for the Bank of England's first QE programme.

For each security, tranches of substitute securities within the sample are constructed. Any security with a remaining maturity within two years of that of security i is deemed its near substitute. A remaining maturity of between two and six years from that of security i is considered a mid-substitute whilst a difference of more than 6 years makes it a far-substitute.

The raw purchase data, which is sourced directly from the Federal Reserve and the Bank of England, is normalised by dividing each purchase amount by the total outstanding supply of the security and its near substitutes. This has a number of benefits. It means the data now shows the fraction of the supply of a given security type that the central bank buys and removes from the market in each operation. The depth of a particular part of the market is thus controlled for. After all, £10mn of purchases in a part of the market where total supply is only £20mn could reasonably be expected to have a different impact than the same £10mn of purchases in an area of the market where the total supply is £20bn. The normalisation also controls for changes in the total supply caused by new issuance by the fiscal authority.[8] Lastly, it makes the coefficients easily comparable.

A final comment to make is that, for the sake of simplicity, all of the analysis is conducted in price space[9]. The implied impact is then converted into an equivalent change in yields using the modified duration of the security.

Ultimately, we are left with a set of time-series on the price, supply and supply of substitute securities for a large number of securities which we use to build both our panel, and also our cross-sectional database. For the panel we bring together the individual time series and maintain the daily frequency of the underlying data. For the cross-section we simply take the change in each time series from the first day of the policy to the last, reducing each time series to a single data point.

[8]This is likely to affect the price and highlights the monetary-fiscal interaction inherent in these policies.

[9]Though D'Amico and King (2010) find that their results are little changed if the same procedure is carried out using yields.

3.2 Flow Effects (Panel Regressions)

Flow effects are those observed on the day of the physical act of purchase of a security or its substitutes.[10] In a world of complete information, frictionless markets and perfectly rational agents flow effects would not exist as investors would have already priced in the impact of the purchases when the policy was announced This may have been true to some degree on the aggregate level, but on the individual security level, agents were unaware exactly how much of exactly which securities would be bought and when. There may also have existed other frictions which caused investors to ineffectively price in the policy until the day of the purchase. It is therefore worthwhile estimating the extent to which flow effects were present.

The flow effects are estimated by a panel regression of the form

$$\frac{\Delta P_{it}}{P_{it-1}} = \alpha_i + \delta_t + \beta_1 q_{it} + \beta_2 q_{it}^N + \beta_3 q_{it}^M + \beta_4 q_{it}^F + \varepsilon_t \qquad (1)$$

The proportional daily change in the price of security i depends on the quantity of that security purchased by the monetary authority, q_{it}, and their purchases of securities considered near-, mid- and far-substitutes, q_{it}^N, q_{it}^M and q_{it}^F respectively.

The panel nature of the regression allows us to control for a security specific fixed effect, α_i, and a daily time dummy, δ_t. The security specific fixed effect enables us to capture any characteristics which are idiosyncratic to a given security but relatively constant across the sample period. The time dummy lets us take account of any unobservable influences which affect prices more uniformly across the full spectrum of securities on a given day such as policy announcements or the release of additional economic information to markets. In essence, a large number of external influences should be captured and controlled for by one or the other of these dummies, though it is not possible to back them out individually.

For the estimation each sample is sub-divided on the basis of two criteria; first, whether an asset had more or less than 15 years of remaining maturity at the time of purchase and second, whether it was eligible or not for the purchases.[11] For instance, the UK QE set eligibility requirements for gilt purchases with an initial purchase range (IPR) of 5-25 year maturity. As investors knew which securities would be targeted, yields might react differently for eligible and ineligible

[10] These flow effects are directly comparable to the auction effects investigated by Breedon et al (2012) and Banerjee et al (2012)

[11] The Federal Reserve and the Bank of England have carried out their asset purchases via reverse auctions, announcing in advance of each auction a range of "eligible" securities which they are willing to buy in that particular operation.

securities.

We then estimate equation 1 for each purchase programme. The sample period for US QE1 runs from 25th March 2009 to 29th November 2009, US QE2 from 4th November 2010 to 30th June 2011 and QE in the UK from 25th March 2009 to 29th October 2009.[12]

Table 5 presents the results of the regressions alongside the original result of D'Amico and King.

Consistent with the result of D'Amico and King, we find that both programmes exhibit significant flow effects and that they are of an approximately similar magnitude. Each billion dollars spent on Treasuries by the Fed lowered the yield of the security purchased by 0.4 basis points in US QE1 and 0.6 basis points in US QE2.[13] In the UK QE, each billion pounds of purchases lowered the yield of the security being purchased by 1.5 basis points, which, taking into account the dollar-sterling exchange rate equates to an almost identical marginal impact for the flow effects across all three policies.

If we then substitute the mean values of each purchase type from our samples into our estimated equations we find that the average operation in US QE1 lowered yields on the average security by 3.5 basis points, in US QE2 by 4.7 basis points and for the UK QE by 1.5 basis points. This highlights an interesting dynamic between the US and UK approach to asset purchase programmes. US QE1 and the UK QE were more or less the same size and had the same pound for pound effectiveness. However, the US programme took place over a much shorter time period and used just over half of the number of operations. Therefore the size of each operation had to be bigger which is reflected in the higher average flow effect per operation.

These results also provide an insight into the transmission mechanism of flow effects. The β_2 coefficients on close substitutes are also significant and almost as large. This suggests that purchases of substitutes directly affected the price of a security, whether the security itself was purchased or not. This implies a high degree of substitutability between securities and seems to support the idea that asset purchases work, at least in part, through portfolio rebalancing. In general the coefficient estimates wane once we move towards "mid" and "far substitutes", indicating that policy effectiveness diminishes as the degree of substitutability

[12] Only days when purchases occured are included in the sample.

[13] The price change per billion dollars of purchases is calculated dividing the coefficient on q_i by the average amount of near substitutes over the sample. This is then converted to a change in yield by multiplying the price change by -φ where φ is the inverse of the average modified duration of the sample.

falls.

It can also be noted that these flow effects are short-lived. When we regress the right hand side of equation 1 on the one-period lead of the price change instead of the contemporaneous price change we find that the coefficients drop in value by around a third in our main samples. However, most do remain significant suggesting the existence of frictions in settlement which may cause some of the flow effects to not be priced in until the following day. This provides support for our use of a two day event window in our event study analysis.

3.3 Stock Effects (Cross-sectional Regressions)

Stock effects, as detailed by D'Amico and King are the change in yields resulting from the more persistent change in supply of an asset. In this sense they are very similar to the supply effects discussed in the first chapter of this thesis and by other authors, including D'Amico et al (2012). In the standard New Keynesian framework, such effects should not exist as complete and perfectly functioning financial markets should lead to full arbitrage and render supply inconsequential for price. However, if there exists imperfect substitutability between assets and market frictions which limit arbitrage such as those discussed in Tobin (1969), Modigliani and Sutch (1966) and Andres et al (2004), then demand will be less than perfectly elastic and supply will have consequences for price.

To test for the existence of these the stock effects, D'Amico and King present a cross-sectional regression which analyses the change in price from the day immediately prior to the first announcement of the programme to the end of the day of its final auction. This is one of the biggest strengths of the methodology. On the first date of the sample it is reasonable to assume that there is little to no information about purchases of individual securities. However, by the time the final auction is completed, by definition there is full information on how much of which securities have been bought by the programme. Agents may internalise this information set at any point between these two dates. In fact, it is likely that they will take on partial information at a number of moments over the period. Identifying and isolating when this happens is one of the key challenges of the event study methodology. The cross-sectional approach allows the change to be captured whenever it becomes internalised, be it at the moment the broad policy is announced, not until the physical purchases are all carried out, or anywhere in between.

One limitation this imposes is that only securities which are in issue for the entirety of the policy can be included in the regression. This reduces our sample sizes to 188 securities for US QE2 and 31 for UK QE.

Another concern highlighted by D'Amico and King (2010) is the potential endogeneity which may arise if the monetary authority targeted particular securities which were underpriced and could be expected to experience a larger change in price regardless. To counter this we use a two-stage least squares regression instrumenting the amount of purchases with the Svensson fitting error of the security on the day prior to the announcement of the policy, Svensson (1994), and a range of other control variables. The Svensson fitting error is the difference between the observed price and the price implied by a yield curve fitted using the methodology outlined by Svensson (1994). Therefore, any security whose price lies above the yield curve is considered overpriced, and any whose price is below is considered underpriced.

The first stage regression therefore takes the form

$$Q_i = \alpha + \beta S_i + \delta_j \phi_{j,i} + \varepsilon \qquad (2)$$

where Q_i is the total quantity of security i bought over the complete duration of the programme, S_i is the Svensson fitting error and ϕ_j is a set of controls, j.

The results of the first stage regression are shown in Table 6. They show that the central bank for all three policies successfully targeted underpriced securities, although in the case of the Bank of England this was applied to underpriced areas of the yield curve, not specific securities themselves.

The results of the first stage are then applied to construct a series of instrumented purchases, $\dot{Q}_i$, and this is in turn used to create a series of purchases of near substitutes. This allows us to control for endogeneity in the near substitute purchases whilst maintaining consistency with own purchases. We then use this to run the second stage regression

$$\frac{\Delta P_i}{P_i^*} = \mu + \kappa_n \dot{Q}_i + \omega_n \dot{Q}_i^N + \gamma_1 M_i + \gamma_2 M_i^2 + \varepsilon \qquad (3)$$

where ΔP_i is the price change for security during the purchase programme, P_i^* is its price on the day before the start of the programme. $\dot{Q}_i$ is the instrumented value of purchases of security i and $\dot{Q}_i^N$ is that of near substitute purchases during the asset purchase programme.[14] M_i is the remaining maturity of security i. We do not include further macroeconomic controls. This is for a number of

[14]Considering the possibility that our coefficients of interest, κ and ω, may vary for different maturities, forUS QE2 we include interaction dummies which separate securities with less than 15 years of remaining maturity from the rest. For UK QE, we use interactive dummies to separate gilts which were within the initial purchase range from those which were not.

reasons, the primary of which is that the methodology does not allow for it. The estimations work on the individual security level in cross-section so any macroeconomic variable we choose to include would posit the same value for all securities in the sample and thus provide no additional information. However, this aside, the fine nature of the data renders further controls unnecessary. To some degree it may be assumed that shifts in inflation and growth expectations, for example, would affect the spectrum of nominal treasury securities more or less uniformly and the extent to which this is not true, the majority of differences in impact will be due to shifts in the slope of the yield curve and thus captured in the remaining maturity controls we do include.

The results of the final regression for each programme are presented in Table 7 and several key points emerge.

We see that purchases did indeed have a positive impact on prices and that there is, as in the flow regressions, some evidence of substitutability and portfolio rebalancing between securities. Introducing the actual data on amounts of each security and its near substitutes purchased we can calculate the change in the price due to the asset purchase programme. We then convert this to a change in yield and subtract it from the observed yield on the last day of each sample to get a counterfactual of what the yield would have been in the absence of the policy. This implies that US QE1 lowered yields by around 30 basis points, but in the 10-15 years to maturity sector this was as high as 50 basis points. For US QE2 it was 21 basis points on average, peaking at 108 basis points in some securities with around 20 years to maturity. The UK QE programme reduced yields by as much as 74 basis points[15], and 27 basis points on average.

When interpreting these estimates one has to bear in mind that the different programmes varied in size. At $300bn, US QE1 was roughly half the size of US QE2 and its maximum effect was approximately half. This suggests a degree of linearity in the impact of asset purchases.[16] However, the average effects across the yield curve was similar between the two programmes, suggesting that US QE2 was less effective per billion dollars spent than US QE1.

One also needs to consider the impact of each programme relative to the quantity of the total stock of securities. In absolute terms, US QE1 and UK QE were of almost an exactly equivalent size. Based on the maximum effect of

[15]In securities with a remaining maturity of around 12 years.

[16]The amount of $600 billion in Treasuries may understate the true extent of supply withdrawn by the Fed as US QE2 was supplemented by additional securities bought by the Fed reinvesting funds originated from other Fed programmes. Taking this into account, the Fed purchases made over the course of US QE2 were just under $750 billion and our results suggest a slight diminishing return to the second programme.

both policies, the British policy was therefore more effective. However, \$300bn represented just 4.7% of the total Treasury stock outstanding in 2009, whilst the £200bn of purchases carried out by the Bank of England accounted for around 29% of the free float of gilts. Even the \$600bn of purchases under US QE2 only amounted to 6.6% of the US Treasury debt stock. Thus to get a broadly similar marginal effect, the Bank of England had to withdraw a much higher proportion of total supply than the Federal Reserve.

These results are largely in line with those of previous work. For instance, Williams (2011) adjusts estimates from a number of existing studies by various authors by the size of asset purchase programmes. For a \$600bn operation, the estimated impact on longer-term bond yields ranges from 14 basis points in Greenwood and Vayanos (2008) and 15 basis points in Krishnamurthy and Vissing-Jorgensen (2011) to 18 basis points in Gagnon et al (2011) for US asset purchases, and 40 basis points in Joyce et al (2011) for UK purchases.

4 Time-series Analysis

The novel nature of quantitative easing in most economies has posed problems for time series analysis, limiting the sample sizes which cover periods of large scale asset purchases. This is a problem which, by its nature, will dissipate over time as the longer the policies last, the more extensive the time series we have at our disposal.

However, although there exists only a relatively short period in which QE has been formally carried out, central banks have been effecting asset purchases in government securities markets for decades. In an open market operation the central bank provides newly created funds to commercial banks in the form of reserves in exchange for collateral in the form of government securities. Whilst the size and duration of these operations may differ from the large scale asset purchases, both actions amount to a manipulation of the quantity of government debt available to the public, and consequently the amount held on the central bank's balance sheet, and both are underpinned by the same theoretical construct. In this respect we have a wealth of time series data to draw upon.

Therefore, using monthly data on the US economy from January 1990 to June 2011 we estimate the following general equation to investigate the historical link between the central bank's balance sheet and interest rates.

$$y_t^{10} = \alpha + \beta_1 M_t^F + \beta_2 M_t^T + \beta_3 Q_t + \beta_4 r_t + \Phi_{tj} \tag{4}$$

y is the yield on a ten year nominal treasury security. Q is the fraction of the total outstanding stock of US treasury securities which is held on the Federal Reserve's balance sheet. This is intended to capture the supply effect of asset purchases in much the same way as the stock effect in the earlier section. To capture the impact of changes in the composition of the balance sheet we also include M^F, which is the average maturity of the Federal Reserve's treasury portfolio. Obviously the Federal Reserve is not the only policymaker engaging in such debt management-style policies, so to control for the effect of debt management by the fiscal authority (over and above the pure supply effects implicitly captured in the denominator underlying QE) we include M^T which is the average maturity of the stock of Treasury debt available for private purchase.[17] To control for traditional monetary policy we include the effective Fed Funds rate, r, and lastly Φ_{tj} represents a set of macroeconomic controls, j.

ADF tests reveal that a number of the variables of interest exhibit non-stationarity, or at the least, heavy persistence, Table 8. To overcome this we adapt the basic regression in equation 4 to error correction form using the two-step Engle Granger process, Engle and Granger (1987).

In the two-step step Engle-Granger process we first establish the long-run relationship for our I(1) variables by regressing our non-stationary independent variables onto our dependent variable. We use AIC and BIC criteria to find which combination of macroeconomic controls yields the best results and find it to be the parsimonious equation

$$y_t^{10} = \alpha + \beta_1 M_t^F + \beta_2 M_t^T + \beta_3 Q_t + \beta_4 r_t + \varepsilon_t \tag{5}$$

We then test the residual from this estimation, ε, and find it to be I(0). This means equation 5 represents our long-run relationship, whilst our residual represents our error correction term and can be written as

$$\varepsilon_t = y_t^{10} - \alpha - \beta_1 M_t^F - \beta_2 M_t^T - \beta_3 Q_t - \beta_4 r_t \tag{6}$$

To formulate our ECM we then regress the one period lag of our error correction term, along with the first differences of our I(1) variables and other I(0) variables on to the first difference of Y to get our short-run dynamics. Now all variables in our regression are I(0). We again use AIC and BIC criteria to establish an appropriate combination of variables in the short-run dynamics leading to the final regression of

[17] We have tried different variants of these policy variables, but the key results remain unchanged.

$$\Delta y_t^{10} = \delta_1 \Delta y_t + \delta_2 \Delta r_t + \delta_2 \Delta \pi_t^e \qquad (7)$$
$$+\delta_4 \left[y_t^{10} - \alpha - \beta_1 M_t^F - \beta_2 M_t^T - \beta_3 Q_t - \beta_4 r_t \right] + \varrho_t$$

The results from both stages of the baseline regression are shown in Table 9. Focussing on the long-run relationship, we can see that each of the policy variables has a significant impact on the ten year treasury rate. As with our previous estimation frameworks, there is a significant inverse relationship between the quantity of assets held by the central bank and interest rates. Specifically, each 1% of the stock of Treasuries outstanding which is removed by Federal Reserve asset purchases leads to a 20 basis points decline in yields. 10 year yields are also inversely related to the average maturity of the supply which is withdrawn so that purchases of longer-dated bonds can reduce 10 year rates even without an expansion in size of the balance sheet. This suggests there is potential for "Operation Twist" style sterilised asset purchases to be effective, changing the relative supply of longer-term securities and flattening the yield curve without burdening the Fed with additional funding requirements. But there is no free lunch, and there are clear limitations and drawbacks from relying solely on the maturity channel.

First, the strategy is limited by the size of current holdings of shorter-maturity assets. Second, carrying a substantial share of longer-maturity Treasury securities or alternative private assets on its portfolio can expose the Fed to potentially much higher risks of future capital losses. There is also the possibility of the Fed becoming a major player in certain segments of financial markets as a consequence of focussing on longer-term debt. This could reduce trading and lead to the exit of some market participants, hampering normal market functioning. What is more, selling short-term securities to sterilise purchases may increase short rates and bank funding costs.

The implied strength of the maturity channel is such that each month increase in the average maturity of the Federal Reserve's security holdings lowers the yield on ten year securities by 3.4 basis points. Looking at the fiscal authority's debt management variable, the same one month increase in the average maturity of the publicly available securities portfolio would actually raise long-term rates by 7 basis points, double the impact of the same move by the monetary authority. The stronger effect is to be expected when one considers the relative size of the portfolio being influenced by the Treasury compared to the portfolio held by the

Fed. Once we control for size, debt management by the monetary authority is much more effective.[18] The opposite signs of the coefficients on the two debt management variables (and implicitly the two supply variables underlying the numerator and the denominator within QE) is logical as each policymaker is acting on the publicly available supply of securities from opposite sides. This highlights the potential for coordination issues between independent monetary and fiscal authorities when both use debt management as a policy tool.

Of note is that the long-run relationship between the ten year yield and the Fed Funds rate is 0.22, which is fully consistent with the consensus in the literature where 0.25 is the canonical figure, inferring that a 100 basis point hike in the policy rate causes long-term rates to rise by 25 basis points.

Using the estimated coefficients and the realised data, we can decompose the change in yield associated with policy factors into its constituent parts, Figure 2. The first thing to note is that the estimates are highly supportive of our estimates derived from other methodologies, around 40 basis points for US QE1 and 120 for US QE2.

Secondly, the driving force in both policies is the balance sheet expansion. This is unsurprising as manipulating the maturity of the Fed's securities portfolio was never an explicit objective of either of the first two asset purchase programmes. In US QE1 the maturity effect is minimal, but perhaps more revealing is that in US QE2 the maturity effect actually works against the Federal Reserve's objective. This tells us that by targeting purchases at assets with an average maturity of less than its existing portfolio, US QE2 actually undoes some of the desired impact resulting from the expansion of the balance sheet. The implications of this result for the design of future policy are that maturity matters. The efficacy of asset purchases can be enhanced if purchases are targeted at a long enough maturity that they lengthen the average maturity of the central bank's own portfolio, causing a complimentary maturity effect.

Thirdly, in both cases the Treasury's debt management policy works against the maturity and quantity effects of Fed policy. This brings us back to the issue of coordination for debt management between the monetary and fiscal authorities. It is worthy of note here that as the Federal Reserve has worked to remove supply and duration from the Treasuries market, the Treasury itself has been increasing both.

[18] See later section for estimation of equation standardised by market share.

4.1 Alternative specifications and robustness

In addition to our baseline regression, a number of alternative specifications were estimated as a means of testing the sensitivity of our result. The more interesting of these are presented in Table 10. Our core result of a significant impact on yields from changes in the size and composition of the Federal Reserve's balance sheet is robust to all alternative specifications, with some variance in the magnitudes of coefficients, but never in sign or significance.

Our first test is to the obvious questions of a structural break in the series caused by the financial crisis. Estimating over a pre-crisis sample which runs from January 1990 to December 2007 we find that the coefficients on the variables associated with Federal Reserve policy are smaller, but remain significant, suggesting that asset purchases had a weaker impact prior to the financial crisis. This makes intuitive sense for a number of reasons. First, the size and maturity profile of the Federal Reserve's balance sheet was not an explicit policy tool prior to 2008, and as such was less likely to be thought to contain information on the stance of policy. Second, theory on the imperfect substitutability of assets would imply that changes in the relative supplies of assets should have a greater impact on prices when there is a greater degree of market segmentation, as is widely accepted of the period since 2008. However, it is an important result that we still find a significant relationship in a non-crisis sample when markets were considered to be operating relatively freely with few frictions. This suggests that the usefulness of asset purchases, and the existence of supply effects, is not limited to periods of crisis.

We control separately for two forward looking variables; inflation expectations and forward interest rates. The first acts to lower the size of the coefficients on policy variables slightly compared with the baseline and is included to capture the impact of change views on future prices will have on the nominal yield. The inclusion of forward rates is intended to proxy for expectations of the path of the policy rate, acknowledging that agents are likely to be concerned with rates over the life of the bond, not just the contemporaneous Fed Funds rate. However, the coefficient turns out to be insignificant and the policy variables are left little changed.

We also look at different ways of measuring the policy variables. Using the fraction of the relevant portfolio with a remaining maturity of 5 years or higher, as in Kuttner (2006) does little to affect the magnitudes implied by the coefficients. We also replace the average maturity of the stock of US Treasuries outstanding with the average maturity of their net issuance, as this is the variable that the fiscal authority can directly control in any one period. The effect is significantly smaller, which is not surprising given that newly issued securities represent a relatively small fraction of the total market.

The last set of controls are adjusting for the relative sizes of the two policy actors. It could be realistically expected that a one month extension in the average maturity of the total supply effected by the US Treasury would have more of an effect than a one month extension of the Federal Reserve's holdings purely because the Federal Reserve is a much smaller player in the market. To counter this we divide the average maturity series by the size of the Federal Reserve's balance sheet and the total supply of Treasuries outstanding respectively. We see that in a like-for-like comparison that debt management carried out by the monetary authority is actually more effective than that carried out by the fiscal authority.

5 Conclusions

Across a range of techniques, we find that central bank asset purchase programmes have had a significant impact, lowering yields in all cases. The magnitude of this impact is broadly consistent across all techniques, and as such represents a robust result which sits within the range of other estimates in the literature.

We find empirical evidence in support of asset purchases affecting yields by shaping expectations of the future path of policy rates, as would be consistent with the canonical New Keynesian view of the world. Alongside this though we also find a significant role for both the supply and composition of asset markets to influence prices. The relevance of these factors is something the baseline New Keynesian model struggles to convincingly reconcile with its assumptions of complete and frictionless financial markets and a representative agent. Our empirical results, combined with the growing body of literature within which they sit, inform the theoretical debate and validate the investigation of breaks in these assumptions, such as the preferred habitat theory.

The implications of our results for policy are that asset purchases can be an effective way to manipulate interest rates, even when the short-term nominal interest rate is constrained. However, they caution on the need for coordination between fiscal and monetary authorities when asset purchases are involved in sovereign debt markets.

References

[1] Andre´s, J., Lopez-Salido, J.D. and Nelson, E. (2004). 'Tobin's imperfect asset substitution in optimizing general equilibrium', Journal of Money, Credit and Banking, vol. 36(4), pp. 665–90.

[2] Banerjee B, Latto D and McLaren N (2012). 'Using changes in auction maturity sectors to help identify the impact of QE on gilt yields'. Bank of England Quarterly Bulletin 2012 Q2, pages 129-137.

[3] Breedon, F., Chadha, J.S.and Waters, A (2012). "The financial market impact of UK quantitative easing," Oxford Review of Economic Policy, Oxford University Press, vol. 28(4), pages 702-728, WINTER.

[4] Caglar, E., Chadha, J.S., Meaning, J., Warren, J. and Waters, A (2011). 'Central Bank Balance Sheet Policies: Three views from the DSGE literature', Chapter in 'Interest Rates, Prices and Liquidity' Cambridge University Press, 2011

[5] D'Amico, S and T King (2010): "Flow and stock effects of large-scale treasury purchases", Federal Reserve Board Finance and Economics Discussion Series, 2010–52

[6] D'Amico, S, English, W, López-Salido, D and Nelson, E (2012). "The Federal Reserve's Large-scale Asset Purchase Programmes: Rationale and Effects," Economic Journal, Royal Economic Society, vol. 122(564), pages F415-F446, November.

[7] Eggertson, G and M, Woodford (2003)'The Zero Bound on Interest Rates and Optimal Monetary Policy' , Brookings Papers on Economic Activity, (1), pp. 139–233

[8] Engle R,F and C. Granger (1987). 'Co-integration and error correction: Representation, estimation and testing'. Econometrica, vol 55, issue 2, pp 251-276

[9] Fic, T. (2013). 'The spillover effects of unconventional monetary policies in major developed countries on developing countries', UN DESA Working Paper No. 131

[10] Gagnon, J, M Raskin, J Remasche and B Sack (2011): "The financial market effects of the Federal Reserve's large-scale asset purchases", International Journal of Central Banking, vol 7, no 1, pp 3–43

[11] Greenwood, R and D Vayanos (2008): "Bond supply and excess bond returns", NBER Working Papers, no 13806

[12] Joyce, M, A Lasaosa, I Stevens and M Tong (2010): "The financial market impact of quantitative easing", Bank of England Working Papers, no 393

[13] Krishnamurthy, A and A Vissing-Jorgensen (2011): "The effects of quantitative easing on interest rates: channels and implications for policy", NBER Working Papers, no 17555

[14] Kuttner, K (2006) "Can Central Banks Target Bond Prices?," NBER Working Papers 12454, National Bureau of Economic Research, Inc.

[15] Meaning, J and Zhu, F (2011). 'The Impact of Recent Central Bank Asset Purchase Programmes', BIS Quarterly Review, December 2011

[16] Meaning, J and Zhu, F (2012). 'The Impact of Federal Reserve Asset Purchase Programmes: Another Twist', BIS Quarterly Review, March 2012

[17] Meier, A, 2009, "Panacea, Curse, or Nonevent? Unconventional Monetary Policy in the United Kingdom," IMF Working Paper 09/163

[18] Modigliani, F. and Sutch, R.C. (1966). "Innovations in interest rate policy", American Economic Review, Papers and Proceedings, vol. 56(2), pp. 178–97.

[19] Svensson, L. (1994). "Estimating and Interpreting Foreward Interest Rates: Sweden 1992–1994", Papers 579 Institute for International Economic Studies.

[20] Tobin, J (1969) "A General Equilibrium Approach to Monetary Theory," Journal of Money, Credit and Banking, 1, February, 15-29.

[21] Wallace, N (1981). "A Modigliani-Miller Theorem for Open-Market Operations," American Economic Review, American Economic Association, vol. 71(3), pages 267-74, June.

[22] Williams, J (2011): "Unconventional monetary policy: lessons from the past three years", Federal Reserve Bank of San Francisco Economic Letter

Table 1: Event Study Policy Announcements

	Date	Policy relevant event
US QE1		
	1st December 2008	Governor Bernanke gives speech declaring the Fed "could purchase longer-term Treasury securities... in substantial quantities"
	16th December 2008	FOMC statement
	28th January 2009	FOMC statement
	18th March 2009	FOMC statement announcing the decision to purchase up $300bn of longer-term US Treasury securities and increase purchases of agency debt and MBS
	29th April 2009	FOMC statement
	24th June 2009	FOMC statement
	12th August 2009	First FOMC statement where "up to" is not included when referring to Treasury security purchases and the pace of purchases is slowed
	23rd September 2009	FOMC statement
	4th November 2009	FOMC statement announcing that purchases of agency debt would be "around $175bn"
US QE2		
	10th August 2010	FOMC statement signalling that "the pace of economic recovery is likely to be more modest in the near term than had been anticipated" and FOMC "will employ its policy tools as necessary"
	27th August 2010	Chairman Bernanke's Jackson Hole speech
	21st September 2010	FOMC statement
	12th October 2010	FOMC minutes released
	15th October 2010	Chairman Bernanke gives speech at Federal Reserve Bank of Boston
	3rd November 2010	FOMC statement announcing a further $600 billion of longer-term Treasury securities by the end of the second quarter of 2011 at a pace of about $75 billion per month
UK QE		
	11th February 2009	Publication of the Inflation Report and press conference in which it is first suggested the Bank of England is likely to embark on a large-scale asset purchase programme
	5th March 2009	First announcement by MPC of £75bn of asset purchases
	7th May 2009	Extension of asset purchases to £125bn
	6th August 2009	Extension of asset purchases to £175bn and expansion of the purchase range to any gilts with a residual maturity of more than 3 years
	5th November 2009	Final extension of QE1 asset purchases to £200bn

Table 2: Event Study for the UK QE Programme

	OIS Rates[1,3]		Government Securities[1,3]				Corporate Bonds[1,3]		Equities[2,4]		
	3 month	1 year	1 year	5 years	10 years	20 years	AAA	BBB	FTSE100	Volatility	NEER[2,4]
11th February 2009	-16	-23	-31	-40	-34	-15	-31	-69	-0.26	-1.28	-1.55
5th March 2009	12	20	5	-34	-68	-89	-91	-41	-3.15	15.2	-0.16
7th May 2009	0	-2	1	9	1	1	10	8	1.49	3.72	-0.84
6th August 2009	2	-4	1	-4	-3	-26	0	-4	1.82	-4.48	-0.80
5th November 2009	2	-3	-3	4	10	5	8	8	0.68	-9.93	-0.05
Sum/Average	0	-12	-27	-65	-94	-124	-104	-98	0.12	0.65	-0.68
Full Sample			-0.2	0.3	0.1	0	-0.5	-0.4	0.22	29.16	0.04

1: basis points change
2: percentage change
3: sum across all events
4: average percentage change per event

Table 3: Event Study for the US QE1 Programme

	OIS Rates[1,3]		Government Securities[1,3]				Corporate Bonds[1,3]		Equities[2,4]		
	3 month	1 year	1 year	5 years	10 years	20 years	AAA	BBB	S&P500	Volatility	NEER[2,4]
1st December 2008	-4	-9	-13	-28	-25	-24	-9		-5.29	13.93	-0.26
16th December 2008	-13	-9	-5	-15	-33	-28	-51		4.13	-12.19	-3.19
28th January 2009	-1	-2	4	28	28	28	13	4	-0.07	0.9	0.2
18th March 2009	0	-4	-9	-36	-41	-35	-31	0	0.76	7.06	-4.64
29th April 2009	0	0	-3	5	11	7	-6	0	2.16	-3.82	-0.72
24th June 2009	-1	1	-3	-13	-1	-5	-4	-24	2.81	-13.8	0.79
12th August 2009	-1	-5	-3	-11	-12	-3	-13	16	1.85	-4.92	-1
23rd September 2009	-1	-2	-3	-8	-6	-4	-8	-4	-1.95	8.1	0.74
4th November 2009	0	-4	-2	-1	7	7	8	5	2.03	-11.73	-0.72
Sum/Average	-21	-34	-37	-79	-72	-57	-101	-3	0.71	-1.83	-0.98
Full Sample			-0.1	0.7	0.9	0.7	0.5	-1.8	0.23	-0.48	-0.08

1: basis points change
2: percentage change
3: sum across all events
4: average percentage change per event

Table 4: Event Study for the US QE2 Programme

	OIS Rates[1,3]		Government Securities[1,3]				Corporate Bonds[1,3]		Equities[2,4]		
	3 month	1 year	1 year	5 years	10 years	20 years	AAA	BBB	S&P500	Volatility	NEER[2,4]
10th August 2010	0	-1	-1	-10	-14	-11	-7	-10	-3.4	14.68	1.37
27th August 2010	1	-1	1	1	4	5	1	15	0.16	-0.58	0.45
21st September 2010	0	-1	-1	-10	-16	-14	-20	-46	-0.74	4.7	-1.35
12th October 2010	1	0	1	2	5	9	14	-3	1.1	0.58	-0.53
15th October 2010	0	0	1	-4	0	3	-27	6	0.93	-3.97	0.56
3rd November 2010	0	0	-1	-11	-10	4	-11	-29	2.3	-14.14	-0.94
Sum/Average	2	-3	0	-32	-31	-4	-50	-67	0.06	0.21	-0.07
Full Sample	0	0	0	-1	-1	0	-1	-1.9	0.24	-0.28	-0.16

1: basis points change
2: percentage change
3: sum across all events
4: average percentage change per event

Table 5: Flow Effect Panel Regression Results

	US QE1				US QE2				UK QE	
	Eligible		Ineligible		Eligible		Ineligible		PR	OPR
	<15 YTM	>15 YTM	<15 YTM	>15 YTM	<15 YTM	>15 YTM	<15 YTM	>15 YTM		
β_1- Direct purchases	0.2763	-0.1063	-	-	0.6072	-0.1850	-	-	0.1220	-
	(0.053)	*(0.098)*			*(0.046)*	*(0.091)*			*(0.021)*	
β_2- Near substitutes	0.2403	-0.1238	0.0665	-0.0268	0.5668	-0.1922	-0.0050	-0.3019	0.0842	0.1681
	(0.048)	*(0.044)*	*(0.018)*	*(0.053)*	*(0.039)*	*(0.038)*	*(0.021)*	*(0.120)*	*(0.014)*	*(0.059)*
β_3- Mid substitutes	0.1700	0.0501	0.0047	-0.007	0.5732	-0.0860	-0.0370	-0.1661	0.0680	0.1806
	(0.045)	*(0.026)*	*(0.0099)*	*(0.021)*	*(0.036)*	*(0.070)*	*(0.017)*	*(0.041)*	*(0.013)*	*(0.024)*
β_4- Far substitutes	-	-	-0.0238	0.0021	-	-	-0.0807	-0.1026	0.0341	-0.0502
			(0.008)	*(0.003)*			*(0.013)*	*(0.010)*	*(0.013)*	*(0.014)*
# observations										
# securities										
R^2										

Note: results for US QE1 are taken from D'Amico and King (2010)
PR= Securities inside the purchase range
OPR= Securities outside the purchase range

Table 6: Stage One Instrumented Variable Regression

	US QE2	UK QE
Remaining maturity	0.0000559	0.00497
	(0.000)	*(0.002)*
Squared remaining maturity	-0.000000126	-0.000013
	(0.000)	*(0.000)*
Svennsson fitting error	-0.0070341	0.4741
	(0.003)	*(0.204)*
Fitting error of near substitutes		-0.3635
		(0.0191)
Old bond dummy	-0.0020395	
	(0.001)	
R^2		

Table 7: Second Stage Stock Effect Regression Results

	US QE2		UK QE	
	<15 YTM	>15 YTM	IPR	OPR
Direct purchases	2.351	3.215	0.1583	-
	(1.049)	*(0.022)*	*(0.062)*	
Near Substitutes	0.031	-0.146	-0.0283	0.010
	(0.022)	*(0.231)*	*(0.025)*	*(0.022)*
Remaining maturity	-0.146			-
	(0.000)			
Squared remaining maturity	0.0000913			-
	(0.000)			
Intercept	0.00128		-0.2589	
	(0.003)		*(0.131)*	
R^2				

PR= Securities inside the purchase range
OPR= Securities outside the purchase range

Table 8: ADF tests

	Number of lags			
	0	1	2	3
y^{10}	-1.812	-2.080	-1.937	-2.018
M^F	0.4161	0.2736	-0.7905	-0.9853
M^T	-1.104	-1.043	-0.9577	-1.256
Q	-1.023	-0.9196	-1.251	-1.580
r	-1.272	-1.768	-2.090	-2.296
π^e	-3.025	-2.968	-2.807	-2.631
VIX	-4.347	-4.510	-3.698	-3.278
GDP^e	-2.626	-2.908	-3.099	-3.212
ADF critical values: 5%=-2.87 1%=-3.46				

Table 9: Error-correction Regression Results

	Variable	Baseline
Short-run dynamics	Δy_{t-1}	0.29
		()
	Δr_t	0.15
		()
Speed of adjustment		-0.11
Long run relationship		()
	M^F_{t-1}	0.034
		(0.004)
	M^T_{t-1}	-0.070
		(0.007)
	Q_{t-1}	0.202
		(0.012)
	r_{t-1}	-0.220
		(0.024)

Table 10: Alternative specifications for time series exercise

	Pre-crisis sample 1990-2007	Inflation expectations	Forward rates	Alternative measure of maturity	Average maturity of issuance	Adjusting for balance sheet size
M^F	–0.080	-0.018	-0.049		-0.034	
	(0.018)	(0.003)	(0.004)		(0.008)	
M^T	0.093	0.052	0.096			
	(0.013)	(0.006)	(0.008)			
Q	–0.126	-0.11	-0.239	-0.011	-0.08	-0.057
	(0.035)	(1.326)	(0.014)	(0.019)	(0.032)	(0.019)
r	0.262	0.15		0.178	0.15	0.233
	(0.027)	(0.021)		(0.024)	(0.051)	(0.025)
π^e		0.86		1.00	0.89	
		(0.080)		(0.093)	(0.156)	
r^e			0.059			
			(0.043)			
alt-M^F				-3.977		
				(0.624)		
alt-M^T				3.909		
				(1.711)		
M^I					0.010	
					(0.006)	
M^F-adj						-0.268
						(0.035)
M^T-adj						0.083
						(0.007)

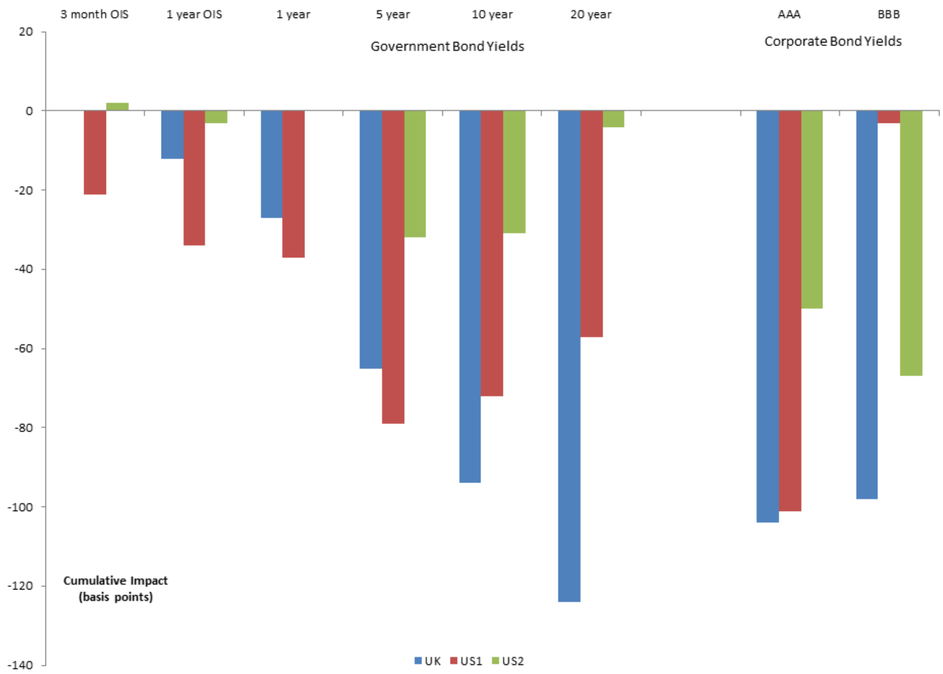

Figure 1: Cumulative Impact of Event Studies

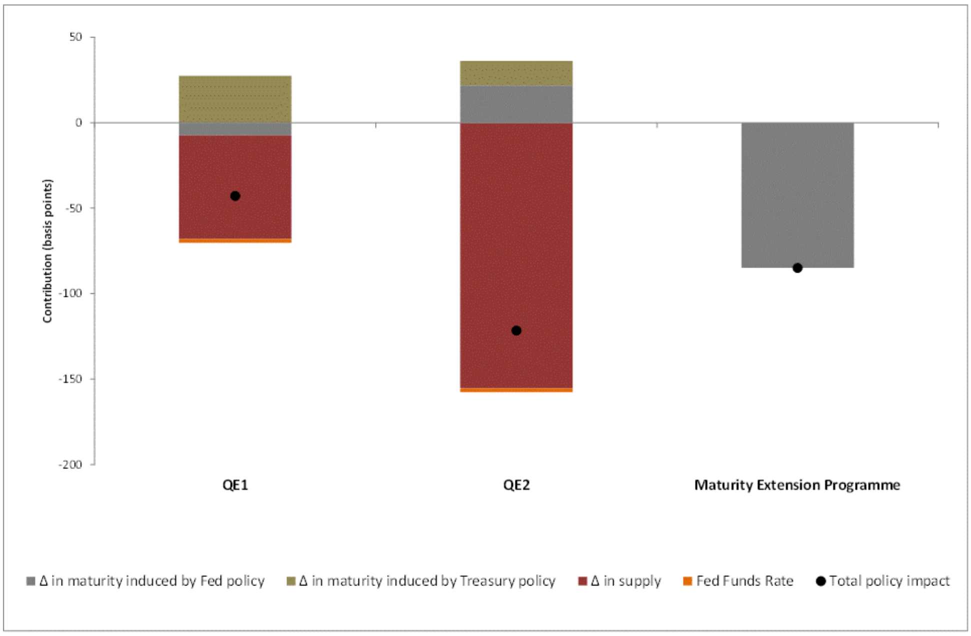

Figure 2: Contributing policy factors determining change in US 10 year treasury yield

The Financial Crisis, QE and the Money Multiplier[*]

Jagjit S. Chadha[†] Luisa Corrado[‡] Jack Meaning[§]

Abstract

We analyse the collapse in the money multiplier since the start of the financial crisis, and the failure of unprecedented expansions of central bank money to bring about large increases in broad money. The multiplier is reframed as an optimal portfolio decision on the part of commercial banks in meeting deposit demand by holding assets as reserves or extending credit, conditioned on the prevalent interest rates and spreads. We show that variation in the bank multiplier can enhance welfare for consumers as a given demand for deposits can be provided with less volatility in financial spreads. Our model shows that it is possible to match the observed fall in the money (bank) multiplier consistent with that observed in the data in response to a composite shock replicating the 2007-2008 crisis. We also run a counterfactual to suggest if the Federal Reserve had not expanded narrow money to such a degree and commercial banks had maintained a fixed multiplier, broad money would have fallen by significantly more and the economy may have experienced a contraction in output 1% greater, inflation would have fallen by an additional 2.5% and employment would have been 2% lower. The recovery would also have been more protracted, taking roughly twice as long to return to equilibrium.

[*]A version of this paper was presented at the Bank for International Settlements, the Bundesbank, the ECB and the Bank of Greece, Bank of Thailand, the University of Cambridge, University of Kent, University of Loughborough, University of Oxford, University of York and the FSA.

[†]Professor of Economics, Chair in Banking and Finance, School of Economics, Keynes College, University of Kent, Canterbury, UK, CT2 9LP. E-mail: jsc@kent.ac.uk.

[‡]Associate Professor of Economics, Department of Economics, University of Rome, Tor Vergata. E-mail: lc242@cam.ac.uk.

[§]Corresponding Author. National Institute of Economic and Social Research, UK. E-mail: j.meaning@niesr.ac.uk.

1 Introduction

The money multiplier, the ratio of narrow money and broader monetary aggregates, is often an integral part of macroeconomic models and a key aspect of the judgement on monetary policy. However, the collapse of the bank money multiplier since 2008, and the failure of huge expansions of narrow money in many advanced economies to bring about a sustained growth in broad money has led some commentators to announce not only the death of the multiplier as a useful concept but also to pronounce on the failure of non-conventional monetary policies. Broadly speaking the bank multiplier has not really grown much from the basic proposition that the commercial banking sector simply lends out some fixed multiple of what is otherwise held as cash and reserves.[1] The fallacy of this assumption is immediately apparent if one looks at any data series on money multipliers, especially since the current financial crisis. In a paper focussed on the empirical evaluation of the relationship between central bank reserves and broader money and bank lending, Carpenter and Demiralp (2012) find little evidence supporting the standard money multiplier story, but suggest that a more complex relationship between monetary policy and the quantity of lending, one where demand for loans and reserves, which are dependent on interest rates, could play a role. Accordingly in this paper we try to impose a more formal optimization problem for commercial banks in the management of the asset side of their balance sheet.

Of course, the majority of central banks conduct monetary policy through manipulation of the short-term nominal interest rate.[2] And many models employing the short-term nominal interest rate do still characterize monetary policy as controlling the quantity of central bank reserves in order to achieve their desired rate i.e. withdrawing supply through open market operations to raise rates and increasing it if they are targeting a lower rate. Disyatat (2008) explains how this amounts to an exercise in reverse causality. What happens in practice is that the policy maker sets the lending rate it will lend at and then provides, through open market operations, as much narrow money as commercial banks demand at that rate. Narrow money is thus demand determined. Commercial banks then supply loans at the market interest rate, which determines broad money. A key insight here is that the evolution of monetary aggregates 'reflects

[1]See Garfinkel and Thornton (1991) for a precursor to our argument that the multiplier cannot be thought to be exogenous and ought to be interpreted with reference to policy and underlying economic conditions, including financial spreads.

[2]The consensus rationale for this can be found in Poole (1970).

prevailing interest rates rather than determines them'.[3]

We suggest thinking about variations in the money multiplier as an optimal portfolio choice on the part of commercial banks that matches deposit demand from households with a choice about assets as reserves or loans.[4] The use of non-conventional monetary policies, which have increased the supply of base money to commercial banks allows us a background against which to judge our model. Accordingly, we build on the money and banking sector model of Goodfriend and McCallum (2007) in which the overall level of deposits is determined by narrow money from the central bank, and loans produced by commercial banks. However unlike theirs which has a fixed ratio between the two monetary aggregates, we derive a demand schedule for narrow money from banks' profit-maximising condition and allow this to determine the level of narrow money, with the central bank meeting this demand perfectly elastically. The quantity of loans is then derived from the bank's optimising condition and loan production function and so we can observe richer dynamics in the money multiplier as the two aggregates evolve in response to macroeconomic and financial conditions. A further innovation is the inclusion of open market operations as a practical mechanism to effect these changes in reserves. We model the central bank exchanging newly created reserves with the private sector for deposits, which are backed up with increased reserve holdings on commercial banks' balance sheets.

The rest of the paper is as follows. In Section 2, we briefly outline recent developments in the money multiplier and discuss what may have been driving them with particular regard to policy in the US. We then provide an exposition of the model. In Section 4 we discuss the implications of our model for the money multiplier, which is the inverse of the reserve-deposit ratio, and how this affects banks and the wider economy. We show that, by allowing the multiplier to vary as the relative returns/costs of narrow and broad money change, banks can remain optimal in how they meet deposit demand. To illustrate these points more clearly in Section 6 we present the impulse responses of the models key variables to both real and financial shocks and show that a flexible money multiplier improves the response of the macroeconomy when compared to a model where the multiplier is held fixed. To strengthen this conceptual point, in Section 7 we carry out welfare analysis by deriving a welfare loss function for the representative household from a second order approximation to utility.

Finally, in Section 8 we apply the model to the recent financial crisis and

[3] See Chadha, Corrado and Holly (2014) for a thorough analysis of this issue.

[4] See Gale (2011) on this point.

find it does a passable job of matching the evolution of key variables in the US economy since 2008. We then investigate what the results of the model imply for the money multiplier and find that following a shock similar to that experienced by the US during the financial crisis, it is optimal for banks to demand much more narrow money and for credit money in the form of loans to contract. This leads to a significant fall in the money multiplier consistent with that observed in the data. We then perform a counterfactual exercise in which we subject the model to the same shock but hold the money multiplier fixed. The model implies that, had the Federal Reserve not massively expanded narrow money and maintained a fixed multiplier, broad money would have fallen by significantly more and the US economy may have experienced a contraction in output 1% greater, inflation would have fallen by an additional 2.5% and employment would have been 2% lower. The recovery would also have been more protracted, taking roughly twice as long to return to equilibrium.

2 Money, Multipliers and Recent Policy

The fall of money multipliers in the World's major economies has been a feature of the recent crisis. Figure 1 shows how the ratio of broad to narrow money in the US fell rapidly as the crisis intensified. However, what this also shows is that the level of broad money proved relatively resilient. So, what has driven this fall in multipliers? The first factor driving it is the contraction of new credit and lending by commercial banks to the wider economy. After suffering heavy losses, and facing increased uncertainty, banks have become more cautious in their new lending and sought to repair their balance sheets and shrink loan books. The second factor, which is clear in Figure 1 that the Federal Reserve has massively expanded the quantity of central bank money through the quantitative easing programmes. Figures 2 and 3 show how the Federal Reserve's balance sheet has evolved with an unprecedented rise in the level of reserves, mostly effected through the Large Scale Asset Purchases (LSAPs) in which the Fed purchased mostly Government bonds and Mortgage-Backed Securities with newly created reserves. In this way, the LSAPs can be thought of as analogous to traditional open market operations, albeit differing in the quantity and type of assets purchased, and the duration for which the assets are held.

In November 2008 the Fed announced it would begin purchasing housing agency debt and mortgage-backed agency securities to the value of $600bn in response to the housing crisis and in order to promote the health of mortgage lending. In March 2009 this was increased to $1.25 trillion. These purchases were largely of maturities between 3 months and 5 years. As they reached maturity, the principal was reinvested to fund the purchase of Treasury securities and maintain the value of the agency debt and agency backed securities section of the LSAP. Accompanying this extension was the announcement that the Fed would buy $300bn of Treasury securities, over 60% of which were of between 3 and 10 year maturities. The purchase of Treasuries was designed to support falling asset prices by acting as a large buyer and through the portfolio balance channel this should spread to other assets in the economy.

These large scale asset purchases were predominantly funded by the creation of several trillion dollars of new reserves, making them the largest quantitative easing programme enacted since the crisis. In November 2010, in light of a continuing weakness in economic forecasts, the purchase of longer-term Treasuries was extended further by $600bn under a second round of quantitative easing (QE2) which took the total LSAP to over $2 trillion. In September 2011, the FOMC announced a maturity extension programme under which it bought an additional $400bn of longer-dated treasuries but simultaneously sterilised these by selling short-term Treasuries to the same value. The goal was to lower longer-term yields without increasing the size of the central bank's balance sheet by "twisting" the yield curve and increasing the average maturity of the Fed's Treasuries portfolio by 25 months. In September 2012 the Fed announced QE3. This heralded an important change in the setting of asset purchase programmes as, unlike the previous policies, QE3 was open-ended, with the Federal Reserve committing to buy $40bn of MBS a month until the outlook for the labour market "improves substantially....in a context of price stability". It was also a direct injection of liquid reserves into the economy aimed at improving confidence and conditions in impaired credit markets. The purchases were eventually halted in October 2014 after accumulating some $4.5Tn in assets. By explicitly tying the duration of the programme to conditions in the macroeconomy the Federal Reserve hoped to relieve uncertainty and return confidence to agents to make longer-term decisions about investment and spending.

If the assumption of a fixed money multiplier were to hold, this huge expansion in narrow money would be expected to lead to an equivalent boom in broad money. Similarly, the contraction in the credit supply and loan issuance of commercial banks would act to reduce broad money. As already observed though, the level of broad money has been reasonably constant, implying that the increase in narrow money has worked to almost exactly offset the contraction in credit money. Whilst these two effects pull in opposite directions on the broad monetary aggregate, they both act to move the money multiplier in the same way, down, as narrow money has formed an increasing fraction of the total money supply. We will go on to show how this fall in the multiplier may be explained by banks' optimising behaviour causing them to rebalance their portfolio between narrow and credit money as the relative returns/costs of each vary and this may have acted to reduce the duration and severity of the economic downturn by offsetting an escalation in market-determined premia.[5]

2.1 The Central Bank Balance Sheet

We suggest a simple framework for analyzing the effect of balance sheet monetary policies.[6] For simplicity, since we abstract from other forms of central bank money and concentrate on bank reserves alone in our model, high powered money is identical to reserves. More traditionally the central bank controls the stock of fiat money (outside money) and financial intermediaries create other forms of money, which are claims on the private sector. As financial intermediation allows alternative assets to serve as money, it offers a close substitute to (outside) fiat money and the ability of the central bank to determine the overall nominal level of expenditure depends on the relationship between outside and inside money. The central bank has a powerful tool to regulate financial intermediaries and to affect the quantity of money in circulation: reserves, which may be either or both of fractional and or voluntary.

Private Sector			
Assets		Liabilities	
Deposits	D	Loans	$(D-r)$
Bonds	γB	Tax	$\sum_{i=0}^{\infty}\beta^i t_i$
Capital	$\gamma_k K$	Capital	K

Government	
Assets	Liabilities
Tax $\sum_{i=0}^{\infty}\beta^i t_i$	Bonds B

Commercial Banks			
Assets		Liabilities	
Reserves	r	Deposits	D
Loans	$(D-r)$		

Central Bank			
Assets		Liabilities	
Bonds	$(1-\gamma)B$	Reserves	r
Capital	$(1-\gamma_k)K$		

[5] See Walsh (2009) for a very interesting attempt to understand policy at the ZLB.
[6] See Chadha, Corrado and Meaning (2012) for more detail.

We first look at the private sector's balance sheet. The private sector has three forms of assets: deposits, D, held at banks and some fraction of bonds, γB, issued by the government and a fraction of total capital.[7] Their liabilities are loans, $D - r$, provided by banks and capital, K. Capital is also on the liability side of households' balance sheet because households own the firms and firms treat capital as a liability. The government sector has liabilities in the form of outstanding public debt, B and assets given by the present discounted value of future taxation. The commercial banks' balance sheet liabilities are deposits, D. Some fraction of liabilities, r, is held as reserves and the rest, $D - r$, is available to be lent to the private sector. The central bank holds assets in the form of some fraction of government bonds, $(1-\gamma)B$, and a fraction of capital, $(1-\gamma_k)K$, with liabilities determined by central bank money, which are reserves in this model.[8] The net assets of commercial banks and of the central bank are both zero. The private sector has net assets given by $D + \gamma B + \gamma_k K - \left(D - r + \sum_{i=0}^{\infty} \beta^i t_i\right)$ and so because $r = (1-\gamma)B + \gamma_k K$ and $\sum_{i=0}^{\infty} \beta^i t_i = B$, we can note that the net private sector assets are also zero.

We can see from this flow of funds the mechanism by which unconventional policies occur. The central bank can perform quantitative easing by increasing the size of its balance sheet. It does this by extending an increased level of reserves to commercial banks which must be backed by an increased holding of either bonds or capital, which in turn must be bought from the private sector. Alternatively, credit easing is conducted through the composition of the balance sheet. With their liabilities unchanged, the central bank can buy capital from the private sector, increasing its own holdings. It funds these purchases by selling bonds back to the private sector, leaving the net effect on the size of both the central bank and private sector's assets at zero. Due to the differing properties of bonds and capital as collateral in loan production, this exchange has implications for levels of deposit demand which we will discuss later.

3 The Model

We now present the model, which is an extended version of that developed by Goodfriend and McCallum (2007) and extended by Chadha and Corrado (2012). Primarily it is a Calvo-Yun monopolistically competitive production economy with sticky prices and three main agents; households, who can work either in

[7]In this example we assume that the private sector is represented by households.

[8]If we operate in an open economy, central bank assets would also include foreign exchange reserves r^f.

the goods producing sector or in the banking sector monitoring loan quality, banks, who meet consumers' deposit demand via reserves and a loans production function, and the monetary authority.[9]

3.1 Households

Following Goodfriend and McCallum (2007) we present the model in terms of the optimising problem of the typical household before then aggregating up under the assumption that there exist many similar households.

The typical household in our model consumes a bundle of differentiated commodities, supplies labor, and saves. They also own and operate the firm and produce for sale a differentiated good. Lastly they also operate a (competitive) banking firm.

Their utility function can be written

$$U = E_0 \sum_{t=0}^{\infty} \beta^t [\phi \log(c_t) + (1-\phi)\log(1 - n_t^s - m_t^s)] \tag{1}$$

where c_t is consumption during period t of Dixit-Stiglitz consumption bundles and $(1-n_t^s-m_t^s)$ represents leisure. n_t^s and m_t^s represent labour provided to goods producing firms and banking firms respectively, though these need not necessarily be those firms owned by the household itself.

The household then maximises this utility subject to three constraints. The first is a budget constraint which states that the household has income in any period equal to the net sale of capital goods, $q_t(1-\delta)K_t - q_tK_{t+1}$, the net sale of government bond holdings $\frac{\gamma B_t}{P_t^A} - \frac{\gamma B_{t+1}}{P_t^A(1+R_t^B)}$, income from changes in deposit balances $\frac{D_{t-1}}{P_t^A} - \frac{D_t}{P_t^A}$, net labour income $w_t(n_t^s + m_t^s) - w_t(n_t + m_t)$ and what it can earn for sale of the differentiated good it produces less the cost of its own consumption basket, $c_t^A(\frac{P_t}{P_t^A})^{1-\theta} - c_t$. There is also a lump-sum tax payment, and a lump transfer from the banking sector. This results in

$$\begin{aligned} &q_t(1-\delta)K_t + \frac{\gamma B_t}{P_t^A} + \frac{D_{t-1}}{P_t^A} + w_t(n_t^s + m_t^s) + c_t^A(\frac{P_t}{P_t^A})^{1-\theta} + \Pi_t \\ &-q_tK_{t+1} - \frac{\gamma B_{t+1}}{P_t^A(1+R_t^B)} - \frac{D_t}{P_t^A} - w_t(n_t+m_t) - c_t - tax_t = 0 \end{aligned} \tag{2}$$

where q_t is the price of capital and K_t is the quantity of capital. P_t is the price

[9]There is also a fiscal authority which runs a balanced budget. We fix levels of government debt as constant unless exogenously shocked, thus the government's role within this set-up is benign.

of household's produced good while P_t^A is the consumption good price index. n_t is the labour demanded by household as producer, m_t, is the labour demanded by household's banking operation. n_t^s and m_t^s are the quantities of labour supplied by the household to each sector for which it is paid the real wage, w_t. D_t is the nominal holding of broad money, tax_t is the real lump-sum tax payment, R_t^B is the nominal interest rate on government bonds purchased in $t+1$, B_{t+1}.

c_t^A is the aggregated consumption in the economy, which the individual household takes as given, but the household has market power in the product that it supplies θ is the elasticity of substitution relating to that good so $c_t^A(\frac{P_t}{P_t^A})^{1-\theta}$ represents the share of the aggregate consumption the typical household supplies. Π_t is a transfer of profits from the banking firm owned by the household. The Lagrange multiplier of this constraint is denoted as λ_t.

The second constraint is a deposit-in-advance constraint which establishes bank deposits as the means of transaction within our model and requires the household to hold deposits with a financial intermediary in a given period in order to effect any consumption expenditure.

$$c_t = v_t D_t / P_t^A \tag{3}$$

This is a key element of the model, and provides a direct link between consumption demand/economic activity and demand for broad money.

As owners of a firm, the typical household is also subject to a sales equals net production constraint

$$K_t^{\eta}(A1_t n_t)^{1-\eta} - c_t^A(P_t/P_t^A)^{-\theta} = 0, \tag{4}$$

where η denotes the capital share in the firm production function and $A1_t$ is a productivity shock in the goods production sector whose mean increases over time at a rate ϱ As above $c_t^A(P_t/P_t^A)^{-\theta}$ reflects the monopolistic competition the firm faces when selling its own differentiated good. Thus the constraint requires that everything the firm produces is sold. The Lagrange multiplier of this constraint is denoted as, ξ_t

Lastly, price adjustment is subject to a nominal rigidity that follows the widely-used Calvo (1983) price mechanism.

$$\frac{P_t}{P_{t-1}} = \beta E_t \frac{P_{t+1}}{P_t} * mc_t^{\kappa} \tag{5}$$

where P_t in this setting refers also to P_t^A as all households and firms face the same economy wide inflation rate. The marginal cost MC_t depends on the lagrangian multipliers from our household budget constraint and the firm's sales equals net production constraint

$$mc_t = \frac{\xi_t}{\lambda_t} \tag{6}$$

3.2 Banks

Deposit demand is determined by the household. It is then met by the financial sector, whih is where our model begins to deviate from that of Goodfriend and McCallum (2007).

Deposits are created in two ways in our model; they can be created by the monetary authority through the creation of narrow money, which appears as central bank reserves on the balance sheets of commercial banks; alternatively deposits can be created by banks themselves who produce loans and in doing so generate a deposit for the household. This view of broad money creation is outlined and discussed in more detail in [reference].

From this, we can write a simple identity which describes both broad money and the balance sheets of the commercial bank.

$$D_t = L_t + r_t \tag{7}$$

where L_t is the total amount of one period loans supplied in time t and r_t is the quantity of narrow money of reserves supplied by the central bank. $\frac{D}{r}$ therefore represents the money multiplier and, as the only source of narrow money in our model is reserves, $\frac{1}{MM} = \frac{r}{D}$, which equals the reserve ratio.[10]

To produce loans, and thus create deposits independently of the central bank, banks apply a technology to collateral posted by households in the form of bonds, b, or capital, qK. This process is captured by a standard Cobb-Douglas production function for loans where collateral is combined with monitoring work, m.

$$L_t/P_t^A = F(\gamma b_{t+1} + A3_t k q_t K_{t+1})^{\alpha}(A2_t m_t)^{1-\alpha} \qquad 0 < \alpha < 1, \tag{8}$$

$A2_t$ denotes a shock to monitoring work, $A3_t$ is a shock on capital as collateral and $b_{t+1} = B_{t+1}/P_t^A(1+R_{t+1}^B)$. The parameter k denotes the inferiority of capital

[10] Under a 100% reserve system, the broad money supply, and thus consumption within our model would be restricted by the creation of narrow money by the central bank. In this variant $D_t = r_t$ and the subsequent problem of reserve demand simplifies to depend purely on demand for consumption at the given policy rate.

as collateral in the banking production function[11], while α is the share of collateral in loan production. Increasing monitoring effort is achieved by increasing the number of people employed in the banking sector and therefore reducing the employment in the goods production sector, for a given level of total employment.

Alternatively, deposits can be created by the central bank. They can do this at no cost to themselves. However, they do charge a rate of interest for supplying these reserves to the banking sector. They do this to be able to control the prevailing benchmark interest rate in the economy, operating much like open market operations at a discount window in practice. In actuality, the rate of importance here is the difference between the rate that must be paid to borrow reserves and the rate paid on reserve holdings stored at the central bank. For simplicity of exposition however we set the rate of interest paid on stored reserves as zero, and so this collapses to be just the rate paid at the discount window, R_t. We assume that the monetary authority supplies reserves perfectly elastically to satiate demand from commercial banks at the given policy rate. We derive the underlying demand function for reserves below.

3.2.1 The bank's problem

The banking sector itself is perfectly competitive and following Baltensperger (1980) banks seek to maximise total returns within period. Thus their optimisation problem can be written

$$\max_{r_t} \Pi_t = R_t^L L_t - R_t r_t - R_t^D D_t, -\frac{1}{2} R_t^T (\bar{r} - r_t)^2 - \tau_t(\bar{r} - r_t) \qquad (9)$$

The latter half of this equation relates to a series of additional costs faced by commercial banks and motivated by concerns over reserve management. We assume that banks have an exogenous target for the level of reserves, $\bar{r}$, perhaps set by custom and practice or by legislation.[12] We assume that any deviation from this target imposes two costs on the bank. The first is symmetric and derives from the bank's desire to smooth the path of reserves and avoid any sharp swings in its asset position as these may signal mismanagement and result in reputational loss. In the model the cost of such deviations from target is the uncollateralised interest rate, R_t^T. This is because if $r_t < \bar{r}$, the commercial bank will fund its shortfall at the penalty rate, and if $r_t > \bar{r}$ the commercial bank will have missed

[11] Capital is considered inferior as there are increased costs to the bank of verifying its physical quality and condition as well as its market price. It is also less liquid should it need to be called upon in the case of default.

[12] In practise, most major economies have a minimum level of required reserves relative to depositis.

the opportunity to lend out those reserves at the same penalty rate and thus incurs the opportunity cost, R_t^T.

The second term relates to the need of commercial banks to hold a certain level of reserves to meet its desired reserve target in any given period. Whilst exogenous in our framework, this target is most likely driven by the level of required reserves set by the regulator (Basel III, 2010), although banks may set a target in excess of this minimum limit if they have heightened precautionary motives for holdings safe, liquid assets, such as central bank reserves. Therefore, the second term can be thought of as an exogenous shift in the ex-ante probability of a reserve shortfall. It represents shifts in the level of reserves necessary to meet the bank's target holdings, so an increase in τ corresponds to bank reserves being below the target level $\bar{r}$.

Lastly, banks are constrained in that they must fully meet the resultant demand for deposits using one of the two methods at their disposal. Therefore

$$D_t = L_t + r_t \tag{10}$$

is a hard constraint. It can be used to substitute out for L_t, leaving the maximisation problem as

$$\max_{r_t} \Pi_t = R_t^L(D_t - r_t) - R_t r_t - R_t^D D_t, -\frac{1}{2}R_t^T(\bar{r} - r_t)^2 - \tau_t(\bar{r} - r_t) \tag{11}$$

Solving this problem for the optimal level of reserves gives us the commercial bank's demand curve for reserves from the central bank, and can be written as

$$r_t = \bar{r} - \left[\frac{R_t + R_t^L}{R_t^T}\right] + \frac{\tau}{R_t^T} \tag{12}$$

t is a positive function of the probability of the commercial bank being short its obligated level of reserves, $\bar{r}$, and negative in $\hat{R}_t$, and, $\hat{R}_t^L$, which is the cost of reserves and the opportunity cost of reserves, the loan rate. We therefore emphasize that the relative cost and returns of the two mechanisms of meeting deposit demand change, so do the bank's optimal quantities of each.

Any consumption above the level equivalent to the quantity of reserves requires households to borrow from the commercial banks, and thus receive a deposit which can be used to effect transactions.[13] Loan demand can thus be

[13] In this way our model differs from others where loans are used to fund investment and have a direct productive purpose. However, by creating broad money and funding consumption they are acting, through the sales equal net production constraint, to achieve a similar effect.

pinned down to the difference between deposit demand and the quantity of reserves demanded by the commercial bank. By combining equations (6), (7) and (10) we can write the quantity of credit money the banking sector should extend in the form of loans as

$$L_t = D_t - \bar{r} + \left[\frac{R_t + R_t^L}{R_t^T}\right] - \frac{\tau}{R_t^T}$$

We can then show how total deposits depend on both money created by the central bank in response to commercial bank demands and that created by the banking sector through loans.

$$D_t = \bar{r} - \left[\frac{R_t + R_t^L}{R_t^T}\right] + \frac{\tau}{R_t^T} + F(\gamma b_{t+1} + A3_t k q_t K_{t+1})^{\alpha}(A2_t m_t)^{1-\alpha}.$$

Finally, we can put this back in to the deposit-in-advance constraint which allows us to neatly show the interconnectivity between the real economy, the financial sector and the bank's decision between monetary aggregates:

$$c_t = v_t \frac{F(\gamma b_{t+1} + A3_t k q_t K_{t+1})^{\alpha}(A2_t m_t)^{1-\alpha}}{P_t^A(1 - rr_t)}, \tag{13}$$

which can be re-written as:

$$c_t = v_t \left[\underbrace{\frac{\tau_t - R_t - R_t^L}{R_t^T P_t^A} + \bar{r}}_{base} + \underbrace{\frac{F(\gamma b_{t+1} + A3_t k q_t K_{t+1})^{\alpha}(A2_t m_t)^{1-\alpha}}{P_t^A}}_{credit}\right].$$

The first two terms inside the square bracket represent narrow money, whilst the third is money created by the banking sector. If it becomes optimal for commercial banks to create less credit money through loans, then the only way to support a given level of broad money, and thus activity, is to increase the quantity of narrow money equivalently. As the central bank is the only agent who can create narrow money, they must increase the supply of reserves, and in our benchmark framework this is exactly what they do, meeting the increased demand for reserves perfectly elastically. This allows commercial banks to always achieve their optimal asset mix. We will later discuss the implications of an alternative scenario in which the central bank only provides reserves as a constant fraction of total deposits (a fixed money multiplier).

3.3 Interest rates and spreads

The inclusion of such a banking sector as outlined above gives rise to a number of interest rates and financial spreads. First, we derive a risk-free rate of interest by imagining a one-period default-free security that cannot be used as collateral by the holder. From the household's problem above we can write the return on this security R_t^T, as a standard intertemporal nominal pricing kernel, priced from expected real consumption growth and inflation. In essence it is a one period Fisher equation.

$$1 + R_t^T = E_t \frac{\lambda_t P_{t+1}}{\beta \lambda_{t+1} P_t} \tag{14}$$

The policy rate of the central bank is the market clearing rate for reserves. Implicitly this is also the interbank rate, although given our assumptions on the banking sector we do not explicitly model interbank transactions. The rate is set by a feedback rule responding to inflation, π_t, and output, y_t, with parameters, ϕ_π and ϕ_y, respectively. Policy rates are smoothed so that $1 > \rho > 0$:

$$R_t = R_{t-1}^{\rho}(1-\rho)(\pi_t^{\phi_\pi} * y_t^{\phi_y}) \tag{15}$$

How this rate is connected to other rates within our model is of clear importance. First, should a bank wish to make an uncollateralised loan at the risk-free rate it would need to incur all of the costs associated with loan production through monitoring work. This cost can be characterised by the factor price of monitoring, which is the real wage $\frac{w_t}{P_t^A}$, divided by the partial derivative of the loan production function with respect to employed monitoring work m_t which is

$$(1-\alpha)\left[\frac{L_t}{P_t}\right] * m_t \tag{16}$$

where loans are defined by the following relationship $L_t = D_t(1 - rr_t) = \frac{c_t P_t^A}{v_t}(1 - rr_t)$. This means the real cost of loan management on an uncollateralised loan can be written as

$$\frac{w_t v_t m_t}{(1-\alpha)(1-rr_t)c_t}. \tag{17}$$

In theory, the commercial bank could obtain funds from the central bank at R_t, applying the required monitoring work as described above and then loaning those funds on at R_t^T. This creates a no-arbitrage condition and means that we can describe the relationship between the risk-free rate and the central bank's policy rate as

$$1 + R_t^T = (1 + R_t)\left[1 + \frac{w_t v_t m_t}{(1-\alpha)(1-rr_t)c_t}\right] \tag{18}$$

which can be approximated to

$$R_t^T - R_t = \frac{w_t v_t m_t}{(1-\alpha)(1-rr_t)c_t} \tag{19}$$

This spread can be thought of as an uncollateralised finance premium.

However, in practice most loans made to households will be collateralised to some extent, taking some of the burden of loan management away from monitoring work. The bank only takes the burden of the share of production costs associated with monitoring and so in a collateralised loan we must multiply our premium derived above by that share, $1-\alpha$. The central bank's rate and the loan rate can then be linked by the following relationship

$$R_t^L - R_t = \frac{w_t v_t m_t}{(1-rr_t)c_t} \tag{20}$$

where

$$\frac{w_t v_t m_t}{(1-rr_t)c_t} \tag{21}$$

is the collateralised external finance premium, EFP_t. This is the real marginal cost of loan management, and it is increasing in velocity, v_t, real wages, w_t, monitoring work in the banking sector, m_t, and the reserve ratio, rr_t, and decreasing in consumption, c_t. Recall that $rr_t = \frac{1}{MM_t}$ so the EFP is also decreasing in the money multiplier, meaning that in this model banks switch to narrow money taking more of the burden of meeting deposit demand, when the external finance premium is higher.

The yield on government bonds is derived by maximising households' utility with respect to bond holdings and rearranging to give the spread.

$$R_t^T - R_t^B = \left[\frac{\phi}{c_t \lambda_t} - 1\right]\Omega_t \tag{22}$$

This spread can be interpreted as a liquidity premium on bonds. $c_t\lambda_t$ measures the household marginal utility relative to households shadow value of funds while Ω_t is the marginal value of the collateral. It is in fact these key margins - the real marginal cost of loan management versus the liquidity service yield - that determine the behavior of spreads. In the above expression, ϕ denotes the consumption weight in the utility function whereas λ_t is the shadow value of consumption, c_t. The interest rate on deposits is the policy rate, R_t, minus a term in the reserve deposit ratio:

$$R_t^D = R_t - \frac{rr}{1-rr} rr_t. \tag{23}$$

As these spreads influence the portfolio decisions of banks they will also impact on the resulting path of consumption. When we come to the analysis of the model we will discuss these premia as a way of understanding our key results.

4 Monetary Policy

This model framework allows us to capture a number of interesting elements of monetary policy. In our benchmark model the direct instrument of monetary policy is the short-term nominal interest rate, which we have seen is set in response to a standard Taylor rule. By varying this rate, the policy maker is changing the cost to commercial banks of obtaining reserves. The endogenously determined external finance premium also changes the return on loans. The liquidity premium impacts on the value of collateral available to households and the deposit rate, a cost of funding, is a negative (positive) function of the reserve-deposit ratio (money multiplier). These effects will change the opportunity cost of meeting deposit demand with narrow money from the central bank rather than extending loans and cause the bank to reset its portfolio mix between narrow money and loans. We will show in Section 6 why this matters for the wider economy.

4.1 Open Market Operations: A Mechanism to Control Reserves

Under non-conventional and conventional monetary policies, the central bank also varies the size of its balance sheet, increasing or decreasing the quantity of reserves in the economy to meet the demand of commercial banks perfectly elastically at its target policy rate.[14] Previous models have lacked a realistic mechanism by which the quantity of reserves in the economy can be controlled by the central bank. We aim to more accurately approximate the practicalities of reserve management by modelling open market operations whereby an asset, primarily bonds, is bought from the private sector in exchange for newly created money. The central bank now holds more bonds on its balance sheet. The private agent from whom the

[14] See Berrospide (2012) for an analysis of the demand for liquidity.

bonds have been purchased receives a newly created deposit in their commercial bank account, whilst their commercial bank's own account with the central bank is credited with an equal increase of freshly created reserves.[15]

To incorporate this mechanism into our model we assume the central bank must match its only liability, reserves, by holding just one class of assets, government bonds, the total supply of which is held fixed unless exogenously shocked.[16] When the central bank buys bonds through an open market operation it increases the fraction of the total bond supply which it holds, and decrease that held by the private sector. We can therefore define total bond holdings as the sum of private sector and central bank bond holdings,

$$b_t = b_t^{CB} + b_t^{P}. \tag{24}$$

As central bank bond holdings must equal reserves, we can substitute and re-arrange to give

$$b_t^p = b_t - r_t \tag{25}$$

It is this newly defined variable b^p which determines the amount of collateral households have available and thus b^p which features in our equations for loan supply and the marginal value of collateralised lending as well as the consolidated government budget constraint[17].

The mechanism outlined here abstracts entirely from sterilised open market operations in which the purchases of assets are funded not by the creation of new reserves, but rather by the sale of other assets on the central bank's balance sheet. Given this paper's focus on the quantitative impact of policy, such sterilised interventions are deemed irrelevant as they do not change the quantity of reserves in the system and would instead act through "credit easing" channels as defined by Bernanke (2009):

'In a pure QE regime, the focus of policy is the quantity of bank reserves, which are liabilities of the central bank; the composition of loans and securities on

[15] We abstract from the possibility of banks themselves holding bonds and acting as the central bank's counterparty in an open market operation. Whilst this would be closer to how traditional open market operations have been carried out, it is not consistent with recent large scale asset purchases carried out by central banks which avoided buying assets directly from banks. In the context of our model, the distinction between the two frameworks holds little importance.

[16] Variants of the model in which the central bank can swap reserves for capital, or even capital for bonds can be explored in future work.

[17] As we deal with a consolidated government budget constraint, the net effect of interest payments on bonds held by the central bank is zero.

the asset side of the central bank's balance sheet is incidental.... In contrast, the Federal Reserve's credit easing approach focuses on the mix of loans and securities that it holds and on how this composition of assets affects credit conditions for households and businesses.'

5 Calibration

Table 2 reports the values for the parameters and Table 3 the steady-state values of relevant variables.[18] Following Goodfriend and McCallum (2007) we choose the consumption weight in utility, ϕ, to yield 1/3 of available time in either goods or banking services production. We also set the relative share of capital and labour in goods production η to be 0.36. We choose the elasticity of substitution of differentiated goods, θ, to be equal to 11. The discount factor, β, is set to 0.99 which is close to the canonical quarterly value while the mark-up coefficient in the Phillips curve, κ, is set to 0.1. The depreciation rate, δ, is set to be equal to 0.025 while the trend growth rate, ϱ, is set to 0.005 which corresponds to 2% per year. The steady-state value of bond holding level relative to GDP, b, is set to 0.56 as of the third quarter of 2005. The steady state of private sector bond holdings relative to GDP is set at 0.50, consistent with holdings of U.S. Treasury securities as of end of year 2006.[19]

The deep parameters linked to money and banking are defined as follows. Velocity at its steady state level is set at 0.276 which is close to the ratio between US GDP and M3 at fourth quarter 2005, yielding 0.31. The fractional reserve requirement, rr, is set at 0.1. This is consistent with the reserve ratio set by the Federal Reserve on all liabilities above the low reserve tranche and approximately equal to the average tier one capital ratio in the US since the mid 2000s.

This leaves us three key deep parameters to manipulate which may influence the rest of the steady state variables. Interestingly these are three financial variables and so are of particular interest to our debate on policies. α is the Cobb-Douglas weight of collateral in loan production. This is the degree to which banks base their lending on collateral as opposed to monitoring work or information based lending. The benchmark calibration in Goodfriend and McCallum (2007) of 0.65 is within a range throughout the literature of 0.6 to 0.89, Zhang (2011), so we follow this. k, is the degree to which capital is less efficient as collateral than

[18] The equations for the steady-state equations are listed in the technical appendix, available on request.

[19] The steady state of the transfer level, the Lagrangian of the production constraint and base money depend on the above parameters. The steady state of the marginal cost is $mc = \frac{\theta-1}{\theta}$.

bonds as it entails higher costs to the bank in order to check its physical condition and market price. It is also less liquid should default occur and the collateral be called upon to repay the value of the loan. We set this parameter to 0.2 which not only follows Goodfriend and McCallum (2007), but is validated by data on the Term Security Lending Facility which found less liquid assets were swapped for bonds in a ratio of 0.21. F, can be thought of as total factor productivity in loan production, or a measure of the efficiency with which banks use the factors of production to produce loans[20]. As in Goodfriend and McCallum (2007) we set this to ensure the rest of our steady state values meet three criteria as closely as possible;

- a 1% per year average short-term real "riskless rate" that is the benchmark in the finance literature,
- a 2% average collateralised external finance premium that is in line with the average spread of the prime rate over the federal funds rate in the postwar United States,
- a share of total U.S. employment in depository credit intermediation as of August 2005 of 1.6% as reported by the Bureau of Labor Statistics.

The value this yields is $F = 9.14$.

With these parameter values we see that the steady state of labour input, n, is 0.31 which is close to 1/3 as required. The ratio of time working in the banking service sector, $\frac{m}{m+n}$, is 1.9% under the benchmark calibration, is not far from the 1.6% share required. As the steady-states are computed at zero inflation we can interpret all the rates as real rates. The riskless rate, R^T, is 6% per annum. The interbank rate, R^{IB}, is 0.84% per annum which is close to the 1% per year average short-term real rate. The government bond rate, R^B, is 2.1% per annum. Finally the collateralised external finance premium is 2% per annum which is in line with the average spread of the prime rate over the federal funds rate in the US. The model is solved using the solution methods of King and Watson (1998) who also provide routines to derive the impulse responses of the endogenous variables to different shocks, to obtain asymptotic variance and covariances of the variables and to simulate the data. For the impulse response analysis and simulation exercise we consider the real and financial shocks described in Table

[20]Some authors have also described it as a measure of credit conditions within the economy. The rationale for this seems plausible as when credit conditions are tight, banks will require more collateral and will employ more monitoring work to provide the same amount of loans to the economy.

4, which reports the volatility and persistence parameters. These are standard parameters in the literature.

6 Why the Money Multiplier Matters

Having outlined a model which incorporates both narrow and broad money, we can begin to think more profoundly about how the balance between the two, and thus the money multiplier (also the inverse of the reserve ratio in this paper), may affect banks and the wider economy. Banks, with two sources for meeting deposit demand, make their choice between obtaining narrow money from the central bank or creating their own deposits by issuing loans. As the policy rate falls, narrow money becomes a relatively cheaper way of meeting deposit demand.[21] Similarly, if the return on loans were to fall, or the cost of producing them to rise, then it becomes less profitable for banks to originate deposits in this way. The optimal quantity of narrow money has now increased, whilst the amount of loans has fallen, meaning that the optimising money multiplier is now lower. By providing more narrow money as a cheaper means of meeting deposit demand when loan production becomes more costly the central banks can effectively subsidize deposit creation. If the central bank were to increase the cost of narrow money by raising the policy rate, or the costs of producing loans were to fall (their returns to rise) then the converse would be true and it would be optimal for the money multiplier to rise.

Using our model we are able to trace this story through to the wider economy by analyzing the response of the model to various shocks. Firstly, Figure 4 shows the impulse responses of the model's key variables to a negative shock to the value of collateral households have available to provide in return for loans. It can be thought of as akin to the shock which hit the US housing market in the second half of the last decade. We first analyse the situation, as in Goodfriend and McCallum where the money multiplier is held fixed. When the shock hits, asset prices fall, eroding the value of collateral available for loan production. To

[21] An interesting point of note here is that, although it is not the purpose of our paper, this framework allows for the possibility of paying interest on reserves. Without a zero lower bound imposed on the policy rate charged at the discount window can go negative. In such a setting, charging a negative rate to borrow reserves is equivalent to paying banks to take reserves. This turns the cost implied by $R_t r_t$ in equation (8) into an income stream, paid by the central bank. This is a useful way of thinking about the expansion of reserves under recent quantitative easing programmes and the need for interest on reserves. With rates charged at discount windows cut to zero, and central banks seeking to expand reserve supply either further, interest on reserves (or a negative cost of holding reserves in our framework) was essential to maintain demand for the increased supply.

produce the same quantity of loans, banks now need to employ more monitoring work, as they cannot rely on collateral. Not only does this push up the cost of loan production, but it also draws resources out of the production sector, causing a fall in production and employment in the real economy. The higher cost of loan production leads to a fall in loan supply, whilst the lower production reduces deposit demand both of which lower the level of broad money in the economy. As the central bank seeks to maintain a fixed money multiplier, it responds by withdrawing narrow money from the economy in the same proportion, amplifying the fall in deposits. Consumption/output and inflation then fall due both to the lack of available funds caused by the contraction in broad money and also the fall in production as resources are drawn in to the financial sector. In response the central bank cuts the policy rate, incentivising consumption by households and the take-up of narrow money by commercial banks, eventually stabilizing the economy.

Alternatively, if when the shock hits commercial banks can optimise the mix of narrow money and loans, the response of broad money, and the macroeconomy is much changed. As before we see that asset prices fall. However, as loans become a relatively more expensive way of creating deposits the central bank steps in to increase the supply of relatively cheaper narrow money, which it injects through open market operations. As mentioned, this depresses the money multiplier and acts to support deposit creation, allowing banks to shed costly monitoring effort. This in turn attenuates the rise in the external finance premium and actually supports loan production. With less resources being drawn away from the production sector into monitoring work and broad money being supported both by increased narrow money and a smaller reduction in loan creation, output and inflation fare much better, falling by less and returning to equilibrium more quickly. In this instance the cut in the central bank policy rate acts to make reserves an even cheaper way for banks to meet deposit demand and further increases their demand. When met by the central bank with increased supply this adds to the shock-attenuating effect of varying the money multiplier.

We observe a similar story when considering a real shock. Figure 5 shows the response of the model to a negative shock to productivity. As productivity falls, so does output. With less economic activity, deposit demand from households also falls, which under a fixed money multiplier necessitates an equivalent reduction in narrow money and loans. However, when the money multiplier is flexible and banks are allowed to optimise, they facilitate the fall in broad money by shedding relatively more narrow money than loans, raising the money multiplier.[22] Although this slightly worsens the output response it militates the inflationary impact of the shock even with a less aggressive interest rate response from the monetary authority and brings the EFP back to equilibrium quicker.

7 Welfare Analysis

Having discussed in the previous section why variance in the money multiplier can improve welfare over the cycle we seek to strengthen this result by quantifying its impact on the representative household. To do this we carry out some more stringent welfare analysis by deriving a welfare loss function from a second order approximation to utility.

7.1 Deriving The Welfare Loss Function

The welfare approximation derived from the canonical New Keynesian model finds that welfare of the representative household only depends on the variance of output and inflation (Galí, 2008). We wish to investigate whether this result continues to hold when applied to our richer class of model. The use of the approximation allows us to quantify precisely the welfare rankings arising from each of our policy rules, possibly allowing some normative statements. Thus, we derive a quadratic loss function using a second-order Taylor approximation to utility by using the labour demand function, marginal cost function and sales-production constraint to substitute for household consumption.[23] Once re-ordered and simplified we are left with a loss function with relevant terms in the variances of consumption, inflation, wages, employment in the goods sector and the marginal cost.

[22] This is due to a combination of the now higher return on loans and the increased cost of obtaining narrow money caused by the policy maker increasing the short-term nominal interest rate in response to higher inflation.

[23] The additive nature of our household's utility function allows us to take a Taylor expansion of each term and substitute it back into the original function. The labour demand function is then rearranged for monitoring work, a second order expansion taken and substitution made. This process is then repeated for the marginal cost equation. Following Galí (2008) we substitute the resulting linear term in goods sector employment for a second order term in inflation using the sales equal net production constraint.

$$U_t - U = -\frac{1}{2}E_0\sum_{t=0}^{\infty}\beta^t L_t + O3 \quad (26)$$

$$\text{with } L_t = \frac{1}{2}\begin{bmatrix} \sigma_c^2 + \frac{\theta}{\chi}\left[\frac{\frac{n}{c}-\frac{w}{c}}{1-\eta}\right]\sigma_\pi^2 - \\ \frac{w}{c}\sigma_w^2 - \frac{n}{c}\sigma_n^2 + \frac{mc}{c}\sigma_{mc}^2 \end{bmatrix},$$

where $\chi = \frac{(1-\theta)(1-\beta\theta)}{\theta}\frac{1-\eta}{1+\eta(\theta-1)}$. Given that $\frac{w}{c} > 0$ *and* $\frac{n}{c} > 0$, more flexible wages and employment improves welfare, whilst $\frac{mc}{c} > 0$ *and* $\frac{\theta}{\chi}\left[\frac{\frac{n}{c}-\frac{w}{c}}{1-\eta}\right] > 0$, so more stable marginal cost, consumption and inflation improves welfare.[24]

Remark The welfare of the representative household in this model, as in the original New Keynesian framework, is approximated by standard variables on the supply side rather than those specifically attributable to financial factors. This means that changes in financial conditions only impact on utility in so far as they impact on the variance of consumption, inflation, wages, labour supply hours and marginal costs.

Table 5 shows the asymptotic standard deviations and the contemporaneous cross-correlations with consumption from a simulation of each. We use these values to calculate the welfare loss in each model. We can see that the welfare loss under a regime where the central bank allows the money multiplier to vary over the cycle, supplying narrow money perfectly elastically to meet the demand of commercial banks is significantly smaller than one in which it maintains a fixed money multiplier. It is also notable that the additional welfare loss associated with the zero lower bound is minimal in both regimes.

8 Capturing the Crisis

We now look at the ability of the model to help us understand the response of the economy and the financial sector to the initial shock of the financial crisis and examine what it may tell us about the behaviour of the multiplier in this period. The obvious problem is how to replicate the shock(s) that afflicted the US economy, culminating in the recession of 2008. This particular crisis had a number of characteristics which are compatible with our model. House prices fell, reducing the value of collateral the private sector had to post against loans from

[24]Whilst it is likely to derive a different loss function through a different sequence of substitutions, ours seems both plausible and parsimonious.

the financial sector. Banks also tightened credit conditions and, due to increased precautionary motives and economic uncertainty, increased their preference for liquidity. Whilst the exact sizes and interconnected nature of each of these factors may for practical purposes be intractable, it seems plausible that some combination of shocks along these lines is an appropriate, though rudimentary way to think of the origins of the crisis, and one that it is within the capabilities of our model to capture. We therefore subject the model to simultaneous shocks to liquidity, collateral, monitoring and velocity. We make a simplifying assumption that each element carries equal weighting on impact, but that the persistence of each shock remains as outlined in Table 4. We also constrain the short-term nominal interest (policy) rate as mentioned above as the ZLB has been a key feature of the recent economic milieu.

Table 7 compares the initial changes of variables within the model in light of our composite 'crisis' shock to those observed in US data over the period. Real GDP in the US fell 4.7%, from peak to trough. Employment, measured by non-farm payrolls, went from 139 million in November 2007 to 130 million in August 2009, a fall of 6.3%. CPI inflation fell to -2.1% as prices contracted[25]. The external finance premium, interpreted as in our model as the prime rate spread over the Federal Funds rate, rose by 0.7% following the crisis. The responses of our model are extremely close in both their matching of each variable to its observed counterparty, but also in their impact relative to the other variables within the economy/model.

Whilst, as with any model of this ilk, there remains a degree of oversimplification, the model we have outlined manages to capture well the general evolution of the US economy in the second half of 2008. And what can this tell us about the money multiplier? Figure 6 plots the impulse responses of variables following the crisis shock. What we see is that, following a shock not unlike the one experienced by the US in 2008, the optimal money multiplier falls dramatically. This is consistent with the observed movement in the multiplier shown in Figure 1. The narrative from our model is that as the external finance premium rose and loans became much more expensive to verify and produce, banks maximising behaviour drove them to reduce loan production and shrink the supply of broad money. However, this effect was attenuated by the increased supply of narrow money, demanded by banks and supplied by the Fed through large scale asset purchase programmes and the massive expansion of its balance sheet.

[25] In our model steady state inflation is set to zero.

To provide a counterfactual we impose the same shock to the fixed multiplier model. The implication is that if the Federal Reserve maintained a fixed multiplier, not only would it not have injected increased narrow money in to support broad money, but it would have had to actively withdraw narrow money from the economy to match the contraction in loan supply effected by banks. This would have worsened the broad money contraction and kept financial spreads higher for longer. The model suggests that this could be associated with a contraction in output 1% greater, inflation would have fallen by an additional 2.5% and employment would have been 2% lower. The recovery would also have been more protracted, taking roughly twice as long to return to equilibrium.

9 Conclusion

It seems that rumours of the death of the multiplier as a useful concept may, as the saying goes, be greatly exaggerated. But this paper provides a way of framing the multiplier for the post-crisis world: it is an optimal choice for the commercial banks as to how they collectively meet deposit demand.[26] In this context, variations in the multiplier, even as dramatic as those experienced by many economies in the last 5 years, can be explained, and may even be seen to be approaching an optimum or, at least, welfare enhancing compared to a fixed multiplier alternative.

We have also go some way to addressing the common criticisms levelled at modelling of both the money multiplier and its role in monetary policy. By modelling reserves as demand determined, conditioned on the prevalent interest rates in the economy and allowing the central bank to set the interest rate and then provide this narrow money perfectly elastically through open market operations to meet the demand of commercial banks we provide a framework which is hopefully more recognizable to monetary policy practitioners.

As far as non-conventional monetary policies are concerned, we do not think they necessarily depart from standard open market operations, other than in their size or duration. In our model, we find that the supply of central bank money, or reserves, whilst not preventing an extended downturn may have played a substantive role in preventing the downturn turning into a sustained depression.

[26] See Kashap and Stein (2012) for separate analysis using interest on reserves. That is not necessary for our result.

References

[1] Baltensperger, E., (1980). "Alternative approaches to the theory of the banking firm", *Journal of Monetary Economics,* 6, pp1-37.

[2] Basel III, (2010). "International Framework for liquidity risk measurement, standards and monitoring", http://www.bis.org/publ/bcbs188.htm.

[3] Bernanke, B., (2009). "The Crisis and the Policy Response", Speech given at the Stamp Lecture, London School of Economics.

[4] Berrospide, J., (2012). "Liquidity Hoarding and the Financial Crisis: An Empirical Evaluation", *Federal Reserve Mimeo.*

[5] Carpenter, S. and M. Demiralp, (2012). "Money, reserves, and the transmission of monetary policy: Does the money multiplier exist?", *Journal of Macroeconomics*, 34(1), pp59-75.

[6] Chadha, J.S., and L. Corrado, (2012). "Macro-Prudential Policy on Liquidity: What Does a DSGE Model Tell Us?", *Journal of Economics and Business*, 64 (1), pp37-62.

[7] Chadha, J.S., L. Corrado and J. Meaning, (2012). "Reserves, Liquidity and Money: An Assessment of Balance Sheet Policies", *BIS Papers, No. 66.*

[8] Chadha, J. S., L. Corrado and S. Holly, (2014). "Reconsidering the Role of Money and Credit in Monetary Policy Making: A Note", *Macroeconomic Dynamics,* 18(8), pp1854-1883.

[9] Disyatat, P., (2008). "Monetary policy implementation: Misconceptions and their consequences", BIS Working Paper 269.

[10] Freeman, S. and J. H. Haslag, (1996). "On the Optimality of Interest-Bearing Reserves in Economies of Overlapping Generations", *Economic Theory*, 7(3), pp557-565.

[11] Gale, D., (2011). "Liquidity and Monetary Policy", Chapter 2 in Jagjit S. Chadha and Sean Holly (Ed) *Interest Rates, Prices and Liquidity,* Cambridge University Press.

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

[13] Garfinkel, M. and D. L. Thornton, (1991). “The multiplier approach to the money supply process: a precautionary note,” *Review*, Federal Reserve Bank of St. Louis, pp47-64.

[14] Goodfriend, M. and B. T. McCallum, (2007). “Banking and Interest Rates in Monetary Policy Analysis: A Quantitative Exploration”, *Journal of Monetary Economics*, 54, pp1480-1507.

[15] Kashyap, A. K. and J. C. Stein, (2012). “The Optimal Conduct of Monetary Policy with Interest on Reserves”, *American Economic Journal: Macroeconomics,* 4(1), pp266-282.

[16] King, R. G. and M. W. Watson, (1998). “The Solution of Singular Linear Difference Systems under Rational Expectations”, *International Economic Review*, 39(4), Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance, pp1015-1026.

[17] Poole, W., (1970). "Optimal Choice of Monetary Policy Instruments in a Simple Stochastic Macro Model," *Quarterly Journal of Economics*, 84(2), pp197-216.

[18] Walsh, C. E., (2009). “Using Monetary Policy to stabilise Economic Activity”, Federal Reserve Bank of Kansas City *Financial Stability and Macroeconomic Policy*, 2009 Jackson Hole Symposium, 2010: 245-296.

[19] Zhang, H., (2011). “Financial Sector Shocks, External Finance Premium and Business Cycles”, *Cardiff Economics Working Papers.*

Table 1: List of Variables

Variable	Description
c	Real Consumption
n	Labour Input
m	Labour Input for Loan Monitoring, or 'Banking Employment'
w	Real wage
q	Price of Capital Goods
P	Price Level
π	Inflation
mc	Marginal Cost
r	Reserves
rr	Reserves/Deposit Ratio
D	Deposits
L	Loans
P^A	Aggregate Prices
b	Real Bond Holding
b^p	Real Private Sector Bond Holdings
Ω	Marginal Value of Collateral
EFP	Uncollateralised External Finance Premium (R^T - R^{IB})
LSY^B	Liquidity Service on Bonds
LSY^{KB}	Liquidity Service on Capital ($kLSY^B$)
R^T	Benchmark Risk Free Rate
R^B	Interest Rate for Bond
R^{IB}	Interbank Rate
R^L	Loan Rate
R^D	Deposit Rate
λ	Lagrangian for Budget Constraint (shadow value of consumption)
ξ	Lagrangian for Production Constraint
T	Real transfer (%)

Table 2: Parameterisation

Parameter	Description	Value
β	Discount factor	0.99
κ	Coefficient in Phillips curve	0.1
α	Collateral share of loan production	0.65
ϕ	Consumption weight in utility	0.4
η	Capital share of firm production	0.36
δ	Depreciation rate of capital	0.025
ϱ	Trend growth rate of shocks	0.015
ρ	Interest rate smoothing	0.8
ϕ_π	Coefficient on Inflation in Policy	1.5
ϕ_y	Coefficient on Output in Policy	0.5
F	Production coefficient of loan	9.14
k	Inferiority coefficient of capital as collateral	0.2
θ	Elasticity of substitution of differentiated goods	11

Table 3: Steady State Parameters

Steady State	Description	Value
m	Banking Employment	0.0063
n	Labour Input	0.3195
R^T	Risk Free Rate	0.015
R^{IB}	Interbank Rate	0.0021
R^L	Loan Rate	0.0066
R^B	Bond Rate	0.0052
b/c	Bond to Consumption Ratio	0.56
b^p/c	Private Sector Bond Holdings to Consumption Ratio	0.50
γ (b^p/b)	Fraction of Bonds Held By Private Sector	0.893
c	Consumption	0.8409
T/c	Transfers Over Consumption	0.126
w	Real Wage	1.9494
λ	Shadow Value of Consumption	0.457
ν	Velocity	0.31
Ω	Marginal Value of Collateral	0.237
K	Capital	9.19
K^P	Private Sector Capital Holdings	9.19
rr	Reserve ratio	0.1
r/c	Reserves to Consumption	0.36

Table 4: Properties of Exogenous Shocks

Shock Name	Standard Deviation	Persistence
Productivity	0.35%	0.95
Monitoring	1.00%	0.95
Collateral	0.35%	0.9
Monetary Policy	0.82%	0.3
Mark Up	0.11%	0.74
Bond Holdings	1.00%	0.9
Velocity	1.00%	0.33
Liquidity	1.00%	0.33

Table 5: Impact on the Economy of Flexible Money Multiplier

Policy	Flexible MM[27]		Flexible MM CIR[28]		Fixed MM[29]		Fixed MM CIR[30]	
	St.Dv	Corr	St.Dv	Corr	St.Dv	Corr	St.Dv	Corr
Real Consumption/Output	1.05	1	1.17	1	0.96	1	1.21	1
Inflation	0.40	0.62	0.42	0.69	0.68	0.75	0.94	0.84
Employment in Monitoring	3.55	-0.74	3.28	-0.83	2.09	-0.58	1.82	-0.61
Employment in Goods Sector	1.60	0.96	1.81	0.97	1.55	0.95	1.98	0.97
Real Wage	1.69	0.99	1.91	0.99	1.65	0.99	2.11	0.99
Private Sector Bond Holdings	1.79	-0.34	1.85	-0.40	1.65	-0.06	1.92	-0.20
Asset Prices	1.02	0.98	1.14	0.99	1.09	0.99	1.42	0.99
Loans	1.10	0.29	1.15	0.40	2.06	0.32	2.97	0.43
Reserves	1.62	0.73	1.83	0.77	2.06	0.32	2.97	0.43
Policy Rate	1.05	-0.08	0.80	-0.36	1.14	-0.07	0.80	-0.24
Loan Rate	0.68	-0.83	0.70	-0.88	0.47	-0.19	0.63	0.12
Bond Rate	3.99	0.49	3.24	0.50	3.13	0.33	2.62	0.38
Deposit Rate	1.04	-0.16	0.85	-0.45	1.14	-0.07	0.80	-0.24
External Finance Premium	1.37	-0.35	0.82	-0.46	1.19	-	0.82	0.33
Liquidity Premium	4.52	-0.56	3.68	-0.62	3.27	-0.35	2.61	-0.34

[27] Model with a flexible money multiplier determined by demand from profit-maximising commercial banks with an unconstrained interest rate policy.

[28] Model with a flexible money multiplier determined by demand from profit-maximising commercial banks with interest rate policy constrained, holding the policy rate constant.

[29] Model with a fixed money multiplier and unconstrained interest rate policy

[30] Model with a fixed money multiplier and with interest rate policy constrained, holding the policy rate constant.

Table 6: Relative Welfare Analysis

Who Determines Reserves	Welfare Loss[31]
Fixed Money Multiplier CIR	10.921
Fixed Money Multiplier	7.829
Flexible Money Multiplier CIR	4.293
Flexible Money Multiplier	3.908

[31] Loss determined by a quadratic loss function derived using a second-order Taylor approximation to utility. See Section 7.1

Table 7: Response of Model to Crisis Shock Vs Stylised Facts

Variable	Response From Model	Observed in Data
Real Output	-4.9%	-4.7%
Inflation	-2.5%	-2.1%
Employment	-7.5%	-6.3%
Liquid Assets/Total	-5%	-5%
External Finance Premium	0.9%	0.7%

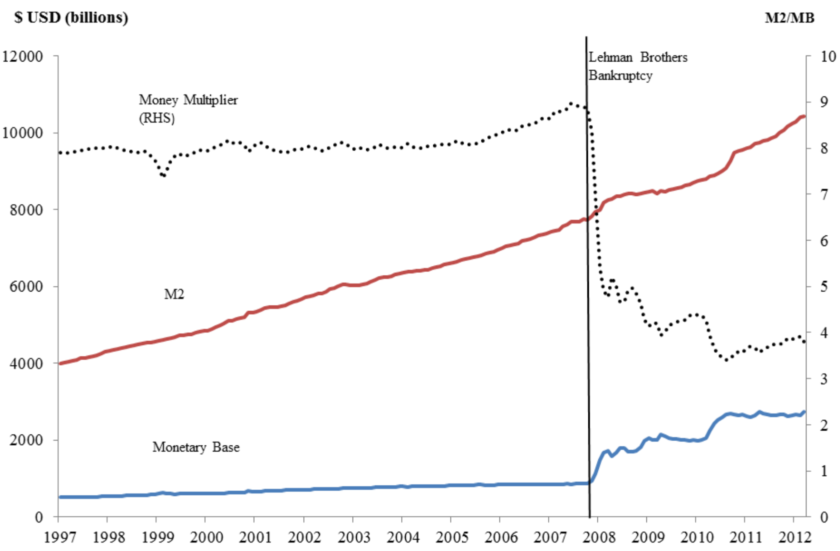

Figure 1: US Money Aggregates and Money Multiplier

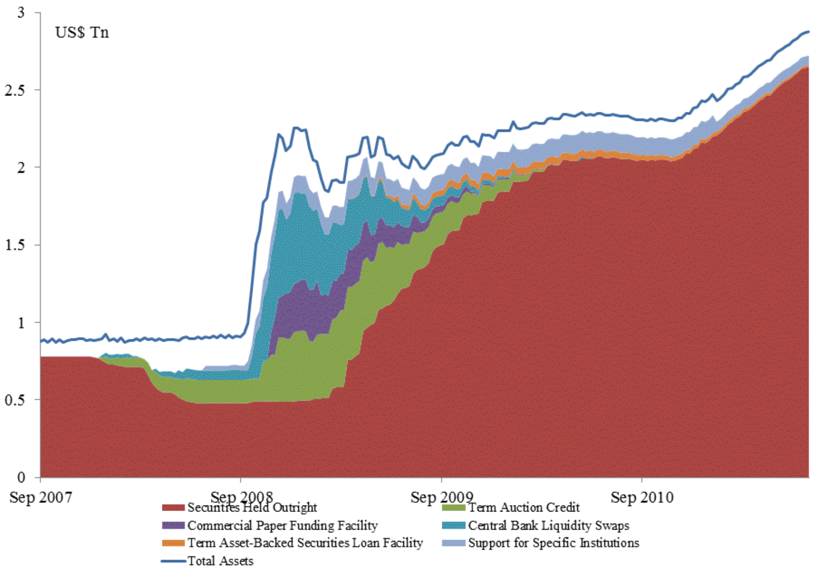

Figure 2: Federal Reserve Assets[32]

[32]Total may differ from constituent parts due to rounding.

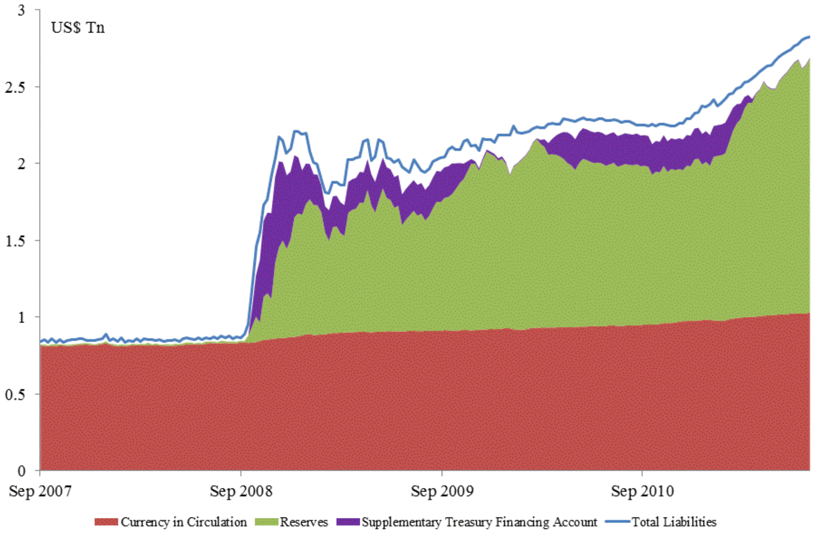

Figure 3: Federal Reserve Liabilities

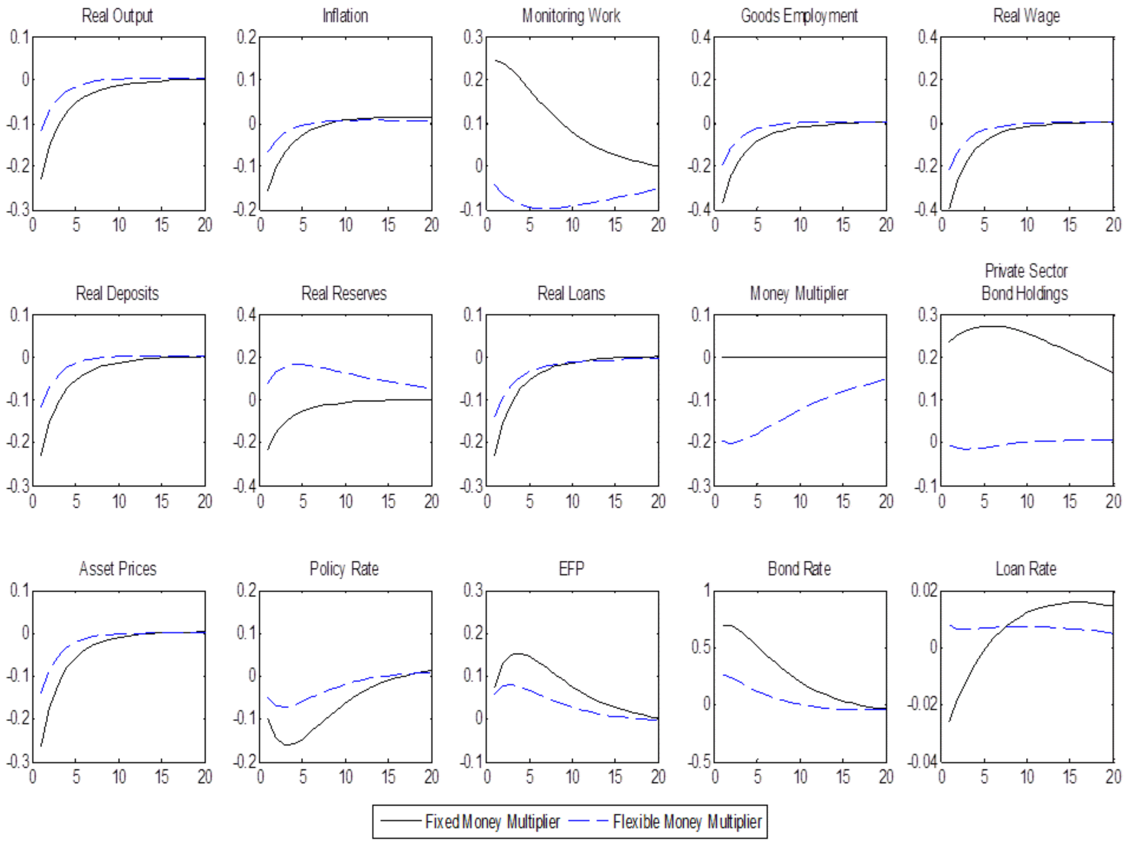

Figure 4: Negative Shock to Collateral

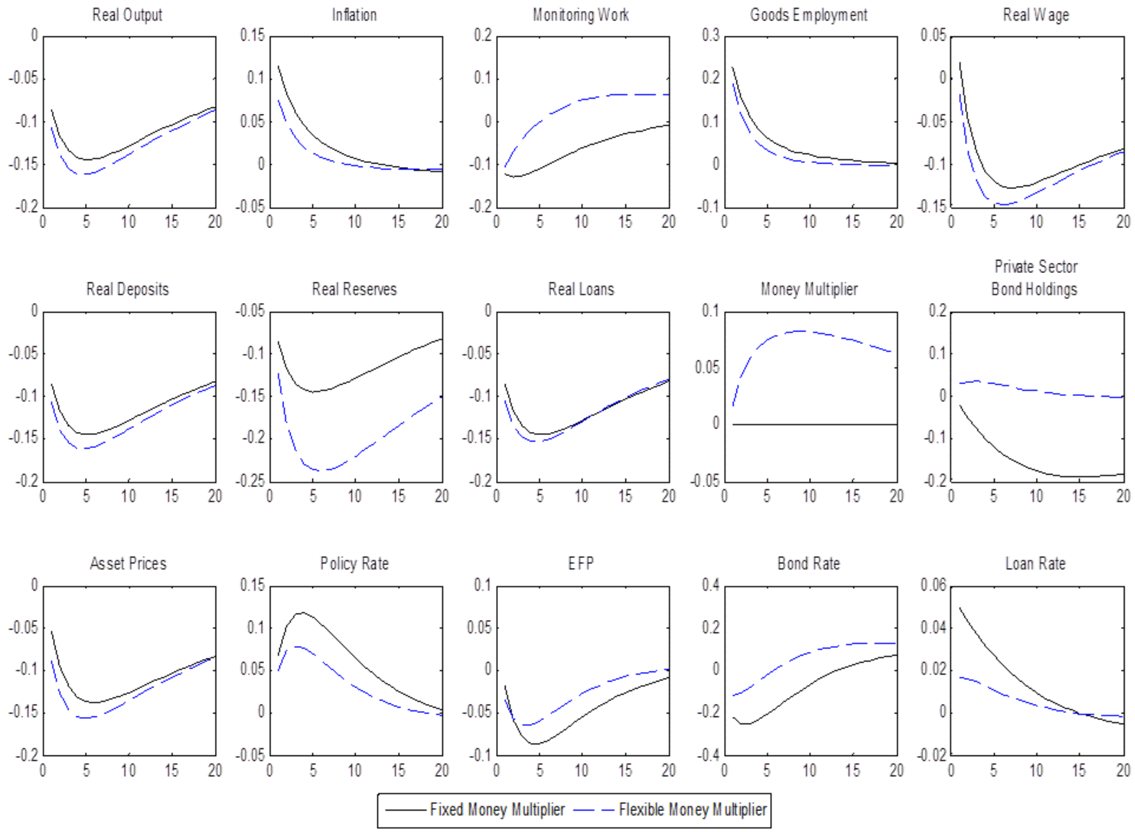

Figure 5: Negative Shock to Productivity

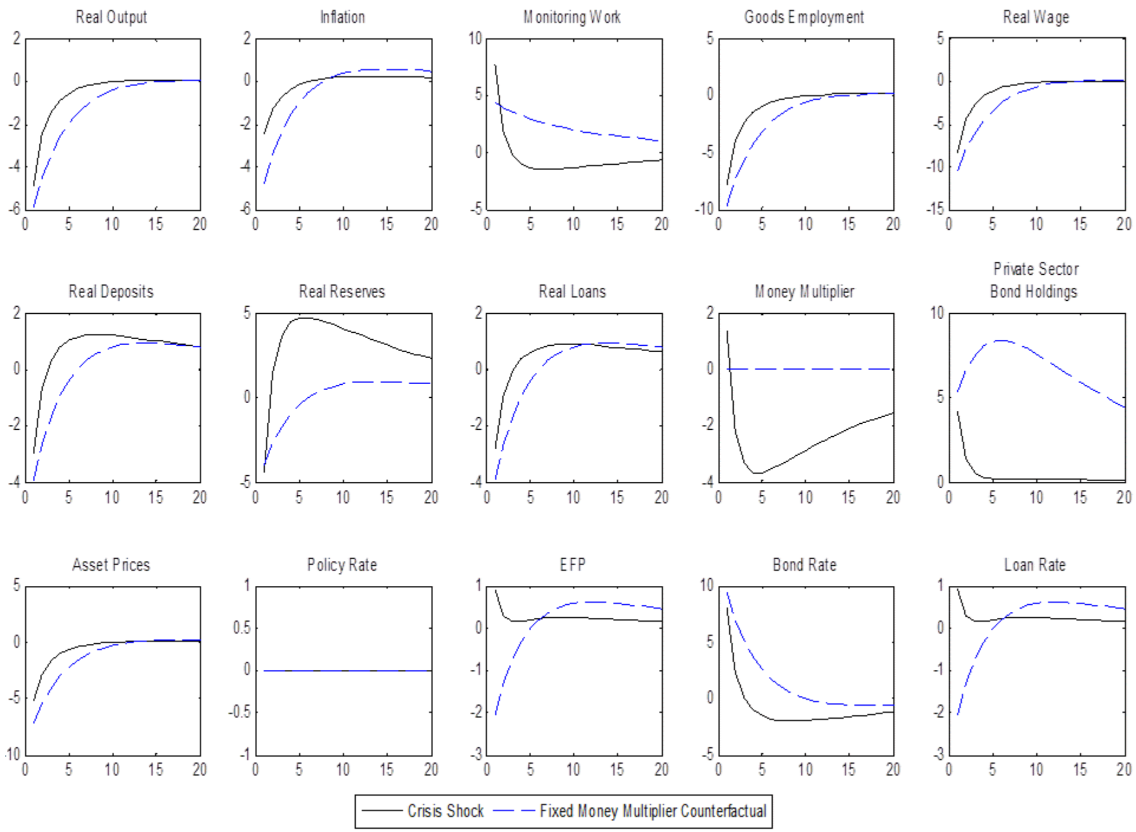

Figure 6: Crisis Shock and Fixed Money Multiplier Counterfactual

Technical Appendix

1 Model Set Up

This is a modified version of the Goodfriend and McCallum's model (2007) incorporating a government (including bank) budget constraint and a cash-in-advance constraint with stochastic velocity of money demand from Chadha and Corrado (2012) to which we add a preference for reserves holdings by bank and a set policies that we characterise as reserve for asset swaps.

- Utility function:

$$U = E_0 \sum_{t=0}^{\infty} \beta^t [\phi \log(c_t) + (1-\phi) \log(1 - n_t^s - m_t^s)] \tag{1}$$

c_t denotes real consumption, n_t^s is supply of labour in goods sector and m_t^s is the supply of monitoring work in the banking sector.

- Budget constraint:

$$\begin{aligned} & q_t(1-\delta)K_t + \frac{\gamma B_t}{P_t^A} + \frac{D_{t-1}}{P_t^A} + w_t(n_t^s + m_t^s) + c_t^A(\frac{P_t}{P_t^A})^{1-\theta} + \Pi_t \\ & -q_t K_{t+1} - \frac{\gamma B_{t+1}}{P_t^A(1+R_t^B)} - \frac{D_t}{P_t^A} - w_t(n_t + m_t) - c_t - tax_t = 0 \end{aligned} \tag{2}$$

q_t is the price of capital, K_t is the quantity of capital, P_t is the price of household's produced good, P_t^A is the consumption good price index, n_t is the labour demanded by household as producer, m_t, is the labour demanded by household's banking operation, w_t is the real wage, D_t is the nominal holding of broad money, tax_t is the real lump-sum tax payment, R_t^B is the nominal interest rate on government bonds purchased in $t+1$, B_{t+1}. We also assume that any profit from the banking sector, Π_t, goes to the households' sector. The Lagrangian multiplier of this constraint is denoted as, λ_t.

- Sales equal net production constraint:

$$K_t^\eta (A1_t n_t)^{1-\eta} - c_t^A (P_t/P_t^A)^{-\theta} = 0 \tag{3}$$

$A1_t$ is a productivity shock in the goods production sector whose mean increases over time at a rate ϱ. In (18) and (19) the superscript A indicates that the variable is an aggregate taken as given from each household. The Lagrangian multiplier of this constraint is denoted as, ξ_t.

- Government (including bank budget constraint):

$$g_t - tax_t = \frac{r_t}{P_t^A(1+R_t)} - \frac{r_{t-1}}{P_t^A} + \frac{\gamma B_{t+1}}{P_t^A(1+R_{t+1}^B)} - \frac{\gamma B_t}{P_t^A} \tag{4}$$

where g_t is real government expenditure. We define:

$$b_{t+1} = B_{t+1}/P_t^A(1+R_{t+1}^B) \tag{5}$$

$$re_t = r_t/P_t^A(1+R_t)$$

So the budget constraint can be rewritten as:

$$T_t = re_t - re_{t-1}\frac{P_{t-1}^A}{P_t^A}(1+R_{t-1}^{IB}) + b_t - b_{t-1}\frac{P_{t-1}^A}{P_t^A}(1+R_t^B) \tag{6}$$

where $T_t = g_t - tax_t$.

- Deposit/money constraint:

$$c_t = v_t D_t/P_t^A \tag{7}$$

v_t denotes velocity and D_t are deposits

- Loans:

$$L_t = D_t(1 - rr_t) \tag{8}$$

where $rr_t = \frac{r_t}{D_t}$ is the reserve/deposit ratio and r_t is high-powered money.

- The bank's problem is to maximize profits within period subject to the returns from loans, L_t, which are lent out at the collateralized interest rate of R^L, reserves obtained from the central bank, r_t, at cost interest rate, R, and the payment of deposit interest, R^D, to deposits:

$$\max_{r_t} \Pi_t = R_t^L L_t - R_t r_t - R_t^D D_t, -\frac{1}{2} R_t^T (\bar{r} - r_t)^2 - \tau_t(\bar{r} - r_t) \tag{9}$$

The hard constraint that deposits must equal the sum of loans and reserves

$$D_t = L_t + r_t \tag{10}$$

means that we can substitute out loans from equation (9), leaving it in terms of reserves and the overall level of deposits, the latter of which is determined by the household.

Production function pertaining to management of loans:

$$L_t/P_t^A = F(\gamma b_{t+1} + A3_t k q_t K_{t+1})^\alpha (A2_t m_t)^{1-\alpha} \qquad 0 < \alpha < 1 \tag{11}$$

From (7):

$$c_t = v_t \frac{F(\gamma b_{t+1} + A3_t k q_t K_{t+1})^\alpha (A2_t m_t)^{1-\alpha}}{P_t^A (1 - rr_t)} \tag{12}$$

$A2_t$ denotes a shock to monitoring work, $A3_t$ is a shock on capital as collateral. The parameter k denotes the inferiority of capital as collateral in the banking production function, while α is the share of collateral in the loan production function. For a complete list of all variables and parameters in the model see Tables 1 and 2 in the main text.

1.1 First Order Conditions

- Derivative with respect to m_t^s and n_t^s of (1), (2) and (3)

$$-\frac{(1-\phi)}{1 - n_t^s - m_t^s} + w_t \lambda_t = 0 \tag{13}$$

- Derivative with respect to m_t.

$$\begin{aligned} \frac{\phi}{c_t} \frac{\partial c_t}{\partial m_t} - \lambda_t w_t - \lambda_t \frac{\partial c_t}{\partial m_t} &= 0 \\ w_t &= \left(\frac{\phi}{\lambda_t c_t} - 1\right) \frac{\partial c_t}{\partial m_t} \end{aligned} \tag{14}$$

given that

$$c_t = \frac{v_t D_t}{P_t^A} = \frac{v_t L_t}{P_t^A (1 - rr_t)} = \frac{v_t F(\gamma b_{t+1} + A3_t k q_t K_{t+1})^\alpha (A2_t m_t)^{1-\alpha}}{(1 - rr_t)} \tag{15}$$

Then

$$\frac{\partial c_t}{\partial m_t} = \frac{1-\alpha}{m_t} c_t$$

So (14) becomes:

$$w_t = (\frac{\phi}{\lambda_t c_t} - 1)\frac{1-\alpha}{m_t} c_t \tag{16}$$

- Derivative with respect to n_t

$$\begin{aligned} \lambda_t w_t &= \xi_t A1_t(1-\eta)(\frac{K_t}{n_t A1_t})^{\eta} \qquad (17) \\ w_t &= \frac{\xi_t}{\lambda_t} A1_t(1-\eta)(\frac{K_t}{n_t A1_t})^{\eta} \end{aligned}$$

- Derivative with respect to K_{t+1}

$$\frac{\phi}{c_t}\frac{\partial c_t}{\partial K_{t+1}} + E_t\lambda_{t+1}q_{t+1}(1-\delta)\beta - q_t\lambda_t - \lambda_t\frac{\partial c_t}{\partial K_{t+1}} + E_t\xi_{t+1}\beta\eta K_t^{\eta-1}(A1_t n_t)^{1-\eta} \tag{18}$$

Given

$$\begin{aligned} \frac{\partial c_t}{\partial K_{t+1}} &= \frac{c_t\alpha A3_t k q_t}{\gamma b_{t+1} + A3_t k q_t K_{t+1}} \qquad (19) \\ &= \Omega_t A3_t k q_t \end{aligned}$$

with

$$\Omega_t = \frac{c_t\alpha}{\gamma b_{t+1} + A3_t k q_t K_{t+1}} \tag{20}$$

So

$$(\frac{\phi}{c_t\lambda_t} - 1)\Omega A3_t k q_t + E_t\frac{\lambda_{t+1}}{\lambda_t} q_{t+1}(1-\delta)\beta - q_t + E_t\beta\eta\left[\frac{\lambda_{t+1}}{\lambda_t}\frac{\xi_{t+1}}{\lambda_{t+1}}(\frac{A1_t n_t}{K_t})^{1-\eta}\right] \tag{21}$$

- Derivative with respect to P_t:

$$\begin{aligned} 0 &= \lambda_t(1-\theta)c_t^A(P_t)^{-\theta}(P_t^A)^{-(1-\theta)} + \theta\xi_t c_t^A(P_t)^{-\theta-1}(P_t^A)^{\theta} \qquad (22) \\ \frac{\xi_t}{\lambda_t} &= \frac{\theta-1}{\theta}\frac{P_t}{P_t^A} \end{aligned}$$

- Derivative with respect to B_{t+1} :

$$\frac{\phi}{c_t}\frac{\partial c_t}{\partial B_{t+1}} - \lambda_t \frac{\partial c_t}{\partial B_{t+1}} + E_t \beta\gamma \frac{\lambda_{t+1}}{P^A_{t+1}} - \frac{\gamma\lambda_t}{P^A_t(1+R^B_t)} = 0$$

where

$$\frac{\partial c_t}{\partial B_{t+1}} = \frac{\gamma\Omega_t}{P^A_t(1+R^B_t)}$$

so

$$= \left[\frac{\phi}{\lambda_t c_t} - 1\right] \frac{\gamma\Omega_t}{P^A_t(1+R^B_t)} + \gamma\beta E_t \frac{\lambda_{t+1}}{P^A_{t+1}} - \frac{\gamma\lambda_t}{P^A_t(1+R^B_t)} \tag{23}$$

$$= \left[\frac{\phi}{\lambda_t c_t} - 1\right] \Omega_t - 1 + \beta E_t \left[\frac{\lambda_{t+1}}{\lambda_t}\frac{P^A_t}{P^A_{t+1}}(1+R^B_t)\right]$$

- Derivative wrt r_t of (9):

$$\frac{\partial \Pi_t}{\partial r_t} = -R^L_t - R_t - R^T_t(\bar{r} - r_t) + \tau_t = 0. \tag{24}$$

$$r_t = \bar{r} - \left[\frac{R_t + R^L_t}{R^T_t}\right] + \frac{\tau}{R^T_t} \tag{25}$$

1.2 Interest Rates

FOC with respect to c_t gives

$$\left(\frac{U_{c,t}}{\lambda_t} - 1\right) = 0 \tag{26}$$

where $U_c = \frac{\phi}{c_t}$. Substituting in (23) gives riskless rate R^T_t:

$$1 + R^T_t = E_t \frac{\lambda_t P^A_{t+1}}{\lambda_{t+1} P_{t+1}} \tag{27}$$

The interest rate on bonds, R^B_t, is derived from (23):

$$R^T_t - R^B_t = \left[\frac{U_{c,t}}{\lambda_t} - 1\right] \Omega_t = \left[\frac{\phi}{c_t \lambda_t} - 1\right] \Omega_t \tag{28}$$

So $\frac{U_c}{\lambda}$ measures the household marginal utility relative to households shadow value of funds while Ω_t is the marginal value of collateral.

While

$$R^L_t - R^B_t = \left[\frac{U_{c,t}}{\lambda_t} - 1\right] k\Omega_t \tag{29}$$

where k determines the degree to which capital is collateralizable.

The central bank's policy rate is a standard taylor rule, of the the form

$$R_t = R_{t-1}^{\rho}(1-\rho)(\pi_t^{\phi_\pi} * y_t^{\phi_y}) \tag{30}$$

To find rate relates to the risk-free benchmark rate we must equate marginal product of loans per unit of labour $(1-\alpha)\frac{L_t}{m_t}$ to their marginal cost $\frac{w_t}{P_t^A}$ where loans are defined as $L_t = D_t(1-rr_t) = \frac{c_t P_t^A}{v_t}(1-rr_t)$. So the difference between rates is equal to real marginal cost of loan management:

$$R_t^T - R_t = \left[\frac{v_t m_t w_t}{(1-\alpha)(1-rr_t)c_t}\right] \tag{31}$$

Since $(1-\alpha)$ is the factor share of monitoring, for collateralised lending the marginal cost of loan production is multiplied by $(1-\alpha)$ and the relevant relationship becomes:

$$R_t^L - R_t = \left[\frac{v_t m_t w_t}{(1-rr_t)c_t}\right] \tag{32}$$

The interest rate on deposits is simply:

$$R_t^D = R_t(1-rr_t). \tag{33}$$

2 Steady-State

For the productivity and monitoring shocks we assume a trend growth rate equal to $A2_t = A1_t = (1+ \varrho)^t$. In steady state $q = 1$, $A2 = A1 = (1+ \varrho)$, λ shrinks at rate ϱ so $\frac{\lambda_{t+1}}{\lambda_t} = \frac{1}{(1+\varrho)}$ and there is no inflation so $P = P^A = 1$ while K is constant.

From (12):

$$1 = \frac{vF}{1-rr}(\frac{b}{c} + \frac{kqK}{c})^{\alpha}(\frac{m}{c})^{1-\alpha} \tag{34}$$

From (20):

$$\Omega = \frac{\alpha}{(\frac{b}{c} + \frac{kqK}{c})} \tag{35}$$

From (13)

$$\frac{1-\phi}{1-n-m} = w\lambda \tag{36}$$

From (16):

$$w = (\frac{\phi}{c\lambda} - 1)\frac{(1-\alpha)\,c}{m} \tag{37}$$

From (22) $\frac{\xi}{\lambda} = \frac{\theta-1}{\theta}$. Replacing in (17):

$$w = \frac{\theta-1}{\theta}(1-\eta)\left(\frac{K}{n}\right)^{\eta} \tag{38}$$

From (21):

$$(\frac{\phi}{c\lambda} - 1)\Omega kq + \frac{1}{1+\varrho}q(1-\delta)\beta - q + E_t\beta\eta\left[\frac{1}{1+\varrho}\frac{\xi}{\lambda}(\frac{n}{K})^{1-\eta}\right] = \tag{39}$$

$$(\frac{\phi}{c\lambda} - 1)\Omega kq - 1 + \frac{\beta}{1+\varrho}\left[(1-\delta) + \eta\frac{\theta-1}{\theta}(\frac{n}{K})^{1-\eta}\right]$$

From the overall resource constraint that incorporates (2), (3) and (4):

$$1 = \left(\frac{K}{c}\right)^{\eta}\left(\frac{n}{c}\right)^{1-\eta} - \frac{\delta K}{c} \tag{40}$$

Equations from (36) to (34) give the steady-state value for m, n, c, K, λ, w, Ω.

The steady-state value for deposits is:

$$D = \frac{c}{v} \tag{41}$$

The steady-state value of reserves is:

$$r = rrD = rr\frac{c}{v} \tag{42}$$

And the steady-state value for re :

$$re = \frac{r}{(1+R)} = \frac{rr\ c}{v\ (1+R)}$$

From the reserve equation setting $r = \bar{r}$ we derive the steady-state value for τ :

$$\tau = R + R^L \tag{43}$$

Finally the uncollateralized and collateralized external finance premium in steady-state are defined as:

$$UEFP = \frac{vmw}{(1-\alpha)(1-rr)c} \tag{44}$$

$$EFP = \frac{vmw}{(1-rr)c} \tag{45}$$

From (28), (44) and (32) we derive the steady-state values for the central bank's policy rate, the loan rate and the bond rate:

$$R_t = \rho R_{t-1} + (1-\rho)(\phi_\pi \pi_t + \phi_y y_t) \tag{46}$$

$$R^T = R + EFP \tag{47}$$

$$R^L = R + CEFP \tag{48}$$

$$R^B = R^T - LSY^B \tag{49}$$

From (6) we derive the steady-state value for transfers:

$$T = r\left(1 - \frac{(1+R)}{(1+\pi)}\right) + b\left(1 - \frac{(1+R^B)}{(1+\pi)}\right) \tag{50}$$

3 The Linearized Model

The model is composed by the following linearized equations.[1]

Supply Labour (from (13)):

$$\frac{n}{(1-n-m)}\widehat{n}_t + \frac{m}{(1-n-m)}\widehat{m}_t - \widehat{\lambda}_t - \widehat{w}_t = 0 \tag{A1}$$

Demand for Labour (from (16)):

$$\widehat{m}_t + \widehat{w}_t + \frac{(1-\alpha)c}{mw}\left(\widehat{c}_t + \frac{\phi}{\lambda}\widehat{\lambda}_t\right) = 0 \tag{A2}$$

Supply of Banking Services (combining (8) and (11))[2]:

$$\begin{aligned} \widehat{c}_t &= \widehat{v}_t c + \widehat{rr}_t c + (1-\alpha)(a2_t + \widehat{m}_t) + \\ &\alpha\left[\frac{bc}{bc+(1+\varrho)kK}(\widehat{c}_t + \widehat{b}_t) + \frac{kK(1+\varrho)}{bc+(1+\varrho)kK}(a3_t + \widehat{q}_t)\right] \end{aligned} \tag{A3}$$

reported in the main text as:

$$c_t = \left\{v_t c + rr_t c + (1-\alpha)(m_t + a2_t) + \alpha\left[\frac{b}{b+k_1}b_t + \frac{k_1}{b+k_1}(q_t + a3_t)\right]\right\}\left(\frac{b+k_1}{b(1-\alpha)+k_1}\right)$$

CIA constraint (from (7)):

$$\widehat{c}_t + \widehat{P}_t = \widehat{D}_t + \widehat{v}_t - \widehat{rr}_t \tag{A4}$$

Aggregate Supply:

$$\widehat{c}_t = (1-\eta)(1+\frac{\delta K}{c})(a1_t + \widehat{n}_t) - \frac{\delta K}{c}\hat{q}_t \tag{A5}$$

Marginal cost:

$$\widehat{mc}_t = \widehat{n}_t + \widehat{w}_t - \widehat{c}_t \tag{A6}$$

Mark-up (from (22)):

$$\widehat{mc}_t = \widehat{\xi}_t - \widehat{\lambda}_t \tag{A7}$$

[1]The model is defined in the Matlab file gmvsys.m. Standard deviation and persistence structure of the stochastic variables are defined in the driver file gmvdrv.m.

[2]The relationship is derived by setting $b = \frac{B}{P(1+R^B)c}$ and $b_{t+1} = b_t c_t$ where b_{t+1} is defined in (5).

Inflation:

$$\widehat{\pi}_t = \widehat{p}_t - \widehat{p}_{t-1} \tag{A8}$$

Calvo pricing:

$$\widehat{\pi}_t = \kappa \widehat{mc}_t + \beta E_t \widehat{\pi}_{t+1} + a5_t \tag{A9}$$

Marginal Value of Collateralized Lending (from (20)):

$$\widehat{\Omega}_t = \frac{kK}{bc + kK}\left(\widehat{c}_t - \widehat{q}_t - a3_t\right) - \frac{bc}{bc + kK}\widehat{b}_t \tag{A10}$$

reported in the main text as:

$$\widehat{\Omega}_t = \frac{k_2}{b + k_2}\left(\widehat{c}_t - \widehat{q}_t - a3_t\right) - \frac{b}{b + k_2}\widehat{b}_t$$

Asset Pricing (from (23))[3]:

$$\begin{aligned}\widehat{q}_t\left[1 - k\Omega(\frac{\phi}{c\lambda} - 1)\right] &= [\frac{\beta(1-\delta)}{1+\varrho} + \frac{\beta\eta mc}{1+\varrho}(\frac{n}{K})^{1-\eta}]\left(E_t\widehat{\lambda}_{t+1} - \widehat{\lambda}_t\right) \\ &+ \frac{\beta(1-\delta)}{1+\varrho}E_t\widehat{q}_{t+1} + \frac{k\Omega\phi}{c\lambda}\left(-\widehat{c}_t - \widehat{\lambda}_t\right) + k\Omega(\frac{\phi}{c\lambda} - 1)\left(\widehat{\Omega}_t + a3_t\right) \\ &+ \left(\frac{\beta\eta mc}{1+\varrho}(\frac{n}{K})^{1-\eta}\right)E_t\left[\widehat{mc}_{t+1} + (1-\eta)\left(\widehat{n}_{t+1} + a1_{t+1}\right)\right]\end{aligned} \tag{A11}$$

ported in the main text as:

$$\begin{aligned}\widehat{q}_t &= (\delta_1 + \gamma_1)\left(E_t\widehat{\lambda}_{t+1} - \widehat{\lambda}_t\right) + \delta_1 E_t\widehat{q}_{t+1} - \frac{k\Omega\phi}{c\lambda}\left(\widehat{c}_t + \widehat{\lambda}_t\right) + \\ & k\Omega(\frac{\phi}{c\lambda} - 1)\left(\widehat{\Omega}_t + a3_t\right) + \gamma_1 E_t\left[\widehat{mc}_{t+1} + (1-\eta)\left(\widehat{n}_{t+1} + a1_{t+1}\right)\right]\end{aligned}$$

Government Budget Constraint[4]:

$$\widehat{TT}_t = \frac{rr\ c}{v(1+R^{IB})}\left(\widehat{re}_t + (1+R)(\widehat{\pi}_t - \widehat{re}_{t-1} - \widehat{R}_{t-1})\right) + b\left(\widehat{b}_t + (1+R^B)(\widehat{\pi}_t - \widehat{b}_{t-1} - \widehat{R}^B_t)\right) \tag{A12}$$

Bond Holding:

$$\widehat{b}_t = a6_t \tag{A13}$$

[3]Note that in steady-state $\frac{\xi}{\lambda} = mc$ and $\frac{\lambda_{t+1}}{\lambda_t} = \frac{1}{1+\gamma}$.

[4]We define the percentage deviation from steady state of flow and stock variables by $\ln x_t - \ln x$, while for interest rates and ratio variables they are $R_t = R + \widehat{R}_t$ (rates) and $r_t = r + \widehat{r}_t$ (ratio, assuming $r_t = x_t/y_t$), respectively. It can be shown the approximation comes from first-order Taylor expansion: $e^x \approx 1 + x$, while for rate variable: $\widehat{R}_t \approx \ln(1+R_t) - \ln(1+R)$ and for ratio: $\widehat{r}_t = r_t - r = \ln(x_t/y_t) - \ln(x/y) = \widehat{x}_t - \widehat{y}_t$.

Riskless Interest Rate (from (27)):

$$\widehat{R}_t^T = \widehat{\lambda}_t + E_t\widehat{\pi}_{t+1} - E_t\widehat{\lambda}_{t+1} \tag{A14}$$

Liquidity Service of Bonds (from (23))[5]:

$$\frac{1+R^B}{1+R^T}\left(\widehat{R}_t^B - \widehat{R}_t^T\right) = \frac{\phi\Omega}{c\lambda}\left(\widehat{c}_t + \widehat{\lambda}_t\right) - \left(\frac{\phi}{c\lambda} - 1\right)\Omega\widehat{\Omega}_t \tag{A15}$$

External Finance Premium (from (31)):

$$\widehat{EFP}_t = \widehat{v}_t + \widehat{w}_t + \widehat{m}_t - \widehat{c}_t + \widehat{rr}_t \tag{A16}$$

Other Interest Rates:

$$\widehat{R}_t = \widehat{R}_t^T - \widehat{EFP}_t \tag{A17}$$

$$\widehat{R}_t^L = \widehat{R}_t + \widehat{EFP}_t \tag{A18}$$

$$\widehat{R}_t^D = \widehat{R}_t - \widehat{rr}_t\frac{rr}{(1-rr)} \tag{A19}$$

Policy Feedback Rule:

$$\widehat{R}_t = (1-\rho)\left(\phi_\pi\widehat{\pi}_t + \phi_y\widehat{mc}_t\right) + \rho\widehat{R}_{t-1} + a4_t \tag{A20}$$

Velocity:

$$\widehat{v}_t = a7_t \tag{A21}$$

Reserves:

$$\widehat{r}_t = \frac{1}{r\ R^T}\left[\left(\tau - R - R^L\right)\widehat{R}_t^T - R\widehat{R}_t - R^L\widehat{R}_t^L + \tau\widehat{\tau}_t\right] \tag{A22}$$

Liquidity:

$$\widehat{\tau}_t = a8_t \tag{A23}$$

Loans:

$$\hat{L}_t = \frac{1}{1-rr}\hat{D}_t - \frac{rr}{1-rr}\hat{r}_t \tag{A24}$$

For notational convenience the relevant log-linearized equations with variables

[5]Log-linearisation of interest rate is defined as difference from steady state: $R_t = R + \widehat{R}_t$.

denoting deviation from steady-state are reported in the main text without $\widehat{\cdot}$.

The benchmark model has 22 endogenous variables $\{c$, n, m, w, q, P, π, mc, D, b, Ω, EFP, R^T, R^B, R, R^L, R^D, λ, ξ, T, r, $re\}$, 6 lagged variables $\{P_{t-1}$, D_{t-1}, c_{t-1}, b_{t-1}, re_{t-1}, $R^B_{t-1}\}$ and 8 exogenous shocks $\{a1, a2, a3, a4, a5, a6, a7, a8\}$. The equations (A1) through (A24), 6 lagged identities construct the model to be solved by King and Watson (1998) algorithm. For the simulation we consider contemporaneous shocks to $a_1, ..., a_8$. To obtain the simulated series we have produced 10,000 draws from a normal distribution, discard the first 500 and considered the middle 100.

4 OMO Model

In the bond-OMO variant as central bank bond holdings must equal reserves, we can substitute and re-arrange to give the log linear relationship

$$b^p \hat{b}_t^p = b\hat{b}_t - r\hat{r}_t \tag{51}$$

which we add to our system of equations. It is this newly defined variable b^p which determines the amount of collateral households have available so we substitute it for b in the loan supply and marginal value of collateralised lending equations.[6]

In the capital-OMO variant, capital could equally be bought by the central bank in exchange for new reserves in the same way that bonds are. To do this we introduce an equation defining total capital holdings as a function of an exogenous shock in the same way we did for bond holdings. The central bank can now hold two assets on its balance sheet, so we hold the level of bonds fixed as before and set the steady state value of capital held by the central bank at zero. By defining private sector capital holdings in a log linear form as

$$k^p \hat{k}_t^p = b\hat{b}_t - r\hat{r}_t \tag{52}$$

what we model is a situation where the central bank buys and sells illiquid assets/capital in exchange for reserves.

The bond-OMO model introduces b_t^p, and its one period lag as additional variables whilst the capital-OMO introduces K_t and K_t^P so each have 31 endogenous variables.

[6] As we deal with a consolidated government budget constraint, the net effect of interest payments on bonds held by the central bank is zero. Therefore, it is appropriate to also change the terms in b to terms in b^p in this equation as well.

5 Taylor Approximation

This section outlines the process of approximating a utility function through a second order Taylor expansion. Our initial utility function is given by;

$$U = E_0 \sum_{t=0}^{\infty} \beta_t \left[\phi \log(c_t) + (1-\phi) \log(1 - m_t^s - n_t^s) \right] \tag{53}$$

As our function is additive we can estimate our Taylor approximations separately for each term and then bring them together. First we derive our approximation of $\log(c_t)$

$$\log(c_t) \approx U_c \breve{c}_t + \frac{U_{cc} \breve{c}_t^2}{2} + O^3 \tag{54}$$

Where O^3 represents all terms higher than second order. This then expands to

$$\approx \frac{1}{c} c(\hat{c}_t + \frac{1}{2}\hat{c}_t^2) - \frac{1}{c^2} \frac{c^2 \hat{c}_t^2}{2} + O^3 \tag{55}$$

We can then cancel out like terms to simplify this to

$$\log c_t \approx \hat{c}_t + O^3 \tag{56}$$

The same process for our second argument yields

$$\log(1 - m_t - n_t) \approx U_{(1-m-n)}(1 - \breve{m}_t - \breve{n}_t) + \frac{U_{(1-m-n)(1-m-n)}(1 - \breve{m}_t - \breve{n}_t)^2}{2} + O^3 \tag{57}$$

which expands to

$$\begin{aligned} \approx \quad & \frac{1}{(1-m-n)}(1-m-n)\left[(1 - \breve{m}_t - \breve{n}_t) + \frac{1}{2}(1 - \breve{m}_t - \breve{n}_t)^2\right] \\ & - \frac{1}{(1-m-n)^2} \frac{(1-m-n)^2(1 - \breve{m}_t - \breve{n}_t)^2}{2} + O^3 \end{aligned} \tag{58}$$

which simplifies to

$$\log(1 - m_t - n_t) \approx (1 - \breve{m}_t - \breve{n}_t) + O^3 \tag{59}$$

Putting these back into our equation (53) we get:

$$U_t - U = \phi \hat{c}_t + (1-\phi)(1 - \hat{m}_t - \hat{n}_t) + O^3 \tag{60}$$

Which is our initial approximation of the deviation of current utility in any given period compared to steady state utility. However, our aim is to find our function in terms of variances (second order terms) so the next step is to simplify this and eliminate as many first order terms as we can through substitution of other equations within our model. We have three first order terms to deal with $\hat{c}_t$ $\hat{m}_t$ and $\hat{n}_t$.

Let us begin with our labour demand function, converting it to log deviations from steady state:

$$\hat{m}_t = -\hat{w}_t - \frac{(1-\alpha)c}{mw}(\hat{c}_t + \frac{\phi}{\lambda}\hat{\lambda}_t) \tag{61}$$

If we assume $\frac{(1-\alpha)c}{mw}$ is equal to one in order to simplify the analysis and substitute this back into equation (60) we get

$$U_t - U = \phi\hat{c}_t + (1-\phi)(1 + \hat{w}_t + \hat{c}_t + \frac{\phi}{\lambda}\hat{\lambda}_t) - \hat{n}_t) + O^3 \tag{62}$$

We can then bring together our terms in $\hat{c}_t$ and this cancels to give

$$U_t - U = \hat{c}_t + (1-\phi)(1 + \hat{w}_t + \frac{\phi}{\lambda}\hat{\lambda}_t) - \hat{n}_t) + O^3 \tag{63}$$

Next we can use our marginal cost function:

$$\hat{c}_t = \hat{w}_t + \hat{n}_t - \widehat{mc}_t \tag{64}$$

If we take a second order approximation of this equation we get:

$$c(\hat{c}_t + \frac{1}{2}\hat{c}_t^2) = w(\hat{w}_t + \frac{1}{2}\hat{w}_t^2) + n(\hat{n}_t + \frac{1}{2}\hat{n}_t^2) - mc(\widehat{mc}_t + \frac{1}{2}\widehat{mc}_t^2) \tag{65}$$

Solving for $\hat{c}$:

$$\hat{c}_t = \frac{w}{c}(\hat{w}_t + \frac{1}{2}\hat{w}_t^2) + \frac{n}{c}(\hat{n}_t + \frac{1}{2}\hat{n}_t^2) - \frac{mc}{c}(\widehat{mc}_t + \frac{1}{2}\widehat{mc}_t^2) - \frac{1}{2}\hat{c}_t^2 \tag{66}$$

Bringing like terms together and ordering our equation so that first order terms are together and second order terms are grouped together we get:

$$\begin{aligned} U_t - U \;=\; & \left(\frac{w}{c} + (1-\phi)\right)\hat{w}_t + \left(\frac{n}{c} - (1-\phi)\right)\hat{n}_t - \frac{mc}{c}\widehat{mc}_t + \frac{(1-\phi)\,\phi}{\lambda}\hat{\lambda}_t \\ & + \frac{w}{2c}\hat{w}_t^2 + \frac{n}{2c}\hat{n}_t^2 - \frac{mc}{2C}\widehat{mc}_t^2 - \frac{1}{2}\hat{c}_t^2 + O^3 \end{aligned} \tag{67}$$

The term in $\hat{n}_t$ can be approximated using the two lemmas described in Galí

(2008) to give:

$$\hat{n}_t = \frac{1}{1-\eta}(\hat{c}_t + \frac{1}{2}\frac{\theta}{\chi}\hat{\pi}_t^2) \tag{68}$$

where $\chi = \frac{(1-\theta)(1-\beta\theta)}{\theta}\frac{1-\eta}{1+\eta(\theta-1)}$. If we substitute this back into our equation we eliminate the term in $\hat{n}_t$ but replace it with a first order term in $\hat{c}_t$.

$$\begin{aligned} U_t - U &= \left(\frac{w}{c} + (1-\phi)\right)\hat{w}_t + \frac{\left(\frac{n}{c} - (1-\phi)\right)}{1-\eta}\hat{c}_t - \frac{mc}{c}\widehat{mc}_t \\ &+ \frac{(1-\phi)\,\phi}{\lambda}\hat{\lambda}_t + \frac{w}{2c}\hat{w}_t^2 + \frac{n}{2c}\hat{n}_t^2 - \frac{mc}{2C}\widehat{mc}_t^2 - \frac{1}{2}\hat{c}_t^2 + \frac{1}{2}\frac{\theta}{\chi}\hat{\pi}_t^2 + O^3 \end{aligned} \tag{69}$$

We can eliminate the term in lambda by using our mark up equation

$$\widehat{mc}_t = \hat{\xi}_t - \hat{\lambda}_t \tag{70}$$

Solving for lambda and noting that there is no deviation in ξ:

$$\hat{\lambda}_t = -\widehat{mc}_t \tag{71}$$

so our equation can be written

$$\begin{aligned} U_t - U &= \left(\frac{w}{c} + (1-\phi)\right)\hat{w}_t + \frac{\left(\frac{n}{c} - (1-\phi)\right)}{1-\eta}\hat{c}_t \\ &- \left[\frac{mc}{c} - \frac{(1-\phi)\,\phi}{\lambda}\right]\widehat{mc}_t + \frac{w}{2c}\hat{w}^2 + \frac{n}{2c}\hat{n}^2 - \\ &\frac{mc}{2C}\widehat{mc}_t^2 - \frac{1}{2}\hat{c}_t^2 + \frac{\left(\frac{n}{c} - (1-\phi)\right)}{1-\eta}\frac{1}{2}\frac{\theta}{\chi}\hat{\pi}_t^2 + O^3 \end{aligned} \tag{72}$$

Leaving only 3 first order terms. We can now replace $\hat{w}$ as a function of terms of $\hat{n}$, $\hat{c}$ and $\widehat{mc}$, leaving us 3 terms still, but one of which is $\hat{n}$. We can convert this $\hat{n}$ term into a term in the volatility of inflation and $\hat{c}$ leaving us with just two first order terms; one in $\hat{c}$ and one in $\widehat{mc}$. We therefore rearrange to make $\hat{w}$ the subject leaving us with a first order term in $\hat{c}$ and a first order term in $\widehat{mc}$ but with everything else being second order or higher.

$$U_t - U = \left[\left(\frac{w}{c} + (1-\phi)\right) + \frac{\left(\frac{n}{c} - (1-\phi)\right)}{1-\eta} - \frac{\left(\frac{w}{c} + (1-\phi)\right)}{1-\eta}\right]\hat{c}_t \tag{73}$$
$$-\left[\frac{mc}{c} - \frac{(1-\phi)\,\phi}{\lambda} - \left(\frac{w}{c} + (1-\phi)\right)\right]\widehat{mc}_t$$
$$+\frac{w}{2c}\hat{w}_t^2 + \frac{n}{2c}\hat{n}_t^2 - \frac{mc}{2c}\widehat{mc}_t^2 - \frac{1}{2}\hat{c}_t^2$$
$$+\left[\frac{\left(\frac{n}{c} - (1-\phi)\right)}{1-\eta}\frac{1}{2}\frac{\theta}{\chi} - \frac{\left(\frac{w}{c} - (1-\phi)\right)}{1-\eta}\frac{1}{2}\frac{\theta}{\chi}\right]\hat{\pi}_t^2 + O^3$$

We can see that the welfare function contains linear terms in $\hat{c}_t$ and $\widehat{mc}_t$. They might tend to dominate the second order terms. We therefore choose weights $(1-\phi)$ and ϕ so that first order terms disappear in the welfare approximation. The particular weights to choose are those that solve the system:

$$\begin{array}{c} \left(\frac{w}{c} + (1-\phi)\right) + \frac{\left(\frac{n}{c} - (1-\phi)\right)}{1-\eta} - \frac{\left(\frac{w}{c} + (1-\phi)\right)}{1-\eta} = 0 \\ -\frac{mc}{c} + \frac{(1-\phi)\phi}{\lambda} + \left(\frac{w}{c} + (1-\phi)\right) = 0 \end{array} \tag{74}$$

Leaving us with the welfare approximation:

$$U_t - U = +\frac{w}{2c}\hat{w}_t^2 + \frac{n}{2c}\hat{n}_t^2 + \frac{mc}{2c}\widehat{mc}_t^2 - \frac{1}{2}\hat{c}_t^2 \tag{75}$$
$$+\left[\frac{\left(\frac{n}{c} - (1-\phi)\right)}{1-\eta}\frac{1}{2}\frac{\theta}{\chi} - \frac{\left(\frac{w}{c} - (1-\phi)\right)}{1-\eta}\frac{1}{2}\frac{\theta}{\chi}\right]\hat{\pi}_t^2 + O3$$

where $(1-\phi) = \frac{\frac{w}{c}(1-\eta) + \left(\frac{n}{c} - \frac{w}{c}\right)}{1-\eta} = \frac{\frac{n}{c} - \eta\frac{w}{c}}{1-\eta}$ and $\phi = \lambda\left[(1+\eta)\frac{\frac{mc}{c} + \frac{w}{c}}{\frac{n}{c} - \eta\frac{w}{c}} + 1\right]$. Therefore our welfare approximation can be written as:

$$U_t - U = \frac{w}{2c}\hat{w}_t^2 + \frac{n}{2c}\hat{n}_t^2 - \frac{mc}{2c}\widehat{mc}_t^2 - \frac{1}{2}\hat{c}_t^2 \tag{76}$$
$$-\frac{1}{2}\left[\frac{\theta}{\chi(1-\eta)}\left(\frac{w}{c}\left(1+\eta^2\right) - \frac{n}{c}\right)\right]\hat{\pi}_t^2 + O^3$$

The above welfare function can be expressed in terms of quadratic loss function:

$$
\begin{aligned}
U_t - U &= -\frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t L_t + O^3 \qquad (77) \\
\text{with } L_t &= \frac{1}{2} \left[\begin{array}{c} \sigma_c^2 + \left[\frac{\theta}{\chi(1-\eta)} \left(\frac{w}{c} \left(1+\eta^2\right) - \frac{n}{c} \right) \right] \sigma_\pi^2 - \\ \frac{w}{c}\sigma_w^2 - \frac{n}{c}\sigma_n^2 + \frac{mc}{c}\sigma_{mc}^2 \end{array} \right]
\end{aligned}
$$

This is equation (27) in the text.

References

[1] Chadha, J.S., L, Corrado (2012). "Macro-Prudential Policy on Liquidity: What Does a DSGE Model Tell Us?", *Journal of Economics and Business*, 64 (1), pp37-62.

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

[3] Goodfriend, M. and B. T. McCallum (2007). "Banking and Interest Rates in Monetary Policy Analysis: A Quantitative Exploration", *Journal of Monetary Economics*, 54, pp. 1480-1507.

[4] King, R. G. and Watson, M. W. (1998). "The Solution of Singular Linear Difference Systems under Rational Expectations", *International Economic Review*, 39(4), Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance, pp. 1015-1026.