We provide a critical review of macroeconomic models used for monetary policy at central banks from a finance perspective. We review the history of monetary policy modeling, survey the core monetary models used by major central banks, and construct an illustrative model for those readers who are unfamiliar with the literature. Within this framework, we highlight several important limitations of current models and methods, including the fact that local-linearization approximations omit important nonlinear dynamics, yielding biased impulse-response analysis and parameter estimates. We also propose new features for the next generation of macrofinancial policy models, including: a substantial role for a financial sector, the government balance sheet and unconventional monetary policies; heterogeneity, reallocation, and redistribution effects; the macroeconomic impact of large nonlinear risk-premium dynamics; time-varying uncertainty; financial sector and systemic risks; imperfect product market and markups; and further advances in solution, estimation, and evaluation methods for dynamic quantitative structural models.

Keywords: Macrofinancial Models; Dynamic Stochastic General Equilibrium Models; Macropudential Policy; Systemic Risk; Monetary Policy Analysis

JEL Classification: G12, G29, C51

*We thank Tobias Adrian, Richard Berner, Fernando Duarte, Mark Gertler, Lars Peter Hansen, Nobuhiro Kiyotaki, Nellie Liang, Debbie Lucas, Johannes Pfeifer, David Skeie and seminar participants at Macro Financial Modeling Group 2013 Meeting at Chicago and Macro Financial Modeling Group 2015 Meeting at New York for helpful discussion and comments, and we thank Jayna Cummings for editorial assistance. We appreciate the generous support from the Macro Financial Modeling Group, the Alfred P. Sloan Foundation, and the MIT Laboratory for Financial Engineering.

†Dou is at the Finance Department, Wharton School, University of Pennsylvania, 2318 Steinberg Hall-Dietrich Hall, 3620 Locust Walk, Philadelphia, PA 19104, USA, wdou@wharton.upenn.edu. Lo is at the MIT Sloan School of Management, 100 Main Street, E02–618, Cambridge, MA 02142, USA, alo-admin@mit.edu. Muley is at the Department of Economics, MIT, 50 Ames Street, Cambridge, MA, 02139, USA, muley@mit.edu. Uhlig is at the Department of Economics, University of Chicago, 1126 East 59th Street, Chicago, IL 60637, USA, huhlig@uchicago.edu.
1 Introduction

The recent financial crisis and the Great Recession of 2007–2009 revealed serious gaps in commonly used approaches to define, measure, and manage financial sector activities that pose risks to the macroeconomy as a whole.

One emerging narrative is that macroeconomic models commonly employed at policy institutions for evaluating monetary policy lack the analytical specificity to account for important financial sector influences on the aggregate economy. A new generation of enhanced models and advanced empirical and quantitative methodologies are needed by policymakers and need to be provided by researchers to better study the impact of shocks that are initially large or build endogenously over time.

This paper presents a review of these macroeconomic models and their empirical methods. Through this review, we hope to clarify the most important challenges faced by existing macroeconomic models for monetary policy analysis, and summarize some recent advances in new modeling and quantitative techniques. The primary goal of this paper is to provide insight, guidance, and motivation for the next generation of young scholars—especially those at the intersection of macroeconomics and financial economics—to develop more effective macroeconomic models for policy decisions.

There has been a remarkable evolution of macroeconomic models used for monetary policy at major central banks around the world, in aspects such as model formulation, solution methods, estimation approaches, and the communication of results between central banks. Central banks have developed many different classes and variants of macroeconomic models in the hopes of producing a reliable and comprehensive analysis of monetary policy. Early types of models included quantitative macroeconomic models,\(^1\) reduced-form statistical models, structural vector autoregressive models, and large-scale macroeconometric models, a hybrid form combining the long-run structural relationships implied by a partial equilibrium treatment of theory (e.g., the decision rule for aggregate consumption) and reduced-form short-run relationships employing error-correcting equations.

Over the past 20 years in particular, there have been significant advances in

\(^1\)For example, the Wharton econometric model and the Brookings model.
the specification and estimation for New Keynesian Dynamic Stochastic General Equilibrium (New Keynesian DSGE) models. Significant progress has been made to advance policymaking models from the older static and qualitative New Keynesian style of modeling to the New Keynesian DSGE paradigm. The New Keynesian DSGE model is designed to capture real world data within a tightly structured and self-consistent macroeconomic model. The New Keynesian DSGE model has explicitly theoretical foundations, allowing it to circumvent the Sims critique (see Sims, 1980) and the Lucas critique (see Lucas, 1976), and therefore it can provide more reliable monetary policy analysis than earlier models.

A consensus baseline New Keynesian DSGE model has emerged, one that is heavily influenced by estimated impulse response functions based on Structural Vector Autoregression (SVAR) models. In particular, a baseline New Keynesian DSGE model has recently been shown by Christiano et al. (2005) to successfully account for the effects of a monetary policy shock with nominal and real rigidities. Similarly, Smets and Wouters (2003, 2007) show that a baseline New Keynesian DSGE model can track and forecast time series as well as, if not better than, a Bayesian vector autoregressive (BVAR) model. New Keynesian DSGE models have been developed at many central banks, becoming a crucial part of many of their core models.\(^2\) Sbordone et al. (2010) have emphasized that an advantage of New Keynesian DSGE models is that they share core assumptions about the behavior of agents, making them scalable to relevant details to address the policy question at hand. For example, Smets and Wouters (2007) introduced wage stickiness and investment frictions into their model, Gertler et al. (2008) and Blanchard and Galí (2010) incorporated labor market search and wage bargaining, and Bernanke et al. (1999), Chari et al. (1995) and Christiano et al. (2008) studied the interaction between the financial sector and macroeconomic activity. Interestingly, DSGE models with richer structures have been included in core models of several central banks.

However, the devastating aftermath of the recent financial crisis and the Great

\(^2\)The Bank of Canada, the Bank of England, the Central Bank of Chile, the European Central Bank, the Norges Bank, the Sveriges Rikbank, and the U.S. Federal Reserve have all incorporated New Keynesian DSGE models into their core models.
Recession has prompted another rethink of monetary and central banking policies, which are now facing many new challenges. Most macroeconomists and many policy-makers and regulators have called for a new generation of DSGE models. The first and foremost critique of the current state of standard New Keynesian DSGE models is that these models lack an appropriate financial sector with a realistic interbank market, and as a result, these models fail to fully account for an important source of aggregate fluctuations, such as systemic risk from the financial system. Second, the linkage between the endogenous risk premium and macroeconomic activity is crucial for policymakers to understand the transmission mechanism of monetary policy, especially in financially stressed periods. In models that lack a coherent endogenous risk premium, policy experiments become unreliable in stressed periods, and the model cannot provide a consistent framework for conducting experimental stress tests regarding financial stability or macroprudential policy. Third, heterogeneity among the players in the economy is essential to our understanding of inefficient allocations and flows between agents. These inefficiencies may have important effects on the equilibrium state of the economy.

Finally, policy makers will need a new generation of models with a unified and coherent framework for both conventional and unconventional monetary policies. For example, at the onset of the financial crisis, the zero lower bound for short-term interest rates went from a remote possibility to reality with frightening speed. This led central banks to quickly develop unconventional measures to provide economic stimulus, including credit easing, quantitative easing, and extraordinary forward guidance. These unconventional measures require a proper platform to be analyzed. Furthermore, these measures have blurred the boundary between monetary policy and fiscal policy. Through these policies, central banks gave preference to some debtors over others (e.g., industrial companies, mortgage banks, governments), and some sectors over others (e.g., export versus domestic). In turn, the distributional effects of monetary policy were much stronger than in normal times; hence, these measures are sometimes referred to as quasi-fiscal policy. As Sims emphasized, a reliable monetary policy experiment cannot ignore the effect of ongoing fiscal policy. In order to implement unconventional measures during the crisis, central banks put
much more risk onto government balance sheets than ever before, which had the potential to lead to substantial losses. Thus the government balance sheets in these models should be forward-looking, and its risk characteristics are crucial to the success of the model.

Methodological and empirical challenges have arisen along the way. First, advanced nonlinear solution methods and estimation approaches are necessary, if one wishes to guarantee that key nonlinear dynamics in the financial market and the macroeconomy are eventually captured in quantitative analysis. Second, data availability and risk measurement are always a central challenge in macroeconomic modeling, but especially so in the wake of the global financial crisis and the subsequent global economic recession. Brunnermeier et al. (2012) pointed out that our current measurement systems are outmoded, leaving regulators, academics, and risk managers in a dangerous position. Assessing systemic risk requires viewing data on the financial sector through the lens of a macroeconomic model. However, macroeconomics in particular frames questions and builds models based on available data, and we have so far lacked the data to construct macro-financial models. New infrastructure for detailed micro-level financial data collection is necessary and critical for further risk measurement development and model construction. In fact, the Office of Financial Research (OFR) at the U.S. Department of the Treasury already has this mandate, and the first steps toward a new, comprehensive financial data collection system are already underway.

These are the issues we hope to address in this review. We begin in Section 2 with a brief summary of goals and mechanisms of central banking monetary policy, a history of macroeconomic policy models, and some motivation for the most popular framework today: the DSGE model. To accelerate progress in macrofinancial policy modeling, we present a fully specified canonical example of New Keynesian DSGE model with financial sectors in Section 3 and Appendix C that readers can work with immediately (an open-source software implementation is provided at MFMWEBSITE). The model in Section 3 is solved globally. We hope the contribution of our code and global solution method to this review may be of general interest to a broader group of researchers in the macrofinancial and monetary economics community. We calibrate and estimate this model using historical data in Section D, and explore the
empirical implications of our canonical model. These implications will uncover clear weaknesses of the DSGE framework, and in Section 4 we consider a series of critiques of this framework as well as suggestions for future directions for developing the next generation of DSGE policy models. We conclude in Section 5 and, to motivate readers to take a more active interest in practical applications of policy modeling, we present a survey in the Appendix of the core models employed by the U.S. Federal Reserve (Fed), the European Central Bank (ECB), the Bank of England, and the Bank of Canada. Most of these models are well documented and young scholars are encouraged to develop improvements that could have enormous impact for macroeconomic policy and society.

2 Central Banks, Monetary Policy, and Models

According to a 1977 amendment to the Federal Reserve Act, the U.S. Federal Reserve’s monetary policy has three basic objectives. These are to promote a “maximum” sustainable output and employment, a moderate long-term interest rate, and a level of “stable” prices. These three basic goals of monetary policy are shared by most major central banks. For example, the Treaty on the Functioning of the European Union also promotes the primary objective of supporting stable prices.

Price stability is an economic environment that avoids both prolonged inflation and deflation. In such an environment, households and firms can make financial decisions without worrying about where prices are headed.

It has previously been argued that a monetary policy directed at maintaining aggregate price level stability will lessen both the incidence and severity of financial instability, most famously in the Schwarz Hypothesis (see e.g. Schwartz, 1988; Schwarz, 1995). Obviously, price stability does not guarantee financial stability, and the recent financial crisis provides an excellent case in point. Some have even argued that periods

---

3 The terms “price stability” and “inflation stability” are often used synonymously, and we shall do the same in this review.

4 The ECB’s Governing Council has announced a quantitative definition of price stability: “Price stability is defined as a year-on-year increase in the Harmonized Index of Consumer Prices (HICP) for the euro area of below 2%.”

---
of tranquility can aid the build-up of financial instability, for example, Borio and Lowe (2002), Brunnermeier and Sannikov (2014), Brunnermeier et al. (2012) and Duarte and Eisenbach (2013).

There has been a long debate on whether the central bank is the natural guarantor of the stability of the financial system. In this paper, we emphasize the natural responsibility of the central bank in this matter for the following reasons. First, the central bank is the only provider of the legal means of payment, and therefore it is the only provider of immediate liquidity during a financial crisis. Second, a natural role of the central bank is to ensure the smooth functioning of the national payment system. As such, it is centrally positioned to monitor and combat systemic risk, defined here as the risk of the serious impairment of a large or crucial part of the financial system. Third, the financial system is the transmission mechanism through which monetary policy has its effect on the real economy. The status of the financial system is critical for the central bank to have any desirable impact and achieve its monetary objectives. For this reason alone, central banks have a natural interest in maintaining a sound financial system. Finally, financial stability may play an important role in guaranteeing price stability, already a basic role of the central bank. More recent discussions on the tradeoff between financial conditions and financial stability can be found in Adrian and Liang (2014), among others.

It is also one of a central bank’s main responsibilities to maintain a sound central bank balance sheet. Central bank balance sheets have proved crucial in designing and understanding policies pursued in the wake of financial crises in recent years. In particular, large-scale asset purchase programs became the primary tools in efforts to prevent any renewal of the financial meltdown as the effective zero lower bound for interest rates was reached. With short-term interest rates near zero, and the effectiveness of conventional monetary policies constrained as a result of a liquidity trap, these policies sought to provide additional monetary stimulus by lowering the long-term interest rate on government bonds. A loss of confidence in banks and in many financial products in the advanced economies disrupted global financial markets. This occurred in large part because the normal operations of financial markets became impaired, blocking the transmission of lower policy rates to the real economy. Central
banks countered this by buying unconventional assets on a large scale. They started with short-term lending, or by buying short-term assets, but progressively moved towards buying long-term assets. At present, the aggregate size of central bank balance sheets in advanced countries is nearly $8 trillion, the equivalent of more than 20% of GDP. In some cases, balance sheets are still growing.

Whether the risks that may be inherent in these large balance sheets matter, and if so, how much, merits detailed attention going forward. A country is surely better off if the central bank has the full financial strength needed to carry out its functions.

### 2.1 The Mechanisms and Tools of Monetary Policy

The scope of monetary policy is limited in terms of what variables the central banks can directly control, to what extent, and for how long the impact of monetary policy will last. The process of how the shock in monetary policy leads to changes in aggregate economic variables—including inflation, output, employment, consumption, and investment— is known as the monetary policy transmission mechanism. In order to analyze the monetary policy transmission mechanism, an admittedly superficial, but nonetheless helpful, approach is to think about the particular channels of how monetary policies operate and affect the real economy. This approach is superficial, since ultimately, it is the general equilibrium dynamic interdependence of a solution for all the variables that determines the outcome. Examining individual “channels” can be misleading for that reason. Furthermore, examining channels is unlikely to be of much help for a precise, quantitative analysis. On the other hand, this approach is helpful because thinking about the complexities of general equilibrium can be hard, and thinking about channels can be intuitive and instructive. Furthermore, it most easily expresses conventionally held views about how monetary policy affects the economy.

The most common channels arising in discussions of monetary policy are the interest rate channel, the inflation expectations channel, the balance sheet channel, the bank credit channel or bank lending channel, the exchange rate channel, and the asset price channel. Model builders often seek to construct models which incorporate these
channels, or else provide theories with a more contrarian view. Here we only seek to describe the conventional perspective on these channels in admittedly rather loose and imprecise language, and not to raise questions about the validity of these conventional perspectives per se. One should, of course, not confuse models constructed so that they exhibit these channels with evidence that these channels are actually present. It is conceivable that considerable progress can be made by breaking through the apparent circularity of such discussions. Figure 1 presents an overview, together with a list of commonly used shocks as sources of fluctuations and a stylized description of the economy overall.

Let us first consider the interest rate channel. The conventional interest rate channel means that lower nominal short-term interest rates lead to lower real interest rates because prices are sticky. These lower interest rates in turn promote investment and consumption, but discourage savings. Conversely, higher interest rates stimulate savings and lower consumption and investment in the short run. Matters are more complicated when examined in detail, of course. There is no such thing as a single interest rate. A change in the official short-run interest rates directly affects money market interest rates, but it only indirectly affects lending and deposit rates, which are set by commercial banks to their debtors and depositors, respectively. While the central bank can control short-term interest rates, the real economy is mainly affected by the medium- and long-term deposit and lending rates charged by these commercial banks to their customers. These rates depend not only on the interest rate set by the monetary authority, but also on a number of other determinants, such as inflation expectations and the risk premium of other channels. The benchmark consensus stipulates that monetary policy has almost no influence on the long-run aggregate supply. Thus, in the short-to-medium run, monetary policy can influence only the difference between the actual level of economic activity and the one that is sustainable over the long run, the potential output. This difference is called the output gap.

We next consider the inflation expectations channel. The inflation expectations channel means that looser monetary policy will result in more inflation down the road, which in turn determines economic choices in the present. At given nominal
Figure 1: Illustrative Graph for Transmission Channels.
interest rates, increases in expected inflation lowers the real rate, resulting in higher consumption and fewer savings. The resulting inflation dynamics will reinforce inflation expectations. Higher demand for consumption or investment goods will put pressure on their marginal cost of production. Faced with an increase in production costs, some firms might decide to reduce their profit margins, but plausibly, many firms will gradually transfer these costs onto the final price, which will eventually generate realized inflation, out of inflation expectations. Inflation expectations are heavily affected by the perceptions of economic agents regarding the central bank’s commitment towards achieving its primary objectives. Anchoring inflation expectations can be one of the most powerful and efficient channels of monetary policy transmission, provided that it is transparent, and its actions are regarded as credible. This is easier said than done. As pointed out by Blanchard (2009), “…although we very much want to believe that monetary policy can anchor inflation expectations, I am not sure we actually understand whether and how it can actually achieve it.”

The balance sheet channel is deeply associated with the external-finance premium, which is defined as the wedge between the cost of capital internally available to firms and the cost of raising capital externally by issuing equity or borrowing from corporate debt markets. The balance sheet channel means that an interest rate increase worsens the balance sheets of firms and raises their external finance premium, as it raises the firm debt burden through higher interest payments on floating rate debt, lowers the value of the firms business via the overall reduction in demand, and decreases the value of the firm’s collateral through decreased asset prices. An increase in the external finance premium in turn makes firms more reluctant to invest and to expand, resulting in a decrease in aggregate economic activity. The balance sheet channel is potentially dangerous since it could amplify and propagate small fluctuations via a pecuniary externality or an adverse feedback loop, as is emphasized in Kiyotaki and Moore (1997). The balance sheet channel likewise affects households, as a rise in interest rates lowers the value of their assets, reducing their demand.

The asset price channel is closely related to the balance sheet channel. The asset price channel means that lower interest rates will cause more capital to flow into stocks and consequently raise the stock prices, leading to higher investment, as it generates a
higher Tobin’s $q$ as well as make it easier for firms to obtain outside equity financing. Note that this channel is at work even for firms that do not hold debt or assets and for whom the value of their business remains unchanged: it is therefore distinct from the balance sheet channel.

The bank lending channel or bank credit channel means that looser monetary policy will enable banks and other monetary financial institutions to lend and provide credit more easily, which in turn stimulates economic activity, once these lending activities come to pass. One part of the bank credit channel is essentially the balance sheet channel as applied to the operations of lending institutions. The uniqueness of the bank credit channel for monetary policy transmission is mainly a result of the special role of the financial sector in the economy relative to other sectors. Other parts of the bank lending channel concern the funds available to banks per se. Monetary policies may change the supply of loanable funds available to banks, and consequently the total amount of credit they can extend to borrowers, including both firms and households. It has been argued that the most direct way monetary policy is able to affect the willingness and capacity of bank lending is to control the supply of bank reserves. Some believe that the bank lending channel has been the most important channel for monetary policy during the recent financial crisis and the Great Recession. An important study of the bank lending channel in monetary policy was by Bernanke (1983). Since then, there has been an extensive academic literature on the topic; see, for example, Bernanke and Blinder (1988), Kashyap and Stein (1994), and Bernanke and Gertler (1995), among others.

We finally consider the exchange rate channel. As monetary policy affects the quantity of domestic currency as well as the opportunity costs for holding it, it can affect exchange rates. The exchange rate channel means that these exchange rate effects in turn lead to changes in export and import decisions as well as international portfolio choices, affecting aggregate economic activity. For example, a depreciation reduces the disposable income that is left after servicing the regular payments on a foreign currency loan, since economic agents with revenues denominated in domestic currency would have to pay a greater amount following a depreciation of the currency. Exchange rates appear to move in response to many influences, however, and the
relative effect of monetary policy here seems to be rather modest.

2.2 A Brief History of Macroeconomic Models

According to Gali and Gertler (2007), economists and policymakers began to be skeptical about large-scale macroeconometric modeling during the 1970s for two related reasons. First, some existing models, like the Wharton econometric model and the Brookings model, failed to forecast the stagflation in the 1970s. These traditional large-scale macroeconometric models were originated by Klein (1985, 1991) and have been in use for decades. Second, leading macroeconomists leveled harsh criticisms about their underlying framework. Lucas (1976) and Sargent (1981) in particular argued that the absence of an optimization-based approach to the development of the structural equations meant that the estimated model coefficients were likely not invariant to shifts in policy regimes, or to other types of structural changes. Similarly, Sims (1980) argued that the absence of convincing identification assumptions to sort out the vast simultaneity among macroeconomic variables meant that one could have little confidence that the parameter estimates would be stable across different policy regimes. More precisely, Sims (1980) argued that large-scale macroeconometric models may fit the data well, but that they will provide misleading answers due to non-credible identification restrictions.

Despite the criticisms by Lucas (1976) and Sims (1980), many central banks continued to use large-scale macroeconometric models and reduced-form statistical models in the 1980s and 1990s to produce forecasts of the economy that presumed no structural change, perhaps for lack of suitable alternatives. Over the past two decades, however, new quantitative and micro-founded macroeconomic frameworks for monetary policy evaluations have made large inroads. The building blocks for the development of this new framework were two independent literatures that emerged in response to the downfall of traditional macroeconometric modeling: the New Keynesian approach and the real business cycle (RBC) theory. The New Keynesian paradigm arose in the 1980s as an attempt to provide microfoundations for key Keynesian concepts such as the inefficiency of aggregate fluctuations, nominal price stickiness, and the non-neutrality
of money (see, e.g. Mankiw and Romer, 1991). The models of this literature, however, were typically static and designed mainly for qualitative, as opposed to quantitative, analysis. By contrast, RBC theory, which was developed concurrently, demonstrated how it was possible to build quantitative macroeconomic models exclusively from the “bottom up”—that is, from explicit optimizing behavior at the individual level (see, e.g. Prescott, 1986). The RBC models usually abstracted monetary and financial factors, or assigned a negligible influence on real activity to monetary policy. New frameworks arising late in the 1990s and more prominently in the early 2000s reflected a natural synthesis of the New Keynesian and the RBC approaches. A variety of labels have been used for this new framework. For example, Goodfriend and King (1997) employ the term “New Neoclassical Synthesis”, while Woodford (2003) uses “NeoWicksellian” and Clarida et al. (1999) uses “New Keynesian”. Now, they are most often referred to as New Keynesian DSGE models, as they incorporate nominal stickiness and the resulting monetary non-neutrality into a fully specified dynamic general equilibrium framework. It is important to keep in mind, however, that a substantial portion and backbone of these models is a real business cycle model: it then is a matter of quantity which of the forces dominate, not of principle.

Central banks nowadays use a wide range of macroeconomic models and tools for forecasting and monetary policy analysis, including large-scale macroeconometric models, reduced-form statistical models, structural autoregressive models, and New Keynesian DSGE models. The characteristics of the various models are summarized in Table 1. Large-scale macroeconometric models constrain purely data-driven models in such a way that the long-run dynamic behavior of the variables converges to the theoretical long-run steady state. In econometric terms, the macroeconometric models developed and used by central banks are essentially large-scale restricted vector error-correction models (VECM). This approach puts less emphasis on theory, insofar as short-run dynamics are largely data-driven and long-run relations implied by theory still have to be confirmed by empirical work. For instance, the modeler would not insist that the model has a balanced growth equilibrium, but instead would test whether the cointegrating relation implied by this is present in the data. Examples of this type of macroeconometric model include the Bank of England’s earlier Medium-Term
Macroeconometric Model (MTMM), the Bank of Canada’s Quarterly Projection Model (QPM), the Fed’s MIT-Penn-Social Science Research Council (MPS) and FRB/US model, and the ECB’s Area-Wide Model (AWM) model.

Although the large-scale macroeconometric model still plays an active role at major monetary authorities such as the Fed, there has been a steady shift towards models that place greater emphasis on theoretical consistency. For example, the Bank of Canada’s shifted its principal model from the QPM model to the Terms-of-Trade Economic Model (ToTEM) in late 2005, the ECB replaced its AWM with a New Area-Wide Model (NAWM), and the Fed started to build various DSGE models such as SIGMA and ODE. This vintage of new macroeconometric models uses a calibrated theoretical model to pin down a set of steady-state attractors to describe an error-correcting relationship. Dynamics are driven by assuming that there are adjustment costs between current and long-run levels for variables on a partial equilibrium basis. Higher orders of adjustment costs introduce a role for forward-looking expectations. The full model is a mixture of structural relations implied by a partial equilibrium treatment of theory, such as the decision rule for aggregate consumption, and some reduced-form relations, such as their trade equations, which employs error-correcting relationships. Finally, SVAR models, as first introduced by Sims (1980) as an alternative to traditional large-scale macroeconometric models, enjoy a large degree of popularity at central banks.

One benefit of having multiple models is the opportunity to examine the robustness of policy strategies across models with quite different foundations. According to Tovar (2009) and Chung et al. (2010), central bankers emphasize that in their experience, model-based policy analysis is enhanced by considering multiple models, and indeed, they often learn as much when models disagree as when they agree.

In the next section, we shall focus on recent advances in the development of New Keynesian DSGE models, which now serve as core models and workhorses at several major central banks. Recent efforts on the academic side include the incorporation of financial frictions (i.e., the financial accelerator channel), financial intermediation (i.e., the bank funding channel), nontrivial fiscal policies, and the government/central bank balance sheet in order to analyze unconventional monetary policies. We shall
first lay out a canonical simple New Keynesian DSGE model, of the type which has been the core component of all central bank DSGE models. Following this, we shall discuss extending the model by adding an imperfect credit market and financial intermediation.

A graphical timeline for the generations of models at major central banks is given in Figure 3.

### 2.3 Why DSGE Models?

Dynamic Stochastic General Equilibrium models or DSGE models have become increasingly attractive to the central banks over the past two decades. The reasons why this has happened also naturally provide answers to how we may use DSGE
Figure 2: Generations of Models at Major Central Banks.
models, and we list three of the most important reasons here.

First, its explicit account of the role of expectations and its identification of deep structural parameters makes the DSGE analysis less subject to the Lucas critique, and more suitable for policy analysis and counterfactual experiments. DSGE models emphasize the important role of expectations in assessing alternative policy actions. A DSGE model is able to relate the reduced-form parameters to deeper structural parameters, which makes the use of the model for policy analysis less subject to the Lucas critique (see Lucas, 1976), as those structural parameters are less likely to change in response to changes in policy regime. Therefore, a DSGE model provides a solid organizing framework for understanding and analyzing the economy and policy impacts.

Second, impulse-response analysis allows a DSGE model to identify and decompose economic and policy structural shocks on the quantitative level. A reasonable identification of structural shocks greatly improves the reliability of policy analysis and counterfactual experiments, making the analysis less subject to the Sims critique (see Sims, 1980). The nature of the DSGE model’s structure, not only in terms of its parameters, but also in the way exogenous shocks drive the economy according to the model, makes it possible to tell coherent stories and structure forecasts around it.

Third, the capacity of DSGE models to link model implications to time-series and cross-sectional data makes it particularly useful to discover deep structural parameters. Recent advances in the construction, simulation, and estimation of DSGE models have made it possible to combine a rigorous microeconomic derivation of the behavioral equations of macroeconomic models with an empirically plausible calibration, an estimation which fits the main features of a macroeconomic time series. Beginning with a series of seminal papers, including Mehra and Prescott (1985) and Hansen and Singleton (1982, 1983), it has been shown that asset pricing data are extremely useful in understanding the deep structural parameters of DSGE models. In addition, these structural parameters can be calibrated/estimated using off-model information, especially when time series are short. In terms of the accuracy of the estimation of structural parameters, the DSGE model reduces the risk of overfitting by helping identify parameters and shocks hitting the economy. The danger of overfitting returns
in disguise, however, in the considerable freedom of devising such models and the particular time series used for their estimation.

An increasing number of central banks and policy institutions have started to use New Keynesian DSGE models as their core models, including the Fed, the ECB, the Bank of England, the Bank of Canada, the Bank of New Zealand, and the International Monetary Fund. The most prominent version of a model of that type is the Smets and Wouters (2007) model or its earlier version, Smets and Wouters (2003), which both are close cousins of Christiano et al. (2005). Starting with the basic stochastic neoclassical growth model and its real business cycle (RBC) extensions, this generation of New Keynesian DSGE models have stochastic ingredients, real and nominal frictions such as the cost of capital adjustment, nominal wage and price rigidity, and monopolistic competition. An excellent introduction to the basics of New Keynesian DSGE models can be found in Galí (2008) and Woodford (2003). A common approach used by the central banks is to start with a benchmark New Keynesian DSGE model such as Smets and Wouters (2007) and then to incorporate additional components such as

(i) exogenous shocks, including preference shocks, marginal efficiency shocks, global shocks, risk premium shocks, fiscal policy shocks, etc.

(ii) frictions in the financial market, including collateral constraints, information-based frictions, moral-hazard-based frictions, and limited commitment, see, for example, Kiley and Sim (2011a,b), Brunnermeier and Pedersen (2009), Chari et al. (1995), and Bernanke et al. (1999). There are other recent papers exploring the policy implications of both nominal rigidities and credit frictions, including collateral-based borrowing constraints (see, e.g. Iacoviello, 2005) and limited access to financial markets (see, e.g. Galí et al., 2004, 2007).

(iii) financially constrained (occasionally) intermediaries, see, for example, Gertler and Kiyotaki (2010), He and Krishnamurthy (2013), Christiano et al. (2010), Adrian and Shin (2010a,b), and Adrian et al. (2010).

(iv) agent heterogeneity and redistribution effects of monetary policies, see, for example, Algan and Ragot (2010), Gornemann et al. (2012), Eggertsson and
Krugman (2012), Jermann et al. (2014), and Auclert (2016).

(v) international trade, see, for example, Gal and Monacelli (2005), Corsetti and Pesenti (2005), Lombardo and Ravenna (2014), and Leibovici and Santacreu (2015).

A survey of current models used by the largest central banks is provided in the Appendix. In the next section, we present a simple canonical model of this kind. It can be used to develop intuition and run experiments. Software for estimating this model is also provided at http://bfi.uchicago.edu/mfm/research-tools.

3 A Canonical Model for Unconventional Monetary Policies

The purpose of this section is to provide a benchmark DSGE model for unconventional monetary policy analysis. This model will incorporate two defining features: the nonlinear dynamics of risk premia and the endogenous financial risks originating from imperfect intermediation. The financial crisis of 2008 and the accompanying Great Recession have highlighted the need for such models. Monetary authorities have become particularly aware of nonlinear risk premia and the real investment dynamics caused by dysfunctional financial intermediaries. As a result, unconventional monetary policies have been brought into the limelight by the monetary authorities following the financial crisis, and their role has fast become a focus of academic research.

Our model is a simplified version of the New Keynesian DSGE model proposed by Gertler and Kiyotaki (2010), yet extending it with regard to asset pricing dynamics. Importantly, instead of determining local solutions, we solve our model globally, emphasizing that the nonlinear features of the system are vital for optimal policy design, rather than focusing on a linear approximation around the small perturbations of the deterministic steady state. We believe the global method developed here is of general interest as well as an useful asset for the macro finance and monetary economics communities. In fact, one of our purposes is to show that global solutions are necessary; they capture the crucial nonlinear features designed into models, which are
usually missed by local-linearization approximations. Nonlinear dynamics, especially those due to the instability of financial intermediation, require an unconventional monetary policy whose intensity directly depends on the observable indicators of financial intermediaries to stabilize the capital markets and the greater economy.

We also extend the simple model into a full-fledged standard New Keynesian DSGE model, incorporating a New Keynesian component similar to Christiano et al. (2005) and Smets and Wouters (2003). We include the full New Keynesian DSGE model in Appendix C; there we introduce the model details, estimate the parameters, and analyze the impact of shocks and implications of policies numerically.

3.1 Households

We begin with a description of households in the model, and then turn to firms and intermediaries. There is a continuum of households of unit mass. The members of each household are either workers or bankers. Although there are two types of household members, and certain portfolio constraints among them, we assume the representative household framework following Gertler and Kiyotaki (2010) and Gertler and Karadi (2011) by assuming the household members to be part of a large family, sharing everything or, equivalently, assuming that the full set of Arrow-Debreu securities are available to the members within each household (but not across households), so that the idiosyncratic consumption risks can be fully insured, and all household members have identical preferences. A fraction $\varpi$ of the members of the household are bankers. At any time, a fraction $1 - \theta$ of randomly selected existing bankers exit and become workers, and return their net worth to their household. At the same time, an equal number of workers become bankers within each household, so the proportion of workers and bankers remains fixed. The new bankers receive some start-up funds from their household, which we describe below. The “perpetual youth” assumption in our model is purely technical, with the purpose of guaranteeing the survivorship of workers and preventing the economy from evolving into a degenerate situation.\footnote{Another way to prevent the over-accumulation of intermediary net worth is to assume efficiency losses, as in Bolton et al. (2011).} It can be seen analytically from the condition (38).
Each banker within a household manages a financial intermediary. Workers deposit funds into these financial intermediaries. Household members do not hold capital directly by themselves. Instead, these financial intermediaries hold equity claims on a firm’s capital; their funding, in turn, comes partly from the deposits put down by household members. At the same time, all household members provide labor to the firm for production. The firm and intermediaries will be described in details in Sections 3.2 and 3.4, respectively.

Since all household members have identical preferences, there are no incentives for bankers to pay dividends from their financial intermediaries. Rather, bankers would choose to accumulate the net worth of the financial intermediary up to a critical level from which the financial intermediation will be out of the credit constraints and stay there forever. If the critical value is the total value of all assets, the workers will be eliminated from the economy in the long run. As mentioned previously, to avoid this outcome, we assume that bankers and workers switch roles with probabilities $1 - \theta$ and $(1 - \theta)\frac{\varpi}{1 - \varpi}$, respectively. When a banker switches roles, she pays all the accumulated net worth to her household. On the other hand, when a worker becomes a banker, she needs funds to operate. To be precise, the start-up fund is transferred from the household, and it is equal to a fraction $\frac{N}{(1 - \theta)\varpi}$ of the aggregate asset value for each new banker. The parameter $\mathcal{N} > 0$ characterizes the intensity of funding transfers from workers to bankers. Therefore, the sum of net transfers of funds due to the exogenous switches from bankers to workers is

$$\Pi_t = (1 - \theta)N_t - \mathcal{N}Q_tK_{t+1}$$

(1)

where $N_t$ is aggregate net worth of intermediaries, $K_t$ is the aggregate physical capital installed at the end of period $t - 1$ and used for production in period $t$, and $Q_t$ is the marginal value of shares. As $\mathcal{N}$ increases, the net transfer $\Pi_t$ becomes smaller. The net transfer from bankers to workers, as an exogenous force, fights off the endogenous tendency of diminishing wealth of workers.

There are two points worth mentioning. First, the members of each household in
this economy are divided into bankers and workers, both of whom supply labor, but only bankers own capital. In this way, bankers decide how much capital to accumulate given the capital adjustment costs. The heterogeneity of agents, however, mainly serves as an interpretational device.\(^6\) Second, a potentially important deficiency is that the role of bankers is hardwired into the model: there is no other way to provide capital finance in financial markets. This sidesteps potentially important opportunities for flexibility in funding sources, an issue discussed more substantially in de Fiore and Uhlig (2011) and de Fiore and Uhlig (2015), for example.

The preferences of the household are given by

\[
\mathbb{E}_t \left[ \sum_{\tau=0}^{\infty} \beta^\tau C_t^{1-\gamma} \right],
\]

where \(C_t\) is the consumption and \(L_t\) is the labor supply at time \(t\).

We denote by \(R_{f,t}\) the real interest rate. Let \(B_t\) denote the quantity of the risk-free debt held by the household at the end of period \(t\). The household then faces a state-by-state budget constraint

\[
C_t = W_t L_t + Pr_t + T_t + (1 + R_{f,t-1})B_{t-1} - B_t,
\]

where \(W_t\) is the real wage, \(Pr_t\) is the profit from risky financial holdings, including Arrow-Debreu securities, \(T_t\) is the real lump sum taxes, and \(R_{f,t-1}\) is the net real risk-free rate from the end of period \(t - 1\) to the end of period \(t\). We assume each household provides one unit of labor inelastically, and thus the total labor supply remains \(L_t \equiv 1\). The intertemporal Euler equation for risk-free bond holding is

\[
1 = \mathbb{E}_t \left[ \frac{\Lambda_{t+1}}{\Lambda_t} (1 + R_{f,t}) \right], \quad \text{where} \quad \Lambda_t \equiv \beta^t C_t^{-\gamma}.
\]

\(^6\)For example, the equilibrium and its implications should not be affected if all households are homogeneous. Each household manages an intermediary. Households can invest in a firm’s equity only through their intermediaries. Each household randomly terminates its intermediary and transfers its net worth to all households. Afterwards, they immediately start new intermediaries with funds collected from the households.
Here, $\Lambda_t$ is the marginal utility of consumption $C_t$ at date $t$.

### 3.2 Consumption Goods Sector

There is a continuum of firms of mass unity in the consumption goods sector. Each firm produces its output using an identical Cobb-Douglas production function with capital and labor as its input. The labor market is perfectly competitive, and labor is perfectly mobile across firms. As a result, there exists a representative firm with the same Cobb-Douglas production function:

$$Y_t = A_{c,t} K_t^\alpha L_{c,t}^{1-\alpha}, \quad 0 < \alpha < 1,$$

(5)

where $A_{c,t}$ is an exogenous and stochastic total factor productivity (TFP) parameter for consumption goods production, $K_t$ is aggregate capital installed at the end of period $t-1$, and $L_{c,t}$ is aggregate labor demand in the consumption goods sector. Denote $a_t \equiv \ln A_{c,t}$. The TFP and the capital quality evolve as

$$a_t = a_{t-1} + \sigma_a \epsilon_{a,t},$$

(6)

where $\epsilon_{a,t}$ are i.i.d. standard normal variables.

Firms are owned by intermediaries. There is no friction between firms and intermediaries. A firm’s investment is still subject to constraints, however, since intermediaries can be financially constrained. Firms always choose to pay their earnings to their intermediaries since the marginal value of cash for intermediaries is never less than that for firms. In addition, the capital structure of firms is irrelevant, since the firm’s leverage may always be neutralized by the intermediary leverage, and it is the total leverage that is eventually subject to the credit constraint. Therefore, similar to most macroeconomic and asset pricing models, it is assumed that firms are all-equity firms, and pay out all their earnings. Such a firm has no wealth of its own, i.e. retained earnings. In period $t$, it issues new equity to intermediaries and uses the proceeds to
purchase capital \( I_t \), to be used for production in the next period\(^7 \) \( t + 1 \). The number of shares issued by the firm is normalized to one, although equity issuances occur over time.

Let us describe more details about the firm’s production, hiring, and investment decisions along the timeline. Shocks are realized at the very beginning of each period. Observing the shocks, the firm hires labor at a perfectly competitive wage \( W_t \) and uses the capital \( K_t \) chosen at the end of period \( t - 1 \) to produce the consumption goods using the production function specified in (5).

After production takes place, the firm makes its investment by converting investment goods into new capital, and trading with other firms in a capital spot market. Together with the newly created capital, the depreciated old capital is traded freely in the spot market, and the amount of capital stock is optimally chosen for the next period. We denote the aggregate investment as \( I_t \) and the depreciation rate as \( \delta \). The law of motion for aggregate capital stock is given by

\[
K_{t+1} = I_t + (1 - \delta)K_t. \tag{7}
\]

There are convex adjustment costs for the rate of investment, \( I_t/K_t \). We assume that the cost for creating \( I_t \) units of new capital in terms of investment goods is

\[
\Upsilon_t = \Upsilon(I_t; K_t) \equiv I_t + g(I_t, K_t), \quad \text{where } g(I_t, K_t) \equiv \frac{\vartheta}{2} \left( \frac{I_t}{K_t} \right)^2 K_t, \tag{8}
\]

where \( \vartheta > 0 \) is a constant. The investment goods are produced by investment good firms. In the simpler cases adopted by macroeconomic asset pricing models (e.g. Gomes et al., 2003; Uhlig, 2007; Guvenen, 2009), a firm’s investment converts consumption goods directly into new capital. By introducing an investment goods sector which exogenously maintains a stable scale relative to the whole economy, we can show that

\(^7\)Note that we adopt the more conventional timing assumption, and index capital with the date when it is used in production, not with the date of the decision. As is well known, one needs to be careful when implementing this in solution software such as Dynare or Uhlig’s Toolkit. See Uhlig (1999).
there exists a competitive equilibrium fluctuating around the balanced growth path. Similar methods have been adopted by Kogan et al. (2015) and Dou (2016). Details about the investment goods sector are introduced in Section 3.3.

Let us introduce the payout and valuation of firms. Arriving in period $t$ with capital stock $K_t$ chosen in $t - 1$, the firm will choose labor $L_t$ and investment $I_t$ to maximize the cash flow (net payout)

$$X_t = \left[ A_t K_t t \alpha L_t^{1-\alpha} - W_t L_{c,t} \right] + \left[ Q_t I_t - P_t Y(I_t; K_t) \right], \quad (9)$$

where $Q_t$ and $P_t$ are the equilibrium spot prices of capital and investment goods, respectively. Here, the modeling of a firm’s investment and payout follows the standard macro view of measuring cash flow for investors who own the entire corporate sector (e.g., Larrain and Yogo, 2008). This assumption is commonly adopted in standard production-based asset pricing models for stock returns (e.g., Papanikolaou, 2011). From a macro point of view, net repurchases of equity and debt are cash outflows from the corporate sector. The net capital gain at the end of period $t$ is $Q_t (1 - \delta) K_t = Q_t K_{t+1} - Q_t I_t$, since all firms trade in a perfectly competitive spot market of capital with equilibrium spot price $Q_t$. In effect, there is zero net trade among firms in equilibrium, since all firms are assumed to be homogeneous, even in the ex post situation.

The value of capital after depreciation, $Q_t (1 - \delta) K_t$, can be viewed as the net capital gain of holding the “corporate sector”. We do not explicitly model assets in place and growth options separately on firm balance sheets. More precisely, the capital in stock $K_t$ implicitly contains both assets in place and growth options, and thus the value $Q_t$ contains two components: the value of assets in place and the value...
of growth opportunities. Thus, the stock return can be represented as follows:

\[
1 + R_{k,t+1} = \frac{Y_{t+1} - W_{t+1}L_{c,t+1} + Q_{t+1}(1 - \delta)K_{t+1}}{Q_tK_{t+1}} + \frac{Q_{t+1}I_{t+1} - P_{t+1}\Upsilon(I_{t+1}, K_{t+1})}{Q_tK_{t+1}}.
\]

where \(Q_tK_{t+1}\) is the value of assets at the end of period \(t\). Capital gains from holding corporate shares are included in the net payout \(X_t\). To sort out the consumption component in the return, we introduce the “dividend” of firms:

\[
D_t \equiv Y_t - W_tL_{c,t} - P_t\Upsilon(I_t, K_t).
\]

The stock return can be rewritten in terms of dividends as

\[
1 + R_{k,t+1} = \frac{D_{t+1}}{Q_tK_{t+1}} + \frac{Q_{t+1}K_{t+2}}{Q_tK_{t+1}}.
\]

It is worth mentioning that the quantitative model in full (see the appendix) follows Gertler and Kiyotaki (2010) and Gertler and Karadi (2011), and includes a separate capital goods sector, a sector which produces capital and sells to consumption good firms. Basically, the model, rather than taking the macro view of stock returns, takes the portfolio view, as defined in Larrain and Yogo (2008).\(^8\) Here in this model, consumption good firms are assumed to be short-lived, and the value of their

\(^8\)The difference between two views as regards stock returns is the growth option component. To be more precise, the stock return under the macro view is

\[
1 + R_{k,t+1}^m \equiv \frac{Y_{t+1} - W_{t+1}L_{c,t+1} + Q_{t+1}(1 - \delta)K_{t+1}}{Q_tK_{t+1}} + \frac{Q_{t+1}I_{t+1} - \Upsilon(I_{t+1}, K_{t+1})}{Q_tK_{t+1}}.
\]

and Gertler and Karadi (2011) which takes the portfolio view, the stock return is

\[
1 + R_{k,t+1}^p \equiv \frac{Y_{t+1} - W_{t+1}L_{c,t+1} + Q_{t+1}(1 - \delta)K_{t+1}}{Q_tK_{t+1}}.
\]

Effectively, the gap between \(R_{k,t+1}^m\) and \(R_{k,t+1}^p\) is the net return due to growth options.
outstanding shares is assumed to be equal to the size of capital.

The adjustment cost function has no intertemporal feature in itself; as a result, the intertemporal and dynamic aspects of investment decisions are captured by the forward-looking capital price $Q_t$. The trading in the competitive spot market of capital breaks the direct link between the current investment $I_t$ and the capital stock for next period’s production $K_{t+1}$ for each firm. Thus, the current decision $I_t$ has no effect on the following decisions $I_{t+1}$ through $K_{t+1}$. As a result, the investment decision is not dynamic, and the standard q theory of Hayashi (1982) holds. To see this more clearly, consider the firm’s optimization problem at the end of period $t$:

$$K_{t+1}Q_t \geq \max_{L_{c,t+1}, I_{t+1}} E_t \{ M_{t,t+1}^I [D_{t+1} + Q_{t+1}(1 - \delta)K_{t+1}] \} \quad (13)$$

where $M_{t,t+1}^I$ is the effective intertemporal marginal rate of substitution (IMRS) of financial intermediaries. The equilibrium asset pricing condition depends on inequality (i.e. the supermartingale condition), instead of equality (i.e. the martingale condition), due to the credit constraints. Furthermore, due to the intermediary’s credit constraint, the intermediary’s IMRS can be different from the household’s actual IMRS $M_{t,t+1} \equiv \Lambda_{t+1}/\Lambda_t$. We denote $\Omega_{t,t+1}^I$ as the wedge between the intermediary’s effective IMRS ($M_{t,t+1}^I$) and the household’s IMRS ($M_{t,t+1}$); more precisely,

$$M_{t,t+1}^I = \Omega_{t,t+1}^I M_{t,t+1}. \quad (14)$$

The wedge $\Omega_{t,t+1}^I$ is derived and discussed in Section 3.4. Briefly, the wedge $\Omega_{t,t+1}^I$ characterizes the economic tightness of the credit constraint. The wedge $\Omega_{t,t+1}^I$ becomes larger when the credit constraint is binding more tightly for financial intermediaries. In addition, when the credit constraints are not binding for intermediaries, the equality holds in the asset pricing condition (13). In fact, the Hamilton-Jacob-Bellman (HJB)
equation in (13) can be rewritten in terms of stock returns as

\[ 1 \geq \mathbb{E}_t \left[ M^I_{t,t+1}(1 + R_k,t+1) \right]. \quad (15) \]

Finally, let us characterize the equilibrium relationships between optimal investment, hiring, and consumption. Given the capital price \( Q_{t+1} \) and the price of investment goods \( P_{t+1} \), the problem of optimal investment for firms can be decomposed into a sequence of state-by-state static (intratemporal) optimization problems:

\[ \max_{I_{t+1}} Q_{t+1}I_{t+1} - P_{t+1}Y(I_{t+1}; K_{t+1}). \quad (16) \]

The state-by-state first-order condition with respect to \( I_{t+1} \) gives

\[ \frac{Q_{t+1}}{P_{t+1}} = 1 + \varphi i_{t+1}, \quad \text{where} \quad i_{t+1} \equiv \frac{I_{t+1}}{K_{t+1}}. \quad (17) \]

This is the standard q theory of investment developed by Hayashi (1982), in which the investment decision \( I_{t+1}/K_{t+1} \) is directly linked to the marginal q (marginal value) of the capital.

Similarly, the optimal labor demand can also be derived from the state-by-state (static) optimization problem:

\[ \max_{L_{c,t+1}} A_{t+1}K_{t+1}^{\alpha} L_{c,t+1}^{1-\alpha} - W_{t+1}L_{c,t+1} \quad (18) \]

The first-order condition with respect to \( L_{c,t+1} \) gives

\[ L_{c,t+1} = \left[ (1 - \alpha) \frac{A_{t+1}}{W_{t+1}} \right]^{1/\alpha} K_{t+1}. \quad (19) \]

The consumption goods are non-durable, and thus the market clearing condition
implies the characterization for aggregation consumption goods:

\[ Y_t = D_t + W_t L_{c,t} + W_t L_{i,t}. \]  

(20)

3.3 Investment Goods Sector

There is a continuum of investment good firms which produce investment goods using labor. These firms are identical, and they have the same production function:

\[ \Upsilon_t = A_{i,t} L_{i,t} \]  

(21)

where \( A_{i,t} \) is the productivity of investment goods production, and \( L_{i,t} \) is the labor demand in the investment goods sector. We assume that the scale of the investment goods section is co-integrated with the scale of the consumption goods sector. More precisely, we simply assume that \( A_{i,t} = Z_{i,t} K_t \) where \( Z_{i,t} \) follows a stationary stochastic process. For simplicity, the process \( Z_{i,t} \) is assumed to be constant. It is worth mentioning that \( K_t \) is the total physical capital stock that cannot be internalized by single investment good firms, and thus it is the exogenous scale of the investment goods sector. This guarantees that the balanced growth path is \( A_t K_t^{\alpha} \).

The market clearing condition for labor market requires

\[ L_{c,t} + L_{i,t} = L_t \text{ for all } t. \]  

(22)

All the investment good firms produce and sell investment goods competitively. As a result, they are zero-profit firms. We denote the competitive price of investment goods as \( P_t \).

3.4 Financial Intermediaries

Financial intermediaries borrow funds from households at a risk-free rate, pool the funds with their own net worth and invest the sum in the equity of the representative
consumption good firm. The balance sheet of intermediary $j$ at the end of time $t$ is given by

$$Q_t K_{t+1} S_{j,t} = N_{j,t} + B_{j,t}, \quad (23)$$

where $Q_t$ is the price of the firm’s equity, $S_{j,t}$ is the quantity of equity held by the intermediary, $N_{j,t}$ is the net worth, and $B_{j,t}$ is the deposits raised from households. The intermediary earns a return $R_{k,t+1}$ from the equity investment at time $t + 1$, and must pay the interest, $R_t$, on the deposit. The net worth of the intermediary, therefore, evolves as

$$N_{j,t+1} = (1 + R_{k,t+1}) Q_t K_{t+1} S_{j,t} - (1 + R_t) B_{j,t}$$

$$= (R_{k,t+1} - R_t) Q_t K_{t+1} S_{j,t} + (1 + R_t) N_{j,t}. \quad (24)$$

where $N_{j,t+1}$ is intermediary $j$’s net worth at the end of period $t + 1$.

The intermediaries face a constraint on raising deposits from households. They cannot raise deposits beyond a certain level, which is determined endogenously in the equilibrium. We shall describe this constraint in more detail below. Since bankers own the intermediaries, we use the bankers’ IMRS, which coincides with the IMRS of the representative household, $M_{t,t+1} \equiv \Lambda_{t+1}/\Lambda_t$, to determine the value of assets held by an intermediary according to the cash flows received by the bankers.

The following schematic (Figure 3) is the timing convention for financial intermediaries and firms, to help explain the ordering of these events within the model.

Since an intermediary exits exogenously in each period with probability $1 - \theta$, the value of intermediary $j$’s terminal wealth to its household is given by

$$V_{j,t} = \max \left\{ S_{j,t+r}, B_{j,t+r} \right\}_{r \geq 1} \mathbb{E}_t \left[ \frac{\Lambda_{t+\bar{\tau}_j}}{\Lambda_t} N_{j,t+\bar{\tau}_j} \right], \quad (25)$$

where $\bar{\tau}_j$ is the stochastic stopping time for the financial intermediary $j$ to exit and
pay out the net worth $N_{\tilde{\tau}_j}$ to its banker. Thus, the value of the financial intermediary $j$ can be expressed as weighted average of discounted possible “payouts” (net worth $N_{t+\tau}$):

$$V_{j,t} = \max_{\{s_{j,t+\tau},B_{j,t+\tau}\}_{\tau \geq 1}} \sum_{\tau = 1}^{+\infty} \mathbb{P}(\tilde{\tau}_j = \tau) \mathbb{E}_t \left[ \frac{\Lambda_{t+\tau}}{\Lambda_t} N_{j,t+\tau} \right]$$

$$= \max_{\{s_{j,t+\tau},B_{j,t+\tau}\}_{\tau \geq 0}} \sum_{\tau = 1}^{+\infty} (1 - \theta) \theta^{\tau-1} \mathbb{E}_t \left[ \frac{\Lambda_{t+\tau}}{\Lambda_t} N_{j,t+\tau} \right]$$

$$= \max_{s_{j,t+1},B_{j,t+1}} \mathbb{E}_t \left\{ \frac{\Lambda_{t+1}}{\Lambda_t} \left[ (1 - \theta) N_{j,t+1} + \theta V_{j,t+1} \right] \right\}. \quad (26)$$

The equation (26) is the HJB equation for the value of financial intermediary $j$.

In order to motivate the borrowing constraint faced by financial intermediaries, we introduce a simple moral hazard/costly enforcement problem following Gertler and Kiyotaki (2010) and Gertler and Karadi (2011). We assume that the banker can choose to liquidate the financial intermediation and divert the fraction of available funds, $\lambda_t$, from the value of the financial intermediation.

The borrowing constraint is modeled as follows. At any time $t$, the banker of the intermediary can divert a fraction $\lambda_t$ of the intermediary’s assets for his own benefit, where $\lambda_t$ is an exogenous parameter. The log of margin parameter $\ln \lambda_t$ follows a first-order Markov chain with long-term mean $\overline{\lambda}$, autocorrelation $\rho_{\lambda}$, and
long-term variance $\sigma^2_{\lambda}$. The quantity $1 - \lambda$ measures the steady-state pledgeability of the intermediary’s asset. If the value of the intermediary falls below $\lambda_t Q_t S_{j,t}$, the banker will simply divert the assets and terminate the intermediary. In such a case, households will get a zero gross return from their deposits. In order for the households to have an incentive to deposit cash with the intermediary, the following condition must hold:

$$ V_{j,t} \geq \lambda_t Q_t K_{t+1} S_{j,t}. \tag{27} $$

Because utility functions are homothetic, the optimal portfolios are linear in terms of the net worth. Thus, the value of financial intermediaries is also linear in the net worth and therefore can be characterized as follows:

$$ V_{j,t} = \Omega_t N_{j,t} \tag{28} $$

where $\Omega_t$ is the marginal value of net worth for financial intermediaries. Since each financial intermediary is atomistic, it does not affect the equilibrium. Furthermore, the cross-sectional distribution of the net worths of the intermediaries does not affect the equilibrium either, due to the linearity of the optimal portfolio holdings of intermediaries guaranteed by the homothetic utilities. Instead, the total net worth share $n_t \equiv \frac{N_t}{Q_t K_{t+1}}$ is the only endogenous state variable needed for characterizing the equilibrium, where $N_t \equiv \int_j N_{j,t} d j$ is the aggregate net worth in all intermediaries. Let $S_{p,t} \equiv \int_j S_{j,t} d j$ be the aggregate outstanding shares of firms held by private financial intermediaries. The total supply of $S_{p,t}$ is determined by the credit policy of the government.

We conjecture that $\Omega_t$ only depends on the aggregate exogenous state $z_t \equiv (\xi_t, \lambda_t)$ and the aggregate endogenous state variable $x_t$. The aggregate net worth characterizes the average leverage of financial intermediaries, and thus the incentives for bankers to walk away from their financial intermediaries. As a consequence, it also determines intuitively the tightness of the credit constraint, and in turn, the expected returns.
to financial holdings. The multiplier $\Omega_t$ can be interpreted as the marginal value ("marginal q") of the net worth held by intermediaries.

Following (26) and (27), the portfolio problem of the financial intermediary with credit constraints can be written as,

$$
\Lambda_t \Omega_t N_{j,t} = \max_{S_{j,t+1},B_{j,t+1}} \mathbb{E}_t [\Lambda_{t+1} (1 - \theta + \theta \Omega_{t+1}) N_{j,t+1}] + \mu_{j,t} \Lambda_t (\Omega_t N_{j,t} - \lambda_t S_{j,t} Q_t K_{t+1})
$$

subject to

$$
N_{j,t} = S_{j,t} Q_t K_{t+1} - B_{j,t}, \quad \text{and} \quad N_{j,t+1} = S_{j,t} Q_t K_{t+1} (1 + R_{k,t+1}) - B_{j,t} (1 + R_{f,t})
$$

and $\mu_{j,t} \geq 0$ and $\Omega_t N_{j,t} \geq \lambda_t S_{j,t} Q_t K_{t+1}$. Here, $\mu_{j,t}$ is the Lagrangian multiplier normalized by $\Lambda_t$; it is non-negative, and becomes positive if and only if the credit constraint becomes binding for intermediary $j$. This is the HJB equation which formulates the optimization problem of the intermediaries. Because the intermediaries are the only channel to hold risky assets of the economy, and they face a leverage constraint, the marginal value of cash of the intermediaries $\Lambda_t (1 - \theta + \theta \Omega_t)$ is always larger than the marginal value of cash outside the intermediaries $\Lambda_t$. As a result, the intermediaries are the natural borrowers in the economy. That is, $B_{j,t} \geq 0$ for each intermediary $j$. In aggregate, it holds that $0 < N_t \leq S_p Q_t K_{t+1}$ and thus $0 < n_t \leq 1$.

The first-order condition of substituting between $S_{j,t}$ and $B_{j,t}$ gives

$$
0 \leq \mu_{j,t} \lambda_t \Omega_{t+1} = \mathbb{E}_t [M_{t,t+1} (R_{k,t+1} - R_{f,t})],
$$

where $\Omega_{t,t+1} = \frac{1 - \theta + \theta \Omega_{t+1}}{\Omega_t}$. The wedge $\Omega_{t,t+1}$ is the core component of the so-called "intermediary asset pricing theory" and the effective IMRS of intermediaries is $M_{t,t+1} \equiv M_t \Omega_{t,t+1}$. The condition shows that $\mathbb{E}_t [M_{t,t+1} (R_{k,t+1} - R_{f,t})] > 0$ when the credit constraint is binding. This condition by itself appears to violate the supermartingale condition of self-financed cash flows. Thus, there appears to be a
possible arbitrage opportunity by going long on equity and going short on risk-free bonds. The absence of arbitrage still holds since the intermediary cannot further increase its leverage when its credit constraint is binding.

Plugging (31) into the HJB equation for intermediaries leads to pricing rules for risk-free bonds and firm equity, respectively:

\[
1 \geq 1 - \mu_{j,t} = \mathbb{E}_t \left[ M_{t,t+1}^0 (1 + R_{f,t}) \right], \quad \text{and} \quad (32)
\]

\[
1 \geq 1 - \mu_{j,t} (1 - \lambda_t \Omega_t^{-1}) = \mathbb{E}_t \left[ M_{t,t+1}^0 (1 + R_{k,t+1}) \right]. \quad (33)
\]

The inequality in (33) holds because \( \lambda_t \) is between 0 and 1. The supermartingale conditions hold for equity returns and risk-free rates separately. From (32) and (33), we can see that the Lagrangian multipliers \( \mu_{j,t} \) for intermediaries should be the same. Upon reflection, we turn our focus to a symmetric equilibrium. Denote \( \mu_t \equiv \mu_{j,t} \) for all \( j \). In the symmetric equilibrium, each intermediary chooses the shares of equity to hold proportionally to its net worth in the following sense:

\[
S_{j,t} Q_t K_{t+1} = s_t N_{j,t}, \quad (34)
\]

for all \( j \) and \( s_t \) only depending on aggregate state variables. Given equilibrium asset prices and returns, the optimal holding \( s_t \) can be characterized by only the market-clearing condition in the equity market. The supply to private financial intermediaries is \( S_{p,t} = 1 - S_{g,t} \) and the market-clearing condition is

\[
S_{p,t} Q_t K_{t+1} = s_t N_t. \quad (35)
\]

Thus, the optimal holding increases in the total supply of equity \( S_{p,t} \), and decreases in the total net worth of intermediaries \( (x_t) \):

\[
s_t = \frac{S_{p,t}}{n_t}. \quad (36)
\]
3.5 Net Worth Evolution

After integrating the dynamic equations in (24) over all intermediaries and accounting for the net fund transfer (1), the aggregate net worth evolves as

\[ N_{t+1} = (R_{k,t+1} - R_{f,t})Q_tK_{t+1}S_{p,t} + (1 + R_{f,t})N_t - \Pi_t \]

\[ = [(R_{k,t+1} - R_{f,t})S_{p,t} + \mathbb{N}] Q_tK_{t+1} + (R_{f,t} + \theta)N_t. \]

Thus, the net worth share of intermediaries evolves as

\[ \frac{n_{t+1}}{n_t} = \left[ (R_{k,t+1} - R_{f,t})S_t + \mathbb{N}/n_t + (R_{f,t} + \theta) \right] / G_{k,t+1} \] (37)

where \( G_{k,t+1} \equiv \frac{Q_{t+1}K_{t+2}}{Q_tK_{t+1}} \) is the total capital gain of equity.

Let \( \mu^* \) be the upper bound of the dividend-price ratio of stocks in the frictionless economy. Then the net worth share \( n_t \) is always less than 1 when

\[ \mu^* < (1 - \theta) - \mathbb{N}. \] (38)

The specifications of \( \mu^* \) can be found in Appendix A. In the case when (38) holds, there exists \( n_t \in (0, \lambda_tS_{p,t}) \) characterizing the constraint-binding boundary such that

\[ \Omega_t = \left\{ \begin{array}{ll} \frac{\lambda_tS_{p,t}}{n_t}, & \text{when } n_t \in (0, n_t]; \\ \Omega(n_t, \lambda_t) > \max\left\{ \frac{\lambda_tS_{p,t}}{n_t}, 1 \right\}, & \text{when } n_t \in (n_t, 1). \end{array} \right. \]

The net worth share \( n_t \) never reaches the limit 1 since there is no efficiency loss attached to the intermediary net worth. Solving the equilibrium is effectively the functional form of \( \Omega(n_t, \lambda_t) \).
3.6 Government Policies

The ultimate goal of our model is to analyze the effectiveness of unconventional monetary policies in fighting financial crises and their destructive impact on the macroeconomy as a whole. Importantly, we demonstrate that the global solution of our model, which captures its nonlinear features, is vital to guarantee the proper and useful analysis of its policy implications.

According to Section 13.3 of the Federal Reserve Act, the Fed is allowed to take risky positions through making loans in the private sector (provided that they are not unduly so), under “unusual and exigent circumstances.” This legislation basically makes the Fed the lender of last resort of the economy. Meanwhile, the Treasury, the Fed, the FDIC, and the bailout bills passed by Congress together took unconventional policy measures, including equity injection into the private sector, asset purchases from distressed banks, the lifting of caps on deposit insurance for certain bank accounts, and lending guarantees for certain types of bank loans. All these policies and interventions were intentionally designed to encourage firms to bring in private capital. For instance, it was intentionally designed that firms returning capital to the government by certain dates would get better terms for the government’s stake. The central plank of all these unconventional measures was to attract private capital.

These different measures work together in practice, and were intentionally designed to complement each other. As a result, it is unrealistic to discuss them individually in a unified framework. Given their primary goal and common ideas, however, our model adopts a single abstract unconventional policy as a modeling device, yet one still relevant enough to serve an illustrative purpose. We assume the government is willing to buy the shares of the firm directly to facilitate lending. Such policies were studied by Gertler and Kiyotaki (2010), and Gertler and Karadi (2011). This captures the unconventional policy of purchasing risky, privately managed, non-government assets, implemented in the U.S., the U.K. and the eurozone in the wake of the financial crisis. The U.S. Federal Reserve’s program of buying $600 million of mortgage-backed securities in 2008-09 (QE-1) and the European Central Bank’s Covered Bond Purchase Programs (CBPPs) for buying private sector debt are examples of such policies. Our
intention of appealing to such a simple form of unconventional monetary policy (or a credit policy) is to develop a baseline model for analysis.

More precisely, in our model, the government buys a fraction $S_{g,t}$ of the total outstanding shares of firms (normalized to one), so that

$$Q_t K_{t+1} = S_{p,t} Q_t K_{t+1} + S_{g,t} Q_t K_{t+1}, \quad (39)$$

where $S_{p,t} = \int S_{j,t} d\tilde{j}$ is the total share of equity held privately, and the share of government-held equity is $S_{g,t} = 1 - S_{p,t}$. To conduct the credit policy, the government issues government debt to households that pay the risk-free rate $R_{f,t}$ and then lends the funds to firms or purchases the equity stakes of firms with returns $R_{k,t+1}$. The government credit has an efficiency cost of $\tau > 0$ units per unit of credit supplied. This deadweight loss may reflect the government’s fundraising costs or its investment search costs.

We then introduce the key assumption which makes the government’s balance sheet, and thus the credit policy, non-neutral as regards its macroeconomic implications. This is the only special feature of government intermediation in our model. A general discussion about the special characteristics that make a government’s balance sheet relevant can be found in Section 4.1. More precisely, government intermediation is not financially constrained in our model, unlike private financial intermediation. This can be justified by the assumption that the government always honors its debt, and thus incurs no agency problems between it and its household creditors.

We define the total leverage ratio $\phi_{c,t}$ as follows

$$Q_t K_{t+1} = \phi_{c,t} N_t. \quad (40)$$

The leverage ratio, $\phi_{c,t}$, is the leverage ratio for total intermediated funds, public as well as private, and has the following relation with the private leverage ratio,
\[
\phi_t \equiv S_{p,t} Q_t K_{t+1}/N_t, 	ext{ and the intensity of government credit intervention, } S_{g,t},
\]

\[
\phi_{c,t} = \frac{\phi_t}{1 - S_{g,t}}. 
\]  

(41)

The government issues government bonds \( B_{g,t} = S_{g,t} Q_t K_{t+1} \) to fund the purchase of these shares. From this activity, the government thus earns an amount \( (R_{k,t+1} - R_{f,t}) B_{g,t} \) every period.

We assume that at the onset of a crisis, which is defined loosely to mean a period when the log risk premium \( \Xi_t \equiv E_t \left[ \ln \left( 1 + R_{k,t+1} \right) \right] - \ln \left( 1 + R_{f,t} \right) \) rises sharply and becomes much higher than the frictionless benchmark \( \Xi^* \equiv E_t \left[ \ln \left( 1 + R^*_k \right) \right] - \ln \left( 1 + R^*_f \right) \), the government injects credit in response to movements in risk premia. Similar to a standard Lucas-tree economy, the log risk premium is

\[
\Xi^* \approx \gamma \sigma^2_{a} - \frac{1}{2} \sigma^2_{a},
\]

where \( \frac{1}{2} \sigma^2_{a} \) is Jensen’s term for the log return. The frictionless benchmark is described in Appendix A. We consider the credit policy that follows the rule for \( S_{g,t} = 1 - S_{p,t} \):

\[
S_{p,t} = \frac{1}{1 + \nu_g \times (\Xi_t - \Xi^*)},
\]  

(42)

where the sensitivity parameter, \( \nu_g \), is positive. According to (42), the government expands credit as the risk premium gap increases. Our specification is a global version of the credit policy considered by Gertler and Kiyotaki (2010) and Gertler and Karadi (2011). In the local-linear approximation when \( \Xi_t - \Xi^* \) is small,

\[
S_{g,t} = 1 - S_{p,t} \approx \nu_g \times (\Xi_t - \Xi^*). 
\]  

(43)

The rationale behind this policy specification is as follows. In the absence of financial friction that prevents the financial intermediaries from leveraging too much, the equilibrium outcome is efficient. The inefficiency arises due to the inability of
households to buy the risky assets directly, and to the limit on the leverage of their financial managers. This inefficiency manifests itself in the form of large risk premia, since the financial intermediaries must be compensated adequately in the absence of high leverage. The government does not intervene when the risk premium is at its steady-state level, but it does intervene when the premium rises to increasingly inefficient levels beyond it.

We shall show that the global solution of this nonlinear system allows for a state-dependent sensitivity coefficient. For example, we can specify a policy rule as follows:

\[ \nu_g = \nu_{g,0} + \nu_{g,1} \times \left( \frac{1}{n_t} - 1 \right), \]  

(44)

with \( \nu_{g,0} \geq 0 \) and \( \nu_{g,1} \geq 0 \). The idea is that it should be better for the government to conduct more aggressive credit policy (i.e., the sensitivity \( \nu_g \) is larger) when the financial system is already more fragile (i.e., \( n_t \) is smaller).

From (41), it is clear that when the private leverage ratio \( \phi_t \) is kept fixed, the expanding credit policy \( S_{g,t} \) increases the total leverage of intermediation, i.e., \( \phi_{c,t} \) rises. This captures the idea that the government’s balance sheet acts as an intermediary to channel household funds to the asset market when the financial intermediaries are constrained. The government’s intermediation prevents asset prices from becoming overly distressed when this is caused by the inefficiency of financial intermediaries after a sequence of negative shocks.

### 3.7 Resource and Government Budget Constraints

The resource constraint for the final good in our model is given by

\[ Y_t = C_t + G_t + \tau S_{g,t} Q_t K_{t+1}. \]  

(45)
The government spends a fraction $\bar{g}$ of output $Y_t$ in period $t$, where $\bar{g}$ is an exogenously specified constant. That is,

$$G_t = \bar{g}Y_t. \hspace{2cm} (46)$$

In addition to funding government expenditure, the government also needs to fund the central bank’s purchase of shares by issuing purchasing bonds worth $B_{g,t} = S_{g,t}Q_tK_{t+1}$ and the efficiency loss by taxes. Its revenues include taxes, $T_t$, and the government’s income from intermediation, $S_{g,t-1}Q_{t-1}K_t(R_{k,t} - R_{f,t-1})$. Thus, the government budget constraint is

$$G_t + (1 + \tau)S_{g,t}Q_tK_{t+1} = T_t + S_{g,t-1}Q_{t-1}K_t(R_{k,t} - R_{f,t-1}) + B_{g,t}. \hspace{2cm} (47)$$

Since the taxation $T_t$ effectively takes up any slack that shows up on the government balance sheet, and given the existence of representative agents in the economy, the intertemporal budget constraint of the representative household and the intertemporal budget constraint of the government can be combined with taxes left out. Intuitively, then, by Walras’ Law, both budget constraints are redundant in determining the equilibria. However, this is very different from saying that the size and composition of the government balance sheet are irrelevant for pinning down the equilibrium under efficient financial market conditions in the sense of Wallace (1981). This is simply because not all investors can purchase an arbitrary amount of the same assets at the same market prices as the government in this model. Put more precisely, unlike private financial intermediation, government intermediation is not constrained by the balance sheet.

### 3.8 Quantitative Analysis

Our model can be used to understand the response of the economy to various shocks. We highlight the role of nonlinear dynamics of risk premia in determining the intensity of the credit policy. Furthermore, we show that it requires global solutions to properly
characterize these nonlinear dynamics and to analyze the effectiveness of policies.

We use a calibrated version of the model, basing our parameter choices mainly on those in Gertler and Karadi (2011) and Gourio (2012), and the estimated dynamic parameters in Smets and Wouters (2007). Further discussion on parameter calibration can be found in Appendix D. Using the dynamic parameters estimated in the full model in the appendix, we calibrate them here for illustrative purposes. In our numerical analysis, the exogenous autoregressive processes are discretized into homogeneous Markov chains according to the method proposed by Rouwenhorst (1995). The calibrated parameters are summarized in Table 2. The parameter choices are basically close to those picked in Gertler and Karadi (2011).

Equilibrium We first demonstrate the equilibrium policy functions of the endogenous variables. Here, we assume the baseline parameter specifications, except that of the margin of financial intermediaries, stay constant. That is, the volatility of margin is zero: \( \sigma_\lambda = 0 \). The nonlinearity of this equilibrium is severe, even having kinks over the endogenous state space. We compare the equilibrium outcomes with the frictionless economy equilibrium.

In Figure 4, we show the equilibrium financial variables. The red dashed curves in Panels A – E are endogenous variables in equilibrium in a frictionless economy, whereas the blue solid curves characterize the equilibrium of an economy with financial friction. The only deviation of the economy with financial friction from the frictionless economy is the leverage constraint faced by intermediaries. The equilibrium of the frictionless economy does not depend on the net worth of financial intermediaries, as the equilibrium policy functions are all constant across different net worth share \( n_t \). In contrast, the equilibrium of the economy with financial friction relies heavily upon intermediary net worth, i.e. the condition of how well the financial intermediaries are capitalized. Importantly, the dependence is largely nonlinear. The intermediary net worth, as an endogenous state variable, is endogenously driven by exogenous shocks hitting the economy, and in turn, it poses endogenous risks to the economy.

Panel D of Figure 4 shows the Lagrangian multiplier of the leverage constraints faced by intermediaries. In equilibrium, the constraint becomes binding if and only if
Table 2: Baseline Parameters (Annualized)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household preference</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount rate</td>
<td>$\beta$</td>
<td>0.99</td>
<td>Standard</td>
</tr>
<tr>
<td>Relative risk aversion</td>
<td>$\gamma$</td>
<td>6</td>
<td>Standard</td>
</tr>
<tr>
<td>Total labor supply</td>
<td>$\bar{L}$</td>
<td>1</td>
<td>Standard</td>
</tr>
<tr>
<td><strong>Financial intermediaries</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steady-state fraction of divertible capital</td>
<td>$\lambda$</td>
<td>0.381</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td>Proportional transfer to new bankers</td>
<td>$\xi$</td>
<td>0.002</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td>Survival rate of bankers</td>
<td>$\theta$</td>
<td>0.5</td>
<td>Non-degenerate condition</td>
</tr>
<tr>
<td><strong>Consumption-good firms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective capital share</td>
<td>$\alpha$</td>
<td>0.33</td>
<td>Standard</td>
</tr>
<tr>
<td>Depreciation rate</td>
<td>$\delta$</td>
<td>0.06</td>
<td>Standard</td>
</tr>
<tr>
<td>Adjustment cost coefficient</td>
<td>$\vartheta$</td>
<td>5</td>
<td>Standard</td>
</tr>
<tr>
<td><strong>Government policies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government expenditure ratio</td>
<td>$\bar{y}$</td>
<td>20%</td>
<td>Standard</td>
</tr>
<tr>
<td>Government efficiency loss</td>
<td>$\tau$</td>
<td>10%</td>
<td>Calibration</td>
</tr>
<tr>
<td>Sensitivity coefficient</td>
<td>$\nu_{g,0}$</td>
<td>5</td>
<td>Calibration</td>
</tr>
<tr>
<td>Sensitivity coefficient</td>
<td>$\nu_{g,1}$</td>
<td>0</td>
<td>Calibration</td>
</tr>
<tr>
<td><strong>Dynamic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volatility of TFP</td>
<td>$\sigma_a$</td>
<td>0.04</td>
<td>Standard</td>
</tr>
<tr>
<td>Persistence of margin</td>
<td>$\rho_{\lambda}$</td>
<td>0.66</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td>Volatility of margin</td>
<td>$\sigma_{\lambda}$</td>
<td>0.267</td>
<td>Gertler and Karadi (2011)</td>
</tr>
</tbody>
</table>

Note: All parameters are standard except the credit-policy related parameters $\tau$, $\nu_{g,0}$, and $\nu_{g,1}$. We pick $\nu_{g,0}$ and $\tau$ here to provide reasonable private holding shares of risky assets and average risk premia. The parameter $\nu_{g,1}$ is chosen to be zero in the baseline calibration, though we emphasize that the non-zero $\nu_{g,1}$ is important as part of the optimal policy given the nonlinear dynamic of risk premia and the economy. The closest comparison in the literature is Gertler and Karadi (2011); however, their model is solved using log-linearization techniques. The local approximation significantly understates the magnitude and the volatility of the conditional risk premium, even though the model is capable of generating both quantitatively. The implication of biased asset pricing makes quantitative discussion of unconventional monetary policies itself biased. For example, the suppressed risk premium and its nonlinear dynamics require extremely sensitive unconventional monetary policy in order to have a quantitatively significant stabilization effect on the aggregate quantities. In Gertler and Karadi (2011), the credit policy sensitivity parameter $\nu_{g,0}$ is chosen about 100 and the cost parameter of the intervention $\tau$ is chosen at an extremely small value 0.1%. 

42
the intermediary net worth becomes less than a critical value $n$ that is about 0.4 under this calibration. Panel F displays the stationary distribution of intermediary net worth. The economy fluctuates mainly between 0.1 and 0.3. The median of the distribution is about 0.19. Comparing Panels D and F, we can see that the leverage constraint is almost always binding in equilibrium. Comparing Panels D and C, we can see that the marginal value of intermediary net worth $\Omega_t$ is higher than the frictionless benchmark $\Omega^* = 1$, even when $n_t$ is far into the region where the leverage constraint is not binding for intermediaries. This is the expectation effect of financial constraints. From Panels A – E, it can be seen that the equilibrium of the frictional economy converges to that of the frictionless economy as the intermediary net worth approaches 1.

The equilibrium risk premium $\Xi_t$ is shown in Panel A of Figure 4. Its unconditional expectation is about 10% each year, and it skyrockets as the intermediary net worth declines. Panel B shows the equity price $q_t$, which decreases rapidly as the intermediary net worth declines. The marginal value $\Omega_t$, as the asset pricing wedge between intermediaries and households, increases quickly as $n_t$ declines. The constraint of the Lagrangian multiplier $\mu_t$ shown in Panel D is intuitive: when the net worth declines, the leverage constraint becomes more binding, and thus the Lagrangian multiplier becomes larger. We can see, in Panel E, that the equilibrium government credit intervention is significant, even the intensity parameter $\nu_g^0$ is 5, much less than the choice of Gertler and Karadi (2011). Again, it shows the importance of nonlinear global solutions for the quantitative discussion on models of unconventional monetary policies. On average, the government holds about 30% of the risky assets in the economy.

We now turn to the equilibrium macroeconomic quantities demonstrated in Figure 5. Panel A illustrates how the investment rate $i_t$ varies when the intermediary net worth fluctuates. The investment rate decreases rapidly as $n_t$ declines. The average investment rate at equilibrium is about 8.5%. Panels B and C shows wage income and consumption, respectively. Consistent with intuition, the dynamics of wage income and consumption are much more smooth relative to investment and other financial variables. More precisely, in this model, the wage income declines from about 0.7 to about 0.69, and the consumption declines from about 0.71 to about 0.7, as $n_t$ declines.
Figure 4: Financial Variables at Equilibrium.

A. Risk Premium

B. Equity Price

C. Marginal Value

D. Constraint Multiplier

E. Government Intervention

F. Stationary Distribution

Note: Panels A – E show the equilibrium variables and Panel F displays the stationary distribution at equilibrium. The red dashed curves characterize the equilibrium policy functions for the frictionless economy. The blue solid curves are equilibrium policy functions of the frictional economy with $\lambda = 0.381$. The economy mainly fluctuates between 0.1 and 0.3 in terms of net worth share $n_t$ at equilibrium.
Financial Shocks  We now describe the equilibrium of the economy with time-varying leverage constraints for intermediaries. Here, we assume that the log of margin parameter $\ln \lambda_t$ follows a first-order Markov chain with long-term mean $\bar{\lambda}$, autocorrelation $\rho_\lambda$, and long-term variance $\sigma^2_\lambda$. Those parameters’ values are chosen according to the baseline calibration in Table 2. We appeal to the discretization scheme proposed by Rouwenhorst (1995) to consider a two-state discretized process for $\lambda_t$. In particular, the margin parameter $\lambda_t$ takes two values $\lambda_H$ and $\lambda_L$ with $\lambda_H > \lambda_L > 0$. The states of higher margin $\lambda_H$ characterize the periods in which intermediaries are stressed by tightened leverage constraints. More precisely, according to the baseline calibration in Table 2 and the discretization method, we take $\lambda_L$ to be 0.2917 and $\lambda_H$ to be 0.4976. The annualized transition matrix of the two states is

$$
\begin{bmatrix}
\lambda_L & \lambda_H \\
\lambda_L & 0.8 & 0.2 \\
\lambda_H & 0.2 & 0.8 \\
\end{bmatrix}
$$

In Figure 6, we describe the equilibrium financial variables. The dashed blue curves are endogenous variables when $\lambda = \lambda_L$, whereas the solid blue curves characterize the equilibrium of the economy when $\lambda = \lambda_H$. The equilibrium of the economy relies upon the intermediary net worth. Moreover, the risk premium $\Xi_t$, the asset price $q_t$, the marginal value of net worth $\Omega_t$, the constraint multiplier $\mu_t$, and the government intervention scale $S_g,t$ are dramatically affected by financial shocks (i.e., shocks on $\lambda_t$). The effects on those financial variables become larger as the intermediary net worth declines. Intuitively speaking, when the financial system is already fragile (i.e., when intermediaries are not well capitalized), the adverse impact of financial shocks becomes especially substantial.

Panel D of Figure 6 shows the Lagrangian multiplier of the leverage constraints faced by intermediaries. The endogenous thresholds of binding leverage constraints are different in different states of $\lambda_t$. When $\lambda_t = \lambda_L$, the threshold of financial binding...
Figure 5: Quantity Variables at Equilibrium.

A. Investment Rate

B. Wage Income

C. Consumption

D. Stationary Distribution

Note: Panels A – C show the equilibrium variables and Panel F displays the stationary distribution of equilibrium. The red dashed curves characterize the equilibrium policy functions for the frictionless economy. The blue solid curves are equilibrium policy functions of the economy with financial friction and $\lambda = 0.381$. The economy is mainly fluctuating between 0.1 and 0.3 in terms of net worth share $n_t$ at equilibrium.
is about \( n_L = 0.14 \), whereas when \( \lambda_t = \lambda_H \), the threshold of financial bidding is \( n_H = 0.34 \). It is intuitive that financial intermediaries become more fragile (i.e., leverage constraints become more binding) when the margin is higher at \( \lambda_H \).

Panel F displays the stationary distribution of intermediary net worth. The economy fluctuates mainly between 0 and 0.3. Compared to Panel F of Figure 4, we can see that the financial shocks make the stationary distribution spread wider. The median of the distribution is about 0.13. From Panels D and F, we can see that the leverage constraint is almost always binding when \( \lambda_t = \lambda_H \) in equilibrium, while interestingly enough, the leverage constraint occasionally becomes binding when \( \lambda_t = \lambda_L \). From Panels D and C, we can see that the marginal value of intermediary net worth \( \Omega_t \) is higher than the frictionless benchmark \( \Omega^* = 1 \), even when \( \lambda_t = \lambda_L \) and the leverage constraint is not binding at all. This is a manifestation of the expectation effect of financial constraints. From Panels A – E, it can be seen that, no matter if \( \lambda_t \) is equal to \( \lambda_H \) or \( \lambda_L \), the equilibrium of the frictional economy converges to that of the frictionless economy as the intermediary net worth approaches 1.

The equilibrium risk premium \( \Xi_t \) is shown in Panel A of Figure 6. The conditional average risk premium given \( \lambda_t = \lambda_L \) is about 4% per year, and it skyrockets to about 13% per year as the economy is hit by an adverse financial shock (i.e. \( \lambda_t \) jumps to \( \lambda_H \) from \( \lambda_L \)). Conversely, Panel B shows that the equity price \( q_t \) suffers from a sizable drop as \( \lambda_t \) increases to \( \lambda_H \). The marginal value \( \Omega_t \), shown in Panel C, also increases as \( \lambda_t \) jumps to \( \lambda_H \). The Lagrangian multiplier constraint \( \mu_t \) shown in Panel D indicates that the financial system becomes much more constrained when the margin \( \lambda_t \) rises. Finally, in Panel E, we can see that the equilibrium government credit intervention is very sensitive to the adverse financial shock. On average, the government holds about 17% of the risky assets in the economy when \( \lambda_t = \lambda_L \), while it holds about 49% when \( \lambda_t = \lambda_H \).

We now turn to the equilibrium macroeconomic quantity demonstrated in Figure 7. Panel A illustrates how investment rate \( i_t \) varies when financial margin \( \lambda_t \) and intermediary net worth \( n_t \) fluctuates. The investment rate drops rapidly as \( \lambda_t \) increases. The conditional average investment rate is about 11.6% in the state of \( \lambda_t = \lambda_L \), while it becomes about 8.3% in the state of \( \lambda_t = \lambda_H \). Panels B and C shows wage income and
Figure 6: Financial Variables at Equilibrium of Economy with Financial Shocks.

Panels A – E show the equilibrium variables and Panel F displays the stationary distribution at equilibrium. The dashed blue curves characterize the equilibrium policy functions of states with lower margin $\lambda_L$, and the solid blue curves are equilibrium policy functions of states with higher margin $\lambda_H$. The economy mainly fluctuates between 0 and 0.3 along the dimension of net worth share $n_t$ at equilibrium.
consumption, respectively. They are also negatively affected by the financial shock. However, their declines are much smaller than that of the investment rate.

**Government Intervention and the Fragility of Intermediaries** From the equilibrium cases in Figure 6 and Figure 7, showing the economy with time-varying leverage constraints for its financial intermediaries, we can see that (rare) adverse financial shocks can have dramatic and devastating effects on financial markets and the real economy. The negative effects become especially severe when intermediary net worth level \( n_t \) is already low. This is one of the key insights of this class of models: the fragility of financial intermediaries is largely captured by their level of capitalization. When the intermediaries are poorly capitalized (when net worth \( n_t \) is low in the model), the intermediaries are more fragile, and an exogenous financial shock can cause tragic financial turmoil and drag the economy into a recession. Therefore, government intervention should not only battle financial shocks, but also combat the fragility of the financial system. It is important when considering financial fragility that this model has global solutions, so it can be used as a laboratory for the analysis of simple nonlinear government intervention.

Here, we extend the credit policies considered in the previous analysis, where the intensity of intervention depends only on the observed risk premium, and the sensitivity \( \nu_g \) is constant, no matter the current fragile state of the financial system. Here, we allow the sensitivity \( \nu_g \) to be state-dependent, depending on the intermediary net worth level \( n_t \) as described in (44). Specifically, we choose \( \nu_{g,0} = 5 \) and \( \nu_{g,1} = 2 \) and keep the other parameter calibrations unchanged.

In Figure 8, we describe the equilibrium financial variables. The dashed blue curves are endogenous variables when \( \lambda = \lambda_L \), whereas the solid blue curves characterize the equilibrium of the economy when \( \lambda = \lambda_H \).

Panel D of Figure 8 shows the Lagrangian multiplier of the leverage constraint faced by financial intermediaries. It is obvious that the leverage constraint binds much less compared to the case \( \nu_{g,1} = 0 \) in Figure 6. Interestingly, with nonlinear government intervention, financial intermediaries are almost unbound in the state \( \lambda = \lambda_L \). This is because the intervention can also be intensive in the normal state \( \lambda = \lambda_L \) when
Figure 7: Quantity Variables at Equilibrium of Economy with Financial Shocks.

A. Investment Rate

B. Wage Income

C. Consumption

D. Stationary Distribution

Note: Panels A – C show the equilibrium variables and Panel F displays the stationary distribution of equilibrium. The blue dashed curves characterize the equilibrium policy functions of states with lower margin $\lambda_L$, and the blue solid curves are equilibrium policy functions of states with higher margin $\lambda_H$. The economy is mainly fluctuating between 0 and 0.3 along the dimension of net worth share $n_t$ in equilibrium.
intermediaries are fragile ($n_t$ is low), and it decreases the demand of risk sharing from intermediaries when $n_t$ is low. As a result of the relaxed leverage constraints, we can see in Panel C of Figure 8 that the marginal value of the intermediary net worth becomes much smaller even when $n_t$ is low, compared to Figure 6. The equilibrium government intervention can be seen clearly by comparing Panel E of Figure 8 to that of Figure 6: while the intensity of intervention in the state $\lambda = \lambda_H$ is similar, the government intervention in the state $\lambda = \lambda_L$ becomes more intense under the nonlinear credit policy. The nonlinear policy (the case of $\nu_{g,1} = 2$) makes the risk premium more stable relative to the case of linear policy (the case of $\nu_{g,0} = 0$). This can be seen by comparing Panel A of Figure 8 to that of Figure 6.

We now turn to the equilibrium macroeconomic quantities of Figure 9. Comparing Panels A and B of Figure 9 to those of Figure 7, it can be seen that the levels of investment and wages are higher under the nonlinear policy. This is due to the greater stability guaranteed by the nonlinear policy. However, Panel C of Figure 9 shows that the consumption level is lower than that of Figure 7. This is the result of the higher efficiency costs incurred by the more intensive government intervention suggested by the nonlinear policy. As can be seen clearly here, there is a crucial tradeoff between stability and efficiency in an optimal (nonlinear) government intervention – that is to say, through unconventional credit policies. To find more reliable quantitative answers to this important question will be a challenge. It will require a richer framework, and global solutions, to chart out a promising research agenda that will have a huge influence both in academia and among practicing monetary authorities.

4 Challenges and Opportunities for DSGE Modeling

There are a number of model features and quantitative methodologies that are crucial to our understanding of the financial market and the macroeconomy that the standard New Keynesian DSGE models of the current generation (such as the simple canonical example in the previous section) simply do not incorporate. The recent crisis and
Figure 8: Financial Variables at Equilibrium of Economy with Financial Shocks and Nonlinear Government Intervention.

Note: Panels A – E show the equilibrium variables and Panel F displays the stationary distribution at equilibrium. The dashed blue curves characterize the equilibrium policy functions of states with lower margin $\lambda_L$, and the solid blue curves are equilibrium policy functions of states with higher margin $\lambda_H$. The economy mainly fluctuates between 0 and 0.2 along the dimension of net worth share $n_t$ at equilibrium. The y-axis scale of the plots is kept the same as the corresponding ones in Figure 6 for comparison purposes.
Figure 9: Quantity Variables at Equilibrium of Economy with Financial Shocks and Nonlinear Government Intervention.

Note: Panels A – C show the equilibrium variables and Panel F displays the stationary distribution at equilibrium. The dashed blue curves characterize the equilibrium policy functions of states with lower margin $\lambda_L$, and the solid blue curves are equilibrium policy functions of states with higher margin $\lambda_H$. The economy mainly fluctuates between 0 and 0.2 along the dimension of net worth share $n_t$ at equilibrium. The y-axis scale of the plots is kept the same as the corresponding ones in Figure 6 for comparison purposes.
recession have put many of these missing pieces into the spotlight. It is evident that these missing pieces have a first-order impact on the economy as whole, and have profoundly affected how governments have conducted their policies. In this section, we discuss these major missing components and methodological challenges. We hope to shed some light on the path along which researchers may advance current New Keynesian DSGE models to the next generation, one which will be more useful to monetary authorities. The issues of the current generation of New Keynesian DSGE models and the challenges of future improvements to these models are fundamentally and deeply interconnected. Therefore, in order to truly improve these models in one dimension, we may need to simultaneously tackle all the others to some degree.

4.1 Government Balance Sheet Irrelevance

Classic monetary macroeconomic theory, as used in modern macroeconomic models, taught in graduate school textbooks, and employed by major central banks all over the world, starts from the simple national income accounting identity

\[ Y = C + I + G + X, \]

where \( Y \) is the aggregate output of the economy, \( C \) is the aggregate household consumption, \( I \) is the aggregate investment, \( G \) is the government spending, and \( X \) is the net export. The only role played by government in this model is through government spending, the dynamics of which are specified exogenously. In other words, the effects of the government balance sheet and any intertemporal budget constraint on government are totally abstracted out of the analysis. This omission is not just some reduced-form modeling trick to simplify the analysis of monetary policy. In fact, the omission of the government balance sheet is completely justifiable in terms of both legislative practice and fundamental economic principles.

In legislative practice, monetary policy decisions by law are independent of government, i.e., the fiscal anchor is independent of the monetary anchor, although the monetary anchor and the fiscal anchor inevitably interact. These monetary-fiscal
interactions mainly include: (1) interest rate changes, leading to changes in the “interest expense” item in the government budget, thereby leading to changes in the growth rate of government debt, which of course depends on whether taxes and expenditures react to the original changes in interest rate, and if so, by how much; (2) central banks holding earning assets (usually bonds) to back the currency they issue (which does not earn interest), giving the banks a stream of revenue (so-called “seigniorage”), which they generally turn over to the treasury (i.e., the government); and (3) increased inflation reducing the real burden of the stream of future payments specified in long-term government bonds. As emphasized by Sims (2008), monetary independence could be sustained on a fair level because, up to 2007, there had been little risk on the Fed’s balance sheet. Its liabilities were mainly currency outstanding and reserve balances, and its assets were mainly short-term U.S. government debt. More precisely, before 2007, there was little risk on the Fed’s balance sheet because (1) while exchange rate movements or inflation can change the value of the dollar, since assets and liabilities were all in dollars, there was no effect on net worth; (2) changes in long-term interest rates can change the market value of long bonds, but since the assets were mainly short term, this effect had a minor effect; and (3) the U.S. government was extremely unlikely to default outright on its nominal bonds, in part because under conditions where this might be an attractive possibility, inflation to reduce the value of the debt would be easier and more efficient. Therefore, it is fair to assume that the government balance sheet played a very limited role in the Fed’s monetary policy decisions before 2007.

The justification from fundamental economic principles is more involved. The efficiency of the financial market is the key. More precisely, the financial market needs to be efficient enough so that the following assumptions are satisfied:

1. assets are valued only for their pecuniary returns. This means that assets only fail to be perfect substitutes from the standpoint of investors due to their different risk characteristics, but not due to any other reasons.

9However, the story for Europe is very different. The ECB’s assets and liabilities are denominated in different currencies because they have large non-euro reserves. Many other major economies in the world also face the same situation. In addition, it is unique to the ECB that there is no single fiscal counterpart to pressure them over seigniorage or interest expense.
(2) all investors can purchase arbitrary amounts of the same assets at the same market prices as the government.

(3) the government conducts a Ricardian fiscal policy, indicating that the government budget constraint must be satisfied for all realizations of the price level (see, e.g. Woodford, 1995). In the presence of multiple equilibria, a non-Ricardian spending or tax policy can trim the set of monetary policy-derived equilibria, as we discuss in Section 4.2.

Under these assumptions, the government balance sheet has no impact on the equilibrium of the economy, and hence neither does the open-market purchase of securities by the government. Thus, monetary policy models need only assume a government printing press which creates additional “money” at a greater or lesser rate, which is then put in the hands of private parties, perhaps by dropping it from helicopters. These assumptions lie at the heart of the classic monetarist view: the amount of monetary liabilities by the central bank matters for macroeconomic equilibrium, but it does not matter at all what kinds of assets might back those liabilities on the other side of the balance sheet, or how the base money gets to be in circulation.

The irrelevance or neutrality of the government balance sheet in determining market equilibrium is essentially the theoretical macroeconomic analog to the Modigliani-Miller Theorem in corporate finance, as noted in the seminal work by Wallace (1981). In that paper, the author emphasized that this result of irrelevancy implies that both the size and the composition of the central bank or government balance sheet should be irrelevant for market equilibrium in a world with frictionless financial markets (or more precisely, a world in which the above postulates hold). Similar to Wallace (1981), Eggertsson and Woodford (2003) derive a neutrality result in a New Keynesian model. In their framework, which assumes Ricardian fiscal policies, the portfolio of assets held by the central bank is irrelevant towards determining the set of equilibrium output and price levels. This does not, however, mean that monetary policy is irrelevant in such a world, as is sometimes thought; it simply means that monetary policy cannot be implemented through open-market operations whenever the neutrality result holds and the fiat money has zero value (see, e.g. Kiyotaki and Moore, 2012). Control of the
short-term nominal interest rate by the central bank remains possible in a frictionless environment. The central bank is still free to determine the nominal interest rate on overnight balances at the central bank. This interest rate must then be linked in equilibrium to other short-term interest rates, through arbitrage relations; and hence the central bank can determine the level of short-term nominal interest rates in general. Moreover, the central bank’s adjustment of nominal interest rates matters for the economy as a whole. Even in an endowment economy with flexible prices for all goods, the central bank’s interest rate policy can determine the evolution of the general level of prices in the economy. In a production economy with sticky prices and/or wages, it can have important real effects as well. However, even in this classic model, the effectiveness of short-term nominal interest rate policies depends heavily on the absence of arbitrage in the financial market, a condition that can be significantly violated.

The irrelevance result can be easily understood in a representative agent setting, although the result does not depend on the representative agent assumption. In representative-household theory, the market price of any asset should be determined by the present value of the random returns to which it is a claim, where the present value is calculated using an asset pricing kernel (an SDF) derived from the representative household’s marginal utility of income in different future states of the world. Insofar as a mere reshuffling of assets between the central bank and the private sector should not change the real quantity of resources available for consumption in each state of the world, the representative household’s marginal utility of income in different states of the world should not change. Hence the pricing kernel should not change, and neither should the market price of one unit of a given asset, assuming that the risky returns to which the asset represents a claim have not changed. More intuitively, if the central bank takes more risky securities onto its own balance sheet and allows the representative household to hold only securities that pay as much in the event of a crash as in other states, this does not make the risk disappear from the economy. The central bank’s earnings on its portfolio will be lower in the crash state as a result of the asset exchange, and this will mean lower earnings distributed to the treasury, which will in turn mean that higher taxes will have to be collected by the government from
the private sector in that state; so the representative household’s after-tax income will be just as dependent on the risk as before. This explains why the asset pricing kernel does not change, and why asset prices are unaffected by open market operations.

A similar result can also be derived when there are heterogenous agents in the economy. If the central bank buys more of asset X by selling shares of asset Y, private investors should wish to purchase more of asset Y and divest themselves of asset X by exactly the amounts that undo the effects of the central bank’s trades. They optimally choose to do this is to hedge the additional tax/transfer income risk that they take on as a result of the change in the central bank’s portfolio. If share $\theta_h$ of the returns on the central bank’s portfolio are distributed to household $h$, where the $\{\theta_h\}$ are a set of weights that sum to 1, then household $h$ should choose a trade that cancels exactly fraction $\theta_h$ of the central bank’s trade to afford exactly the same state-contingent consumption stream as before. Summing over all households, the private sector chooses trades that, in aggregate, precisely cancel the central bank’s trade. In fact, the representative household assumption is not essential here. As long as it is assumed that agents can fully undo the central bank’s trade, the result holds even if different households have very different attitudes toward risk, different time profiles of income, different types of non-tradeable income risk that they need to hedge, and so on, and also regardless of how large or small the set of marketed securities may be. One can easily introduce heterogeneity of the kind that is often invoked as an explanation of time-varying risk premia without implying that any “portfolio balance” effects of central bank transactions should exist.

In fact, the portfolio balance effect is contrary to the proposition that the balance sheet size and composition are irrelevant. The portfolio balance effect of central bank transactions means that if the central bank holds less of certain assets and more of others, then the private sector is forced to hold more of the former and less of the latter as a requirement for equilibrium, and a change in the relative prices of the assets will almost always be required to induce the private parties to change the portfolios that they prefer. Therefore, portfolio balance effects imply that open market purchases of securities by the central bank must inevitably affect the market prices of those securities and hence other prices and quantities as well.
However, the recent financial crisis and the Great Recession taught us that all the assumptions that guarantee the irrelevance of the government balance sheet can be violated. First, we can see in Table 3 that from October 2007 to October 2008 the size, composition, and risk characteristics of the Fed’s balance sheet changed dramatically. By October 22, 2008, its assets were no longer mainly government bonds. Through the open market purchase programs, the Fed had built up a new balance sheet with assets mainly consisting of risky loans from the private sector. These assets could potentially have suffered substantial capital loss not offset by reductions in the liabilities. On the liability side, we see that more than 25% of its liabilities were in the form of special deposits from the U.S. Treasury. This made the Fed’s independence fragile, and the government balance sheet began to play a potentially important role in monetary policy.

From Figure 10(a), we can see that the component of the Fed’s liabilities constituted by reserves held by depository institutions changed in an especially remarkable way: by the fall of 2008, reserves were more than 100 times larger than they had been only a few months earlier. This explosive growth led some commentators to suggest that the main instrument of U.S. monetary policy had changed from an interest rate policy to one often described as “quantitative easing”. It seems that quantitative easing became the important monetary policy decision once the overnight rate (the federal funds rate) reached the zero lower bound, as it effectively has in the U.S. since December 2008. In Figure 10(b), we see that the past two years have also seen dramatic developments in regard to the composition of the asset side of the Fed’s balance sheet. Whereas prior to the fall of 2007, the Fed had largely held Treasury securities on its balance sheet, other kinds of assets have rapidly grown in importance. A variety of new “liquidity facilities”, new programs under which the Fed essentially became a direct lender to certain sectors of the economy, targeted purchases of certain kinds of assets, including more than a trillion dollars’ worth of mortgage-backed securities. Decisions about the management of these programs have occupied much of the attention of policymakers during the recent period.

Moreover, financial market frictions apparently affected the transition dynamics of monetary policy in a nontrivial way. Indeed, there is some evidence suggesting
Table 3: The U.S. Federal Reserve Balance Sheet – Assets and Liabilities

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reserve Bank Credit</strong></td>
<td>2,154,356</td>
<td>1,803,300</td>
<td>944,345</td>
</tr>
<tr>
<td>Securities Held Outright</td>
<td>1,692,177</td>
<td>490,633</td>
<td>- 288,947</td>
</tr>
<tr>
<td>U.S. Treasury</td>
<td>774,552</td>
<td>476,528</td>
<td>- 303,052</td>
</tr>
<tr>
<td>Federal Agency &amp; MBS</td>
<td>917,626</td>
<td>14,105</td>
<td>14,105</td>
</tr>
<tr>
<td>Repurchase Agreements</td>
<td>0</td>
<td>80,000</td>
<td>42,286</td>
</tr>
<tr>
<td>Term Auction Credit</td>
<td>139,245</td>
<td>263,092</td>
<td>263,092</td>
</tr>
<tr>
<td>Other Loans</td>
<td>107,630</td>
<td>418,580</td>
<td>481,050</td>
</tr>
<tr>
<td>Primary Credit</td>
<td>22,578</td>
<td>105,754</td>
<td>105,612</td>
</tr>
<tr>
<td>Primary Dealer and Other Broker-Dealer Credit</td>
<td>0</td>
<td>111,255</td>
<td>111,225</td>
</tr>
<tr>
<td>Asset-Backed Commercial Paper Money Market</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutual Fund Liquidity Facility</td>
<td>0</td>
<td>114,219</td>
<td>114,219</td>
</tr>
<tr>
<td>Credit extended to AIG</td>
<td>42,786</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Term Asset-Backed Securities Loan Facility</td>
<td>41,818</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other Credit Extensions</td>
<td>0</td>
<td>87,332</td>
<td>87,332</td>
</tr>
<tr>
<td>Net Portfolio Holdings of Maiden Lane LLC</td>
<td>26,381</td>
<td>29,137</td>
<td>29,137</td>
</tr>
<tr>
<td>Net Portfolio Holdings of Maiden Lane LLC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Float</td>
<td>-2,476</td>
<td>-1,048</td>
<td>-558</td>
</tr>
<tr>
<td>Central Bank Liquidity Swaps</td>
<td>33,315</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other Federal Reserve Assets</td>
<td>90,476</td>
<td>522,906</td>
<td>481,050</td>
</tr>
<tr>
<td><strong>Gold Stock</strong></td>
<td>11,041</td>
<td>11,041</td>
<td>0</td>
</tr>
<tr>
<td><strong>Special Drawing Rights Certificate Account</strong></td>
<td>5,200</td>
<td>2,200</td>
<td>0</td>
</tr>
<tr>
<td><strong>Treasury Currency Outstanding</strong></td>
<td>42,605</td>
<td>38,773</td>
<td>92</td>
</tr>
<tr>
<td><strong>Liability Items</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Currency in Circulation</strong></td>
<td>913,756</td>
<td>854,517</td>
<td>41,706</td>
</tr>
<tr>
<td><strong>Reverse Repurchase Agreements</strong></td>
<td>65,737</td>
<td>98,110</td>
<td>61,384</td>
</tr>
<tr>
<td>Foreign Official and International Accounts</td>
<td>65,737</td>
<td>73,110</td>
<td>36,384</td>
</tr>
<tr>
<td>Dealers</td>
<td>0</td>
<td>25,000</td>
<td>25,000</td>
</tr>
<tr>
<td><strong>Treasury Cash Holdings</strong></td>
<td>284</td>
<td>276</td>
<td>- 46</td>
</tr>
<tr>
<td><strong>Deposits with FR Banks</strong></td>
<td>86,496</td>
<td>554,927</td>
<td>542,435</td>
</tr>
<tr>
<td>U.S. Treasury, General Account</td>
<td>43,241</td>
<td>23,166</td>
<td>18,120</td>
</tr>
<tr>
<td>U.S. Treasury, Supplementary Financial Account</td>
<td>29,992</td>
<td>524,771</td>
<td>524,771</td>
</tr>
<tr>
<td>Foreign Official</td>
<td>2,297</td>
<td>254</td>
<td>255</td>
</tr>
<tr>
<td>Service-Related</td>
<td>3,237</td>
<td>6,138</td>
<td>- 441</td>
</tr>
<tr>
<td>Required Clearing Balances</td>
<td>3,237</td>
<td>6,138</td>
<td>-441</td>
</tr>
<tr>
<td>Adjustments to Compensate for Float</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>7,730</td>
<td>598</td>
<td>289</td>
</tr>
<tr>
<td><strong>Other Liabilities and Capital</strong></td>
<td>61,537</td>
<td>46,213</td>
<td>4,273</td>
</tr>
</tbody>
</table>

60
Figure 10: Liabilities (a) and assets (b) of the U.S. Federal Reserve (source: Federal Reserve Board).
that at least some of the Fed’s special credit facilities, and similar programs of the
other central banks, have affected asset prices. As a simple example, Figure 11 shows
the behavior of the spreads between yields on various categories of commercial paper
and the one-month overnight interest-rate swap rate (essentially, a market forecast
of the average federal funds rate over that horizon) over the period just before and
after the introduction of the Fed’s Commercial Paper Funding Facility (CPFF) at the
beginning of October 2008. (The darkest solid line shows the quantity of purchases
of commercial paper by the Fed, which spikes up sharply at the introduction of the
new facility.) The reason for the introduction of the new facility was the significant
disruption of the commercial paper market, indicated by the explosion of spreads in
September 2008 for all four types of commercial paper shown in Figure 11. The figure
also shows that spreads for three classes of paper (all except A2/P2 paper) came back
down again immediately with the introduction of the new facility, these three series
being precisely the ones that qualified for purchases under the CPFF. In contrast,
the spread for A2/P2 paper remained high for several more months, though this
spread also returned to more normal levels eventually with the general improvement
of financial conditions. Spreads did not decline in the case of paper not eligible for
purchase by the new facility, suggesting that targeted asset purchases by the Fed did
change the market prices of the assets.

During the recent crisis, conventional monetary policy measures, such as targeting
a short-term nominal interest rate, had negligible effect given the zero lower bound.
As a result, unconventional policy measures have become of great importance. Uncon-
tventional measures which use the central bank balance sheet as an instrument include:
(1) changes in the supply of bank reserves beyond those required to achieve an interest
rate target; (2) changes in the assets acquired by central banks (e.g., quantitative
easing and credit easing); and (3) changes in the interest rate paid on reserves. In order
to analyze these unconventional monetary policies, we need to extend the standard
New Keynesian DSGE model to allow a role for the government balance sheet in
equilibrium determination, and we need to consider the connections between these
alternative monetary policy measures and traditional interest rate policy. For example,
Cúrdia and Woodford (2011) extended the standard New Keynesian DSGE model by
Figure 11: Spreads between yields on four different classes of commercial paper and the 1-month OIS rate, together with the value of paper acquired by the Fed under CPFF. (Source: Federal Reserve Board.)
allowing a transactions role for central bank liabilities and heterogeneous households to guarantee that the government balance sheet has a nontrivial effect on determining equilibrium. This allows Cúrdia and Woodford (2011) to provide a framework to analyze unconventional monetary policy measures. In addition, Gertler and Karadi (2011) develop a quantitative monetary DSGE model with financial intermediaries that face endogenously determined balance sheet constraints to evaluate the effects of the central bank using unconventional monetary policy to combat a simulated financial crisis.

One important channel through which the balance sheet takes effect is the intertemporal budget constraint:

\[
GL_t = \sum_{k=t}^{\infty} E_t \left[ \frac{\Lambda_{t+k}}{\Lambda_t} \left( s_k + \frac{i_k}{1+i_k} \frac{M_k}{P_k} \right) \right]
\]

where \(GL_t\) denotes the real value of net government liabilities in periods \(t\), \(\Lambda_t\) is state price density in period \(t\), \(M_k\) is the money supply, \(P_k\) is the price level, and \(s_k\) is the real primary government budget surplus, which is the difference between revenues from taxes, real investments, the premium from insurance/guarantees, the assets held by the central bank and the treasury, etc., on the one hand, and the costs of capital for interest expenses on government debt, real investments, insurance payments, etc. on the other. Lucas (2012) reviews the theoretical and practical rationale for treating market risk as a cost to governments, presenting an interpretive review of the growing literature that applies the concepts and tools of modern finance to evaluating the costs of government policies and projects. Lucas (2012) stresses that governments typically understate their cost of capital because they identify it with their borrowing costs, rather than with a rate of return commensurate with the risk of a project. A consequence is that the official cost estimates for many government investment and financial activities are significantly understated. However, in a few cases risk adjustment lowers estimated costs relative to official estimates. Lucas (2012) emphasizes that when the financial market is incomplete, the choice of appropriate state price density, \(\Lambda_k\), becomes critical, and in practice rather tricky. In such a
complex case, different cash flows could require different SDFs. However, there is still a debate on whether the government balance sheet is constrained.

4.2 Fiscal Theory of the Price Level

Many specifications of monetary policy by themselves fail to determine a unique equilibrium in their inflation dynamics. These multiple equilibria arise when monetary policy is the sole focus of these models, the government budget constraint is ignored, and it is assumed that fiscal policy is completely accommodating to monetary policy. However, as pointed out by Leeper (1991), Sims (1994), and Woodford (1994, 1995), an active fiscal policy will be able to, in the words of Cochrane (2011), trim the set of equilibria, and achieve not just a determinate solution for inflation, but also for the price level. We elaborate this point below using a simple model from Cochrane (2011), which we describe in detail in Appendix E.

The equilibrium of the simple flexible price economy with a constant output $Y_t = Y$ and government spending $G_t = 0$ in each period can be summarized by the Fisher relation

$$1 + i_t = (1 + r)\Pi_{t+1},$$  \hspace{1cm} (49)$$

the monetary policy rule

$$1 + i_t = (1 + r)\Phi(\Pi_t),$$  \hspace{1cm} (50)$$

and the government’s present value budget constraint

$$\frac{B_{t-1}}{P_t} = \sum_{j=0}^{\infty} \frac{T_{t+j}}{(1 + r)^j},$$  \hspace{1cm} (51)$$

where $i_t$ is the nominal interest rate at time $t$, $r$ is the constant real interest rate, $P_t$ is the price level at time $t$, $\Pi_t \equiv P_t/P_{t-1}$ is the inflation, $B_{t-1}$ is the gross government
debt due at time $t$ and $T_t$ is the real tax at time $t$. It is assumed that $P_{-1}$ and $B_{-1}$ are given.

Combining Equations (49) and (50) we get the equilibrium path of inflation

$$
\Pi_{t+1} = \Phi(\Pi_t).
$$

Figure 12: Inflation dynamics with a Taylor rule. Source: Cochrane (2011).

Clearly, in general, without an initial or terminal condition for $\Pi_t$, monetary policy by itself will not determine the inflation rate, let alone the price level. A Taylor rule implies a locally deterministic equilibrium if we ignore equilibria with explosive inflation dynamics. Figure 12 illustrates this point. If $\Pi^*$ is the desired optimum level of inflation, a Taylor rule should have $\Phi'(\Pi^*) > 1$. However, this is an unstable equilibrium. As pointed out by Cochrane (2011), the only reason that the inflation $\Pi^*$ will hold is that for any other starting $\Pi_0 \neq \Pi^*$, the monetary authority would threaten to “blow up” the economy by creating explosive inflation. Cochrane (2011) points out that ruling out such explosive equilibria has no basis either in economic theory (since it is only inflation, or the nominal side, that is blowing up, not the real side, such as the asymptotic value of the debt), or in economic history, which has many recorded instances of hyperinflation. Thus, such explosive equilibria are
also valid solutions in addition to $\Pi^*$. Moreover, since $i_t \geq 0$, we must have a stable equilibrium $\Pi_L$ where $\Phi'(\Pi_L) < 1$, and to which any path of inflation starting from $\Pi_0 < \Pi^*$ must converge. This adds to the multiplicity of equilibria.

All such equilibria, however are valid only because we have ignored the government budget constraint (51), and assumed a Ricardian fiscal policy, which implies that the budget constraint must hold for all price levels and taxes must adjust to accommodate the price level obtained from monetary policy dynamics. An “active” (in the sense of Leeper, 1991) or non-Ricardian (Woodford, 1994) fiscal policy would force exogenous values of $\{T_t\}$ such that only one price level given by the budget constraint (51) can hold as an equilibrium. The policy is non-Ricardian because it does not satisfy the government budget constraint for all realizations of the price level. Non-Ricardian policies trim the set of equilibria, or in this case, select one equilibrium from the many that are implied by the monetary policy. With reference to Figure 12, the non-Ricardian fiscal policy would select $P_0$, following which inflation would follow a deterministic path to either $\Pi_L$ or $\Pi^*$.

To avoid explosive inflation dynamics (rather than ruling them out ad hoc), the following combination of active fiscal policy and passive monetary policy is very effective. The active fiscal policy takes the form of non-Ricardian policies as above, while the passive monetary policy in this case means that the interest rate should react less than one-for-one to inflation, or $\Phi'(\Pi^*) < 1$. This would ensure stability and convergence to $\Pi^*$ of any equilibrium path, the initial $P_0$ of which would be determined by the active fiscal policy. (See Figure 13.)

The multiplicity of equilibria is a common feature in New Keynesian models. The key to determining a unique equilibrium, when monetary policy is unable to determine the equilibrium alone, is the specification of additional policies to the monetary policy. Optimal policy prescriptions in zero-lower-bound situations, such as Werning (2011), must rely heavily on the Taylor rule to select the right equilibrium path. The Taylor rule, however, generates uniqueness only locally, not globally, as seen in Figure 12, and only after ruling out ad hoc explosive inflation dynamics.

This begs the question whether a non-Ricardian policy for long-term interest rates
coupled with a passive monetary policy can similarly be used to select equilibria. Quantitative easing, if we interpret it as an active policy for long-term interest rates, can help select the appropriate non-deflationary equilibrium. The crux of the ability to set long-term interest rates independently of the short-term rates can only be seen in a non-linearized model with uncertainty, with a policy for short-term rates that does not depend on the realizations of any of the endogenous variables. Changing the long-term interest rate independently of the short-term rate implies re-weighting the future short term rates by a different stochastic discount factor in an economy with variable output. We leave this question for future research.

4.3 Heterogeneity, Reallocation, and Redistribution Effects

In the core monetary models employed by major central banks, the existence of a representative agent is assumed, as seen in the simple canonical model in Section 3. It has been widely recognized that important features of macroeconomic data are difficult or impossible to explain within the representative agent framework. These include: the cyclical behavior of the factor shares of national income; the cyclical behavior of the large risk premium (see, e.g. Basak and Cuoco, 1998; Guvenen, 2009); the cyclical behavior of cross-sectional stock returns (see, e.g. Fama and French, 1992); the cyclical
behavior of the distribution of income, wealth, and leverage of households; and the
cyclical behavior of the distribution of leverage, asset, and cash holdings in firms. The
data also show strong cyclical patterns of capital, labor and credit reallocation among
firms, along with cyclical behavior in the bankruptcy rates, entries and exits, and
mergers and acquisitions among firms (see, e.g. Eisfeldt and Rampini, 2006, 2008).
These behaviors all should have significant equilibrium effects on aggregate quantities.
One important feature in the data is that households do not mainly save by themselves
to finance their consumption; instead, one side of households finances the other side’s
consumption (see, e.g. Guerrieri and Lorenzoni, 2012).

Theoretically, the representative agent can be justified by the assumption of
complete markets in an economy without externality. Complete market conditions
imply perfect insurance for agents in the economy, or that insurance in the economy
is costless. However, in reality, the profits of insurance companies in the U.S. and
costly transactions and portfolio constraints demonstrate the violation, or at least the
poor approximation, of the complete market assumption. Also, under assumptions
like a time-separable utility function and the stationarity of the economy, dynamically
complete market models predict the capital structure of firms and the asset portfolio of
households to be constant, even when the investment opportunity varies over time and
there are heterogeneous agents (see e.g. Judd et al., 2003). This is largely contradicted
by the real data. Moreover, Sims (2006) argues that there is no overall aggregate
capital and no aggregate consumption good, and that the real economy has a rich
array of financial markets, which models have so far not included in a widespread or
successful manner. Aggregate identities are problematic when we realize that different
agents in the economy deal with uncertainty and risk differently.

There are three main reasons for incorporating rich heterogeneity into macroeco-
nomic models. First, it is crucial not to ignore the significant equilibrium effects of the
distribution of wealth, income, leverage, cash holdings, etc. on the aggregate quantities
and the transitional dynamics of various monetary policies (see e.g. Guerrieri and
Lorenzoni, 2012). Moreover, from first principles, the cross-sectional distribution
serves as an infinite-dimensional important endogenous state variable, and hence
it has a strong impact on equilibrium. Second, heterogeneity makes it possible to
analyze the equilibrium effect of extensive margins on the aggregate quantities and the transitional dynamics of monetary policies. Third, the heterogeneity of agents provides a framework to assess the welfare properties of different monetary policy measures, fiscal policies, and other government policies, including unemployment insurance and social security programs.

The key feature of models with heterogeneous agents that makes them different from representative agent models is that the set of possible trades available for agents is restricted. The trading restrictions are usually modeled as an incomplete set of Arrow-Debreu securities, or portfolio constraints, or trading frictions. This prevents various aggregation results from holding (see e.g. Deaton, 1992). For example, Constantinides and Duffie (1996) show that the heterogeneous agent model with incomplete markets can be represented as several homogeneous agent models as long as the shocks in the model are such that agents do not gain anything by trading, even in the presence of those markets. Therefore, computing the equilibrium requires keeping track of the distribution of agents. In each period of time, the state of the economy is characterized by exogenous state variables driven by exogenous shocks, and by endogenous state variables whose law of motions are endogenously determined in the economy. The endogenous state variables usually include the distribution of agents. The difference between these distributions and distributions from ordinary endogenous state variables such as the capital stock is that they are usually infinite-dimensional mathematical objects. The equilibrium prices and quantities are functions of the (potentially infinite-dimensional) endogenous and exogenous state variables. The law of motion for endogenous state variables is usually called the transition map. Because the law of motion for endogenous state variables is an equilibrium output, it is intrinsically difficult to solve the transition map for a infinite-dimensional state variable. The aggregate level of prices and quantities is not enough to characterize the state of the economy or to predict future endogenous state variables. It is the key feature distinguishing the economy of heterogeneous agents and incomplete markets from the representative agent economy with complete markets. As Hall (1978, corollary 1, page 974) indicates, "No information available in period $t$ apart from the level of consumption, $c_t$, helps predict future consumption, $c_{t+1}$, in the sense of affecting
the expected value of marginal utility. In particular, income or wealth in period $t$ or earlier [is] irrelevant, once $c_t$ is known.”

Solving the heterogeneous-agent model entails solving the policy functions and the transition map simultaneously. Mathematically, this amounts to solving a fixed-point problem for an infinite-dimensional object. The standard numerical methods include discretization of the state space, parameterization of distributions, backward and recursive methods, and so on. To see the extra computational complexity caused by an endogenous transition map more clearly, we note that solving the problem of the agent for a given law of motion of the endogenous state variables (i.e., the transition map) is not enough. The correct transition map has to be found at the same time. This requires a double-layer iteration algorithm. To circumvent the complexity of the double-layer iteration algorithm, or to avoid iterations on the transition map, one needs to prevent the distribution of agents from affecting relative prices. Preventing this dramatically simplifies the computations. One example is Aiyagari (1994) which focuses on the steady state of an economy without aggregate fluctuations. In similar fashion, Imrohoruglu (1989) utilizes a storage technology that pins down the rate of return of savings exogenously, while Díaz-Giménez (1990) assumes that the government commits itself to a specific inflation rate policy that does not depend on the asset distribution.

**Heterogeneity and New Keynesian Models**

Heterogeneous agent models have lately been incorporated into the New Keynesian DSGE framework to study the effects of monetary policy. Algan and Ragot (2010) show the importance of a new precautionary savings motive in an incomplete market model in which the traditional redistributive effects of inflation are also introduced. The paper also shows that the long-run neutrality of inflation on capital accumulation obtained in complete market models does not hold under household binding credit constraints. They demonstrate that there is a quantitative rationale for the observed hump-shaped relationship between inflation and capital accumulation. Borrowing-constrained households are not able to adjust their money holdings differently compared
to unconstrained households since they cannot rebalance their financial portfolio when fluctuations in inflation become large. Inflation therefore increases capital accumulation due to the precautionary saving motive under heterogeneity.

It is necessary to understand heterogeneity to better study the redistribution effects of monetary policy. Gornemann et al. (2012) show that heterogeneous workers vary in their employment status due to search and matching frictions in the labor market, their potential labor income, and their amount of savings. This New Keynesian model quantitatively assesses who stands to gain or lose from unanticipated monetary accommodation, and who benefits the most from systematic monetary stabilization policy. This paper finds substantial redistribution effects from monetary policy shocks. A contractionary monetary policy has opposing effects on the wealthiest 5% versus the rest of the population. The top 5% enjoy increases in income and welfare, while the remaining 95% suffer under a contractionary monetary policy shock. Consequently, the negative effect of a contractionary monetary policy shock to social welfare is larger if heterogeneity is taken into account. In an influential paper, Eggertsson and Krugman (2012) theoretically discuss the importance of the redistribution effect of monetary policy between creditors and debtors in understanding the economic difficulties during 2007–2009. As for the redistribution effect of monetary policies, the firm heterogeneity is also important when it comes to inflation dynamics, investments, and risk premia (see, e.g. Jermann et al., 2014).

**Heterogeneity is Not New in Macroeconomics**

The effects of heterogeneity have long been studied in macroeconomics, leading to its serious adoption in New Keynesian DSGE models. Imrohoruglu (1989) is perhaps the first published paper to compute the equilibrium of a model with heterogeneous agents, and to calibrate it to match key U.S. observations. Imrohoruglu (1989) considers different institutional market arrangements under three different environments and also evaluates the welfare difference across institutional market arrangements. Similar welfare levels indicate that the existence of liquidity constraints in an economy is trivial for welfare considerations. Díaz-Giménez (1990) explores the business cycle
implications of alternative insurance technologies using a similar methodology to Imrohoruglu (1989), which could be easily adjusted to study the welfare effects of monetary and fiscal policy. Díaz-Giménez (1990) compares perfect insurance and monetary arrangements with pervasive liquidity constraints, finding that the welfare costs of monetary arrangements were 1.25% of output in zero-inflation economies. Hansen and Imrohoroglu (1992) find that the optimal level of unemployment insurance is very low, even when there is a very small amount of moral hazard.

Huggett (1993) explains the puzzle of very low risk-free interest rates in the postwar period in the U.S. by assessing the importance of the role played by the lack of insurance. Huggett (1993) does not have aggregate uncertainty and assumes an economy in which agents, subjected to idiosyncratic labor market shocks of the same type as in Imrohoruglu (1989), can lend and borrow up to certain limits at a rate that is endogenously determined by nontrivial market-clearing conditions, which are necessary to solve for the equilibrium of this economy.

Aiyagari (1994) describes two features: the first, an endowment economy that has no possibilities to save as a whole; the second, the level of aggregate savings affecting the society’s ability to produce goods. Aiyagari (1994) incorporates these features by using the standard neoclassical growth model with production. In order to measure the size of the role of precautionary savings, especially those motivated by self-insurance against idiosyncratic risk, Aiyagari (1994) has to deviate from the endowment economy setting from Huggett (1993). Aiyagari (1994) finds that with moderate and empirically plausible parameter values, uninsured idiosyncratic risk accounts for a 3% increase in the aggregate savings rate.

Krusell and Smith (1998) propose an important method for solving models with heterogeneity and aggregate uncertainty. When there are aggregate shocks in the model, the entire wealth distribution is an endogenous state variable, but its distribution can be approximated by its first few moments. The authors find that this approximate aggregation is reasonable, and to forecast future prices and quantities, it is enough to use the mean wealth instead of the entire cross-sectional wealth distribution. The distribution of wealth is unimportant to aggregate quantities such as aggregate consumption when most agents have the same marginal propensity to consume after
aggregate shocks. Most agents achieve good self-insurance in the model, which is equivalent to saying that the consumption policy functions are roughly linear. Aggregate capital by design is three times larger than output, and therefore most agents are rich enough to almost completely smooth out shocks. Only very poor agents, who account for a small fraction of aggregate consumption, do not have self-insurance. Krusell and Smith (1998) conduct an experiment to compare the model under complete market and incomplete market conditions, and find that heterogeneity has little effect on the model’s business cycle properties.

Liquidity and Heterogeneous Firms

It is well known that efficient trade and the reallocation of resources among different agents and sectors have a crucial impact on the macroeconomic performance and transitional dynamics of monetary policy (see, e.g. Walsh, 2012). However, the data show that resource mobility is far from frictionless, and the intensity of resource reallocation has strong cyclical patterns (see, e.g. Eisfeldt and Rampini, 2006). The imperfect nature of resource mobility plays a surprisingly small role in most policy models in major central banks. In those core New Keynesian DSGE models, for example, it is costly for firms to adjust their selling prices, but those same firms can hire and fire workers without cost, and both workers and capital can frictionlessly shift from one firm to another.

Theoretically, ignoring the potential costs associated with shifting real economic resources is consistent with a standard economy with one sector of homogeneous firms and representative households. However, real world economies consist of multiple sectors and heterogeneous agents, and the data shows that different sectors of the economy and different firms and households behave very differently over the course of the business cycle. For example, durable goods producing sectors are more cyclically sensitive than service sectors. Economic fluctuations may be associated with shifts in relative prices across sectors, or with persistent shifts in relative demand that may require labor and capital to shift from contracting to expanding sectors of the economy and from low productivity firms to high productivity firms. These shifts
require resources to transfer, yet differences in labor skills or in the type of capital employed in different occupations or sectors may make sectoral reallocations costly. The costs that arise because resources are not fully mobile may have consequences for policies on aggregate demand. Monetary policy shocks will definitely alter the transitional dynamics of the demand shock. For example, Walsh (2012) concludes that resource mobility matters for both the transitional dynamics of monetary policy shocks and the goals of monetary policy. Resource mobility affects the transmission mechanism that links monetary policy instruments to inflation and the real economy, thereby affecting the tradeoffs faced by the policy authority and the way policymakers weigh their objectives.

One important type of resource reallocation is capital reallocation. Eisfeldt and Rampini (2006) define the ease of capital reallocation between firms as capital liquidity and show that the amount of capital reallocation between U.S. firms is procyclical. In contrast, the benefits to capital reallocation appear countercyclical. The benefits to capital reallocation are approximated by the dispersion among the productivity of firms. This is intuitive because smart capital should flow out of low productivity firms into high productivity firms. They document that capital mobility is far from frictionless and particularly difficult in bad economic times. In order to quantify the cost of capital reallocation, they calibrate a simple model economy in which capital reallocation is subject to a standard adjustment cost function and impute the cost of reallocation. They find that reallocation costs need to be substantially countercyclical to be consistent with the observed joint cyclical properties of reallocation and productivity dispersion. Eisfeldt and Rampini (2008) provide one possible microfounded explanation for this endogenous inefficient capital reallocation. The authors argue that when managers have private information about the productivity of assets under their control and receive private benefits, substantial bonuses are required to induce less productive managers to declare that capital should be reallocated. Capital is less productively deployed in downturns because agency costs make reallocation more costly.

Another important type of resource reallocation is labor reallocation. Work by Davis et al. (1998) has been central to the surge of interest in this area. Their empirical
analysis is based on data for manufacturing plants covering the period from the early 1970s to the mid-1980s. After defining employment increases at new and growing plants as job creation, and decreases at dying and shrinking plants as job destruction, they pointed out a number of empirical regularities. One striking feature is that the data is marked by a high rate of job creation and destruction. On average, close to one out of ten manufacturing jobs disappeared in a given year, while the rate of new job creation is slightly lower. These changes are quite persistent: a year later, nearly seven out of ten newly created jobs were still in existence, and about eight in ten lost jobs were still lost. In addition, job creation and destruction tended to be concentrated at plants that experienced large changes in employment (those associated with plant shutdowns and startups, for instance). Another finding is that job destruction varied more noticeably over the cycle than job creation. The data show that job destruction tended to increase sharply during a recession and then fall back, while job creation did not move as much.

Some questions have been raised about these results. For instance, some economists have cautioned against relying on data for a single sector of the economy, especially manufacturing, where employment has been shrinking so noticeably. Furthermore, the data cover a relatively limited span (the 1970s and the 1980s), and it is possible that the recessions of this period differ fundamentally from previous (or subsequent) recessions in terms of restructuring and reallocation. Though the issue is not settled yet, some of the findings in Davis et al. (1998) have been replicated elsewhere. For instance, Blanchard and Diamond (1990) rely mainly on data from the Current Population Survey, which is not restricted to manufacturing alone, and confirm the finding about the relative volatility of job creation and destruction. For example, they find that “[...]booms are times of low job destruction rather than high job creation” (Blanchard and Diamond, 1990, p. 87); similar patterns have been discovered in data for foreign countries as well.

More recently, Kuehn et al. (2012) argue that frictions in the labor market are important for understanding the equity premium in the financial market. The authors embed the Diamond-Mortensen-Pissarides search framework into a DSGE model with recursive preferences. The model produces realistic equity premium and stock market
volatility, as well as a low and stable interest rate. In particular, they show that in their model the job flows and matching friction can help generate disasters in employment, output, and consumption along the lines of Rietz (1988) and Barro (2009). Moreover, when incorporated into otherwise standard RBC models, it has been shown to improve significantly their empirical performance. More importantly, it allows one to analyze the cyclical behavior of unemployment, job vacancies, and job flows, important phenomena which general equilibrium models based on Walrasian labor markets are not designed to address. For example, see Merz (1995), which tries to explain some cyclical behavior in the U.S. labor market by introducing a two-sided search in the labor market as an economic mechanism propagating technological shocks into a standard business cycle model; Andolfatto (1996), which shows that the labor market search is a quantitatively important propagation mechanism in generating business cycles; den Haan et al. (2000), which stresses the economic importance of the interaction between the capital adjustment cost and the labor destruction rate in propagating technology shocks; Gertler and Trigari (2009), which extends period-by-period Nash bargaining to staggered multiperiod wage contracts, and shows that it can account for the volatile behavior of labor market activities; and Hall (2005), which generates endogenous wage stickiness under a matching framework, and shows that sticky wages in turn make labor market activities realistically sensitive to aggregate shocks.

Given the significant equilibrium effects of job market reallocation, it is reasonable for us to speculate that job market mobility should have an important impact on the transitional dynamics of monetary policy shocks. In fact, there is an extensive literature that focuses on the positive implications of labor market friction in New Keynesian models, i.e., how search and matching frictions affect the empirical performance of the New Keynesian model and the transitional dynamics of monetary policy.\footnote{See, for example, Cheron and Langot (2000), Walsh (2005), Trigari (2006), Moyen and Sahuc (2005), Christoffel and Linzert (2005), and Krause and Lubik (2007), among many others.} Thomas (2008) analyzes the optimal monetary policy under the New Keynesian framework with search and matching frictions. Monetary policy shocks should affect job market flows in a nontrivial way.
Finally, the most important resource reallocation is the reallocation of credit or funding among firms or agents. The reallocation of funding is crucial, partly because it can possibly explain capital reallocation and labor reallocation, as is discussed in Eisfeldt and Rampini (2006). In Bernanke et al. (1999) and Kiyotaki and Moore (1997), essentially the representative firm or the productive agent is impatient enough such that the firm or agent does not save very much, and does not escape its financial constraints. Consequently, the models have two salient features: first, the firm or the agent saves by itself and uses the savings to invest later; and second, the economy as a whole is financially constrained in the steady state. However, these two implications are both inconsistent with the data. On the contrary, the data suggest that only a fraction of firms are occasionally bound by financial constraints, and that firms also finance each other’s investment. Fund reallocation among firms is one of the key functions of the financial sector.

Chari et al. (2008) stress two facts that have been underappreciated. First, non-financial corporations in the aggregate can pay their capital expenditures entirely from their retained earnings and dividends without borrowing from banks or households. Second, in the aggregate, increases in non-financial corporate debt are roughly matched by increases in their share repurchases. More precisely, Figure 14(a) shows that in the aggregate, without any funds from the rest of the economy, the cash available to these firms from their operations can easily pay for their investment expenditures. Figure 14(b) shows that equity repurchases are roughly matched by funds raised through credit market instruments. The data suggest that in the aggregate, firms raise debt to buy back their shares, and not to finance investment.

However, it can be misleading to conclude purely from the aggregate macroeconomic time series that a deep financial crisis is not observed or that the poor condition of the financial system did not affect the corporate sector much during 2007–2009. Among many others, Shourideh and Zetlin-Jones (2012) emphasize the role that financial markets play in reallocating funds from cash-rich, low productivity firms to cash-poor, high productivity firms. In their calibrated model, they find that a shock to the collateral constraints, generating a one standard deviation decline in the debt-to-asset ratio, leads to a 0.5% decline in aggregate output on impact, roughly comparable to
Figure 14: (a) Retained Earnings, Dividends, and Capital Expenditure; (b) New Debt and Net Repurchases of Equity.
the effect of a one standard deviation shock to aggregate productivity in a standard RBC model. They find that disturbances in financial markets are a promising source of business cycle fluctuations when non-financial linkages across firms are sufficiently strong.

4.4 Risk Premium Dynamics

Does the risk premium matter for macroeconomic dynamics and the transitional dynamics of monetary policy? The key features for risk premium dynamics include high levels of volatility, nonlinearity, and countercyclicality. However, as explained in Jermann (1998), Lettau and Uhlig (2000), and Kaltenbrunner and Lochstoer (2010), it is often difficult to generate endogenously a large and time-varying market price of risk in a production economy.

Rouwenhorst (1995) shows that the standard RBC model fails to explain the equity premium because of consumption smoothing. Using models with internal habit preferences, Jermann (1998) and Boldrin et al. (2001) use capital adjustment costs and cross-sector immobility, respectively, to restrict consumption smoothing to explain the equity premium. However, both models struggle with excessively high interest rate volatilities. Using recursive preferences to curb interest rate volatility, Tallarini (2000) and Kaltenbrunner and Lochstoer (2010) show that baseline production economies without labor market frictions can explain the Sharpe ratio, but still fail to match the equity premium and the stock market volatility.

Uhlig (2007) shows that wage rigidity helps explain the Sharpe ratio and the interest rate volatility in an external habit model, but in this model the equity premium and the stock market volatility are close to zero. Gourio (2011) shows that operating leverage derived from labor contracting helps explain the cross-section of expected returns, but aggregate asset prices are not studied. Favilukis and Lin (2012) quantify the role of infrequent wage renegotiations in an equilibrium asset pricing model with long-run productivity risk and labor adjustment costs. They argue that, in standard models, highly procyclical and volatile wages are a hedge against adverse shocks of productivity for the shareholder. The residual—profit or dividends—becomes
unrealistically smooth, as do returns. Smoother wages act like operating leverage, making profits more risky. Bad times and unproductive firms are especially risky because committed wage payments are high relative to output. Instead of specifying the wage rule exogenously, Kuehn et al. (2012) differ from the prior studies by using the search framework to derive equilibrium wages. Because dividends equal output minus wages minus total vacancy costs (in an analogous manner to investment), providing a microfoundation for equilibrium wages makes the dividends truly endogenous in a production economy.


Gourio (2012) shows that an increase in disaster risk leads to a decline in employment, output, investment, stock prices, and interest rates, and an increase in the expected return on risky assets. The model matches the data well on quantities, asset prices, and particularly the relationship between quantities and prices, suggesting that variation in aggregate risk plays a significant role in some business cycles. More precisely, the mechanism is that an increase in the disaster probability affects the economy by lowering expectations and increasing risk. Because investors are risk averse, this higher risk leads to higher risk premia, with significant implications for both business cycles and asset prices: stock prices fall, employment and output contract, and investment especially declines. Demand for precautionary savings increases, leading the yield on less risky assets to fall, while expected excess returns on risky securities increase. These dynamics occur in the absence of any change in TFP.

Risk premia are important in understanding many macroeconomic questions, for instance, why investment is often low despite low riskless interest rates. Here, the relevant user cost of capital may well be high if the riskless interest rate is low precisely because of high disaster risk. This will directly affect the transitional dynamics of
monetary policies.

Introducing time-varying risk requires solving a model using nonlinear methods, i.e., going beyond the first-order approximation and considering higher-order terms in the Taylor expansion. Researchers disagree on the importance of these higher-order terms, and a fairly common view is that they are irrelevant for macroeconomic quantities. In his presidential address, Lucas (2003, p. 7) summarizes this perspective: “Tallarini uses preferences of the Epstein-Zin type, with an intertemporal substitution elasticity of one, to construct an RBC model of the U.S. economy. He finds an astonishing separation of quantity and asset price determination: The behavior of aggregate quantities depends hardly at all on attitudes toward risk, so the coefficient of risk aversion is left free to account for the equity premium perfectly.”

Gourio (2012) shows, however, that when the risk is large and varies over time, risk aversion affects macroeconomic dynamics in a significant way. In a similar spirit, but using a two-country open economy setting, Dou and Verdelhan (2014) show that the time-varying risks generate rich joint volatile dynamics of international asset prices and capital flows.

The following are some particular examples of the potential importance of the time-varying risk premium on macroeconomic dynamics and the transitional dynamics of monetary policy. Gilchrist and Zakrajsek (2012) show that the default premium, rather than the default probability, is the informative variable about macroeconomic conditions. Gilchrist et al. (2010) show that an uncertainty shock can boost the default premium strongly, without increasing the default probability of firms significantly. The extremely high default risk premium prevents firms from investing optimally, even when the risk-free rate is low. The term premium is crucial to accurately characterize the aggregate demand relationship (the IS curve). According to Galí and Gertler (2007), the aggregate demand depends on the gap between the long-term interest rate and its natural correspondence in a model economy with flexible prices. The relationship between the long-term and short-term interest rate is captured by the term premium, which depends on the risk premium of investors. According to Galí and Gertler (2007), the IS curve is also characterized by the relationship between aggregate demand and marginal $q$. The dependence between marginal $q$ and the short-term
interest rate also largely captured by financial friction and the risk premium. However, to the best of our knowledge, generating a realistic term premium is still a challenging task in a model production economy. Reasonable risk premia, including currency and sovereign risk premia, are crucial to understanding international financial linkages and capital flow dynamics which, in turn, have a nontrivial impact on the implications of monetary policy.

Finally, as emphasized in Section 2.3, the availability of rich financial data makes DSGE models particularly useful in learning deep structural parameters. However, the absence of reasonable risk premium dynamics in DSGE models wastes the information embedded in asset prices.

4.5 Uncertainty

The uncertainty shock has been shown to have an adverse effect on macroeconomic quantities and can even drive business cycles. For example, the Federal Open Market Committee minutes repeatedly emphasize uncertainty as a key factor driving the 2001 and 2007–2009 recessions, while Stock and Watson (2012, p. 26) conclude that “The main contributions to the decline in output and employment during the [2007–2009] recession are estimated to come from financial and uncertainty shocks.”

In addition, in recent studies (see, e.g. Christiano et al., 2010; Del Negro and Schorfheide, 2012), economists have empirically found that the two most important shocks that drive aggregate fluctuations are “financial disruption” and “heightened uncertainty”. However, as emphasized by Hansen (2012), it is crucial to have a better understanding of the sources of financial and uncertainty shocks in macroeconomic models and their endogenous interactions.

In fact, there has been a fast-growing literature studying the aggregate effects of such uncertainty shocks (e.g., Pástor and Veronesi, 2006, 2009; Bloom, 2009; Arellano et al., 2011; Bloom et al., 2013a; Bachmann and Bayer, 2014; Christiano et al., 2010, 2014; Bundick and Basu, 2014; Gilchrist et al., 2010; Herskovici et al., 2014). It should be noted that the use of the term uncertainty here is different from Knightian uncertainty, which emphasizes situations where agents cannot know all the information
they need to set accurate odds (e.g., Knight, 1921; Hansen and Sargent, 2008). In addition, the use of uncertainty here is also different from aggregate volatility, which has also been extensively studied in the literature (Bansal and Yaron, 2004; Drechsler and Yaron, 2011; Shaliastovich, 2015; Campbell et al., 2013, 2015; Fernandez-Villaverde et al., 2011; Nakamura et al., 2014; Segal et al., 2013; Gourio et al., 2015; Ai and Kiku, 2015).

Impact of Uncertainty Shocks

Since the recent financial crisis and the Great Recession, policy authorities and academic researchers have been engaging in a vigorous debate on the impact of uncertainty shocks on the joint dynamics of macroeconomic quantities and asset prices. Policy authorities, including the Fed and the ECB, have claimed that uncertainty has had an adverse effect on their economy, and have built uncertainty shocks into their core DSGE models as a main driver of aggregate fluctuations (see, e.g. Christiano et al., 2010, 2014). Moreover, there is an extensive academic literature showing the adverse effect of uncertainty, e.g., Bloom (2009); Bloom et al. (2013b), Gilchrist et al. (2010), and Basu and Bundick (2011), among others.

However, there are two major concerns with this narrative. First, the causal relationship between fluctuations in uncertainty and fluctuations in the economy is far from clear to policymakers and researchers. Although the correlation between fluctuations in uncertainty and the economy is evident, it is still undetermined whether the heightened uncertainty partially caused the Great Recession, and whether it should be blamed for prolonging the recovery process out of the Great Recession. This is due both to the lack of crystal-clear empirical evidence, and the lack of comprehensive theoretical studies on the equilibrium feedback effect between fluctuations in uncertainty and the economy, as suggested by Bloom (2013), a review paper on uncertainty. Second, it has been argued that uncertainty could have a positive effect on investment and the stock market. Pastor and Veronesi (2006) use a simple calibrated

---

11 For example, Federal Reserve Bank of Dallas President Richard Fisher gave a formal speech titled “Uncertainty matters. A lot.” emphasizing that uncertainty might worsen the Great Recession and on-going recovery, at the 2013 Causes and Macroeconomic Consequences of Uncertainty Conference.
stock valuation model with uncertainty to show that the fundamental value of a firm increases with uncertainty about its average future profitability, and this uncertainty was extremely high in the late 1990s. Bar-Ilan and Strange (1996) showed that in a high uncertainty environment, the benefits from investment, including the growth opportunity caused by investment lags and abandoned project options, can dominate the cost of investment, the loss of the real option value of waiting. As a result, high uncertainty can sometimes promote investment.

In fact, there is a rich, if contradictory, literature on the relationship between uncertainty and macroeconomic quantities including consumption and investment. Different theories emphasize different channels, some showing a positive relationship and others showing a negative relationship. As a whole, the impact of uncertainty is still ambiguous. The basic channels under consideration include the real option channel (i.e., the option to wait), the risk premium channel, the precautionary savings channel, the growth opportunity channel, the Oi-Hartman-Abel-Caballero channel, and the learning-by-doing channel.

The first channel under consideration, the real option channel, appears to be the most direct channel through which uncertainty can potentially affect a firm’s investment and hiring decisions. The idea is that the sizable adjustment cost in investment and hiring (see, e.g. Ramey and Shapiro, 2001; Cooper and Haltiwanger, 2006) and its irreversibility (see, e.g. Pindyck, 1991; Kogan, 2001) together make the investment decision effectively a decision on exercising call options. This real option can be viewed as an option to wait, and the opportunity cost of delay is the foregone income from the project, which is unaffected by uncertainty. This asymmetric effect of uncertainty on the benefits and costs of waiting captures the essence of the real option effect. This is referred to as the “bad news principle” by Bernanke (1983). However, the real option effect can be alleviated or even overturned when environmental variables shift. For example, when projects have a liquid reallocation market (i.e., reversible), the real option effect is negligible. Another, more relevant, example is when firms are financially constrained. As demonstrated in Bolton et al.

---

12 There is a long literature on the real option effect, including Bernanke (1983), Brennan and Schwartz (1985), McDonald and Siegel (1986), and Dixit and Pindyck (1994).
(2013), for financially constrained firms, the uncertainty shock can have both a positive and a negative effect on a firm’s investment and financing decisions.

The idea behind the second channel, the risk premium channel, is that uncertainty reduces aggregate investment, hiring, and growth through a sharp increase in the risk premium. The risk premium channel plays a key role in linking asset pricing to the interaction between uncertainty and investment/hiring, an idea which has been missing in the uncertainty literature, although with a few exceptions, including Gilchrist et al. (2010), Arellano et al. (2011), and Christiano et al. (2014). The key idea is that in an economy with corporate debt and costly default, higher uncertainty raises the default probability for those firms that are already near default boundaries, and hence the cost of debt financing increases. This in turn reduces investment, and increases the default probabilities for firms originally further from the default boundaries, and accordingly diminishes hiring, which in turn leads to lower consumption of households. This adverse feedback loop causes a ripple effect, dragging the whole economy into recession, while creating sky-high credit spreads. It is clear that if financial intermediaries are strong, and very few firms are close to their financial binding constraints, the risk premium effect on an economic downturn will largely be dampened. This is a nontrivial point in generating rich and realistic endogenous uncertainty dynamics.

The third channel, the precautionary saving channel, focuses on households. It is evident that higher uncertainty depresses household consumption expenditures (see, e.g. Bansal and Yaron, 2004). In a full-closed economy, the motivation to increase precautionary savings will also reduce contemporaneous consumption, and at the same time increase investment. However, the investment will also drop when uncertainty is high, assuming price rigidity (see, e.g. Basu and Bundick, 2011; Leduc and Liu, 2012).

Growth opportunities, the fourth channel we consider, are the major force generating a positive association between uncertainty and investment. This idea is usually implemented in two ways in the literature. Following Bar-Ilan and Strange (1996), the first method assumes that there is an investment lag with a time-to-build, $h > 0$, and an abandonment option available for each project. The abandonment option means the loss is bounded below in bad states, while the time-to-build feature forces the firm to invest earlier in order to be able to capture opportunities in the near future. The
two components together cause the rational firm to invest sooner in a high uncertainty environment. The second method is to model two capital goods: traditional capital called “trees”, and investment options called “seeds” (see, e.g. Jovanovic, 2009). In a high uncertainty environment, the investment in seeds experiences a gradual boom.

Fifth, the Oi-Hartman-Abel-Caballero channel is based on the work by Oi (1961), Hartman (1972), Abel (1983), and Caballero (1991). The key idea of these models is that the adjustment cost of capital makes investment less flexible than labor adjustment. This concept, combined with a constant-return-to-scale technology, makes the marginal product of capital a convex function of output price. It follows from Jensen’s inequality that uncertainty in output price leads to a high marginal product of capital, and hence to a high intensity of investment.

Finally, the learning-by-doing channel assumes that investors or firms have imperfect information about the underlying state of the economy, and that the only way receive extra signals about its true state is by a sequence of investments. It naturally follows that in a high uncertainty environment, firms conduct earlier and more intensive investment to learn the true state (see, e.g. Roberts and Weitzman, 1981; Pindyck, 1993; Pavlova, 2002).

An important and still unanswered question is which channel dominates under which economic conditions. It is possible that the sign and magnitude of the impact of uncertainty shocks on investment and asset prices depend on the soundness of the financial system and the prevailing external financing costs. When financial intermediaries are strong and the risk premium is low, negative effect channels such as the real option channel and the risk premium channel will have limited impact because investment options are deep out of the money, and it is hard to trigger a crash in the financial market. In contrast, positive effect channels are given full play in this environment. Therefore, higher uncertainty should lead to earlier and more intensive investment and create a stock market boom. When the financial intermediaries are fragile, however, the real option channel with liquidity hoarding and the risk premium channel will dominate, while positive effect channels will play a very limited role.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
Measuring Uncertainty

It is impossible to measure uncertainty directly given that it is unobservable, conceptual, and ex ante. After all, it lives in people’s minds and has no direct material instantiation. Therefore, a range of proxies have been employed to study the impact of uncertainty shocks. Aggregate market volatility and aggregate TFP volatility are among the most popular proxies for uncertainty shock in the existing literature (see, e.g. Bloom, 2009; Bansal et al., 2014; Campbell et al., 2012). These aggregate volatility proxies are usually referred to as “macro uncertainty”. In other papers (see, e.g. Bloom, 2009; Gilchrist et al., 2010), the uncertainty shock is approximated by an increase in the cross-sectional dispersion among agents. These dispersion-based measures are referred to as “micro uncertainty”. There are also measures based on survey data. These measures include forecaster disagreement and news mentions of uncertainty. Empirically, they are all believed to be reasonably good proxies because they co-move over time. This co-movement itself is a nontrivial puzzle to solve, and its solution should shed light on providing a better proxy for uncertainty, and therefore a better equilibrium impact of uncertainty shocks.

4.6 The Financial Sector and Systemic Risk

Systemic risk is believed to be a key driver of the recent financial crisis and the Great Recession. Systemic risk is rooted in the financial sector, and through contagion it has a strong adverse effect on the whole economy. However, the key players in systemic risk, financial intermediaries, are missing in the New Keynesian DSGE models used by the central banks. When incorporating financial intermediation into macroeconomic models, it is essential to consider two crucial effects: the balance-sheet effect for financial intermediaries and the effect of imperfections in the interbank lending market. First, in order to analyze how shocks in the financial sector spill over to the whole economy, neither the nonfinancial corporate balance sheet nor the integrated bank-firm balance sheet is satisfactory. Second, in order to model endogenous systemic risk, properly modeling the interbank lending market and the interconnections between financial intermediaries is crucial. Incorporating the financial sector into these models
is necessary to allow us to study endogenous systemic risk in the economy. This is critical for analyzing conventional monetary policy, unconventional monetary policy, and macroprudential policy since the primary task of the monetary authority is to maintain a healthy financial system in normal times and restore a distorted financial system in times of crisis.

Most of the recent macroeconomics literature on financial frictions focuses on credit market constraints on nonfinancial corporate borrowers without any real role for financial intermediation. The most recent advances include Bernanke and Gertler (1989), Kiyotaki and Moore (1997), and Bernanke et al. (1999). Among them, Bernanke et al. (1999) introduces a financial accelerator into the New Keynesian DSGE framework. However, they all focus on the balance-sheet effect of nonfinancial firms, while ignoring the unique properties of financial intermediaries. The financial accelerator channel from the balance sheet of nonfinancial corporate borrowers is definitely a relevant financial friction, but it is just one aspect of many possible financial frictions.

One of the first papers which tried to incorporate the financial sector into macroeconomic models, and studied the effects of financial intermediary balance sheets and the interbank lending market is Gertler and Kiyotaki (2010). There are also other papers analyzing monetary policies with liquidity risks in the interbank markets (e.g. Freixas et al., 2011). The authors focus on understanding how disruptions in financial intermediation can induce a crisis that affects the real economy. The credit market constraints on financial intermediation are incorporated into an RBC framework, modified with habit formation and flow investment adjustment costs instead of capital adjustment costs. More precisely, the financial intermediaries are assumed to play three unique roles in the economy, as discussed extensively in the literature. First, the financial intermediaries are delegated monitors and specialists. They are conduits that channel funds from households to nonfinancial firms. Second, the financial intermediaries engage in maturity transformation. In the model, they are assumed to issue short-term debts and hold long-term assets. Third, the financial intermediaries facilitate liquidity provision via the interbank lending market. In the model, it is assumed that there is a continuum of “banks” which fund the goods producers and
finance their investment from both a wholesale market, i.e., an interbank lending market, and a retail market, where banks hold deposits from households. To simplify this analysis, the authors assume constant returns to scale production, perfect labor mobility, and goods producers without financial constraints. With these assumptions, there is no need to keep track of the distribution of capital held by producers or their net worth. It is also assumed that the banks and the nonfinancial firms are “buddies”, in the sense that there is no financial friction in their funding relationship. In other words, it is essentially assumed that the producers’ balance sheet can be viewed as part of the banks’ balance sheet. To achieve such a simplification, following Gertler and Karadi (2011), the authors assume complete consumption insurance among workers and bankers, and independent and identically distributed random turnovers between workers and bankers. By doing so, this guarantees that a representative household exists to determine aggregate consumption and prices, with no need to track the wealth distribution of households. In this complete market economy, there is unique SDF while the agents are actively borrowing and lending in equilibrium.

In contrast, He and Krishnamurthy (2013) incorporate the financial intermediary into a standard macro-finance model with segmentation between bankers and workers and reasonable terms of macroeconomic calibration in order to generate extremely nonlinear risk premium dynamics. They model the dynamics of risk premia during crises in asset markets where the marginal investor is a financial intermediary. In this model, intermediaries face an equity capital constraint. Risk premia rise when the constraint binds, reflecting the capital scarcity. The calibrated model matches the nonlinearity of risk premia during crises and the speed of reversion in risk premia from crisis levels back to pre-crisis levels. They evaluate the effect of three government policies: reducing intermediaries’ borrowing costs, injecting equity capital, and purchasing distressed assets. Injecting equity capital is particularly effective because it alleviates the equity capital constraint that drives the model’s crisis. However, it is still far from satisfactory for monetary policy decision making because the model simplifies many important features, and hence there is no way to see how the constrained financial intermediaries would affect the economy as a whole in such a model.

If the theory of He and Krishnamurthy (2013) is correct, the marginal value
of wealth for financial intermediaries should therefore provide a more informative SDF than that of a representative consumer. Empirically, Adrian et al. (2010) use shocks to the leverage of securities broker-dealers to construct an intermediary SDF. Intuitively, deteriorating funding conditions are associated with de-leveraging and a high marginal value of wealth. Their single-factor model prices size, book-to-market, momentum, and bond portfolios with an $R^2$ of 77% and an average annual pricing error of 1%, performing as well as standard multi-factor benchmarks designed to price these assets. It empirically documents that financial intermediary balance sheets contain strong predictive power for future excess returns on a broad set of equity, corporate, and Treasury bond portfolios. They also show that the same intermediary variables that predict excess returns forecast real economic activity and various measures of inflation. Their findings point to the importance of financing frictions in macroeconomic dynamics and provide quantitative guidance for preemptive macroprudential and monetary policies.

Moreover, Gilchrist and Zakrajsek (2012) empirically relate the predictive power of bond premia for the business cycle to the risk-bearing capacity of the marginal investors in these bonds. These investors act in a more risk-averse way when their capital becomes impaired, which translates to an increase of the bond premium and a reduction in the supply of credit available to potential borrowers.

However, the literature of occasionally binding financial constraints on financial intermediation, including Brunnermeier and Sannikov (2014), He and Krishnamurthy (2012), and Danielsson et al. (2011), stresses that precautionary effects can generate endogenous tightening of margins. This is in contrast to the literature of the financial accelerator, including Bernanke and Gertler (1989), Bernanke et al. (1999), Kiyotaki and Moore (1997), Gertler and Kiyotaki (2010), and Christiano et al. (2010) in which the financial constraints are always binding. The precautionary effect means that even if the borrowing constraint is not currently binding, an increase in likelihood that it could be binding in the future, possibly due to increased uncertainty, can induce a tightening of margins. While the papers on occasionally-binding financial constraints on financial intermediation provide reasonable asset pricing dynamics with enough nonlinearities, they generally fail to generate enough nonlinearities
in macroeconomic quantities. Dewachter and Wouters (2012) add a reduced-form working capital constraint on goods producers, which quantitatively improves the macroeconomic dynamics of the model.

Recently the CMR framework, introduced by Christiano et al. (2010), has been adopted by the ECB. However, the CMR framework is still subject to some major concerns. First, while a crisis in this framework can now originate in the financial sector, rather than through risk in the production sector, the interbank market is still missing in this version of the model. Second, the model poorly accounts for the external financing premium variables (e.g., the credit spread), negating the important advantage of modeling the financial sector separately from the real sectors. Third, the log-linearization method fails to capture key dynamics of the financial sector and the real economy. Fourth, the absence of precautionary effect makes it hard to generate realistic nonlinear dynamics in asset prices.

4.7 Goods Market and Markups

In the baseline New Keynesian DSGE model, the desired markup of price over marginal cost is constant. This is mainly due to two factors: a constant elasticity of substitution among differentiated goods, and the validity of the Modigliani-Miller theorem. As observed by Blanchard (2009), however, the desired markup appears to be anything but constant, and how it varies in response to other factors is still unknown territory in macroeconomics.

One popular model for the desired markup is the customer market mechanism, in which a firm that lowers its current price not only sells more to its existing customers, but also expands its customer base, leading to higher future sales at any given price. This idea was first introduced by Phelps and Winter (1970), and formalized by Gottfries (1986), Klemperer (1987), Farrell and Shapiro (1988), and Bils (1989),

14For example, some evidence comes from findings on the pass-through effects of exchange rate movements. Recent empirical work on the U.S., using disaggregated prices, shows that, when import prices are denominated in dollars at the border, exchange rate movements have minimal effect on the prices for these imports in the U.S. Conversely, if import prices are denominated in foreign currency at the border, exchange rate movements lead to nearly one-for-one effects on prices for those imports (see e.g. Gopinath et al., 2010).
among others. Several strands in the marketing and industrial organization literature provide empirical support for the customer model. For example, Houthakker and Taylor (1966) use 80 detailed items of consumer expenditure and find that current demand depends positively on existing inventory, suggesting some habit formation. Guadagni and Little (1983) use a multinomial logit model of brand choice, calibrated on 32 weeks of purchases of regular ground coffee by 100 households, and show high statistical significance for the explanatory variables of brand loyalty. Erdem (1996) finds that for margarine, peanut butter, yogurt, and liquid detergent, accounting for habit formation improves both in-sample and out-of-sample fit. More recently, Genesove and Mullin (1997) find that the American Sugar Refining Company sharply cuts its price to maintain market share and to deter entry of competitors. Bronnenberg et al. (2009) find an early entry effect on a brand’s current market share and perceived quality across U.S. cities. There is also direct evidence provided by firm-level surveys in several OECD countries, all pointing out that price stickiness is mainly driven by customer relationships. Dou and Ji (2014) build on the idea of the customer market, and analyze how financing decisions interact with strategic pricing when the financial market is imperfect. One major focus of Dou and Ji (2014) is the endogenous relationship between financing and the price-setting behavior of firms. A closely related paper is Chevalier and Scharfstein (1994), which studies the impact of imperfect financial markets on firms’ price-setting decisions. However, the authors use limited supermarket data to test the causal effect of liquidity shocks on goods price.

Chevalier and Scharfstein (1994) were the first to introduce capital-market imperfection into a customer market model in an attempt to interpret countercyclical markups. They focus on how liquidity constraints affect pricing behavior and find that liquidity-constrained firms have an incentive to raise prices in order to boost current profits to meet their liabilities and finance investment. A recent paper by Sim et al. (2013) provides more evidence using product-level price data. They find that during the recent crisis, firms with weak balance sheets increased prices significantly relative to industry averages, whereas firms with strong balance sheets lowered prices. A general

\[^{15}\text{See, e.g., Hall et al. (1997), Aucremanne and Druant (2005), Fabiana et al. (2005), and Amirault et al. (2006).}\]
equilibrium model with financial market distortions is proposed to rationalize these findings. The idea of Dou and Ji (2014) is related to Chevalier and Scharfstein (1994) and Sim et al. (2013), but they look for more evidence on firms’ pricing and financing behavior in normal times across industries and on the frequency of price resetting over business cycles. Moreover, Dou and Ji (2014) focus more on the interaction between financing and pricing. Their DSGE model provides a unified theory on Tobin’s $q$, corporate investment, financing, price setting, and asset pricing. Weber (2013) shows that firms that adjust their product prices infrequently earn a cross-sectional return premium of more than 4% a year.

The cyclicality of markups has significant implications for economic fluctuations, since countercyclical markups would tend to dampen fluctuations in economic activity, whereas procyclical markups would amplify fluctuations. To generate procyclical factor prices, Rotemberg and Woodford (1991) propose that markups should be countercyclical and they provide evidence using aggregate data. The countercyclical markup is also found in Bils (1987) and in the supermarket industry (Chevalier and Scharfstein, 1994). However, other studies find that markups are procyclical using different industry-level data.\footnote{See, e.g., Domowitz et al. (1986), Machin and Van Reenen (1993), Ghosal (2000), Nekarda and Ramey (2013).}

Based on the customer market model introduced by Phelps (1998), Gilchrist et al. (2012) investigate the effect of financial conditions on price-setting behavior during the recent financial crisis by assuming a deep habit component in the model. In their model, firms have an incentive to set a low price to invest in market share. In other words, the loss from setting a lower price can be viewed as an investment cost for positive net present value projects (i.e., market shares). When financial distortions are severe, firms forgo these investment opportunities and maintain high prices because the marginal value for cash dominates the profits from investment in market share. The model with financial distortions implies a substantial attenuation of price dynamics in response to contractionary demand shocks relative to the baseline without financial distortions, which has important policy implications. Empirically, the authors find theory-consistent evidence that, at the peak of the crisis, firms with relatively weak
balance sheets increased prices, while firms with strong balance sheets lowered their prices.

4.8 Solution, Estimation, and Evaluation

The proper methodologies for the solution, estimation, and evaluation of New Keynesian DSGE models are critically important in economics, yet extremely hard to do technically. Without proper methods, the credibility of monetary policies based on these models will be dramatically compromised, and their results may be extremely misleading, even if the modeler constructs a perfect model, one which incorporates all the mechanisms discussed in previous sections. In this section, we shall review the methodologies used by central banks, and point out their principal issues.

Solution Methods

Solving the DSGE model with heterogeneous agents in incomplete markets and severe nonlinearity is mathematically equivalent to solving a large system of nonlinear equations. The nonlinearity and infinite dimensionality of the model makes the problem extremely challenging, even for mathematicians and computer scientists. Given these technical and computational challenges, economists must make difficult trade-offs between complexity and tractability when specifying the model.

Because of these trade-offs, macroeconomists at central banks prefer to use simpler models and the log-linearization solution method. The DSGE model relies on log-linearization around the steady state. As pointed out by Tovar (2009), due to the computational burden often associated with the likelihood evaluation for the solution of the nonlinear expectation equations implied by DSGE models, the empirical literature has concentrated its attention on the estimation of first-order linearized DSGE models. First-order approximations have been, until recently, the main tool employed for empirically evaluating DSGE models and for forecasting. However, as Judd (1997, p. 911) observes, “If theoretical physicists insisted on using only closed-form solutions or proofs of theorems to study their models, they would spend their time examining the hydrogen atom, universes with one star, and other highly simplified cases and ignore
most interesting applications of physical theories.”

The log-linearization approximation method has several important drawbacks. First, the solution methodology makes it impossible to model and study systemic risk. The most recent papers on modeling financial intermediaries, such as Brunnermeier and Sannikov (2014) and He and Krishnamurthy (2013), show that the nonlinearity of the amplification effect is a key aspect of systemic risk. Second, first-order approximations fail to be appropriate for evaluating welfare across policies that do not affect the steady state of the economy, e.g., when asset prices and the risk premium are taken into consideration. Log-linearization around a constant steady state is not applicable to asset pricing because, by construction, it eliminates all risk premia in the model. In fact, the risk premium is zero in a first-order approximation, and constant in the case of a second-order approximation, therefore higher-order approximations are required.17 Third, Fernández-Villaverde et al. (2006) consider log-linearization approximation to be unsatisfactory as they argue that second-order approximation errors in the solution of the model can have first-order effects on the likelihood function approximation. Ackerberg et al. (2009) made important asymptotic corrections to a theoretical result in Fernández-Villaverde et al. (2006), arguing that the approximation error on the classical maximum likelihood estimation of the approximate likelihood function has the same magnitude as the approximation error of equilibrium policy functions. When exact yet highly nonlinear policy functions are approximated by local linear ones, the likelihood implied by the linearized model can diverge greatly from that implied by the exact model, and similarly the likelihood-based point estimation.

**Estimation Methods**

Today most central banks have adopted Bayesian likelihood estimation methods instead of the more traditional equation-by-equation estimation used for large macro models. The main reasons are as follows. First, as shown in Canova (2009), the likelihood function of DSGE models is often flat and irregular in a number of parameters. Prior information helps overcome such identification issues. (However, there are 17See, for example, Schmitt-Grohé and Uribe (2004) and Kim et al. (2005), and An and Schorfheide (2007) for a discussion of second-order approximations.)
general issues on justifying the correct choice of priors, and it is dangerous to use too strong of a prior.) Second, the Bayesian approach can deal explicitly with measurement errors, unobservable state variables, large data sets, and different sources of information. Third, the Bayesian approach allows for decision making under uncertainty for policymakers. Fourth, although the Bayesian method is exposed to the “stochastic singular” problem that occurs when the number of variables is more than the number of the shocks, there are some useful techniques to tackle the problem.\textsuperscript{18} This can be viewed as an example of the models lying between the data-driven and structural ends of the model spectrum.

The other main reason macroeconomists at central banks resort to log-linearization approximation is to make estimation easier. Since Smets and Wouters (2003), the Bayesian estimation method has become the most popular estimation approach at central banks. However, as is well known, the standard Bayesian method requires full specification of the likelihood function of the model. This seems implausible for complex New Keynesian DSGE models without log-linear approximation. However, advanced Bayesian computing techniques such as the Approximate Bayesian Computing (ABC) method can be adopted. This method is able to work with the Dynare software platform which allows higher-order approximations of the model.\textsuperscript{19} Finally, it is important when using these methods to match the impulse response instead of only matching the moments of the model.

**Evaluation Methods**

The traditional method of evaluating DSGE models is to compare the simulated subset of moments with those observed in the data. More cautious researchers have conducted sensitivity analysis to check the fragility of the model. However, they mostly conduct this robustness check in an informal way, by perturbing parameters one by one and measuring the difference in model effects. The choices of the parameters and the magnitude of disturbance are \textit{ad hoc}. Chen et al. (2013) observe that even when the

\textsuperscript{18}See, for example, Harrison and Oomen (2010) where “structural shocks” were added into the baseline model to overcome the stochastic singular issue and improve the fitting of the data.

\textsuperscript{19} A simple example using the ABC method to estimate a dynamic macroeconomic finance model can be found in Chen et al. (2013).
model is stable in each parameter, it could be the case that the model is fragile in
a combination of multiple parameters. Zin (2002) points out that a primary goal of
characterizing asset market data using a tightly parameterized general equilibrium
model is to try to uncover deep structural parameters for policy purposes. However,
he emphasizes that it is not an easy task, mainly because the aggregate historical
data are usually not enough to provide an informative statistical test on the models
or their structural stability. To demonstrate this idea more explicitly, he discusses
a simple asset pricing model which cannot be rejected by data if the asset pricing
moment restrictions depend on high-order moments (e.g., the fifth moment) of the
distribution of the fundamental process (i.e., the endowment process). It is difficult to
visualize or even to describe high-order moments of fundamental processes, and hence
it is difficult to believe these asset pricing explanations would be deemed structural or
useful. However, these high-order moments could be macroeconomic “dark matter”,
as in Chen et al. (2013). The solution proposed in Zin (2002) is to augment the
statistical tests with subjective non-sample-based judgments about the reasonableness
of the assumptions. Chen et al. (2013) significantly improve Zin’s argument by
explicitly defining and quantitatively measuring dark matter, while Zin (2002) only
demonstrates the idea qualitatively. More importantly, Zin (2002) only focuses on the
“weak identification” side of dark matter, while missing the more important side to
dark matter: that it may cause model implications to become extremely sensitive to
parameters. While Zin (2002) and Chen et al. (2013) both stress the insufficiency of
current statistical tests for structural model evaluation, Chen et al. (2013) propose an
explicit, quantitative, and implementable method focus on “dark matter” or “fragility”
in models to augment conventional statistical specification tests for model evaluation.

5 Conclusion

The depth and length of the financial crisis of 2007–2009 has given new urgency and
relevance to macrofinancial economists around the world. Just as the Great Depression
and its aftermath inspired Tinbergen and Klein, and the recession and stagflation of
the 1970s inspired Lucas, Kydland, and Prescott, the current macroeconomic milieu
has prepared the way for a major shift in macroeconomic modeling for policy. Although this challenge seems daunting, it can also be viewed as an extraordinary opportunity to effect dramatic change in how we conduct macroeconomic policy. Three major themes seem to be emerging in what needs to be done.

The first theme is to take risk seriously in macroeconomic models and incorporate individual, institutional, and regulatory responses to changing risks—both actual and perceived—in them. Thanks to early attempts to model the macroeconomy, the critical field of national income accounting emerged, and transformed macroeconomics from armchair quarterbacking to a scientific endeavor with enormously practical implications. We now measure many aspects of the economy such as inflation, output, and unemployment, but we currently have no measure of aggregate risk in the economy. The old adage that one cannot manage what one does not measure is particularly relevant when it comes to risk in the macroeconomy. In the same spirit of Keynes’ hope that economic policy would some day be as effective and prosaic as going to the dentist, we can hope that systemic risk measurement would some day be as effective as hurricane forecasts and flood warnings issued by the National Weather Service.

The second theme is to incorporate the intricacies of the financial sector more effectively into existing DSGE models. Given the complexity of today’s financial system, this challenge may seem hopeless and naive. However, the very essence of macroeconomics is to distill complex phenomena into macroscopic narratives that can be grasped and managed by human cognition. Together with new technologies such as massive data sets, new computational and statistical methods, and social media, the potential for creating even greater information compression for macroeconomic policy decisions has never been more promising. The ability to measure business activity, leverage, inflation, and employment in real time at the level of the individual is close at hand, and the aggregation of such micro-level measures will surely transform macroeconomics.

The third theme is perhaps the most radical, which is to challenge the physics- and theory-based orthodoxy of macroeconomic modeling in reexamining the micro-foundations of the DSGE framework. To bring models closer to reality, it may be necessary to let go of the deeply cherished conviction that agents always optimize
their behavior according to rational expectations, and allow for certain predictable irrationalities in their behavior. These agents would still reflect the spirit of the Lucas critique by adapting to economic circumstance, but not necessarily in an instantaneously and fully optimal way. One of the positive aspects of financial crises may be to provide motivation for economists to revisit their assumptions of optimizing, forward-looking behavior and adjust them to reflect the realities of decision making in a complex, uncertain, and changing environment with limited information and cognitive abilities. This may also require us to abandon our predilection for simple models with elegant closed-form solutions in favor of less elegant but more practically relevant computational and numerical approaches to macroeconomic analysis.

When Albert Einstein was criticized for the complexity of his theory of relativity, he responded that “A theory should be as simple as possible, but no simpler.” The same can be said about the theories of the macroeconomy. We are discovering—as Keynes discovered over half a century ago—that from a policy perspective, being precisely wrong is not as useful as being approximately right.
A Frictionless Benchmark

A frictionless economy is used as a benchmark (1) for government policy in the larger model, (2) to calibrate the parameters, (3) as a starting point of the time-iteration algorithm, (4) as a check of the solution method for the full model, and (5) for economic analysis.

The equilibrium is the balanced growth path with stochastic trend $A_tK_t^\alpha$. Because $\tau > 0$, the government should not conduct credit policy in the frictionless economy, i.e. $\nu_g \equiv 0$.

The prices are

$$Q_t = qA_tK_t^{\alpha - 1} \quad \text{and} \quad P_t = pA_tK_t^{\alpha - 1}. \quad (53)$$

The optimal investment rate satisfies

$$q/p = 1 + \vartheta i. \quad (54)$$

The optimal wage is

$$W_t = (1 - \alpha)A_tK_t^\alpha L_c^\alpha. \quad (55)$$

Taking out the trend in $W_t$, the optimal normalized wage is

$$w = (1 - \alpha)\ell_c^{\alpha}, \quad (56)$$

where $\ell_c \equiv L_{c,t}$. Denote $\ell_t \equiv L_{i,t}$. The labor market clearing condition implies

$$\ell_c + \ell_i = 1. \quad (57)$$

The dividend is

$$D_t = \alpha A_tK_t^\alpha \ell_c^{1-\alpha} - P_t \left( iK_t + \frac{\vartheta}{2} i^2 K_t \right). \quad (58)$$
We characterize the equilibrium dividend and consumption as respectively

\[ D_t = dA_tK_t^\alpha \quad \text{and} \quad C_t = cA_tK_t^\alpha. \]  

(59)

The market clearing condition for consumption goods implies

\[ y = d + w\ell_c + w\ell_t. \]  

(60)

And the relationship (58) can be rewritten as

\[ d = \alpha \ell_c^{1-\alpha} - p \left( i + \frac{\vartheta}{2}i^2 \right). \]

The equilibrium resource constraint implies that the consumption goods output \( y \) and the household consumption \( c \) are

\[ y = \ell_c^{1-\alpha} \quad \text{and} \quad c = (1 - \varpi)\ell_c^{1-\alpha}. \]  

(61)

The zero-profit condition in the investment goods sector is

\[ W_tL_{t,t} = P_tK_tL_{t,t} \]  

(62)

which implies that \( w = p \), and thus (56) can be rewritten as

\[ p = (1 - \alpha)\ell_c^{-\alpha}. \]  

(63)

The market clearing condition for the investment goods is

\[ i + \frac{\vartheta}{2}i^2 = \ell_t. \]  

(64)

The IMRS of household is

\[ M_{t,t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} = \beta \left( \frac{A_{t+1}K_{t+1}^\alpha}{A_tK_t^\alpha} \right)^{-\gamma} = \beta e^{-\gamma\sigma_{\alpha,t+1} (i + 1 - \delta)}^{-\gamma\alpha} \]  

(65)
The equilibrium interest rate satisfies

$$1 = \mathbb{E}_t [M_{t,t+1}] (1 + R)$$

(66)

It implies that

$$\ln (1 + R) = -\ln \beta + \gamma \alpha \ln(i + 1 - \delta) - M(\gamma, \sigma_a),$$

(67)

where $M(\gamma, \sigma_a) \equiv \ln \left( \frac{1}{2} e^{-\gamma \sigma_a} + \frac{1}{2} e^{\gamma \sigma_a} \right)$. As in the standard Lucas-tree economy, there are three components to the equilibrium interest rate. The first is the time value captured by the subjective discount rate $-\ln \beta$. The second is the sensitivity to the growth of consumption $\gamma \times \alpha \ln(i + 1 - \delta)$. The third component is the precautionary saving motive term $M(\gamma, \sigma_a)$. The interest rate depends heavily on the growth rate $i + 1 - \delta$, where the investment rate $i$ is mainly governed by the adjustment cost coefficient $\vartheta$.

The equilibrium stock return is

$$1 + R_{k,t+1} = \frac{D_{t+1} + Q_{t+1}K_{t+2}}{Q_tK_{t+1}} = e^{\sigma_a \epsilon_{a,t+1} (i + 1 - \delta)^{\alpha - 1} \left( \frac{d}{q} + i + 1 - \delta \right)},$$

(68)

with log return

$$\ln (1 + R_{k,t+1}) = (\alpha - 1) \ln(i + 1 - \delta) + \ln \left( \frac{d}{q} + i + 1 - \delta \right) + \sigma_a \epsilon_{a,t+1}.$$ (69)

Therefore, the conditional expected log return is

$$\mathbb{E}_t [\ln (1 + R_{k,t+1})] = (\alpha - 1) \ln(i + 1 - \delta) + \ln \left( \frac{d}{q} + i + 1 - \delta \right).$$ (70)

The Euler equation for equity return is

$$0 = \ln \beta + (\alpha - 1 - \gamma \alpha) \ln(i + 1 - \delta) + \ln \left( \frac{d}{q} + i + 1 - \delta \right) + M_k(\gamma, \sigma_a),$$

(71)

where $M_k(\gamma, \sigma_a) \equiv \ln \left( \frac{1}{2} e^{(\gamma - 1)\sigma_a} + \frac{1}{2} e^{-(\gamma - 1)\sigma_a} \right)$. Combining (67), (69), and (71), the
equilibrium risk premium can be derived without solving for the other equilibrium variables. The equilibrium risk premium is

$$\Xi^* \equiv \mathbb{E}[\ln (1 + R_k^*)] - \ln (1 + R^*) = M(\gamma, \sigma_a) - M_k(\gamma, \sigma_a) \approx \gamma \sigma_a^2 - \frac{1}{2} \sigma_a^2. \quad (72)$$

This coincides with the equilibrium risk premium in the Lucas-tree economy with Jensen’s term $\frac{1}{2} \sigma_a^2$. Similarly, the risk premium is independent of the growth rate of the economy.

Other equilibrium variables $c^*, i^*, \ell^*, p^*, q^*$, and $d^*$ can be solved from the system of equations including (54), (57), (60), (61), (63), (64), and (71).

## B Frictional Economy

### B.1 Equilibrium Conditions

There is a leverage constraint for each intermediary. The Lagrangian multiplier is $\mu(\lambda, n)$ and is associated with the financial constraint. It is dual to the slackness in the balance sheet, and thus we define $\mu(\lambda, n) \equiv \max \{0, -\mu_b(\lambda, n)\}^2$. The slackness of the balance sheets of the intermediaries is

$$\Omega(\lambda, n)n - \lambda S_p(\lambda, n) = \max \{0, \mu_b(\lambda, n)\}^2\quad (73)$$

These two conditions imply that

$$\mu(\lambda, n) [\Omega(\lambda, n)n - \lambda S_p(\lambda, n)] = 0. \quad (74)$$

The resource constraint of the economy follows from (45) and (46) that

$$(1 - \bar{g})y(\lambda, n) = c(\lambda, n) + \tau S_g(\lambda, n)q(\lambda, n). \quad (75)$$

Importantly, the expectation correspondence characterizes how the next period’s aggregate normalized net worth $n'$ as an endogenous state variable depends on the current states $(\lambda, n)$ and exogenous state variables and shocks in the next period. The expectation correspondence is also an equilibrium to be solved:

$$n' = \Gamma(\lambda, n; \lambda', \epsilon_a'). \quad (76)$$
Following (37), the expectation correspondence can be expressed as

\[ n' = \Gamma(\lambda, n; \lambda', \epsilon'_a) \]

\[ = \frac{\left\{ G_r(\lambda, n; \lambda', \epsilon'_a) - [1 + R_f(\lambda, n)] \right\} S_p(\lambda, n) + \kappa + [R_f(\lambda, n) + \theta] n}{G_k(\lambda, n; \lambda', \epsilon'_a)}. \]

Here \( G_r(\lambda, n; \lambda', \epsilon'_a) \equiv 1 + R_k(\lambda, n; \lambda', \epsilon'_a) \) is the total stock return, and \( G_k(\lambda, n; \lambda', \epsilon'_a) \equiv Q'K'/(QK) \) is the total return of capital gain on stocks whose expression can be found in (78). The stock return described in (12) can be rewritten as

\[ 1 + R_k(\lambda, n; \lambda', \epsilon'_a) = \frac{d(\lambda', n')}{q(\lambda, n)} e^{\sigma_a \epsilon'_a} [i(\lambda, n) + 1 - \delta]^{\alpha-1} + G_k(\lambda, n; \lambda', \epsilon'_a) \] (77)

where the capital gain return is

\[ G_k(\lambda, n; \lambda', \epsilon'_a) = \frac{q(\lambda', n')}{q(\lambda, n)} e^{\sigma_a \epsilon'_a} [i(\lambda, n) + 1 - \delta]^{\alpha-1} [i(\lambda', n') + 1 - \delta] \] (78)

According to (33), the Euler equation for stock returns is

\[ \Omega(\lambda, n)c(\lambda, n)^{-\gamma} - \mu(\lambda, n)c(\lambda, n)^{-\gamma} [\Omega(\lambda, n) - \lambda] \]

\[ = \beta [i(\lambda, n) + 1 - \delta]^{-\gamma \alpha} E \left\{ c(\lambda', n')^{-\gamma} e^{-\gamma \sigma_a \epsilon'_a} [1 - \theta + \theta \Omega(\lambda', n')] G_r(\lambda, n; \lambda', \epsilon'_a) \bigg| \lambda, n \right\} \]

Define

\[ \tilde{G}_r(\lambda, n; \lambda', \epsilon'_a) = G_r(\lambda, n; \lambda', \epsilon'_a)/q(\lambda, n). \] (79)

To obtain the Euler equation above, the key intermediate step is

\[ \frac{\Lambda'}{\Lambda} = \beta \left[ \frac{c(\lambda', n')}{c(\lambda, n)} \right]^{-\gamma} e^{-\gamma \sigma_a \epsilon'_a} [i(\lambda, n) + 1 - \delta]^{-\gamma \alpha}. \] (80)
According to (32), the intermediary Euler equation for the risk-free rate is

\[
[1 - \mu(\lambda, n)] \Omega(\lambda, n)c(\lambda, n)^{-\gamma} = \beta \left[ i(\lambda, n) + 1 - \delta \right]^{-\gamma \alpha} [1 + R_f(\lambda, n)] \mathbb{E} \left\{ c(\lambda', n')^{-\gamma} e^{-\gamma \sigma \xi' a} \left[ 1 - \theta + \theta \Omega(\lambda', n') \right] \big| \lambda, n \right\}
\]

The household Euler equation for the risk-free rate is

\[
1 = \beta \left[ i(\lambda, n) + 1 - \delta \right]^{-\gamma \alpha} [1 + R_f(\lambda, n)] \mathbb{E} \left\{ e^{-\gamma \sigma \xi' a} \big| \lambda, n \right\}
\]  
(81)

The investment goods production is

\[
u(\lambda, n) = \ell_i(\lambda, n).
\]  
(82)

The investment goods sector market clearing condition is

\[
u(\lambda, n) = i(\lambda, n) + \frac{\sigma}{2} i(\lambda, n)^2.
\]  
(83)

And the labor market clearing condition is

\[
\ell_c(\lambda, n) + \ell_i(\lambda, n) = 1.
\]  
(84)

The zero-profit condition for investment good firms is

\[
w(\lambda, n) = p(\lambda, n).
\]  
(85)

The optimal demand of labor in consumption goods sector is

\[
w(\lambda, n) = (1 - \alpha) \ell_c(\lambda, n)^{-\alpha}.
\]  
(86)

The total dividend paid out from the consumption goods sector is

\[
d(\lambda, n) = \alpha \ell_c(\lambda, n)^{1-\alpha} - p(\lambda, n) u(\lambda, n),
\]  
(87)
where $\alpha c(\lambda, n)^{1-\alpha}$ is the total consumption goods minus the labor cost and $p(\lambda, n)u(\lambda, n)$ is the expenditure of purchasing investment goods. The optimal investment decision is characterized by the traditional $q$ theory relationship:

$$q(\lambda, n) = [1 + \vartheta_i(\lambda, n)] p(\lambda, n).$$  (88)

The consumption goods production is

$$y(\lambda, n) = \ell c(\lambda, n)^{1-\alpha}.$$  (89)

The log risk premium $\Xi(\lambda, n)$ is defined as

$$\Xi(\lambda, n) + \log [1 + R_f(\lambda, n)] = E_t \{ G_r(\lambda, n; \lambda', \epsilon'_a) | \lambda, n \}. $$  (90)

The credit policy can be written as

$$S_p(\lambda, n) [1 + \nu_g(\Xi(\lambda, n) - \Xi^*)] = 1,$$

where $\nu_g = \nu_{g,0} + \nu_{g,1} \left( \frac{1}{n_t} - 1 \right)$ is state-dependent.

### C The Full Benchmark DSGE Model

The purpose of this section is to provide an extension of a benchmark New Keynesian DSGE model, incorporating financial intermediation, a role for asset markets and unconventional monetary policy. In light of the financial crisis of 2008, such extensions appear to be in particularly high demand. The model here is kept simple on purpose, and can serve as a base for constructing more elaborate versions. We view this model as a multi-purpose tool rather than a model designed to answer a specific question at hand. Our model is a simplified version of Christiano et al. (2010) and Christiano et al. (2014), which are state-of-the-art New Keynesian DSGE models with nontrivial financial intermediation. This model is meant for illustrative purposes, and we have left out many exogenous shocks that would be studied in a full-scale DSGE model.

The New Keynesian component of the model is a simplified version based on Christiano et al. (2005) and Smets and Wouters (2003). A simple New Keynesian DSGE model featuring monopolistic competitive firms and rigid nominal prices without endogenous capital accumulation can be found in Gali (2008). Kollman (1997) and Erceg et al. (2000) both introduced nominal sticky wages which are adjusted according
to the Calvo rule (see Calvo, 1983). In terms of endogenous capital accumulation, we follow Christiano et al. (2005), which incorporates endogenous capital accumulation with an adjustment cost characterized by the relative level of investment, rather than the investment-capital stock ratio, is commonly assumed in the RBC literature. We adopt their external habit formation in consumption, which helps generate persistence in the consumption process in the data.

The financial intermediation component in our model is based on Gertler and Karadi (2011) and Gertler and Kiyotaki (2010). We allow for time-varying rare disaster risk and its premia (see, e.g. Gourio, 2012). We also incorporate “capital quality shocks” similar to Smets and Wouters (2003) and Christiano et al. (2005) in a New Keynesian setting and Gourio (2012) in an asset pricing setting, inspired by Greenwood et al. (1988) and King and Rebelo (1999) in the RBC literature.

In this simplified but recognizable version of the real economy, households maximize their individual utility function with consumption and labor over an infinite horizon. The utility function is characterized by external habit formation. The habits depend on lagged aggregate consumption that is unaffected by any single household’s decision. Abel (1990) calls this the “catching up with the Joneses” effect. For simplicity, we assume households face flexible nominal wages.

The members of each household are divided into bankers and workers. Bankers and workers can both supply labor, but only bankers own capital in this economy and rent capital to the intermediate goods firm to extract rents. In this way, bankers decide how much capital to accumulate given the capital adjustment costs. A potentially important deficiency of the model is that the role bankers is hard-wired into the model: there is no other way for provide capital finance. This side-steps potentially important opportunities for flexibility in funding sources, an issue discussed more substantially in de Fiore and Uhlig (2011) and de Fiore and Uhlig (2015), for example.

The representative intermediate goods firm produces one kind of intermediate good, which can be used for investing and creating capital goods, or sold wholesale to retailers who simply convert the intermediate goods into differentiated goods for consumption. The retailers produce differentiated goods, giving them some monopoly power over goods prices, with a downward sloping demand for goods from households. The intermediate goods firm decides on labor and capital inputs, and at the same time, the retailers re-optimize goods prices according to the Calvo rule. Eventually, under the assumption that the bankers and the workers can fully insure their idiosyncratic risks in consumption, the representative households have the full claim on the dividends paid out by the intermediate goods firm and the retailers.

It should be noted that in benchmark New Keynesian DSGE models such as Christiano et al. (2005) and Smets and Wouters (2003), the financial and credit markets play no role in determining asset prices except for the term structure of real interest rates and expectations of future payouts, and have no impact on the real economy, beyond the repercussions of shocks added to the Euler equation. An equivalent statement is that these models adopt the assumptions underlying the Modigliani and Miller (1958) Theorem, which implies that financial structure is both
indeterminate and irrelevant to real economic outcomes. In order to quantitatively study how credit market imperfections influence the transmission of monetary policy, Bernanke et al. (1999) incorporated a countercyclical credit-market friction into a dynamic model, which is endogenously generated from first principles (i.e., using agent optimization). This countercyclical credit-market friction, first emphasized in Kiyotaki and Moore (1997), is shown to amplify and propagate productivity shocks. In more recent work, Christiano et al. (2010) extend the model in Bernanke et al. (1999) in many dimensions, including financially-constrained intermediations, but the key financial and credit market imperfections are not far removed from the financial accelerator mechanism in Bernanke et al. (1999). The crucial feature of constrained financial intermediation has been formulated by Gertler and Karadi (2011) in a simple but transparent model, on which we build here. A drawback of this approach is that the particular friction is hard-wired into the model, with little scope for agents in the model to get around them. By contrast, real-life participants on financial markets often appear to be particularly creative in circumventing restrictions.

Our goal here is to understand and illustrate the effects of three shocks in this model economy: (1) a financial sector shock impairing the ability of banks to borrow and hold productive assets, (2) a technology shock that impairs the quality of the physical capital, and (3) a risk shock. We also desire to show how the economy would respond to various policy responses to these shocks.

C.1 Households

We begin with a description of households in our canonical model. There is a continuum of households of unit mass. The members of each household are either workers or bankers. Though there are two groups of agents, and certain portfolio constraints among them, we assume the representative framework following Gertler and Karadi (2011) and Gertler and Kiyotaki (2010) by assuming the household members to be part of a large family, sharing everything or, equivalently, assuming that the full set of Arrow-Debreu securities are available to the members within each household (but not across households), so that the idiosyncratic consumption risks can be fully insured, and the agents in two groups have identical preferences. At any time, a fraction $f$ of the members of the household are bankers. Bankers live for a finite number of periods with probability 1. At any time, a fraction $1 - \theta$ of randomly selected existing bankers exit and become workers, and return their net worth to their household. At the same time, an equal number of workers become bankers within each household, so the proportion of workers and bankers remains fixed. The new bankers receive some start-up funds from their household, which we describe below. The “perpetual youth” assumption in our model is purely technical, with the purpose of guaranteeing the survivorship of both groups of agents.
The preferences of the household are given by

$$
\mathbb{E}_t \left[ \sum_{\tau=0}^{\infty} \beta_t^{t+\tau} \left( \frac{(C_{t+\tau} - hC_{t+\tau-1})^{1-\gamma}}{1-\gamma} - \frac{\chi}{1+\varphi} L_t^{1+\varphi} \right) \right],
$$

where $C_t$ is the consumption and $L_t$ is the labor supply at time $t$. $\varphi$ is the Frisch elasticity of labor supply and it is positive. The subject discount rate $\beta \in (0, 1)$ and habit parameter $h \in (0, 1)$. Also, we assume that $\chi > 0$. The logarithm of the discount factor $\beta_t$ follows an AR(1) process

$$
\log \beta_t = (1 - \rho_{\beta}) \log \beta_{ss} + \rho_{\beta} \log \beta_{t-1} + \sigma_{\beta} u_{\beta,t},
$$

where $\beta_{ss}$ is the long term mean, and $u_{\beta,t} \sim$ i.i.d. $N(0, 1)$.

Both bankers and workers within each household can hold nominally risk-free debt issued by the government, and can deposit its cash with a financial intermediary that pays a nominally risk-free rate. Assuming that both assets are perfect substitutes, we denote by $R_{f,t+1}$ the real gross interest rate paid by either of these assets. It should be noted that $R_{f,t+1}$ is possibly random up to the information set at time $t$ because the debt contract is written on the nominal term and the inflation, $\Pi_{t+1}$, is random up to the information set at time $t$. Let $B_{t+1}$ denote the quantity of this debt held by the household at the end of period $t$. The household then faces a state-by-state budget constraint

$$
C_t = W_t L_t + Pr_t + T_t + R_{f,t} B_t - B_{t+1},
$$

where $W_t$ is the real wage, $Pr_t$ is the profits from the various firms the household owns (which we describe below), and $T_t$ is the real lump-sum taxes. The first-order conditions to the household’s utility maximization problem include the intertemporal Euler equation for working hours

$$
\Lambda_t = \chi \frac{L_t^\varphi}{W_t},
$$

and the intertemporal Euler equation for risk-free bond holding

$$
1 = \mathbb{E}_t \left[ \beta \frac{\Lambda_{t+1}}{\Lambda_t} R_{f,t+1} \right],
$$
where
\[
\Lambda_t \equiv (C_t - hC_{t-1})^{-\gamma} - \beta h \mathbb{E}_t (C_{t+1} - hC_t)^{-\gamma}.
\] (95)

is the marginal utility of consumption \( C_t \) at date \( t \).

**C.2 Financial Intermediaries**

Financial intermediaries borrow funds from households at a risk-free nominal rate, pool this with their own net worth or wealth and invest the sum in the equity of the representative intermediate goods firm. We describe the intermediary using real variables in what follows. The balance sheet of intermediary \( j \) at the end of time \( t \) is given by

\[
Q_t S_{j,t} = N_{j,t} + B_{j,t+1},
\] (96)

where \( Q_t \) is the price of the intermediate goods firm’s equity, \( S_{j,t} \) is the quantity of equity held by the intermediary, \( N_{j,t} \) is the net worth, and \( B_{j,t+1} \) is the deposits raised from households. The intermediary earns a gross return \( R_{k,t+1} \) from the equity investment at time \( t + 1 \), and must pay the gross interest, \( R_{f,t+1} \), on the deposit. The net worth of the intermediary, therefore, evolves as

\[
N_{j,t+1} = R_{k,t+1} Q_t S_{j,t} - R_{f,t+1} B_{j,t+1} = (R_{k,t+1} - R_{f,t+1}) Q_t S_{j,t} + R_{f,t+1} N_{j,t}.
\]

The intermediaries face a constraint on raising deposits from households. They cannot raise deposits beyond a certain level, which is determined endogenously in the equilibrium. We shall describe this constraint in more detail below. Since the bankers own the intermediaries, we use the bankers’ stochastic discount factor (SDF), which coincides with the SDF of the representative agent, \( \beta^i \Lambda_{t+\tau}/\Lambda_t \), to compute the value of assets to the intermediary. The presence of the borrowing constraints implies

\[
\mathbb{E}_t \left[ \beta^\tau \frac{\Lambda_{t+\tau+1}}{\Lambda_t} (R_{k,t+\tau+1} - R_{f,t+\tau+1}) \right] \geq 0, \quad \forall \tau \geq 0,
\] (97)

with equality if and only if the intermediary faces no borrowing constraint. Note that, so far, the returns and SDF are all real.

Since the intermediary ceases being a banker each period with probability \( 1 - \theta \),
the value of intermediary \( j \)'s terminal wealth to its household is given by

\[
V_{j,t} = \max_{\{S_{j,t+\tau}, B_{j,t+\tau+1}\}_{\tau \geq 0}} \mathbb{E}_{t} \left[ \sum_{\tau=0}^{\infty} (1 - \theta) \theta^{\tau+1} \frac{\Lambda_{t+\tau+1}}{\Lambda_{t}} N_{j,t+\tau+1} \right]
\]

\[
= \max_{\{S_{j,t+\tau}, B_{j,t+\tau+1}\}_{\tau \geq 0}} \mathbb{E}_{t} \left[ \sum_{i=0}^{\infty} (1 - \theta) \theta^{\tau+1} \frac{\Lambda_{t+\tau+1}}{\Lambda_{t}} \times \right.
\]

\[
\left. [(R_{k,t+\tau+1} - R_{t+\tau+1})Q_{t+\tau}S_{j,t+\tau} + R_{f,t+\tau+1}N_{j,t+\tau}] \right].
\]

(98)

In order to motivate the borrowing constraint faced by financial intermediaries and following Gertler and Karadi (2011) and Gertler and Kiyotaki (2010), we introduce a simple moral hazard/costly enforcement problem. We assume that the banker can choose to liquidate the financial intermediation and divert the fraction of available funds, \( \lambda_{t} \), from the value of the financial intermediation.

The borrowing constraint is modeled as follows. At any time \( t \), the manager of the intermediary can divert a fraction \( \lambda_{t} \) of the intermediary’s assets to his household for his own benefit, where \( \lambda_{t} \in [0, 1] \) is an exogenous parameter. To make sure that \( \lambda_{t} \) is between zero and one, we use the transformation

\[
\lambda_{t} = \frac{1}{1 + \tilde{\lambda}_{t}}.
\]

The logarithm of \( \tilde{\lambda}_{t} \) follows an AR(1) process

\[
\log \tilde{\lambda}_{t} = (1 - \rho_{\tilde{\lambda}}) \log \tilde{\lambda}_{ss} + \rho_{\tilde{\lambda}} \log \tilde{\lambda}_{t-1} + \sigma_{\tilde{\lambda}} u_{\tilde{\lambda},t},
\]

(99)

where \( \tilde{\lambda}_{ss} \) is the long term mean, and \( u_{\tilde{\lambda},t} \sim \text{i.i.d.} \ N(0, 1) \). If the value of the intermediary falls below \( \lambda_{t}Q_{t}S_{j,t} \), the intermediary will simply divert the assets and the households will get a zero gross return from their deposits. In order for the households to have an incentive to deposit cash with the intermediary, the following condition must hold:

\[
V_{j,t} \geq \lambda_{t}Q_{t}S_{j,t}.
\]

(100)

We conjecture that the value of the intermediary is linear in its net worth and the
value of the assets it holds:

\[ V_{j,t} = \nu_t Q_t S_{j,t} + \eta_t N_{j,t}. \]  

(101)

for some \( \nu_t \) and \( \eta_t \). We see that the incentive constraint binds only if \( 0 < \nu_t < \lambda_t \), otherwise the marginal value to the intermediary of increasing the assets is larger than the marginal value of diverting them, and the intermediary has an incentive to increase its assets. As in the equilibrium in Gertler and Karadi (2011), we assume that the incentive constraint (100) always binds in the local region of the long run mean, \( \lambda_{ss} \), of \( \lambda_t \). When the constraint binds, we have the condition

\[ \nu_t Q_t S_{j,t} + \eta_t N_{j,t} = \lambda_t Q_t S_{j,t}, \quad \text{or} \quad \frac{\eta_t}{\lambda_t - \nu_t} N_{j,t} = \phi_t N_{j,t}. \]  

(102)

\[ Q_t S_{j,t} = \frac{\eta_t}{\lambda_t - \nu_t} N_{j,t} = \phi_t N_{j,t}. \]  

(103)

Using the definition of \( \phi_t \), we can rewrite the evolution of the intermediary’s net worth as

\[ N_{j,t+1} = N_{j,t} [(R_{k,t+1} - R_{f,t+1})\phi_t + R_{f,t+1}]. \]  

(104)

This is the standard wealth or net worth law of motion with leverage, where \( \phi_t \) can be viewed as the share of net worth invested in the risky asset (i.e., equity).

We verify the guess to the solution of the value function of the intermediary when the incentive constraint binds, and obtain

\[ \nu_t = \mathbb{E}_t \left[ (1 - \theta)\beta \frac{\Lambda_{t+1}}{\Lambda_t} (R_{k,t+1} - R_{f,t+1}) + \theta \beta \frac{\Lambda_{t+1}}{\Lambda_t} \frac{\phi_t}{\phi_{t+1}} \nu_{t+1} ((R_{k,t+1} - R_{f,t+1})\phi_t + R_{f,t+1}) \right] \]  

(105)

and

\[ \eta_t = \mathbb{E}_t \left[ 1 - \theta + \theta \beta \frac{\Lambda_{t+1}}{\Lambda_t} \eta_{t+1} ((R_{k,t+1} - R_{f,t+1})\phi_t + R_{f,t+1}) \right]. \]  

(106)

Since \( Q_t S_{j,t} = \phi_t N_{j,t} \), and since \( \phi_t \) does not depend on intermediary-specific factors,
we can aggregate over the equation to get

\[ Q_t S_t = \phi_t N_t, \quad (107) \]

where \( S_t \) is the aggregate investment in the equity and \( N_t \) is the aggregate wealth of the intermediaries.

Finally, we determine the evolution of the aggregate net worth of the intermediaries. The aggregate net worth is the sum of the net worth of the existing intermediaries, \( N_{n,t} \), and the net worth of the new entrants, \( N_{e,t} \):

\[ N_t = N_{e,t} + N_{n,t}. \quad (108) \]

Since a fraction \( \theta \) of the bankers from \( t - 1 \) survive up to \( t \), we have

\[ N_{e,t} = \theta N_{t-1} [(R_{k,t} - R_{f,t})\phi_{t-1} + R_{f,t}]. \quad (109) \]

The new entrants receive funds from the households to “start up”. As in Gertler and Karadi (2011), we assume that each entering intermediary receives a fraction \( \omega \frac{1}{1-\theta} \) of the value of the final period assets of the exiting intermediaries, which is \( (1-\theta)Q_t S_{t-1} \), the remainder being distributed to the households. This gives

\[ N_{n,t} = \omega Q_t S_{t-1}. \]

Thus, we have the evolution of the aggregate net worth

\[ N_t = \theta N_{t-1} [(R_{k,t} - R_{f,t})\phi_{t-1} + R_{f,t}] + \omega Q_t S_{t-1}. \quad (110) \]

### C.3 Firms

There are three types of firms in our model: intermediate-goods, capital-producing, and retail firms.

#### Intermediate-Goods Firms

The financial intermediaries invest in the equity of the intermediate-goods firm. We assume that the representative intermediate-goods firm produces the intermediate goods used by the retail firms to produce differentiated goods. Such a firm has no
Wealth of its own. It is alive for two periods, \(t - 1\) and \(t\), say. In period \(t - 1\), it issues shares to banks and uses the proceeds of the investment by the intermediaries to purchase capital \(K_t\) from the capital producing firm, to be used for production in the next period\(^{20}\). The number of shares issued by the intermediate goods firm is equal to the number of units of capital purchased, and we assume no frictions or wedges (such as those induced by an agency problem between the financial firms and the managers of the intermediate goods firms), so that

\[
Q_{t-1}S_{t-1} = Q_{t-1}K_t.
\]  

or, stated as an equation for period \(t\),

\[
Q_tS_t = Q_tK_{t+1}.
\]  

The intermediate firm faces no informational or incentive problems. In period \(t\), it hires labor at a wage \(W_t\) and uses the capital \(K_t\) chosen in period \(t - 1\) to produce the intermediate goods using the production function

\[
Y^I_t = A_t(\xi_tK_t)^\alpha L_t^{1-\alpha},
\]

where \(A_t\) is an exogenous and stochastic total factor productivity (TFP) parameter, \(\xi_t\) is an exogenous and stochastic quality of capital parameter, and \(K_t\) is the capital stock determined at period \(t - 1\). The TFP and the capital quality evolve as

\[
\log A_t = (1 - \rho_A) \log A_{ss} + \rho_A \log A_{t-1} + \sigma_A u_{A,t},
\]

\[
\log \xi_t = (1 - \rho_\xi) \log \xi_{ss} + \rho_\xi \log \xi_{t-1} + \sigma_\xi u_{\xi,t},
\]

where the subscripts \(ss\) denote the long term steady states, and \(u_{A,t}, u_{\xi,t} \sim \text{i.i.d } N(0,1)\). After production takes place, the firm sells the quality-adjusted and depreciated capital back to the capital-producing firms at a price of \(Q_t\) per unit. Thus, the proceeds from the capital sale are \(Q_t(1 - \delta)\xi_tK_t\), where \(\delta\) is the depreciation rate.

Denote by \(P_{m,t}\) the market price of the output of the intermediate goods, taken as given by the firm. Arriving in period \(t\) and given the amount of capital \(K_t\) chosen in \(t - 1\), the firm — or, perhaps better, the firm-owning banker — will choose labor \(L_t\)

\(\text{Note that we adopt the more conventional timing assumption of indexing capital with the date when it is used in production, not with the date of the decision. As is well known, one needs to be careful with that, when implementing this in solution software such as Dynare or Uhlig’s Toolkit, see Uhlig (1999).}\)
to maximize the cash flow

\[ cf_t = P_{m,t}A_t(\xi_t K_t)^\alpha L_t^{1-\alpha} + Q_t(1 - \delta)\xi_t K_t - W_t L_t \tag{116} \]

The first-order condition with respect to labor gives

\[ P_{m,t}(1 - \alpha)\frac{Y^I_t}{L_t} = W_t. \tag{117} \]

Since the intermediate firm has no wealth of its own and is owned entirely by its banks, it pays out the cash flow \( cf_t \) in proportion to the underlying shares or units of capital to the banks. The return per unit of resources for a bank investing at date \( t \) is therefore

\[ R_{k,t+1} = \frac{P_{m,t+1}Y^I_{t+1} - W_{t+1}L_{t+1} + Q_{t+1}(1 - \delta)\xi_{t+1}K_{t+1}}{Q_tK_{t+1}}, \tag{118} \]

where \( \delta \) is the depreciation rate. Plugging in the first-order condition for labor, we get

\[ R_{k,t+1} = \frac{P_{m,t+1}Y^I_{t+1} + Q_{t+1}(1 - \delta)\xi_{t+1}K_{t+1}}{Q_tK_{t+1}}. \tag{119} \]

**Capital Producing Firms**

We assume that the representative capital-producing firm purchases the depreciated capital \( (1 - \delta)\xi_t K_t \) from the intermediate goods firm at the end of every period. It also produces new capital or, possibly, turns existing old capital back into consumption goods, denoted as gross capital investment \( I_t \). It then sells the total new capital to the new generation of intermediate goods firms at price \( Q_t \) per unit. The aggregate production function for new capital is

\[ K_{t+1} = I_t + (1 - \delta)\xi_t K_t. \tag{120} \]
We state the production function for new capital or gross capital investment as a resource cost function per unit of investment,

\[ x_t = z_t \left( I_t + \Phi \left( \frac{I_t}{\xi_t K_t} \right) \xi_t K_t \right) \]  

(121)

where \( x_t \) is in units of the aggregate consumption good, \( z_t \) is an exogenous and stochastic investment cost parameter, and \( \Phi(\cdot) \) is often called an adjustment cost function. The function \( \Phi \) is assumed to satisfy \( \Phi(\delta) = \Phi'(\delta) = 0 \) and \( \Phi''(\delta) > 0 \). Specifically, we assume the quadratic adjustment cost function

\[ \Phi(x) \equiv \frac{\vartheta}{2} (x - \delta)^2, \text{ with } \vartheta > 0. \]  

(122)

The profits of the capital producing firms in period \( t \) are

\[ Q_t(K_{t+1} - (1 - \delta) \xi_t K_t) - x_t = Q_t I_t - z_t \left( I_t + \Phi \left( \frac{I_t}{\xi_t K_t} \right) \xi_t K_t \right) \]  

(123)

We assume that the capital producing firms are directly owned by the households, and that there is free entry. Thus, the discounted value of the profits of the capital goods producer is

\[ \max_{\{I_t\}_{t \geq 0}} \sum_{\tau = t}^{\infty} \mathbb{E}_t \left[ \beta^{\tau - t} \frac{\Lambda_\tau}{\Lambda_t} \left( Q_\tau I_\tau - z_\tau I_\tau - z_\tau \Phi \left( \frac{I_\tau}{\xi_\tau K_\tau} \right) \xi_\tau K_\tau \right) \right]. \]  

(124)

Maximising this gives the first-order condition

\[ Q_t = z_t (1 + \Phi'(.)). \]  

(125)

(126)

The exogenous and stochastic investment parameter evolves as

\[ \log z_t = \rho_z \log z_{t-1} + \sigma_z u_{z,t}, \]  

(127)

with \( u_{z,t} \sim \text{i.i.d. } N(0,1) \).
Retail Firms

Our model includes a unit-mass continuum of retail firms. Retail firms use the (single) intermediate good to produce differentiated goods. Each firm $f$ can use one unit of the intermediate good to produce one unit of the differentiated good. If $Y_{f,t}$ denotes the final output of the retail firm, the final output composite good is the constant elasticity of substitution (CES) aggregator

$$Y_t = \left[ \int_0^1 Y_{f,t}^{(\varepsilon_t-1)/\varepsilon_t} \, df \right]^{\varepsilon_t/\varepsilon_t-1},$$  \hspace{1cm} (128)

where $\varepsilon_t > 1$ is the elasticity of substitution. The steady state of the elasticity of substitution is $\varepsilon_{ss}$. Define $\bar{\varepsilon} \equiv \varepsilon - 1$, which evolves as

$$\log \bar{\varepsilon}_t = (1 - \rho \bar{\varepsilon}) \log \bar{\varepsilon}_{ss} + \rho \bar{\varepsilon} \log \bar{\varepsilon}_{t-1} + \sigma \bar{\varepsilon} u_{\bar{\varepsilon},t},$$  \hspace{1cm} (129)

with $u_{\bar{\varepsilon},t} \sim$ i.i.d. $N(0,1)$.

The retail firms are monopolistically competitive. Equation (128) and household optimization implies that retail firms face a downward sloping demand for their goods. More precisely, if $P_{f,t}$ is the nominal price that each retail firm charges for its good, then cost minimization for the users of the final good gives

$$Y_{f,t} = \left( \frac{P_{f,t}}{P_t} \right)^{-\varepsilon_t} Y_t,$$  \hspace{1cm} (130)

where $P_t$ is the ideal nominal price index

$$P_t = \left[ \int_0^1 P_{f,t}^{1-\varepsilon_t} \, df \right]^{1/(1-\varepsilon_t)},$$  \hspace{1cm} (131)

We assume that, at each time $t$, a random fraction $1 - \varsigma$ of the retail firms can reset their price, while all other firms are stuck at the price from the previous period\textsuperscript{21}. Retail firms are owned by the households, who instruct these firms to maximize the net present value of future cash flows, using the household discount factor. A retail firm which can reset its price at time $t$ to $P_{f,t} = P_t^*$ therefore sets it to maximize the

\textsuperscript{21}Often, this framework is extended to allow for price indexation for these firms, thereby largely reducing the effect of average inflation on the equilibrium. To keep matters simple, we do not include indexation here.

118
discounted value of its future profits as long as it is stuck with that price, i.e.,

$$\max_{P_t^*} \mathbb{E}_t \left[ \sum_{\tau=0}^{\infty} \beta^\tau \zeta^\tau \Lambda_{t+\tau} \left( \frac{P_t^*}{P_{t+\tau}} Y_{f,t+\tau} - P_{m,t+\tau} Y_{f,t+\tau} \right) \right] \equiv (132)$$

$$\max_{P_t^*} \mathbb{E}_t \left[ \sum_{\tau=0}^{\infty} \beta^\tau \zeta^\tau \Lambda_{t+\tau} \left( \frac{P_t^*}{P_{t+\tau}} \left( \frac{P_t^*}{P_{t+\tau}} \right)^{-\epsilon_t} - P_{m,t+\tau} \left( \frac{P_t^*}{P_{t+\tau}} \right)^{-\epsilon_t} \right) Y_{t+\tau} \right].$$

where it is useful to note that $P_{m,t}$ is the real price for the intermediate input good or real cost per unit of retail output, and where the second line follows, because $Y_{f,t+\tau}$ is a function of $P_{f,t+\tau} = P_t^*$ per (130). The first-order condition with respect to $P_t^*$ is:

$$\mathbb{E}_t \left[ \sum_{\tau=0}^{\infty} \beta^\tau \zeta^\tau \Lambda_{t+\tau} \left( \frac{P_t}{P_{t+\tau}} \right)^{1-\epsilon_t} Y_{t+\tau} \right] \frac{P_t^*}{P_t} - \mathbb{E}_t \left[ \beta^\tau \zeta^\tau \Lambda_{t+\tau} \frac{\epsilon_t}{\epsilon_t - 1} P_{m,t+\tau} \left( \frac{P_t}{P_{t+\tau}} \right)^{-\epsilon_t} Y_{t+\tau} \right] = 0.$$

Defining

$$J_t = \mathbb{E}_t \left[ \sum_{\tau=0}^{\infty} \beta^\tau \zeta^\tau \Lambda_{t+\tau} \left( \frac{P_t}{P_{t+\tau}} \right)^{1-\epsilon_t} Y_{t+\tau} \right]$$

$$H_t = \mathbb{E}_t \left[ \sum_{\tau=0}^{\infty} \beta^\tau \zeta^\tau \Lambda_{t+\tau} \frac{\epsilon_t}{\epsilon_t - 1} P_{m,t+\tau} \left( \frac{P_t}{P_{t+\tau}} \right)^{-\epsilon_t} Y_{t+\tau} \right]$$

we have:

$$\frac{P_t^*}{P_t} = H_t / J_t.$$

At any time $t$ a fraction $\zeta$ of the retail firms will be unable to change their prices and their aggregate price index will be simply $P_{t-1}$. A fraction $1 - \zeta$ will be able to reset the prices, and they will all reset their price to $P_t^*$. Thus, by the law of large numbers, we have $P_{t-1}^{1-\epsilon_t} = \zeta P_{t-1}^{1-\epsilon_t} + (1 - \zeta) P_t^{1-\epsilon_t}$. Rearranging, and substituting for $P_t^*$ from
the first-order condition, we get

\[
\frac{1 - \varsigma \Pi_t^{\varepsilon_t-1}}{1 - \varsigma} = \left( \frac{H_t}{J_t} \right)^{1-\varepsilon_t},
\]

(133)

where \( \Pi_{t+1} = P_{t+1}/P_t \) is the inflation. We can rewrite \( J_t \) and \( H_t \) recursively as

\[
J_t = Y_t + \varsigma \beta \mathbb{E}_t \left[ \frac{\Lambda_{t+1}}{\Lambda_t} \Pi_{t+1}^{\varepsilon_t-1} J_{t+1} \right],
\]

(134)

\[
H_t = \frac{\varepsilon_t}{\varepsilon_t - 1} P_{m,t} Y_t + \varsigma \beta \mathbb{E}_t \left[ \frac{\Lambda_{t+1}}{\Lambda_t} \Pi_{t+1}^{\varepsilon_t} H_{t+1} \right].
\]

(135)

The relationship between the intermediate goods output and the final goods output is:

\[
Y_t^I = \int_0^1 Y_{f,t} df = \frac{Y_t}{P_{t-\varepsilon_t}^{\varepsilon_t}} \int_0^1 P_{f,t}^{-\varepsilon_t} df = \frac{Y_t}{P_{t-\varepsilon_t}^{\varepsilon_t}} \left[ (1 - \varsigma)(P_t^*)^{-\varepsilon_t} + \varsigma P_{t-1}^{-\varepsilon_t} \right] = Y_t \left[ (1 - \varsigma)^{1-\varepsilon_t} (1 - \varsigma \Pi_t^{\varepsilon_t-1}) + \varsigma \Pi_t^{\varepsilon_t} \right].
\]

C.4 Government Policies

In our simple model, the central bank sets the short-term nominal risk-free interest rate, \( i_t \), at time \( t \). This gives the expression for the ex post real return on one period safe nominal bonds

\[
R_{f,t+1} = \frac{1 + i_t}{\Pi_{t+1}}.
\]

(136)

The central bank uses a Taylor rule to set the short-term nominal interest rate:

\[
i_t = \max\{ (1 - \rho_t) [i_{ss} + \kappa_t \log \Pi_t + \kappa_y (\log Y_t - \log Y_t^*)] + \rho_t i_{t-1} + m_t, 0 \},
\]

(137)

The interest rate is bounded below by zero. We assume that the central bank normally follows the Taylor rule to set the interest rate, but sets the interest rate to zero when the Taylor rule indicates a negative interest rate. This simple treatment of monetary policy abstracts from considerations of other policy tools such as forward guidance, as in Eggertsson and Woodford (2003) and Werning (2011).
In Equation (137), \( Y^*_t \) is the natural level of output that would hold in a flexible price equilibrium, \( i_{ss} \) is the steady-state nominal interest rate, the smoothing parameter, \( \rho_i \), lies between zero and one, and \( \kappa_\pi \) and \( \kappa_y \) are constants which satisfy certain conditions, such as those in Woodford (2003), so that there is a unique equilibrium with non-exploding inflation. Here, \( m_t \) is an exogenous and stochastic component of the monetary policy rule. It evolves as

\[
m_t = \rho_m m_{t-1} + \sigma_m u_{m,t},
\]

with \( u_{m,t} \sim \text{i.i.d. } N(0,1) \).

Now we specify the credit policy. The central bank is also willing to buy the shares of the intermediate goods firm to facilitate lending. Such policies were studied by Gertler and Kiyotaki (2010), and Gertler and Karadi (2011). They capture the unconventional policies of purchasing risky, privately managed, non-government assets, implemented in the U.S., the U.K. and the Eurozone in the wake of the financial crisis. The U.S. Federal Reserve’s program of buying $600 million of Mortgage Backed Securities in 2008-09 (QE-1), and the European Central Bank’s Covered Bond Purchase Programs (CBPPs) for buying private sector debt, are examples of such policies.

The central bank buys a fraction \( \psi_t \) of the total outstanding shares of the intermediate goods firm, so that

\[
Q_t S_t = \phi_t N_t + \psi_t Q_t S_t,
\]

or (139)

where \( \phi_t \) is the leverage ratio for the privately-held asset, i.e., \( Q_t S_{p,t} \equiv \phi_t N_t \) and the government-held asset is \( Q_t S_{g,t} = \psi_t Q_t S_t \).

We define the total leverage ratio \( \phi_{c,t} \) as follows

\[
Q_t S_t = \phi_{c,t} N_t.
\]

The leverage ratio, \( \phi_{c,t} \), is the leverage ratio for total intermediated funds, public as well as private, and has the following relation with the private leverage ratio, \( \phi_t \), and the intensity of government credit intervention,

\[
\phi_{c,t} = \frac{\phi_t}{1 - \psi_t}.
\]

The central bank issues government bonds \( B_{g,t} = \psi_t Q_t S_t \) to fund the purchase of these shares. From this activity, the central bank thus earns an amount \( (R_{k,t+1} - R_{f,t+1}) B_{g,t} \) every period. The central bank credit has an efficiency cost of \( \tau > 0 \) units per unit of
credit supplied.

We assume that at the onset of a crisis, which is defined loosely to mean a period when the credit spread rises sharply, the central bank injects credit in response to movements in credit spreads, according to the following rule for $\psi_t$:

$$
\psi_t = \psi_{ss} + \nu \mathbb{E}_t \left( [\log R_{k,t+1} - \log R_{f,t+1}] - (\log R_{k,ss} - \log R_{f,ss}) \right),
$$

(142)

where $\psi_{ss}$ is the steady-state fraction of intermediation, $\log R_{k,ss} - \log R_{f,ss}$ is the steady-state risk premium, and the sensitivity parameter, $\nu$, is positive. According to (142), the central bank expands credit as the credit spread increase relative to the the steady-state credit spread. The rationale behind this policy specification, used by Gertler and Kiyotaki (2010) and Gertler and Karadi (2011), is as follows. In the absence of the financial friction that prevents the financial intermediaries from leveraging too much, the equilibrium outcome is efficient. The inefficiency arises due to the inability of the households to directly buy the risky assets, and the limit on their financial managers’ leverage. The inefficiency manifests itself in the form of large risk premium, since the financial intermediaries must be compensated adequately in the absence of high leverage. The government does not intervene when the risk premium is at its steady state level, but does when it rises to increasingly inefficient levels beyond it.

From (141), it is clear that, when the private leverage ratio, $\phi_t$, is kept fixed, the expanding credit policy increases the total leverage of financial intermediaries, i.e., $\phi_{c,t}$ rises.

### C.5 Resource and Government Budget Constraints

The resource constraint for the final good in our model is given by

$$
Y_t = C_t + I_t + \Phi \left( \frac{I_t}{\xi_t K_t} \right) \xi_t K_t + G_t + \tau \psi_t Q_t K_{t+1} K_{t+1}.
$$

The government spends a fraction $g_t$ of output $Y_t$ in period $t$, where $g_t$ is an exogenous and stochastic parameter. That is,

$$
G_t = g_t Y_t.
$$

(143)

It also funds the central bank’s purchase of shares by issuing purchasing bonds worth $\psi_t Q_t K_{t+1}$. Its revenues include taxes, $T_t$, and the central bank’s income from
intermediation, $\psi_{t-1}Q_{t-1}K_t(R_{k,t} - R_{f,t})$. Thus, the government budget constraint is

$$G_t + \tau \psi_t Q_{t+1} = T_t + \psi_{t-1}Q_{t-1}K_t(R_{k,t} - R_{f,t}).$$  \hspace{1cm} (144)$$

exogenously fixed at a constant fraction of output. We denote steady-state government spending share of output by $g_{ss}$. Government spending evolves as

$$\log g_t = (1 - \rho_g) \log g_{ss} + \rho_g \log g_{t-1} + \sigma_g u_{g,t},$$  \hspace{1cm} (145)$$

with $u_{g,t} \sim \text{i.i.d } N(0, 1)$.

Since the taxation, $T_t$, effectively takes up any slack that shows up on the government balance sheet, and given the existence of representative agents in the economy, the intertemporal budget constraint of the representative household and the intertemporal budget constraint of the government can be combined with taxes left out. Intuitively, then, by Walras’ Law, both budget constraints are redundant in determining the equilibria. However, this is very different from saying that the size and composition of the government balance sheet are irrelevant for pinning down the equilibrium under efficient financial market conditions, as was proposed by Wallace (1981). This is simply because not all investors can purchase an arbitrary amount of the same assets at the same market prices as the government in this model. Put more precisely, unlike private financial intermediation, government intermediation is not balance-sheet constrained.

### D Calibration and Estimation

We use our model to understand the response of the economy to various shocks. We use a calibrated version of the model, basing our parameter choices mainly on those in Gertler and Karadi (2011) and Gourio (2012), and the estimated dynamic parameters in Smets and Wouters (2007). In particular, we set the steady-state government expenditure to be $g_{ss} = 20\%$ and the steady-state government credit intervention to be $\psi_{ss} = 0$. These values are close to the average government expenditure and investment in the U.S. over the time period 1934 to 2010. We note that credit intervention by the U.S. government has historically been negligible, only becoming substantial after the recent crisis. Nevertheless, we include such intervention in our analysis to understand the effects of modern policy responses. The parameter values are summarized in Tables 4 and 5. We use Dynare to perform the analysis. For the baseline analysis, we use the first-order approximation around the steady state.

We allow the nominal interest rate to have a lower bound at zero. The impulse response functions with the zero lower bound are computed using the algorithm from Holden (2011). (As demonstrated by Cochrane (2013), Duarte (2016), Holden (2016) etc. the zero lower bound leads to the existence of multiple equilibria, even with the
Table 4: Static Parameter Calibration (Quarterly)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Household preference</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discount rate</td>
<td>$\beta$</td>
<td>0.99</td>
<td>Standard</td>
</tr>
<tr>
<td>Relative risk aversion</td>
<td>$\gamma$</td>
<td>2</td>
<td>Standard</td>
</tr>
<tr>
<td>Habit parameter</td>
<td>$h$</td>
<td>0.815</td>
<td>Standard</td>
</tr>
<tr>
<td>Relative weight of labor</td>
<td>$\chi$</td>
<td>3.409</td>
<td>Standard</td>
</tr>
<tr>
<td>Inverse Frisch elasticity of labor supply</td>
<td>$\psi$</td>
<td>0.276</td>
<td>Standard</td>
</tr>
<tr>
<td><strong>Financial intermediaries</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steady-state fraction of divertible capital</td>
<td>$\lambda_{ss}$</td>
<td>0.381</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td>Proportional transfer to new bankers</td>
<td>$\omega$</td>
<td>0.002</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td>Survival rate of bankers</td>
<td>$\theta$</td>
<td>0.972</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td><strong>Capital producing firms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depreciation rate</td>
<td>$\delta$</td>
<td>0.025</td>
<td>Standard</td>
</tr>
<tr>
<td>Adjustment cost coefficient</td>
<td>$\psi$</td>
<td>1</td>
<td>Standard</td>
</tr>
<tr>
<td><strong>Retail firms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elasticity of substitution</td>
<td>$\epsilon$</td>
<td>4.167</td>
<td>Standard</td>
</tr>
<tr>
<td>Probability of keeping prices fixed</td>
<td>$\varsigma$</td>
<td>0.779</td>
<td>Standard</td>
</tr>
<tr>
<td>Price indexation</td>
<td>$\varsigma_p$</td>
<td>0</td>
<td>Simplification</td>
</tr>
<tr>
<td><strong>Government policies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflation coefficients of Taylor rule</td>
<td>$\kappa_\pi$</td>
<td>1.5</td>
<td>Standard</td>
</tr>
<tr>
<td>Output coefficients of Taylor rule</td>
<td>$\kappa_y$</td>
<td>0.125</td>
<td>Standard</td>
</tr>
<tr>
<td>Persistence of interest rate</td>
<td>$\rho_i$</td>
<td>0.8</td>
<td>Standard</td>
</tr>
<tr>
<td>Government expenditure ratio</td>
<td>$g_{ss}$</td>
<td>20%</td>
<td>Standard</td>
</tr>
<tr>
<td>Steady-state government share of capital</td>
<td>$\psi_{ss}$</td>
<td>0</td>
<td>Standard</td>
</tr>
</tbody>
</table>

The presence of the financial friction implies that the steady state of the model features a non-zero risk premium. The low supply of risk-free assets causes the interest rate to be low. The inability of the financial intermediaries to finance the intermediate goods producers also lowers the price of capital, which increases the expected return on capital, further increasing the risk premium.

D.1 Calibration Experiments

We conduct two calibration experiments involving shocks to capital quality and margin. The capital quality shock allows us to examine the reaction of the economy
Table 5: Dynamic Parameter Calibration (Quarterly)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Capital Quality</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td>$\rho_\xi$</td>
<td>0.66</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td>Volatility</td>
<td>$\sigma_\xi$</td>
<td>0.05</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td><strong>Margin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td>$\rho_\lambda$</td>
<td>0.66</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td>Volatility</td>
<td>$\sigma_\lambda$</td>
<td>0.20</td>
<td>Gertler and Karadi (2011)</td>
</tr>
<tr>
<td><strong>Monetary Policy</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td>$\rho_m$</td>
<td>0.15</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td>Volatility</td>
<td>$\sigma_m$</td>
<td>0.24</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td><strong>Government Spending</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td>$\rho_g$</td>
<td>0.97</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td>Volatility</td>
<td>$\sigma_g$</td>
<td>0.53</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td><strong>Markup</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td>$\rho_\varepsilon$</td>
<td>0.89</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td>Volatility</td>
<td>$\sigma_\varepsilon$</td>
<td>0.14</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td><strong>TFP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td>$\rho_A$</td>
<td>0.95</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td>Volatility</td>
<td>$\sigma_A$</td>
<td>0.45</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td><strong>Investment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistence</td>
<td>$\rho_z$</td>
<td>0.75</td>
<td>Smets and Wouters (2007)</td>
</tr>
<tr>
<td>Volatility</td>
<td>$\sigma_z$</td>
<td>0.45</td>
<td>Smets and Wouters (2007)</td>
</tr>
</tbody>
</table>
to productivity shocks, whereas the margin shock allows us to examine the impact of financial market frictions on the real economy. A comparison of the two responses will provide insights into whether inclusion of the financial sector generates amplifications that do not appear in conventional DSGE models such as Smets and Wouters (2007).

**Experiment 1: Capital Quality Shock**

In our first experiment, we examine the effect of a capital quality shock on our model economy. We argue that an initial adverse disturbance of capital quality can approximately capture a decline in the quality of intermediary assets, leading to a severe decline in the net worth of the financial intermediaries. There are two major effects of the quality shock, one exogenous and one endogenous. The first effect is the exogenous impact of the destruction of capital on output and asset values. The second effect is endogenous. The balance sheet of intermediaries is weakened by the decline of asset values, and hence intermediaries reduce their demand for investment goods, which suppresses the price of capital, $Q$. The endogenous feedback effect of a decline in $Q$ is to further weaken the balance sheet of the financial intermediaries.

We choose the shock size to be a 5% deviation from the steady-state $\xi$ level (i.e., $\sigma_\xi = 0.05$). Conditional on occurring, the shock obeys an AR(1) decaying path with a persistent parameter $\rho_\xi = 0.66$. The shock path is displayed in the top-left corner of Figure 15. The real economy’s responses to this capital quality shock are shown in Figure 15. The capital quality shock triggers dramatic drops in output, investment, labor, and capital stock. Also, the intermediate and capital goods prices decline due to the weakened demand for investment. Interestingly, however, we can see that credit policies do not show any significant power to combat the capital quality shock, which can be seen even more clearly in Figure 16. In Figure 16, we see that the change in credit policy lasts for a long period, although the total amount of credit intervention is low overall. The credit policy does bring down the leverage and the risk premium for the intermediaries, but only by a moderate amount.

**Experiment 2: Margin Shock**

In our second experiment, we examine the effect of a margin shock on our model economy. The crisis scenario of a sudden collapse of funding can be roughly captured by an adverse shock in $\lambda_t$, meaning the deterioration of the pledgeability of the financial intermediaries’ assets. This can also be interpreted as a margin shock, as is emphasized by Geanakoplos (2001, 2009), among others.

The adverse margin shock affects asset value only through the endogenous channel, by deteriorating the capacity of the intermediaries to take on leverage. Like the capital quality shock, it also triggers a feedback effect via the balance sheet of intermediaries.

We choose the shock size to be one standard deviation from the steady-state $\lambda$ level (i.e., $\sigma_\lambda = 0.20$). We use such a large shock to get the impulse response to the real variables to be roughly the size they were during the recent financial crisis.
Figure 15: Real quantities’ response to capital quality shock: 5% deviation from steady state. The thick solid curve represents the case in the absence of financial frictions. The solid curve is for the case of zero credit policy intervention ($\nu = 0$). The dashed curve is for the case of moderate credit policy intervention ($\nu = 10$). The dash-dotted curve is for the case of intensive credit policy intervention ($\nu = 100$). The nominal interest rate has a zero lower bound.
Figure 16: Financial variables’ response to capital quality shock: 5% deviation from steady state. The solid curve is for the case of zero credit policy intervention ($\nu = 0$). The dashed curve is for the case of moderate credit policy intervention ($\nu = 10$). The dash-dotted curve is for the case of intensive credit policy intervention ($\nu = 100$). The nominal interest rate has a zero lower bound.
Conditional on occurring, the shock obeys an AR(1) decaying path with a persistent parameter $\rho = 0.66$, the same as in experiment I. The shock path is displayed in the top-left corner of Figure 17. The real economy’s responses to this margin shock are shown in Figure 17. The margin shock causes severe but very temporary drops in output, investment, and employment. However, it only generates a small decline in capital stock, different from the response of capital quality shock. Also, the prices of the intermediate and capital goods decline due to the weakened demand for investment. In contrast to the case of capital quality shock, credit policies are very efficient in alleviating the adverse impact of the margin shock, which can be seen even more clearly in Figure 18. In Figure 18, we find that the credit policy only lasts for about 10 quarters, with a level similar to that of the policy response to the 5% capital quality shock. From Figure 18, it is obvious that the credit policy is a powerful tool in maintaining the stability of the financial system when a sudden funding crisis occurs.

D.2 Estimation Analysis

In this section, we perform an estimation exercise to demonstrate the inadequacy of current estimation methods, which rely on linearized approximations and Kalman filtering, in incorporating the non-linear dynamics of asset pricing and risk premia. We estimate our model using macroeconomic and financial data. Since the static parameters of Table 4 are reasonably well understood and estimated, we fix them in place, and only estimate the dynamic parameters for exogenous shocks using a Bayesian method (see, e.g. Smets and Wouters, 2007). For simplicity, we assume that the nominal interest rate does not have a zero lower bound.

We use quarterly data on the data variables

$$Y_t = \begin{pmatrix}
\log(Y_t) - \log(Y_{t-1}) \\
\log(C_t) - \log(C_{t-1}) \\
\log(I_t) - \log(I_{t-1}) \\
\log(W_t) - \log(W_{t-1}) \\
\log(L_t) - \log(L_{t-1}) \\
\log \Pi_t \\
i_t \\
R_{k,t}
\end{pmatrix}$$

from 1951Q4 to 2013Q4 for our estimation. For $R_{k,t}$ we use the quarterly return on the S&P 500 index, and unlever it using data on aggregate corporate leverage from the Flow of Funds tables. We ignore the zero lower bound on the nominal interest rate. We also add an AR(1) shock to the discount rate to identify the parameters (8 variables and 8 shocks). Formally, we calibrate the parameters $\Theta$ in Table 4 and estimate...
Figure 17: Real quantities’ response to intermediary margin shock: one standard deviation from steady state. The thick solid curve represents the case in the absence of financial frictions. The solid curve is for the case of zero credit policy intervention ($\nu = 0$). The dashed curve is for the case of moderate credit policy intervention ($\nu = 10$). The dash-dotted curve is for the case of intensive credit policy intervention ($\nu = 100$). The nominal interest rate has a zero lower bound.
Figure 18: Financial variables’ response to intermediary margin shock: one standard deviation from steady state. The solid curve is for the case of zero credit policy intervention ($\nu = 0$). The dashed curve is for the case of moderate credit policy intervention ($\nu = 10$). The dash-dotted curve is for the case of intensive credit policy intervention ($\nu = 100$). The nominal interest rate has a zero lower bound.
the dynamic parameters $\Psi$. The estimation is performed using Dynare. We use the Metropolis-Hasting algorithm to numerically compute the posterior distribution

$$f(\Psi|Y_t, \Theta) \propto g(Y_t|\Psi, \Theta) \cdot p(\Psi),$$

where $f(\Psi|Y_t, \Theta)$ is the posterior distribution of the parameters, $g(Y_t|\Psi, \Theta)$ is the likelihood function or the conditional distribution of the observables given the parameters, and $p(\Psi)$ is the prior distribution of the parameters. The prior means are based on the calibrations of Tables 4 and 5. We simulate the posterior using a sample of 20000 draws after dropping 40% of the draws. We report the priors and the estimated mean and the 90% HPD interval. In Table 6 we report the results of the estimation using all eight variables, and in Table 7 we report the results after dropping the variable for the return on assets $R_{k,t}$.

It is evident that the estimated shock sizes with asset return data are unreasonably large, even after imposing tight priors on the volatilities of the shocks. Particularly large is the volatility of the margin shock. This is because the likelihood function obtained using the Kalman filter in the Bayesian estimation is based on the first-order approximation around the deterministic steady state. The first-order approximated likelihood function suppresses the significant nonlinear structure of the model design. The nonlinear structural components are critical to capture the large fluctuations in risk premia and their important equilibrium feedback in the real economy. The Bayesian estimation based on first-order approximated likelihood functions performs successfully in DSGE models without financial intermediaries or risk premia data in Smets and Wouters (2007). However, it fails in DSGE models with financial frictions, trying to capture the volatile dynamics of risk premia in the data by ignoring the nonlinear features in the model.

When we include the risk aversion and habit parameters in the estimation, the results are similar, with very large estimated values of the shock sizes.

## E Fiscal Theory of the Price Level

In this Appendix, we elucidate the simple model that we use in Section 4.2 to illustrate the role of fiscal policy in determining equilibria.

Consider a deterministic, perfect foresight model with the representative consumer maximizing the utility function:

$$\sum_{t=0}^{\infty} \beta^t u(C_t).$$

(146)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Priors</th>
<th>Posteriors</th>
<th></th>
<th></th>
<th>90% HPD interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Dist.</td>
<td>Mean</td>
<td>Stdev.</td>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>Financial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda^h_{ss}$</td>
<td>Steady state</td>
<td>Inv Gam</td>
<td>1.625</td>
<td>2</td>
<td>0.1523</td>
<td>0.1475</td>
</tr>
<tr>
<td>$\lambda_{ss}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{\lambda}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
<td>0.7190</td>
<td>0.7167</td>
</tr>
<tr>
<td>$\sigma_{\lambda}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
<td>0.6821</td>
<td>0.6192</td>
</tr>
<tr>
<td>Capital Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{\xi}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
<td>0.6033</td>
<td>0.5996</td>
</tr>
<tr>
<td>$\sigma_{\xi}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
<td>1.3698</td>
<td>1.3020</td>
</tr>
<tr>
<td>Discount Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{\beta}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
<td>0.9404</td>
<td>0.8929</td>
</tr>
<tr>
<td>$\sigma_{\beta}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
<td>1.4846</td>
<td>1.4778</td>
</tr>
<tr>
<td>Monetary Policy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{mp}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
<td>0.5867</td>
<td>0.5860</td>
</tr>
<tr>
<td>$\sigma_{mp}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
<td>1.2175</td>
<td>1.2150</td>
</tr>
<tr>
<td>Spending</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{g}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
<td>0.3163</td>
<td>0.2877</td>
</tr>
<tr>
<td>$\sigma_{g}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
<td>0.2523</td>
<td>0.2476</td>
</tr>
<tr>
<td>Markup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{mk}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
<td>0.9821</td>
<td>0.9720</td>
</tr>
<tr>
<td>$\sigma_{mk}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
<td>0.3807</td>
<td>0.3694</td>
</tr>
<tr>
<td>TFP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{a}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
<td>0.6758</td>
<td>0.6645</td>
</tr>
<tr>
<td>$\sigma_{a}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
<td>1.0078</td>
<td>0.9667</td>
</tr>
<tr>
<td>Investment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{z}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
<td>0.5070</td>
<td>0.4693</td>
</tr>
<tr>
<td>$\sigma_{z}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
<td>0.1461</td>
<td>0.1150</td>
</tr>
</tbody>
</table>

133
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Priors</th>
<th>Posterior Mean</th>
<th>90% HPD interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dist.</td>
<td>Mean</td>
<td>Stdev.</td>
<td></td>
</tr>
<tr>
<td>Financial</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{ss}^h$</td>
<td>Steady state</td>
<td>Inv Gam</td>
<td>1.625</td>
<td>2</td>
</tr>
<tr>
<td>$\lambda_{ss}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_\lambda$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
</tr>
<tr>
<td>$\sigma_\lambda$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Capital Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_\xi$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
</tr>
<tr>
<td>$\sigma_\xi$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Discount Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_\beta$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
</tr>
<tr>
<td>$\sigma_\beta$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Monetary Policy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{mp}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
</tr>
<tr>
<td>$\sigma_{mp}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Spending</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_g$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
</tr>
<tr>
<td>$\sigma_g$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Markup</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{mk}$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
</tr>
<tr>
<td>$\sigma_{mk}$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>TFP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_a$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
</tr>
<tr>
<td>$\sigma_a$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Investment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_z$</td>
<td>Persistence</td>
<td>Beta</td>
<td>0.66</td>
<td>0.2</td>
</tr>
<tr>
<td>$\sigma_z$</td>
<td>Volatility</td>
<td>Inv Gam</td>
<td>0.5</td>
<td>0.2</td>
</tr>
</tbody>
</table>
Every period, the consumer can buy (or sell) one-period bonds which pay the nominally risk-free interest rate, $i_t$. Denote by $B_t$ the number of such bonds bought by the consumer at time $t$. The consumer maximizes utility subject to the present value budget constraint:

$$\sum_{t=0}^{\infty} Q_{0,t} P_t C_t = B_{-1} + \sum_{t=0}^{\infty} Q_{0,t} P_t (Y_t - T_t).$$

(147)

Here, $Y_t$ is the exogenous output, which we assume to be a constant $Y$, $C_t$ is the consumption, $P_t$ is the price level, $T_t$ is the (real) lump-sum taxes. The government has no expenditures. For $t < s$, the nominal discount rate is:

$$Q_{t,s} = \prod_{j=0}^{s-t-1} \left( \frac{1}{1 + i_{t+j}} \right),$$

where $i_t$ is the nominal interest rate set by the central bank.

The SDF used by the consumer to value nominal claims is

$$\frac{\Lambda_{t+1}}{\Lambda_t} = \beta \frac{u'(C_{t+1})}{u'(C_t)} \frac{P_t}{P_{t+1}}.$$

Since $C_t = Y$ for all $t$, we have

$$\frac{\Lambda_{t+1}}{\Lambda_t} = \beta \frac{1}{\Pi_{t+1}},$$

where $\Pi_{t+1} = P_{t+1}/P_t$ is the inflation rate. The consumer Euler equation for the risk-free bond is

$$1 = \frac{\Lambda_{t+1}}{\Lambda_t} (1 + i_t),$$

which gives the Fisher relation by defining the real interest rate as $1 + r = 1/\beta$,

$$1 + i_t = (1 + r) \Pi_{t+1}.$$

(148)
Substituting the Fisher relation into the budget constraint (147), we get

\[
\frac{B_{t-1}}{P_t} = \sum_{j=0}^{\infty} \frac{T_{t+j}}{(1+r)^j}. \tag{149}
\]

This equation represents the government budget constraint, which must hold by Walras’ Law. It states that the real value of government debt must be equal to the net present value of its tax revenues.

We now have three sequences of variables, \(\{i_t, B_t, T_t\}\), that can be controlled by the government. However, only two of these can be independent because the government budget constraint

\[
(1 + i_{t-1})B_{t-1} = B_t + P_tT_t
\]

must hold period by period. Suppose, then, that the government controls \(\{i_t, T_t\}\), with the monetary authority setting, \(\{i_t\}\), and the fiscal authority setting, \(\{T_t\}\).

The monetary policy is given by

\[
1 + i_t = (1 + r)\Phi(\Pi_t). \tag{150}
\]

Combining (148) and (150), we get the difference equation for inflation:

\[
\Pi_{t+1} = \Phi(\Pi_t). \tag{151}
\]

F Summary of Central Banking Macro Models

In this Appendix, we review the core models used by major central banks to analyze monetary policy. As discussed in the main text, they include both large-scale macroeconometric models and New Keynesian DSGE models. In the following tables, we summarize the key models that the major monetary authorities have used in the past or are currently using.

F.1 The U.S. Federal Reserve Board

Large-scale Macroeconometric Models
The U.S. Federal Reserve Board (“the Fed”) makes its monetary policy principally
Table 8: The Core Open-Economy DSGE Models at Central Banks

<table>
<thead>
<tr>
<th></th>
<th>U.S. Fed</th>
<th>ECB</th>
<th>BOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Full Name</td>
<td>SIGMA</td>
<td>NAWM</td>
<td>ToTEM</td>
</tr>
<tr>
<td>Managed by</td>
<td>FOMC</td>
<td>Governing Council</td>
<td>Governing Council</td>
</tr>
<tr>
<td>Adjustment Friction</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Financial Sector</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Habit Utility?</td>
<td>Habit</td>
<td>Habit</td>
<td>Habit</td>
</tr>
<tr>
<td>Open Economy?</td>
<td>Yes Multicountry</td>
<td>Yes Small</td>
<td>Yes Small</td>
</tr>
<tr>
<td>Estimation?</td>
<td>Est. &amp; Calibration</td>
<td>Est. &amp; Calibration</td>
<td>Est. &amp; Calibration</td>
</tr>
<tr>
<td>Linearization?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing Market?</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Endo.Risk Premium</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>About 36</td>
<td>About 66</td>
<td>About 54</td>
</tr>
<tr>
<td>Number of Equations</td>
<td>About 37</td>
<td>About 102</td>
<td>About 72</td>
</tr>
<tr>
<td>Number of Shocks</td>
<td>About 16</td>
<td>About 18</td>
<td>About 7</td>
</tr>
<tr>
<td>Frequency of Data/Updates</td>
<td>Quarterly</td>
<td>Quarterly</td>
<td>Quarterly</td>
</tr>
<tr>
<td>Short-run Fluctuation</td>
<td>Supply/Demand</td>
<td>Supply/Demand</td>
<td>Supply/Demand</td>
</tr>
<tr>
<td>Long-run Steady State</td>
<td>Supply</td>
<td>Supply</td>
<td>Supply</td>
</tr>
<tr>
<td>Expectation Formation</td>
<td>Rational</td>
<td>Rational</td>
<td>Rational</td>
</tr>
<tr>
<td>Microfounded</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Table 9: The Hybrid Open-Economy Models at Central Banks

<table>
<thead>
<tr>
<th></th>
<th>BOE</th>
<th>BOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>COMPASS</td>
<td>BEQM</td>
</tr>
<tr>
<td>Model Full Name</td>
<td>Bank of</td>
<td>Bank of</td>
</tr>
<tr>
<td></td>
<td>England COMPASS</td>
<td>England Quarterly</td>
</tr>
<tr>
<td></td>
<td>Model</td>
<td>Model</td>
</tr>
<tr>
<td>Managed by</td>
<td>MPC</td>
<td>MPC</td>
</tr>
<tr>
<td>Meeting</td>
<td>once</td>
<td>once</td>
</tr>
<tr>
<td>Frequency</td>
<td>per year</td>
<td>per year</td>
</tr>
<tr>
<td>References</td>
<td>Burgess et al. (2013)</td>
<td>Harrison et al. (2005)</td>
</tr>
<tr>
<td>Adjustment</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Friction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial Sector</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Utility</td>
<td>Habit</td>
<td>Habit</td>
</tr>
<tr>
<td>Open?</td>
<td>Yes Small</td>
<td>Yes Small</td>
</tr>
<tr>
<td>Estimation?</td>
<td>Estimation &amp; Calibration</td>
<td>Estimation &amp; Calibration</td>
</tr>
<tr>
<td>Linearization?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing Market?</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Endo.Risk Premium</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>About 106</td>
<td>About 147</td>
</tr>
<tr>
<td>Frequency of Data/Updates</td>
<td>Quarterly</td>
<td>Quarterly</td>
</tr>
<tr>
<td>Short-run Fluctuation</td>
<td>Supply/Demand</td>
<td>Supply/Demand</td>
</tr>
<tr>
<td>Long-run Steady State</td>
<td>Supply</td>
<td>Supply</td>
</tr>
<tr>
<td>Expectation Formation</td>
<td>Rational</td>
<td>Rational</td>
</tr>
<tr>
<td>Microfounded</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>for the core model</td>
<td>for the core model</td>
</tr>
</tbody>
</table>
Table 10: The Core Closed-Economy DSGE Models at Central Banks

<table>
<thead>
<tr>
<th></th>
<th>U.S. Fed</th>
<th>ECB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Full Name</td>
<td>Fed’s Estimated, Dynamic, Optimization-based model</td>
<td>Christiano-Motto-Rostagno model</td>
</tr>
<tr>
<td>Managed by Foundation</td>
<td>FOMC</td>
<td>Governing</td>
</tr>
<tr>
<td>Adjustment</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Friction Financial</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Sector</td>
<td></td>
<td>financial accelerator</td>
</tr>
<tr>
<td>Friction Financial</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Utility</td>
<td>Habit</td>
<td>Habit</td>
</tr>
<tr>
<td>Open?</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Estimation?</td>
<td>Bayesian</td>
<td>Bayesian</td>
</tr>
<tr>
<td>Calibrate SS</td>
<td>Calibrate SS</td>
<td>Calibrate SS</td>
</tr>
<tr>
<td>Linearization?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Housing Market</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Endo.Risk Premium</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>About 43</td>
<td>About 74</td>
</tr>
<tr>
<td>Number of Equations</td>
<td>About 66</td>
<td>About 49</td>
</tr>
<tr>
<td>Number of Shocks</td>
<td>About 11</td>
<td>About 16</td>
</tr>
<tr>
<td>Frequency of Data/Updates</td>
<td>Quarterly</td>
<td>Quarterly</td>
</tr>
<tr>
<td>Short-run Dynamics</td>
<td>Supply/Demand</td>
<td>Supply/Demand</td>
</tr>
<tr>
<td>Long-run Steady State</td>
<td>Supply</td>
<td>Supply</td>
</tr>
<tr>
<td>Expectation Formation</td>
<td>Rational</td>
<td>Rational</td>
</tr>
<tr>
<td>Microfounded</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Model</td>
<td>U.S. Fed</td>
<td>U.S. Fed</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>FRB/US</td>
<td>FRB/Global</td>
<td>AWM</td>
</tr>
<tr>
<td>Foundation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjustment</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Friction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Sector</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Calibrate for steady state</td>
<td>Calibrate for steady state</td>
</tr>
<tr>
<td>Linearization?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Endo.Risk Premium</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Number of Core Equations</td>
<td>about 300</td>
<td>about 1700</td>
</tr>
</tbody>
</table>
using a core model called FRB/US (Reifschneider et al., 1997). At the same time, however, it uses a class of “periphery” models, mostly reduced-form econometric models, such as vector autoregression (VAR) models, and some medium- and small-scale calibrated New Keynesian DSGE models, focusing on a few particular mechanisms in equilibrium. As pointed out by David J. Stockton, the former director of the Fed’s Division of Research and Statistics, one reason these periphery models are separated from the core model at the Fed is because of the difficulty incorporating them into the core model in a robust way. Caballero (2010) has emphasized that it is very dangerous to add micro-level insights or mechanisms into the core model in a brute force manner. The \textit{ad hoc} manner of incorporating the Fed’s periphery models with its core model exposes it to the danger of being “too brutal”, as stressed by Caballero (2010).

The Fed’s first generation large-scale macroeconometric model, which still serves as the primary formal model of the U.S. economy for the Federal Open Market Committee, is the MIT-Penn-Social Science Research Council (MPS) model, which was adopted from the late 1960s until the beginning of 1996. An overview of the MPS model can be found in Brayton and Mauskopf (1985). This model employed quarterly data and had about 125 stochastic behavioral equations and more than 200 identities.

The MPS model was replaced by the FRB/US model in mid-1996 in an effort to improve the expectation formation dynamics and the long-run equilibrium component of the model. The FRB/US model specifies a neoclassical long-run steady state and dynamic behavior designed to address the Lucas critique by considering the influence of expectations and other sources of dynamics. A key feature of the FRB/US model, compared to the MPS model, is that expectations of future economic conditions are explicitly specified in many of its equations. For example, the FRB/US model can show how the anticipation of future events, such as a policy shift, may affect the economy today. The adoption of the rational expectations assumption alleviates the Lucas critique and is regarded as a major paradigm shift. Rational expectations allow agents in the model to look at policy changes as a contingent plan rather than as a one-time shock. In the FRB/US model, rational expectations are only a baseline assumption, and hence the policymakers can easily add learning on the top of the baseline.

Another key feature of the FRB/US model, compared to the MPS model, is that the FRB/US model incorporates polynomial adjustment frictions. The equations based on these adjustment costs are difficult to interpret because there is little microeconomic justification motivating them. Theoretically, these frictions imply decision rules in an error-correction format. The significant coefficients on lagged changes in the variables suggest adjustment costs that are of an order higher than quadratic (see, e.g. Tinsley, 1993). This led Taylor (1997) to call for new models in which the polynomial adjustment cost functions could be replaced by more microfounded frictions and structures.

The steady-state properties of the FRB/US model are close to those of the MPS model. The steady state is characterized using the neoclassical framework. In particular, all markets clear and the marginal product of each factor of production
is equal to its relative price in the long run. The growth of output in the long run depends on the exogenous population growth and the exogenous productivity growth of the production factors, following the assumption of a constant return to scale production technology. The short-run properties of the model are Keynesian in spirit. For example, output and employment are mainly determined by the aggregate demand because wages and prices are assumed to be sticky.

The FRB/US model has about 30 stochastic behavioral equations and about 300 identities. The behavioral equations can be categorized into four fundamental building blocks: arbitrage equilibria for financial variables, equilibrium planning for variables not determined in financial markets, dynamic adjustments for activities in nonfinancial sectors, and the formation of expectations. A detailed list of equations and identities is unmanageable for this paper; for more details, please refer to Brayton and Tinsley (1996) and Tinsley (1993). However, the following are some representative examples for the behavioral equations in all categories.

The aggregate consumption equations include the long-run equilibrium and short-run dynamic adjustment equations:

\[ c^* = 1.0v + 0.62s_{trans} - 0.15s_{prop} + 0.52s_{stock} + 1.28s_o + 0.13\tilde{x}, \]  \hspace{1cm} (152)

and

\[ \Delta c_t = -0.12(c_{t-1} - c^*_{t-1}) + 0.17\text{lags}_1(\Delta c_{t-i}) + 0.75\text{leads}_\infty(\Delta c^*_t) + 0.09\Delta y_t. \]  \hspace{1cm} (153)

The equilibrium (152) belongs to the category of equilibrium planning for aggregate consumption. Here, \( v \) is the log of wealth (\( V \)), that is, the present value of permanent income. The income consists of three main components: labor income (denoted by \( s_{labor} \) once normalized by \( V \)), transfer income (denoted by \( s_{trans} \) once normalized by \( V \)), and property income (denoted by \( s_{prop} \) once normalized by \( V \)). The variables \( s_{stock} \) and \( s_o \) are normalized by \( V \), and represent the value of corporate equity and other net financial/tangible assets, respectively. \( \tilde{x} \) is the aggregate output gap.

The dynamic adjustment equation (153) is categorized in the group of dynamic adjustments. Here, the lag operator \( \text{lags}_k(\cdot) := \sum_{i=1}^k w_i a_{t-i} \) with \( \sum_{i=1}^k w_i = 1 \), and the lead operator \( \text{leads}_k(\cdot) := \sum_{i=0}^{\infty} w_i a_{t+i} \) with \( \sum_{i=0}^{\infty} w_i = 1 \). The superscript “e” indicates the current market forecasts based on information available in period \( t \). The expectations of future variables are approximated by using small-scale VAR models. Finally, \( y \) is the log total income.

The coefficients of (152) and (153) are estimated using quarterly U.S. data from 1963Q1 to 1995Q4 (see Brayton and Tinsley, 1996, for details).
Another example is the no-arbitrage equation for 10-year government bond rates:

\[ r_{10,t} = 0.46 + 1.0 \text{leads}_{40}(r_t^e) - 0.79 \text{leads}_{40}(\tilde{x}_t^e) + 0.85 \text{lags}_{1}(\tilde{\mu}_{10,t}). \]  

Here, \( r_{10,t} \) is the 10-year government bond rate and \( r_t \) is the federal funds rate. The term \( \tilde{\mu}_{10,t} \) represents the term premium of the U.S. Treasury yield curve between the 10-year rate and the 3-month rate. The coefficients are estimated based on U.S. quarterly data from 1965Q1 to 1995Q4. Again, for more details, please refer to Brayton and Tinsley (1996).

In fact, equations for three long-term interest rates and the stock market comprise the core of the financial market sector in the FRB/US model. Unlike nonfinancial behavior, where frictions make it too costly to move immediately to equilibrium values, asset prices are assumed to be in equilibrium continuously. The financial market equations are all exact non-arbitrage conditions.

The Fed also uses a multi-country large-scale macroeconometric model, the FRB/MCM model. The FRB/MCM model consists of about 1400 equations. The model treats each country in a roughly symmetric manner. The countries modeled include the G-7 economies, Mexico, non-G-7 OECD economies, the newly industrialized economies, OPEC countries, and the rest of the world. The Fed also employs a large-scale macroeconometric model, called the World or FRB/Global model, which merges the non-U.S. parts of the FRB/MCM model and the FRB/US model. A more detailed introduction to the FRB/Global model can be found in Levin et al. (1997).

New Keynesian DSGE models at the Fed

We review two of the major New Keynesian DSGE models constructed and employed at the Fed for monetary policy analysis and decision making. The first is a multi-country open economy model named SIGMA, and the second is a model of the U.S. economy called the Federal Reserve Board’s Estimated Dynamic Optimization-based (FRD/EDO) model. Both models are extensions of the Christiano et al. (2005) and Smets and Wouters (2007) models whose core ingredients were reviewed in the canonical model of Section 3. We shall point out key features of the models added on top of that canonical model.

The SIGMA model, compared to the canonical New Keynesian DSGE model with habit persistence in consumption and adjustment costs in investment, incorporates the open economy framework of Obstfeld and Rogoff (1995). Furthermore, the SIGMA model also incorporates international pricing and trading frictions, such as the local currency pricing adjustment cost (see e.g. Betts and Devereux, 1996; Devereux and Engel, 2002) and the cost of adjusting trade flows. Another feature of the SIGMA model is that the agents have incomplete information about the persistence of shocks. More precisely, the agents learn the nature of the shocks using the Kalman filter. This learning mechanism produces gradual responses of the economy to shocks. The third important feature of the SIGMA model is that there are non-Ricardian households
who are simply assumed to consume their current after-tax disposable income. The main goal of introducing information frictions and non-Ricardian households is to generate a high persistence in the fiscal multiplier. Erceg et al. (2006) provide a review of the SIGMA model, comparing the short-run responses of it to those of the FRB/Global model, and show that they are quantitatively close.

The FRB/EDO model moves beyond the canonical two categories of private demand, consumption and investment, as found in Section 3. It divides private consumption into two categories: consumer durable goods, and consumer non-durable goods and non-housing services. It also separates residential and non-residential investment. The model features two final-goods sectors to capture key long-run growth patterns and produce different cyclical patterns of different durable expenditures such as those in consumer durable goods, residential investment, and non-residential investment. One sector produces goods mainly for consumption, and the other sector produces goods that are used for investment or capital accumulation. Like Christiano et al. (2005), this model allows for the variable utilization of capital. The non-residential capital is assumed to be owned by specialists who make decisions on non-residential investment, and hence non-residential capital accumulation. The model incorporates exogenous risk premia shocks trying to capture the financial accelerator effect in Bernanke et al. (1999). For further details on the FRB/EDO models please see Chung et al. (2010).

F.2 The European Central Bank

Macroeconometric Models at the European Central Bank

The Area-Wide Model (AWM) is a traditional macroeconometric model of the euro area that has been extensively used at the ECB over the past fifteen years. Like the FRB/US model at the Fed, the AWM model describes the dynamics of the economy through two major components. One is the long-run component, which characterizes the steady state of the economy and is consistent with neoclassical theory, while the other is the short-run component, which captures the demand-driven short-run dynamics in the data justified by the sluggish adjustment of prices and quantities. As a macroeconometric model, the short-run dynamics are not explicitly derived from an optimization framework, but are instead specified in a more \textit{ad hoc} form and estimated on the basis of historical data. Importantly, like other macroeconometric models, the dynamics are “disciplined” by the need to fulfill long-run steady-state properties by the use of error-correction terms and appropriate homogeneity properties.

Similar to the FRB/US model, the rest of the world is not explicitly modeled in the AWM framework. More precisely, the AWM model does not include any equations for variables that describe the rest of the world, which are instead treated as exogenous shocks to the model. There are two important drawbacks to the AWM model. First, a number of important channels are ignored. For example, there is no explicit role of financial and credit markets in shaping the transmission dynamics of monetary policy because financial quantities and credit variables have no explicit
impact on the decisions made by agents in the model. Second, in most equations, expectations are treated implicitly by the inclusion of lagged values of the variables (i.e., adaptive expectations), with the exception of some equations for financial variables (e.g., exchange rates and long-term interest rates). The backward-looking expectation formation method is unrealistic and clearly not satisfactory for policy analysis.

Because of the high level of aggregation in its data, the size of this model is relatively small compared to the FRB/US model. It contains about 84 equations, of which only 15 are estimated behavioral equations. We provide a simple illustration of the equations in the following example. For more detailed documentation, please refer to Fagan et al. (2001).

The aggregate real consumption of households is a function of real GDP, real disposable income, and real wealth:

\[
\Delta c_t = 0.77\Delta y_t - 0.066 \times \left[ 0.74 + c_{t-1} - 0.8(s_{\text{dis},t-1} - \pi_{c,t-1}) - 0.199(s_{\text{wealth},t-1} - \pi_{c,t-1}) \right]
\]  

(155)

where \(c_t\) is log real consumption, \(y_t\) is log real output, \(s_{\text{dis},t}\) is nominal households’ disposable income, \(\pi_{c,t}\) is the consumption deflator, and \(s_{\text{wealth},t}\) is nominal wealth which is defined as the sum of the capital stock, net foreign assets and public debt. The coefficients in (155) are estimated based on quarterly data from 1980Q1 to 1997Q4.

New Keynesian DSGE Models at the ECB

The ECB has developed several DSGE models, which it uses to analyze the economy of the eurozone as a whole rather than country by country. The models are intended as alternatives to the AWM, a more traditional macroeconometric model that the ECB has been using for fifteen years.

According to Smets et al. (2010), the core models at the ECB include two different models. The first is the New Area-Wide Model (NAWM), which is mainly based on Christiano et al. (2005), Smets and Wouters (2003), and Adolfson et al. (2007). Smets and Wouters (2003) originally estimated a closed-economy DSGE model of the euro area using Bayesian techniques, while Adolfson et al. (2007) estimated a small open-economy DSGE model of the euro area using Bayesian methods. The second model is the Christiano, Motto, Rostagno (CMR) model based on Christiano et al. (2008, 2010) that incorporates the New Keynesian components in Smets and Wouters (2003) and Christiano et al. (2005) with the imperfect credit market mechanism in Bernanke et al. (1999) and Chari et al. (1995).

The NAWM model is similar to the Fed Board’s calibrated open-economy model, SIGMA (see Section F.1). More precisely, besides the long-run neoclassical nature and short-run Keynesian features of nominal stickiness, it incorporates real frictions
such as consumption habit persistence and investment adjustment costs. Moreover, it also incorporates frictions relevant in an open-economy model, including local currency pricing, which generates an imperfect exchange rate pass-through in the short run, and costs of adjusting trade flows. Using Bayesian estimation methods, the model is estimated on 18 key macroeconomic variables, including real GDP, private consumption, total investment, government consumption, exports and imports, a number of deflators, employment and wages, and the short-term nominal interest rate. In addition, data for the nominal effective exchange rate, euro area foreign demand, euro area competitors’ export prices, and oil prices are used, which are deemed important variables in projections capturing the influence of external developments. 18 structural shocks are considered in the estimation. The NAWM model assumes that households are all Ricardian. An important feature (or limitation) of the NAWM model is that it distinguishes between producers of tradable differentiated intermediate goods and producers of three non-tradable final goods (a private consumption good, a private investment good, and a public consumption good). In addition, there are foreign intermediate goods producers that sell their differentiated goods in domestic markets, and a foreign retail firm that combines the exported domestic intermediate goods. International linkages arise from the trade of intermediate goods and international assets, allowing for imperfect risk sharing and limited exchange-rate pass-through on the import side. For detailed documentation on the NAWM model, please refer to Christoffel et al. (2008).

A distinguishing feature of the CMR model, compared to our canonical New Keynesian DSGE model (see Section 3), is its incorporation of the financial accelerator channel with an imperfect credit market, as emphasized in Bernanke et al. (1999), and the banking system of Chari et al. (1995). In the CMR model, firm investment in physical capital is leveraged, giving rise to the need for external financing. In particular, part of the working capital has to be financed prior to when revenues from selling current production become available. That is, firms need to pay for working capital in advance of production. Another main feature of the model is that the savers and the lenders do not interact directly, but via financial intermediaries. Intermediaries have their own balance sheet with liabilities, mainly different types of deposits, making it possible to construct aggregates such as M1 and M3, and assets, mainly different types of loans. The production of deposits requires resources in terms of capital, labor, and excess reserves. The presence of excess reserves captures the intermediaries’ need for maintaining a liquidity buffer to accommodate unexpected withdrawals. In this model, intermediaries cannot default. Financial contracts are denominated in nominal terms; given that borrowers and lenders are ultimately interested in the real value of their claims, shifts in the price level that were unanticipated at the time the financial contract was signed have real effects. This is a way to include the Fisher debt-deflation channel in the model. A detailed illustration and analysis of the CMR model can be found in Christiano et al. (2010).
F.3 The Bank of England

Macroeconometric Models at the Bank of England

The Medium-Term Macroeconometric Model (MTMM, or just MM) is a traditional macroeconometric model of the British economy that has played a central role at the Bank of England. Like the FRB/US model at the Fed and the AWM model at the ECB, the MTMM model is built around a number of estimated econometric relationships between important variables and is simultaneously disciplined by long-run properties consistent with economic theory. More precisely, as a macroeconometric model, its short-run dynamics are not explicitly derived from an optimization framework, but are instead specified in a more ad hoc form, estimated on the basis of historical data. Its long-run steady-state properties are imposed in the form of parameter restrictions that are implied by theory.

Like the FRB/US model and the AWM model, the rest of the world is not explicitly modeled in the MTMM model. The MTMM model treats the British economy and the rest of the world in an asymmetric manner, with variables for the rest of the world not appearing in equations as endogenous variables, but only as exogenous shocks in the equations. For example, in the MTMM model, which models Britain as an open economy, aggregate demand can be met from overseas as well as from domestic supply, and domestic supply can be sold overseas to meet foreign demand. So a stylized IS-curve model of aggregate demand can be written as:

$$c = \gamma_0 + \gamma_1 s^h + \gamma_2 s^w + \gamma_3 r + \gamma_4 x,$$

where $c$ is the real aggregate demand, $s^h$ is the real domestic income, $s^w$ is the real income of the rest of the world, $r$ is the real interest rate, and $x$ is the real exchange rate. Here, the variable $s^w$ shows up as an exogenous variable.

The MTMM model has two drawbacks similar to the AWM model. First, a number of important channels are missing, such as imperfect financial and credit markets. Second, expectations of the exchange rate one period ahead are assumed to be formed in a forward-looking manner, which implies that the exchange rate will jump in response to unexpected changes in interest rate differentials or in the long-run exchange rate level. However, other asset prices are not treated in a forward-looking manner, but are assumed to move in ways that are broadly consistent with the long-run growth path of the economy. Inflation expectations are assumed to exhibit a degree of inertia: wage-setters, for example, take time to respond to new information (i.e., adaptive expectations).

The MTMM model is, to some extent, a restricted vector error-correction model (VECM). It consists of about 20 key behavioral equations determining endogenous variables and about 90 identities defining relationships between variables. We provide a simple illustration of the equations in the following example. For more details, please refer to the Bank of England official documentation BOE (2010).
The aggregate households’ consumption is described by:

\[
\Delta c_t = -0.036 + 0.19\Delta s_{labor,t} + 0.052\Delta(s_{nonlabor,t-1} - \pi_{c,t-1}) - 0.068\Delta ur_{t-1}
\]

\[
+ 0.14\Delta(s_{housing,t} - \pi_{c,t}) + 0.014\Delta(s_{fin,t} - \pi_{c,t}) - 0.0016\Delta r_t - 0.0017\Delta r_{t-1}
\]

\[
-0.17 \times \left[ c_{t-1} - 0.89s_{labor,t-1} - 0.11(s_{wealth,t-1} - \pi_{c,t-1}) + 0.0028(r_{t-2} - \pi_i^{e,t-2}) \right]
\]

where \( c_t \) is real log aggregate consumption, \( s_{labor,t} \) is real post-tax log labor income, \( s_{nonlabor,t} \) is nominal log non-labor income, \( \pi_{c,t} \) is the log total final consumers’ expenditure deflator, \( ur_{t} \) is the log unemployment rate, \( s_{housing,t} \) is the log total housing wealth in nominal terms, \( s_{fin,t} \) is the nominal log net financial wealth, \( r_{t} \) is the base log interest rate, and \( s_{wealth,t} \) is nominal log total household sector wealth. The coefficients in (157) are estimated based on quarterly U.K. data from 1975Q1 to 1998Q1, and expected inflation, \( \pi_i^{e,t} \), can be estimated using past inflation rates:

\[
\pi_i^{e,t} = 1.1 + 1.2\pi_{t-1} - 0.6\pi_{t-2}.
\]

**Hybrid Models at the Bank of England**

The Bank of England, like other major monetary authorities, has developed a macroeconometric model for use in preparing its Monetary Policy Committee quarterly economic projections and inflation reports. Motivated by fears of potential technical insufficiency and the demand for tractability, the Bank of England has built a model with two distinct layers. Since 2003, the Bank of England Quarterly Model (BEQM) has become the main tool in the suite of models employed by the staff and the Monetary Policy Committee in the construction of the projections contained in its quarterly inflation report. The core layer is a tightly specified theoretical model containing dynamic decision rules derived from the solution of standard New Keynesian DSGE models. The non-core layer consists of equations that include additional lags and variables to match dynamics that are not modeled formally in the core. These non-core equations also allow the imposition of judgments based on “off-model” information or the judgment of the monetary authorities. The final forecast path can be thought of as a combination of theoretical insight from the structural core model, and the direct application of judgment or ad hoc estimated behavioral dynamics. A detailed illustration of the BEQM is provided in Harrison et al. (2005).

The core of the BEQM is a standard New Keynesian DSGE model for a small open economy. The model can be used to analyze a wide range of economic issues. Some standard features in its theoretical structure are designed to help match dynamic responses in the data, including consumption habits, labor adjustment costs, capital and investment adjustment costs, inertia in prices and nominal wages, wage and price
inflation stickiness, and slow import price pass-through. Because of the size of the core model, it does not fully capture all of the economic channels and dynamic relationships affecting the observed correlations between economic variables. This, in part, reflects the choice not to include in the core model certain features of the economy which could make the core model too large and complex to be tractable, such as credit market frictions. Moreover, its theoretical assumptions, such as Calvo mechanisms and price adjustment costs, which try to match some aspects of these correlations (for example, the degree of persistence of many nominal variables), are not yet well understood because they model components that are still “reduced form” on some level.

The non-core layer of the BEQM consists of ad hoc or “data-driven” dynamics on top of its theoretical structure. Incorporating the additional structure in the core model consistently would make the full model much more complicated and potentially difficult to run. Additionally, there are some effects that seem empirically robust, but are very difficult to model formally. For these reasons, the BEQM tries to embed the additional structures for data coherence, while at the time making the model sufficiently flexible and tractable for forecasting applications. For example, one can think of a neoclassical story about consumption being combined with proxies for credit effects for investment, supplemented by terms for firms preparing for the short run. The only restriction on the structure of ad hoc non-core equations is that the projected path for a given variable should always converge to the long-run equilibrium imposed by the core theory.

The full model is a hybrid combination of core and non-core elements, which matches past movements in the data better than either type of element on its own and enables a straightforward application of judgment to the forecast. One interpretation of this hybrid approach is that the final projections are a weighted average of three types of information: a structural story coming from the core model, extra short-run correlations from the non-core model, and judgment applied by the user through the non-core model (the relative weights on these types of information will vary across different parts of the model).

The model has the general format for a non-core equation as follows

\[ A(L)y_t = B(L)y^{\text{core}}_{t} + C(L)z_t + \epsilon_t \]  

where \( A, B, \) and \( C \) are polynomials in the lag operator \( L \). The variable \( y \) is the endogenous variable, and the prediction of the variable in the core model is denoted by \( y^{\text{core}} \). The variable \( z \) represents a vector of selected endogenous and exogenous variables, and \( \epsilon \) is an error term.

In addition, the model has another slightly different form of non-core equation, which simply follows the idea of combining forecasts. For example,

\[ y_{t+1} = \gamma_0 + \gamma_1 y^{\text{core}}_{t+1} + \gamma_2 y^{SR}_{t+1}. \]  

149
Here $y$ is an endogenous variable, $\hat{y}^{\text{core}}_{t+1}$ is the one-step-ahead forecast generated by the core model, and $\hat{y}^{\text{SR}}_{t+1}$ is a one-step-ahead forecast produced by a statistical “short-term” model.

A key feature of this approach is the strict separation between the core and non-core elements of the model. If the ad hoc elements are introduced into the core, it would risk violating the underlying theoretical assumptions of the core model, and it could also produce an unstable system. One way of viewing this hybrid approach is that it treats the path from the core model as a regressor, along with additional variables and ad hoc dynamics, in the full model equation, as in (159).

Projections from non-core equations feeding back into the core model are not allowed because this would bring about similar problems of instability and an undermining of the microfoundations of the core theory. Instead, the model uses a “non-feedback” approach, which maintains the distinction between the values from the core and the full forecasting models. This also facilitates the direct application of judgment to the forecast model, so that it is easy to impose desired paths for particular variables.

However, there are several concerns about this hybrid approach. First, the ad hoc component of the model is subject to the Lucas critique for monetary policy analysis. Second, there is no transparent interpretation for the parametric form of the dependence of the endogenous variables on their correspondence in the core model. Third, the one-way causal relationship between the projections from the non-core model and the projections from the core model makes the full model theoretically inconsistent.

Following the November 2011 Inflation Report, the forecast process at the Bank of England has been supplanted by the COMPASS platform. The detailed structure of the COMPASS platform is documented in Burgess et al. (2013). The COMPASS platform essentially uses the same idea as the BEQM model, consisting of four components: (1) the Central Organizing Model for Projection Analysis and Scenario Simulation (COMPASS) which is the core theoretical model with microeconomic foundations; (2) the suite of modes alongside the core model; (3) the Model Analysis and Projection System (MAPS), a MATLAB toolkit built and maintained by economists at the Bank of England; and (4) the Economic Analysis and Simulation Environment (EASE), a new IT user interface consisting of two components: a modeling toolbox called MAPS and a user interface called EASE.

The core model, COMPASS, is the platform’s main organizing framework for forecast production. COMPASS is an open-economy New Keynesian DSGE model, sharing many features with earlier models at other central banks, such as the ECB’s NAWM (see, e.g. Christoffel et al., 2008). As a DSGE model, it is stochastic by definition in the sense that exogenous random shocks to preferences, technologies, and constraints will affect agent decisions. In the absence of shocks, the model settles on a balanced growth path where all variables grow at constant (but possibly different) rates, reflecting exogenous population and technology trends. Shocks push the variables in the model away from the balanced growth path temporarily, with the speed at which
they return to the balanced growth path governed by the persistence of the shocks and the strength of the model’s propagation mechanisms, which in turn depend on the specific frictions in the model. COMPASS follows rational expectations as its baseline assumption (i.e., “model-consistent” expectations). The MAPS toolkit can assist analyzing COMPASS using alternative expectation formation assumptions.

Given the recent arguments that the current generation of New Keynesian DSGE models are ill-suited to analyzing the causes and consequences of financial crises, using a model like COMPASS may seem incomplete, particularly given that the model does not include a financial sector. Economists at the Bank of England believe that, at current levels of understanding, the benefits of adding a financial sector to COMPASS would be outweighed by the costs of the added complexity. It is possible that they will come to a different view in the future, as this rapidly developing area advances. Although COMPASS does not include an explicit role for a banking sector, there are several models in its suite that can be used to consider the impact of credit on the economy and explore the effects of an impaired banking sector.

According to Burgess et al. (2013), there are other main economic channels missing from COMPASS beside the financial sector. For example, the COMPASS model does not explicitly account for energy as an input to production or consumption. Changes in energy prices impact the marginal cost and inflation in a substantially different way than the changes in the prices of other goods and services. Further, fiscal policy is only modeled in a very simple way. Government spending is assumed not to affect household utility, distortionary taxes play little role, and households behave in a way which guarantees Ricardian equivalence holds. Also, there is only a single, short-term interest rate in the COMPASS model, which renders the core model silent on the effects of unconventional policies such as quantitative easing.

However, these missing channels are included in the suite of models. The suite consists of more than 50 separate specific models, covering a wide range of different channels and ways of thinking about the economy, which are not as yet included in the core COMPASS model. Different models can be selected from the suite, depending on what insight is required. The suite provides the means to cross-check the projections in COMPASS, expand the forecast to cover more variables, incorporate potentially critical mechanisms into analysis, and challenge the key judgments in the forecast. The non-core models are incorporated in the core model in an ad hoc manner. The suite also includes various extensions of COMPASS to incorporate financial sector channels. For example, one model introduces credit spreads into COMPASS. These drive a wedge between the official policy rate and the effective marginal interest rates faced by households and firms. The household rate enters the consumption equation in both the sticky and flexible price models (in the same way as the risk premium shock), so a rise in credit spreads has a similar effect to a negative demand shock. Meanwhile, a working capital channel is included on the production side. Firms have to borrow to pay for their labor and capital in advance of sales, so a higher credit spread increases their marginal cost. This means that a shock that increases spreads faced by firms leads to higher inflation and a fall in output. The model also allows for
a monetary policy response to credit spread shocks.

F.4 The Bank of Canada

Macroeconometric Models at the Bank of Canada

The Quarterly Projection Model (QPM) has been one of the core models of the Bank of Canada since September 1993. A detailed documentation of the QPM can be found in Poloz et al. (1994) and the related papers therein. The QPM as a system has two formal components: one, the steady-state model based on economic theory at some level of rigor, and the other, a set of short- to medium-run dynamic relationships that provide paths linking the starting conditions to solutions implied by the steady state.

The long-run equilibrium component is called SSQPM. The SSQPM contains several interesting structural features not shared by other steady-state components of macroeconometric models. First, households are modeled using a theoretical device known as “overlapping generations”. Consumers live an uncertain length of time and must plan their consumption and savings over that unknown lifetime. In doing so, they must balance the desire for current consumption with the incentive to save to generate higher consumption levels later in life. The QPM provides solutions for both the desired financial wealth of consumers in the long run, and the consumption/savings paths that will sustain that level. Second, it is an “almost small open economy”. A typical small open economy is characterized by exogenous prices for its exports and borrowing costs. The SSQPM model relaxes the assumption of exogenous exports prices. The idea is that the Canadian economy as a whole has some effect on the price of exports, even though individual firms act in a competitive manner. Such aggregate market power may arise from the fact that Canada is a large exporter of certain goods—wheat, lumber, and natural gas, for example. If the supply of these goods increases, the price falls, since the foreign demand curve for these products is not perfectly elastic. While this phenomenon is judged to be important enough to be included in the Canadian model, the effect is assumed to be too small to influence the general level of prices in the rest of the world. Moreover, Canada is assumed to have no influence on the world price level of imported goods. Third, the SSQPM introduces an exogenous risk premium into the firm’s specification. In particular, the risk premium is put into the cost of capital first-order condition as a wedge.

The dynamic structure of the QPM consists of three distinct types of equations. First, there are adjustment dynamics originating from the real and nominal frictions in the economy, including the investment adjustment costs and the labor market contracts. The adjustment features give rise to a gradual response to disturbances. Second, there are separate formation dynamics for expectations. The expectations in QPM are modeled as a mixture of backward- and forward-looking components. The model user can change the relative weights on the two components to generate the sort of stylized facts that are desired. In the core version of QPM, considerable weight is put on the backward-looking portion in order to capture the slow adjustment of expectations apparent in economic data. The forward-looking component is solved conceptually as
described above, while the backward-looking portion is usually specified as a simple weighted average of recent historical data. Third, there are automatic policy reactions to disturbances. Accordingly, QPM is specified with inflation control targets and rules of behavior that the monetary authorities will follow, should projected inflation deviate from those targets. Specifically, QPM includes a monetary policy reaction function, according to which a rise in anticipated inflation above target produces a rise in interest rates intended to move inflation back towards its target level over a horizon of six or seven quarters.

In sum, QPM has 27 behavioral equations. There are a total of 329 equations in the model, not counting the satellite structures. There are 155 equations describing expectations; most of the rest are identities. There are only 10 variables for which expectations are required. The large number of expectations equations is needed because the model must keep track of a number of leading terms for each of them.

New Keynesian DSGE models at the Bank of Canada
The Terms-of-Trade Economic Model (ToTEM) replaced the QPM in December 2005 as the Bank’s principal projection and policy analysis model for the Canadian economy. ToTEM is an open-economy, New Keynesian DSGE model. Interestingly, ToTEM contains producers of four distinct finished products: consumption goods and services, investment goods, government goods, and export goods. ToTEM also contains a separate commodity producing sector. Commodities are either used in the production of finished products, purchased directly by households as a separate consumption good, or exported on world markets. The law of one price is assumed to hold for exported commodities, whereas temporary deviations from the law of one price are permitted for commodities that are purchased domestically.

Recall that QPM only went partway towards incorporating fully rational expectations. Expectations in QPM are a weighted average of model-consistent expectations, or expectations based on forecasts that use the entire structure of the model, and adaptive expectations, which are based only on extrapolations of past values of the variable in question. In traditional macroeconometric models, adaptive expectations are utilized to yield the persistence inherent in the macroeconomic data, including inflation persistence. In ToTEM, a rational expectations DSGE model, expectations can be sticky if monetary policy is viewed as being less than fully credible. A detailed technical description of ToTEM can be found in Fenton and Murchison (2006).
References


