What Explains the Great Recession and the Slow Recovery?*

Pablo Cuba-Borda†

University of Maryland

Job Market Paper

November 2014

Please check here for the latest version

Abstract

This paper studies aggregate dynamics near the zero lower bound (ZLB) of nominal interest rates in a medium-scale New Keynesian model with capital. I solve a quantitatively realistic model of the U.S. economy with a ZLB constraint and use Sequential Monte Carlo methods to uncover the shocks that pushed the U.S. economy to the ZLB during the Great Recession. I investigate the interaction between shocks and frictions in generating the contraction of output, consumption and investment during 2008:Q3-2013:Q4 and find that a combination of shocks to the marginal efficiency of investment and to households' discount factor generated the prolonged liquidity trap observed in this period. A comparison between these two sources suggests that investment shocks played a more important role during this period. Fiscal and monetary policy stimulus help explain why the U.S. did not fall into a deflationary spiral despite a binding zero bound.

* I am indebted to Borağan Aruoba, John Shea and Luminita Stevens, for continuous support and advice in this project. I thank Felipe Saffie, Federico Mandelman, Jerónimo Carballo, Yiqun Chen, Luis Gonzáles and seminar participants at the University of Maryland, the Federal Reserve Bank of Atlanta, and Universidad Privada Boliviana for helpful comments and suggestions. I gratefully acknowledge the Federal Reserve Bank of Atlanta for their hospitality. All remaining errors are my own.

† Correspondence: Department of Economics, University of Maryland, College Park, MD 20742. Website: http://www.econ.umd.edu/~cuba. Email: cuba@econ.umd.edu.

1 Introduction

From 2007 to 2009 the U.S. economy was caught up in the throes of a severe recession. The Great Recession was the worst episode of economic contraction since the Great Depression, with consumption and investment plunging after the collapse of large financial institutions in September 2008. It has taken over half a decade for the economy to climb back to pre-recession levels. This episode is noteworthy not only because of its depth and subsequent slow recovery, but also because monetary policy quickly became constrained by the Zero Lower Bound (ZLB) on nominal interest rates. The Federal Reserve responded swiftly, lowering the Federal Funds Rate to nearly zero by the first quarter of 2009. The policy rate has remained at the ZLB for over five years to this date. Two natural questions arise: what made the Great Recession so severe? and why was the recovery slow?

Five years after the end of the recession, modern macroeconomic models continue to struggle to generate a coherent story about the events that caused such a severe contraction in economic activity. On one hand, existing medium scale DSGE models with richer structure that are able to account for the dynamics of aggregate quantities and prices abstract from the ZLB because of the computational complexity of solving rational expectation models with nonlinearities. On the other hand, existing DSGE models that explicitly account for the ZLB adopt highly stylized frameworks that abstract from investment or assume very simple structures for the shocks in the model, diminishing the quantitative capability of such models. This paper focuses on bridging this gap. I use the structure of a modern macro model commonly used for quantitative analysis and a set of computational techniques that help me paint a full picture of the causes of the Great Recession and the slow recovery. A key contribution is to uncover the shocks that pushed the nominal interest rate to the ZLB and understand their role in shaping the evolution of aggregate demand, in particular investment which is often ignored in models with a ZLB constraint.

Given the length of the ZLB spell, the ability to solve quantitative models that incorporate this fundamental constraint of monetary policy is essential for our understanding of the Great Recession. This is precisely the goal of this paper. I look at U.S. data from the perspective of a Dynamic Stochastic General Equilibrium (DSGE) model that incorporate the ZLB constraint. My model builds on the work of Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007), which were widely adopted as the common framework for important quantitative analysis among researchers and policy institutions. I use Sequential Monte Carlo methods (SMC) to structurally estimate the shocks that explain the recession and the slow recovery. Moreover, I use the estimated shocks to investigate how fiscal and monetary policies contributed during the recovery.

Compared to similar work that investigates the causes of the Great Recession in models that allow for the ZLB, this paper provides a broader answer for two reasons. First, I use a medium scale DSGE model that incorporates investment. Adding investment is crucial in order to understand the importance of shocks related to financial frictions, which have been argued to be at the root of the recession. Second, I take a novel approach to the structural estimation of shocks, combining nonlinear solution methods with SMC techniques. To the best of my knowledge, this paper is the first to apply both computational techniques simultaneously in a medium scale New Keynesian DSGE model of the type that is commonly used for policy analysis, and unveil the underlying drivers of the Great Recession. In addition, I use the estimated structural shocks to conduct counterfactual exercises. Among the five disturbances included in the model economy, I consider a shock to the marginal efficiency of investment (investment shock) and a shock to households' subjective discount factor (preference shock). Both shocks represent, in reduced form, deeper frictions in the financial sector of the economy.[1]

[1] This modeling approach is commonly used in DSGE models. For example, Smets and Wouters (2007) introduce a 'risk-premium' shock that affects the relative price of the nominal bond. In contrast, Justiniano, Primiceri and Tambalotti (2010) use a shock to households' subjective discount factor instead of the 'risk-premium' shock. Up to a first order, both frictions enter in the consumption Euler equation in the same way, but the risk-premium shock also affects the spread between the return on capital and bonds directly.

I find that the Great Recession originated in a decline in the marginal efficiency of investment. This decline started in the second half of 2007 and worsened in the third quarter of 2008, after the bankruptcy of important financial institutions. The U.S. economy encountered the ZLB as monetary policy responded to a large negative shock to households' ability to borrow and the ability of the financial system to channel resources to investment. In the absence of these shocks the recession would have been milder; with output falling 20% less with respect to its pre-recession trend and consumption and investment recovering fully to pre-crisis levels by the end of 2010. In the aftermath of the recession, my results indicate that the U.S. economy remained at the ZLB because of stimulative monetary policy that kept the nominal interest rate pegged at zero. During the liquidity trap, fiscal policy provided substantial stimulus, in particular during 2009:Q2-2011:Q2, and its stimulative effect on output helped the economy stave off deflation. Without the fiscal stimulus inflation would have been close to -3% when the economy hit rock bottom in 2009:Q1 and would have remained negative for another two quarters. The unwinding of the fiscal stimulus program in 2011 and political struggles that resulted in a reversion in the stance of fiscal policy held back the recovery.

A general consensus has emerged among economists that the recession originated in the financial system. However, the source and relative importance of frictions that caused the financial system to fail remains open to debate. For example, Mian, Rao and Sufi (2013), show that households' deleveraging reduced consumption in early 2007 and 2008, leading the collapse of the financial system. On the other hand, Gilchrist and Zakrajšek (2012) point towards sharp increases in borrowing costs for firms that depressed investment as the leading cause of the recession. Whether shocks affecting households consumption are more important than frictions that disrupt investment or not, is of particular interest for policy evaluation. Should policy have focused in alleviating households' mortgage debt and rehabilitate the housing market? Or should it have concentrated in providing resources to replenish bank capital and avoid the collapse of financial intermediaries and investment banks? My results indicate that shocks and frictions affecting investment played a prominent role in explaining the Great Recession. I also find that this friction remains elevated compared to their pre-recession which explains the sluggishness in the economic recovery.

The rest of the paper is organized as follows. Section 2 discusses related literature. Section 3 shows some important features of the data and a potential interpretation of the shocks that explain the dynamics around the Great Recession. The DSGE model used for the quantitative analysis is spelled out in detail in Section 4. In Section 5, I discuss the parameterization and the solution strategy of the nonlinear model that explicitly incorporates the ZLB. Section 6 describes how to uncover the structural shocks and the causes that pushed the economy to the ZLB. Section 7 presents a series of counterfactual exercises to understand the dynamics during and after the Great Recession. Section 8 concludes.

2 Related Literature

This paper fits within the literature that investigates macroeconomic dynamics in the presence of the Zero Lower Bound constraint. Eggertsson and Woodford (2003) were the first to study the response of the economy at the ZLB in a New Keynesian DSGE model. However, to maintain analytical tractability and characterize optimal policy, their setup abstracts from capital accumulation. To take the economy to the ZLB they study the effect of a temporary, unanticipated rise in households' discount factor that increase the real interest rate and lowers consumption. Their setup delivers sharp insights on the mechanics of ZLB events but it is not suited for quantitative analysis. In this regard, my paper is different because I incorporate capital accumulation and investment, and allow for five different shocks to drive the dynamics of the economy, bringing models that study ZLB and business cycle dynamics closer together.

Much of the work that emerged on the ZLB adopted the Eggertsson and Woodford (2003) modeling environment. For example Eggertsson (2009) investigates the effects of alternative fiscal policies at the ZLB, while Christiano, Eichenbaum and Rebelo (2011) studies the size of the fiscal multiplier. In these papers, it is assumed that a shock to households' discount factors is what causes the ZLB to bind, and hence the narrative around liquidity trap episodes has been centered on frictions that affect mostly consumption. Compared to this line of work, in my model there are two shocks that can push the economy to the ZLB, one that affects households' discount factors and works exactly in the same way as in related literature, and another that disturbs aggregate investment dynamics directly. Because I want to quantify the forces that took the U.S. economy to the ZLB during the Great Recession, I let actual data uncover the role of each shock, and investigate their contribution to the depth of the recession and the slow economic recovery. Shocks to the marginal efficiency of investment have a long tradition in business cycle analysis since Greenwood, Hercowitz and Huffman (1988), and have recently been rekindled as a dominant source of business cycles fluctuations by Justiniano, Primiceri and Tambalotti (2011). However, my paper is the first to quantify the relative importance of these alternative shocks in generating a liquidity trap.

This paper uses a medium-scale DSGE model along the lines of Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007). Variants of such models have been used to study business cycle dynamics as in Justiniano, Primiceri and Tambalotti (2010) or investigate trade-offs in monetary policy stabilization as in Justiniano, Primiceri and Tambalotti (2013). Moreover, variants of such models have been widely adopted in policy making institutions in the U.S. and around the world. Compared to this literature, my paper is among the few that solves the full nonlinear dynamics of a medium-scale model subject to a ZLB constraint. Christiano, Eichenbaum and Rebelo (2011) was an early attempt to bring the ZLB into medium-scale DSGE models but only to study the size of the fiscal multiplier. In related work Christiano, Eichenbaum and Trabandt (2014) solve a DSGE model that accounts for labor market variables as well as aggregate demand and prices.

I differ from Christiano, Eichenbaum and Trabandt (2014) in a key aspect. I use a particle filter to perform a formal estimation of the shocks that explain the data as seen through the structure of the model. In contrast, they take an informal approach in this dimension. They exploit the first order conditions of their model to map certain observables in the data with unobserved wedges in the model. In doing so they need to take a stand on the observables that better correspond to their proposed wedges and impose additional restrictions in order to map the observables to the model equilibrium conditions. I do not impose neither restrictions, instead I let the data speak freely about the driving forces that caused the recession. [2] In my application the estimation of the shocks respects the nonlinear equilibrium conditions of the model at all times, hence the estimation and interpretation of the shocks is more transparent.

This paper also sheds light on the importance of financial frictions during the Great Recession. Although, I do not incorporate an explicit mechanism like the financial accelerator of Bernanke, Gertler and Gilchrist (1999), or model financial amplification through banks balance sheets as in Gertler and Karadi (2011), I take a reduced form approach that is useful for measuring the strength and persistence of financial frictions. An additional advantage of my approach is that the frictions that I recover from the data can be rationalized with different microeconomic mechanisms for financial frictions. I consider two shocks that can be interpreted as disruptions in financial markets. In this regard, my paper also echoes the main result in Christiano, Eichenbaum and Trabandt (2014), that attributes most of the fluctuations during the Great Recession to a *financial wedge*. However, I do not need to assume that such wedge can be recovered directly from data on credit spreads only.[3] In my filtering exercises, I back out the equivalent to a *financial wedge* directly from observed data on consumption, investment and output growth. It turns out that my reduced form measure of financial frictions is closely related to fluctuations in the cost of borrowing for nonfinancial firms during the worst part of the crisis, but that remained persistently high even after credit spreads returned to pre-recession levels.

[2]To facilitate a direct measurement of the consumption and financial *wedges*, Christiano, Eichenbaum and Trabandt (2014) assume away the covariance between the stochastic discount factor and the ex-ante real interest rate.

[3]Their results are sensitive to the particular measure of spreads. When they use Gilchrist and Zakrajšek (2012) measure of credit spreads, their financial wedge is not persistent enough to produces a long lasting recession as the one observed in the U.S.

In terms of methodology this paper builds on the solution methods developed in Judd, Maliar and Maliar (2012), Aruoba, Cuba-Borda and Schorfheide (2013), Gust, Lopez-Salido and Smith (2012), and Fernández-Villaverde et al. (2012) to characterize the full nonlinear equilibrium dynamics of DSGE models subject to occasionally binding constraints.[4] Compared to the discrete state-space solution method based on policy function iteration reviewed in Richter, Throckmorton and Walker (2011), my paper uses a combination of projection and simulation techniques to find the global approximation to the model decision rules. The advantage is that my solution strategy is more suitable for medium-scale models with many state-variables. To extract the unobserved shocks that drive the dynamics during the Great Recession I implement a particle filter adapted from the work in Aruoba, Cuba-Borda and Schorfheide (2013).

3 The Great Recession

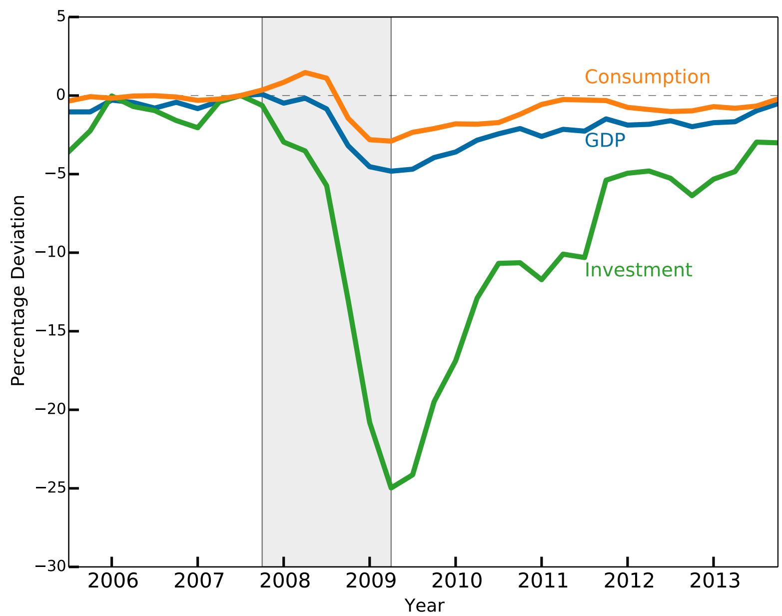
Before discussing the model I briefly review the evolution of key macroeconomic aggregate during and after the Great Recession. Figure 1 shows the comovement of key macroeconomic variables before and after the Great Recession. I look at the cyclical components of quarterly data on Gross Domestic Product (GDP), consumption and investment.[5] All series are expressed in annualized real per-capita terms. I extract the cyclical component using the Hodrick-Prescott filter with a standard smoothing parameter for quarterly observations. I normalize the data to 2007:Q3, which is the quarter prior to the official start of the recession according to the NBER. The y-axis in the figure is expressed in terms of the percent change relative to the peak of the NBER cycle. Output, investment and consumption all experienced a severe and prolonged contraction. Investment fell below trend together with GDP at

[4]Gust, Lopez-Salido and Smith (2012) are the first to estimate model parameters in a small New Keynesian with a ZLB but their setup abstracts from capital which prevents them from studying the evolution of investment in the data.

[5]My measure of consumption includes private personal consumption expenditure of non-durable goods and services, whereas my measure of investment combines fixed private investment and the private consumption of durable goods. Additional details on the data series are provided in Section 5.2.

the start of the recession.[6] Consumption remained above its pre-recession level for another three quarters, then fell rapidly along with investment as the financial crisis intensified. By the trough in the first quarter of 2009, detrended investment had fallen 25% from its peak, while detrended output fell 5% and detrended consumption about 3%.[7]

Figure 1: U.S. Great Recession: Macroeconomic Comovement



Notes: Output, consumption and investment are expressed in annualized real per-capita terms. All series are detrended using the HP filter ($\lambda = 1600$) *and normalized to 2007:Q3. The shaded region indicates the NBER recession*

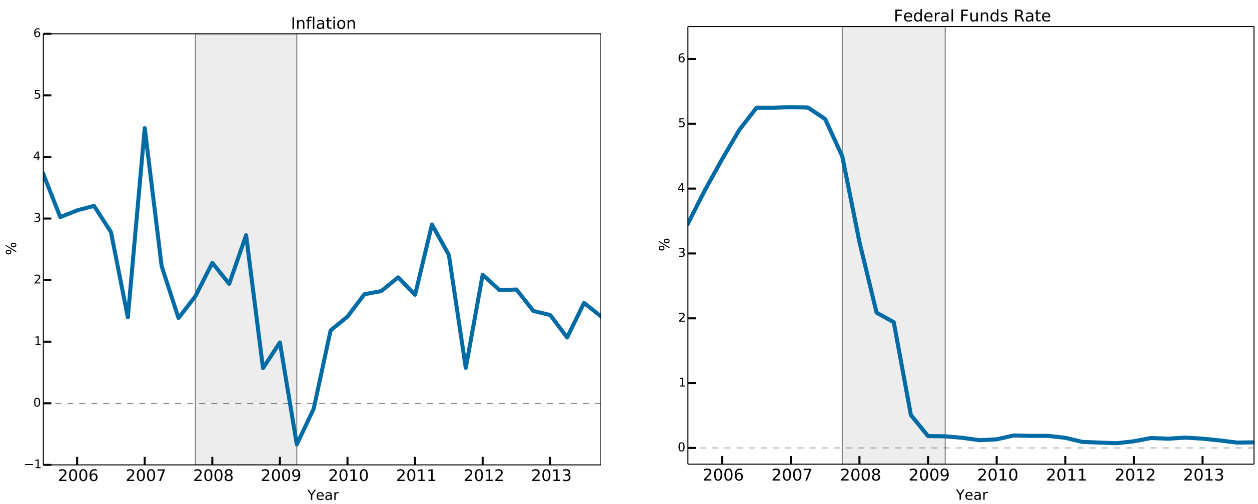
Figure 2 shows the evolution of prices and interest rates during the same time period. The figure shows inflation of the GDP deflator and the annualized effective federal funds rate, both expressed in percentage terms. At the onset of the recession and before 2008:Q3, inflation remained roughly around 2% while the nominal interest rate fell from 4.75% to 2%. As economic conditions worsened inflation fell rapidly and became negative in the first quarter of 2009. At the same time the nominal interest fell below 0.20%, effectively reaching

[6] Net exports did not contract until 2008:Q3. In fact, the value of total exported goods and services increased 4.5% between 2007:Q3 and 2008:Q2, while imports declined 0.1%. During the same period, government consumption increased 0.2% with respect to its pre-recession trend.

[7] In terms of levels, output fell by 7.3% from peak to through, consumption fell 5%, and investment fell 29%. Additional measures of economic activity also deteriorated sharply. Average weekly hours fell by almost 2% and the civilian unemployment rate rose from 4.4% to 9.8%

its lower bound. There is no doubt that from 2007:Q3-2008:Q2 the U.S. economy was in a recession. However up, to that point the evolution of macroeconomic aggregates can be dubbed a plain vanilla recession. As is evident from both figures, from 2008:Q3 onward the story changes substantially, with consumption turning around and quickly falling below pre-recession levels, and investment contracting even at a faster rate. Trying to uncover the forces behind this latter period is challenging precisely because the zero bound became binding.

Figure 2: Inflation and Interest Rates



Notes: The shaded region indicates the NBER recession

3.1 Potential causes of the Great Recession

I present some informal discussion of the micro foundations for the investment and preference shocks that play a central role in my results. I focus on these shocks, because the timing of the dramatic decline in consumption and investment shown in Figure 1, points to the prominence of disruptions originating in the financial sector of the economy. In addition, there is a long tradition in the ZLB literature, in particular in small scale New Keynesian models without capital, that relies on shocks to preferences as a simple mechanism to cause contractions in aggregate demand and push the economy to the ZLB.

Preference shocks. This type of shock affects the growth rate of aggregate consumption through movements in the real interest rate that tilt the consumption Euler equation.[8] Where do these movements in the real interest rate come from? Guerrieri and Lorenzoni (2011) provide a possible explanation based on tightening of borrowing constraints. A sudden reduction in the debt limit forces household near the constraint to reduce consumption and repay debt; the increased desire for savings induced by precautionary motives puts downward pressure on the nominal interest rate. From this perspective a tightening of borrowing constraints in a heterogeneous agent economy provides a rationale for an increase in the desire to save and reduction in the nominal interest rate that can be captured by shocks to households' preferences in the representative agent economy.

Investment shocks. With respect to shocks that distort the intertemporal margin of capital accumulation there are various reduced form interpretations. For example, Justiniano, Primiceri and Tambalotti (2010) and Justiniano, Primiceri and Tambalotti (2011) document that a type of investment shock represented as a wedge in the transformation of current investment into installed capital (marginal efficiency of investment) is the main source of business cycle fluctuations and they provide evidence that it played a significant role in the run-up of the Great Recession.[9] However, they cannot provide estimates of the investment shock after 2008:Q3 because their solution methods are unable to capture the ZLB constraint. A shock to the marginal efficiency of investment can be interpreted as a disruption in financial intermediation that will affect the supply of capital and generate fluctuations in its rate of return. Justiniano, Primiceri and Tambalotti (2010) point out that a costly monitoring friction in the spirit of Carlstrom and Fuerst (1997) also gives rise to a wedge that affects the transformation of investment into new capital. This wedge looks similar to a shock that shifts the cost of adjusting investment and interpret it as a disturbance that raises the cost of monitoring investment projects. This observation is consistent

[8] A typical linearized Euler equation expressed in percentage deviations and assuming logarithmic utility is: $\hat{c}_{t+1} - \hat{c}_t = \mathbb{E}_t\{\hat{R}_t - \hat{\pi}_{t+1} + \hat{\epsilon}_{t+1}\}$. Here ϵ_t are the shocks to households' subjective discount factor.

[9] Shocks to the marginal efficiency of investment were originally introduced by Greenwood, Hercowitz and Huffman (1988) in a real business cycle framework.

with the large increase in corporate credit spreads observed between 2008:Q1-2010:Q2, as documented by Gilchrist, Yankov and Zakrajšek (2009) and Gilchirst and Zakrajšek (2012).

4 The model

This section describes the model used for evaluating the macroeconomic dynamics observed during the Great Recession period. The model economy contains several frictions that introduce nominal and real rigidities that have been shown to be successful in capturing the dynamics of macroeconomic aggregates (Christiano, Eichenbaum and Evans (2005), Smets and Wouters (2007) and Justiniano, Primiceri and Tambalotti (2010)). The frictions in the model include price rigidity in the form of convex cost of price adjustment, habit formation in consumption, variable capital utilization, and investment adjustment costs. The dynamics are driven by exogenous shocks to the growth rate of technological progress, shocks to preferences, shocks to the marginal efficiency of investment, a shock to aggregate demand in the form of government purchases, and a shock to the monetary policy rule.

4.1 Households

Preferences. There is a representative household that consumes, and supplies labor L_t. Preferences are separable over consumption and labor (hours worked) and take the following functional form

$$\max_{C_t, I_t, u_t, \bar{K}_t, B_t, L_t} \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s d_{t+s} \left[\ln \left(C_{t+s} - hC_{t+s-1} \right) - \psi_L \frac{L_{t+s}^{1+\nu}}{1+\nu} \right]$$

The representative household maximizes expected discounted utility, where $\mathbb{E}_t$ denotes the expectation operator conditional on information available in the current period and β is the discount factor. The utility specification allows for external habits in consumption, where the parameter h controls the strength of the habit. The utility cost of labor is controlled by the term ψ_L, and ν represents the inverse of the Frisch elasticity of labor supply. The term

d_t is an intertemporal shock that follows a stationary first order autoregressive process

$$\ln d_t = \rho_d \ln d_{t-1} + \varepsilon_t^d, \quad \text{with} \quad \varepsilon_t^d \sim N(0, \sigma_d^2)$$

The d_t shock captures exogenous changes in the desire to increase or decrease consumption in the present compared to the future, and in what follows I refer to it simply as the *preference shock*.

Budget constraint. Households receive a nominal wage W_t as compensation for the labor they supply to intermediate firms. The capital stock of the economy $\bar{K}_{t-1}$ is owned by the households who rent it to intermediate firms every period in exchange for a nominal return R_t^k. In addition to the quantity of capital rented to firms, the household also chooses the intensity of capital utilization in the production process, denoted by u_t, such that the amount of capital that firms use to produce is equal to $K_t = u_t \bar{K}_{t-1}$. A higher intensity of operation of the capital stock entails a real cost for the household denoted by $\mathcal{A}(u_t)$, and expressed in terms of the final consumption good. In addition to factor income, the representative household collects interest from holding a one period risk-free nominal bond B_{t-1} issued by the government. This asset pays R_{t-1} dollars in period t. In addition the household pays a lump sum tax T_t, and receives the profits generated by firms Π_t. Income is allocated to consumption (C_t), investment (I_t) and the purchase of government bonds issued in the current period (B_t). Altogether the period budget constraint is given by

$$P_t C_t + P_t I_t + B_t \leq W_t L_t + R_t^k u_t \bar{K}_{t-1} - P_t \mathcal{A}(u_t) \bar{K}_{t-1} + R_{t-1} B_{t-1} - T_t + \Pi_t \qquad (1)$$

Investment frictions. Investment decisions are subject to an adjustment cost function $S(.)$ and a shock to the marginal efficiency of investment μ_t. As in Justiniano, Primiceri and Tambalotti (2011), I interpret this shock as a reduced form representation of a friction that disturbs the process of financial intermediation and affects the efficiency with which investment goods are transformed into capital

$$\bar{K}_t = (1-\delta)\,\bar{K}_{t-1} + \mu_t \left[1 - S\left(\frac{I_t}{I_{t-1}}\right)\right] I_t \tag{2}$$

The marginal efficiency of investment evolves according to the process

$$\ln \mu_t = \rho_z \ln \mu_{t-1} + \varepsilon_t^{\mu}, \quad \text{with} \quad \varepsilon_t^{\mu} \sim N(0, \sigma_{\mu}^2)$$

Let Ξ_t denote the multiplier associated with the capital accumulation equation and Λ_t the multiplier associated with the budget constraint of the household. The optimal investment allocation implies that the relative price of installed capital in terms of consumption goods is a function of current and future realizations of the investment efficiency shock

$$\frac{\Xi_t}{P_t\Lambda_t} = \beta\mathbb{E}_t \frac{P_{t+1}\Lambda_{t+1}}{P_t\Lambda_t}\left\{\left[u_{t+1}\frac{R^k_{t+1}}{P_{t+1}} - \mathcal{A}(u_{t+1})\right] + (1-\delta)\frac{\Xi_{t+1}}{P_{t+1}\Lambda_{t+1}}\right\}$$

The marginal efficiency of investment, μ_t, affects the transformation of current investment into installed capital directly through the capital accumulation equation and also indirectly through Tobin's Q that affects the rate on return of capital.

Labor Supply. Labor services produced by the representative household are sold to a perfectly competitive labor market at the aggregate nominal wage rate W_t. The optimality condition for the intratemporal allocation is[10]

$$W_t = d_t \psi_L \frac{L_t^{\nu}}{\Lambda_t}$$

[10]The shock d_t enters the intratemporal condition because it also affects the marginal disutility of labor. This specification of the preference shock help the model generate a positive correlation between consumption and hours, to match business cycle facts without the need of a separate shock to preferences for leisure. However, the equilibrium response of hours depends on the wealth effects that the preference shocks generates through changes in consumption.

4.2 Firms

The production side of the economy consists of perfectly competitive final good producers that buy intermediate goods $Y_{i,t}$ from a continuum of firms that operate in a monopolistically competitive market. The intermediate firms are indexed by $i \in [0,1]$.

Final good producers. The final good firms buy intermediate inputs from producers and aggregate the intermediate goods using a technology with constant elasticity of substitution to produce the consumption good Y_t. Taking the prices of inputs $P_{i,t}$ and the price at which they sell the final good P_t as given, the final-good firm chooses its demand for each intermediate input $Y_{i,t}$ to maximize profits

$$\max_{Y_{i,t}} P_t Y_t - \int_0^1 P_{i,t} Y_{i,t} di, \quad s.t. \quad Y_t \leq \left[\int_0^1 {Y_{i,t}}^{1-\lambda_p} di \right]^{\frac{1}{1-\lambda_p}}$$

where λ_p is the inverse of the elasticity of substitution across intermediate inputs, which controls the steady state markup of price over marginal cost. The optimal demand for intermediate goods satisfies $Y_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\frac{1}{\lambda_p}} Y_t$.

Intermediate-goods firms. Intermediate firms operate a technology that combines labor and capital to produce the intermediate good

$$Y_{i,t} = \begin{cases} K_{i,t}^{\alpha} \left(A_t L_{i,t}\right)^{1-\alpha} - A_t F & \text{if } K_{i,t}^{\alpha} \left(A_t L_{i,t}\right)^{1-\alpha} > A_t F \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Here $K_{i,t}$ and $L_{i,t}$ denote the firm's demand for effective units of capital and composite labor services respectively. A_t is an aggregate technology shock with growth rate $z_t \equiv A_t / A_{t-1}$. The growth rate z_t follows an exogenous autoregressive process

$$\ln(z_t/z) = \rho_z \ln(z_{t-1}/z) + \varepsilon_t^z, \quad \text{with} \quad \varepsilon_t^z \sim N(0, \sigma_z^2)$$

The term F represents a fixed cost that is calibrated to ensure zero profits in steady state. The growth rate of technology along the balanced growth path is given by $z = \gamma$.

Marginal costs. Intermediate firms rent labor and capital in perfectly competitive markets taking factor prices W_t and R_t^k as given. Each firm solves the following program

$$\min_{K_{i,t}, L_{i,t}} R_t^k K_{i,t} + W_t L_{i,t}$$

Subject to the production technology (3). Cost minimization entails the following equilibrium condition

$$\frac{K_{i,t}}{L_{i,t}} = \frac{\alpha}{1-\alpha} \frac{W_t}{R_t^k}$$

Because capital is traded in an economy-wide market, all intermediate producers take as given the aggregate rental rate of capital R_t^k. As a consequence the optimal factor allocation depends only on aggregate prices, so that aggregation is straightforward. Using the definition of aggregate demand for inputs, $K_t = \int_0^1 K_{i,t} di$ and $L_t = \int_0^1 L_{i,t}$, it is easy to show that the following expression for the marginal cost of production holds[11]

$$MC_t = \alpha^{-\alpha}(1-\alpha)^{\alpha-1} \frac{W_t^{1-\alpha} R_t^{k,\alpha}}{A_t^{1-\alpha}}$$

Price setting. Intermediate firms face a cost of adjusting prices in every period. The cost is expressed as a fraction of firms' revenue and is controlled by the convex function $\Phi_p(P_t/P_{t-1})$. Taking the marginal cost MC_t as given, each intermediate firm solves the following price

[11] Appendix A shows that because optimal factor allocation is identical across firms, so are marginal costs.

setting problem

$$\max_{\{P_{i,t}\}} \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s \frac{\Lambda_{t+s}}{\Lambda_t} \left\{ \left[1 - \Phi_p \left(\frac{P_{i,t+s}}{P_{i,t+s-1}} \right) \right] P_{i,t+s} Y_{i,t+s} - MC_{t+s} Y_{i,t+s} \right\}$$

$$s.t. \quad Y_{i,t} = \left(\frac{P_{i,t}}{P_t} \right)^{-\frac{1}{\lambda_p}} Y_t$$

4.3 Government

Monetary Policy. The monetary authority controls the short term interest rate following an operational rule that responds to deviations of inflation with respect to the central bank's desired level of inflation and the gap of observed output with respect to the non-stochastic level of output along the balanced growth path of the economy, Y^*. Based on the results in Aruoba, Cuba-Borda and Schorfheide (2013), I do not consider equilibria with deflationary dynamics, and I let the central bank's desired level of long-run inflation coincides with the steady-state level of inflation $\pi^* > 0$.[12] The main difference with respect to the standard analysis is that I impose the zero lower bound constraint on the nominal interest rate:

$$R_t = \max \left\{ 1, \left[(r^* \pi^*) \left(\frac{\pi_t}{\bar{\pi}} \right)^{\psi_1} \left(\frac{Y_t}{Y^*} \right)^{\psi_2} \right]^{\rho_R} R_{t-1}^{1-\rho_R} \exp(\varepsilon_t^r) \right\},$$

I assume that the response of the interest rate is smoothed with respect to the previously observed nominal interest rate. The parameter ρ_r controls the speed of the adjustment, while ε_t^R is a monetary policy shock, that is normally distributed with mean zero and standard deviation σ_r^2.

Fiscal Policy. The government issues bonds B_t every period to satisfy its flow budget constraint $P_t T_t - P_t G_t = R_{t-1} B_{t-1} - B_t$. The term G_t, is government expenditure, and evolves exogenously according to: $G_t = \zeta_t Y_t = \left(1 - \frac{1}{g_t}\right) Y_t$, where g_t is an exogenous autoregressive

[12] Aruoba, Cuba-Borda and Schorfheide (2013) find that the U.S. economy remained in a targeted-inflation regime where the steady-sate level of inflation is positive.

process with mean $\bar{g} = 1/(1-\zeta)$ that follows

$$\ln(g_t) = (1-\rho_g)\ln\bar{g} + \rho_g \ln(g_{t-1}) + \varepsilon_t^g, \quad \text{with} \quad \varepsilon_t^g \sim N(0, \sigma_g^2)$$

Here, ζ is the government expenditure to GDP ratio in the steady state.

4.4 Market Clearing

The market clearing conditions for this economy are as follows. I consider a symmetric price equilibrium, $P_{i,t} = P_{j,t} = P_t \quad \forall i, j \in [0,1]$, that satisfies: *(i)* the market for capital clears: $\int_0^1 K_{i,t} = K_t \equiv u_t \bar{K}_{t-1}$, *(ii)* the market for labor services clear: $\int_0^1 L_{j,t} = L_t$ *(iii)* Installed capital $\bar{K}_t$ evolves according to (2), such that the market for final goods clears

$$\left[\frac{1}{g_t} - \Phi(\pi_t)\right] Y_t = C_t + I_t + \mathcal{A}(u_t)\bar{K}_{t-1}$$

By Walras' Law the market for government bonds clears if all other markets clear.

4.5 Functional Forms

For estimation and the subsequent quantitative analysis I specify specific functional forms for the price adjustment cost function $\Phi(.)$, the investment adjustment cost function $S(.)$ and the capacity utilization function $\mathcal{A}(.)$ given by

$$\begin{aligned}
\Phi\left(\frac{P_{i,t}}{P_{i,t-1}}\right) &= \frac{\phi_p}{2}\left(\frac{P_{i,t}}{P_{i,t-1}} - \pi^*\right)^2 \\
S\left(\frac{I_t}{I_{t-1}}\right) &= \frac{\xi}{2}\left(\frac{I_t}{I_{t-1}} - \gamma\right)^2 \\
\mathcal{A}(u_t) &= \rho^* \frac{u_t^{1+\chi} - 1}{1+\chi}
\end{aligned}$$

where ϕ_p and ξ are parameters that control the magnitude of the adjustment costs of prices and investment. The parameter χ controls the curvature of the capacity utilization function.

The parameters ρ^* and π^* denote the steady state values of the rental rate and the inflation rate respectively.

4.6 Equilibrium Conditions and Solution Strategy

The characterization of the equilibrium conditions of the model is relatively standard and is relegated to Appendix A.1. The stochastic process for aggregate technology A_t introduces a source of long-run growth in the model. The equilibrium conditions are transformed into a stationary representation by dividing all real variables by the technology factor A_t and all nominal variables by the factor $P_t A_t$. Further details are presented in Appendix A.2. In what follows I use small case letters to refer to detrended variables, e.g. $x_t \equiv \frac{X_t}{A_t}$.

The computational strategy adopted in this paper relies on the concept of *Functional Rational Expectations Equilibrium (FREE)* employed in Krueger and Kubler (2004) and Malin, Krueger and Kubler (2007). The idea consists of finding a suitable set of functions defined over a compact set that satisfy the first order equilibrium conditions of the model. More precisely, I characterize the equilibrium in terms of the following five policy functions: $\mathcal{C} = \{L(\mathbb{S}), q(\mathbb{S}), \lambda(\mathbb{S}), i(\mathbb{S}), \pi(\mathbb{S})\}$, that correspond to hours worked, Tobin's q, marginal utility of wealth, investment and inflation, respectively.

The solution is assumed to be a time-invariant function of a minimum set of state variables $\mathbb{S}$. The state vector is formed by $\mathbb{S} = \left[R_{-1}, c_{-1}, \bar{k}_{-1}, i_{-1}, \mu, d, z, g, \varepsilon^r\right]$, where x_{-1} corresponds to the lagged value of the variable x, x denotes its current realization, and x' denotes future realizations. In total the model has $n = 9$ state variables. The choice of $\mathbb{S}$ is fundamental for the characterization of the equilibrium. Because the ZLB creates a kink in the monetary policy rule, the model has two steady-states, opening the possibility of multiple equilibrium dynamics (i.e. the control functions that satisfy the equilibrium conditions may not be unique). In fact this is the case in Aruoba, Cuba-Borda and Schorfheide (2013), who show that it is possible to construct a deflationary equilibrium and many non-fundamental equilibria in which the state vector is augmented by an extraneous stochastic process (a sunspot) that moves the equilibrium dynamics from the equilibrium with positive inflation to the equilibrium with deflationary dynamics. They find no evidence that the U.S. economy switched away from the target-inflation equilibrium during the Great Recession. For this reason I focus solely on the targeted-inflation equilibrium in my quantitative analysis.

Definition 1 *A FRE Equilibrium is defined by the compact set $\mathbb{S} \in \mathbb{R}^n$ and the set of control functions: $\mathcal{C} = \{L(\mathbb{S}), q(\mathbb{S}), \lambda(\mathbb{S}), i(\mathbb{S}), \pi(\mathbb{S})\}$ such that:*

$$\lambda(\mathbb{S}) = \beta R_t \mathbb{E} \frac{\lambda(\mathbb{S}')}{\pi(\mathbb{S}')} \frac{1}{\gamma e^{z'}} \tag{4}$$

$$\lambda(\mathbb{S}) = \frac{\gamma e^{d+z}}{\gamma c e^z - h c_{-1}} + h\beta \mathbb{E} \frac{e^{d'}}{\gamma c' e^{z'} - hc} \tag{5}$$

$$q(\mathbb{S}) = \beta \mathbb{E} \frac{\lambda(\mathbb{S}')}{\gamma e^{z'} \lambda(\mathbb{S})} \{\rho(\mathbb{S}')u' - \mathcal{A}(u') + (1-\delta) q(\mathbb{S}')\} \tag{6}$$

$$1 - e^{\mu} q(\mathbb{S}) \left[1 - S(\Delta i) - dS(\Delta i) x\right] = \beta \mathbb{E} q(\mathbb{S}') \frac{\lambda(\mathbb{S}')}{\lambda(\mathbb{S})} \frac{1}{\gamma e^{z'}} e^{\mu'} dS(\Delta i') x'^2 \tag{7}$$

$$(\frac{1}{\lambda_p} - 1) \left[1 - \Phi_p(\pi(\mathbb{S}))\right] - \frac{mc}{\lambda_p} + d\Phi_p(\pi(\mathbb{S}))\pi(\mathbb{S}) = \beta \mathbb{E} \frac{\lambda(\mathbb{S}')}{\lambda(\mathbb{S})} \Phi_p(\pi(\mathbb{S}'))\pi(\mathbb{S}') \frac{y'}{y} \tag{8}$$

$$w = \psi_L \frac{e^d L(\mathbb{S})^{\nu}}{\lambda(\mathbb{S})} \tag{9}$$

$$\rho = \frac{\alpha}{1-\alpha} \frac{L(\mathbb{S})}{\bar{k}_{-1} u} \gamma e^z \tilde{w} \tag{10}$$

$$u = d\mathcal{A}^{-1}(\rho) \tag{11}$$

$$mc = \alpha^{-\alpha} (1-\alpha)^{\alpha - 1} w^{1-\alpha} \rho^{\alpha} \tag{12}$$

$$c = \left[\frac{1}{\bar{g} e^g} - d\Phi_p(\pi(\mathbb{S}))\right] y - \mathcal{A}(u) \frac{\bar{k}_{-1}}{\gamma e^z} - i(\mathbb{S}) \tag{13}$$

$$\bar{k} = (1-\delta) \frac{\bar{k}_{-1}}{\gamma e^z} + \mu \left[1 - S(\Delta i)\right] i(\mathbb{S}) \tag{14}$$

$$y = k^{\alpha} L(\mathbb{S})^{1-\alpha} - \mathcal{F} \tag{15}$$

$$R = \max \left\{ 1, \left[(r^* \pi^*) \left(\frac{\pi(\mathbb{S})}{\pi^*} \right)^{\psi_1} \left(\frac{y}{y^*} \right)^{\psi_2} \right]^{\rho_R} R_{-1}^{1-\rho_R} e^{\varepsilon^r} \right\} \tag{16}$$

To simplify notation I use: $\Delta i = \frac{i(\mathbb{S})}{i_{-1}} e^z \gamma$, $d\Phi_p = \partial \Phi_p(x)/\partial x$,$dS = \partial S(x)/\partial x$ and $d\mathcal{A} =$

$\partial\mathcal{A}(x)/\partial x$, where x stands in for the argument of each function as defined in Section 4.5.

5 Quantitative Results

5.1 Computational Strategy

The FRE Equilibrium definition requires the solution of an infinite dimensional nonlinear system of equilibrium conditions in order to characterize the functions in $\mathcal{C}(\mathbb{S})$. The solution strategy adopted here uses two approximations. First, the compact state space $\mathbb{S}$ is represented using ergodic set methods as in Judd, Maliar and Maliar (2012) and Aruoba, Cuba-Borda and Schorfheide (2013). This reduces the problem to finding the best approximation to the true functions in $\mathcal{C}(\mathbb{S})$ using a grid $m_i \in \mathcal{M} \subset \mathbb{S}, \quad i = 1, \ldots, M$, that represent the region of the state space that is relevant to characterize the solution. Section 5.3 explains how to construct this grid such that it contains enough nodes where the ZLB binds. Second, the unknown equilibrium functions in $\mathcal{C}^{(j)} \in \mathcal{C}(\mathbb{S})$ are approximated by piece-wise continuous functions characterized by a set of coefficients $\theta \in \mathbb{R}^{2\times N}$. These coefficients are used to construct linear combinations of basis functions $\mathcal{T}_j : \mathbb{S} \rightarrow \mathbb{R}$ that are evaluated in each solution node. In particular I use Chebyshev polynomials, defined as $T_j(x) = \cos(j \times \arccos(x))$, where $x \in [-1, 1]$, which are combined using a *complete polynomial* rule in order to form the multidimensional basis function $\mathcal{T}_j$.[13]

Because I will look for an approximate solution to the functional equations that satisfy the equilibrium conditions, a criterion that informs about how close my approximation is to the “true” solution is needed.[14] The metric for the approximation is given by a set of residual functions $\mathcal{R}(\mathbb{S})$ that are obtained from the equilibrium conditions (4)-(8). For example, the

[13] Additional details of the construction of the basis functions is provided in Appendix B.

[14] Note here that because of the multiplicity of equilibria, constructing the approximate function numerically requires an initial guess that converges to the desired equilibrium. It turns out that using the decision rules of the linearized model, as the initial guess, the nonlinear decision rule converge to the dynamics of the equilibrium with positive steady state inflation.

residual for Equation 4 is:

$$\mathcal{R}_1(\mathbb{S}) = \lambda(\mathbb{S}) - \beta R \mathbb{E} \frac{\lambda(\mathbb{S}')}{\pi(\mathbb{S}')} \frac{1}{\gamma e^{z'}}.$$

Appendix B explains how to construct all the residual functions used to solve the nonlinear model. To evaluate the expectations that appear in the residual functions, I use a sparse-grid approximation based on the integration rules discussed in Heiss and Winschel (2006).

5.2 Parameter Estimation with Pre-ZLB data

A estimation exercise of the model subject to the ZLB constraint is a computationally intensive enterprise because it requires the nonlinear solution to be computed for a large number of parameter vectors. Instead I follow a two-step procedure. First I estimate a log-linearized version of the model using data prior to the ZLB episode from *1984:Q1 - 2008:Q3*. I estimate the model parameters using a first-order approximation of the DSGE model equilibrium conditions, and characterize the posterior distribution of the parameters using the Random Walk Metropolis-Hastings algorithm described in An and Schorfheide (2007). Conditional on the parameters obtained from the pre-ZLB period, I solve the model enforcing the ZLB and use Sequential Monte Carlo methods to extract the underlying states and shocks corresponding to the period *2008:Q2 - 2013:Q4*.

Some parameters are fixed before the estimation because the likelihood is not informative with respect to them. The parameter ζ is set to 0.22 in order to match the long-run average ratio of government consumption expenditure to Gross Domestic Product observed in NIPA data from 1960-2013. The parameter λ_p is fixed at 0.1667, implying a steady state price markup of 20%. This value is slightly lower than that estimated in medium-scaled DSGE models, which find a steady price-markup of 28%. Since I do not use data on hours worked for estimation I set the parameter $\nu = 1$, implying a Frisch elasticity of labor supply equal to one. This value is large with respect to the microeconomic evidence for the elasticity along the intensive margin reported in Chetty et al. (2013). Nevertheless it is within the range of estimated values obtained by Rios-Rull et al. (2012). I normalize the steady state level of hours worked to 1/3 using the parameter ψ_L.

Data. I use quarterly data on five macroeconomic variables covering the period 1984:Q1 to 2008:Q3.[15] I map the model variables to data on output growth, consumption growth, investment growth, inflation and the nominal interest rate. I use data on Gross Domestic Product (GDP) to measure output growth. Consumption is the sum of personal consumption of non-durable goods (PCND) and services (PCND). Investment includes the personal consumption of durable goods (PCDG), fixed private investment (FPI) and the change in inventories (CBI). All these series are scaled by civilian non-institutionalized population over sixteen years (CNP16OV) and deflated using the implicit GDP price deflator (GDPDEF). Growth rates are computed as the one period log difference expressed in percentages. Inflation is computed as the percentage log-difference of the implicit price deflator. Finally the nominal interest rate is measured using the quarterly average of the Federal Funds Rate (FEDFUNDS).

Table 1 presents parameter estimates based on the results of the MH simulator. I obtain 100,000 draws of parameters from the posterior distribution and construct summary statistics of the posterior distribution based on the last 50,000 draws of the sequence. A few results from the estimation are discussed next. The estimated value of the share of capital in the intermediate firms' production function is 0.18. In estimated DSGE models this parameter is usually below the commonly used value of 0.33 obtained from long-run averages of the capital share in aggregate output. With respect to the parameter h that controls the consumption habit and the persistence of consumption with respect to nominal shocks, I obtain a value of 0.55. This degree of habit persistence is relatively modest compared to the most recent estimates of Christiano, Eichenbaum and Trabandt (2014).

The parameter χ, which controls the elasticity of capacity utilization with respect to the

[15] All the series were extracted from the FRB St. Louis FRED Database, with the original name of the data series shown in parenthesis. Additional details are provided in Appendix C.

rental rate of capital, is an important parameter in determining the persistence of inflation in response to demand shocks. There is wide range of variation in previous estimates of this parameter. For example, Christiano, Eichenbaum and Evans (2005) assume a very small value of 0.01 whereas Justiniano, Primiceri and Tambalotti (2010) obtain an estimate in the range of 3 to 7. My estimate for χ is closer to the latter, implying that the rental rate is very sensitive to changes in capacity utilization. This implies that, all else equal, marginal costs will respond strongly to movements in capital utilization, making inflation less persistent.

The parameter controlling nominal rigidities is an important one for the transmission of shocks in the model and deserves additional attention. I estimate the price adjustment cost parameter ϕ_p indirectly using the implied slope of the Phillips curve $\kappa(\phi_p)$. The estimate of the slope in the Phillips curve is steep, with a value of $\kappa = 0.21$ at the posterior mean. To give a sense of the degree of price stickiness implied by the model, I compute the associated frequency of price adjustment in a first order approximation of the Phillips curve derived under Calvo pricing. My estimates imply that firms in a Calvo setting would adjust prices roughly every three quarters, which is on the lower end of price rigidity, compared to the four to six quarters commonly obtained in the DSGE literature. Nonetheless, the posterior credible set of the estimated Phillips curve parameter is consistent with the range wide range of values reported in Schorfheide (2008).

I use informative priors to estimate the parameters of the monetary policy rule. There is a significant amount of persistence in the determination of the interest rate reflected in the estimate of ρ_R. The response of the nominal rate to inflation is within the range of estimated parameters in the literature. The output gap response parameter seems low because the policy rule is expressed in terms of quarterly percentage deviations of output with respect to its balanced growth path. I also estimate the parameters that control the growth rate of technology and the inflation rate along the balanced growth path. They imply a long-run rate of growth of 2% and a long run inflation rate of 2.3% in annualized terms.

Table 1: DSGE Model Parameters

Parameter	Description	Prior Mean	Posterior Mean	90% Credible Sets		Prior	Prior SD
h	Habit persistence	0.5	0.5467	0.4582	0.6328	B	0.10
α	Capital share	0.3	0.1806	0.1607	0.1991	N	0.05
ξ	Inv. Adj. Cost	4.0	4.0491	2.5816	5.5861	G	1.00
χ	Capital Utilization Cost	5.0	5.2988	3.7511	6.9456	G	1.00
$\kappa(\phi_p)$	Phillips Curve	0.3	0.2127	0.0994	0.3262	G	0.20
Taylor Rule:							
ρ_r	Smoothing	0.5	0.7310	0.6756	0.7833	B	0.20
ψ_1	Inflation Response	1.5	1.6758	1.5299	1.8157	N	0.10
ψ_2	Output Gap Response	0.005	0.0748	0.0297	0.1154	N	0.05
Shock Process:							
ρ_z		0.4	0.0933	0.0134	0.1669	B	0.20
ρ_g		0.6	0.9886	0.9794	0.9983	B	0.20
ρ_{mu}		0.6	0.7000	0.5982	0.8117	B	0.20
ρ_d		0.6	0.9475	0.9170	0.9829	B	0.20
$100\sigma_z$		0.2	0.9171	0.7333	1.1097	IG	1.00
$100\sigma_g$		0.5	0.2746	0.2428	0.3052	IG	1.00
$100\sigma_\mu$		0.5	3.8763	2.4248	5.3372	IG	1.00
$100\sigma_d$		0.2	1.2436	0.8029	1.6882	IG	1.00
$100\sigma_r$		0.2	0.1746	0.1476	0.2004	IG	1.00
Balanced Growth Path:							
$\gamma^{(q)}$	Long Run Growth	0.5	0.4978	0.4604	0.5343	N	0.025
$\pi^{(q)}$	Inflation Rate	0.5	0.5734	0.4302	0.7169	N	0.10
$100(\beta^{-1}-1)$	Discount rate	0.3	0.1848	0.1010	0.2646	G	0.10
Implied Parameters:							
β	Discount Factor		0.9982				
ϕ_p	Price Adj. Cost		15.1062				
$rr^{(q)}$	Real rate		2.7255				

Notes: The parameters were estimated using 1984:Q1-2008:Q3 data. The credible sets are obtained from the 5th and 95th percentiles of the posterior distribution.

Fit of the Estimated Model. Because I estimate the model using a Bayesian framework, one way to check the empirical fit of the model is through posterior predictive checks. I use simulated trajectories from the model and draws from the posterior distribution of parameters to construct a set of statistics $\mathcal{S}(\tilde{\mathbb{Y}}^T) \in \mathbb{R}^n$. The same sample statistic can be constructed using observed data $\mathcal{S}(\mathbb{Y}^T)$. If the observed sample statistic lies far in the tail of the predictive distribution, this indicates that the model has trouble capturing the data along that dimension.

In Figure 3 I show the posterior predictive checks for the empirical distribution of four statistics $\mathcal{S}(\tilde{\mathbb{Y}}^T)$: the mean, standard deviation, first order autocorrelation and the correlation with GDP growth.[16] The posterior predictive checks reveal that the estimated model captures well the mean of all the observed series except that of investment growth. The reason is that neutral technology shocks are the only source of long-run growth in the model economy. Hence all variables grow at the rate (γ) along the balanced growth path. However, in the data investment has a different long-run growth rate. A simple way to resolve this discrepancy is to introduce an additional source long-run technological progress that only affects investment and that can disentangle the long-run growth rate of both variables. With respect to standard deviations and first order autocorrelations the DSGE model does well in matching the empirical counterparts, although on average it tends to over predict the autocorrelations. A similar picture emerges for the correlation with output growth.

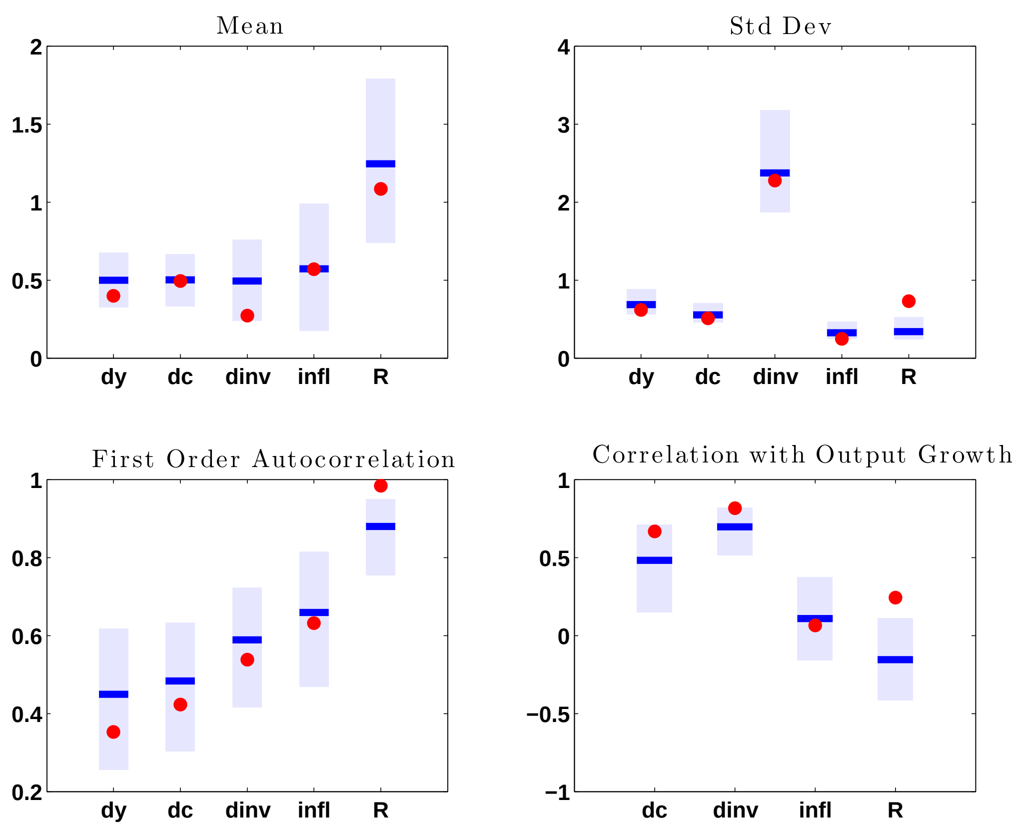
Lastly the model tends to underpredict the second moments of the nominal interest rate. This could be explained if in practice if the monetary authority responds more strongly to periods of low economic activity compared to periods of high economic activity. For example, in the face a negative output gap, the Fed may decide to lower the nominal interest rate faster. This would generate greater volatility in the observed nominal interest. Similarly, when the economy is in midst of a recovery, the Fed may decide no to increase the nominal rate too fast to avoid halting the expansion. This would introduce a higher autocorrelation

[16]The algorithm used to construct the predictive distribution is discussed in Appendix B.5.

in observed interest rate series. In fact, Aruoba, Bocola and Schorfheide (2013) find evidence that supports the view that the Fed adjusts the nominal interest rate asymmetrically. My specification of the policy rule cannot capture none of these dynamics. Nonetheless, the estimated model with pre-ZLB data captures important features of the dynamics of output, consumption, investment, inflation and the nominal rate.

Taking as given the estimated parameters at their posterior mean, I move to explain how to incorporate the ZLB into the solution of the model, which is the backbone of the quantitative exercise of Section 7.

Figure 3: Posterior Predictive Checks



Notes: The red dot corresponds to the observed statistic. The dark-blue horizontal bar is the mean of the simulated statistic. The light-blue bands correspond to the 5th and 95th percentiles of the posterior predictive density.

5.3 Incorporating the Zero Bound

Before moving to the quantitative analysis, I solve the nonlinear model that incorporates the zero bound. Using the computational strategy described in Section 4.6 and the posterior mean of the parameters in Table 1, I approximate the FRE equilibrium on a grid constructed using simulation based methods. As in Judd, Maliar and Maliar (2012) I construct a representation of the ergodic set of the model using a clustered grid algorithm. However, as emphasized in Aruoba, Cuba-Borda and Schorfheide (2013), the essential ergodic set does not capture events where the ZLB is binding. Following their computational strategy, I augment the essential ergodic set with grid points that capture binding ZLB periods in the U.S. experience from 2009:Q1-2013:Q4. The additional grid points are obtained from the filtered distribution of states during this period.[17]

Decision rules. The functional equations that characterize the equilibrium dynamics during normal times when the ZLB is not binding (nb) and when the ZLB is binding (b) are parameterized by the vector of unknown coefficients $\Theta = \{\theta^{L,r}, \theta^{q,r}, \theta^{\pi,r}, \theta^{\lambda,r}, \theta^{i,r}\}$, where $r = \{b, nb\}$ denotes one of the two possible regimes of the nominal interest rate.

I use piece-wise smooth functions as in Aruoba, Cuba-Borda and Schorfheide (2013) because they provide a flexible approximation that allows each of the control functions to inherit the kink induced by the zero bound constraint. For example, consider the approximation of the marginal utility of wealth

$$\lambda(\mathbb{S}) \approx \begin{cases} \sum \theta_j^{\lambda,nb} \mathcal{T}_j(\mathbb{S}) & \text{if } R(\mathbb{S}) > 1, \quad j = 1, \ldots, N \\ \sum \theta_j^{\lambda,b} \mathcal{T}_j(\mathbb{S}) & \text{if } R(\mathbb{S}) = 1, \quad j = 1, \ldots, N \end{cases}$$

A total of $2 \times N$ coefficients and basis functions are used to approximate the decision rules over the grid of points $\mathcal{M}$. The piece-wise smooth approximations consist of using one set

[17] The construction of the filtering distributions is explained in Section 6.1, and additional details are provided in Appendix D.

of coefficients $\theta_j^{\lambda,nb} \in \mathbb{R}^N$ to approximate the functional equation in the regions of the state space where the zero bound is not binding, while a second set of coefficients $\theta_j^{\lambda,b} \in \mathbb{R}^N$ approximates the decision rule when the constraint is active.

Using a third order approximation with a complete basis of Chebyshev polynomials, each element of Θ contains $N = 220$ unknown coefficients; hence a total of $2,200$ unknowns need to be solved numerically. The objective function that pins down the unknown coefficients is given by the sum of squared residuals of the equilibrium conditions $\mathcal{R}(\mathbb{S}, \Theta)$ evaluated at the $M = 600$ grid points, out of which 440 correspond to the essential ergodic set obtained from simulating the model, and the remaining 160 points are obtained from the filtered states described previously. I use a Newton-based solver to find the coefficients that $\min_\Theta \sum_{i=1}^{M} \mathcal{R}(\mathbb{S}_i, \Theta)^2$.

Accuracy of nonlinear solution After obtaining the vector of coefficients Θ, I check the accuracy of the solution using the bounded rationality measure from Judd (1998), also known as Euler Equation errors. This approach scales the approximation errors $\mathcal{R}(\mathbb{S}, \Theta)$ as a fraction of current consumption and expresses them in terms of unit-free quantities. Table 2 shows a summary of the Euler error accuracy measure computed for all residual functions. The average approximation error is on the order of 10^{-3}, which means that the representative agent's loss from following the approximate decision rules is 0.1 cents for every dollar spent.[18]

Table 2: Euler Equation Errors

	Euler Equation	**Marginal Utility of Wealth**	**Capital Euler Equation**	**Investment Equation**	**Pricing Equation**
Mean	-3.00	-2.63	-3.29	-2.82	-2.35
Min	-6.14	-5.74	-5.33	-5.88	-5.52
Max	-2.14	-1.70	-2.48	-1.80	-1.54

[18] Figure A-1 shows the full distribution of the Euler equation errors.

6 What Caused the Great Recession?

This section investigates the forces that explain the dynamics of quantities and prices during the Great Recession. I use the implied equilibrium dynamics of the model to match U.S. data from 2000:Q1 to 2013:Q4. A key challenge in understanding the forces driving the events in the aftermath of the financial crisis in 2008:Q3 is backing out the structural shocks that rationalize the data. This is a complicated task because the presence of the ZLB renders the model highly nonlinear.

Prior to the Great Recession, the vast majority of DSGE models ignored the ZLB because at the time, it seemed very unlikely that the U.S. economy would ever face with shocks large enough to make this constraint binding.[19] Without the ZLB, and equilibrium dynamics are well approximated with a linear state-space system, the estimation of structural innovations can be performed using the Kalman filter, for example as in the exercise of Bauer, Nicholas and Rubio-Ramirez (2003). In my application the structural innovations are Gaussian but the ZLB causes the dynamics of the economy to be highly nonlinear, which complicates inference about the underlying unobserved variables.[20]

I tackle the inference problem using Sequential Monte Carlo methods (SMC) to numerically approximate the distribution of states that rationalizes the sequence of observed output, consumption, and investment growth as well as inflation and the nominal interest rate during the period 2008:Q3 - 2013:Q4.[21] The exercise is similar to Aruoba, Cuba-Borda and Schorfheide (2013) in terms of extracting the filtered states and shocks, but in addition in this paper I use the estimated structural shocks to perform counterfactual exercises about the evolution of observed macroeconomic variables.

[19]Assuming the ZLB away has the added benefit that linear approximations are enough to characterize equilibrium dynamics, which allows the use of fast solution methods that open the door for estimating medium scale DSGE models.

[20]An example of a New Keynesian model with a linear structure but with non-Gaussian innovations is studied in Curdia, Del Negro and Greenwald (2013).

[21]Herbst and Schorfheide (2014) provide a detailed description of SMC methods, as well as practical guidelines for implementing different algorithms.

6.1 Inference of unobserved states

I briefly discuss the inference problem and sketch the mechanics of the filtering algorithm used to estimate unobserved states in nonlinear models. The solution of the system of equilibrium conditions of the model has the following nonlinear state-space representation of the dynamics of the endogenous variables

$$\mathbb{S}_t = g\left(\mathbb{S}_{t-1}, u_t\right), \quad u_t \sim N(0, \Sigma_u) \tag{17}$$

$$\mathbb{Y}_t = m(\mathbb{S}_t) + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma_\varepsilon), \tag{18}$$

Where (17) is the transition equation that describes the evolution of the state variables $\mathbb{S}_t$ as a function of the previous position of the system $\mathbb{S}_{t-1}$, and the realization of structural shocks u_t. The evolution of observable variables to the model state variables is given by the measurement equation (18). It is augmented with disturbances ε_t which represent measurement error that may create a discrepancy between the model implied series and their observed counterparts. I set the variance of the measurement error in (18) to 10% of the sample variance in the observables. The introduction of measurement error is not an arbitrary device to increase the fit of the model, is essential for the evaluation of the observation density described below.

Consider s_t and y_t as realizations generated by the nonlinear system. Let Y^t denote a time series of observations from $1, \ldots, t$. The state-space system described above induces an *observation density* $p(y_t|s_t, Y^{t-1})$ and a *transition density* denoted $p(s_t|s_{t-1}, Y^t)$. The filtering problem consists of learning about the realizations of the unobserved states and shocks $\{s_t, u_t\}$ given the sequence of observations Y^t. In other words, the objective is to characterize the shape of the *filtering density* $p(s_t|Y^t)$. In a nutshell, SMC methods start with a discrete approximation of the filtering density $p(s_{t-1}|Y^{t-1})$ characterized by a collection of particles (a swarm) $\left\{\pi_{t-1}^{(i)}, W_{t-1}^{(i)}\right\}_{i=1}^{N_p}$, where N_p is the number of particles, and use the information contained in the current observation y_t together with the state transition equation of the model to update the particles, creating a new particle swarm $\left\{\pi_t^{(i)}, W_t^{(i)}\right\}_{i=1}^{N_p}$ such that the following approximation holds for any function $h(s_t)$: $\frac{1}{N_p}\sum_{i=1}^{Np} h(s_t^{(i)})W_t^{(i)} \approx \int h(s_t)p(s_t|Y^t)ds_t$. For this particular application I set $N_p = 100,000$ particles and start the the approximation to $p(s_0|Y^0)$ using simulations from the solution of the model.

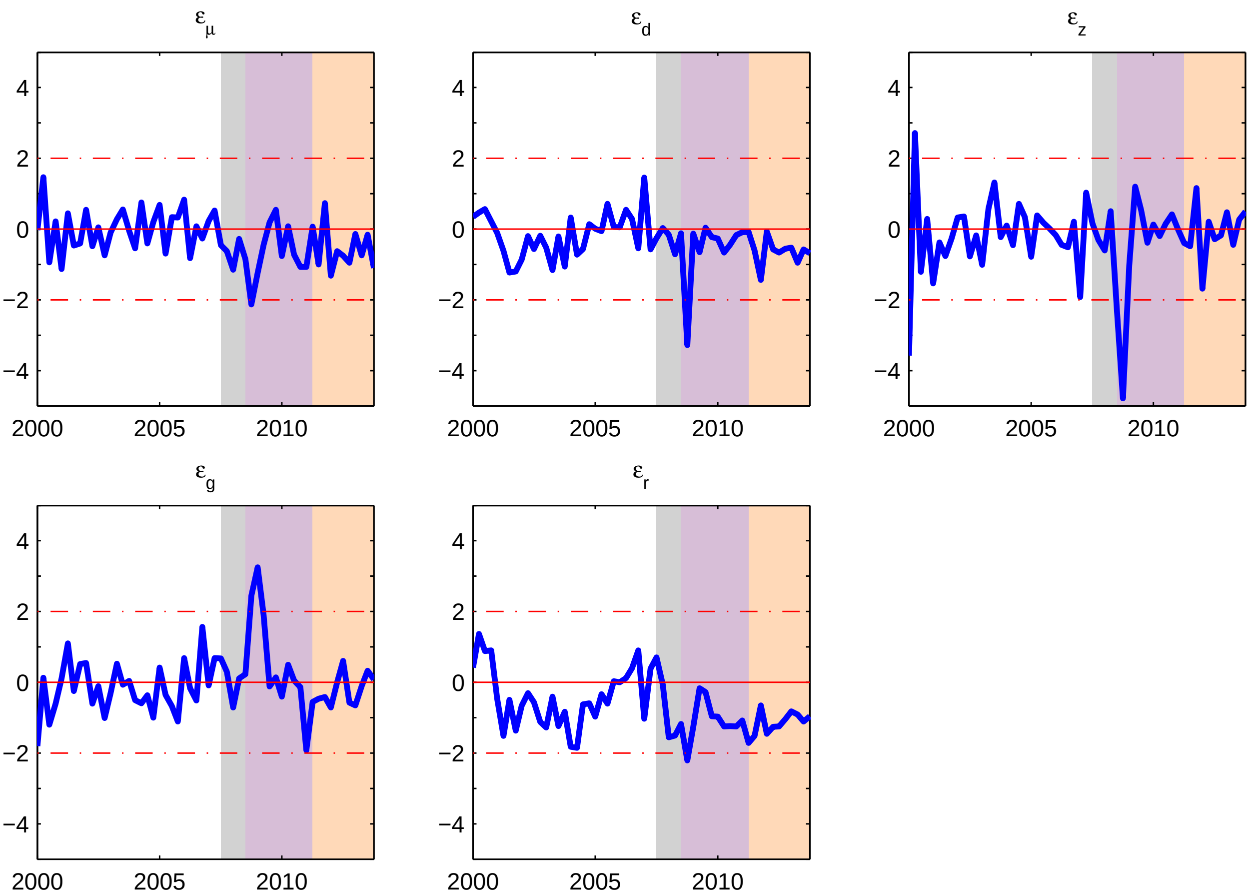
6.2 What shocks explain the Great Recession?

I recover the structural shocks from the approximation of the filtering density $p(s_t|Y^t)$. Figure 4 presents the mean of the filtered shocks $\mathbb{E}(s_t|Y^t)$ for *t=2000:Q1* to *t=2013:Q4*, depicted as thick blue lines. Feeding these shocks through the system (17) - (18) I recover the observed evolution of consumption, investment and output, as well as the dynamics of inflation and the nominal interest rate. With the exception of 2008:Q3, when financial distress was at its highest, all the innovations are within two standard deviations in both the pre and post recession periods. This is a common feature in New Keynesian models, because the ZLB is a low frequency event, occurring about 0.1% of the time in the ergodic distribution of the model, requiring a large initial impulse in order to explain the sharp decline in economic activity.

Preference and Investment Shocks. The left-top panel of Figure 4 shows that the economic downturn originated with a sequence of negative shocks to the marginal efficiency of investment (μ_t). The negative sequence of investment shocks started in the second half of 2007 and continued through the second half of 2008. I interpret this sequence as a slow build-up of frictions affecting financial intermediation prior to the financial crisis. The negative investment shocks prior to 2008:Q3 capture the decline in consumption of durable goods and investment in residential structures which were the components of aggregate demand that were affected with the deterioration in the housing market at the onset of the recession.

In the third quarter of 2008, three negative shocks are necessary to explain the contraction in macroeconomic aggregates. My estimates indicate that the deterioration of the marginal

Figure 4: Structural Innovations (in standard deviations)



Notes: The gray shadow covers the period 2007:Q4 to 2008:Q2. The purple shadow corresponds to the period 2008:Q3 to 2011:Q2. The light orange shadow corresponds to the period 2011:Q3 to 2013:Q4.

efficiency of investment was coupled with a large negative shock to preferences and an even larger shock to total factor productivity. Although the economic downturn began in 2007:Q3, the size and the correlation of the shocks in 2008:Q3 imply that the forces that pushed the economy into a binding zero lower bound and deepened the recession were unusual and unexpected. In models without capital, explaining such an event requires even larger innovations of the structural shocks. For example in Fernández-Villaverde et al. (2012) a ZLB episode of a duration of four quarters is generated with a negative shock to the discount factor of eight standard deviations in size. Similarly Gust, Lopez-Salido and Smith (2012), using a small New Keynesian model without capital, explain the decline in output and the prolonged ZLB episode, using a negative shock to households' discount factor of roughly five standard deviations.

From an ex-ante perspective, the combination of shocks that generate the ZLB event in my model has a probability of (0.0001%).[22] Compared to a single shock of five standard deviations, my results are not very different in terms of the extreme nature of the initial impulses. However, as I discuss next, trying to match the decline in consumption and investment using only shocks to preferences would be impossible.

Why does the model need shocks to both preferences and investment? The reason is that in models with capital, preference shocks trigger wealth effects that generate a comovement problem between consumption and investment. This is illustrated in Figure 5, where I plot the impulse responses to the preference and investment shocks away from the ZLB. In 2008:Q2, before the financial crisis unfolded, the Federal Funds rate was 2.1%. Because negative preference shocks shift output away from consumption towards investment, while negative shocks to the marginal efficiency of investment produce the opposite effect, the model cannot match simultaneously the decline in both components of aggregate demand with a single large shock.

Figure 5 highlights the counterfactual response in investment following negative preference shocks.[23] To further illustrate this point, I re-solve the model, shutting down the investment shock while keeping the other parameters at their estimated values. I then repeat the exercise with the preference shock.[24] Then I filter U.S. data again and calculate the marginal contribution to the log-likelihood of each of the observations in the period 2007:Q4-2009:Q1. I present the results of this exercise in in Table 3. A more negative number indicates a worse fit of the model for a given observation. Note that in all periods trying to explain the data

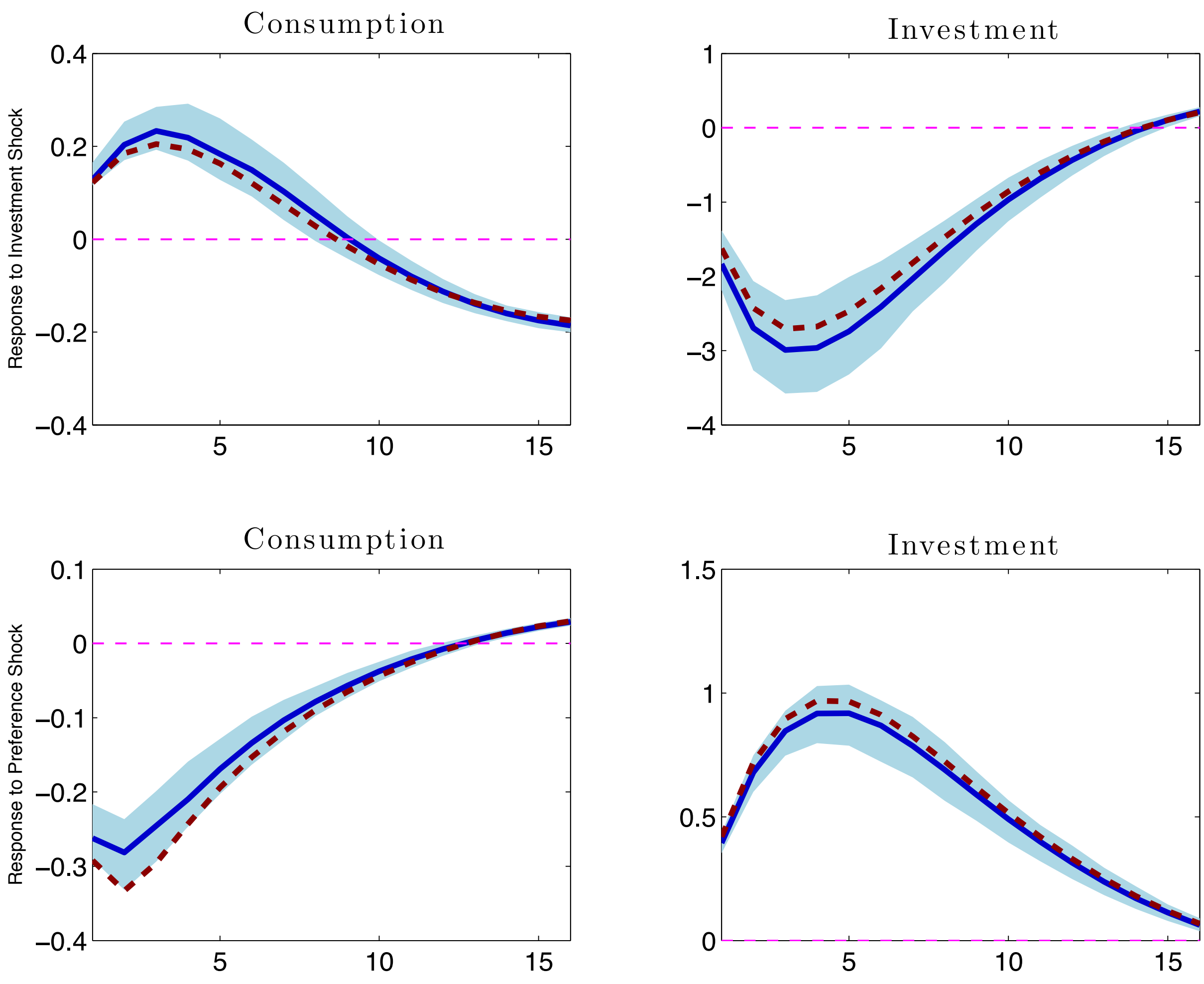
[22]This corresponds to the probability of drawing a negative $2.1\sigma_{\epsilon_\mu}$ and a negative $2.6\sigma_{\epsilon_d}$, which are the shocks that deliver the ZLB episode of 2009:Q1.

[23]Resolving the comovement problem remains a challenge for DSGE models. This issue was originally pointed out by Barro and King (1984) and touches the heart of the debate on the sources of business cycle fluctuations.

[24]Specifically in the first experiment I set $\sigma_\mu = 0.0001$, while for the second experiment I set $\sigma_d = 0.0001$

with only investment shocks or only a preference shock worsens the likelihood. Early in the recession, from 2007:Q4-2008:Q2, the model without the preference shock doesn't do as poorly as the model without the investment shock. The reason is that early in the recession there was a slight increase in consumption but a decline in investment (see Figure 1) so one shock is enough to match both dynamics. In 2008:Q3, both investment and consumption plunged, so trying to match the decline in both variables with a single shock is impossible, as illustrated by the sharp deterioration in the log-likelihood.

Figure 5: Comovement Problem Away from the ZLB



Notes: The dark blue lines correspond to the impulse responses of the nonlinear model. The light blue shade denote the 20%-80% confidence interval. The red dashed lines are obtained from the linear solution of the model ignoring the ZLB. All impulse responses correspond to a one standard deviation shock.

Table 3: Marginal Contribution to the Log-Likelihood

	Baseline	Without Preference Shock	Without Investment Shock
2007:Q4	-1.6	-1.4	-9.6
2008:Q1	-4.8	-3.6	-43.4
2008:Q2	-3.2	-12.3	-24.3
2008:Q3	-6.0	-8.9	-34.8
2008:Q4	-33.4	-207.2	-149.6
2009:Q1	-10.6	-70.7	-121.7

Technology shocks The top right panel of Figure 4 shows the evolution of the innovation to technology growth. There is an enormous negative spike in 2008:Q4, of approximately 4.8 standard deviations, followed by a second negative shock of around 1 standard deviation in 2009:Q1. To put these shocks in context, I compute an equivalent direct measure of innovations to technology using quarterly data on total factor productivity (TFP) for the U.S. business sector.[25] Figure 6 shows the comparison between the filtered innovations to technology and their directly measured counterpart.[26] The correlation between the filtered and observed measure of technology shocks is 0.52, and there is also a clear negative TFP shock in 2008:Q4 observed in the data.

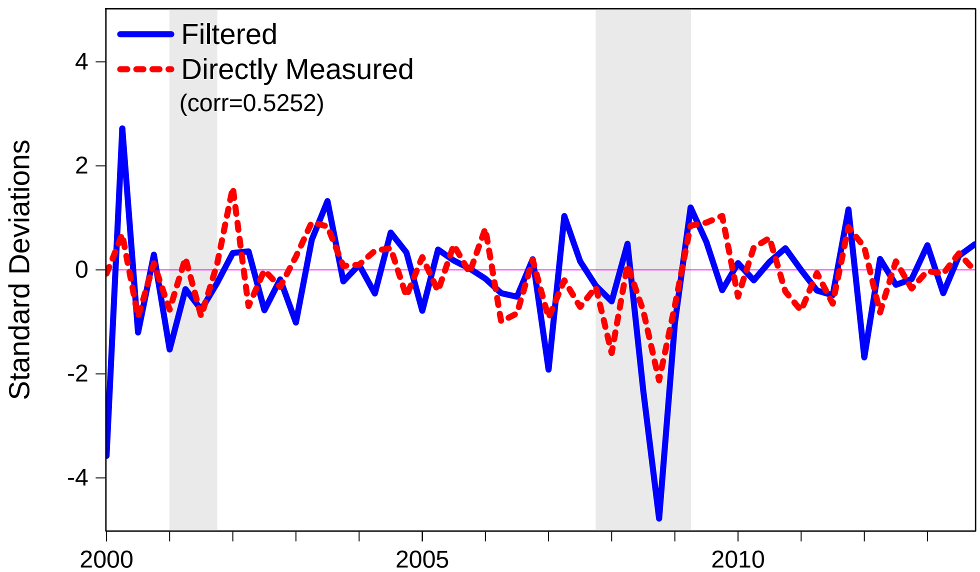
Why are the filtered innovations to technology larger in this particular period? The reason is the absence of persistent deflation in the data. When the ZLB is binding, prices fall due to a self-reinforcing loop of high real interest rates, declining marginal costs and low aggregate demand, so explaining the absence of a more severe deflation in the U.S. remains a puzzle. After 2008:Q3, inflation turned negative only in the first two quarters of 2009 bouncing back to positive territory thereafter. Although inflation has remained below the Federal

[25] This measure is produced by the Federal Reserve Bank of San Francisco, and methodological details are presented in Fernald (2012).

[26] I fit an AR(1) process with drift to the quarterly growth rate of observed TFP, exactly as in the description of the model, and report the fitted residuals.

Reserve's target of 2% in the aftermath of the recession, the U.S. economy has avoided a deflationary spiral. The negative TFP shocks observed in the data explain in part why prices did not fall more dramatically and persistently as the economy approached the ZLB. Negative technology shocks increase marginal costs counteracting the deflationary pressures at the ZLB, and if the shocks are negative enough it is possible to reproduce inflation and align the model prediction with the data.

Figure 6: Filtered and Directly Measured Innovations to TFP



Notes: The dashed red line is the direct measure of technology shocks are obtained from the TFP series discussed in Fernald (2012). The solid blue line is the mean filtered state z_t obtained from $\mathbb{E}(s_t|Y^t)$. The gray areas indicate NBER recession dates.

Using a different time series representation for technology growth, Christiano, Eichenbaum and Trabandt (2014) also rationalize the absence of deflation assuming a one-time negative productivity shock in 2008:Q3. My estimates of the unobserved technology shocks (z_t) square well with those obtained from direct measurement of TFP and using a more common representation of the stochastic process for it. The advantage of my filtering procedure is that I let the data speak through the model to recover the shocks more transparently, for instance I do not need to assume the size of the negative technology shock in 2008:Q3, nor to assume that it has to be one-time innovation in that particular period. The fact that this is actually the case is a result I obtain from the estimation.

Fiscal and monetary policy shocks. Turning to the fiscal policy side, two positive shocks in 2009:Q1 and 2009:Q2, and a large negative shock in 2011:Q2 stand out. These events correspond to well identified events related to changes in discretionary fiscal expenditures. The positive shocks correspond to the *American Reinvestment and Recovery Act*, which was enacted quickly after the collapse in the financial sector.[27] The negative shocks observed from 2010:Q4 and continuing until 2012:Q3, correspond to the prolonged struggle about the stance of fiscal policy between the White House and Congress. Examples of this struggle include the 2010 year-end debate about extending tax cuts and federal unemployment benefits, the fiscal entrapment surrounding the Federal debt limit in the first half of 2011, the downgrade in the credit rating of U.S. federal government debt in August 2011, and concerns about the fiscal cliff in early 2013.

To emphasize that the recovered filtered shocks capture well the developments in actual U.S. fiscal policy, Figure 7 compares the filtered model estimate of g_t to a direct measure of the government shock based on the observed ratio of government consumption and investment to GDP. The empirical measure of g_t is recovered using the same equation that describes the evolution of government expenditure in the model

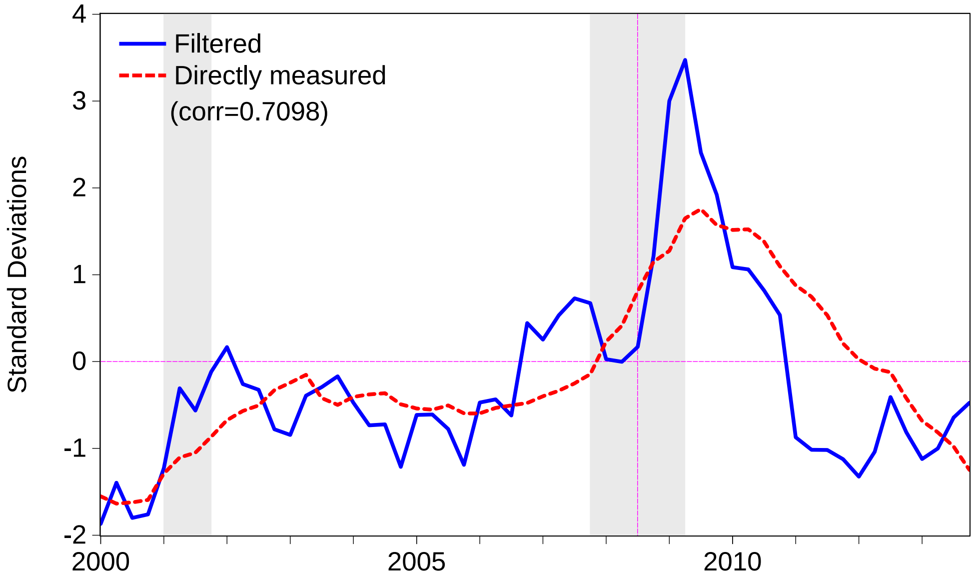
$$\frac{G_t}{Y_t} = \left(1 - \frac{1}{g_t}\right)$$

The correlation between the empirical measure of g_t and the filtered measure is slightly over 0.7.[28] Fiscal policy was clearly expansionary from 2009:Q1 until 2011:Q2, and tightened afterwards. As shown in Section 7, this fiscal swing had a negative effect on the economic recovery.

[27]The "stimulus bill" was signed into law in February 17, 2009, less than a month after the change in U.S. presidential administration.

[28]Note that I do not use the share of government consumption G_t/Y_t when estimating the filtered shocks. I use only the same data on consumption, investment and GDP growth. Nevertheless, the sequence of filtered g_t shocks tracks well its data counterpart. The latter is constructed using quarterly information on government consumption and expenditure including gross investment.

Figure 7: Fiscal Policy Shocks: Filtered vs Data



Notes: The dashed red line corresponds to the empirical measure of the g_t process. The solid blue line is the mean filtered state g_t obtained from $\mathbb{E}(s_t|Y^t)$. Data on government consumption and gross investment used to construct g_t comes from Table 1.1.5 of the National Income and Product Accounts. The gray areas indicate NBER recession dates.

Turning to monetary policy, from 2007:Q3 on, there is a long sequence of negative shocks to the monetary policy rule. During the period 2007:Q3-2008:Q4, these shocks reflect an aggressive response by the Federal Reserve to counteract contractions in aggregate demand. However, once the economy hits the ZLB in 2009:Q1, the continued sequence of negative monetary policy shocks observed thereafter have a more subtle interpretation.

As early as December of 2008, the Federal Open Market Committee adopted a communication language that suggested an active policy decision of maintaining the nominal interest near zero for a substantial period of time.[29] The FOMC adopted more explicit language in 2011:Q3, when it announced that the federal funds rate would likely remain at zero until mid-2013. In 2012:Q1 the reference date for the *forward guidance* policy was extended until late 2014. My model does not take into account the shift towards an explicit forward

[29]For instance, the first FOMC statement of 2009 indicates that *economic conditions are likely to warrant exceptionally low levels of the federal funds rate for some time.*

guidance policy. However, the negative monetary policy shocks obtained after 2009:Q1, and in particular during the 2011:Q3-2013:Q4, reflect the commitment of the Federal Reserve to maintain a zero interest rate policy even if the monetary policy rule would otherwise have called for an earlier *lift-off* in the nominal rate.

6.3 Interpreting the structural shocks

I complement the evidence presented in the previous section by comparing the filtered shocks with some direct measures of the underlying financial frictions that caused the recession. Instead of focusing on the filtered innovations, ϵ_t^{μ} and ϵ_t^{d}, I transform them into the implied paths of the marginal efficiency of investment μ_t and the subjective discount factor adjusted by the preference shock, $\tilde{\beta} = \beta \frac{d_t}{d_{t-1}}$. The latter reflects the effective level of patience or impatience of individuals ex-post and is related to households' desire to save.[30] To aid the interpretation, I compare the path of the marginal efficiency of investment to the observed evolution of credit spreads, and the path of the adjusted discount factor to the observed evolution of U.S. personal savings rate.

In Figure 8, I compare the paths of the marginal efficiency of investment (on an inverted scale) and credit spreads. The measure of credit spreads comes from the corporate bond spread measure constructed by Gilchrist and Zakrajšek (2012) (GZ).[31] I scale both the marginal efficiency of investment and the credit spreads data in terms of standard deviations to plot them on the same scale.

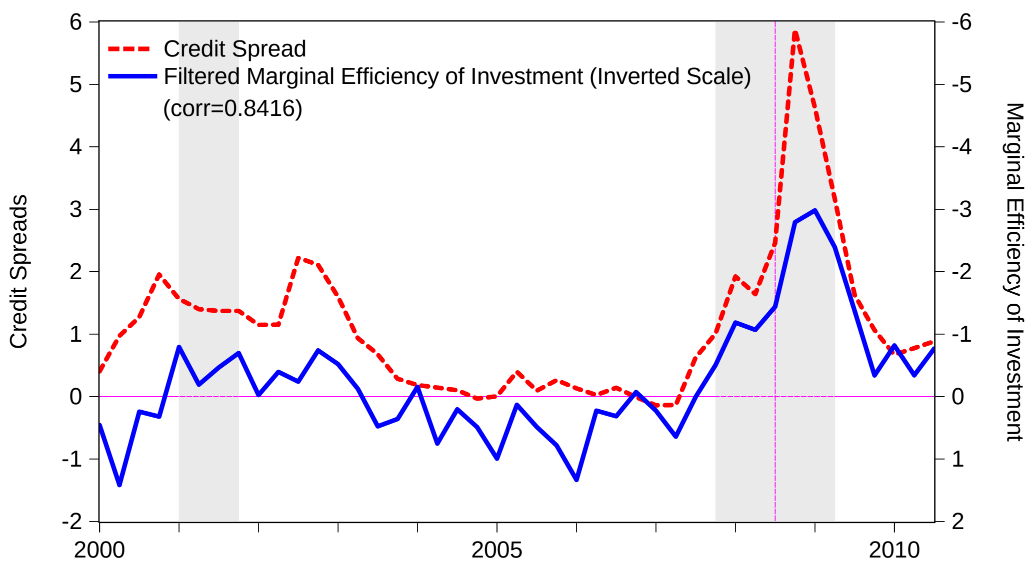
As credit spreads rose with the onset of the recession, the marginal efficiency of investment declined. In fact, in 2007:Q3, μ_t was near its steady state of zero. Within the first three quarters of the recession, 2007:Q4-2008:Q2, the marginal efficiency of investment declined

[30] To derive this object, I use Equation 4 and assume that the habit persistence parameter is zero, $h = 0$. This implies a simple consumption Euler equation of the form: $c_t = \beta \frac{d_t}{d_{t-1}} \frac{R_{t-1}}{\gamma e^{z_t}} c_{t-1}$.

[31] The GZ spread measure is constructed using secondary market prices of senior unsecured fixed coupon corporate bonds of U.S. nonfinancial firms. Compared to the simple measure of Baa-Aaa corporate bond spreads, it has the advantage that it adjusts for the duration mismatch between the cash flow of corporate bonds and the risk-free security.

about 6% below its steady state, to eventually fall 15% below its steady state by 2009:Q1. The spike in the observed cost of borrowing for nonfinancial firms in this period coincides with the deterioration of the marginal efficiency of investment. The tight correlation (-0.84) between the filtered marginal efficiency of investment and the credit spreads is a strong signal that the investment shock I recover from the model captures disruptions in credit markets. Following the end of the recession, credit spreads declined and the marginal efficiency of investment recovered. However, the financial frictions remained at levels that continued to depress aggregate demand. After 2009:Q1, the marginal efficiency of investment remained 6.9% below its steady state.

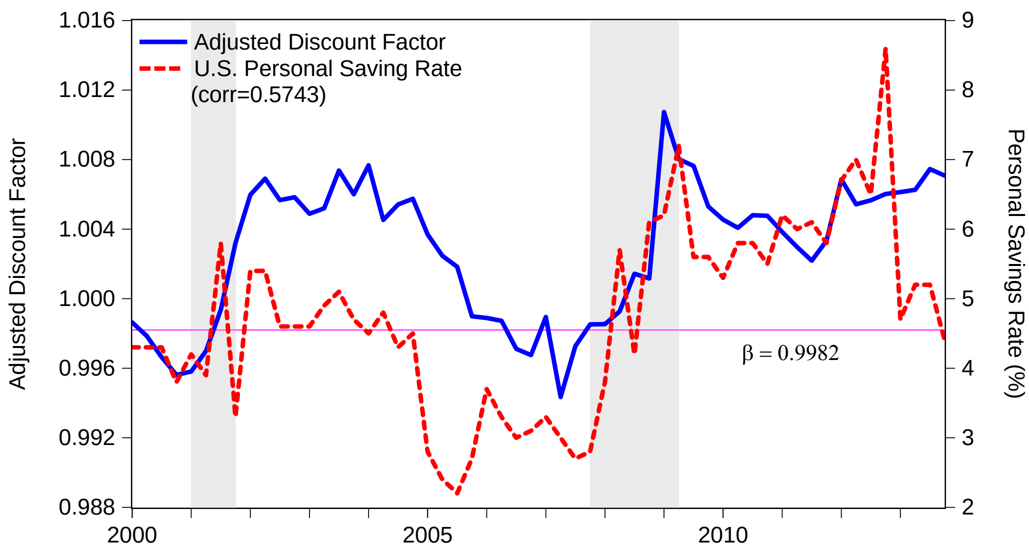
Figure 8: Marginal Efficiency of Investment and Credit Spreads



Notes: The dashed red line corresponds to the credit spread measure of Gilchrist and Zakrajšek (2012). The solid blue line is the mean filtered state, μ_t, obtained from $\mathbb{E}(s_t|Y^t)$. All series are standardized. The gray areas indicate NBER recession dates.

According to the interpretation provided in Section 3.1, another possible cause of the Great Recession that has received considerable attention is the tightening of borrowing constraints for households. For instance Mian and Sufi (2012) favor this view as the leading cause of the economic collapse. In my model it is not possible to capture household borrowing and lending because of the representative agent structure. However, the shock to preferences serves as a stand in for deeper frictions that cause agents to reduce leverage, causing a contraction in aggregate consumption. To illustrate this point, in Figure 9 I plot the adjusted discount factor against the U.S. personal savings rate. A value of the adjusted discount factor above $\beta = 0.9981$ indicates an increased desire to save by households driven by negative shocks to d_t. I interpret an increase in the adjusted discount factor above β as a tightening of borrowing constraints. The measure of the savings rate I use is the ratio of personal saving to disposable personal income. The adjusted discount factor tracks the increase in the personal savings rate which jumped from 2.3% prior to the start of the recession to 7% by the end of the episode.

Figure 9: Subjective Discount Factor and Saving Rates



Notes: The dashed red line corresponds the personal savings rate defined as the ratio of personal savings to disposable personal income obtained from FRED, Federal Reserve Bank of St. Louis. The solid blue line is the mean filtered state, $\tilde{\beta}_t$, constructed from $\mathbb{E}(s_t|Y^t)$. The gray areas indicate NBER recession dates.

A possible objection to interpreting the preference shock as a reduced form measure of tightening borrowing constraints, is that the increase in the personal savings rate could have been driven by a faster decline in disposable income that did not affect borrowing limits directly. This was unlikely to be the case. Table 4 shows the evolution of households' stock of debt using the Federal Reserve Bank of New York Consumer Credit Panel data. Total debt peaked in 2008:Q3 at 12.7 trillion dollars, and declined continuously from that point on. The first row shows the year-over-year change in the stock of debt. The second row shows the cumulative growth with respect to 2008:Q3. The reduction in debt in the data coincides with the increase in households' adjusted discount factor. This provides further support to the interpretation given to this shock.

Table 4: Evolution of Households' Stock of Debt

	2008:Q3	2008:Q4	2009:Q4	2010:Q4	2011:Q4
Annual Change (%)	4.47	2.41	-3.98	-3.72	-1.51
% Change with Respect to 2008:Q3	-	-0.04	-4.02	-7.59	-8.98

Notes: Data comes from the Federal Reserve Bank of New York Consumer Credit Panel, Brown et al. (2010).

7 Counterfactual Experiments

I now turn to analyze the effect of structural shocks on macroeconomic dynamics. I focus attention on Great Recession and its aftermath (2008:Q3 to 2013:Q4). I use the model solution and the filtered states to construct counterfactual responses of output, consumption, investment, the nominal rate and inflation, and compare them to their realized values. In what follows, I use the counterfactual exercises to answer three questions: What caused the ZLB to bind? What was the contribution of each shock to aggregate demand? and What shocks helped the economy stave off deflation?

Constructing counterfactual paths. I briefly explain the algorithm I use to construct the counterfactual paths. First I recover the mean of the filtering density $\bar{s}_t = \mathbb{E}p(s_t|Y^t)$. Feeding the path $\bar{s}_t$ using (18), I can reconstruct the observed U.S. time series. The counterfactual paths are constructed according to the following algorithm:

Algorithm 1 ***Counterfactual Paths***

1. *Construct the posterior mean* ($\bar{s}_t$) *of the state vector* $\bar{s}_t = \mathbb{E}(s_t|Y^t)_{2008:Q2}^{2013:Q4}$*, using the particle approximation of* $p(s^t|Y^t)$*.*

2. *The actual data path* $\mathbb{Y}_t$ *can be reconstructed feeding* $\{\bar{s}_t\}_{2008:Q2}^{2013:Q4}$ *to Equation 17 and Equation 18.*

3. *The counterfactual path* $\tilde{\mathbb{Y}}_t^i$ *is constructed by setting* $\{\varepsilon_t^i = 0\}_{2008:Q2}^{2013:Q4} \in \bar{s}_t$*, for a subset of shocks* $i \subset \{\mu, d, z, g, R\}$

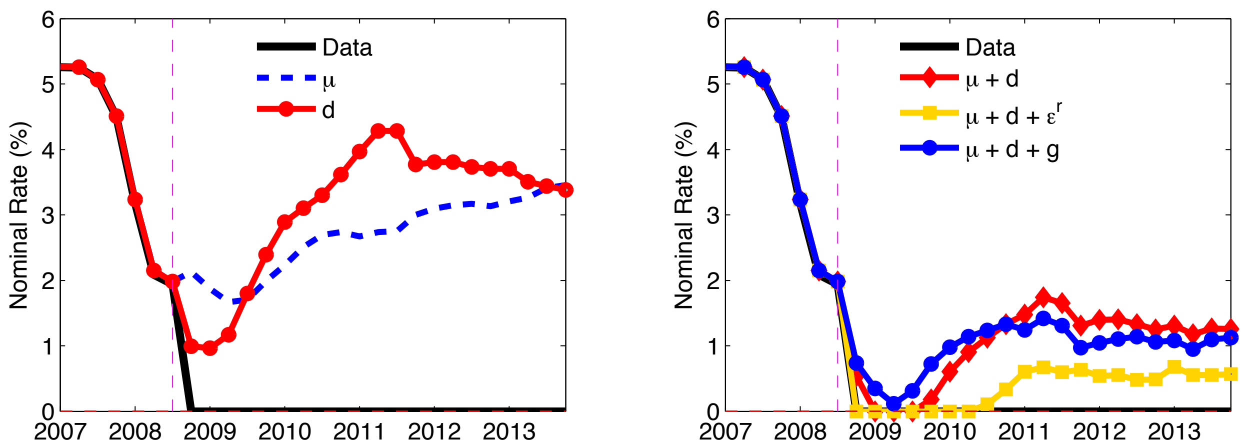
4. *The difference between* $\mathbb{Y}_t$ *and* $\tilde{\mathbb{Y}}_t^i$ *measures the contribution of the* i^{th} *structural innovation to the evolution of the observed time series.*

7.1 What drove the economy to the ZLB, and kept it there?

My results indicate that the initial force that pushed the economy to the ZLB was a combination of negative investment and preference shocks. In Figure 10 I feed only one filtered shock at a time through the model measurement equation from 2008:Q3 onward. The left panel shows that no single shock is able to push the U.S. economy into a liquidity trap. In the right panel I plot counterfactual paths implied by different combinations of shocks.

The combination of investment and preference shocks is sufficient to push the economy to the ZLB and absent any other disturbance would have kept the economy at the ZLB from 2009:Q2-2010:Q1. The figure also shows that monetary policy actions kept the nominal interest rate pegged at zero even after the initial impulse of investment and preference shocks subsided. Turning to fiscal policy, increasing government expenditure would have pushed the nominal interest away from the zero bound. When the nominal interest rate adjusts to changes in government expenditure, fiscal policy is less effective (e.g. Christiano, Eichenbaum and Rebelo (2011)). By committing to keep the nominal interest rate close to the ZLB, the Federal Reserve increased the effectiveness of fiscal stimulus during the recovery.

Figure 10: Counterfactual Path: Nominal Interest Rate



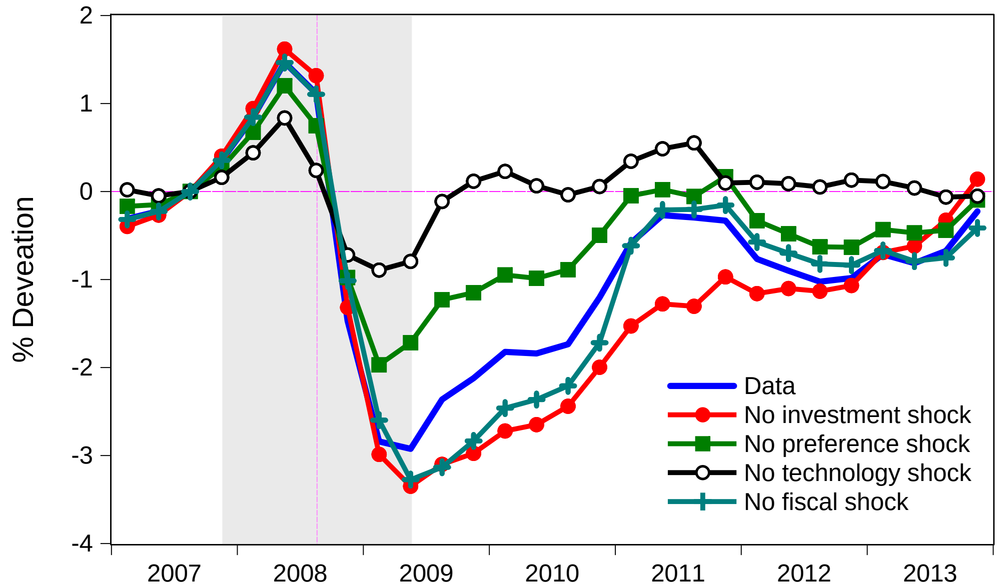
7.2 Response of Aggregate Demand

What explain the dynamics of consumption and investment during and after the Great Recession? Figure 11 shows the evolution of consumption and investment in the data and in their counterfactual paths, all expressed in percentage deviations with respect to their pre-recession values. Panel (a) shows that after the financial crisis, the driving forces in the decline in consumption were technology and preference shocks.

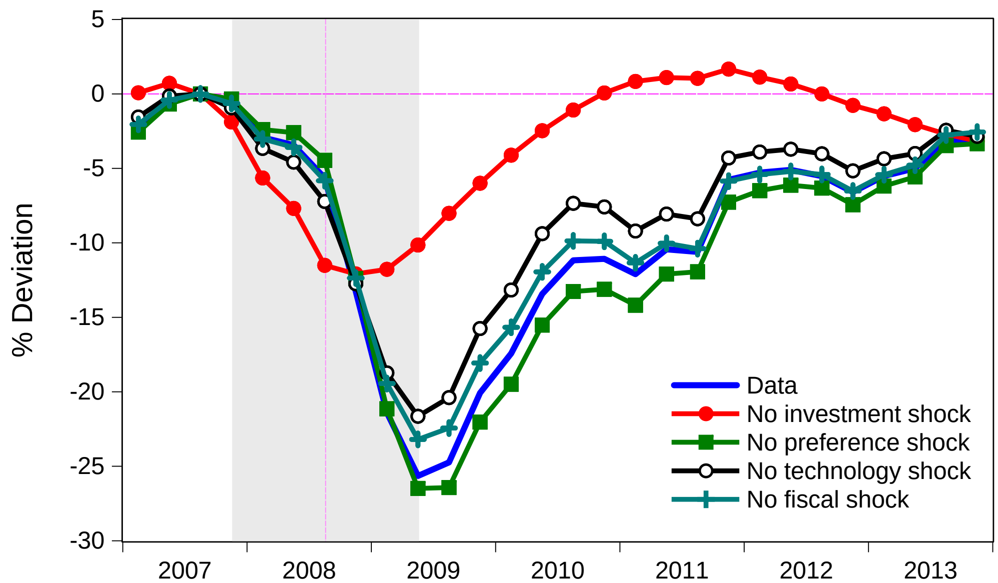
Technology shocks are important to explaining consumption because nominal rigidities are weak in my baseline estimation. Weak nominal rigidities imply that negative demand shocks cause small declines in consumption but put significant downward pressure on prices. Because deflation is not a dominant feature in the data during the recession, a combination of both technology and preference shocks is needed to explain a decline in consumption with low but positive inflation. Meanwhile, the fiscal stimulus during the period 2009-2011 had positive effects on private consumption, because at the ZLB government expenditure lowers the real interest rate and boosts consumption. In the absence of a fiscal stimulus, private consumption would have been between 0.2-0.6% lower during the recovery.

Figure 11: Counterfactual Paths: Consumption and Investment

(a) Consumption



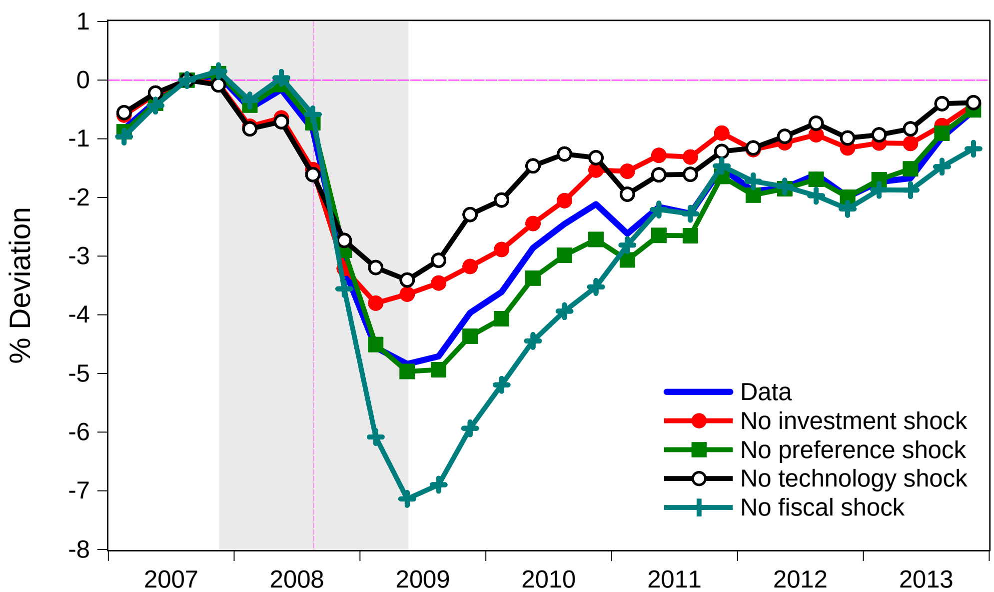
(b) Investment



Notes: All series are detrended using the HP filter ($\lambda = 1600$) *and normalized to 2007:Q3. The shaded region indicates the NBER recession. The vertical dashed line corresponds to the start of the counterfactual exercise in 2008:Q3.*

Panel (b) of Figure 11 shows a similar exercise for investment. The differences are starker. The dominant shock that explains the decline in investment is the deterioration in the marginal efficiency of investment. The protracted recovery in the aftermath of the recession is due to the persistence of financial frictions, as captured by the filtered estimates of μ_t, which remained below trend throughout. Contrary to the effects on consumption, government expenditure did have a crowding out effect on aggregate investment. This presents a trade-off for the effectiveness of fiscal policy during a liquidity trap that cannot be appreciated in New Keynesian models without investment. The economic reason for this effect is that higher government expenditure does not reduce the excess return between risky capital and risk-free bonds, which is driven by the μ_t process. By lowering the real rate increased government spending shifts resources away from already low levels of investment towards consumption.

Figure 12: Counterfactual Paths of Output



Notes: All series are detrended using the HP filter ($\lambda = 1600$) and normalized to 2007:Q3. The shaded region indicates the NBER recession. The vertical dashed line corresponds to the start of the counterfactual exercise in 2008:Q3

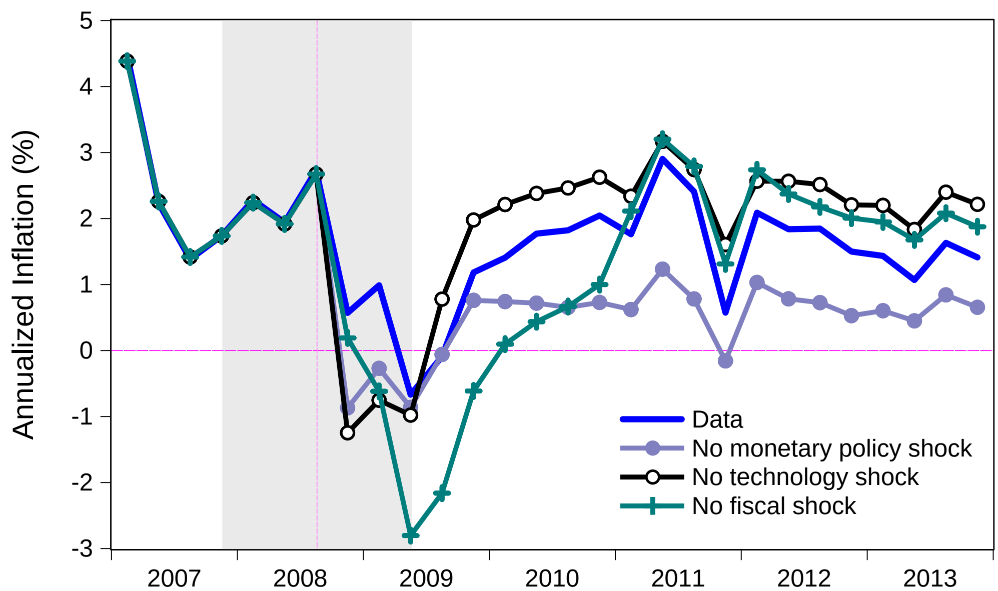
Figure 12 shows the overall effect of shocks on output. I focus on two results. First, financial frictions captured by the investment shock had a significant role in the depth of the recession, and after 2008:Q3 explain the vast majority of the slow recovery in output.

Second, expansionary fiscal policy had a positive effect on output from 2009-2011, and its reversal from 2012 onward held back a faster recovery.

7.3 Where did the deflation go?

An observation that continues to generate debate among macroeconomists is the absence of persistent deflation during and after the Great Recession. As noted by Hall (2011), any model that delivers a *Phillips* curve equation relating prices to some measure of economic activity (e.g. unemployment or marginal costs) would predict persistent deflation when there is substantial slack in the economy. However, despite the severity of the Great Recession, inflation has for the most part remained positive. I try to provide an explanation for the absence of deflation based on the extracted shocks from the model. Figure 13 shows different counterfactual paths for inflation constructed using algorithm 1. I focus only on the shocks that would have predicted deflation.

Figure 13: Counterfactual Paths of Inflation



Notes: Inflation is measured by the annualized quarterly percentage change in the GDP deflator. The shaded region indicates the NBER recession. The vertical dashed line corresponds to the start of the counterfactual exercise in 2008:Q3.

The counterfactuals show that fiscal and monetary policy helped prevent a sustained decline in prices. The negative technology shock helped reduce the deflationary pressures in the immediate aftermath of the financial crisis, but thereafter it was mostly the stimulative effect of fiscal and monetary policy on aggregate demand what kept inflation on positive terrain. Del Negro, Giannoni and Schorfheide (2014) offer an alternative explanation of the deflation puzzle based on estimates of nominal rigidities that imply a very slow frequency of price adjustment.[32] In comparison, my results complement the findings of Aruoba, Cuba-Borda and Schorfheide (2013) calling attention to the role of policies that helped avoid the perils of a switch to an equilibrium with deflationary dynamics.

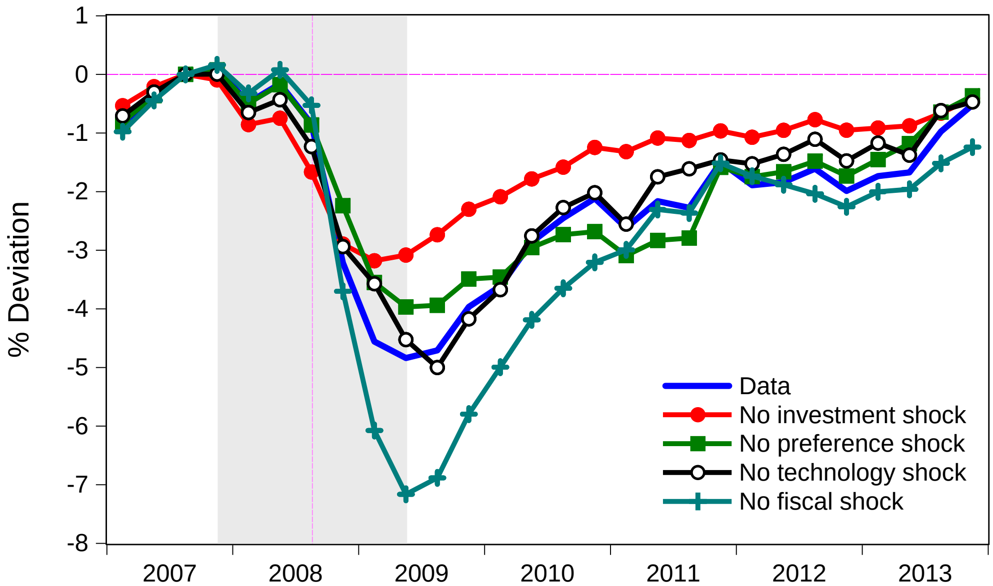
7.4 Role of nominal rigidities

My baseline estimates imply a low degree of nominal rigidities and to rationalize the data, my model relies heavily on particular realizations of shocks, in particular technology. Because nominal rigidities are important to the contribution of different structural shocks during the Great Recession and its recovery, I resolve the model increasing the value of the parameter ϕ_p that controls the cost of price adjustment. The new value of ϕ_p implies a slope of the Phillips curve of $\kappa(\phi_p) = 0.05$, which is close to the degree of nominal price rigidity estimated in Christiano, Eichenbaum and Trabandt (2014) and Justiniano, Primiceri and Tambalotti (2010).

The response of output is enough to highlight the main difference with the baseline parameterization. Figure 14 shows that the main difference is the balance between the contribution of technology and preference shocks. With prices that are costlier to adjust, preference shocks can explain the majority in the drop of consumption and reduce the importance of technology shocks overall. The contribution of investment shocks as a main driver of the recession remains, as well the stimulative effect of fiscal policy during the liquidity trap.

[32] Del Negro, Giannoni and Schorfheide (2014) augment the Smets and Wouters (2007) model with financial frictions and find that the degree of nominal rigidities is stronger when they use credit spread data to identify the structural parameters of the model.

Figure 14: Counterfactual Path: Output - Stronger Nominal Rigidities



Notes: All series are detrended using the HP filter ($\lambda = 1600$) *and normalized to 2007:Q3. The shaded region indicates the NBER recession. The vertical dashed line corresponds to the start of the counterfactual exercise in 2008:Q3*

8 Conclusions

In this paper I examine the potential causes of the U.S. Great Recession and the subsequent slow recovery through the lens of a medium scale New Keynesian model subject to the Zero Lower Bound constraint. Using Sequential Monte Carlo Methods I recover the shocks and unobservable states that caused the Great Recession. I find that the recession started with a decline in the marginal efficiency of investment, reflecting frictions in the process of financial intermediation. When the financial crisis unfolded in 2008:Q3, these frictions were exacerbated pushing the economy to the ZLB.

To explain the economy's dynamics prior and during the recession, the model relies on disturbances to the marginal efficiency of investment and households' subjective discount factor. Both shocks are necessary in order to explain why the economy reached the liquidity trap, although shocks to the marginal efficiency of investment are more important overall.

In particular, the persistent deterioration in the marginal efficiency of investment played the dominant role in explaining the slow recovery. Absent the negative shocks to investment from 2008:Q3 onward, investment would not have contracted about 45% less than with it actually did. Moreover, investment would have recovered faster, returning to its pre-recession levels by mid-2010.

Discretionary government expenditure provided substantial stimulus to reduce the severity of the recession. Government consumption helped sustain aggregate demand, and with the economy at the ZLB the fiscal stimulus helped avoid a deflationary spiral. Monetary policy also helped stimulate the economy by keeping the nominal interest rate at the ZLB even after the frictions affecting consumption and investment started to subdue. However, the impasses between the White House and Congress with regard to the long-run outlook of fiscal policy generated a slowdown and held back the economic recovery from 2011 onward.

In this paper I made progress to bridge the gap between the solution of quantitative models with a ZLB constraint and medium-scale DSGE models commonly used for the study of business cycles and policy analysis. Solving this class of models is important to understand the joint dynamics of aggregate demand, inflation and a nominal interest rate that is subject to the ZLB constraint. I leave extensions like the study of different price setting mechanisms to better understand the dynamics of inflation, and the importance of explicit financial frictions when the economy is at the ZLB for future research.

References

Andrieu, Christophe, Arnaud Doucet, and Roman Holenstein. 2010. "Particle Markov Chain Monte Carlo Methods." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3).

An, Sungbae, and Frank Schorfheide. 2007. "Bayesian Analysis of DSGE Models." *Econometric Reviews*, 26(2-4).

Aruoba, Borağan Aruobagan, Luigi Bocola, and Frank Schorfheide. 2013. "Assessing DSGE Model Nonlinearities." National Bureau of Economic Research NBER Working Papers 19693.

Aruoba, Borağan, Pablo Cuba-Borda, and Frank Schorfheide. 2013. "Macroeconomic Dynamics Near the ZLB: A Tale Two Countries." National Bureau of Economic Research Working Paper 19248.

Barro, Robert, and Robert King. 1984. "Time-separable Preferences and Intertemporal-Substitution Models of Business Cycles." *The Quarterly Journal of Economics*, 99(4).

Bauer, Andy, Haltom Nicholas, and Juan Rubio-Ramirez. 2003. "Using the Kalman filter to smooth the shocks of a dynamic stochastic general equilibrium model."

Bernanke, Ben, Mark Gertler, and Simon Gilchrist. 1999. "The financial accelerator in a quantitative business cycle framework." In *Handbook of Macroeconomics*. Vol. 1 of *Handbook of Macroeconomics*, , ed. J. B. Taylor and M. Woodford, Chapter 21. Elsevier.

Brown, Meta, Andrew Haughwout, Donghoon Lee, and Wilbert van der Klaauw. 2010. "The financial crisis at the kitchen table: trends in household debt and credit." Federal Reserve Bank of New York Staff Reports 480.

Carlstrom, Charles, and Timothy Fuerst. 1997. "Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis." *American Economic Review*, 87(5).

Chetty, Raj, Adam Guren, Day Manoli, and Andrea Weber. 2013. "Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities." *NBER Macroeconomics Annual, University of Chicago*

Press, 27(1).

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,." *Journal of Political Economy.*

Christiano, Lawrence, Martin Eichenbaum, and Mathias Trabandt. 2014. "Understanding the Great Recession." National Bureau of Economic Research Working Paper 20040.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo. 2011. "When Is the Government Spending Multiplier Large?" *Journal of Political Economy*, 119.

Creal, Drew. 2009. "A survey of sequential Monte Carlo methods for economics and finance." VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics Serie Research Memoranda 0018.

Curdia, Vasco, Marco Del Negro, and Daniel Greenwald. 2013. "Rare shocks, Great Recessions." Federal Reserve Bank of San Francisco Working Paper Series 2013-01.

del Negro, Marco, and Frank Schorfheide. 2012. *Bayesian Macroeconometrics. Handbook of Bayesian Econometrics*, Oxford University Press.

Del Negro, Marco, Marc Giannoni, and Frank Schorfheide. 2014. "Inflation in the Great Recession and New Keynesian Models." *American Econ*, Forthcoming.

Eggertsson, Gautti. 2009. "What Fiscal Policy is Effective at Zero Interest Rates?" Federal Reserve Bank of New York?

Eggertsson, Gautti, and Michael Woodford. 2003. "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity.*

Fernald, John. 2012. "A quarterly, utilization-adjusted series on total factor productivity." Federal Reserve Bank of San Francisco Working Paper Series 2012-19.

Fernández-Villaverde, Jesús, Grey Gordon, Pablo Guerr'øn-Quintana, and Juan Rubio-Ramírez. 2012. "Nonlinear Adventures at the Zero Lower Bound." University of Pennsilvania.

Gertler, Mark, and Peter Karadi. 2011. "A model of unconventional monetary policy." *Journal of Monetary Economics*, 58(1).

Gilchirst, Simon, and Egon Zakrajšek. 2012. "Rethinking the Financial Crisis." , ed. Andrew Lo Alan Blinder and Robert Solow, Chapter Credit Supply Shocks and Economic Activity in a Financial Accelerator Model. Russell Sage Foundation.

Gilchrist, Simon, and Egon Zakrajšek. 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review*, 102(4).

Gilchrist, Simon, Vladimir Yankov, and Egon Zakrajšek. 2009. "Credit Market Shocks and Economic Fluctuations: Evidence from Corporate Bond and Stock Markets." National Bureau of Economic Research, Inc NBER Working Papers 14863.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory Huffman. 1988. "Investment, capacity utilization, and the real business cycle." *The American Economic Review.*

Guerrieri, Veronica, and Guido Lorenzoni. 2011. "Credit Crises, Precautionary Savings and the Liquidity Trap."

Gust, Christopher, David Lopez-Salido, and Matthew Smith. 2012. "The Empirical Implications of the Interest-Rate Lower Bound." Federal Reserve Board.

Hall, Robert. 2011. "The Long Slump." *American Economic Review*, 101(2).

Heiss, Florian, and Viktor Winschel. 2006. "Estimation with Numerical Integration on Sparse Grids." University of Munich, Department of Economics Discussion Papers in Economics 916.

Herbst, Edward, and Frank Schorfheide. 2014. *Bayesian Inference for DSGE Models.* Manuscript, University of Pensilvania.

Judd, Kenneth. 1998. *Numerical Methods in Economics.* MIT Press.

Judd, Kenneth, Lilian Maliar, and Sergei Maliar. 2012. "Merging Simulation and Projection Aproaches to Solve High-Dimensional Problems." NBER 18501.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti. 2010. "Investment shocks and business cycles." *Journal of Monetary Economics*, 57(2).

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti. 2011. "Investment Shocks and the Relative Price of Investment." *Review of Economic Dynamics*, 14(1).

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti. 2013. "Is There a Trade-Off between Inflation and Output Stabilization?" *American Economic Journal:*

Macroeconomics, 5(2).

Krueger, Dirk, and Felix Kubler. 2004. "Computing equilibrium in OLG models with stochastic production." *Journal of Economic Dynamics and Control, Elsevier*, 28(7).

Malin, Benjamin, Dirk Krueger, and Felix Kubler. 2007. "Computing Stochastic Dynamic Economic Models with a Large Number of State Variables: A Description and Application of a Smolyak-Collocation Method." National Bureau of Economic Research NBER Technical Working Papers, National Bureau of Economic Research, Inc 0345.

Mian, Atif, and Amir Sufi. 2012. "What explains high unemployment? The aggregate demand channel." National Bureau of Economic Research, Inc NBER Working Papers 17830.

Mian, Atif, Kamalesh Rao, and Amir Sufi. 2013. "Household Balance Sheets, Consumption, and the Economic Slump." *The Quarterly Journal of Economics*, 128(4).

Richter, Alexander, Nathaniel Throckmorton, and Todd Walker. 2011. "Accuracy, Speed and Robustness of Policy Function Iteration." *Computational Economics*.

Rios-Rull, Victor, Frank Schorfheide, Cristina Fuentes-Albero, Maxym Krishko, and Raul Santaeulalia-Llopis. 2012. "Methods versus substance: Measuring the effects of technology shocks." *Journal of Monetary Economics*, 59(8).

Schorfheide, Frank. 2008. "DSGE model-based estimation of the New Keynesian Phillips curve." *Economic Quarterly*, , (Fall).

Smets, Frank, and Rafael Wouters. 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *The American Economic Review*, 97(3).

Appendix

A DSGE model

A.1 Model description

The economy is composed by households, firms and the government. Next I describe the optimization problem of each agent in the economy:

Households. The representative agent in this economy solves the following problem:

$$\max_{C_t, I_t, L_t, \bar{K}_t, B_t, u_t} I\!\!E_0 \sum_{s=0}^{\infty} \beta^s d_{t+s} \left[\ln \left(C_{t+s} - bC_{t+s-1} \right) - \psi_L \frac{L_t^{1+\nu}}{1+\nu} \right]$$

$$s.t.$$

$$P_t C_t + P_t I_t + B_t \leq W_t L_t + R_t^k u_t \bar{K}_{t-1} - P_t \mathcal{A}(u_t) \bar{K}_{t-1} + R_{t-1} B_{t-1} - P_t T_t + \Pi_t$$

$$\bar{K}_{t+1} = (1-\delta) \bar{K}_t + \mu_t \left(1 - S \left(\frac{I_t}{I_{t-1}} \right) \right) I_t$$

Let Λ_t be the multiplier associated with the nominal budget constraint and Ξ_t the multiplier associated with the law of motion of installed capital.

The optimality condition for consumption is:

$$\Lambda_t P_t = \frac{d_t}{C_t - hC_{t-1}} - h\beta \mathbb{E}_t \frac{d_{t+1}}{C_{t+1} - hC_t}$$

And the associated Euler equation is:

$$\Lambda_t = \beta R_t \mathbb{E}_t \Lambda_{t+1}$$

The investment decision is governed by:

$$\Lambda_t P_t = \mu_t \Xi_t \left[1 - S \left(\frac{I_t}{I_{t-1}} \right) - S' \left(\frac{I_t}{I_{t-1}} \right) \frac{I_t}{I_{t-1}} \right] + \beta \mathbb{E}_t \Xi_{t+1} \mu_{t+1} S' \left(\frac{I_{t+1}}{I_{t-1}} \right) \left(\frac{I_{t+1}}{I_t} \right)^2$$

For capital accumulation the optimality condition is:

$$\Xi_t = \beta \mathbb{E}_t \left\{ \Lambda_{t+1} \left[u_{t+1} R^k_{t+1} - P_{t+1} \mathcal{A}(u_{t+1}) \right] + (1-\delta) \Xi_{t+1} \right\}$$

The optimal level of capacity utilization satisfies the condition:

$$R^k_t = P_t \mathcal{A}'(u_t)$$

The first order condition for labor supply is trivial:

$$\Lambda_t W_t = \psi_L L_t^{\nu}$$

Intermediate-goods firms Firms operate a technology that combines labor and capital to produce the intermediate good. Taking the demand for their products as given, intermediate-goods firms have to choose their demand for labor and capital input and set the price at which they sell their product. The problem can be broken in these two stages.

Optimal factor demand First firms takes the price of its output as given and rent capital $K_{i,t}$ and labor $H_{i,t}$ from households to minimize costs subject to its production technology. To hire labor firms pay the real wage W_t and a rental rate of capital R^k_t that are determined at the aggregate level.

$$\begin{aligned} \min W_t H_{i,t} + R^k_t K_{i,t} \\ s.t. \quad Y_{i,t} \leq K^{\alpha}_{i,t} \left(A_t H_{i,t} \right)^{1-\alpha} - A_t F \end{aligned}$$

This problem yields the following first order conditions:

$$\begin{aligned} W_t H_t(i) = \Psi_t (1-\alpha) K^{\alpha}_t(i) \left(H_t(i) \right)^{1-\alpha} \\ R^k_t K_t(i) = \Psi_t \alpha K^{\alpha}_t(i) \left(H_t(i) \right)^{1-\alpha} \end{aligned}$$

In the first stage optimal factor demand yields the following condition:

$$\frac{R_t^k}{W_t} = \frac{\alpha}{1-\alpha}\frac{H_t(i)}{K_t(i)}$$

This implies that all firms choose the same demand for factors of production, and one can write an expression for marginal costs as:

$$\begin{aligned} MC_t &= \alpha^{-\alpha}(1-\alpha)^{\alpha-1} R_t^{k\alpha}(W_t)^{1-\alpha} \\ MC_t &= \alpha^{-\alpha}(1-\alpha)^{\alpha-1}\left(\frac{\alpha}{1-\alpha}\frac{H_t}{K_t}W_t\right)^{\alpha}(W_t)^{1-\alpha} \\ MC_t &= (1-\alpha)^{-1}\left(\frac{H_t}{K_t}\right)^{\alpha} W_t \\ MC_t &= \frac{W_t}{(1-\alpha)(K_t/H_t)^{\alpha}} \end{aligned}$$

Pricing decision . Taking the marginal cost as given, In the second stage firms set prices to maximize (nominal) profits: $D_t = [1-\Phi_p(P_{i,t}/P_{it-1})]\,P_{i,t}Y_t(i) - MC_tY_t(i)$

To the solve the following program:

$$\begin{array}{ll} \max\limits_{\{P_{i,t}\}} & \mathbb{E}_t\sum_{s=0}^{\infty}\beta^s\Lambda_{t+s}\left\{\left[1-\Phi_p\left(\frac{P_{i,t+s}}{P_{it+s-1}}\right)\right]P_{i,t+s}Y_{i,t+s} - MC_{t+s}Y_{i,t+s}\right\} \\ s.t. & Y_t(i) = \left(\frac{P_{i,t}}{P_t}\right)^{-\theta} Y_t \end{array}$$

To ease the notation I use $\theta = \frac{1}{\lambda_p}$. The first order conditions of this problem is:

$$\begin{aligned} &\Lambda_t\left\{(1-\theta)\left[1-\Phi_p\left(\frac{P_{i,t}}{P_{it-1}}\right)\right]\left(\frac{P_{i,t}}{P_t}\right)^{-\theta}Y_t - \Phi_p'\left(\frac{P_{i,t}}{P_{it-1}}\right)\left(\frac{P_{i,t}}{P_t}\right)^{-\theta}Y_t\frac{P_{i,t}}{P_{it-1}}\right. \\ &\left.+\theta MC_t\left(\frac{P_{i,t}}{P_t}\right)^{-\theta}\frac{Y_t}{P_t}\right\} + \beta\mathbb{E}_t\Lambda_{t+1}\left\{\Phi_p'\left(\frac{P_{it+1}}{P_{i,t}}\right)\left(\frac{P_{it+1}}{P_{i,t}}\right)^2 Y_{it+1}\right\} = 0 \end{aligned}$$

Symmetric price equilibrium. In a symmetric equilibrium this reduces to:

$$\Lambda_t\left\{(1-\theta)\left[1-\Phi_p(\pi_t)\right]Y_t + \theta\left(\frac{MC_t}{P_t}\right)Y_t - \Phi_p'(\pi_t)\pi_tY_t\right\} \ + \ \beta\Lambda_{t+1}\Phi_p'(\pi_{t+1})\pi_{t+1}^2Y_{t+1} = 0$$

In a symmetric equilibrium the optimal pricing decision yields the following equilibrium condition:

$$\Phi_p'(\pi_t)\pi_t + (\theta - 1)\left[1 - \Phi_p\left(\pi_t\right)\right] - \theta\left(\frac{MC_t}{P_t}\right) = \beta\mathbb{E}_t\frac{\Lambda_{t+1}}{\Lambda_t}\Phi_p'(\pi_{t+1})\pi_{t+1}^2\frac{Y_{t+1}}{Y_t}$$

Marginal costs. Before setting prices the firm decides the optimal factor demand after minimizing production costs given by the term $W_t L_{i,t} + R_t^k K_{i,t}$ and subject to the production technology described earlier. Here we assume that labor and capital are traded in an economy wide factor market which results in simple solution in which all intermediate firms choose the same capital labor ratio: $K_{i,t}/L_{i,t} = \frac{\alpha}{1-\alpha}\frac{W_t}{R_t^k}$. As a consequence all intermediate firms face identical marginal costs given by:

$$MC_t = \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}W_t^{1-\alpha}R_t^{k\alpha}A_t^{-(1-\alpha)}.$$

A.2 Stationary Equilibrium

I express the equilibrium conditions in terms of stationary transformation of the model's 14 endogenous variables: y_t, k_t, L_t, w_t, ρ_t, mc_t, λ_t, c_t, π_t, u_t, q_t, i_t, $\bar{k}_t$, R_t.

$$\text{Consumption Decision:} \quad \lambda_t = \frac{\gamma e^{d_t+z_t}}{\gamma c_t e^{z_t} - hc_{t-1}} - h\beta\mathbb{E}_t \frac{e^{d_{t+1}}}{\gamma c_{t+1} e^{z_{t+1}} - hc_t} \tag{A.1}$$

$$\text{Euler Equation:} \quad \lambda_t = \beta R_t \mathbb{E}_t \frac{\lambda_{t+1}}{\pi_{t+1}} \frac{1}{\gamma e^{z_{t+1}}} \tag{A.2}$$

$$\text{Capital Decision:} \quad q_t = \beta\mathbb{E}_t \frac{\lambda_{t+1}}{\lambda_t} \frac{1}{\gamma e^{z_{t+1}}} \left\{ u_{t+1}\rho_{t+1} - \mathcal{A}(u_{t+1}) + (1-\delta) q_{t+1} \right\} \tag{A.3}$$

$$\begin{aligned} \text{Investment Decision:} \quad 1 &= \mu_t q_t \left[1 - S(\Delta i)_t - dS(x_t)x_t\right] \\ &\quad + \beta\mathbb{E}_t q_{t+1} \frac{\lambda_{t+1}}{\lambda_t} \frac{1}{\gamma e^{z_{t+1}}} \mu_{t+1} dS(x_{t+1}) x_{t+1}^2 \end{aligned} \tag{A.4}$$

$$\text{Capital Utilization:} \quad \rho_t = d\mathcal{A}(u_t) \tag{A.5}$$

$$\text{Evolution of Installed Capital:} \quad \bar{k}_t = (1-\delta)\frac{\bar{k}_{t-1}}{\gamma e^{z_t}} + \mu_t \left[1 - S(x_t)\right] i_t \tag{A.6}$$

$$\text{Optimal Factor Demand:} \quad \frac{\rho_t}{w_t} = \frac{\alpha}{1-\alpha} \frac{L_t}{k_t} \tag{A.7}$$

$$\text{Marginal Cost:} \quad mc_t = \alpha^{-\alpha}(1-\alpha)^{\alpha-1} w_t^{1-\alpha} \rho_t^{\alpha} \tag{A.8}$$

$$\begin{aligned} \text{Price Setting:} \quad \Gamma_t &= \beta\mathbb{E}_t \frac{\lambda_{t+1}}{\lambda_t} \Phi_p(\pi_{t+1}) \pi_{t+1} \frac{y_{t+1}}{y_t} \\ \Gamma_t &= \left\{ (\frac{1}{\lambda_p} - 1) \left[1 - \Phi_p(\pi_t)\right] - \frac{mc_t}{\lambda_p} + d\Phi_p(\pi_t)\pi_t \right\} \end{aligned} \tag{A.9}$$

$$\text{Labor Leisure:} \quad \tilde{w}_t = e^{d_t} \frac{\psi_L}{\lambda_t} L_t^{\nu} \tag{A.10}$$

$$\text{Resource Constraint:} \quad c_t = \left[\frac{1}{ge^{g_t}} - d\Phi_p(\pi_t)\right] y_t - \mathcal{A}(u_t)\frac{\bar{k}_{t-1}}{\gamma e^{z_t}} - i_t \tag{A.11}$$

$$\text{Effective Capital:} \quad k_t = u_t \frac{\bar{k}_{t-1}}{\gamma e^{z_t}} \tag{A.12}$$

$$\text{Gross Output:} \quad y_t = k_t^{\alpha} L_t^{1-\alpha} - \mathcal{F} \tag{A.13}$$

A.3 Steady state

To determine the steady state first note that:

$$\rho = \frac{\gamma}{\beta} - 1 + \delta$$

From the pricing equation we have:

$$mc = \frac{(\theta - 1)}{\theta} = \frac{1}{\lambda_p}$$

From this point on it will be easier to characterize the steady state as a function of the capital-labor ratio:

$$\begin{aligned} w &= mc(1-\alpha)(k/L)^{\alpha} \\ \rho &= \alpha mc(k/L)^{\alpha-1} \end{aligned}$$

Which implies that,

$$(k/L) = \left(\frac{\rho}{\alpha mc}\right)^{\frac{1}{\alpha-1}}$$

And then we ca replace in the expression for w to obtain,

$$w = mc(1-\alpha)\left(\frac{\rho}{\alpha mc}\right)^{\frac{\alpha}{\alpha-1}} = mc^{1/1-\alpha}(1-\alpha)\left(\frac{\rho}{\alpha}\right)^{\frac{\alpha}{\alpha-1}}$$

From the production function we can obtain,

$$\frac{y}{L} = (k/L)^{\alpha} - \frac{\mathcal{F}}{L}$$

And since $\mathcal{F}$ is set such that profits are zero in steady state and using the result that $\Phi_p(\pi^*) = 0$, we have:

$$\frac{\mathcal{F}}{L} = (k/L)^{\alpha} - \rho(k/L) - w$$

Which implies that,

$$\frac{y}{L} = \rho(k/L) - w$$

Now using the law of motion of effective capital and the definition of effective capital we have:

$$i/k = \gamma - (1 - \delta)$$

And the investment output ratio is:

$$\frac{i}{y} = \frac{i}{k}\frac{k}{L}\frac{L}{y}$$

Finally from the resource constraint:

$$\frac{c}{y} = \frac{1}{\bar{g}} - \frac{i}{y}$$

Now we need to solve for the steady state value of hours worked L in order to recover the rest of the objects of the steady state. From the consumption-leisure optimality condition we obtain:

$$w = \frac{\psi_L L^{\nu}}{\lambda}$$

And from the definition of the marginal utility of wealth we obtain:

$$\lambda = \frac{1}{c}\frac{\gamma - h\beta}{\gamma - b}$$

Hence we can write:

$$\lambda L = \left(\frac{c}{L}\right)^{-1}\frac{\gamma - h\beta}{\gamma - b}$$

And note that $\frac{c}{L} = \frac{1}{g}\frac{y}{L} - \frac{i}{L}\frac{L}{y}\frac{y}{L}$. Combining with the previous equation,

$$w = \frac{\psi_L L^{1+\nu}}{\lambda L}$$

Hence we can recover the steady state level of hours as:

$$L = \left(\frac{w\lambda L}{\psi_L}\right)^{\frac{1}{1+\nu}}$$

With the steady state level of hours we can directly recover the steady state of $k, c, i, y, \mathcal{F}$.

B Computing the Nonlinear Solution

I explain how to solve the nonlinear decision rules. I use the notation $\mathbb{S} = \left[R_{-1}, c_{-1}, \bar{k}_{-1}, i_{-1}, \mu, d, z, g, \varepsilon^r\right]$ to summarize the state variables, and approximate the decision rules $\mathcal{C} = \{L(\mathbb{S}), q(\mathbb{S}), \lambda(\mathbb{S}), i(\mathbb{S}), \pi(\mathbb{S})\}$

B.1 Residuals

To find the policy functions that solves the above system of equilibrium conditions I minimize the sum of squared residuals with respect to the unknown coefficients Θ. To that end, I first define the residual functions that will serve as metric for the solution procedure described later.

$$\mathcal{R}_1(\mathbb{S}) = \lambda(\mathbb{S}) - \beta R_t \mathbb{E}\frac{\lambda(\mathbb{S}')}{\pi(\mathbb{S}')}\frac{1}{\gamma e^{z'}} \tag{B.1}$$

$$\mathcal{R}_2(\mathbb{S}) = \lambda(\mathbb{S}) - \frac{\gamma e^{d+z}}{\gamma c e^{z} - hc_{-1}} - h\beta\mathbb{E}\frac{e^{d'}}{\gamma c' e^{z'} - hc} \tag{B.2}$$

$$\mathcal{R}_3(\mathbb{S}) = q(\mathbb{S}) - \beta\mathbb{E}\frac{\lambda(\mathbb{S}')}{\gamma e^{z'}\lambda(\mathbb{S})}\left\{\rho(\mathbb{S}')u' - \mathcal{A}(u') + (1-\delta)q(\mathbb{S}')\right\} \tag{B.3}$$

$$\begin{aligned}\mathcal{R}_4(\mathbb{S}) &= 1 - e^{\mu}q(\mathbb{S})\left[1 - S(\Delta i) - dS(\Delta i)x\right] \\ &- \beta\mathbb{E}q(\mathbb{S}')\frac{\lambda(\mathbb{S}')}{\lambda(\mathbb{S})}\frac{1}{\gamma e^{z'}}e^{\mu'}dS(\Delta i')x'^2\end{aligned} \tag{B.4}$$

$$\begin{aligned}\mathcal{R}_5(\mathbb{S}) &= \left(\frac{1}{\lambda_p} - 1\right)\left[1 - \Phi_p(\pi(\mathbb{S}))\right] - \frac{mc}{\lambda_p} + d\Phi_p(\pi(\mathbb{S}))\pi(\mathbb{S}) \\ &- \beta\mathbb{E}\frac{\lambda(\mathbb{S}')}{\lambda(\mathbb{S})}\Phi_p(\pi(\mathbb{S}'))\pi(\mathbb{S}')\frac{y'}{y}\end{aligned} \tag{B.5}$$

B.2 Expectations

To evaluate the expectations that form part of the residual equations (B.1) - (B.5), I use deterministic integration methods based on a Gauss-Hermite quadrature rule. The exogenous components of the state vector, $\mathbb{S}'$, is constructed using a non-product using the sparse grid algorithm of Heiss and Winschel (2006). For example, suppose that we want to compute $\mathbb{E}[f(\mathbf{x})]$ for $\mathbf{x} \in \mathbb{R}^D$ where, $\mathbf{x}$ is a vector of random variables distributed according to $N(0, I_D)$.

Define first the Q^{th} order discrete approximation to any univariate function $\mathbb{E}g(x)$ to be:

$$V_Q \equiv \mathbb{E}[g(x)] \approx \sum_{i=1}^{Q} g(x_i) w_i$$

Where x_i and w_i are the Gauss-Hermite nodes and weights as in Judd (1998). Usually to evaluate $\mathbb{E}f(\mathbf{x})$ one would construct a tensor product approximation using the x_i, and w_i for each element in $\mathbf{x}$. However this approach becomes computationally costly as the dimensionality of the function of interest and the accuracy of the approximation Q increases. To simplify, take the case of a bivariate function $D = 2$, and define the set of indexes $\mathcal{I}_Q^{D=2} = \left\{\mathbf{i} \in \mathbb{N}^D : \sum_{d=1}^{D} i_d = D + Q\right\}$, where $\mathbb{N}$ is the set of all positive integers. The *level-k* sparse grid approximation with $D = 2$ dimensions is given by:

$$\mathbb{E}[f(\mathbf{x})] \approx \sum_{q=k-D}^{k-1} (-1)^{k-1-q} \binom{D-1}{k-1-q} \sum_{\mathbf{i}\in\mathcal{I}_Q}^{D} V_{i1} \otimes \cdots \otimes V_{iD}$$

B.3 Computational Algorithm

Algorithm 2 *The solution algorithm proceeds as follows.*

1. *Without loss of generality begin in step j with a guess for the unknown coefficients $\Theta^{(j)}$.*
2. *Construct the approximated decision rules:*

$$\mathcal{C}^{(j)} = \{L(\mathbb{S}; \Theta^{(j)}), q(\mathbb{S}; \Theta^{(j)}), \lambda(\mathbb{S}; \Theta^{(j)}), i(\mathbb{S}; \Theta^{(j)}), \pi(\mathbb{S}; \Theta^{(j)})\}$$

3. *Construct the following objects using the decision rules that correspond to a non-binding ZLB (Θ^{nb}): $w, \rho, u, mc, c, \bar{k}, y$, using equations (9) - (15).*
4. *Compute the notional interest rate using (16). If the the notional rate violates the ZLB go back to step 3, set $R = 1$ and recompute all objects using Θ^b.*
5. *Repeat steps 1 - 4, to construct and evaluate the objects inside the expectations in (B.1) - (B.5).*
6. *Updated the vector of unknown coefficients to $\Theta^{(j+1)}$ using any minimization routine until a solution for $\min_\Theta \sum_{i=1}^{M} \mathcal{R}(\mathbb{S}_i, \Theta)^2$ is found.*

B.4 Accuracy of the Nonlinear Solution

The accuracy of the numerical solution is evaluated with respect to the residual functions defined in Appendix B.1. If one could obtain the actual policy functions instead of the approximated ones, then Equation B.1-Equation B.5 would be satisfied exactly. A measure of the "exactness" of the approximated policy rules can be measured by how much does the approximated decision rules fail to satisfy the residual equations exactly. To take an specific example, take Equation B.1:

$$\mathcal{E}\mathcal{E}_1 = \log 10 \left| 1 - \frac{\beta R_t \mathbb{E} \frac{\lambda(\mathbb{S}')}{\pi(\mathbb{S}')} \frac{1}{\gamma e^{z'}}}{\lambda(\mathbb{S})} \right|$$

The above expression measured in terms of consumption units is expressed in $\log 10$ for ease of interpretation. Figure A-1 shows the distribution of the Euler Errors for all the residual functions.

Figure A-1: Distribution of Euler Equation Errors in $\log 10$ units

B.5 Posterior Predictive Checks

An brief introduction to the use of predictive checks can be found in del Negro and Schorfheide (2012). The predictive checks rests upon the construction of the predictive distribution of a sequence of simulated data $p(\tilde{\mathbb{Y}}^T|\Omega_T)$ where Ω_T is the information set available up to period T and includes the realization of the observed data $\mathbb{Y}^T$ and the draws from the posterior distribution of the parameters of the model $p(\theta|\Omega_T)$. 3 describes how to obtain draws from the predictive distribution. Once these draws are obtained they can be transformed into empirical moments of interest $\mathcal{S}(\tilde{\mathbb{Y}}^T) \in \mathbb{R}^n$, for example sample means, covariances, autocovariances, correlations, etc. I compute the posterior predictive checks for the five observable series used for estimation. The predictive distribution is constructed using $T = 100$ periods initialized at the deterministic steady state of the model and sampling $N = 1,000$ realizations of the posterior distribution of estimated parameters.

Algorithm 3 *Drawing from* $p(\tilde{\mathbb{Y}}^T|\Omega_T)$

1. *Fix a draw of the parameter vector from the posterior distribution,* θ_j*,* $j = 1, \ldots, N$*.*
2. *Use the model solution to simulate a sequence of observables* $\tilde{\mathbb{Y}}_t(\theta_j), \quad t = 1, \ldots, T$*.*
3. *Construct the vector of moments of interest* $\mathcal{S} : \tilde{\mathbb{Y}}^T(\theta_j) \to \mathcal{S}(\tilde{\mathbb{Y}}^T) \in \mathbb{R}^{n\times 1}$*.*
4. *The posterior predictive distribution* $p(\mathcal{S}(\tilde{\mathbb{Y}}^T)|\Omega_T)$ *can be characterized with the empirical distribution of* $\mathcal{S}(\tilde{\mathbb{Y}}^T)$*.*
5. *Compare the distribution of* $\mathcal{S}(\tilde{\mathbb{Y}}^T)$ *with the corresponding statistic based on actual data* $\mathcal{S}(\mathbb{Y}^T)$*.*

C Data for Estimation and Filtering

All information comes from the Federal Reserve Bank of St. Louis FRED data service (mnemonics in parenthesis), for the period 1984:Q1 to 2013:Q4

1. **Price Level.** Is the implicit GDP price deflator (GDPDEF) index.
2. **Population.** Is the civilian non-institutionalized population over sixteen years (CNP16OV).
3. **Per Capita Output Growth** Real per-capita output growth is constructed using data on Gross Domestic Product (GDP) expressed in billions of dollars:

$$\Delta y_t = \left[\ln\left(\frac{GDP_t}{CNP16OV_t \times GDPDEF_t}\right) - \ln\left(\frac{GDP_{t-1}}{CNP16OV_{t-1} \times GDPDEF_{t-1}}\right)\right] \times 100$$

4. **Per Capita Consumption Growth.** Real per-capita consumption growth is constructed first summing Personal Consumption Services (PCESV) + Personal Consumption Durables Divided (PCND). Let $CONS_t = PCESV_t + PCND_t$, expressed in billions of dollars.

$$\Delta c_t = \left[\ln\left(\frac{CONS_t}{CNP16OV_t \times GDPDEF_t}\right) - \ln\left(\frac{CONS_{t-1}}{CNP16OV_{t-1} \times GDPDEF_{t-1}}\right)\right] \times 100$$

5. **Per Capita Investment Growth.** Real per-capita investment growth is constructed first summing Personal Consumption Durable Goods(PCEDG) + Fixed Private Investment(FPI) + Change in Private Inventories(CBI). Let $INV_t = PCEDG_t + FPI_t + CBI_t$, expressed in billions of dollars.

$$\Delta i_t = \left[\ln\left(\frac{INV_t}{CNP16OV_t \times GDPDEF_t}\right) - \ln\left(\frac{INV_{t-1}}{CNP16OV_{t-1} \times GDPDEF_{t-1}}\right)\right] \times 100$$

6. **Inflation.** The inflation rate is measured as the quarterly change of GDPDEF:

$$\pi_t = [\ln GDPDEF_t - \ln GDPDEF_{t-1}] \times 100$$

7. **Interest Rate.** The nominal interest rate is measured using the quarterly rate of the federal funds rate (FEDFUNDS):

$$R_t = FEDFUNDS_t/4$$

D Filtering

Here I describe the algorithm used to approximate the filtering density $p(s_t|\mathcal{Y}^t)$ used to recover the unobserved states of the economy. The exposition follows Creal (2009) and Andrieu, Doucet and Holenstein (2010). An in depth treatment with applications to the New Keynesian models can be found in Herbst and Schorfheide (2014).

To simplify the discussion I consider a single variable dynamic model, whose dynamics can be represented in the form of a nonlinear state-space system:

$$s_t = g\left(s_{t-1}, u_t\right), \quad u_t \sim N(0, \sigma_u) \tag{D.1}$$

$$y_t = m(s_t) + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon) \tag{D.2}$$

This systems gives rise to the measurement density $p(y_t|s_t)$ and the transition density $p(x_t|x_{t-1})$, which inherit their Markov structure from the transition equation of the model. Given a sequence of observations $y^t = \{y_1, \ldots, y_t\}$ we are interested in recovering the sequence of states that generated them $x^t = \{x_1, \ldots, x_t\}$. But because the system is stochastic we can only characterize the joint distribution where the states comes from, $p(x^t|y^t)$, known as the joint filtering distribution.

Filtering Distribution Decomposition. A key step to apply particle filter methods, is to decompose the the joint filtering distribution:

$$\begin{aligned}
p(s^t|y^t) &= \frac{p(y_t|s^t, y^{t-1})p(s^t|y^{t-1})}{p(y_t|y^{t-1})} \\
&= \frac{p(y_t|s^t, y^{t-1})}{p(y_t|y^{t-1})} p(s_t|s^{t-1}, y^{t-1})p(s^{t-1}|y^{t-1}) \\
&= \frac{p(y_t|s_t)p(s_t|s_{t-1})}{p(y_t|y^{t-1})} p(s^{t-1}|y^{t-1}) \propto p(y_t|s_t)p(s_t|s_{t-1})p(s^{t-1}|y^{t-1})
\end{aligned}$$

The last steps follows from the Markov property of the transition and measurement densities. This makes clear that it is possible to characterize the filtering density sequentially starting from some initial joint density of states $p(s^{t-1}|y^{t-1})$. The scaling factor in the last equation is the marginal contribution of observation y_t to the likelihood.

Marginal Distribution Decomposition. An alternative to trying to uncover the joint distributions of states, the filtering problem can be cast in terms of the marginal distribution $p(s_t|y^t)$. This is the key approach taken in the implementation of Sequential Monte Carlo methods. One can thing of this decomposition as a form of sequential learning, yielding greater flexibility to the filtering approach.

Suppose we have access to an initial distribution of states $p(s_0)$. Given the evolution of the system to a known position, y^{t-1}, s_{t-1}, the marginal predictive density $p(s_t|y^{t-1})$ can be obtained integrating out the transition equation of the model according to:

$$p(s_t|y^{t-1}) = \int p(s_t|s_{t-1})p(s_{t-1}|y^{t-1})ds_{t-1}$$

Following the same steps as before, the marginal filtering density is given by:

$$p(s_t|y^t) = \frac{p(y_t|s_t)p(s_t|y^{t-1})}{p(y_t|y^{t-1})} = \frac{p(y_t|s_t)}{p(y_t|y^{t-1})} \int p(s_t|s_{t-1})p(s_{t-1}|y^{t-1})ds_{t-1} \tag{D.3}$$

Given the additive nature of the measurement errors in the observation equation the *observation density:* $p(y_t|s_t)$ can be readily evaluated. However, there is no closed form expression for the objects that form the marginal predictive density $p(s_t|y^{t-1})$ or the marginal likelihood $p(y_t|y^{t-1})$

D.1 Sequential Monte Carlo Approximation

The key challenge in uncovering the marginal distribution of unobserved states is solving for the integrals that shown in the previous section. I will use Monte Carlo methods to approximate these objects. The key idea, is to start with a probability mass function represented by a collection of particles $\{\pi_{t-1}^{(j)}\}_{j=1}^{N}$ with associated weights $\{W_{t-1}^{(j)}\}_{j=1}^{N}$ to approximate the filtering density $p(s_{t-1}|y^{t-1})$ and systematically use the model transition and measurement equations to update this approximation to obtain $p(s_t|y^t)$.

Importance Sampler. If we could simply draw the particles and weights $\{\pi_t^{(j)}, W_t^j\}_{j=1}^N$ from the target distribution, $p(s_t|y^t)$ then for any function $h(.)$ it is possible to construct a

Monte Carlo estimator:

$$\frac{1}{M}\sum_{j=1}^{M} h(s_t^j)W_t^j \xrightarrow{a.s.} \mathbb{E}\left[h(s_t)|y^t\right] = \int h(s_t)p(s_t|y^t)ds_t \tag{D.4}$$

Because this is not possible in practice because the target distribution is unknown, the approximation has to be constructed in a different way. The solution is to use a known distribution, known as importance density $g_t(s_t|s_{t-1}, y^t)$, such that:

$$\mathbb{E}\left[h(s_t)|y^t\right] = \int h(s_t)\frac{p(s_t|y^t)}{g_t(s_t|s_{t-1}, y^t)}g_t(s_t|s_{t-1}, y^t)ds_t$$

Note that now we can approximate this integral drawing from the known importance density $g_t(s_t|s_{t-1}, y^t)$. The importance density is indexed at time t, meaning that it can be adjusted as new information is incorporated into the approximation. Also note that since we draw from the importance density instead of drawing from the target density, the draws are reweighted using the importance weights, $w_t^j = \frac{p(s_t|y^t)}{g_t(s_t|s_{t-1},y^t)}$.

Sequential Importance Sampler. Having defined the idea of an importance sampler, now I address how to approximate $p(s_t|y^t)$ sequentially. Suppose that a swarm of particles $\{\pi_{t-1}^j, W_{t-1}^j\}_{j=1}^N$ approximates $p(s_{t-1}|y^{t-1})$ according to Equation D.4. Using Equation D.3 we can write:

$$\mathbb{E}[h(s_t)|y^t] = \frac{\int h(s_t)p(y_t|s_t)p(s_t|s_{t-1})p(s_{t-1}|y^{t-1})ds_{t-1}}{p(y_t|y^{t-1})} \tag{D.5}$$

Concentrate the numerator, note that the particle approximation to $p(s_{t-1}|y^{t-1})$ is known. Conditional on this approximation, we draw (s_t^j) from an importance density $g_t(s_t|s_{t-1}, y^t)$ to generate the following approximation:

$$\int h(s_t)p(y_t|s_t)p(s_t|s_{t-1})p(s_{t-1}|y^{t-1})ds_{t-1} \approx \frac{1}{N}\sum_{j=1}^{N} h(s_t^j)w_t^jW_{t-1}^j \tag{D.6}$$

Where the *incremental weights*, $w_t = \frac{p(y_t|s_t)p(s_t|s_{t-1})}{g_t(s_t|s_{t-1},y^t)}$, are computed for each of the draws from the importance density. A by-product of this approximation is a expression for the likelihood:

$$p(y_t|y^{t-1}) \approx \frac{1}{N}\sum_{j=1}^{N} w_t^j W_{t-1}^j \tag{D.7}$$

Because these weights are not drawn from the target density, we need to re-scale them in order to update the approximation that we are really interested on. Let the unscaled weights be given by, $\tilde{W}_t^i = w_t^i W_{t-1}^i$. The normalized weights are $W_t^i = \frac{\tilde{W}_t^i}{\sum_{j=1}^{N}\tilde{W}_t^j}$, and then the approximation of the filtering density $p(s_t|y^t)$ is given by:

$$\mathbb{E}\left[h(s_t)|y^t\right] = \frac{\frac{1}{N}\sum_{j=1}^{N} h(s_t^j)\tilde{W}_t^j}{\sum_{j=1}^{N}\tilde{W}_t^j} = \frac{1}{N}\sum_{j=1}^{N} h(s_t^j)\left(\frac{\tilde{W}_t^j}{\sum_{j=1}^{N}\tilde{W}_t^j}\right) \tag{D.8}$$

The approximation $\{\pi_t^i, W_t^i\}_{i=1}^N$ suffers from particle degeneracy, in the sense that some of the draws from the importance density have a negligible weight. To mitigate this problem, the particles are resampled at the end of each step, keeping only those that have positive weights.