Aggregate Precautionary Savings Motives *

Pierre Mabille†

July 1, 2019

Abstract

This paper studies households' precautionary savings when they face *macro*economic shocks, a channel that complements the traditional *micro*economic precautionary savings motive. I incorporate continuous aggregate income and credit supply shocks, two prominent sources of risk, into a Bewley-Huggett-Aiyagari model calibrated to the U.S. economy. I then propose a novel solution method that quantifies how much the economy departs from certainty equivalence. The precautionary motive associated with movements in credit supply is substantial. Its negative effect on the equilibrium risk-free rate is one fourth as large as for idiosyncratic income changes, and much larger than for aggregate income changes. In the long-run, large movements in credit generate a low risk-free rate, low debt environment like the post-Great Recession period. They persistently, albeit mildly, depress consumption and employment, leading to higher estimates of the costs of business cycles. Over time, the model assigns about half of the volatility of consumption and the risk-free rate to credit supply shocks. When inverted to recover the sequence of structural shocks around the Great Recession, it suggests that households' borrowing constraints have remained tight during the recovery, despite rising aggregate consumption.

JEL classification: E30, E44, D52, C63
Keywords: Credit crisis, aggregate risk, volatility, precautionary savings, risk-free rate

*I am grateful to Marco Del Negro, Mark Gertler, Julian Kozlowski, Fernando Leibovici, Virgiliu Midrigan, Juan Pablo Nicolini, Thomas Philippon, Stijn Van Nieuwerburgh, Olivier Wang, and Stanley Zin for helpful comments and discussions, as well as participants at the American Finance Association, the Becker-Friedman Institute MFM summer meeting, the SED Meeting in St. Louis, and multiple seminars at NYU. I thank the Alfred P. Sloan Foundation and the Macro Financial Modeling project of the Becker-Friedman Institute for supporting this research.

†Department of Economics, Stern School of Business, New York University, 44 W. 4th Street, New York, NY 10012, USA; pmabille@stern.nyu.edu.

1 Introduction

Changes in household debt over the business cycle are large, correlated with real variables (consumption, employment) and financial variables (risk-free rates, credit spreads), and they are pervasive. Not only were they quantitatively important during and after the Great Recession, with the protracted decrease in U.S. household debt to GDP from 99% in 2008 to less than 80% in 2018. They are also associated with large changes in GDP and unemployment in longer time series samples that exclude the Great Recession (see for instance Mian, Sufi and Verner (2017) for evidence dating back to 1960). Furthermore, the incidence of economy-wide credit contractions is unevenly distributed across households, with those with less savings, a lower and/or riskier income profile being more affected (Mian, Rao and Sufi (2013)). This makes changes in credit supply at the macroeconomic level a potentially important source of risk, distinct from individual factors affecting credit risk[1], and against which households may wish to insure, increasing the value of safe assets. How strong is such a *macro*economic precautionary savings motive? How does it compare to the traditional *micro*economic motive associated – primarily – with idiosyncratic income shocks? How did it contribute to the decline in risk-free rates around the Great Recession?

In this paper, I build a quantitative macroeconomic model, which I solve with a novel nonlinear method that uniquely allows to address these questions. I extend a workhorse incomplete markets, heterogeneous agents model à la Guerrieri and Lorenzoni (2017), with continuous *aggregate* income and credit supply shocks. Unlike in models with transitional dynamics and perfect foresight experiments, households know the stochastic processes for the shocks, which are estimated by indirect inference using the model and historical data on output and Treasury rates. Accounting for households' expectations about aggregate risk allows to study macroeconomic precautionary savings motives. Solving the model nonlinearly has two benefits[2]. First, the economy departs from certainty equivalence with respect to aggregate shocks. Households' responses are affected not only by the level, but also by higher-order moments of income and credit shocks. Second, it accounts for the unequal incidence of credit shocks across households and over time. To my knowledge, this is the first paper that finds an economically significant difference between the linear and nonlinear solutions of a model with heterogeneous agents and aggregate risk, both in the steady state and in the response to shocks (see the discussion

[1] An example could be bank lenders incurring balance sheet losses unrelated to their portfolio of borrowers, and cutting credit to the latter, even though their characteristics are unchanged. See Huber (2018) for a recent example.

[2] Most recent linear solution methods are based on Reiter (2009).

below for more detail). This has implications for the levels and dynamics of the variables most strongly associated with macroeconomic precautionary savings. In particular, the equilibrium real risk-free rate and household debt are lower than when credit supply shocks are fully unanticipated by households. Figure 1 illustrates those differences.

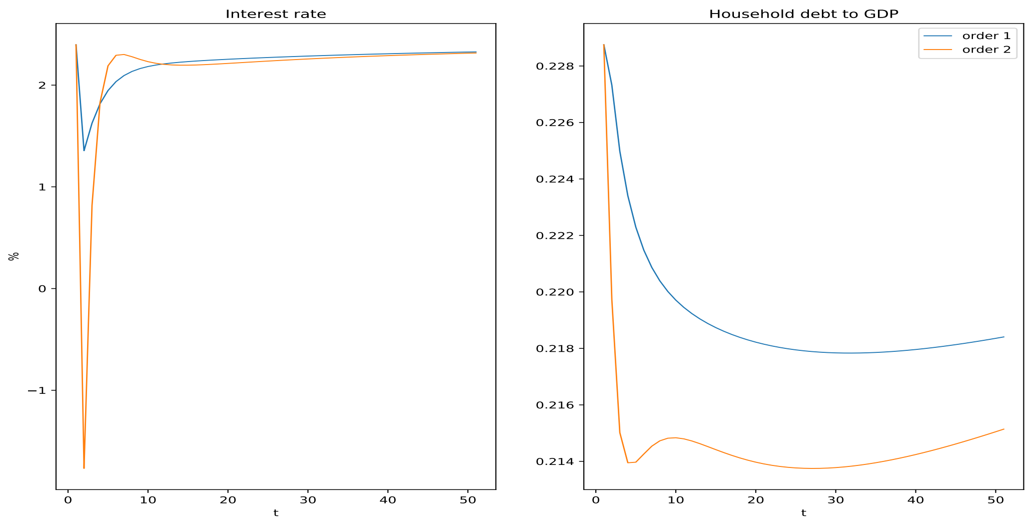

Figure 1: Impulse response functions to a one standard deviation negative credit shock. Initial period: deterministic steady state. Left panel: real risk-free rate (annualized). Right panel: debt to GDP ratio (annualized). Order 1 (blue line): first-order approximate solution. Order 2 (orange line): second-order approximate solution. A period corresponds to a quarter.

I calibrate the model to match level and cross-sectional facts on unsecured household credit in the U.S., to deliver an accurate picture of how households use assets and debt to smooth aggregate shocks. I assume that borrowing constraints are hit by a continuous aggregate credit shock, which is mean-reverting. To capture changes in households' insurance opportunities when their ability to borrow deteriorates, I model a realistic distribution of taxes and transfers across income levels. With flexible prices, changes in the risk-free rate reflect households' precautionary savings motives. In the long-run, stochastic steady state, movements in credit supply generate a low debt, low interest rate environment similar to the post-Great Recession period. When the volatility of credit shocks is calibrated based on historical data going back to 1973, the debt to GDP ratio and the risk-free rate are about 3% lower than if such shocks were fully unanticipated by households. When it is calibrated based on the Great Recession period alone (2005-2015), they are up to 50% lower. Simultaneously, the wealth distribution contracts around its mode, an equilibrium response to higher precautionary savings motives. The latter consist of two components: a microeconomic motive due to idiosyncratic earnings risk, and a macroeconomic motive due to technology (TFP) and credit supply shocks. First, absent aggregate shocks, idiosyncratic income risk increases the demand for safe assets, thus decreasing the rate of return that households require to hold them in equilibrium. This is the traditional precautionary motive of Bewley-Huggett-Aiyagari economies[3]. In line with received wisdom (Krusell and Smith (1998)), this motive turns out to have the largest contribution to the decline in the risk-free rate, compared to an economy without shocks. Second, fluctuations in TFP also affect labor earnings, but their long-run effect on consumption and the risk-free rate turns out to be close to zero. This is consistent with the calculations of Lucas (2003), who finds very low welfare costs of business cycle fluctuations due to aggregate productivity shocks. Third, credit supply shocks generate a sizable precautionary motive, which represents about one fourth of the micro motive, and thus implies higher estimates of the costs of business cycles when those are associated with credit cycles.

Credit shocks affect households' ability to borrow through two channels. In level, they induce borrowing-constrained households to deleverage, and those close to the constraint to increase precautionary savings. Their future volatility, on the other hand, makes constraints more likely to bind going forward. In turn, this results in lower output and welfare. With elastic labor supply, low risk-free rates create an intertemporal substitution effect that increases households' consumption of leisure. In particular, it decreases the labor supply of unconstrained households, which are more productive on average. In addition, a wealth effect (active with most utility functions[4]) induces those households to further decrease their labor supply, as they consume more goods. Indeed, a difference from models with transitional dynamics experiments like Guerrieri and Lorenzoni (2017) is that credit contractions lead to persistently lower employment. Aggregate employment reflects the increase in hours worked by constrained households to pay off their debt, and the decrease in hours worked by unconstrained households. With perfect foresight, the increase dominates because households close to the constraint perfectly anticipate that they will be constrained in the future, and choose to strongly increase their hours. When households have (rational but imperfect) expectations about future credit shocks, the decrease in hours dominates because households expect credit shocks to mean-revert.

To further investigate how credit shocks contribute to macroeconomic volatility, I im-

[3]This motive arises because idiosyncratic income shocks induce some households to become borrowing-constrained, thus unable to smooth consumption, and because of the "prudence" property of the utility function (Kimball (1990)).

[4]As is well-known, preferences of the Greenwood, Hercowitz and Huffman (1988) (GHH) type do not feature this effect.

plement two business cycle exercises in the model. First, a variance decomposition shows that credit shocks contribute to about half of the volatility of macroeconomic aggregates over time. In particular, they explain most of the variation in the real risk-free rate – in the absence of additional factors like external demand, shifts in investors' preferences or demographic changes. Second, I "invert" the model by applying a particle filter to recover the sequence of structural credit and technology shocks that generated the observed paths for the risk-free rate and aggregate consumption around the Great Recession, from 2006 to 2017. The fall in the risk-free rate in the sample from 2.5% to -1.5% (for the annual 5-Year Treasury Inflation-Indexed Security with constant maturity yield) is explained by a strong tightening of borrowing constraints. Their effect is exacerbated by the simultaneous collapse of TFP, which households would like to, but cannot smooth by issuing more debt. Combined, both shocks generate a temporary decline in consumption during the years of the Great Recession, which vanishes when TFP reverts to its pre-crisis level. In contrast, the implied measure of borrowing constraint tightness stays persistently low throughout the sample, limiting household indebtedness.

This paper contributes to the methodological literature by proposing a *nonlinear* solution method to solve heterogeneous agents models with aggregate risk. It improves upon existing linear methods in terms of accuracy, and it incorporates the analysis of macroeconomic risk. By allowing for multiple continuous aggregate shocks, it complements methods based on Krusell and Smith (1998) algorithms, which are often limited to study fewer shocks taking two or three values – interpreted as regimes – because of the curse of dimensionality. The method consists in solving for first- and second-order approximations of the dynamics of the economy around its steady state, described by the set of equilibrium conditions. The constant second-order terms measure the departure from certainty equivalence, and the remaining terms reflect the nonlinearity of the economy. The novelty of my approach is to make a second-order perturbation computationally tractable by combining the projection and perturbation approach of Reiter (2009) with the `gensys2` algorithm of Kim, Kim, Schaumburg and Sims (2008), to reduce the dimensionality of the system of equations for second-order coefficients – and using automatic differentiation to generate that system.

Related literature

This paper relates to the macro-finance literature on credit crises, the literature on precautionary savings, and the methodological literature on solving heterogeneous agents model with aggregate risk using perturbation-based methods.

First, It differs from Midrigan and Philippon (2016) and Justiniano, Primiceri and Tambalotti (2015) in that it focuses on the business cycle effects of credit shocks rather than the transitory effects of large deleveraging episodes. Furthermore, it does so in a model with full cross-sectional heterogeneity. In that, the benchmark model is similar to the economy in Guerrieri and Lorenzoni (2017), with the addition of continuous technology and credit shocks. The focus is different, as I analyze the economy's response to credit shocks when households have (rational but imperfect) expectations about movements in credit supply, as opposed to a tightening of borrowing constraints, which has zero probability before happening, but then whose path is perfectly forecasted. A similar comment applies to Huo and Rios-Rull (2016). Incorporating macroeconomic risk into a similar model, Favilukis, Ludvigson and Nieuwerburgh (2017) and Kaplan, Mitman and Violante (2017) study a two-point shock to aggregate credit conditions. My setting differs from theirs in that I model credit *fluctuations* as continuous, mean-reverting shocks rather than large *regime* changes. While less quantitatively less rich than theirs, my setting allows to tractably analyze macroeconomic precautionary motives. Compared to Boz and Mendoza (2014), who study an economy where households learn about the persistence of credit shocks in a representative agent setting, I solve a general equilibrium model where the risk-free rate is endogenous and reflects households' demand for insurance.

Second, this paper complements the literature on precautionary savings, which has traditionally focused on idiosyncratic income risk, as in Carroll and Samwick (1998) and Parker and Preston (2005). I set up a model to quantify that motive for TFP and credit supply shocks. Consistent with the literature on higher-order perturbations of RBC models, I replicate the result that TFP fluctuations are a negligible source of precautionary savings (Fernandez-Villaverde, Rubio-Ramirez and Schorheide (2016)). In contrast, I highlight the role of credit supply shocks, which is consistent with empirical evidence on the co-movement of credit and business cycles (Mian et al. (2017)), and speaks to the contribution of precautionary savings motive to lowering risk-free rates (Pflueger, Siriwardane and Sunderam (2017)). Theoretically, credit shocks can be viewed as negative shocks to the degree of market completeness (Dávila and Philippon (2017)), which generate a strong deleveraging response from households.

Third, the novelty of the numerical solution method is to combine the independent algorithms in Reiter (2009) and Kim et al. (2008), in order to make second-order perturbations computationally tractable in this class of models. Mathematically, this is achieved by reducing the dimension of the system of equations for second-order coefficients of policy and aggregate functions. It is different from the approach of Schmitt-Grohe and

Uribe (2008) for representative agent models, which would require to solve an intractably large system of equations. It is also different from Winberry (2016) and Ahn, Kaplan, Moll, Winberry and Wolf (2017), which focus on first-order approximate solutions. To my knowledge, the only other paper that implements a second-order order perturbation of such a model is Winberry (2016). Importantly and unlike this paper, I find an economically significant difference with the first-order solution. As I have emphasized, the difference is entirely due entirely to credit supply shocks, because they *directly* affect the cross-distribution of assets. Absent such shocks, linear methods work identically well.

Outline

The rest of this paper is organized as follows. The next paragraph reviews the literature. Section 2 presents the model. Section 3 describes the solution and its interpretation. Section 4 presents the calibration. Section 5 analyzes the long-run and the dynamic effects of credit shocks. Section 6 analyzes the contribution of credit shocks to macroeconomic volatility, with a focus on the Great Recession. Section 7 concludes.

2 Model

The economy is a general equilibrium, Bewley-Huggett-Aiyagari model, in which the assumption of fixed borrowing constraints is relaxed to allow for stochastic borrowing constraints. I assume that they are hit by continuous aggregate credit shocks over time, which affect households' ability to borrow. Continuous fluctuations in firms' total factor productivity (TFP) are another source of aggregate risk, which translates into macroeconomic income risk, on top of idiosyncratic income risk. To match the empirical correlation between output and household debt over the business cycle, I assume that credit and TFP shocks are positively correlated. All the results hold when credit and TFP shocks are independent.

2.1 Household Problem

There is a continuum of measure 1 of heterogeneous, risk-averse households. Households face idiosyncratic labor income risk. They consume a single final good produced by competitive firms, and elastically supply efficiency units of labor to those firms. Profit shares of those firms are distributed equally across households and are non-tradable. To smooth consumption over time and share risk, households borrow and save using risk-free, one-period government bonds, which determine their net asset positions. When borrowing, they face stochastic borrowing constraints. Those are the product of a common component depending on aggregate credit, and of an idiosyncratic component depending on households' current income, a proxy for creditworthiness. Finally, households face distortionary, progressive taxes on labor income, and receive lump-sum transfers from the government conditional on their resources.

Formally, the household's problem is described as follows:

$$\max_{\{c_{it},n_{it},b_{it+1}\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{c_{it}^{1-\gamma}}{1-\gamma} - \psi \frac{n_{it}^{1+\eta}}{1+\eta} \right] \tag{1}$$

$$\text{subject to} \quad c_{it} + \frac{b_{it+1}}{1+r_t} + \tau_t\left(\theta_{it}, n_{it}\right) \leq w_t \theta_{it} n_{it} + b_{it} + T\left(\theta_{it}\right) + \pi_t \tag{2}$$

$$\tau_t\left(\theta_{it}, n_{it}\right) = \tau_{0t} + \tau_1\left(\theta_{it}\right) w_t \theta_{it} n_{it} \tag{3}$$

$$\log \theta_{it} = \rho_\theta \log \theta_{it-1} + \sigma_\theta\left(z_t\right) \epsilon_{it}^{\theta}, \;\; \epsilon^\theta \sim \mathcal{N}(0,1) \tag{4}$$

$$\log z_t = \rho_z \log z_{t-1} + \epsilon_t^z \tag{5}$$

$$b_{it+1} \geq -\overline{\phi}_t \phi\left(\theta_{it}\right) \tag{6}$$

$$\log \overline{\phi}_t - \log \overline{\phi} = \rho_\phi \left(\log \overline{\phi}_{t-1} - \log \overline{\phi}\right) + \epsilon_t^\phi \tag{7}$$

$$\begin{pmatrix} \epsilon^\phi \\ \epsilon^z \end{pmatrix} \overset{iid}{\sim} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \sigma_\phi^2 & \sigma_\phi \sigma_z \rho_{\phi z} \\ \sigma_\phi \sigma_z \rho_{\phi z} & \sigma_z^2 \end{pmatrix}\right) \tag{8}$$

Household idiosyncratic productivity θ follows an AR(1) process in log, chosen to replicate the evolution of the wages of employed individuals in the U.S. Its volatility is a decreasing function of TFP z_t, chosen to match the estimates for the countercyclicality of individual income in Storesletten, Telmer and Yaron (2004). In the numerical solution, income is discretized as a finite discrete Markov chain $\Theta\left(z_t\right) = \left\{\underline{\theta}\left(z_t\right), ..., \overline{\theta}\left(z_t\right)\right\}$ with transition matrix $\Pi_\theta\left(z_t\right)$, using the Rouwenhorst method[5]. The innovations to TFP and credit shocks follow a bivariate normal distribution with mean zero.

The function $\phi\left(.\right)$ maps individual productivity to credit limits. It is calibrated to match the distribution of unsecured household debt. The aggregate credit shock is a mean-reverting process with mean $\overline{\phi}$: $\epsilon^\phi < 0$ induces a tightening of all borrowing constraints, $\epsilon^\phi \geq 0$ a relaxation. For the problem to be well-defined, the support of ϵ must be bounded. If not, the household's choice becomes empty for large values of ϵ, which can be

[5]The Rouwenhorst delivers a more accurate income dynamics than the Tauchen method when the process is persistent. Countercyclical income volatility induces the dispersion of grid points to change over time.

interpreted as the household defaulting on his debt. Even this event it is unlikely, it would be inconsistent with one-period bonds being risk-free[6]. Therefore during the simulation I always verify that households' choice sets are nonempty. Thus households' ability to borrow depends on individual-specific variables (e.g. households' balance sheets, credit scores) and economy-wide variables (e.g. nationwide lending standards).

The solution of the model accommodates alternative specifications for stochastic borrowing constraints, for instance with additive shocks, or individual sensitivities to shocks. My specification nests the model in Ludvigson (1999), which studies borrowing constraints that vary stochastically with individual income, and in Fulford (2015), which studies access to credit at the extensive margin, both under a fixed risk-free rate. A limitation of the setting here is the absence of a financial accelerator due to the absence of an endogenous tradable, durable good in the collateral constraint. An integrated treatment of households' portfolio choices between a risk-free bond and a durable good with stochastic borrowing constraints is left for further research.

2.2 Firm Problem

A continuum of competitive firms hires efficient units of labor from households every period, and combines them using a decreasing returns to scale production technology subject to a TFP shock. Firms are owned by households equally. They solve the following static problem:

$$\max_{N_t} \pi_t = z_t N_t^{\alpha} - w_t N_t \tag{9}$$

$$\tag{10}$$

There is no capital. In equilibrium, firms' profits and the wage bill are constant fractions of output, thus the firm sector transmits TFP shocks to households' wages and shares:

$$\begin{aligned} \pi_t^* &= (1-\alpha)Y_t^* = (1-\alpha)z_t \; N_t^{*\,\alpha} && (11)\\ w_t N_t^* &= \alpha Y_t^* = \alpha z_t \; N_t^{*\,\alpha} && (12) \end{aligned}$$

2.3 Government

The government issues one-period, risk-free bonds in exogenous, positive supply. To finance transfers to households and outstanding debt, it raises a progressive tax on house-

[6] Huo and Rios-Rull (2016) allow for the possibility of unexpected default in an extension, and show that it amplifies the impact of shocks.

holds' labor income and issues new bonds. The progressive income tax is modeled as an affine function of labor income, whose slope depends on households' productivity[7]. The intercept of the tax function adjusts such that the government's budget constraint holds every period.

$$\int \tau_t\left(\theta, n\left(\theta, b\right)\right) d\lambda_t\left(\theta, b\right) + \frac{B_{t+1}}{1+r_t} = \int T\left(\theta\right) d\lambda_t\left(\theta, b\right) + B_t \tag{13}$$

$$\Rightarrow \qquad \tau_{0t} = \int T\left(\theta\right) d\lambda_t\left(\theta, b\right) + B\frac{r_t}{1+r_t} - w_t \int \tau_1\left(\theta\right) \theta n_t\left(\theta, b\right) d\lambda_t\left(\theta, b\right) \tag{14}$$

Government taxes are an equilibrium object that must be solved for as a fixed point consistent with the household problem. Given prices (risk free rate, wage), they depend on households' policy functions (labor supply) and the cross-sectional distribution. Policy functions and the distribution, in turn, depend on taxes. The quantity of government bonds is assumed to be fixed, and is chosen to match unsecured debt and liquid assets to GDP ratios in the deterministic steady state of the model. The analysis of a setting where the supply of bonds changes over time, consistently with the evolution of taxes and the government budget constraint, is left for future work.

2.4 Equilibrium

With aggregate credit and TFP shocks, the household's problem is similar to Krusell and Smith (1998). The state space includes the cross-sectional distribution of households across productivity and bonds, $\{\lambda_t(\theta, b)\}$. Households know the current price of bonds $q_t = 1/(1+r_t)$, but must forecast the price of bonds next period to allocate their consumption between t and $t+1$. Since government bonds are in fixed supply, it amounts to forecasting next period demand for bonds, which depends on the distribution of NAP, which is time-varying because of aggregate shocks affecting wages and collateral values. They must also forecast the current wage, determined by the intersection of firms' labor demand and households' labor supply, itself a function of the wage. Finally, they forecast taxes, which depend on the wage and labor supply. The rational expectation equilibrium of the model is a fixed point from households' expectations to realized values. Given their anticipation of next period distribution and the current wage, households make savings

[7] Heathcote, Storesletten and Violante (2017) model a more realistic progressive tax schedule, but have transfers vary to satisfy the government budget constraint. Since the distribution of transfers across households is key to understand how much extra risk sharing households get from credit, I choose to take the distribution of transfers in the U.S. data as given, and assume a tax schedule that doesn't depend on endogenous labor supply.

and labor choices in the current period that generate that distribution next period and that wage in the current period.

Definition A competitive equilibrium is a sequence of time-varying households' policy functions

$$\left\{c_t(\theta,b), b_t'(\theta,b), n_t(\theta,b)\right\}$$

firms' policy functions N_t, prices r_t, w_t, government taxes τ_t and aggregate shocks $\overline{\phi}_t, z_t$, such that:

(i) Given prices, taxes and aggregate shocks, households' optimality conditions hold:

$$c_t(\theta,b)^{-\gamma} = \beta(1+r_t)\mathbb{E}_t\left[c_{t+1}(\theta,b)^{-\gamma}\right] + \mu_t(\theta,b) \tag{15}$$

$$(1-\tau_1(\theta))\,w_t\theta c_t(\theta,b)^{-\gamma} = \psi n_t(\theta,b)^{\eta} \tag{16}$$

where $\mu_t(\theta,b)$ denotes the multiplier on agent (θ,b)'s borrowing constraint.

(ii) Given prices and aggregate shocks, firms' optimality conditions hold:

$$\alpha z_t N_t^{\alpha-1} = w_t \tag{17}$$

(iii) Given prices and aggregate shocks, the government's budget constraint holds:

$$\tau_{0t} = \int T(\theta)\,d\lambda_t(\theta,b) + B\frac{r_t}{1+r_t} - w_t\int \tau_1(\theta)\,\theta n_t(\theta,b)\,d\lambda_t(\theta,b) \tag{18}$$

(iv) The goods, bonds and labor markets clear:

$$\int c_t(\theta,b)\,d\lambda_t(\theta,b) = z_t N_t^{\alpha} \tag{19}$$

$$\int b_t'(\theta,b)\,d\lambda_t(\theta,b) = B \tag{20}$$

$$\int \theta n_t(\theta,b)\,d\lambda_t(\theta,b) = N_t \tag{21}$$

$$\tag{22}$$

(v) The distribution of households λ_t and aggregate shocks evolve according to their laws of motion. Formally, denote $\Theta\times\mathcal{B}$ the sigma-algebra associated with the Cartesian product of the discrete set of productivity levels and the compact set of bond holdings, and $(\tilde{\Theta},\tilde{\mathcal{B}})$ a subset of the sigma-algebra (in fact, those depend on

t because of countercyclical income volatility, i.e. the set of productivity levels depends on z_t). Then,

$$\lambda_{t+1}\left(\tilde{\Theta},\tilde{\mathcal{B}}\right)=\int_{\Theta\times\mathcal{B}}Q_{\overline{\phi}_t,z_t}\left(\left(\theta,b\right),\left(\tilde{\Theta},\tilde{\mathcal{B}}\right)\right)d\lambda_t\left(\theta,b\right) \tag{23}$$

$$\text{where}\quad Q_{\overline{\phi}_t,z_t}\left(\left(\theta,b\right),\left(\tilde{\Theta},\tilde{\mathcal{B}}\right)\right)=\mathbf{1}\left\{b_t'(\theta,b)\in\tilde{\mathcal{B}}\right\}\sum_{\theta'\in\tilde{\Theta}}\Pi_\theta\left(\theta'|\theta\right) \tag{24}$$

The transition function $Q_{\overline{\phi},z}$ depends on individual productivity, bonds, TFP and aggregate credit conditions. The latter two shift the distribution of bonds over time[8].

The market clearing risk-free rate and the wage are endogenous.The former is solved for numerically, but the wage can be solved for analytically using the firms' optimality condition and labor market clearing:

$$w_t=\alpha z_t N_t^{\alpha-1}=\alpha z_t\left(\int\theta n\left(\theta,b\right)d\lambda_t\left(\theta,b\right)\right)^{\alpha-1} \tag{25}$$

It is directly affected by the TFP shock z_t, and indirectly by the credit shock $\overline{\phi_t}$ through its effect on the distribution λ_t.

2.5 Discussion

When modeling households' expectations, this paper relaxes the assumption of zero-probability credit shocks made in most of the literature on credit crises. This is arguably more relevant for the Great Recession and the subsequent decade (see Veldkamp et al). In particular, it differs from models with transitional dynamics experiments, in which tightening and relaxation of households' borrowing constraints are zero probability events, whose path is perfectly forecasted by households once they hit the economy. Most papers that study the effect of a shock to households' ability to borrow make this assumption for analytic or computational tractability. The solution that I propose extends Guerrieri and Lorenzoni (2017), and keeps track of the entire wealth distribution as a state variable while analyzing the effect of aggregate credit changes anticipated by agents. Households have rational expectations throughout the transition, they know the stochastic process governing the credit shock and its current realization, but not its future realizations. The solution of the model relies on a second-order perturbation of the equilibrium around its

[8]Other papers study how aggregate shocks shift the households distribution over time, for instance Krueger, Mitman and Perri (2016). To my knowledge, this paper is the first to study how the distribution evolves in response to an aggregate credit shock.

deterministic steady state with fixed aggregate credit and TFP. Every variable is decomposed as the sum of a linear and a quadratic term in the state variables' deviations from steady state, and a constant term linear in the variance of the aggregate shocks. Since only the linear term appears in the first-order, we can measure the nonlinearities associated with the various aggregate shocks by computing the difference between first- and second-order terms[9]. On top of that, the constant terms permanently shift the model long-run (stochastic) steady state, and capture precautionary savings motive associated with the various aggregate shocks. Households insure against aggregate credit and TFP shocks based on their knowledge of their processes. That channel is active in models with two aggregate credit shocks like Favilukis et al. (2017) and Kaplan et al. (2017). The limitation of this solution is to be based on perturbation methods, which only makes it accurate in a neighborhood of the steady state. However, as I show below, the approximation remains accurate for large deviations, such as the deleveraging episode of the Great recession.

3 Solution

The model is solved numerically. This paper proposes a solution method for heterogeneous agents models with aggregate shocks that extends the projection and perturbation approach of Reiter (2009), to make a second-order approximation of the equilibrium tractable in models of small to medium dimensions. I describe the algorithm in detail in the computational appendix subsection A.1. In this section, I outline the main steps and the economic interpretation.

3.1 Algorithm

There are three main steps. The first two steps, with variants, are common to perturbation-based solutions of heterogeneous agents models. First, approximate the infinite-dimensional equilibrium variables by finite-dimensional objects: policy functions with linear splines, pricing functions, cross-sectional distribution with histogram weights. I obtain a finite parametrization that I refer to as the discrete model. Second, compute the stationary steady state of the discrete model. The steady state solution is global, and fully nonlin-

[9] This measure is not exact because it is based on a second-order perturbation of the model, abstracting from higher-order terms. Solving for higher-order terms is computationally very challenging, and most solutions of heterogeneous agents models with aggregate shocks restrict themselves to *first*-order perturbations (e.g. Reiter (2009), McKay and Reis (2016), Ahn et al. (2017)). Moreover, it may be reasonable to assume that households react to the average and the volatility of aggregate shocks but not to higher-order moments, as in approaches based on bounded rationality (Favilukis et al. (2017) make a similar remark).

ear with respect to idiosyncratic income risk, but abstracts from aggregate shocks. Third, perturb the discrete model around its steady state and compute the rational expectations solution of the perturbed system. The Jacobian and Hessian of the equilibrium system of equations, used as input to the perturbation method, are computed exactly with automatic differentiation (Revels, Lubin and Papamarkou (2016)).

Then, I compare the results of a first- to a second-order perturbation. In the first-order solution, equilibrium variables depend linearly on the lagged values of the states, and the economy features certainty equivalence. That is, the volatility of aggregate shocks is irrelevant for households' decisions and aggregates, it only scales the impact of current shocks. In the second-order solution, variables depend nonlinearly on the lagged values, and they are affected by the volatility of aggregate shocks.

First step Equilibrium conditions are stacked in a multivariate, vector-valued function $\mathcal{F}(.)$ that represents the nonlinear system of equations defining the equilibrium:

$$\mathbb{E}_t \left[\mathcal{F} \left(\mathbf{y_t}, \mathbf{y_{t+1}}, \mathbf{x_t}, \mathbf{x_{t+1}}, \epsilon^{\phi}_{t+1}, \epsilon^{z}_{t+1} \right) \right] = 0 \tag{26}$$

The vector of non-predetermined variables $\mathbf{y}$ contains the policy function coefficients (for labor and next period bonds; consumption is backed out from the budget constraint), the bond price, the wage, aggregate consumption and employment. The vector of predetermined variables $\mathbf{x}$ contains the histogram weights, aggregate TFP and credit shocks.

Second step Solving for the deterministic steady state boils down to solving a large nonlinear system of equations,

$$\mathcal{F}(\mathbf{y}, \mathbf{y}, \mathbf{x}, \mathbf{x}, 0, 0) = 0. \tag{27}$$

In theory, it could be solved directly using a nonlinear equation solver. In practice, there is no guarantee that numerical equation solvers will converge when we use projection methods to approximate policy functions. In addition to solving the households' consumption problem, the difficulty comes from having endogenous labor supply on top of it, endogenous government taxes, and solving for two prices in equilibrium (the wage and the interest rate). I also solve for the value of the disutility of labor ψ that normalizes steady state output Y to 1.

Therefore, to make the problem more stable, I use the following variant of policy time iteration. First, given a guess for $\mathbf{x}$ and $\mathbf{y}$[10], compute government taxes for all agents. Given taxes and the guess, solve for households' labor supply policy. Given that policy, solve then for households' savings policy. Using the policy functions, compute the implied stationary distribution (the eigenvector method is efficient), and the new taxes. The process is repeated until policy functions converge. I use Broyden's method every time a numerical solver is needed, and automatic differentiation to compute exact derivatives. Since the convergence of the numerical solver is not guaranteed under any initial guess and parameter combination, I calibrate the steady state of the model with a homotopy method. That is, I slowly change parameters until the target is reached, starting from a combination under which the model steady state is easily computed. If needed, I modify the state space boundaries over that process.

Third step Then, denote η the perturbation parameter scaling the amount of aggregate uncertainty in the economy. The solution to the equilibrium expectational difference equation $\mathbb{E}_t\left[\mathcal{F}\left(.\right)\right]=0$ is of the form (Schmitt-Grohe and Uribe (2008)):

$$\mathbf{x_{t+1}} = \mathbf{h}\left(\mathbf{x_t},\eta\right)+\eta\begin{pmatrix}\mathbf{0}\\ \epsilon^{\phi}_{t+1}\\ \epsilon^{z}_{t+1}\end{pmatrix} \tag{28}$$

$$\mathbf{y_t} = \mathbf{g}\left(\mathbf{x_t},\eta\right) \tag{29}$$

The first-order approximation of the solution, in steady state deviations, is

$$\widehat{\mathbf{x_{t+1}}} = \mathbf{h_x}\left(\mathbf{x},0\right)\widehat{\mathbf{x_t}}+\eta\begin{pmatrix}\mathbf{0}\\ \epsilon^{\phi}_{t+1}\\ \epsilon^{z}_{t+1}\end{pmatrix} \tag{30}$$

$$\widehat{\mathbf{y_t}} = \mathbf{g_x}\left(\mathbf{x},0\right)\widehat{\mathbf{x_t}} \tag{31}$$

There are various ways to solve for the coefficients $\mathbf{h_x}\left(\mathbf{x},0\right)$ and $\mathbf{g_x}\left(\mathbf{x},0\right)$. I use the `gensys` algorithm of Sims (2001) (I verify that the results coincide with those obtained with another method, Klein (2000)). This step involves computing the Jacobian of $\mathcal{F}\left(.\right)$.

[10] A good guess is usually obtained by using the endogenous grid method of Carroll (2006) to iterate backwards on the household's optimality conditions, starting from any feasible guess. This preliminary step is used in McKay and Reis (2016), and is especially needed with endogenous labor supply.

Then, the second-order approximation of the solution writes:

$$\widehat{\mathbf{x_{t+1}}} = \mathbf{h_x}(\mathbf{x},0)\,\widehat{\mathbf{x_t}} + \frac{1}{2}\mathbf{h_{xx}}(\mathbf{x},0)\,\widehat{\mathbf{x_t}}^2 + \frac{1}{2}\mathbf{h}_{\eta\eta}(\mathbf{x},0)\,\eta^2 + \eta\begin{pmatrix}\mathbf{0}\\ \epsilon^{\phi}_{t+1}\\ \epsilon^{z}_{t+1}\end{pmatrix} \quad (32)$$

$$\widehat{\mathbf{y_t}} = \mathbf{g_x}(\mathbf{x},0)\,\widehat{\mathbf{x_t}} + \underbrace{\frac{1}{2}\mathbf{g_{xx}}(\mathbf{x},0)\,\widehat{\mathbf{x_t}}^2}_{\text{nonlinear term}} + \underbrace{\frac{1}{2}\mathbf{g}_{\eta\eta}(\mathbf{x},0)\,\eta^2}_{\text{non-certainty equivalence term}} \quad (33)$$

This step uses the Jacobian and the Hessian of $\mathcal{F}(.)$, which are usually sparse. It involves solving for second-order coefficients in a large linear system of equations involving the Jacobian and the Hessian of the system, and the first-order coefficients obtained earlier. Usual second-order solution methods for representative agents models (Schmitt-Grohe and Uribe (2008)) cannot be applied here, because they require constructing and inverting a matrix too large to be stored in most computers. To circumvent the problem, I use the `gensys2` algorithm of Kim et al. (2008)[11]. It applies a series of linear operations to the original system (Schur and singular value decomposition) to reduce its dimensionality and solve for the coefficients.

Computation parameters and time are in appendix. I use a log-spaced grid for bonds to increase the precision of policy functions at values closer to borrowing constraints. I use the smallest possible bond grid for the histogram, which leaves the first-order approximation of the model unchanged compared to finer grids, to reduce the dimension of the system as much as possible. The first-order solution takes longer than first-order approximations in continuous-time (e.g. Ahn et al. (2017)), but is much faster than discrete-time algorithms based on Krusell and Smith (1998). There is no benchmark for the second-order approximation, since this paper is one of the first to implement it (Winberry (2016) uses a smooth density approximation, but does not report computing time).

3.2 Stochastic steady state and impulse response functions

In the first-order approximation, deviations from the steady state are zero in the absence of aggregate shocks, so the steady state in the first-order approximation coincides with the deterministic steady state. In the second-order approximation, even in the absence of aggregate shocks, deviations from the deterministic steady state are nonzero and depend on the volatility of aggregate shocks, generating a stochastic steady state. The latter can be interpreted as the average level of the economy when it is hit by a typical sequence of

[11] I thank Jinill Kim and Sunghyun Kim for sharing their code.

aggregate shocks in the long-run. In the second-order (and higher) approximation, the effects of aggregate shocks on equilibrium states and policy functions do not cancel out in the long-run, because of the deviation from certainty equivalence. To compute the deviations of the stochastic steady state from the deterministic one, I compute a fixed point of the pruned laws of motion of the economy[12]. I use pruned laws of motion to alleviate the well-known problem that iterating on second-order laws of motion gives rise to higher-order terms that do not in general increase the accuracy of the approximation and are likely to lead to explosive paths (Kim et al. (2008)). Pruning essentially computes first-order projections of second-order terms, based on a first-order expansion of the conditional expectation of the system's deviation from steady state.

The impulse response functions (IRF) to credit and TFP shocks are computed by feeding the pruned laws of motion with nonzero innovations in the first period and iterating on them. Over the simulated paths, I measure the accuracy of the first- and second-order approximations by computing the residuals of equilibrium conditions, in particular market clearing conditions for bonds, consumption and labor (Table 1). They are very small in the first-order approximation of the model, and further decrease towards zero in the second-order approximation, showing the good fit of the model.

Market:	Bonds	Good	Labor
order 1	0.01% (0.03%)	0.04% (0.04%)	0.01% (0.01%)
order 2	0.00% (0.02%)	0.00% (0.00%)	0.00% (0.00%)

Table 1: Average market clearing errors for IRFs (sup norm in parentheses), computed as percentage differences normalized by the steady state value of the variable (when it applies), or by the initial value of the series.

4 Calibration

The model is calibrated at quarterly frequency to match annualized targets at the deterministic steady state. The benchmark calibration is summarized in Table 2. To deliver a realistic picture of households' use of liquid assets and short-term debt to smooth aggregate income and credit shocks over the business cycle, the model is calibrated to match level and cross-sectional facts on unsecured household credit in the U.S., government taxes and transfers, and the risk-free rate.

[12]This gives more accurate results than feeding the pruned system with a series of randomly generated shocks, since even with pruning the system may diverge if shocks are large.

Parameter	Explanation	Value	Target/source
β	Discount factor	0.9925	Risk-free rate $r = 2\%$ (FRB)
γ	Coefficient of relative risk aversion	5	–
η	Curvature of disutility of working	2	Frisch elasticity $= 1/2$
ψ	Disutility of working	11.5	Normalize $Y = 1$
$\overline{\phi}$	Average credit shock	2.6	Unsecured debt-to-GDP 0.18 (FRB)
$\phi(\theta)$	Credit limit function	$(1, 1.03, 1.06, 1.08, 2.33)$	Debt dispersion across incomes (SCF)
α	Cobb-Douglas parameter	$2/3$	Labor share of $2/3$
$\tau_1(\theta)$	Tax function	$(0.05, 0.13, 0.17, 0.20, 0.28)$	Tax dispersion across incomes (CPS)
$T(\theta)$	Transfer function	$(1, 0.43, 0.24, 0.17, 0.13)$	Transfer dispersion across incomes (CPS)
B	Bond supply	6	Liquid assets-to-GDP 1.78 (FRB)
ρ_θ	Persistence of productivity shock	0.977	AC wage process (Kopecky and Suen (2010))
σ_θ	Std. dev. of productivity shock	0.12	Std. dev. wage process (Kopecky and Suen (2010))
ρ_ϕ	AC credit shock	0.99	AC risk-free rate 0.65 (FRB)
σ_ϕ	Std. dev. credit shock	0.025	Std. dev. risk-free rate 1.9% (FRB)
ρ_z	Persistence TFP shock	0.86	AC TFP
σ_z	Std. dev. TFP shock	0.0128	Std. dev. TFP
$\rho_{\phi z}$	Cor. credit and TFP shocks	0.5	Cor. debt-income 0.9 (FRB, BEA)

Table 2: Parameters, quarterly frequency. Targets are annualized. The top panel lists parameters for the household's problem, the middle-top panel for the firm's problem, the middle-bottow panel for the government, and the bottom panel for aggregate shocks.

Household credit For unsecured debt and liquid assets to GDP, I target values of 0.18 and 1.78, and obtain 0.22 and 1.73. Those are the same as in Guerrieri and Lorenzoni (2017), from the Flow of Funds of the Federal Reserve Board (FRB) for 2006[13]. For that, I choose $\overline{\phi}$ to be equal to 2.4, and the sum of all households NAP, B, equal to 6 (1.5 quarterly).

To calibrate variables' distributions across income types, I map the income distribution in the model to the data by constructing the corresponding productivity quintiles. I then compute the empirical counterpart and distribution of model variables. In the steady state, 6.25% of households are in the first income bin, 25% in the second, 37.5% in the third, 25% in the fourth, and 6.25% in the fifth.

In the 2007 Survey of Consumer Finances (SCF), I compute total unsecured credit as total household debt minus the total value of debt secured by primary residence (including mortgages and HELOC) and the total value of debt for other residential properties. This leaves other lines of credit, credit card balances, installment loans (including education and vehicle loans), and other debt. In the five wage income bins, average (median) unsecured debt in 2013 dollars is respectively: 130,190 (0); 17,150 (450); 25,340 (6,512); 141,190 (9,431); 302,920 (0). Replicating this distribution would require lower debt limits

[13]This is respectively data from the Federal Reserve Board Flow of Funds (Z.1) table B.100, sum of lines 9, 16, 19, 20, 21, 24, and 25; and from table B.100, line 34, which essentially corresponds to total household liabilities minus mortgage debt.

for θ_2 and θ_3 agents than for θ_1, at odds with the empirical evidence that credit limits typically increase with income (a key determinant of credit limits together with credit scores). The reason for this empirical pattern is likely due to facts not captured by the model, which lead very low income agents to accumulate a lot of debt and very high income agents to do the same[14]. Therefore, I smooth the distribution by assuming that debt limits are uniformly distributed from θ_1 to θ_4, and thus are increasing with income. I interpret these amounts as rough measures of the relative debt limits faced by the various income types, and therefore calibrate the idiosyncratic component of debt limits as $\phi(\grave{}) = (1, 1.03, 1.06, 1.08, 2.33)$ (values normalized by the value for θ_1). This allows to match the dispersion of average debt across income types (for the same reason as above, I choose not to replicate the distribution of median debt, which shows a lot of heterogeneity within each income bin). The dispersion of debt limits at the bottom of the income distribution, where more households are constrained, interacts with consumption smoothing, and determines the strength of the precautionary savings motive.

In each wage income bin, average debt to average income is respectively: Inf[15], 3.13, 0.56, 0.78, 0.15. In the model, household debt is positively related to their debt limits, and is endogenously chosen. Households occasionally borrow up to their debt limits. In particular, households with lower incomes are able to borrow a lower maximum amount than richer households, they borrow less on average when they do, and they are more frequently constrained. The fractions obtained are: 5.15, 1.09, 0.47, 0.05, 0.01. Overall, it matches the decreasing profile of average debt to average income, and the fact that low income households hold on average disproportionately more debt relative to their income.

Taxes and transfers Similarly, I choose the tax and the transfer functions to replicate the empirical distribution of taxes and transfers across incomes. In Congressional Budget Office data for 2006 ("The Distribution of Federal Spending and Taxes in 2006", Exhibit 18 in Congressional Budget Office (2006)), average total transfers to nonelderly households for the 5 quintiles of pretax market income (productivity in the model) are respectively

[14] It is beyond the scope of this paper to incorporate these facts in the model. Those probably include large liquidity shocks (such as health expenditures), the cost of college education, etc.

[15] Average income is zero in the first bin because wage income is zero for unemployed agents, so the ratio goes to infinity. I do not replicate average debt to income by income bin, because the very low values of income and the very high values of debt for some of the low income households, and the fact that some households have both zero debt and income, give the following extreme respective fractions: NaN, 3,187; 0.686, 0.643, 0.169.

(in 2006 dollars per households[16]): 15,200; 6,600; 3,700; 2,600; 2,000. Normalizing by the value for θ_1, it delivers a transfer function $T(\grave{}) = (1, 0.43, 0.24, 0.17, 0.13)$. Transfers represent 6.9% (9%) of total average (median) market income. To match that share in the model, I use a multiplicative factor identically applying to all transfers, which is set to 1. Taxes by income quintiles are 2,600; 6,500; 11,800; 19,700; 68,100. This delivers a tax function $\tau_1(\grave{}) = (0.05, 0.13, 0.17, 0.20, 0.28)$.

Risk-free rate The risk-free rate is a key price in the model because it clears the bond market, whose equilibrium is directly impacted by credit fluctuations. The discount factor β is chosen equal to 0.9925 to match the annual Treasury Inflation Indexed long-term average yield of about 2.5% in 2009Q3 (based on the unweighted average bid yields for all TIPS with remaining terms to maturity of more than 10 years; source: FRB, H.15 Selected Interest Rates). Pflueger et al. (2017) reports a value of 2.4% for the period 1973Q1-2014Q4, using the one-year Treasury bill rate net of one-year survey expectations of the inflation (GDP deflator) from the Survey of Professional Forecasters, expressed in percentage terms and linearly detrended. The value obtained using the lending rate adjusted for inflation using the GDP deflator is 2.5% in 2009 (source: IMF International Financial Statistics for the lending rate, and World Bank for the GDP deflator.). Taxes and transfers give relatively more resources to θ_1 agents with the lowest income than to θ_2 and θ_3 agents with slightly higher incomes. Under that calibration, the model would generate a lower fraction of θ_1 agents constrained than of θ_2 and θ_3 agents, at odds with empirical evidence that low income agents are on average more financially constrained. Therefore, I assume that agent with the lowest income θ_1 are also more impatient, with a 20% lower discount factor than the rest of agents.

The annual persistence and volatility of the risk-free rate in the data are 0.65 and between 1.50% (FRB) and 1.90% (Pflueger et al. (2017)). To match those, I reverse-engineer the quarterly autocorrelation and volatility of the credit shock process, respectively 0.99 and 2.5%. The model counterpart for the risk-free rate persistence and volatility are estimated in a $T = 10,000$ period simulation of the linearized version of the model with random TFP and credit shocks, and then annualized[17]. These estimates from the model

[16]The dollar denomination is irrelevant, since I only use these numbers to calibrate the dispersion of variables across income types.

[17]I do not use the second-order approximation because the variable paths may diverge in long-run simulations even with pruning, a well-known problem with high-order approximations. The second-order approximation matters more for the average level than for the autocorrelation and standard deviation of the risk-free rate. Thus even if computing the stochastic steady state based on a long-run simulation was possible, the estimated values would likely be very similar to those of the first-order simulation, which correspond to the deterministic steady state.

are well-identified (Figure 8). In theory, an alternative would be to use unsecured credit changes from the data to estimate the parameters of the $\overline{\phi_t}$ process. In practice, however, there is little change over time in credit limits for given (household,loan) observations, at odds with what the model would imply. In particular, falls in debt limit for given (household,loan) observations are infrequent, while changes in aggregate credit conditions continuously happen. The latter only indirectly induce changes in household debt, through mechanisms absent from the model (for instance, stricter application requirements, lower debt limits on new loans[18], long-term debt). On the other hand, the model has clear implications about how credit shocks drive the risk-free rate required by households to hedge against deviations from consumption smoothing. Thus using them to calibrate the $\overline{\phi_t}$ process is natural and avoids taking a stance on the ultimate nature of credit shocks, still an open question in the macro-finance literature. Finally, I calibrate the correlation between TFP and credit shocks to match the empirical correlation between unsecured household debt and household income in the data. Using various linearly detrended measures for consumption (personal consumption expenditures) or income (personal income) from the Bureau of Economic Analysis (BEA), and outstanding total consumer credit owned and securitized from the FRB (G.19 Consumer Credit), this correlation is about 0.90. It is 0.71 in a 10,000 period simulation of the first-order model.

Households The coefficient of relative risk aversion γ determines the strength of the precautionary savings motive. I set it equal to 5, an intermediate value between Guerrieri and Lorenzoni (2017) (4) and Favilukis et al. (2017) (8), and experiment with different values. Larger values induce a lower risk-free rate and less debt in the stochastic steady state, and more amplification in the economy's response to shocks. For some parameter combinations, values below 4 imply an *increase* in aggregate consumption together with a decrease in debt to GDP following a *tightening* of borrowing constraint, a possibility so far undocumented in models of credit crises. The reason is that if a low number of agents are constrained, and all agents have low precautionary savings motives, then the decrease in consumption by (close to) constrained households who mechanically deleverage may be more than offset by the corresponding increase in consumption by richer households, who decrease their NAP in order for the bonds market to clear when poorer households demand more bonds (this is facilitated by the sharp decrease in the risk-free rate), and thus save less and consume more. Those effects would change if one allowed for aggregate NAP B (outstanding government bonds) to vary over time instead of taxes τ_{0t}, a

[18]An additional cost on new loans for households with previously positive NAP could be easily added to the model, at the expense of additional complexity.

possibility that would allow the model to speak to the issue of safe assets scarcity. The curvature of the disutility from labor is chosen to obtain a Frisch elasticity of labor supply of 1/2, a standard value in the macroeconomic literature.

The persistence and variance of the discretized productivity process are chosen to match those of the wage process in Kopecky and Suen (2010), 0.977 and 0.12. Unlike Guerrieri and Lorenzoni (2017), the model does not have exogenous unemployment risk, so the volatility of income is larger to proxy for the otherwise missing income volatility[19]. To calibrate the countercyclicality of income volatility, I use the estimate of Storesletten et al. (2004) that the standard deviation of individual income increases by 0.09 from peak to trough. In the model, a one standard deviation negative shock to TFP induces a -0.5% deviation from steady state output. For an average trough of -1.5% deviation from steady state, a three standard deviation TFP shock is needed. Thus I assume that a one standard deviation negative TFP shock raises the income volatility by 0.09/3=0.03 (credit shocks do not have any effect). Countercyclical income risk does not affect the deterministic steady state where TFP shocks are zero, but it slightly amplifies the economy's response to shocks in both the first- and the second-order solution of the model. The reason it affects the first-order solution of the model despite only changing moments of order 2 is that it increases the dispersion of individual income, hence the probability that households hit their borrowing constraints and are prevented to smooth consumption, an effect accounted for in the first-order model. Finally, the TFP persistence and volatility parameters are in line with those used in the macroeconomic literature.

Net worth distribution: model versus data The shape of the distribution determines how the economy responds to shocks. Unlike other aggregate shocks that affect it indirectly, credit shocks directly shift the distribution: to the right when borrowing limits are tightened, to the left when they are relaxed. Table 3 displays summary statistics for the unconditional distribution $\lambda(\mathbf{b})$ of liquid net worth. By construction, the model matches both aggregate assets and debt to income. It generates a realistic distribution, with a mass of constrained households, and a decreasing fraction of households with larger wealth. The fraction of constrained households lies between two of its empirical counterparts, the fraction of poor Hand-to-Mouth households in Kaplan and Violante (2014) (between 3.5% and 21% in 2001) and the fraction of households reporting to be without savings in The Pew Charitable Trusts (2015) (about 33%). It also generates realistic wealth inequality as

[19] It is more technically involved to add an unemployment state when solving for households' labor supply in the system of equilibrium conditions, because the system becomes not differentiable at that point. In the model, zero labor supply can be interpreted as unemployment, but there is no involuntary unemployment.

measured by the mean to median ratio of net worth, but too much inequality as measured by the ratios of various net worth percentiles to income.

Statistic:	Data	Model
aggregate liquid assets/aggregate income	1.78	1.73
aggregate liquid debt/aggregate income	0.18	0.23
mean/median	4.60	4.90
share of Hand-to-Mouth households	0.035-0.33	0.35
P50 liquid net worth/aggregate income	0.15	0.30
P75 liquid net worth/aggregate income	0.68	2.77
P90 liquid net worth/aggregate income	1.69	5.64

Table 3: Distribution of liquid net worth in the model versus the data. Income is in annual terms. Data on net worth percentiles are from Gorea and Midrigan (2017).

5 Results: the Long-Run and the Dynamic Effects of Credit Shocks

The model has static and dynamic implications for how an economy with time-varying credit differs from one with fixed borrowing constraints. Because of the departure from certainty equivalence, households anticipate future credit shocks and insure every period against future binding borrowing constraints. The economy's stochastic steady state shifts in response to increased precautionary motives: on average, households accumulate less debt and the risk-free rate is lower. For low magnitudes of the credit shocks, the large negative financial response only mildly affects real variables, because prices are flexible and adjust to clear markets. When credit shocks are large, however, financial conditions affect real variables because of a composition effect that induces less productive households to work more and more productive ones to work less, and because of the wealth effect on labor supply. This gives rise to an economy with persistently low output, employment, debt and risk-free rate. Therefore, in terms of long-run aggregate consumption, credit changes lead to substantially larger costs of business cycles than those estimated by Lucas (2003) with only TFP shocks.

Credit shocks also affect the economy dynamically, from the time they hit households' borrowing constraints to the time they revert back to steady state. Nonlinearities in the second-order perturbation of the model amplify the economy's response to a tightening of borrowing constraints, by capturing the different responses of households depending on their debt levels. However, they are close to zero for TFP shocks, which affect all households identically. Thus over the business cycle, when both shocks are present, the economy departs from the "near-aggregation" property of Krusell and Smith (1998) and a representative agent setting. This result differs from Winberry (2016), the only other paper I am aware of to study a second-order perturbation of a heterogeneous agents model, in which the linear and the nonlinear solutions almost coincide. In my model, adding credit shocks make nonlinearities play a key role in shaping the economy's dynamics, because of the unequal effect of a tightening of borrowing constraints across the wealth distribution.

In the following section, I cover those results in more detail.

5.1 Precautionary Savings Motives

Stochastic steady state By taking the difference between the second- and the first-order solution of the model, we obtain a numerical measure of nonlinearities and deviations from certainty equivalence due to aggregate shocks. Table 4 displays the values of macroeconomic aggregates in the deterministic steady state and in the stochastic steady state. While other variables remain identical, there is a slight decrease in the risk-free rate and household debt in the stochastic steady state. This effect is small in the baseline calibration because the magnitude of aggregate shocks is small, households have a moderate risk-aversion, and the impatience to consume of the income-poorest households – who are more constrained – conflicts with their precautionary savings motives that would further lower debt and the risk-free rate. The strength of the precautionary motive is increasing in households' risk aversion, in the persistence and volatility of the credit shock, in the dispersion of debt limits across income types and in the countercyclicality of income risk. As the volatility of credit shocks increases, the risk-free rate required by agents to hold liquid safe assets decreases and so does aggregate debt. As implied by the distributions of net worth conditional on income in the previous section, debt decreases more for low-income and households close to the constraint.

There are three types of precautionary savings motives in the model: at the microeconomic level, at the macroeconomic level, and at both. First, absent aggregate fluctuations, idiosyncratic income risk at the micro level increases the demand for insurance directed towards safe assets. One reason this motive arises is "prudence" (Kimball (1990)): $u'''(c) > 0$, so the volatility of income increases future expected marginal utility because of Jensen's inequality, which implies a decrease in current consumption and an increase in savings in households' Euler equation. Another reason is the presence of borrowing con-

Variable	Deterministic SS	SS w/ $\sigma_\phi = 0.025$	$\sigma_\phi = 0.05$	$\sigma_\phi = 0.075$	$\sigma_\phi = 0.10$
r	2.397%	-0.1%	-1.4%	-7.4%	-25.4%
w	1.491	0%	+0.07%	+0.3%	+0.9%
π	0.333	0%	0%	-0.6%	-1.5%
N	0.447	0%	-0.2%	-0.7%	-2.5%
Y	1.000	0%	0%	-0.5%	-1.6%
C	1.000	0%	0%	-0.5%	-1.6%
$B^-/(4Y)$	0.229	-0.4%	-2.7%	-12.5%	-45%

Table 4: Steady state summary statistics: risk-free rate, wage, profits, employment, output, consumption, aggregate NAP, debt/GDP, assets/GDP. Risk-free rate, debt/GDP and asset/GDP in annual terms. Columns 3-6 are percentage deviations from steady state values (column 2). Note that as σ_ϕ increases the normalization that $Y = 1$ doesn't hold any more, because it applied to the deterministic steady state. Market clearing errors slightly increase, but remain low.

straints: households fear that a sequence of negative income shocks will lead them to hit their borrowing constraints, preventing consumption smoothing; this effect is stronger for households at or near their constraints. Absent this first motive, the steady state risk-free rate would be equal to $400 \times \left(\frac{1}{0.0625\times0.80\times0.9925+(1-0.0625)\times0.9925} - 1\right) = 8.33\%$ (in annualized percentage terms)[20]. With this motive, it is equal to 2.40%, that is 71% lower.

Second, with aggregate fluctuations, a precautionary motive due to TFP shocks arises at the macro level. Its mechanism is the same as the first motive, since TFP shocks result in labor income changes. The difference is that TFP changes across periods are smaller and less persistent than changes in idiosyncratic productivity, and they impact households uniformly. As a result, the resulting decrease in the risk-free rate and aggregate consumption is close to zero. The quasi-absence of an effect on long-run consumption is consistent with the very low estimates of the cost of macroeconomic fluctuations in Lucas (2003), who focuses on TFP shocks.

Third, in the benchmark calibration there is a small positive precautionary motive due to aggregate credit shocks. They contribute to lowering the steady state risk-free rate by an extra 0.13%, to 2.39%. This represents about 0.20% of the magnitude of precautionary savings for pure micro motives. In economies or times with volatile credit ($\sigma_\phi = 0.10$), they contribute to lowering the steady state risk-free rate by an extra 25%, more than a third of micro precautionary savings. In that case, household credit is not neutral in the

[20] The benchmark discount factor is $\beta = 0.9925$. A fraction 0.0625 of agents have a 20% lower discount factor, so the average discount factor is equal to $0.0625 \times 0.80 \times 0.9925 + (1 - 0.0625) \times 0.9925 \approx 0.979$.

long-run, and its changes substantially increase the cost of macroeconomic fluctuations, by decreasing long-run aggregate consumption by 1.6%.

Policy functions To understand how credit fluctuations affect macroeconomic aggregates in the long-run, I plot the policy functions of the median income household when $\sigma_\phi = 0.10$ in Figure 2. The blue line is for the deterministic steady state, with zero probability credit shocks, and the orange line is for the stochastic steady state, where households anticipate credit shocks. When this is the case, households consume less, save and work more to achieve higher precautionary savings.

A comparison of policy functions across income levels (available upon request) reveals that this is especially true for more patient, low-income households (θ_2 and θ_3). Those benefit less than the lowest income households (θ_1) from the progressivity of taxes and transfers, and they are less rich than more productive households (θ_4 and θ_5), so their precautionary motive dominates their impatience to consume. Finally, unlike other households, the most productive ones (θ_5) consume slightly more when anticipating credit shocks. Accordingly, they also save less and work less. This is due to the lower risk-free rate, which creates incentives for them to dissave and consume more of the consumption good and of leisure. With a fixed supply of bonds B, the lower risk-free rate clears the bond market when the demand for bonds from poorer households is higher, and that from richer households is lower.

Comparison with TFP shocks The same sensitivity exercise as earlier, for the volatility of TFP shocks (holding other parameters at their benchmark values), reveals that the differences between the deterministic and the stochastic steady states are close to zero, even for large TFP volatility (results available upon request). For instance, increasing $\sigma_z = 0.0128$ to $\sigma_z = 0.03$ does not affect variables compared to the deterministic steady state. Increasing it by a factor of five, to $\sigma_z = 0.06$, slightly decreases the risk-free rate by 0.10% (to 2.394%), while other variables remain unchanged. Thus credit volatility is a more important source of precautionary savings than TFP fluctuations. The lesser importance of TFP shocks for precautionary savings is consistent with the literature on high-order approximations of representative agent RBC models. In those models, the steady state and the dynamics of first- and second-order approximations of the economy are usually nearly identical (Fernandez-Villaverde et al. (2016)).

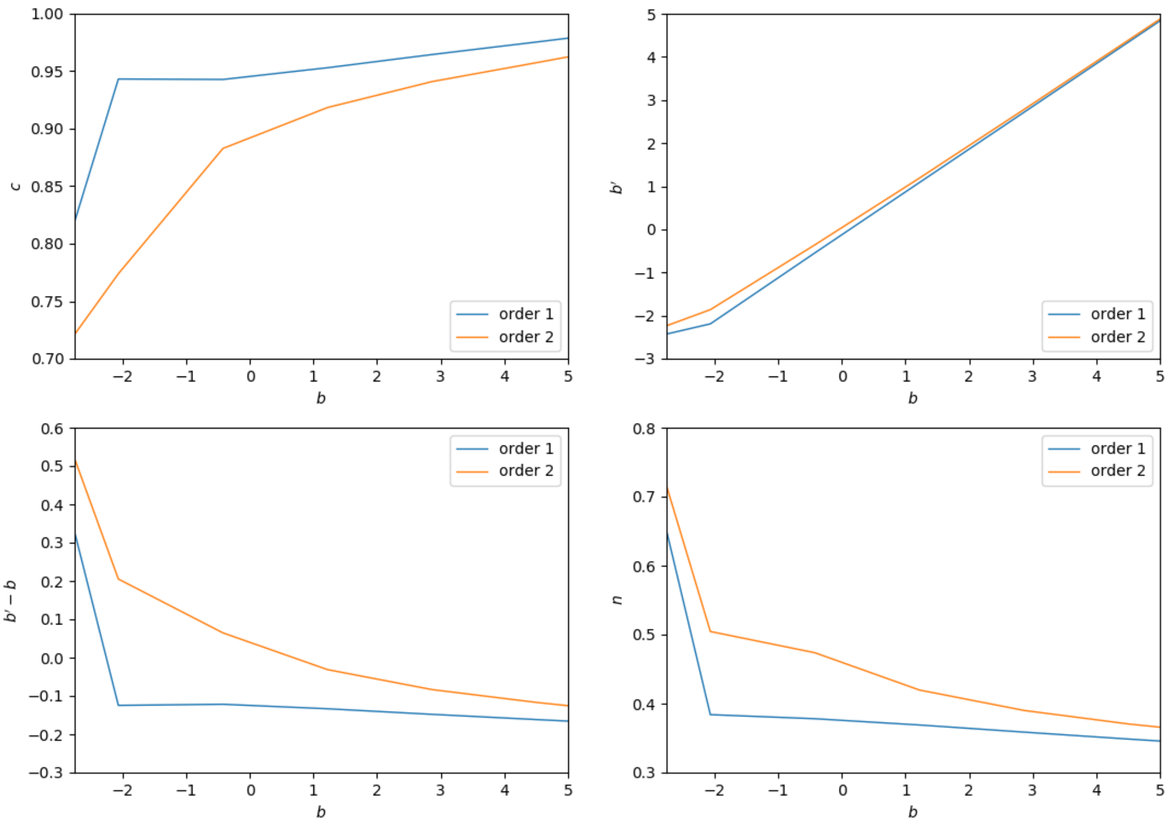

Figure 2: Policy function of the median income household as a function of net worth, starting at the corresponding steady state credit limit $\overline{\phi}\phi\left(\theta\right)$ for: consumption, next period bonds, savings and labor supply. Order 1 (blue) vs order 2 (orange).

5.2 The Dynamic Response to Credit Shocks

I now compare the economy's impulse response functions (IRFs) to credit and TFP shocks, in the first- and in the second-order approximations of the model. Nonlinearities amplify the economy's response to credit shocks, so that financial and real variables stay longer at lower values following a tightening of households' borrowing constraints. In the first-order approximation, households respond to the levels of current and expected future shocks, but not to their volatility, because of the certainty equivalence property of the model. Policy functions, the cross-sectional distribution and prices respond linearly to shocks, and are linear functions of their lagged values. In the second-order approximation, households anticipate aggregate shocks, whose volatility enter linearly in the IRF.

In addition, the economy evolves nonlinearly with respect to the level of shocks, with variables being linear and quadratic functions of their lagged values.

Amplification Figure 3 plots the economy's response to a one standard deviation credit shock, under the linear dynamics with certainty equivalence (order 1), versus the nonlinear dynamics and households anticipating credit shocks (order 2). Aggregates are computed using the time-varying path of individual policy functions and of histogram weights. Deviations are from the deterministic steady state (results are close if instead I compute deviations from the stochastic steady state). All households' debt limits fall, but lower income households are able to borrow much less than the richest ones, reflecting idiosyncratic differences in their ability to borrow. As a result, constrained households are forced to reduce their debt and increase their NAP, thereby decreasing their consumption of goods and leisure. They trade-off working more to smooth consumption against the disutility of labor. As a result, aggregate debt to GDP decreases, and stays persistently low, mostly because of the large persistence of the credit shock. The decrease in total consumption results from the composition of low-income constrained households decreasing their spending, and richer unconstrained households increasing theirs because they earn a lower return on their bonds. The decrease in the risk-free rate allows to balance a larger bond demand from the former with a lower demand from the latter, and to clear the bond market. Aggregate employment (not measured in efficiency units), decreases, causing a decline in output (I come back to the employment effect of credit shocks below). As implied by the economy's resource constraint, consumption falls like output.

Accounting for nonlinearities significantly amplifies the response of aggregates to a credit shock, contrasting with the findings of Winberry (2016) for TFP shocks. The initial response is amplified by a factor of 5 for debt to GDP, of 4 for the risk-free rate, of 1.5 for consumption, output and profit, and of 1.4 for the wage. While the risk-free rate quickly reverts back to steady state, other variables stay persistently low. The sharp decline in the rate causes consumption and employment to rebound (simultaneously, profits slightly increase and the wage slightly decreases). However, the rebound is short-lived, and the large persistence of the credit shock that induces borrowing constraints to stay persistently low, further decreases consumption and employment. Debt to GDP stays persistently low and barely rebounds. The price adjustment (the decrease in the risk-free rate) cannot offset the quantity restriction imposed by tighter borrowing constraints, which mechanically force constrained households to hold less debt. At the household level, these changes can be decomposed into changes in the cross-sectional distribution and in policy functions.

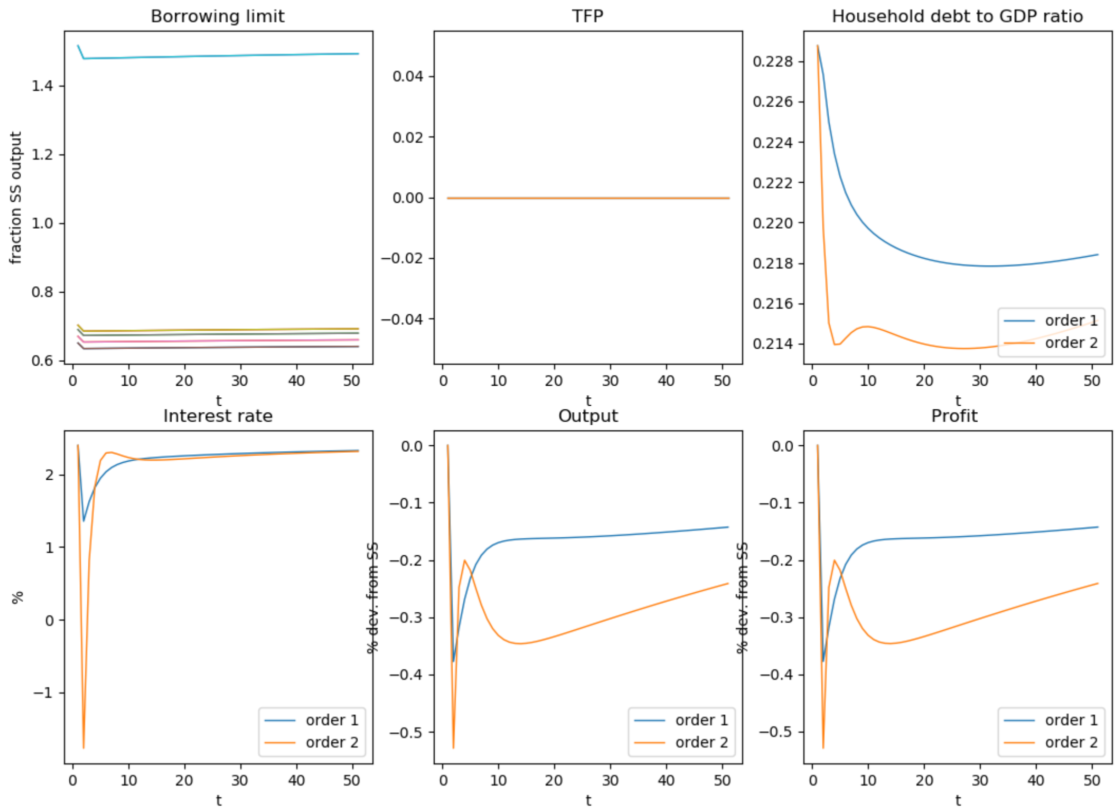

Figure 3: IRF to a one standard deviation credit shock. Borrowing limits (upper left panel) as a fraction of annual steady state output (upper left panel) are for θ_1 (lowest line) to θ_5 households (highest line). Other panels plot IRFs in the 1st versus the 2nd order approximation of the model. Initial period: deterministic steady state.

Employment The employment response in Figure 4 results from the composition of less productive, constrained households that increase their labor supply to smooth consumption when they are forced to deleverage, and of more productive, unconstrained households that consume more leisure as they decrease their savings (wealth effect). In addition, the sharp decline in the risk-free rate creates an intertemporal substitution effect, which induces all households to consume more leisure in the current period. The sign of the labor response depends on which effects dominates. Here output declines, mainly because more productive agents work less, despite less productive agents working more. This result differs from Guerrieri and Lorenzoni (2017), where employment increases after a credit shock in their benchmark calibration (however, they mention alternative calibrations where it decreases). The difference between employment responses can be explained as follows. In this paper, the credit shock magnitude is relatively small and it is mean-reverting. In their paper, the economy is hit by a large and permanent deleveraging shock that unfolds over several periods, and is perfectly foreseen by agents. As a result, constrained households and those expecting to be constrained in the near future (because of the anticipated path of the constraint) choose to work more. This effect is dampened in my model, because agents expect the credit shock to mean-revert. Thus even with flexible prices, smaller and mean-reverting credit shocks can generate larger negative employment responses than unanticipated, large and permanent deleveraging episodes[21]. This interpretation is warranted by the fact that the negative employment response is barely amplified in order 2. That is, nonlinearities only affect the labor supply of those households that are (close to) constrained, and amplify the *increase* in their labor supply, as seen in their policy functions earlier. This additional increase in their labor supply almost exactly offsets the decrease in the labor supply of all households induced by the sharper fall of the risk-free rate in order 2, and the decrease in the labor supply of the richest households due to the intertemporal substitution effect and the wealth effect. As a result, the employment response is almost identical in order 1 and in order 2. Another distinctive feature of my model is the presence of a firm sector hiring labor at a competitive wage, which negatively depends on aggregate employment in efficiency units, closely related to output. The decrease in aggregate employment is consistent with the increase in wages, which induces firms to demand less labor.

The employment response also differs from Midrigan and Philippon (2016), where employment slightly increases after a mean-reverting credit shock on a given "island". The difference with my model is that they effectively study a representative agent economy: to increase the island's NAP and smooth consumption, the agent is forced to increase her labor supply when her borrowing constrained is tightened. Here, the employment response is the composition of individual employment responses across the entire distribution of productivity and NAP. Because the sum of NAP, B, is fixed, poorer constrained agents increasing their NAP implies richer unconstrained agents decreasing theirs, consuming more good and leisure. Consistently with the respective employment responses, the wage increases here, while it decreases in their model.

Departure from "near-aggregation" To compare, I plot the economy's response to a one standard deviation negative TFP shock (Figure 9 in appendix), in the first- versus

[21] In the benchmark calibration of Guerrieri and Lorenzoni (2017), a smaller, mean-reverting credit shock would generate a decline in employment.

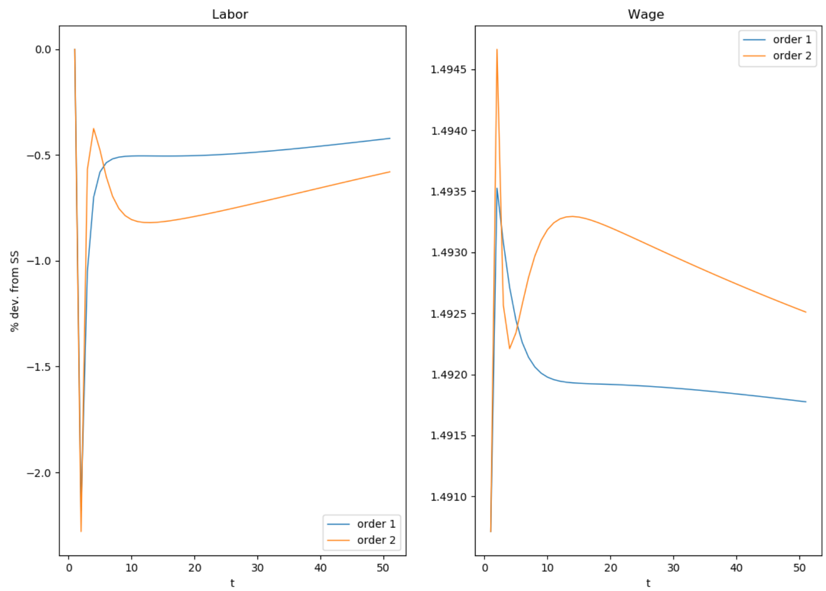

Figure 4: IRF to a one standard deviation credit shock for aggregate employment and wages, in order 1 vs order 2. Initial period: deterministic steady state.

the second-order approximation of the model. Unlike with credit shocks, the economy's dynamics response is the same as in representative agent RBC models. Moreover, the economy's response in order 1 versus order 2 is close to identical (except for debt to GDP and the risk-free rate, for which we observe small differences), and is linear. In a model with household credit, the "near-aggregation" property of Krusell and Smith (1998) holds with respect to TFP, but not credit shocks.

With a shock to TFP alone, borrowing constraints are fixed. Because the economy becomes less productive, a decrease in TFP induces a fall in output (hence in profits). Because it induces a fall in the wage (the marginal productivity of labor decreases), households' income decreases and they supply more labor to compensate for the decrease. The wage decline induces a higher labor demand from firms. Summing the two effects, aggregate employment increases. To smooth their consumption when their incomes decline, lower income unconstrained households issue more debt, while richer households increase their savings. This increases wealth inequality. The risk-free rate increases to clear the bond market, with a larger demand for debt by low-income households and a larger demand for safe assets by richer ones. Together with the decline in output, this induces debt to GDP to increase.

Finally, the combination of credit and TFP shocks generate a realistic business cycle dynamics (see Figure 10 in appendix), which I exploit in the next section to back out the sequence of structural shocks that explain the Great Recession period. Both households' borrowing constraints and TFP fall. The effects of the TFP shock dampen those of the credit shock. Debt to GDP decreases, but less than implied by the single credit shock, since low-income, unconstrained households resort to debt to smooth their consumption when their labor income declines. This partly offsets the precautionary savings effect implied by tighter borrowing constraints. Similarly, the risk-free rate and employment decrease, but much less than with just a credit shock. This is partly due to the decrease in the wage resulting from lower TFP, which makes hiring workers more attractive to firms. However, the negative effect of lower TFP and tighter borrowing constraints on consumption (hence output) is cumulative, so that it decreases twice as much as with only one shocks. Firm profits follow the decline in output. Overall, accounting for non-linearities amplifies the recessionary dynamics triggered by credit shocks, and combining credit with TFP shocks makes the risk-free rate and debt to GDP dynamics over the business cycle less extreme than with credit shocks only.

6 Credit Shocks and Macroeconomic Volatility

To conclude, I investigate how changes in household credit contribute to macroeconomic volatility over the business cycle. First, I do a variance decomposition exercise using the economy's nonlinear laws of motion, to distinguish the contribution of credit from that of TFP. Except for wages, the model assigns about half of business cycle volatility to credit shocks, making it an important source of uncertainty for households. Then, I apply a particle filter to the model to back out the sequence of structural credit and TFP shocks that generated the observed paths for the risk-free and aggregate consumption around the Great Recession, from 2006 to 2017. The nonlinear solution of the model allows to match the data even in times of high volatility, including data on debt and employment that were not targeted. While TFP only fell during the Great Recession itself (2008-2009) and quickly reverted to its pre-crisis level, the implied measure of aggregate credit (borrowing constraint tightness) kept decreasing until 2014, and has stayed persistently low throughout the post-Great Recession period.

6.1 Volatility Contribution: Credit versus TFP

For the variance decomposition exercise, I set the correlation between the credit and the TFP shocks to zero to interpret them as structural shocks. I detail the exercise in appendix (subsection A.4), which uses the nonlinear laws of motion of the economy. As Table 5 shows, the model assigns about or more than half of the variation of macroeconomic aggregates to credit shocks – except for wages, which scale one by one with TFP. This happens even in the absence of price rigidities, because of the adverse reaction of employment analyzed in the previous section. This effect is strengthened by the volatility and persistence of credit shocks, much higher than their TFP counterparts. This is consistent with the IRF seen earlier, where the effect of a credit shock on output was much more persistent than for a TFP shock.

Variable:	Credit shock	TFP shock
q	59%	41%
N	52%	48%
w	21%	79%
π	59%	41%
Y	59%	41%
τ_0	57%	43%

Table 5: Variance decomposition: shares of the variance of variables in the first column accounted for by credit (second column) and TFP shocks (third column): bond price, employment, wages, profits, output, taxes. Variance shares are computed by bootstrap, as the Monte-Carlo average of the variance decompositions of generalized forecast errors at a large forecasting horizon ($H = 1000$ periods). Computations use $N = 500$ simulations.

6.2 Structural Shocks: the Great Recession

I now use the model to recover measures of structural credit and TFP shocks implied by data on the risk-free rate and aggregate consumption around the Great Recession episode, from 2006Q3 to 2017Q2. This period of high volatility is a good test of the ability of the nonlinear model with household credit to match large changes in financial (risk-free rate, debt) and real variables (output, employment).

My empirical measure for the risk-free rate is the 5-Year Treasury Inflation-Indexed Security (Constant Maturity rate; source: Board of Governors of the Federal Reserve System), and Real Personal Consumption Expenditures (source: U.S. Bureau of Economic Analysis) for consumption. Because the deviation of consumption from its initial value in 2006Q3 is nonstationary over this period, I detrend it using a Hodrick-Prescott filter, and subtract the resulting initial value to normalize the detrended deviation to zero in the first period of the sample.

Because the model is nonlinear, I use a particle filter to recover the sequence of structural shocks leading to the dynamics of the risk-free rate and consumption over this period. The computations are detailed in appendix (subsubsection A.5.2). As Figure 5 shows, the model matches the data over the sample. It only slightly misses periods when the risk-free rate takes extreme values, and in general the model forecast is slightly higher than its empirical counterpart. This may be due to parsimony of the credit shock process, modeled as an AR(1), and the absence of other sources of low risk-free rates in the model. The tight link between consumption and TFP through the goods market clearing condition allows to match the sequence of consumption deviations exactly.

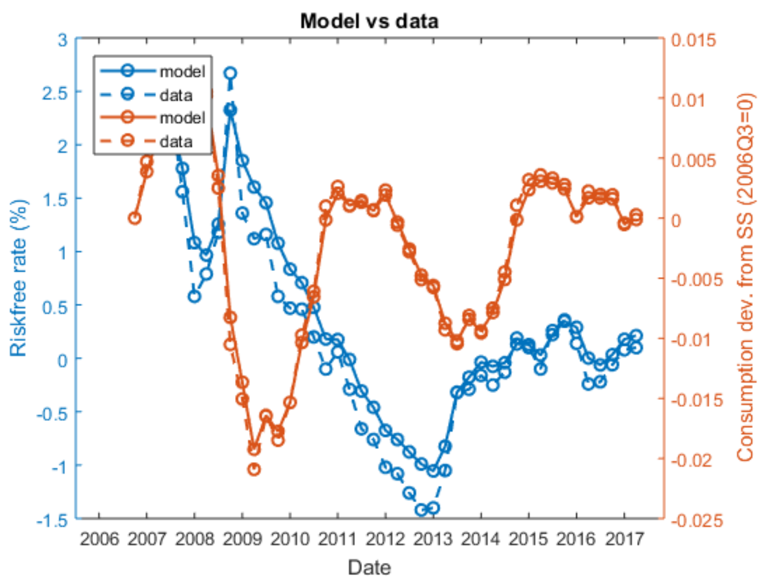

Figure 5: Risk-free rate (annualized, percentage) and consumption deviation from 2006Q3 value, predicted by particle filtering in the nonlinear approximation of the model (solid line) vs the data (dashed line). $N = 20,000$ particles, computations parallelized over the N dimension. The risk-free rate (left axis, blue) is measured as the 5-Year Treasury Inflation-Indexed Security, Constant Maturity rate, not seasonally adjusted (source: Board of Governors of the Federal Reserve System). Consumption (right axis, orange) is measured as Real Personal Consumption Expenditures, Billions of Chained 2009 Dollars, quarterly, seasonally adjusted (source: U.S. Bureau of Economic Analysis). Quarterly sample, 2006Q3-2017Q2.

The model-implied path for debt/GDP and employment, which were not targeted by the particle filter, also closely match the data, demonstrating the good fit of the model (Figure 6). Here, the empirical measure for household debt is Total Revolving Credit Owned and Securitized (source: Board of Governors of the Federal Reserve System), it is Personal Consumption Expenditures (source: U.S. Bureau of Economic Analysis) for GDP (in the model, output and consumption coincide), and Civilian Employment-Population Ratio (source: U.S. Bureau of Labor Statistics) for employment (the model features a continuum of measure 1 of households, so N is the ratio of employed to the entire population). The model replicates the hump-shaped pattern of household credit/GDP, with the run-up to the crisis until 2008, the decline and then the increase in credit around 2015. However, it fails to match the data in the last part of the sample, by overstating the decline in debt/GDP and in employment, for which the large persistence of credit shocks in the benchmark calibration and the wealth effect on labor supply may be responsible.

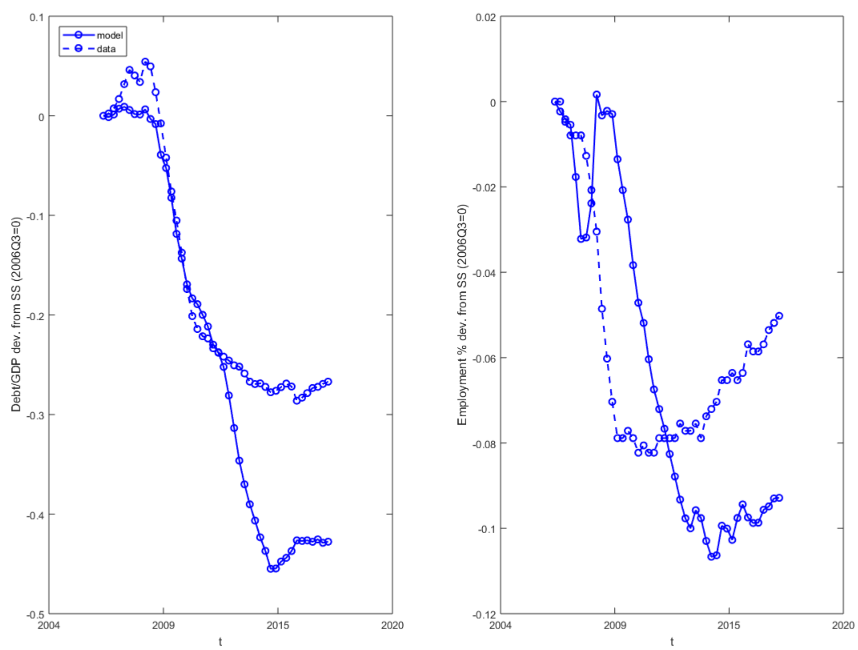

Figure 6: Debt/GDP (left panel) and employment (right panel) implied by risk-free rate and consumption data, recovered by particle filtering. Model (solid line) versus data (dashed line). $N = 20,000$ particles. Variables are in log-deviations from their 2006Q3 values. Quarterly sample, 2006Q3-2017Q2.

Finally, I plot the implied measures of structural credit and TFP shocks around the Great Recession period in Figure 7. By seeking to match the risk-free rate path over the sample, the model-based measure of credit shocks reflects how households' insurance opportunities through short-term debt and precautionary motives changed with the availability of credit. In the model, the fall in the real rate over the sample from 2.5% to -1.5% (annually) is rationalized by a large tightening of households' borrowing constraints. Compared to its 2006Q3 value, the availability of credit decreases by more than 15%, and consistently stays 10% lower throughout the sample. The tightening prevents households from using debt to smooth consumption fluctuations, which are exacerbated by the collapse of TFP, and it forces constrained households to quickly deleverage. To clear the bond market, the risk-free rate collapses and stays persistently low, as long as borrowing constraints remain tight. Credit shocks are a slow-moving state variable, with successive regime changes, inducing lower frequency fluctuations in debt and the risk-free rate. In contrast, TFP changes are more frequent, and track the higher frequency fluctuations in consumption. The 2% decline in TFP happens around 2008, to help match the fall in aggregate consumption, and then reverts to its pre-crisis level within less than two years, and is more than 2% higher at the end of the sample. As I show in appendix, the model-based measure of TFP strongly correlates with empirical measures of TFP over the sample, which can be interpreted as an external check of the validity of the model (Figure 14).

7 Conclusion

This paper attempts to measure macroeconomic precautionary savings motives, a complementary channel to the traditional microeconomic motive. As far as those are concerned, credit supply shocks have very different effects from technology shocks. They generate a precautionary motive with a large negative effect on the risk-free rate, and have unequal effects on the wealth distribution, which generate strongly nonlinear responses of aggregates. In an economy with flexible prices, they account for about half of the volatility of real and financial variables over the business cycle. In an economy with nominal rigidity, the response of aggregate quantities may be further magnified. An interesting extension may thus be to incorporate such frictions into the model, and connect with the recent heterogeneous agents, New Keynesian literature.

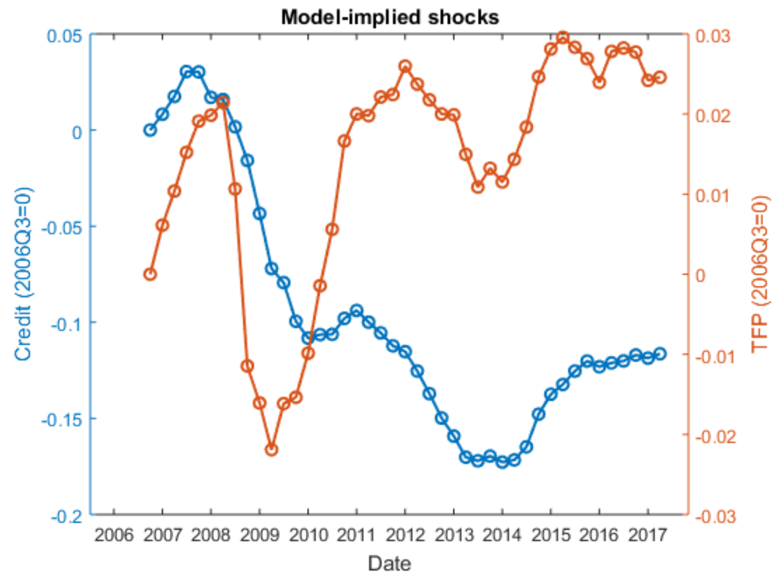

Figure 7: Structural credit (left axis, blue) and TFP shocks (right axis, orange) recovered by particle filtering. Variables are in log-deviations from their 2006Q3 values. Quarterly sample, 2006Q3-2017Q2.

References

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf, "When Inequality Matters for Macro and Macro Matters for Inequality," June 2017.

Boz, Emine and Enrique Mendoza, "Financial innovation, the discovery of risk, and the U.S. credit crisis," *Journal of Monetary Economics*, 2014, *62*, 1–22.

Carroll, Christopher D., "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, June 2006, *91*, 312–320.

_ **and Andrew A. Samwick**, "How Important is Precautionary Saving?," *The Review of Economics and Statistics*, August 1998, *80* (3), 410–419.

Congressional Budget Office, "The Distribution of Federal Spending and Taxes in 2006," Technical Report 2006.

Dávila, Eduardo and Thomas Philippon, "Incompleteness Schocks," December 2017.

Favilukis, Jack, Sydney C. Ludvigson, and Stijn Van Nieuwerburgh, "The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium," *Journal of Political Economy*, 2017, *125* (1), 140–223.

Fernandez-Villaverde, Jesus, Juan F. Rubio-Ramirez, and Frank Schorheide, "Solution and Estimation Methods for DSGE Models," in John B. Taylor and Harald Uhlig, eds., *Handbook of Macroeconomics*, Vol. 2A, Elsevier B.V., 2016.

Fulford, Scott L., "How important is variability in consumer credit limits?," *Journal of Monetary Economics*, 2015, *72*, 42–63.

Gorea, Denis and Virgiliu Midrigan, "Liquidity Constraints in the U.S. Housing Market," April 2017.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman, "Investment, Capacity Utilization, and the Real Business Cycle," *The American Economic Review*, 1988, *78* (3), 402–417.

Guerrieri, Veronica and Guido Lorenzoni, "Credit Crises, Precautionary Savings, and the Liquidity Trap," *The Quarterly Journal of Economics*, 2017, *132* (3), 1427–1467.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "Optimal Tax Progressivity: An Analytical Framework," *The Quarterly Journal of Economics*, June 2017, *1* (62).

Huber, Kilian, "Disentangling the Effects of a Banking Crisis: Evidence from German Firms and Counties," *American Economic Review*, 2018, *108* (3), 868–898.

Huo, Zhen and Jose-Victor Rios-Rull, "Financial Frictions, Asset Prices, and the Great Recession," Staff Report 526, Federal Reserve Bank of Minneapolis February 2016.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti, "Household leveraging and deleveraging," *Review of Economic Dynamics*, 2015, *18*, 3–20.

Kaplan, Greg and Giovanni L. Violante, "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, July 2014, *82* (4), 1199–1239.

_, Kurt Mitman, and Giovanni L. Violante, "The Housing Boom and Bust: model Meets Evidence," August 2017.

Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A. Sims, "Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models," *Journal of Economic Dynamics and Control*, 2008, *32*, 3397–3414.

Kimball, Miles S., "Precautionary Savings in the Small and in the Large," *Econometrica*, 1990, *58*, 53–73.

Klein, Paul, "Using the generalized Schur form to solve a multivariate linear rational expectations model," *Journal of Economic Dynamics and Control*, 2000, *24*, 1405–1423.

Kopecky, Karen and Richard Suen, "Finite State Markov-chain Approximations to Highly Persistent Processes," *Review of Economic Dynamics*, July 2010, *13* (3), 701–714.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri, "Chapter 11 - Macroeconomics and Household Heterogeneity," in "in," Vol. 2 of *Handbook of Macroeconomics*, "Elsevier",, 2016.

Krusell, Per and Anthony Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, October 1998, *106* (5), 867–896.

Lanne, Markku and Henri Nyberg, "Generalized Forecast Error Variance Decomposition for Linear and Nonlinear Multivariate Models," *Oxford Bulletin of Economics and Statistics*, 2016, *78* (4), 595–603.

Lucas, Robert, "Macroeconomic priorities," *American Economic Review*, 2003, *93* (1), 1–14.

Ludvigson, Sydney, "Consumption and credit: a model of time-varying liquidity constraints," *The Review of Economics and Statistics*, August 1999, *81* (3), 434–447.

McKay, Alisdair and Ricardo Reis, "The Role of Automatic Stabilizers in the U.S. Business Cycle," *Econometrica*, 2016, *84* (1), 141–194.

Mian, Atif, Amir Sufi, and Emil Verner, "Household Debt and Business Cycle Worldwide," *The Quarterly Journal of Economics*, November 2017, *132* (4), 1755–1817.

_ **, Kamalesh Rao, and Amir Sufi**, "Household Balance Sheets, Consumption, and the Economic Slump," *The Quarterly Journal of Economics*, 2013, *128* (4), 1687–1726.

Midrigan, Virgiliu and Thomas Philippon, "Household Leverage and the Recession," May 2016.

Parker, Jonathan A. and Bruce Preston, "Precautionary Savings and Consumption Fluctuations," *American Economic Review*, September 2005, *95* (4), 1119–1143.

Pflueger, Carolin, Emil Siriwardane, and Adi Sunderam, "Does Precautionary Savings Drive the Real Interest Rate? Evidence from the Stock Market," August 2017.

Reiter, Michael, "Solving Heterogeneous-Agent Models by Projection and Perturbation," *Journal of Economic Dynamics and Control*, 2009, *33* (3), 649–665.

Revels, J., M. Lubin, and T. Papamarkou, "Forward-Mode Automatic Differentiation in Julia," *arXiv:1607.07892 [cs.MS]*, 2016.

Schmitt-Grohe, Stephanie and Martin Uribe, "Solving dynamic general equilibrium models using a second-order approximation to the policy function," *Journal of Economic Dynamics and Control*, 2008, *28*, 755–775.

Sims, Christopher A., "Solving linear rational expectations models," *Computational Economics*, 2001, *20* (1-2), 1–20.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron, "Cyclical Dynamics in Idiosyncratic Labor Market Risk," *The Journal of Political Economy*, 2004, *112* (3), 695–717.

The Pew Charitable Trusts, "Americans' Financial Security," Brief March 2015.

Winberry, Thomas, "A Toolbox for Solving and Estimating Heterogeneous Agents Macro Models," June 2016.

Appendix

A Computational Appendix

A.1 Algorithm

1. (a) Variables are index by time t to denote their dependence on aggregate states $\left(\overline{\phi_t}, z_t, \lambda_t\right)$. The distribution of households over $\Theta \times \mathcal{B}$ is approximated as a histogram by a finite number of mass points on the Cartesian product of $\Theta = \{\theta_i\}_{i=1}^{N_\theta}$ and a fine bond grid $\{b_j\}_{j=1}^{N_b^f}$. $\Phi_t(\theta_i, b_j)$ denotes the fraction of households with productivity θ_i and bonds b_j. Its evolution is implied by policy functions according to:

$$\Phi_{t+1}\left(\theta_{i'}, b_{j'}\right) = \sum_{i=1}^{N_\theta}\sum_{j=1}^{N_b^f} \Pi_\theta\left(\theta_{i'}|\theta_i\right)\omega_{i,j,j',t} \times \Phi_t\left(\theta_i, b_j\right) \tag{34}$$

$$\text{where} \quad \omega_{i,j,j',t} = \begin{cases} \frac{b'-b_{j'-1}}{b_{j'}-b_{j'-1}} & \text{if } b'_t\left(\theta_i, b_j\right) \in \left[b_{j'-1}, b_j\right] \\ \frac{b_{j'+1}-b'}{b_{j'+1}-b_{j'}} & \text{if } b'_t\left(\theta_i, b_j\right) \in \left[b_{j'}, b_{j'+1}\right] \\ 0 & \text{otherwise,} \end{cases} \tag{35}$$

where $b_{j'-1}, b_{j'}, b_{j'+1}$ are bond points on the fine grid that bracket the value of next period bonds implied by the policy function. ω depend on t because policy functions depend on the aggregate state, i.e. $b'_t\left(\theta_i, b_j\right) = b'\left(\theta_i, b_j; \overline{\phi_t}, z_t, \lambda_t\right)$. For instance, if credit shocks $\overline{\phi_t}$ are low, tightening borrowing constraints, this distorts and shifts upwards the function $b'(.)$ because households are forced to save more, which through its impact on ω results in less mass on low bond values.

(b) Household saving and labor supply policy functions are interpolated using linear splines with respectively N_b and N_n knots. Households' saving function $b'(.)$ is characterized by a critical level of bonds χ_θ at which their borrowing constraints start binding, which depends on productivity. For every $\theta \in \Theta$, let $b_{\theta,j} = \chi_\theta + x_j$, with $0 = x_1 < ... < x_{N^b}$ denote the splines' knots for b' at which households' Euler equations hold with equality. For $b \leq \xi_\theta$, savings $b'(\theta, b) = -\overline{\phi}\phi(\theta)h(\theta)$ are determined by the borrowing limit ($\overline{\phi_t} = \overline{\phi}$ in the deterministic steady state). It defines the collocation nodes at which we force households' optimality conditions to hold to solve for policy functions.

For a given aggregate state $\left(\overline{\phi}, z, \Phi\right)$, the saving function is finitely represented by $N_\theta \times (N_b + 1)$ coefficients giving the value of savings at the knots and the threshold below which households are constrained. So is the labor supply function, with $N_\theta \times N_n$ values at the knots for labor (which may differ from the knots for savings). The consumption function at the saving knots is backed out from the budget constraint:

$$c_t\left(\theta, b_{\theta,j}\right) = b_{\theta,j} + \left(1 - \tau_1\left(\theta\right)\right) w_t \theta n_t\left(\theta, b_{\theta,j}\right) + T\left(\theta\right) + \pi_t - \tau_{0t} - \frac{b'_t\left(\theta, b_{\theta,j}\right)}{1 + r_t} \quad (36)$$

(c) Equilibrium conditions for the discrete model are listed below. The first set of equations and the following two involve predetermined variables: the histogram weights (because weights should sum to 1, we keep only track of the number of weights minus 1), the credit and TFP shocks. The next sets of equations involve jump variables: the bond price, aggregate labor demand, the wage, profits, aggregate output, aggregate consumption, and the (discretized versions of) policy functions for labor and savings (including values of coefficients at knot points and borrowing constraint thresholds). The inclusion of some variables among jump variables, whose dynamics we want to solve for, is not strictly speaking necessary (it is the case for aggregate labor demand, the wage, profits, aggregate output and aggregate consumption). Their equation counterparts are definitional, and their values can be backed out from the other jump variables without including them explicitly in the equilibrium system of equation. However, including them makes the system dynamics better behaved numerically, because it provides more information to the code when taking derivatives with automatic differentiation.
In words, these equations are: the laws of motion for the distribution, credit and TFP; the market clearing conditions for bonds and labor; the definitions of aggregate output[22], consumption, the wage and profits; the intratemporal optimality condition for households' labor supply, and the intertemporal optimality condition for savings/consumption (Euler equation). In the Euler equations, the t-conditional expectation is about $t+1$ values of aggregate shocks

[22]Given the goods market clearing condition implied by the remaining equilibrium conditions and Walras law, aggregate output should equal aggregate consumption. During simulations, I recompute aggregate output fully nonlinearly using the policy functions and distributions implied by the perturbed solution, as $Y_t = z_t K^{1-\alpha}\left[\sum_{i=1}^{N_\theta}\sum_{j=1}^{N_b^f} \theta_i n_t\left(\theta_i, b_j\right)\Phi_t\left(\theta_i, b_j\right)\right]^\alpha$. I check that the deviation from goods market clearing is close to 0.

(next period borrowing constraints and wage influence current decisions), and is taken with respect to their values at t.

$$\Phi_{t+1}\left(\theta_{i'}, b_{j'}\right) - \sum_{i=1}^{N_\theta}\sum_{j=1}^{N_b^f}\Pi_\theta\left(\theta_{i'}|\theta_i\right)\left(\omega_{i,j,j',t}\Phi_t\left(\theta_i, b_j\right)\right) = 0,\ \ i' \in [|1, N_\theta|], j' \in \left[|1, N_b^f|\right] \quad (37)$$

$$\log\overline{\phi}_{t+1} - \log\overline{\phi} - \rho_\phi(\log\overline{\phi}_t - \log\overline{\phi}) - \epsilon^q_{t+1} = 0 \quad (38)$$

$$\log z_{t+1} - \rho_z \log z_t - \epsilon^z_{t+1} = 0 \quad (39)$$

$$\begin{pmatrix}\epsilon^\phi \\ \epsilon^z\end{pmatrix} \overset{iid}{\sim} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix}\sigma^2_\phi & \sigma_\phi\sigma_z\rho_{\phi z} \\ \sigma_\phi\sigma_z\rho_{\phi z} & \sigma^2_z\end{pmatrix}\right) \quad (40)$$

$$B - \sum_{i=1}^{N_\theta}\sum_{j=1}^{N_b^f} b_j\Phi_{t+1}\left(\theta_i, b_j\right) = 0 \quad (41)$$

$$N_t - \sum_{i=1}^{N_\theta}\sum_{j=1}^{N_b^f}\theta_i n_t\left(\theta_i, b_j\right)\Phi_t\left(\theta_i, b_j\right) \quad (42)$$

$$Y_t - \sum_{i=1}^{N_\theta}\sum_{j=1}^{N_b^f} c_t\left(\theta_i, b_j\right)\Phi_t\left(\theta_i, b_j\right) \quad (43)$$

$$C_t - \sum_{i=1}^{N_\theta}\sum_{j=1}^{N_b^f} c_t\left(\theta_i, b_j\right)\Phi_t\left(\theta_i, b_j\right) \quad (44)$$

$$w_t = \alpha z_t\left(\frac{1}{N_t}\right)^{1-\alpha} \quad (45)$$

$$\pi_t = (1-\alpha)z_t K^{1-\alpha}N_t^\alpha \quad (46)$$

$$(1-\tau_1(\theta_i))\, w_t\theta_i c_t\left(\theta_i, b_j\right)^{-\gamma} - \psi n_t\left(\theta_i, b_j\right)^\eta = 0,\ \ i \in [|1, N_\theta|], j \in [|1, N_b|] \quad (47)$$

$$c_t\left(\theta_i, b_j\right)^{-\gamma} - \beta(1+r_t)\mathbb{E}_t\left\{\sum_{i'=1}^{N_\theta} c_{t+1}\left(\theta_{i'}, b'\left(\theta_i, b_j\right)\right)^{-\gamma}\right\} = 0,\ \ i \in [|1, N_\theta|], j \in [|1, N_b|] \quad (48)$$

Denote as $\mathbf{y_t}$ the $6 + N_\theta \times (N_n + N_b + 1)$ vector of current jump (control) variables. Denote as $\mathbf{x_t}$ the $N_\theta \times N_b^f - 1 + 2$ vector of current state (predetermined) variables. Equilibrium conditions are stacked in a multivariate, vector-valued function $\mathcal{F}(.)$ that represents the nonlinear system of equations that defines the equilibrium:

$$\mathbb{E}_t\left[\mathcal{F}\left(\mathbf{y_t}, \mathbf{y_{t+1}}, \mathbf{x_t}, \mathbf{x_{t+1}}, \epsilon^q_{t+1}, \epsilon^z_{t+1}\right)\right] = 0 \quad (49)$$

2. Solving for the deterministic steady state of the economy (without aggregate shocks) amounts to finding $\mathbf{y}, \mathbf{x}$ that solve the following system of equation, which has as many unknowns as equations:

$$\mathcal{F}\left(\mathbf{y}, \mathbf{y}, \mathbf{x}, \mathbf{x}, 0, 0\right) = 0 \quad (50)$$

In theory, it could be solved directly using a nonlinear equation solver. In practice, there is no guarantee that numerical equation solvers will converge when we use projection methods to approximate policy functions. In addition to solving the households' consumption problem, the difficulty comes from having in addition a labor choice, endogenous government taxes, and solving for two prices in equilibrium (the wage and the interest rate). I also solve for the value of the disutility of labor ψ that normalizes steady state output Y to 1. I therefore use the following algorithm to make the problem more stable.

(a) Start with a guess for the bond price and labor demand $\left(p^{(0)}, N^{(0)}, \psi^{(0)}\right)$, for policy function values $\left(\mathbf{b}'^{(0)}(.), \mathbf{n}^{(0)}(.)\right)$, and the cross-sectional distribution $\mathbf{\Phi}^{(0)}(.)$ (it is only needed to compute the first iterate of government taxes). It is easier to solve for the bond price and labor demand demand, and back out the interest rate $1/p - 1$ and the wage (from the firm's optimal labor choice) than solving directly for the latter. Thus having $\left(p^{(0)}, N^{(0)}\right)$ is equivalent to having $\left(r^{(0)}, w^{(0)}\right)$.

(b) Given those, use the endogenous grid method of Carroll (2006) to iterate backwards on the household's optimality conditions (the Euler and the labor intratemporal equations), and obtain a new guess for policy functions that will be supplied to the nonlinear policy solver solving the household's problem, $\left(\mathbf{b}'^{(1)}(.), \mathbf{n}^{(1)}(.)\right)$[23]. This requires computing endogenous government taxes (fixed every period because we are at the steady state), which is why we need a guess for the cross-sectional distribution.

(c) The guess for prices is supplied to a second nonlinear solver wrapped around the policy solver, which solves for the prices clearing the bond and the labor market, and for the disutility of labor normalizing steady state output to 1. Within the price solver, I ensure that prices and labor disutility are positive ($p^{(n)}, N^{(n)}, \psi^{(n)} > 0$), and the stability condition $\beta / p^{(n)} \leq 1$ holds at every iteration n. The following steps occur within the price solver, and their iterates start at $n = 1$.

(d) Given the exogenous law of motion for idiosyncratic income and the policy functions, compute the associated stationary distribution of households $\mathbf{\Phi}^{(1)}(.)$ (I use the eigenvector method). Also compute the wage and profits from the firm's optimality condition: $w^{(1)} = \alpha \left(\frac{1}{N^{(1)}}\right)^{1-\alpha}$, and $\pi^{(1)} = (1-\alpha)\left(N^{(1)}\right)^{\alpha}$.

[23] This step is used by McKay and Reis (2016).

Then, given prices, policy functions and the distribution, compute endogenous government taxes $\tau_0^{(1)}$.

(e) Given prices, profits, taxes, and savings policies $\mathbf{b}'^{(1)}(.)$, solve the household's labor supply equation (using $\mathbf{n}^{(1)}(.)$ as a guess), and denote $\mathbf{n}^{(2)}(.)$ the new labor supply policy. It should always be non-negative. Here I use a nonlinear equation solver with Broyden's method, and supplies it with the Jacobian of the system of intratemporal equations. Here and later, derivatives are computed exactly with automatic differentiation, implemented with Julia's `ForwardDiff` package (Revels et al. (2016)).

(f) Back out the associated consumption function from the budget constraint. If it has a non-positive entry at a point in the state space, adjust $\mathbf{n}^{(2)}(.)$ at that point such that the household consumes $c_{min} = 0.001$. This step helps with convergence of the solver when solving for savings in the next step.

(g) Given prices, profits, taxes and the new labor policy $\mathbf{n}^{(2)}(.)$, solve the household's Euler equation (using $\mathbf{b}'^{(1)}(.)$ as a guess), and denote $\mathbf{b}'^{(2)}(.)$ the new savings policy. Use the same solver as for labor.

(h) This completes one iterate in the loop solving for policy functions given prices. If the new policy functions $\left(\mathbf{n}^{(2)}(.)\, \mathbf{b}'^{(2)}(.)\right)$ are close enough to the previous ones $\left(\mathbf{n}^{(2)}(.)\, \mathbf{b}'^{(2)}(.)\right)$, then stop and we have solved the household's problem given prices $\left(p^{(0)}, N^{(0)}, \psi^{(0)}\right)$. Otherwise, iterate on steps (d)-(g). That is, given $\left(p^{(0)}, N^{(0)}, \psi^{(0)}\right)$ (hence the same wages and profits), compute new government taxes $\tau_0^{(n+1)}$. Then solve for new policy functions $\left(\mathbf{n}^{(n+1)}(.)\, \mathbf{b}'^{(n+1)}(.)\right)$, compare them to the previous ones $\left(\mathbf{n}^{(n)}(.)\, \mathbf{b}'^{(n)}(.)\right)$, and stop when they are close enough. This completes the solution of the household's problem given prices.

(i) Using the law of motion of the exogenous income shock and the optimal savings function, compute the stationary distribution $\mathbf{\Phi}^{(2)}$. Use it with policy functions to compute aggregate values for savings, labor supply and output. The price solver then chooses new values for prices and disutility of labor, $\left(p^{(1)}, N^{(1)}, \psi^{(1)}\right)$, to solve the following three equations:

$$B - \sum_{i=1}^{N_\theta} \sum_{j=1}^{N_b^f} b_j \Phi_{t+1}^{(2)}\left(\theta_i, b_j\right) = 0$$

$$N^{(1)} - \sum_{i=1}^{N_\theta} \sum_{j=1}^{N_b^f} \theta_i n\left(\theta_i, b_j\right) \Phi^{(2)}\left(\theta_i, b_j\right)$$

$$Y^{(1)} - 1 = 0 \Leftrightarrow \left(N^{(1)}\right)^{\alpha} - 1 = 0$$

(j) Then go back to step (a) with the new prices, and iterate until convergence, i.e. policy functions and the stationary distribution have converged, and the three equations are satisfied. We then obtain prices, policy functions and a distribution that solve the model in the deterministic steady state.

3. Do a first- and a second-order perturbation of the discrete model around its steady state. The solutions to the equilibrium expectational difference equation $\mathbb{E}_t\left[\mathcal{F}\left(.\right)\right] = 0$ are of the following form (Schmitt-Grohe and Uribe (2008)):

$$\mathbf{x_{t+1}} = \mathbf{h}\left(\mathbf{x_t}, \eta\right) + \eta \begin{pmatrix} \mathbf{0} \\ \epsilon^q_{t+1} \\ \epsilon^z_{t+1} \end{pmatrix} \quad (51)$$

$$\mathbf{y_t} = \mathbf{g}\left(\mathbf{x_t}, \eta\right) \quad (52)$$

where η is the perturbation parameter (there is only one such parameter) scaling the amount of aggregate uncertainty in the economy. The goal is to solve for approximations of the functions $\mathbf{h}, \mathbf{g}$.

(a) For the first-order approximation of the model, several methods can be used. I check existence and uniqueness, and verify that I obtain identical results using Sims' `gensys` (Sims (2001)) and Klein's methods (Klein (2000)), commonly used in the macro literature. I briefly describe the input and the output of Klein's method because it has a clear interpretation in terms of jump and predetermined variables. We solve for a first-order approximation of $\mathbf{g}, \mathbf{h}$. Writing variables in deviations from their steady state values (denoted as $\widehat{x}, \widehat{y}$) and linearizing equilibrium conditions around 0 (where variables equal their steady state values), we obtain

$$\mathcal{F}_{\mathbf{y_t}}\widehat{\mathbf{y_t}} + \mathcal{F}_{\mathbf{y_{t+1}}}\mathbb{E}_t\left[\widehat{\mathbf{y_{t+1}}}\right] + \mathcal{F}_{\mathbf{x_t}}\widehat{\mathbf{x_t}} + \mathcal{F}_{\mathbf{x_{t+1}}}\mathbb{E}_t\left[\widehat{\mathbf{x_{t+1}}}\right] + \mathcal{F}_{\epsilon^q_{t+1}}\mathbb{E}_t\left[\widehat{\epsilon_{t+1}}^q\right] + \mathcal{F}_{\epsilon^z_{t+1}}\mathbb{E}_t\left[\widehat{\epsilon_{t+1}}^z\right] = 0 \quad (53)$$

where the derivatives of $\mathcal{F}$ are evaluated at the steady state. They are submatrices of the Jacobian of $\mathcal{F}$, computed exactly with automatic differentiation. $\widehat{\mathbf{y}}, \widehat{\mathbf{x}}$ terms are vectors, so their (matrix) products with the derivative matrices of $\mathcal{F}$ are vectors. The Jacobian is a matrix of dimension

$$\begin{aligned} & \left\{\left[N_\theta \times N_b^f - 1 + 2\right] + \left[6 + N_\theta \times (N_n + N_b + 1)\right]\right\} \\ \times \quad & \left\{2 \times \left[N_\theta \times N_b^f - 1 + 2\right] + 2 \times \left[6 + N_\theta \times (N_n + N_b + 1)\right] + 2\right\} \end{aligned} \tag{54}$$

First-order approximations of the solution have the following form:

$$\widehat{\mathbf{x_{t+1}}} = \mathbf{h_x}\left(\mathbf{x}, 0\right)\widehat{\mathbf{x_t}} + \eta \begin{pmatrix} \mathbf{0} \\ \epsilon_{t+1}^q \\ \epsilon_{t+1}^z \end{pmatrix} \tag{55}$$

$$\widehat{\mathbf{y_t}} = \mathbf{g_x}\left(\mathbf{x}, 0\right)\widehat{\mathbf{x_t}} \tag{56}$$

(b) For the second-order approximation of the model, I do a second-order approximation of equilibrium conditions around the steady state. It involves the Hessian of $\mathcal{F}$, a large three-dimensional array computed by automatic differentiation, of dimension:

$$\begin{aligned} & \left\{\left[N_\theta \times N_b^f - 1 + 2\right] + \left[6 + N_\theta \times (N_n + N_b + 1)\right]\right\} \\ \times \quad & \left\{2 \times \left[N_\theta \times N_b^f - 1 + 2\right] + 2 \times \left[6 + N_\theta \times (N_n + N_b + 1)\right] + 2\right\}^2 \end{aligned} \tag{57}$$

The second-order approximation of the solution has the form:

$$\widehat{\mathbf{x_{t+1}}} = \mathbf{h_x}\left(\mathbf{x}, 0\right)\widehat{\mathbf{x_t}} + \frac{1}{2}\mathbf{h_{xx}}\left(\mathbf{x}, 0\right)\widehat{\mathbf{x_t}}^2 + \frac{1}{2}\mathbf{h}_{\eta\eta}\left(\mathbf{x}, 0\right)\eta^2 + \eta \begin{pmatrix} \mathbf{0} \\ \epsilon_{t+1}^q \\ \epsilon_{t+1}^z \end{pmatrix} \tag{58}$$

$$\widehat{\mathbf{y_{t+1}}} = \mathbf{g_x}\left(\mathbf{x}, 0\right)\widehat{\mathbf{x_t}} + \frac{1}{2}\mathbf{g_{xx}}\left(\mathbf{x}, 0\right)\widehat{\mathbf{x_t}}^2 + \frac{1}{2}\mathbf{g}_{\eta\eta}\left(\mathbf{x}, 0\right)\eta^2 \tag{59}$$

where the terms equal to zero (in $h_\eta, g_\eta, h_{x\eta}, h_{\eta x}, g_{x\eta}, g_{\eta x}$) were canceled. $\widehat{\mathbf{x}}, \widehat{\mathbf{y}}$ terms are vectors, $\mathbf{g_x}, \mathbf{h_x}$ terms are matrices, $\mathbf{h_{xx}}, \mathbf{g_{xx}}$ are 3-dimensional arrays, and $\mathbf{h}_{\eta\eta}, \mathbf{g}_{\eta\eta}$ are vectors. Thus products of $\widehat{\mathbf{x}}, \widehat{\mathbf{y}}$ vectors with first-order derivative matrices are matrix products, those with second-order arrays are tensor products, and those with η are simple constant times vectors products. I use Kim et al. (2008)'s `gensys2` method to solve for the unknown coefficients. Schmitt-Grohe and Uribe (2008) propose instead to solve for the second-order coefficients in a linear system of equations involving the Jacobian and the Hessian of $\mathcal{F}$, and the first-order coefficients. While most papers with representative agent models use this method, it is not tractable in a setting with heterogeneous agents where the cross-sectional distribution is discretized as a histogram, since it involves constructing and inverting a matrix whose dimensions blow up with the number of state variables[24]. `gensys2` allows to reduce the dimensionality of the system of equation to solve by applying a sequence of linear operations to the original system (Schur and singular value decompositions).

A.2 Computing the Steady State in the Second-Order Approximation of the Model

Pruning essentially computes first-order projections of second-order terms, based on a first-order expansion of the conditional expectation of the system's deviation from steady state, according to the following steps.

First, `gensys2` solves a linearly transformed system, where original variables $\begin{pmatrix}\widehat{\mathbf{x}} & \widehat{\mathbf{y}}\end{pmatrix}'$ that solve $\mathbb{E}_t\left[\mathcal{F}\left(.\right)\right] = 0$ are replaced by $Z'\begin{pmatrix}\widehat{\mathbf{x}} & \widehat{\mathbf{y}}\end{pmatrix}'$, where Z is a square, non-singular matrix. To simplify notation, denote the transformed variables as $\begin{pmatrix}\widehat{\mathbf{x}} & \widehat{\mathbf{y}}\end{pmatrix}'$ too. The second-order solution to the transformed system has the form (see the paper for details):

$$\begin{aligned}
\widehat{\mathbf{x_{t+1}}} &= F_1\widehat{\mathbf{x_t}} + F_2\eta\mathbf{ffl_{t+1}} + F_3\eta^2 + \frac{1}{2}F_{11}\widehat{\mathbf{x_t}}^2 + F_{12}\widehat{\mathbf{x_t}}\mathbf{ffl_{t+1}}\eta + \frac{1}{2}F_{22}\eta^2\mathbf{ffl_{t+1}}^2 && (60)\\
\widehat{\mathbf{y_t}} &= \frac{1}{2}M_{11}\widehat{\mathbf{x_t}}^2 + M_2\eta^2 && (61)
\end{aligned}$$

The presence of cross-derivative terms in the transformed solution does not contradict their absence in the original solution, since they can be canceled by Z. Then, it implies that for $s > 0$:

$$\begin{aligned}
\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s}}}\right] &= F_1\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s-1}}}\right] + F_3\eta^2 + \frac{1}{2}F_{11}\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s-1}}^2}\right] + \frac{1}{2}F_{22}\eta^2\Omega_s \\
&= F_1\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s-1}}}\right] + F_3\eta^2 + \frac{1}{2}F_{11}\left(\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s-1}}}\right]^2 + \Sigma_{s-1}\right) + \frac{1}{2}F_{22}\eta^2\Omega_s && (62)\\
\mathbb{E}_t\left[\widehat{\mathbf{y_{t+s}}}\right] &= \frac{1}{2}M_{11}\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s}}^2}\right] + M_2\eta^2 \\
&= \frac{1}{2}M_{11}\left(\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s}}}\right]^2 + \Sigma_s\right) + M_2\eta^2 && (63)\\
\Sigma_{s+1} &= \eta^2F_2\Omega_tF_2 + F_1\Sigma_sF_1 && (64)
\end{aligned}$$

[24]Despite the Jacobian and Hessian of $\mathcal{F}$ being typically sparse matrices. A potential way around would be to use state space reduction techniques, as in Ahn et al. (2017).

where Ω_s is the t-conditional variance-covariance matrix of $\mathbf{ffl_{t+s}}$, and Σ_s is the t-conditional variance-covariance matrix of $\widehat{\mathbf{x_{t+s}}}$, defined recursively by a discrete Lyapunov equation (from the law of motion of $\widehat{\mathbf{x_{t+1}}}$).
Then, projecting $\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s-1}}}\right]$ terms on their first-order counterparts, denoted $\mathbb{E}_t^1\left[\widehat{\mathbf{x_{t+s-1}}}\right]$, we obtained the pruned law of motion of the transformed solution:

$$\begin{aligned}
\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s}}}\right] &= F_1\mathbb{E}_t\left[\widehat{\mathbf{x_{t+s-1}}}\right] + F_3\eta^2 + \frac{1}{2}F_{11}\left(\mathbb{E}_t^1\left[\widehat{\mathbf{x_{t+s-1}}}\right]^2 + \Sigma_{s-1}\right) + \frac{1}{2}F_{22}\eta^2\Omega_s && (65)\\
\mathbb{E}_t\left[\widehat{\mathbf{y_{t+s}}}\right] &= \frac{1}{2}M_{11}\left(\mathbb{E}_t^1\left[\widehat{\mathbf{x_{t+s}}}\right]^2 + \Sigma_s\right) + M_2\eta^2 && (66)\\
\mathbb{E}_t^1\left[\widehat{\mathbf{x_{t+s}}}\right] &= F_1\mathbb{E}_t^1\left[\widehat{\mathbf{x_{t+s-1}}}\right] && (67)\\
\Sigma_{s+1} &= \eta^2 F_2\Omega_t F_2 + F_1\Sigma_s F_1 && (68)
\end{aligned}$$

To compute the steady state of the second-order solution to the original system, we first compute the steady state of the transformed system using its laws of motion. In particular, we solve for the steady state value of expected deviations of transformed variables from their steady state (set $\eta = 1$):

$$\begin{aligned}
\mathbb{E}\left[\widehat{\mathbf{x}}\right] &= (I - F_1)^{-1}\left(F_3 + \frac{1}{2}F_{22}\Omega + \frac{1}{2}F_{11}\Sigma\right) && (69)\\
\mathbb{E}\left[\widehat{\mathbf{y}}\right] &= \frac{1}{2}M_{11}\Sigma + M_2 && (70)\\
\text{where}\quad \Sigma &= F_2\Omega_t F_2 + F_1\Sigma F_1 && (71)
\end{aligned}$$

Finally, we back out the steady state values of original variables as $Z'^{-1}\begin{pmatrix}\mathbb{E}\left[\widehat{\mathbf{x}}\right] & \mathbb{E}\left[\widehat{\mathbf{y}}\right]\end{pmatrix}'$.

A.3 Computing IRFs in the Second-Order Approximation of the Model

To compute the economy's impulse response functions, we use the pruned version of the law of motion for transformed variables (for $\eta = 1$), for $t \geq 0$:

$$\begin{aligned}
\widehat{\mathbf{x_{t+1}}} &= F_1\widehat{\mathbf{x_t}} + F_2\mathbf{ffl_{t+1}} + F_3 + \frac{1}{2}F_{11}\widehat{\mathbf{x_t^1}}^2 + F_{12}\widehat{\mathbf{x_t^1}}\mathbf{ffl_{t+1}} + \frac{1}{2}F_{22}\mathbf{ffl_{t+1}}^2 && (72)\\
\widehat{\mathbf{y_t}} &= \frac{1}{2}M_{11}\widehat{\mathbf{x_t^1}}^2 + M_2 && (73)\\
\widehat{\mathbf{x_{t+1}^1}} &= F_1\widehat{\mathbf{x_t^1}} + F_2\mathbf{ffl_{t+1}} && (74)
\end{aligned}$$

We then back out the path of original variables as $\left\{Z'^{-1}\begin{pmatrix}\widehat{\mathbf{x_t}} & \widehat{\mathbf{y_t}}\end{pmatrix}'\right\}_t$.

A.4 Variance Decomposition

A.4.1 First-Order

The vector $Y = \begin{pmatrix} x & y \end{pmatrix}$ of equilibrium objects contains the predetermined and the jump variables. It is in deviation from steady state, but it doesn't matter for this exercise because we can just add the steady state vector, which will cancel out when taking variances. The output from gensys is a law of motion for Y, consisting of an AR(1) matrix Φ and an impact matrix Z:

$$(I - \Phi L)\, Y_{t+1} = Z\epsilon_{t+1} \tag{75}$$

where $\epsilon_{t+1} = \begin{pmatrix} \epsilon^{\phi}_{t+1} & \epsilon^{z}_{t+1} \end{pmatrix}'$ is the vector of the two shocks, with covariance matrix $\tilde{\Sigma}_{\epsilon} = \begin{pmatrix} 1 & \rho_{\phi,z} \\ \rho_{\phi,z} & 1 \end{pmatrix}$, and where the rows of Z corresponding to ϵ^{ϕ}_{t+1} and ϵ^{z}_{t+1} are $\begin{pmatrix} \sigma_{\phi} & 0 \\ 0 & \sigma_z \end{pmatrix}$.
Thus $\text{Var}\left(\begin{pmatrix} \sigma_{\phi} & 0 \\ 0 & \sigma_z \end{pmatrix} \tilde{\Sigma}_{\epsilon}\right) = \begin{pmatrix} \sigma^2_{\phi} & \rho_{\phi,z}\sigma_{\phi}\sigma_z \\ \rho_{\phi,z}\sigma_{\phi}\sigma_z & \sigma^2_z \end{pmatrix} = \Sigma_{\epsilon}$.
First, we transform the shocks with covariance $\tilde{\Sigma_{\epsilon}}$ so that they are orthogonal, i.e. their covariance matrix is the identity matrix. This is done by Cholesky factorization. The new orthogonal shocks are defined as $\nu_t = Q\epsilon_t$, with Q such that $\mathbb{E}\left[\nu_t \nu_t'\right] = I$. Denoting $S = Q^{-1}$, $\epsilon_t = S\nu_t$ and $SS' = \tilde{\Sigma_{\epsilon}}$. S is a lower triangular matrix given by the Cholesky factorization of $\tilde{\Sigma}_{\epsilon}$.
Then, we transform the economy's law of motion from an AR(1) to an MA(∞) representation, using the fact that the eigenvalues of Φ are within the unit circle (we denote L the lag operator). We also substitute for $\epsilon_{t+1} = S\nu_{t+1}$.

$$\begin{aligned}
(I - \Phi L)\, Y_{t+1} &= Z\epsilon_{t+1} \\
\Rightarrow Y_{t+1} &= (I - \Phi L)^{-1} ZS\nu_{t+1} \\
Y_{t+1} &= \sum_{k=0}^{\infty} \Phi^k L^k ZS\nu_{t+1} \\
Y_{t+1} &= \sum_{k=0}^{\infty} \Phi^k ZS\nu_{t+1-k} \\
\Rightarrow Y_{t+h} &= \sum_{k=0}^{\infty} \tilde{\Phi}^{(k)} \nu_{t+h-k}
\end{aligned} \tag{76}$$

for any forecasting horizon $h > 0$, and where $\tilde{\Phi}^{(k)} = \Phi^k ZS$ is a matrix of dimension (number of variables,number of shocks). Here we consider N variables and 2 shocks.

Then, forecast errors at horizon $h > 0$ are:

$$
\begin{aligned}
e_{t+h} &= Y_{t+h} - \mathbb{E}_t\left[Y_{t+h}\right] && (77)\\
&= \tilde{\Phi}^{(0)}\nu_{t+h} + \tilde{\Phi}^{(1)}\nu_{t+h-1} + \tilde{\Phi}^{(2)}\nu_{t+h-2} + \ldots + \tilde{\Phi}^{(h-1)} && (78)\\
&= \sum_{i=1}^{h} \tilde{\Phi}^{(h-i)}\nu_{t+i} && (79)
\end{aligned}
$$

For variable $Y_j, j \in \{1, \ldots N\}$,

$$
\begin{aligned}
e_{j,t+h} &= \sum_{i=1}^{h} \tilde{\Phi}^{(h-i)}_{j,.}\nu_{t+i} && (80)\\
&= \sum_{i=1}^{h} \left(\tilde{\Phi}^{(h-i)}_{j,1}\nu_{1,t+i} + \tilde{\Phi}^{(h-i)}_{j,2}\nu_{2,t+i}\right) && (81)
\end{aligned}
$$

So the total forecast error variance at horizon $h > 0$ for variable Y_j is, using the fact that ν's are mutually independent, identically distributed and serially uncorrelated:

$$
\Rightarrow \text{Var}\left(e_{j,t+h}\right) = \sum_{i=1}^{h} \left(\left(\tilde{\Phi}^{(h-i)}_{j,1}\right)^2 + \left(\tilde{\Phi}^{(h-i)}_{j,2}\right)^2\right) \tag{82}
$$

Finally, the share of the forecast error variance of variable Y_j at horizon $h > 0$ accounted for by ν^1 and ν^2 (transformed versions of the original shocks ϵ^{ψ} and ϵ^{z}) are respectively:

$$
\frac{\sum_{i=1}^{h}\left(\tilde{\Phi}^{(h-i)}_{j,1}\right)^2}{\sum_{i=1}^{h}\left(\left(\tilde{\Phi}^{(h-i)}_{j,1}\right)^2 + \left(\tilde{\Phi}^{(h-i)}_{j,2}\right)^2\right)} \quad \text{and} \quad \frac{\sum_{i=1}^{h}\left(\tilde{\Phi}^{(h-i)}_{j,2}\right)^2}{\sum_{i=1}^{h}\left(\left(\tilde{\Phi}^{(h-i)}_{j,1}\right)^2 + \left(\tilde{\Phi}^{(h-i)}_{j,2}\right)^2\right)} \tag{83}
$$

Results are sensitive to whether the matrix obtained from the Cholesky factorization is lower or upper triangular. A lower triangular S implies that ν^2 has no effect on nu^1. Note that because of the factorization, the ν shocks are not clearly interpretable as credit and TFP shocks. To have that interpretation for the results in the main text, I resolve the model with $\rho_{\phi,z} = 0$, and omit the factorization step. Results are displayed in Table 6.

A.4.2 Second-Order

I use the generalized forecast error variance decomposition for nonlinear models described by Lanne and Nyberg (2016). The starting point is the nonlinear (quadratic) model

Variable:	Credit shock	TFP shock
$1/(1+r)$	51%	49%
N	56%	44%
w	6%	94%
π	70%	30%
Y	70%	30%

Table 6: Variance decomposition, order 1: share of variance in left handside variables accounted for by credit and TFP shocks. Variance computed as the variance of forecast errors at a large forecasting horizon ($H = 1000$). See appendix for computations.

given by `gensys2`, which can be written as

$$Y_{t+1} = f\left(Y_t, \epsilon_{t+1}\right) \tag{84}$$

where G is a nonlinear function of the equilibrium vector and of innovations. As above, the interpretation of shocks is clearer when $\rho_{\phi,z} = 0$.
The generalized impulse-response function (GIRF) at horizon $i > 0$ (i.e. at date $t+i$) of variable Y_j, with respect to a credit shock (or TFP shock) of magnitude $\delta_{\phi,t+1}$ (or $\delta_{z,t+1}$) hitting at date $t+1$, conditional on history of states $\omega_t = y_t$, is defined as:

$$\begin{aligned} GI_j\left(i, \delta_{\phi,t+1}, \omega_t\right) &= \mathbb{E}_t\left[Y_{j,t+i} | \epsilon^{\phi}_{t+1} = \delta_{\phi,t+1}, \omega_t\right] - \mathbb{E}_t\left[Y_{j,t+i} | \omega_t\right] &(85)\\ \text{and } \; GI_j\left(i, \delta_{z,t+1}, \omega_t\right) &= \mathbb{E}_t\left[Y_{j,t+i} | \epsilon^{z}_{t+1} = \delta_{z,t+1}, \omega_t\right] - \mathbb{E}_t\left[Y_{j,t+i} | \omega_t\right] &(86)\end{aligned}$$

Then, the generalized forecast error variance decomposition (GFEVD) of variable Y_j at horizon $h > 0$, is between the fraction of variance explained by credit shocks, and that explained by TFP shocks, respectively:

$$\begin{aligned} GFEVD_j(h, \delta_{\phi,t}) &= \frac{\sum_{i=0}^{h} GI_j\left(i, \delta_{\phi,t+1}, \omega_t\right)^2}{\sum_{i=0}^{h} GI_j\left(i, \delta_{\phi,t+1}, \omega_t\right)^2 + \sum_{i=0}^{h} GI_j\left(i, \delta_{z,t+1}, \omega_t\right)^2} &(87)\\ GFEVD_j(h, \delta_{z,t}) &= \frac{\sum_{i=0}^{h} GI_j\left(i, \delta_{z,t+1}, \omega_t\right)^2}{\sum_{i=0}^{h} GI_j\left(i, \delta_{\phi,t+1}, \omega_t\right)^2 + \sum_{i=0}^{h} GI_j\left(i, \delta_{z,t+1}, \omega_t\right)^2} &(88)\end{aligned}$$

Because GIRF are nonlinear, GFEVD depend on the sign and size of the innovations δ. I therefore compute average GFEVD using bootstrap. First, because the solution of the model is based on perturbations around the steady state, we can get rid of the history dependence in ω. Then, I simulate a history of credit and TFP innovations of length

$T = 1000$, $\left\{\epsilon_t^{\phi}, \epsilon_t^{z}\right\}_{t=0}^{T} = \left\{\delta_{\phi,t}, \delta_{z,t}\right\}_{t=0}^{T}$ using $\begin{pmatrix} \epsilon_t^{\phi} \\ \epsilon^{z} \end{pmatrix} \overset{iid}{\sim} \mathcal{N}(0, I_2)$ (with `gensys2` the innovation variances σ_{ϕ}^2 and σ_z^2 are incorporated in the GIRF matrices). For each innovation $\delta_{\phi,t}$, I compute the associated $GFEVD_j(h, \delta_{\phi,t})$ for variable Y_j at horizon h. Finally, the average GFEVD is obtained by averaging over individual $GFEVD_j(h, \delta_{\phi,t})$'s by using the probability associated to each $\delta_{\phi,t}$ by the standard normal p.d.f. (Because $\mathcal{N}(0,1)$ is symmetric, we should get something like an average of the GFEVD for a shock $\delta = -1$ and a shock $\delta = +1$.) Computations are parallelized over the N dimension. It takes about 17 hours to run the case $N = 500, H = 1000$ using 28 cores.

A.5 Recovering Structural Shocks

A.5.1 Kalman Filter (order 1)

A linear state space representation of the model is obtained from `gensys`. Using the above notation, the transition and the measurement equations are respectively:

$$\begin{aligned} Y_{t+1} &= \Phi Y_t + Z\epsilon_{t+1}, \;\; \epsilon_{t+1} \overset{iid}{\sim} \mathcal{N}(0, Q) \\ Y_{t+1}^{obs} &= H' Y_{t+1} + v_t, \;\; v_{t+1} \overset{iid}{\sim} \mathcal{N}(0, R) \end{aligned}$$

Φ and Z are readily obtained from `gensys` and $Q = I_2$ (variance-covariance terms are in Z by design). H is a selection matrix filled everywhere with zeros, and with ones for the entries corresponding to the observable variables in Y_{t+1} (bond price and consumption). There is no noise in the measurement equation, i.e. $R = 0_{2\times 2}$: the bond price and consumption are perfectly observed.

Using standard notation, denote $Y_{t|t-1} = \mathbb{E}\left[Y_t | Y^{obs,t-1}\right]$ (best linear predictor of Y_t given the history of observables Y^{obs} until $t-1$), $Y_{t|t-1}^{obs} = \mathbb{E}\left[Y_t^{obs} | Y^{obs,t-1}\right]$, and $Y_{t|t} = \mathbb{E}\left[Y_t | Y^{obs,t}\right]$. Also denote $\Sigma_{t|t-1} = \mathbb{E}\left[\left(Y_t - Y_{t|t-1}\right)\left(Y_t - Y_{t|t-1}\right)' | Y^{obs,t-1}\right]$ (predicting error variance-covariance matrix of Y_t given the history of observables until $t-1$), $\Omega_{t|t-1} = \mathbb{E}\left[\left(Y_t^{obs} - Y_{t|t-1}^{obs}\right)\left(Y_t^{obs} - \right.\right.$

and $\Sigma_{t|t} = \mathbb{E}\left[\left(Y_t - Y_{t|t}\right)\left(Y_t - Y_{t|t}\right)' | Y^{obs,t}\right]$.

The goal of the Kalman filter here is to back out the sequences of forecasted observable variables and underlying states $\left\{Y_{t|t-1}^{obs}, Y_{t|t}\right\}$ implied by the model, given a sequence of observable variables $\left\{Y_t^{obs}\right\}$ taken from the data. The algorithm proceeds as follows:

1. At $t = 1$, initial conditions $Y_{1|0}, \Sigma_{1|0}$ are set equal to their (deterministic) steady state values. That is, $Y_{1|0} = 0$ (the initial system of equations was written in log de-

viations from steady state), and $\Sigma_{1|0}$ is the solution to the Riccati equation $\Sigma_{1|0} = \Phi\Sigma_{1|0}\Phi' + ZI_2Z'$, which is solved by iterating on a symmetric, positive definite guess $\Sigma_{1|0}^{(0)}$ (using the stability of the system). I verify that the solution $\Sigma_{1|0}^{(\infty)} = \Sigma_{1|0}$ is symmetric and positive definite too. Following steps are for $t \geq 1$.

2. Given $\Sigma_{t|t-1}, Y_t^{obs}, Y_{t|t-1}^{obs}$, compute $\Omega_{t|t-1} = H'\Sigma_{t|t-1}H + R = H'\Sigma_{t|t-1}H$.

3. Compute $\text{Cov}_{t-1}\left(Y_t^{obs}, Y_t\right) = \mathbb{E}\left[\left(Y_t^{obs} - Y_{t|t-1}^{obs}\right)\left(Y_t - Y_{t|t-1}\right)' | Y^{obs,t-1}\right] = H'\Sigma_{t|t-1}$.

4. Compute the Kalman gain $K_t = \Sigma_{t|t-1}H\left(H'\Sigma_{t|t-1}H + R\right)^{-1} = \Sigma_{t|t-1}H\Omega_{t|t-1}^{-1}$.

5. Compute $Y_{t|t} = Y_{t|t-1} + K_t\left(Y_t^{obs} - H'Y_{t|t-1}\right)$ ("nowcast" of the state).

6. Compute $\Sigma_{t|t} = \Sigma_{t|t-1} - K_tH'\Sigma_{t|t-1}$ (variance-covariance matrix associated with the "nowcast" error).

7. Compute $\Sigma_{t+1|t} = \Phi\Sigma_{t|t}\Phi' + ZQZ' = \Phi\Sigma_{t|t}\Phi' + ZZ'$ (next period forecast error variance-covariance matrix).

8. Finally, compute $Y_{t+1|t} = \Phi Y_{t|t}$ and $Y_{t+1|t}^{obs} = H'Y_{t+1|t}$ (next period implied state, and next period forecasted observables).

A.5.2 Particle Filter (order 2)

A nonlinear state space representation of the model is obtained from `gensys2`. Using the above notation, the transition and the measurement equations are respectively:

$$\begin{aligned} Y_{t+1} &= f\left(Y_t, W_{t+1}\right), \;\; W_{t+1} \overset{iid}{\sim} \mathcal{N}(0, Q) \\ Y_{t+1}^{obs} &= H'Y_{t+1} + v_t, \;\; v_{t+1} \overset{iid}{\sim} \mathcal{N}(0, R) \end{aligned}$$

f is the quadratic mapping (from `gensys2`) used to compute impulse responses in the second-order solution of the model (see above). $Q = I_2$ (variance-covariance terms are in the matrices part of f by design), and H is a selection matrix filled everywhere with zeros, and with ones for the entries corresponding to the observable variables in Y_{t+1} (bond price and consumption). I assume that there is some but very little noise in the measurement equation, i.e. $R = 10^{-6} \times I_2$: the bond price and consumption are close to perfectly observed. This is because the joint density of measurement errors is needed in the algorithm, so R cannot be zero.

Particles are i.i.d. draws $\{Y_{t-1}^i, W_{t-1}^i\}_{i=1}^N$ from the joint density $p\left(W_{t-1}, Y_{t-1}|Y_{t-1}^{obs}\right)$. Proposed particles are i.i.d. draws $\left\{Y_{t|t-1}^i, W_{t|t-1}^i\right\}_{i=1}^N$ from the joint density $p\left(W_t, Y_{t-1}|Y_{t-1}^{obs}\right)$. There are N of each of them. Here, the structural innovations W are independent of the vector of predetermined and jump variables Y. Therefore, drawing from the proposed joint density boils down to drawing from the innovations' density, and then applying the nonlinear mapping f to the previous proposed Y and the new innovations w, to get the new proposed particle Y. As before, the sequence of observable variables $\{Y_t^{obs}\}_{t=0}^T$ is taken from the data, with $Y_0^{obs} = 0$. That is, I assume w.l.o.g. that the beginning of the sample represents the deterministic steady state (hence log-deviations are zero). The algorithm proceeds as follows.

1. At $t = 1$, set the initial condition $Y_{0|0}^i = Y_0^i = W_0^i = 0$ for all $i = 1, ...N$, i.e. the log-deviation from the deterministic steady state is assumed to be zero at $t = 0$.

2. Generate N i.i.d. draws of proposed particles $\left\{Y_{t|t-1}^i, W_{t|t-1}^i\right\}_{i=1}^N$ from $p\left(W_t, Y_{t-1}|Y_{t-1}^{obs}\right)$. That is, draw $w_{t|t-1}^i$ innovations from $\mathcal{N}(0, I_2)$ and obtain the associated $Y_{t|t-1}^i$ from f.

3. Evaluate the conditional density $p\left(Y_t^{obs}|w_{t|t-1}^i, Y_{t-1}^{obs}, Y_{t|t-1}^i\right)$ using the measurement equation and the distribution of measurement errors v. That is,

$$p\left(Y_t^{obs}|w_{t|t-1}^i, Y_{t-1}^{obs}, Y_{t|t-1}^i\right) = \phi\left(Y_t^{obs} - H'Y_t|w_{t|t-1}^i, Y_{t-1}^{obs}, Y_{t|t-1}^i\right)$$

   where ϕ is the (conditional) density of the multivariate standard normal distribution.

4. Evaluate the relative weights $q_t^i = \frac{p\left(Y_t^{obs}|w_{t|t-1}^i, Y_{t-1}^{obs}, Y_{t|t-1}^i\right)}{\sum_{j=1}^N p\left(Y_t^{obs}|w_{t|t-1}^j, Y_{t-1}^{obs}, Y_{t|t-1}^j\right)}$, normalized to be probabilities.

5. Re-sample, with replacement, N values $\left\{Y_{t|t-1}^i, W_{t|t-1}^i\right\}_{i=1}^N$ from the sample we had so far, now using the $\{q_t^i\}_{i=1}^N$ as probabilities. These new values are the particles, denoted $\{Y_t^i, W_t^i\}_{i=1}^N$.

6. Go back to step 2 for $t+1$, generate new innovations and use the new swarm of particles $\{Y_t^i, W_t^i\}_{i=1}^N$ to generate a new swarm of proposed particles $\left\{Y_{t+1|t}^i, W_{t+1|t}^i\right\}_{i=1}^N$. Then iterate until reaching the end of the sample $t = T$.

Thus we obtain a sequence of swarms of particles $\left\{\left\{Y_t^i, W_t^i\right\}_{i=1}^N\right\}_{t=0}^T$, which represent empirical conditional densities at every point in time for the state Y, which are implied by the model, given the sequence of observables $\left\{Y_t^{obs}\right\}_{t=0}^T$ from the data. In the main text, I plot the sample averages of these empirical conditional densities at $t = 0, ...T$. This paper is to my knowledge the first paper to apply nonlinear filtering to the perturbation-based solution of a heterogeneous agents model with aggregate shocks. Computations are parallelized over the N dimension. It takes about 12 hours to run the case $N = 20,000, T = 44$ using 28 cores.

B Calibration

B.1 Numerical Parameters

Parameter	Explanation	Value
N_θ	Nb. idiosyncratic income states	5
N_b^f	Length bond grid for distribution	60
N_b	Length bond grid for savings	20
N_n	Length bond grid for labor supply	20
$\overline{b}$	Max. bond grid	90
x_1	Min. x added to χ	0.001
c_{min}	Min. consumption	0.001
–	Nb. iterations endogenous grid for initial guess	150
–	Solver tolerance for policy functions	10^{-6}
–	Solver tolerance for prices and ψ	10^{-6}

Table 7: Parameters for computations.

On a 3.4 GHz Intel Core i5-7500 desktop with 8 GB of RAM, it takes 55s to solve for the model steady state, 7s and 821s to compute the Jacobian and the Hessian using automatic differentiation (in Julia), 8s and 170s to call `gensys` and `gensys2` (in MATLAB). Overall, the model is solved in 15-20min.

B.2 Parameter Identification

In Figure 8, I separately plot surfaces for the risk-free rate autocorrelation and volatility, as functions of the credit shock autocorrelation and volatility, to show that the latter are well-identified, i.e. surfaces are not flat at the point that represents the model calibration.

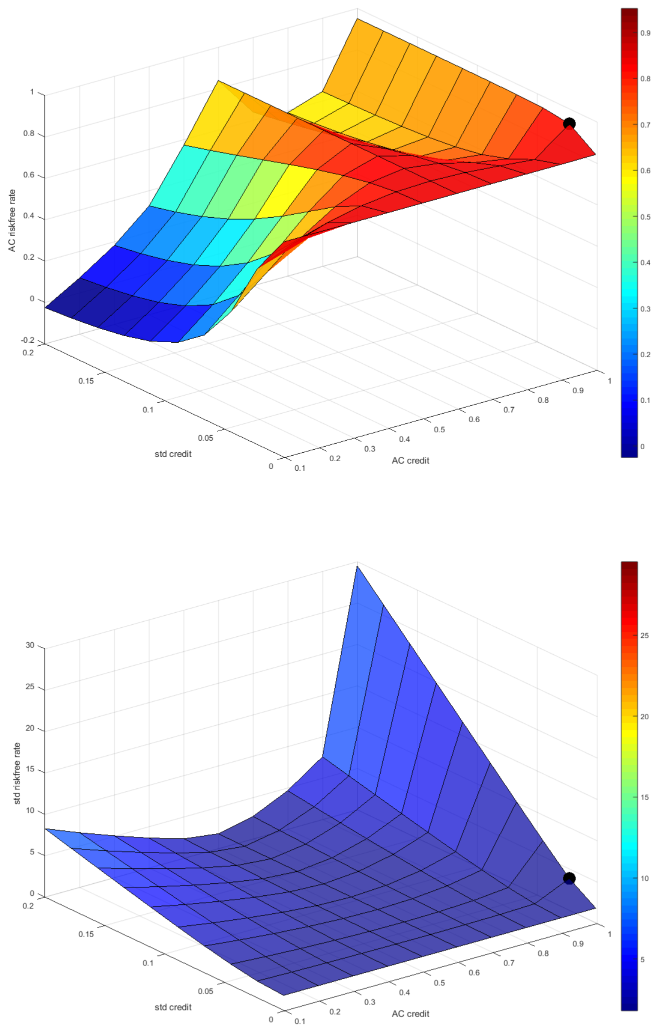

Figure 8: Risk-free rate autocorrelation (upper panel) and annual % volatility (lower panel), as functions of the credit shock autocorrelation and volatility, estimated in a simulation of the linearized model with $T = 10,000$ periods. In each graph, the black dot is at the model calibration for the credit shock process.

C TFP Shocks

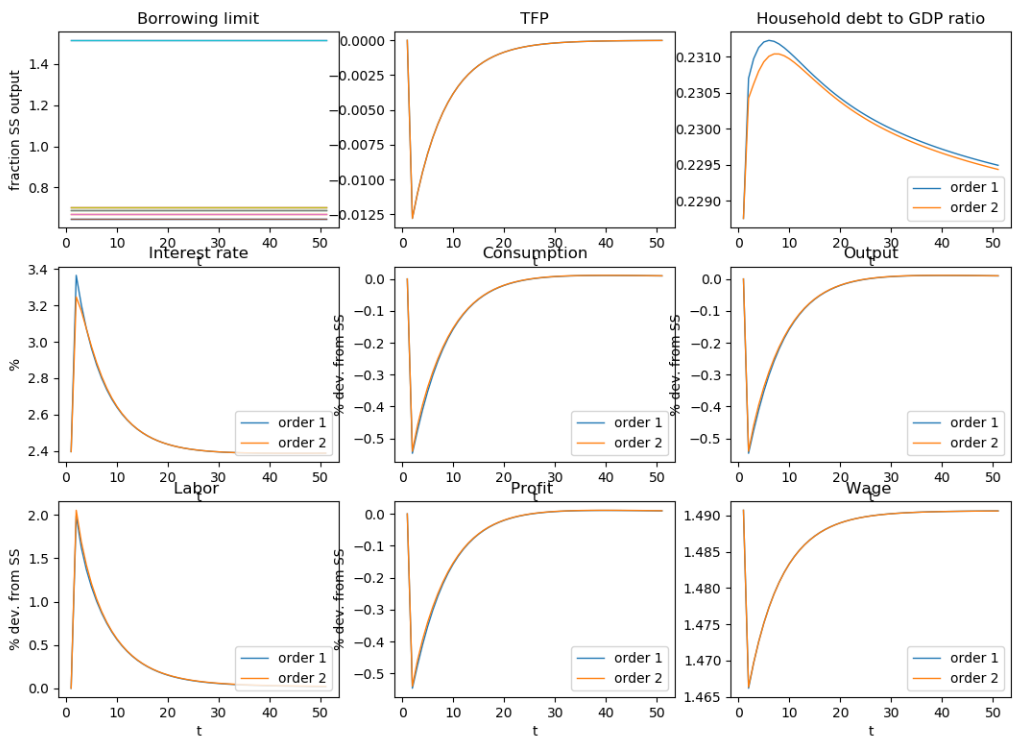

Figure 9: IRF to TFP shock: order 1 vs 2. The upper left panel plots the response of boorowing constraints to output for all income types (θ_1 for the lowest line, θ_5 for the highest), here zero. Initial period: deterministic steady state.

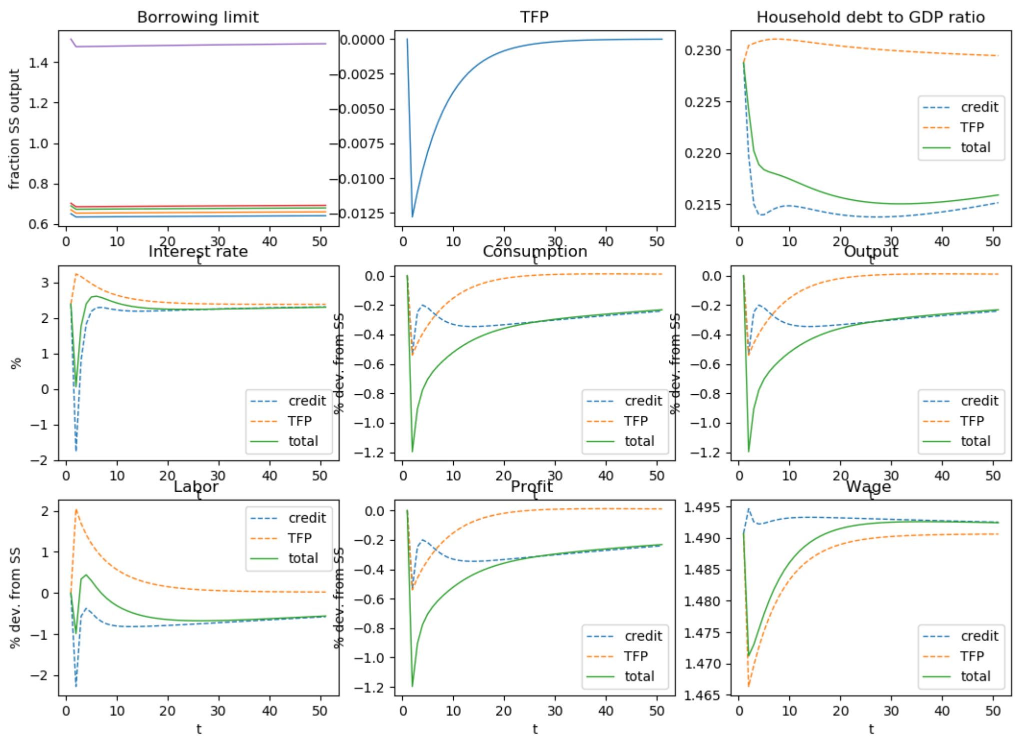

Figure 10: Decomposition of the total response (green line) between credit shocks (dotted blue) and TFP shocks (dotted orange): order 2. The upper left panel plots the response of boorowing constraints to output for all income types (θ_1 for the lowest line, θ_5 for the highest)

D Data

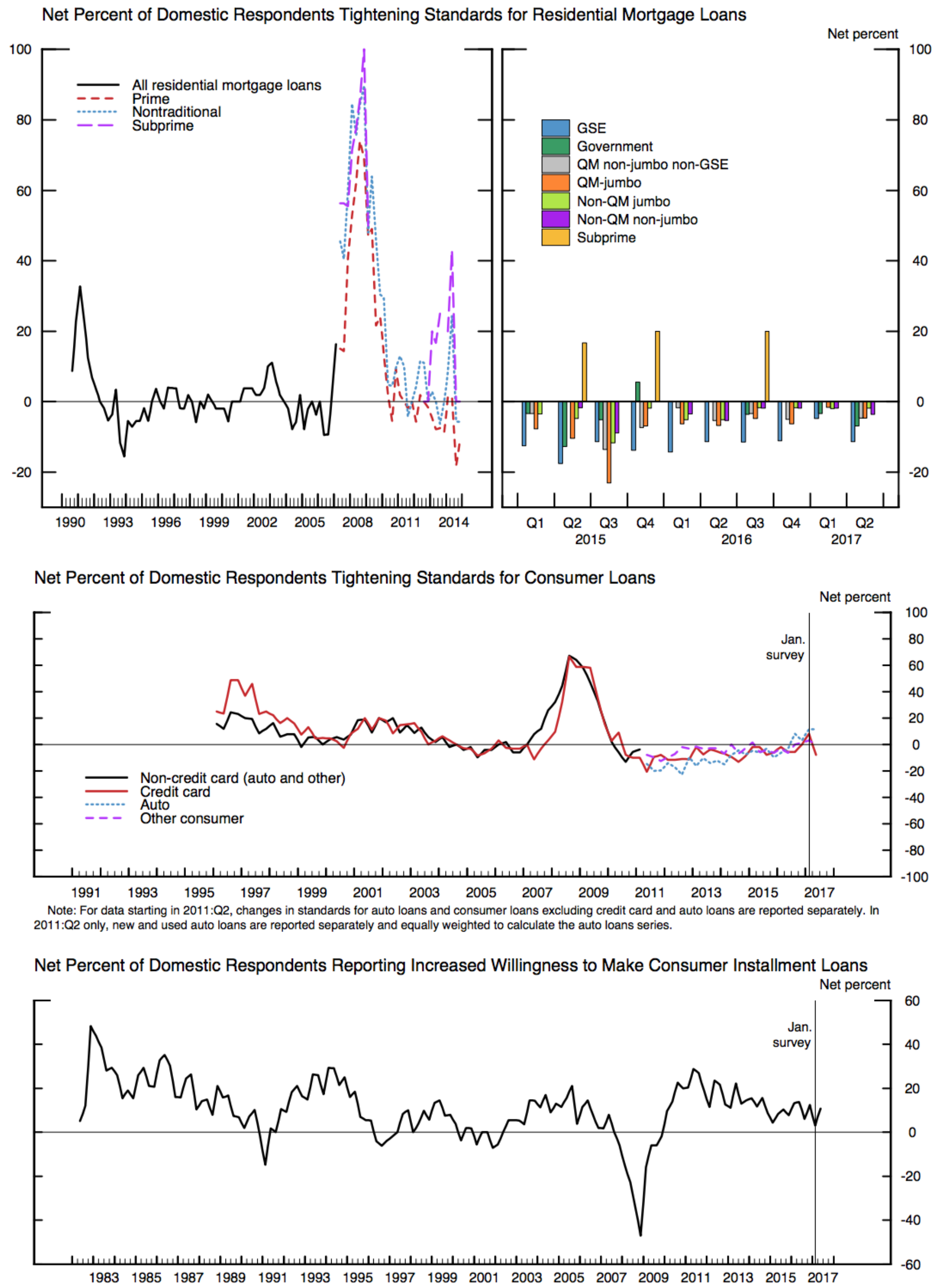

Figure 11: Source: Ferederal Reserve Board, April 2017 Senior Loan Officer Opinion Survey on Bank Lending Practices. Quarterly frequency.

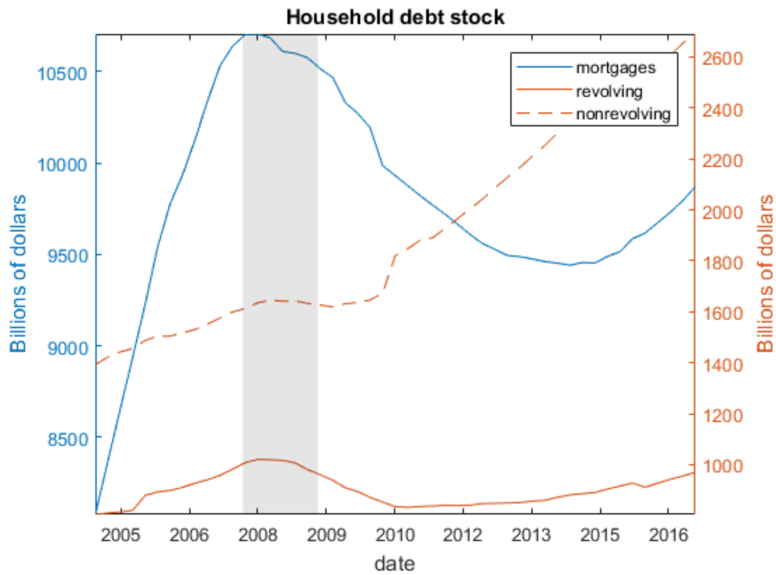

Figure 12: Left axis: Households and Nonprofit Organizations; Home Mortgages; Liability, Level, Billions of Dollars, Quarterly, Seasonally Adjusted (source: FRB, Z.1 Financial Accounts of the United States). Right axis: Total Revolving and Nonrevolving Credit Owned and Securitized, Outstanding, Billions of Dollars, Quarterly (originally series is monthly), Seasonally Adjusted (source: FRB, G.19 Consumer Credit). Shaded areas represent NBER recession.

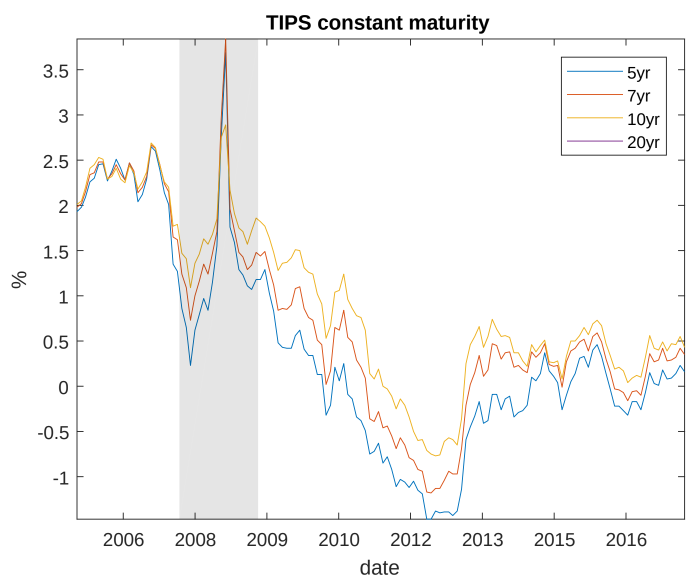

Figure 13: Treasury Inflation-Indexed Security, Constant Maturity, Percent, Monthly, Not Seasonally Adjusted. Maturity from 5 to 20 year. Source: FRB, H.15 Selected Interest Rates. Shaded areas represent NBER recession.

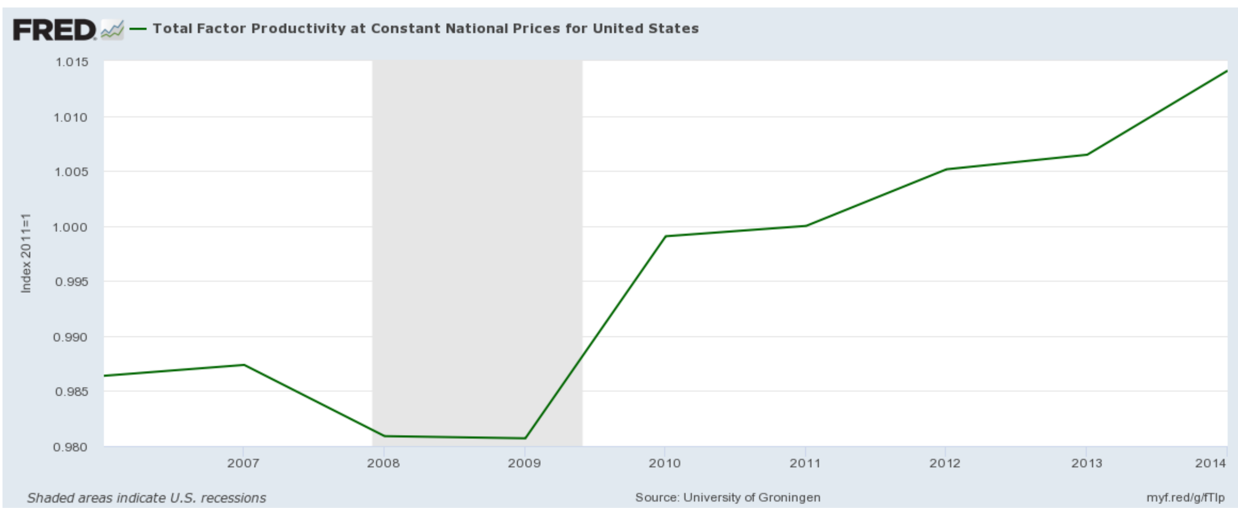

Figure 14: Total Factor Productivity at Constant National Prices for United States. Source: Penn World Table 9.0. Shaded areas represent NBER recession.