Macroprudential Policy, Balance Sheets, and Tail Risk

Yucheng Zhou*

March 13, 2026

Abstract

This paper develops a continuous-time general equilibrium model in which risk-tolerant financial intermediaries borrow from risk-averse households to hold levered positions in productive capital, subject to Value-at-Risk (VaR) constraints on their risk exposure. Adverse aggregate shocks erode intermediary capital and depress asset prices and investment through a balance sheet channel—precisely the conditions under which VaR constraints bind, triggering forced deleveraging that amplifies these effects. The quantitative analysis reveals a central tension in macroprudential regulation: tighter constraints mitigate tail risk, but simultaneously increase the overall volatility of the financial system by magnifying the impact of shocks when constraints bind. The crisis analysis sharpens this tension: tighter constraints reduce crisis frequency but prolong crisis duration, while a state-dependent constraint that tightens with financial conditions comes closest to matching NBER-dated recession statistics.

Keywords: General equilibrium, heterogeneous investors, leverage, VaR constraints.
JEL Codes: E13, E22, E32, E44, G11, G12, G28.

*Gies College of Business, University of Illinois Urbana-Champaign. Email: yz33@illinois.edu.

1 Introduction

Macroprudential policies, such as Value-at-Risk (VaR) constraints and leverage limits on financial intermediaries, have become central tools for financial stability regulation in the aftermath of the 2008 financial crisis. While these policies aim to reduce systemic risk, their effects on aggregate investment and economic growth remain a subject of active debate (Adrian and Shin, 2014). Recent studies in macro-finance show that financial constraints on intermediaries can persistently amplify systemic shocks through the *balance sheet channel*, as in He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014) and Di Tella (2017). However, the quantitative consequences of macroprudential regulations for aggregate investment dynamics—and in particular, the distributional trade-offs they create—remain relatively underexplored in the literature.

This paper examines how macroprudential policies, in particular VaR-type constraints on intermediaries' risk exposure, affect aggregate investment and economic growth through the balance sheet channel. The central finding is a tension in macroprudential regulation: tighter constraints reduce the negative skewness and excess kurtosis of the stationary distribution of asset valuations and growth, thereby mitigating tail risk, but simultaneously increase the overall volatility of asset valuations by amplifying the impact of shocks when constraints bind. The crisis analysis sharpens this tension further: tighter constant constraints reduce crisis frequency by suppressing the volatility of intermediary wealth, but once a crisis occurs, the amplification mechanism prolongs the downturn, resulting in longer average crisis duration. A state-dependent constraint that tightens endogenously with financial conditions produces intermediate outcomes, illustrating how the cyclicality of regulation shapes the entire distribution of macroeconomic outcomes.

To formally develop these results, I build a continuous-time general equilibrium model with two main ingredients: (i) a productive economy subject to aggregate shocks and investment adjustment costs, and (ii) two types of investors with recursive preferences who differ in risk aversion and face portfolio constraints. Risk-tolerant financial intermediaries borrow from risk-averse households to hold levered positions in productive capital. Following Adrian and Shin (2014) and Kargar (2021), intermediaries face a VaR-type constraint that caps their risk exposure, representing macroprudential policy. The model is solved globally using projection methods, which is essential for capturing the occasionally binding nature of the constraints and the resulting nonlinear dynamics.

The production side builds on the framework of production economies with investment adjustment costs. The technology produces consumption goods using capital and is subject to aggregate capital "quality" shocks. Capital adjustment costs prevent instantaneous adjustment of investment, creating a tight link between the price of capital (Tobin's q) and the investment rate through the firm's optimality condition. The economy is populated by two types of investors with recursive preferences, who differ in their level of risk aversion. Financial intermediaries (labeled A) are relatively risk-tolerant and hold levered positions in the risky assets financed by borrowing from more risk-averse investors, i.e., households (labeled B).[1] Investors face financial frictions in the form of VaR constraints that limit their risk exposure. I consider both constant constraints of varying tightness and a state-dependent constraint that tightens as intermediary wealth declines. Crucially, in the absence of these frictions, investors would share risks efficiently and aggregate investment would be at its first-best level.

The equilibrium dynamics are driven by one endogenous state variable: the wealth share of risk-tolerant intermediaries as a fraction of total wealth. The mechanism operates as follows. When the economy is hit by a negative aggregate shock, the wealth share of intermediaries falls because they hold leveraged positions. Since households are now more dominant in clearing markets, aggregate risk aversion rises, the market price of risk increases, and Tobin's q declines. Through the firm's optimality condition, a lower q translates directly into lower investment and aggregate growth. This is the standard balance sheet channel.

A feedback loop amplifies this mechanism when VaR constraints bind. As intermediary balance sheets deteriorate and return volatility rises, the constraint forces intermediaries to shed risk, widening the gap between the constrained and unconstrained equilibrium. The forced deleveraging further depresses asset prices and investment, reinforcing the deterioration of intermediary wealth shares and the rise in aggregate risk aversion. Even though the underlying aggregate shocks are *i.i.d.*, the effects on investment are persistent due to the occasionally binding nature of the constraints and the endogenous dynamics of the wealth distribution.

The quantitative analysis reveals that macroprudential policies reshape the *entire distribution* of macroeconomic outcomes, not just its center. Comparing across constraint scenarios, tighter VaR constraints compress the left tail and reduce excess kurtosis in the stationary distribution of asset valuations, consistent with the goal of mitigating systemic risk. However, they simultaneously increase the overall dispersion of valuations—a paradoxical finding that arises because forced deleveraging during downturns magnifies the impact of aggregate shocks on Tobin's q. The state-dependent constraint, which tightens endogenously as financial conditions deteriorate, produces outcomes intermediate between the constant-constraint scenarios, illustrating how the cyclicality of regulation governs the trade-off between tail risk and volatility. Impulse response functions confirm that the VaR constraint has negligible effects

[1] Intermediaries should be interpreted as representing leveraged institutions, such as commercial banks, investment banks, and hedge funds.

during expansions but amplifies and prolongs the adjustment process during downturns, generating endogenous persistence through the balance sheet channel.

I complement the distributional analysis by defining crisis events based on the simulated growth rate and computing their frequency and duration across policy scenarios. Tighter constant constraints reduce crisis frequency—as the suppressed volatility prevents the economy from drifting into the crisis region—but increase crisis duration, as the amplification mechanism slows the recovery of intermediary wealth. Compared to NBER-dated recession statistics, the model captures the right order of magnitude for both crisis probability and duration, with the state-dependent constraint coming closest to the empirical crisis frequency.

The main contributions of this study are (i) presenting a dynamic framework to quantitatively analyze the effects of macroprudential policies on aggregate investment through the balance sheet channel, (ii) showing that VaR constraints create a fundamental trade-off between tail risk reduction and volatility amplification, (iii) demonstrating that tighter constraints reduce crisis frequency but prolong crisis duration, creating a tension between prevention and amplification, and (iv) illustrating how the cyclicality of regulation—constant versus state-dependent constraints—shapes the distribution of macroeconomic outcomes.

Literature review. This paper fits into several strands of literature. First, it contributes to the voluminous literature going back to the seminal contributions of Bernanke and Gertler (1989), Bernanke et al. (1999), and Kiyotaki and Moore (1997). I extend the framework of He and Krishnamurthy (2013) and Brunnermeier and Sannikov (2014, 2016), who study the role of financial sectors in amplifying systemic shocks into fluctuations of risk premia and asset prices. My paper blends these insights by incorporating heterogeneous investors with recursive preferences and state-dependent portfolio constraints into a production economy with investment frictions. The focus is on how VaR-type macroprudential regulations affect aggregate investment efficiency through the balance sheet channel.

Di Tella (2017), Gopalakrishna (2023), Maxted (2024), and Krishnamurthy and Li (2025) are recent examples that build on Brunnermeier and Sannikov (2014) to study additional amplification mechanisms. A key difference of my paper is the focus on macroprudential policy: I show how VaR-type constraints, intended to reduce systemic risk, can paradoxically amplify economic fluctuations by limiting aggregate investment during downturns. The model demonstrates that when intermediaries face binding constraints, the effects work through discount-rate changes and reduced investment rather than physical capital transfers.

This paper contributes to the literature on macroprudential regulation and financial stability. Adrian and Shin (2014) document how VaR-based risk management at financial institutions leads to procyclical leverage. Kargar (2021) studies state-dependent margin constraints in an intermediary asset pricing framework. My contribution is to embed these constraints in a production economy with endogenous investment, showing how macroprudential policies affect aggregate growth through the balance sheet channel. The model provides a laboratory for evaluating the trade-offs inherent in financial stability regulations—in particular, the tension between tail risk mitigation and volatility amplification.

On a technical note, this paper also contributes to the heterogeneous-agent asset pricing literature, for which Panageas (2020) provides an excellent survey. The seminal contributions of Dumas (1989, 1992), Wang (1996), Chan and Kogan (2002), Bhamra and Uppal (2009, 2014), Longstaff and Wang (2012), Gârleanu and Panageas (2015) study equilibrium in economies with two heterogeneous agents and different preference assumptions. Basak and Cuoco (1998), Gârleanu and Pedersen (2011), He and Krishnamurthy (2012, 2013), Chabakauri (2013), and Rytchkov (2014) examine the asset pricing implications of exogenous and endogenous portfolio constraints in economies populated by two heterogeneous agents with one or many assets where constrained agents have logarithmic or constant relative risk aversion (CRRA) preferences. More recent examples include Drechsler et al. (2018), Silva (2020), Kargar (2021), Schneider (2022), who study the role of heterogeneous agents with recursive preferences in various settings. This paper follows by allowing for investors with recursive preferences and endogenous portfolio constraints in a production economy with real frictions.

The remainder of the paper is organized as follows. Section 2 lays out the model. Section 3 discusses the model's solution and the characterization of the equilibrium. Section 4 presents the quantitative analysis. Section 5 examines the model's implications for financial crises and macroeconomic stability. Finally, Section 6 concludes.

2 Model

In this section, I outline a general equilibrium model in continuous time. A representative firm uses capital to produce consumption goods, subject to investment adjustment costs. The real technology employed by the firm is subject to aggregate shocks. The economy is populated by a continuum of investors with recursive preferences, categorized into two classes: financial intermediaries and households. Financial intermediaries are less risk averse than households, although both types of investors have access to the financial market subject to portfolio constraints. The information structure of the economy is standard. Uncertainty is characterized in a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$, with a standard filtration $\mathscr{F}_t, t \in [0, \infty), \mathscr{F}_t \subset \mathscr{F}$.

In the remainder of this section, I first describe the economic setup and decision problems faced by the firm and investors, and then define the competitive equilibrium.

2.1 Production

The productive side of the economy consists of a representative firm using capital as the sole factor of production. Let K, I, and Y denote the capital stock, gross investment, and output, respectively. I first introduce the firm's technology, followed by its optimization problem.

2.1.1 Technology

The representative firm employs a constant returns-to-scale production technology, specifically an "AK" type.[2] The stock of capital K_t produces output at rate

$$Y_t = AK_t \tag{1}$$

per unit of time, where A is a constant governing the productivity of capital.

Capital stock is exposed to exogenous aggregate Brownian shocks, introducing instantaneous riskiness. The evolution of capital stock follows

$$\frac{dK_t}{K_t} = g_t dt + \sigma dB_t, \tag{2}$$

where g_t denotes the expected growth rate of capital at time t. $B = \{B_t \in \mathbb{R}; \mathscr{F}_t, t \geq 0\}$ is a standard Brownian motion in the probability space $(\Omega, \mathscr{F}, \mathbb{P})$, representing aggregate shocks to the capital stock. Such shocks may reflect stochastic depreciation, investment-specific, or capital quality shocks.[3] The constant $\sigma > 0$ is the exposure of capital to the aggregate shock.

Following Hayashi (1982), Jermann (1998) and others, investment incurs adjustment cost that is homogeneous of degree one in investment and capital stock.[4] To achieve a given expected growth rate of capital $g(\iota)$, the firm invests $I = \iota K$, where ι denotes the investment-capital ratio. Function $g(\cdot)$ satisfies $g'(\cdot) > 0$ and $g''(\cdot) \leq 0$. I assume a quadratic adjustment cost, and the investment-capital ratio is given by

$$g = \iota - \delta - \frac{\theta}{2}\iota^2, \tag{3}$$

where θ is the adjustment cost parameter and δ is the depreciation rate of capital.

[2]See, for example, Kogan (2001, 2004), Brunnermeier and Sannikov (2014), and Pindyck and Wang (2013).

[3]See, for example, Brunnermeier and Sannikov (2014), and Maxted (2024).

[4]Homogeneous adjustment cost functions have been widely used in the q theory of investment literature. Hayashi (1982) showed that with homogeneous adjustment costs and perfect capital markets, marginal and average q are equal. Jermann (1998) integrated this type of adjustment costs into a real business cycle model.

2.1.2 Optimization problem

The firm pays output net of investment as dividends to shareholders. It chooses the investment rate[5] to maximize the expected discounted value of dividends, given the *stochastic discount factor (SDF)*. The value of the firm is given by

$$S_t = \sup_{g_s} \mathbb{E}_t \left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t} (AK_s - \iota(g_s)K_s) ds \right]$$
$$\text{s.t. capital evolution (2) and adjustment cost (3).} \tag{4}$$

where Λ_t denotes the SDF that is determined in equilibrium. The evolution of SDF can be written as

$$\frac{d\Lambda_t}{\Lambda_t} = -r_t dt - \eta_t dB_t \tag{5}$$

where r_t is the risk-free rate and η_t is the market price of risk, both determined in equilibrium.

2.2 Investors

The economy is populated by a continuum of investors in two classes, A and B. I start from the demographic structure and preference of the investors in the economy, followed by their budget and financial constraints. Then I define the optimization problems faced by investors.

2.2.1 Demographics and preferences

The aggregate population remains constant and is normalized to one. To obtain stationarity in the model, I follow Gârleanu and Panageas (2015) and consider a simple overlapping-generations (OLG) framework. All investors, regardless of age, preference, or wealth, face an exogenous and constant mortality rate $\kappa > 0$, and a fraction κ of new investors are born per unit of time.[6] In aggregate, newborn investors inherit the wealth of their deceased parents on an equal per-capita basis. Of the newly born investors, a constant fraction $\bar{u}$ is of type A, while $1 - \bar{u}$ is of type B. The OLG setup ensures the existence of a stationary distribution of wealth among investors.

All investors of two classes have recursive preferences with constant coefficient of relative risk aversion (RRA) and elasticity of intertemporal substitution (EIS).[7] Specifically, the preference is described by the stochastic differential utility (SDU) introduced by Duffie and Epstein

[5] Or equivalently, the desired expected capital growth rate.

[6] The total population size is thus $\int_{-\infty}^{t} \kappa e^{-\kappa(t-s)} ds = 1$.

[7] See, for example, Kreps and Porteus (1978), Epstein and Zin (1989), and Weil (1989).

(1992a,b). That is, an investor of type i for $i \in \{A, B\}$ has lifetime utility defined by

$$J_{i,t} = \mathbb{E}_t \left[\int_t^{\infty} f_i(C_{i,s}, J_{i,s}) ds \right], \tag{6}$$

where function $f(C,J)$ is known as the aggregator of the recursive utility for consumption C and the continuation value J. The aggregator function $f(C,J)$, common form across investors, is given by

$$f(C,J) = \frac{1-\gamma}{1-1/\psi} J \left[\left(\frac{C}{[(1-\gamma)J]^{1/(1-\gamma)}} \right)^{1-1/\psi} - (\hat{\rho} + \kappa) \right], \tag{7}$$

where the RRA is denoted as γ, and the EIS is denoted as ψ. For simplicity, assume $\gamma \neq 1$. The time-additive separable CRRA utility is a special case of the above recursive utility specification, when the RRA γ is equal to the reciprocal of the EIS ψ, i.e., $\gamma = 1/\psi$. Gârleanu and Panageas (2015) show that, due to mortality risk, the (effective) subjective discount rate is defined as $\rho \equiv \hat{\rho} + \kappa$, where $\hat{\rho}$ is the time impatience.

The two classes of investors have different RRA γ. I use A to denote financial intermediaries and B to denote households. In particular, households are more risk averse than financial intermediaries, i.e., $\gamma_A < \gamma_B$. It follows that in equilibrium financial intermediaries, including banks, hedge funds, and broker-dealers, etc., raise funds from households via short-term debt financing and use these funds to make levered investments. To focus on the risk aversion heterogeneity, I assume the same EIS across all investors.

2.2.2 Budget and financial constraints

All investors in the economy can trade claims on dividends from the firm and a risk-free bond in zero net supply, subject to portfolio constraints. Investors dynamically choose their consumption and portfolio weights. Let s denote the birth date of an investor and $W_{i,t}(s)$ denote the wealth of an investor of type i born in date s at time t. Since investors of the same type face the same optimization problems (defined below), the decisions do not depend on the age of each investor. Thus, the notation can be simplified by dropping s in the wealth dynamics. Let $\omega_{i,t}$ be the portfolio weight in the risky asset. Then the financial wealth of an investor of type i evolves according to the dynamic budget constraint

$$\frac{dW_{i,t}}{W_{i,t}} = [\omega_{i,t} \lambda_{R,t} + r_t - c_{i,t}] dt + \omega_{i,t} \sigma_{R,t} dB_t, \tag{8}$$

where $c_{i,t} \equiv C_{i,t}/W_{i,t}$ is investor's consumption-wealth ratio. r_t is the risk-free rate, $\lambda_{R,t}$ is the expected excess return of the risky asset, and $\sigma_{R,t}$ is the loading on the aggregate shock of asset returns (all defined in equilibrium in Section 3).

The leverage of an investor is defined as the ratio of asset over equity. In equilibrium, intermediaries lever up their balance sheets by borrowing from households. That is, when the portfolio weight in the risky asset exceeds one, the intermediaries lever up by issuing instantaneous risk-free debt to households. I follow Kargar (2021) and assume investors face an occasionally binding state-dependent margin constraint. Specifically, at time t, the margin constraint is assumed as

$$\omega_{i,t} \le \zeta_t, \tag{9}$$

where ζ_t determines the form of margin constraints, which is linked to endogenously determined equilibrium objects. In particular, I assume that margin requirements depend inversely on the volatility of the risky asset return. Because return volatilities also depend on the state of the economy, margin requirements are state dependent as well.

I consider a change of variables that transforms the characterization to simplify the exposition. Define the risk exposure of an investor of type i as $\sigma_{i,t} \equiv \omega_{i,t}\sigma_{R,t}$. Given the absence of arbitrage, the expected excess return of the risky asset is given by $\lambda_{R,t} = \sigma_{R,t}\eta_t$. The dynamic budget constraint can then be rewritten as

$$\frac{dW_{i,t}}{W_{i,t}} = [\sigma_{i,t}\eta_t + r_t - c_{i,t}]dt + \sigma_{i,t}dB_t. \tag{10}$$

Following the change of variables, I assume financial constraints in terms of the risk exposure as

$$\sigma_{i,t} \le \hat{\zeta}_t, \tag{11}$$

where $\hat{\zeta}_t$ denotes the tightness of the constraint on the risk exposure.[8] Therefore, the investors effectively face a VaR rule, as in Danielsson et al. (2012); Adrian and Shin (2014); Kargar (2021). This VaR constraint represents macroprudential policy in the model. Although all investors face margin constraints, in equilibrium, consistent with empirical evidence, only financial intermediaries face constraints that occasionally bind.

2.2.3 Optimization problems

Investors within each type face identical optimization problems, as mentioned above. They start with initial wealth $W_{i,0} > 0$, decide on consumption as a fraction of wealth $c_{i,t}$, and risk exposure $\sigma_{i,t}$ to maximize lifetime utility. The optimization problem of an investor of type i is

[8]Note that Kargar (2021) assume the portfolio weight is bounded by $\frac{1}{\alpha\sigma_{R,t}}$ and thus the risk exposure is bounded by a constant $\frac{1}{\alpha}$.

given by

$$V_{i,t} = \max_{\{c_{i,s}\geq 0, \sigma_{i,s}\}} J_{i,t}$$
$$\text{s.t. dynamic budget constraint (10) and financial constraint (11),} \tag{12}$$

and a solvency constraint $W_{i,t} \geq 0$.

2.3 Equilibrium

The following defines the competitive equilibrium.

Definition 1 *A competitive equilibrium is a set of adapted stochastic processes for interest rate r_t, market price of risk η_t, stochastic discount factor Λ_t; firm's output Y_t, capital K_t, investment I_t, and value S_t; investors' wealth $W_{i,t}$, consumption $C_{i,t}$, risk exposure $\sigma_{i,t}$ for $i \in \{A, B\}$ such that*

i. *the firm solves an optimal investment problem* (2.1.2)*;*

ii. *investors maximize utility by choosing consumption and risk exposure* (2.2.3)*;*

iii. *markets clear:*

$$\begin{aligned} \bar{u}C_{A,t} + (1-\bar{u})C_{B,t} &= Y_t - I_t \quad \text{(goods market)}, \\ \bar{u}W_{A,t}\sigma_{A,t} + (1-\bar{u})W_{B,t}\sigma_{B,t} &= S_t\sigma_{R,t} \quad \text{(risk exposures market)}. \end{aligned} \tag{13}$$

The bond market clears by Walras' law and it implies

$$\bar{u}W_{A,t} + (1-\bar{u})W_{B,t} = S_t. \tag{14}$$

3 Solution

The equilibrium can be characterized in a recursive formulation, where all equilibrium objects are functions of the aggregate state variable. In this section, I first define the endogenous state variable and derive its dynamics. Subsequently, I derive the Hamilton-Jacobi-Bellman (HJB) equations for the scenario without portfolio constraints, characterize the solutions to the optimization problems of the firm and investors, and provide intuition regarding the optimal policy functions. Lastly, I define a recursive Markov equilibrium and briefly outline the numerical algorithm used to solve the system of equations.

3.1 State variable

Due to the homogeneity of Epstein-Zin preferences and the capital accumulation process, all variables can be rescaled by the total wealth or total capital. Therefore, it suffices to track only one aggregate state variable: the share of the aggregate wealth that belongs to the investors of type A. The equilibrium conditions can be derived as functions of the following endogenous state variable:

$$w_t = \frac{\bar{u}W_{A,t}}{\bar{u}W_{A,t} + (1-\bar{u})W_{B,t}}. \tag{15}$$

State variable w is the key state variable in recent intermediary asset pricing models (He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; Kargar, 2021).[9] The dynamics of w capture the redistribution of wealth between risk-tolerant intermediaries and risk-averse households, which is the main driver of the equilibrium objects in this model. The following proposition gives the law of motion of w.

Proposition 1

$$dw_t = \kappa(\bar{u} - w_t)dt + w_t(1-w_t)[\mu_{w,t}dt + \sigma_{w,t}dB_t], \tag{16}$$

where $\mu_{w,t}$ and $\sigma_{w,t}$ are given by

$$\begin{aligned} \mu_{w,t} &= (\sigma_{A,t} - \sigma_{B,t})\eta_t + c_{B,t} - c_{A,t} - (\sigma_{A,t} - \sigma_{B,t})\sigma_{R,t}, \\ \sigma_{w,t} &= \sigma_{A,t} - \sigma_{B,t}. \end{aligned}$$

Proof. See Appendix A. ■

Notice that the law of motion of the intermediaries' wealth share w_t has an exogenous component in the drift function due to the demographic structure assumed earlier. This term ensures a non-degenerate long-run distribution of w, i.e., $w \in (0,1)$. The endogenous component (in brackets) is driven by the consumption and risk exposure choices of investors of both types. Term $\mu_{w,t}$ depends on the consumption-wealth ratio of households relative to that of intermediaries, and on the differential aggregate risk exposure between intermediaries and households. Term $\sigma_{w,t}$ only depends on this latter force. As will be shown in (25), the wealth share dynamics depend on asset prices, which themselves depend on wealth share dynamics-generating a two-way feedback loop that amplifies return volatility (Brunnermeier and Sannikov, 2014).

[9]Note that the combination of heterogeneity and recursive preferences gives rise to the wealth share as a state variable even under perfect risk sharing.

3.2 Hamilton-Jacobi-Bellman equations

The recursive formulation of the equilibrium is characterized by the value function of the firm and investors. The value functions satisfy the HJB equations.

3.2.1 Firm's problem

The firm's problem is to maximize the expected discounted value of dividends D_t, which can be written recursively,

$$0 = \max_{g_t} \left\{ D_t dt + \mathbb{E}_t \left[\frac{d(S_t \Lambda_t)}{\Lambda_t} \right] \right\}. \tag{17}$$

The homogeneity property implies that marginal q equals Tobin's average q, following Hayashi (1982). The following proposition gives the firm's value function.

Proposition 2 *The value of the firm has the form*

$$S_t(K_t, w_t) = K_t q_t(w_t), \tag{18}$$

where q_t captures future investment opportunities of the firm. The dynamics of q_t is given by

$$\frac{dq_t}{q_t} = \mu_{q,t} dt + \sigma_{q,t} dB_t, \tag{19}$$

where $(\mu_{q,t}, \sigma_{q,t})$ is determined in equilibrium. The expressions for $\mu_{q,t}$ and $\sigma_{q,t}$ are given in the Appendix A.

The HJB equation for the firm's problem is given by the following differential equation:

$$0 = \max_{g_t} \left\{ \frac{A - \iota(g_t)}{q_t} + g_t + \mu_{q,t} + \sigma_{q,t}\sigma - r_t - \eta_t(\sigma_{q,t} + \sigma) \right\}. \tag{20}$$

Proof. See Appendix A. ■

The first-order condition (FOC) gives the optimal growth rate as follows

$$\iota'(g_t) = q_t. \tag{21}$$

Consistent with Q-theory of investment, the optimal investment rate ι_t is an increasing function of q_t, which summarizes the marginal benefit of increasing capital stock. The marginal cost, on the other hand, is given by $\iota'(g_t)$. In equilibrium, the representative firm optimally chooses to equate marginal cost to marginal benefit.

The return of holding claims on the firm is given by the dividend yield plus capital gains, as follows

$$dR_t = \frac{A - \iota(g_t)}{q_t} dt + \frac{dS_t}{S_t}, \tag{22}$$

$$\equiv \mu_{R,t} dt + \sigma_{R,t} dB_t, \tag{23}$$

where

$$\mu_{R,t} = \frac{A - \iota(g_t)}{q_t} + g_t + \mu_{q,t} + \sigma_{q,t}\sigma, \tag{24}$$

$$\sigma_{R,t} = \sigma_{q,t} + \sigma. \tag{25}$$

$\mu_{R,t}$ and $\sigma_{R,t}$ denote the expected return and the volatility term (loading on the aggregate shock) of the asset's return, respectively. Notice that the volatility of return is driven by an exogenous component σ and an endogenous component $\sigma_{q,t}$, which captures the firm's investment opportunities. The expected excess return is $\lambda_{R,t} \equiv \mu_{R,t} - r_t$.

3.2.2 Investors' problems

The investor's problem of each type is to maximize lifetime utility, which can be written recursively,

$$0 = \max_{c_{i,t}, \sigma_{i,t}} \left\{ f_i(c_{i,t} W_{i,t}, V_{i,t}) dt + \mathbb{E}_t[dV_{i,t}] \right\}. \tag{26}$$

With homothetic preferences, the following proposition gives investors' value functions.

Proposition 3 *The value function of investor $i \in \{A, B\}$ has the form*

$$V_{i,t}(W_{i,t}, w_t) = \frac{W_{i,t}^{1-\gamma_i}}{1-\gamma_i} \xi_{i,t}(w_t)^{\frac{1-\gamma_i}{1-\psi}}, \tag{27}$$

where $\xi_{i,t}$ captures the investor i's future investment opportunities. The dynamics of $\xi_{i,t}$ is given by

$$\frac{d\xi_{i,t}}{\xi_{i,t}} = \mu_{\xi_i,t} dt + \sigma_{\xi_i,t} dB_t, \tag{28}$$

where $(\mu_{\xi_i,t}, \sigma_{\xi_i,t})$ is determined in equilibrium. The expressions for $\mu_{\xi_i,t}$ and $\sigma_{\xi_i,t}$ are given in the Appendix A.

The HJB equation for the investors' problem is given by the following differential equation:

$$
\begin{aligned}
0 = \max_{c_{i,t},\sigma_{i,t}} \Bigg\{ & \frac{1}{1-1/\psi}\left(c_{i,t}^{1-1/\psi}\xi_{i,t}^{1/\psi}-\rho\right)+\left(\sigma_{i,t}\eta_t+r_t-c_{i,t}\right) \\
& +\frac{1}{1-\psi}\mu_{\xi_i,t}-\frac{\gamma_i}{2}\sigma_{i,t}^2 \\
& +\frac{1}{2}\frac{1}{1-\psi}\left(\frac{1-\gamma_i}{1-\psi}-1\right)\sigma_{\xi_i,t}^2 \\
& +\frac{1-\gamma_i}{1-\psi}\sigma_{i,t}\sigma_{\xi_i,t}\Bigg\}.
\end{aligned} \tag{29}
$$

Proof. See Appendix A. ■

The first-order conditions give the optimal consumption and risk exposure as follows

$$
c_{i,t} = \xi_{i,t}, \tag{30}
$$

$$
\sigma_{i,t}^{*} = \frac{1}{\gamma_i}\eta_t + \frac{1}{\gamma_i}\frac{1-\gamma_i}{1-\psi}\sigma_{\xi_i,t}. \tag{31}
$$

In equilibrium, the consumption-wealth ratio, which equals $\xi_{i,t}$, responds to changes in investment opportunities. The optimal unconstrained risk exposure has two components, as in standard intertemporal capital asset pricing model (ICAPM) result of Merton (1973). The first term in (31) is the myopic demand of a one-period mean-variance investor, and is driven by the market price of risk. Since intermediaries are assumed to be less risk averse, they tend to be more sensitive to changes in the market price of risk. The second term is the hedging demand capturing the variations in the investor's investment opportunity set.

Taking together the optimal unconstrained risk exposure (31) and the financial constraint (11), the optimal risk exposure is

$$
\sigma_{i,t} = \min\left(\sigma_{i,t}^{*}, \hat{\zeta}_t\right). \tag{32}
$$

3.3 Markov equilibrium

I define a stationary recursive Markov equilibrium with state variable w_t, in which all equilibrium objects are expressed as a function of w.

Definition 2 *A Markov equilibrium in state variable w is the set of functions: equilibrium prices (r, η), firm's valuation ratio q and policy function g, investors' marginal value of wealth ξ_i and policy functions c_i and σ_i for $i \in \{A, B\}$ such that* *i. valuation ratio q solves HJB equation for the firm, and g is the corresponding policy function, taking (r,η) and law of motion for w as given;*

ii. marginal value of wealth ξ_i solves HJB equation for investor of type i, and c_i and σ_i are the corresponding policy functions, taking (r,η), q, and law of motion for w as given;

iii. markets clear:

$$w c_A + (1-w) c_B = \frac{A-\iota(g)}{q} \quad \text{(goods market)},$$
$$w\sigma_A + (1-w)\sigma_B = \sigma_R \quad \text{(risk exposures market)};$$

iv. the law of motion for w satisfies (16).

3.4 Asset pricing implications from unconstrained economy

In this section, I discuss the equilibrium implications for risk taking and asset pricing. However, in the absence of constraints, the equilibrium does not permit closed-form solutions. Therefore, I derive the following propositions to characterize the mechanisms, assuming the economy is unconstrained. The constrained equilibrium will be analyzed in Section 4 through numerical methods.

The following proposition characterizes the type A investors' demand for risks.

Proposition 4 *In the absence of constraints, type A investors' demand for risk, i.e., their risk exposure σ_A^*, is given by*

$$\sigma_A^* = \frac{R_w - \frac{\gamma_B}{w(1-w)}}{(\gamma_B-\gamma_A)\frac{q_w}{q} - \frac{\gamma_B}{1-w} - \frac{\gamma_A}{w} + R_w}\sigma \tag{33}$$

where

$$R_w = \frac{1-\gamma_A}{1-\psi}\frac{\xi_{A,w}}{\xi_A} - \frac{1-\gamma_B}{1-\psi}\frac{\xi_{B,w}}{\xi_B}. \tag{34}$$

Proof. See Appendix A. ■

The above proposition warrants several remarks. First, the risk exposure of type A investors depends on the variable R_w, which captures the risk sharing mechanism through the wealth distribution channel. The variable R_w characterizes the investors' risk sharing motive that arises from the redistribution of wealth between the two types. Mechanically, it can be written as the difference in the sensitivity of the investors' value functions with respect to the state variable w, that is

$$R_w = \frac{\partial \log V_A}{\partial w} - \frac{\partial \log V_B}{\partial w}. \tag{35}$$

A negative (positive) R_w implies that a marginal increase in intermediary wealth share improves the utility of type B (A) investors relatively more. Notice that R_w would be zero if there were no motives to share aggregate risks.

Second, the risk exposure depends on how capital valuation responds to wealth distribution, captured by q_w/q, which is the sensitivity of Tobin's Q with respect to w. This captures that firms' valuations respond to changes in the state variable through the general equilibrium effect on discount rates. Thus the risk exposure of type A investors depends on the interaction of R_w and q_w/q, which allows for a tractable characterization of the risk sharing mechanism in the economy.

It is worth noting that the risk is concentrated in the type A investors, who are less risk averse. Thus, in equilibrium, $\sigma_A^* > \sigma_B^*$ as long as $\gamma_A < \gamma_B$. This implies that intermediaries will be subject to the binding constraint as the volatility of the risky asset increases, which will be shown in Section 4.

The expressions for the market price of risk η_t and the interest rate r_t, which depend on the risk-sharing dynamics of the economy, are given in the following propositions to highlight key implications.

Proposition 5 *In the absence of constraints, the market price of risk η_t can be expressed as follows*

$$\begin{aligned}\eta_t &= \gamma_t \sigma_{R,t} \\ &\quad - \gamma_t \left[w_t \left(\frac{1}{\gamma_A}\right)\left(\frac{1-\gamma_A}{1-\psi}\right)\sigma_{\xi_A,t} + (1-w_t)\left(\frac{1}{\gamma_B}\right)\left(\frac{1-\gamma_B}{1-\psi}\right)\sigma_{\xi_B,t}\right],\end{aligned} \tag{36}$$

where γ_t denotes aggregate risk aversion, and is given by

$$\gamma_t = \left(\frac{w_t}{\gamma_A} + \frac{1-w_t}{\gamma_B}\right)^{-1}. \tag{37}$$

Proof. See Appendix A. ■

The market price of risk (36) is the product of aggregate risk aversion and the difference between the supply of risk and the average hedging demand.[10] Aggregate risk aversion (37) is the reciprocal of the wealth-weighted average of the risk tolerance of the two types of investors. Hence, periods during which intermediaries are relatively less capitalized, so aggregate risk aversion is high, tend to experience a high market price of risk. This is consistent with the heterogeneous agents asset pricing literature (Panageas, 2020). The second term, wealth-weighted average of marginal value of wealth, with the weight also accounting for differences

[10] The term $\sigma_{R,t}$ is the volatility of the risky asset return.

in preferences, captures the idea that if the average hedging demand is high, then a lower market price of risk is required to induce investors to bear the risk.

The risk premium of the asset is given by

$$\lambda_{R,t} = \sigma_{R,t}\eta_t. \tag{38}$$

Following the expression for market price of risk, the risk premium consists of two terms as well. The first term is driven by the volatility of the risky asset. An asset with high volatility is risky and commands a high risk premium. The second term captures the fact that when the average hedging demand is high, a lower risk premium is required to induce investors to bear risk.

Proposition 6 *In the absence of constraints, the interest rate r_t can be expressed as follows*

$$\psi\rho + \mu_t + (1-\psi)(r_t + \Phi_t) = \frac{A - \iota(g_t)}{q_t}, \tag{39}$$

where $\mu_t = w_t\mu_{\xi_A,t} + (1-w_t)\mu_{\xi_B,t}$, and $\Phi_t = w_t\Phi_{A,t} + (1-w_t)\Phi_{B,t}$. The expressions for $\mu_{\xi_i,t}$ and $\Phi_{i,t}$ are given in the Appendix A.

Proof. See Appendix A. ■

The left-hand side of (39) represents the (normalized) aggregate demand for goods, while the right-hand side represents the (normalized) aggregate supply of goods. The effects of the interest rate on aggregate demand depends crucially on the EIS: if $\psi > 1$, then an increase in the interest rate decreases aggregate demand (the opposite is true when $\psi < 1$).

Before proceeding to the quantitative analysis, it is worth stressing that under the representative agent benchmark of this model, the equilibrium outcomes are constant. This is a useful benchmark because it indicates that all variation in the equilibrium in the heterogeneous setup is endogenously driven by investors' risk-sharing activities (subject to the constraints).

4 Quantitative analysis

In this section, I explore the quantitative properties of the model and analyze the effects of macroprudential policies on aggregate investment and economic growth. I first present the calibration of the model. Then I compare the equilibrium functions across different policy scenarios, varying the tightness of the VaR constraint $\hat{\zeta}$. Finally, I trace the dynamic adjustment of the economy following a negative shock using impulse response functions.

4.1 Numerical Solution

The model is analytically tractable: the equilibrium dynamics can be fully characterized by a system of differential equations that are solved numerically. The computation of equilibrium requires solving the HJB equations of the firm and investors simultaneously. The value functions $q(w)$ and $\xi_i(w)$ for $i \in \{A, B\}$ can be found by solving a system of second-order ordinary differential equations (ODEs) in w. To do so, all equilibrium objects need to be expressed in terms of these value functions and their derivatives. Unfortunately, the system of nonlinear differential equations does not admit a closed-form solution, so I rely on numerical techniques. I use projection methods, specifically orthogonal collocation using Chebyshev polynomials, to solve for equilibrium. Unlike a log-linearized representation around the steady state, this method provides a global solution and a full characterization of the whole dynamic system. This is particularly important as the model features regions where constraints occasionally bind. In Appendix B, I discuss the algorithm used to solve the system of differential equations.

4.2 Calibration

Table 1 reports the calibration of the model. Note that the values correspond to annual values as the model is cast in continuous time.

Table 1: Calibration

Parameters	Symbol	Values
Subjective discount rate	ρ	0.001
Mortality rate	κ	0.02
Elasticity of intertemporal substitution	ψ	1.5
RRA of type A investors	γ_A	3
RRA of type B investors	γ_B	21
Population share of type A investors	$\bar{u}$	0.1
Productivity of capital	A	0.2
Volatility of capital	σ	0.04
Depreciation rate of capital	δ	0.04
Adjustment cost	θ	8

This table reports the parameter values used in calibrating the model. The values correspond to annual values as the model is cast in continuous time.

The subjective discount rate ρ and the mortality rate κ of agents are set to 0.001 and 0.02, respectively. This means the effective discount rate is $\rho + \kappa = 0.021$, which is consistent with a real interest rate of around 2% in annual terms. This is comparable to calibrations used in the asset pricing literature.[11] I set EIS to $\psi = 1.5$, common across both investor types. The RRA of intermediaries (type A) is $\gamma_A = 3$, while households (type B) have $\gamma_B = 21$, capturing the substantially greater risk tolerance of financial intermediaries. The population share of type A agents is $\bar{u} = 0.1$. On the production side, productivity is $A = 0.2$, aggregate volatility is $\sigma = 0.04$, the depreciation rate is $\delta = 0.04$, and the adjustment cost parameter is $\theta = 8$. These parameters are broadly consistent with the macro-finance literature (Brunnermeier and Sannikov, 2014; Silva, 2020).

I consider several scenarios for the tightness of the VaR constraint $\hat{\zeta}$ defined in (11). The baseline scenarios set $\hat{\zeta}$ to a constant: $\hat{\zeta} \in \{0.10, 0.12, 0.15\}$. A lower $\hat{\zeta}$ represents progressively tighter macroprudential regulation. A tighter constraint limits intermediaries' risk exposure more severely, causing the constraint to bind over a wider range of the state space. In addition, I consider a state-dependent constraint $\hat{\zeta}(w)$ that varies with the intermediary wealth share. Specifically, the state-dependent constraint takes the following form:

$$\hat{\zeta}(w) = \begin{cases} 0.10 + 0.05 \cdot \dfrac{w}{a} & \text{if } w \leq a, \\ \infty & \text{if } w > a. \end{cases} \tag{40}$$

This specification captures the countercyclical nature of VaR-based risk management: when intermediary wealth is high, the constraint is relatively loose, but it tightens endogenously as financial conditions deteriorate (a is chosen to match the point in the state space that the constraint of constant 0.15 starts to bind) The unconstrained economy ($\hat{\zeta} \to \infty$) serves as the benchmark for comparison.

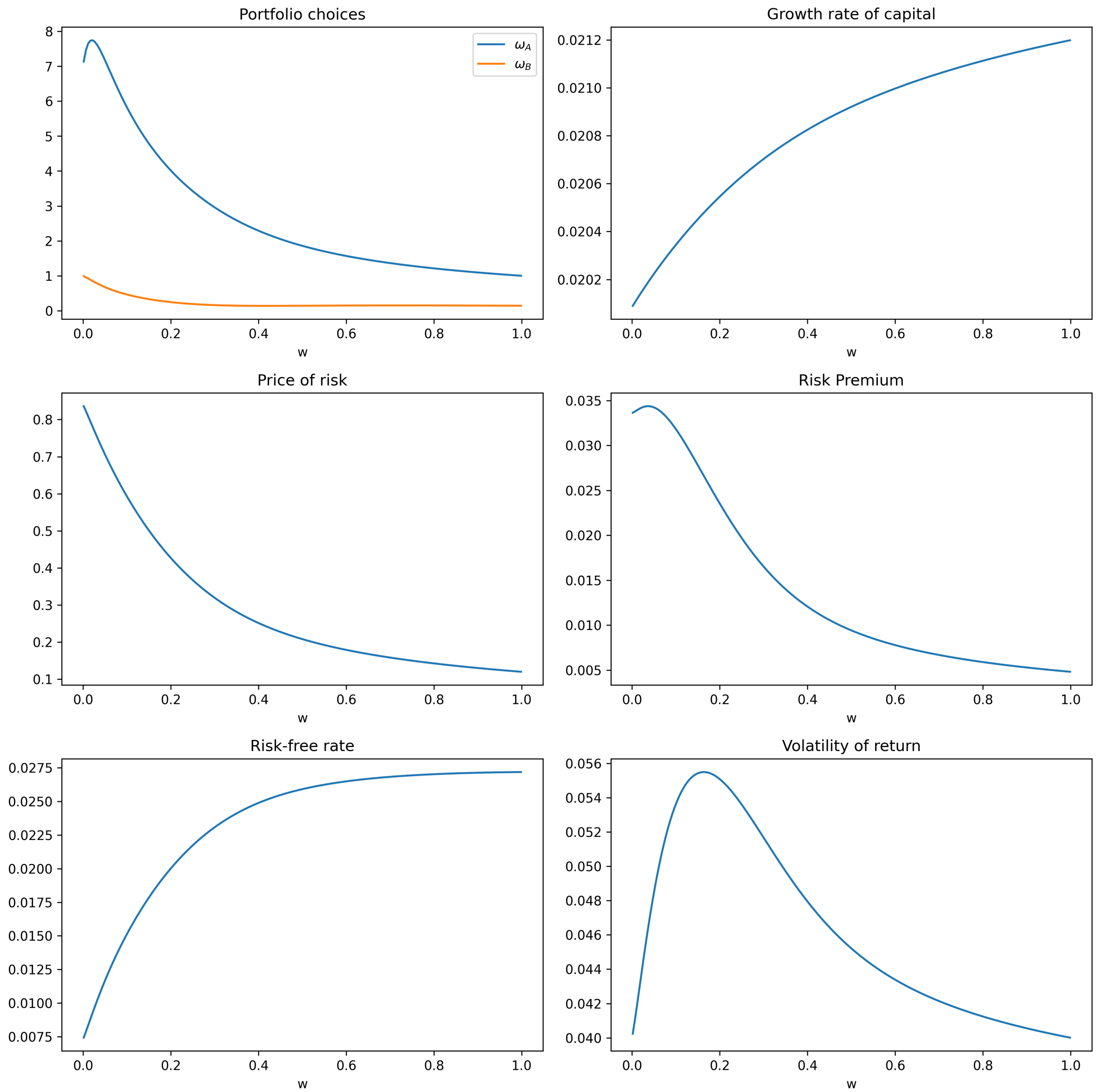
4.3 Model results

I begin by examining how the VaR constraint affects equilibrium objects as functions of the state variable w, the intermediary wealth share. Figure 1 displays the equilibrium of the unconstrained benchmark economy. In this economy, investors freely choose their optimal risk exposure as characterized by (31), determined by their myopic and hedging demands. As the intermediary wealth share w varies, the equilibrium smoothly adjusts: when w is low, aggregate risk aversion (37) is high (reflecting the dominance of households), the market price of risk is elevated (Proposition 5), and the price of capital is depressed. Conversely, when w is high, intermediaries dominate the market, which decreases the market price of risk and raises the price of capital. Through the optimality condition (21), this translates into higher invest-

[11] For instance, the effective discount rate is close to that in Gârleanu and Panageas (2015).

ment and aggregate growth. The volatility of the risky asset shows a hump-shaped pattern, and when the economy is populated by only one type of investor, it coincides with the fundamental volatility. The volatility amplification, consistent with the balance sheet channel, is most pronounced when intermediaries are relatively less capitalized.

Figure 1: Unconstrained benchmark economy



This figure presents the portfolio choices, growth rate of capital, price of risk, risk premium, risk-free rate and volatility of return of the unconstrained benchmark economy. Each quantity is plotted against w, the wealth share of intermediaries.

Figures 2-4 provide comparisons for each constraint scenario against the unconstrained benchmark. Each figure plots the constrained equilibrium (dashed lines) alongside the unconstrained benchmark (solid lines). The parameters are from Table 1, with the only difference being the value of $\hat{\zeta}$. A higher $\hat{\zeta}$ corresponds to a looser constraint.

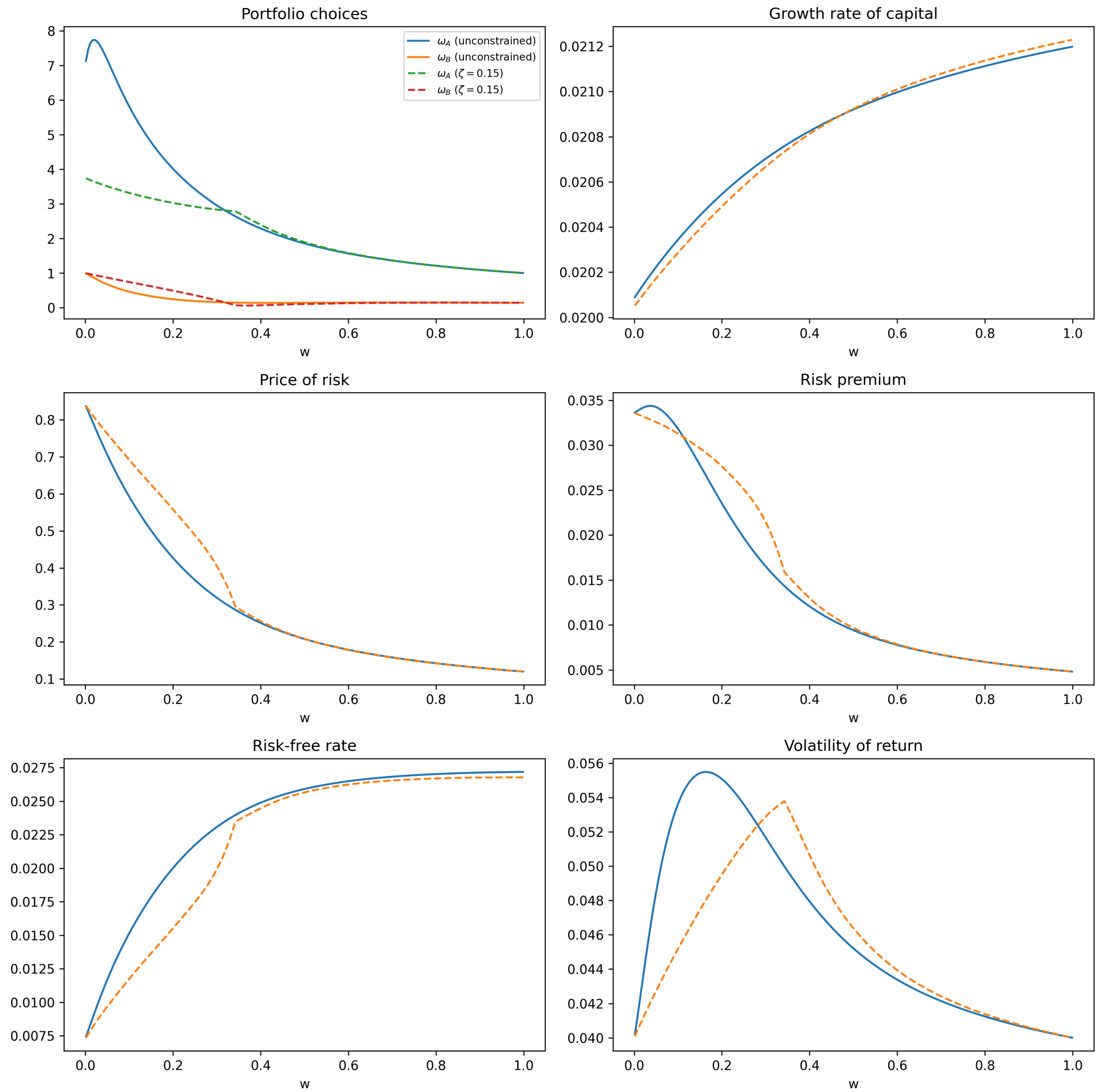
Figure 2 illustrates the effects of a moderately loose VaR constraint ($\hat{\zeta} = 0.15$) relative to the unconstrained benchmark across six equilibrium objects. The risk exposure of intermediaries (σ_A) is capped at $\hat{\zeta}$ once the unconstrained optimum σ_A^* exceeds the constraint, which occurs in the low-w region; households (σ_B) correspondingly take on more risk to clear the market. This can be seen from portfolio shares as well, as shown in the top left panel. The top right panel shows that the constrained economy achieves a lower growth rate of capital when the constraint binds, reflecting the depressed price of capital q and the resulting reduction in investment via (21); at higher values of w where the constraint is slack, growth coincides with the benchmark, or slightly exceeds it.

The middle left panel shows that price of risk rises more sharply in the constrained region when intermediaries are forced to shed risk. During such periods, households must bear more risk, demanding higher compensation. The middle right panel mirrors this pattern: the risk premium on the risky asset widens in the constrained region due to higher η. The bottom left panel shows that risk-free rate declines in the constrained region, reflecting a precautionary savings channel: as intermediaries deleverage and face tighter constraints, the reduced demand for borrowing and increased desire for safe assets push the equilibrium interest rate downward. Finally, the bottom right panel displays that the volatility of returns still behaves in a hump-shape. More importantly, it is much lower in the constrained region: volatility is attenuated when intermediaries are forced to reduce risk exposure. Importantly, all six panels show that deviations from the benchmark are concentrated in the low-w region where the constraint binds.

Across Figures 2, 3 and 4, a pattern emerges: tighter constraints amplify the deviation from the benchmark, particularly in the low-w region. Under $\hat{\zeta} = 0.15$ (Figure 2), the distortions are moderate and concentrated in a narrow region of the state space. As the constraint tightens to $\hat{\zeta} = 0.12$ (Figure 3) and further to $\hat{\zeta} = 0.10$ (Figure 4), the constrained region expands significantly: the market price of risk is elevated over a broader range of w, asset prices are depressed by a larger magnitude, and the growth rate of capital falls more sharply. Notably, the volatility of return is substantially lower under the tighter constraint. However, such decrease in volatility comes at the cost of a higher risk premium. These results illustrate that macroprudential policies designed to limit risk-taking have a quantitatively significant impact on equilibrium outcomes, particularly during periods of intermediary distress.

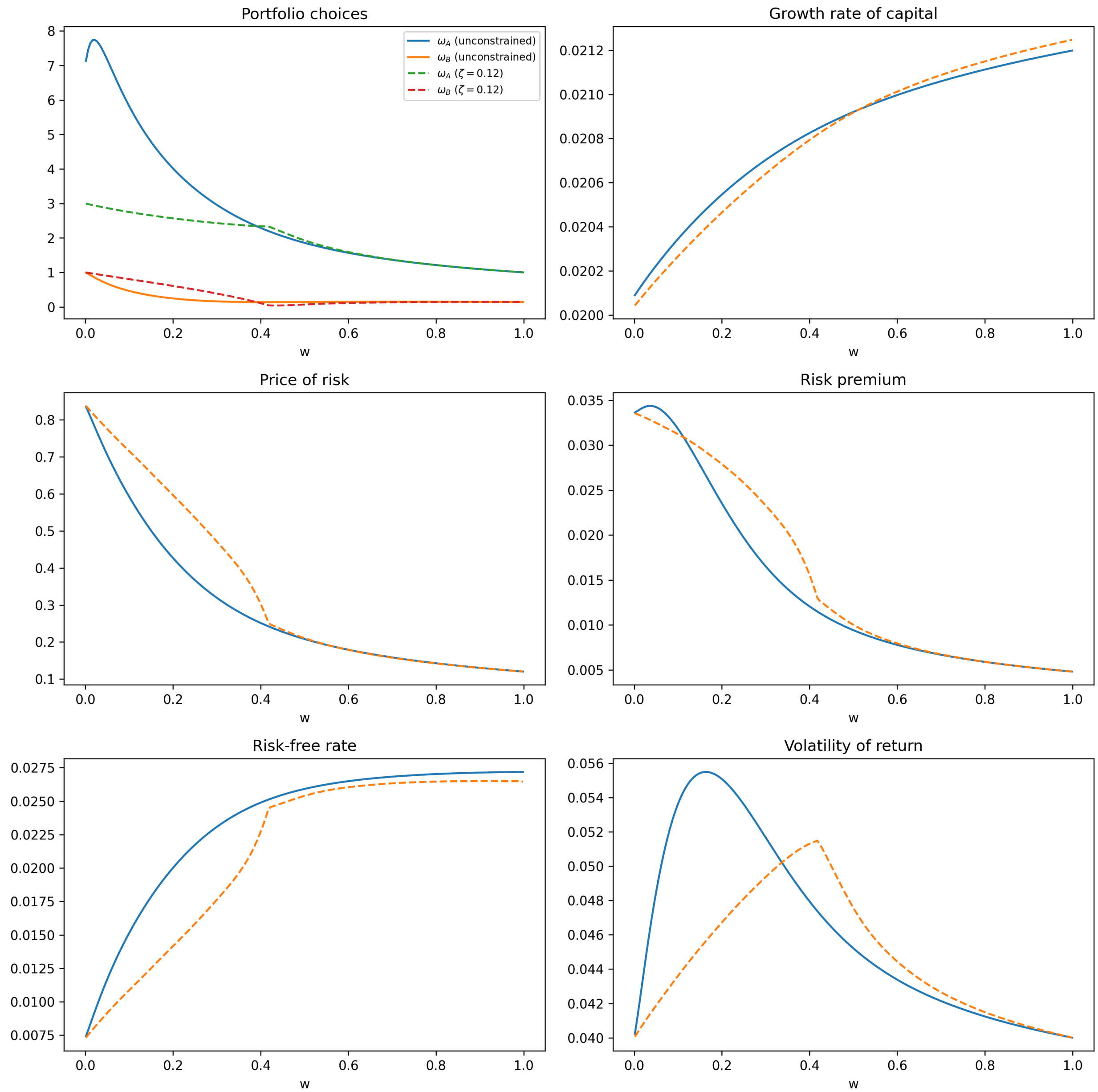
Figure 5 shows the model results when $\hat{\zeta}(w)$ is countercyclical. The leverage (portfolio

Figure 2: Constrained economy ($\hat{\zeta} = 0.15$) vs. unconstrained economy



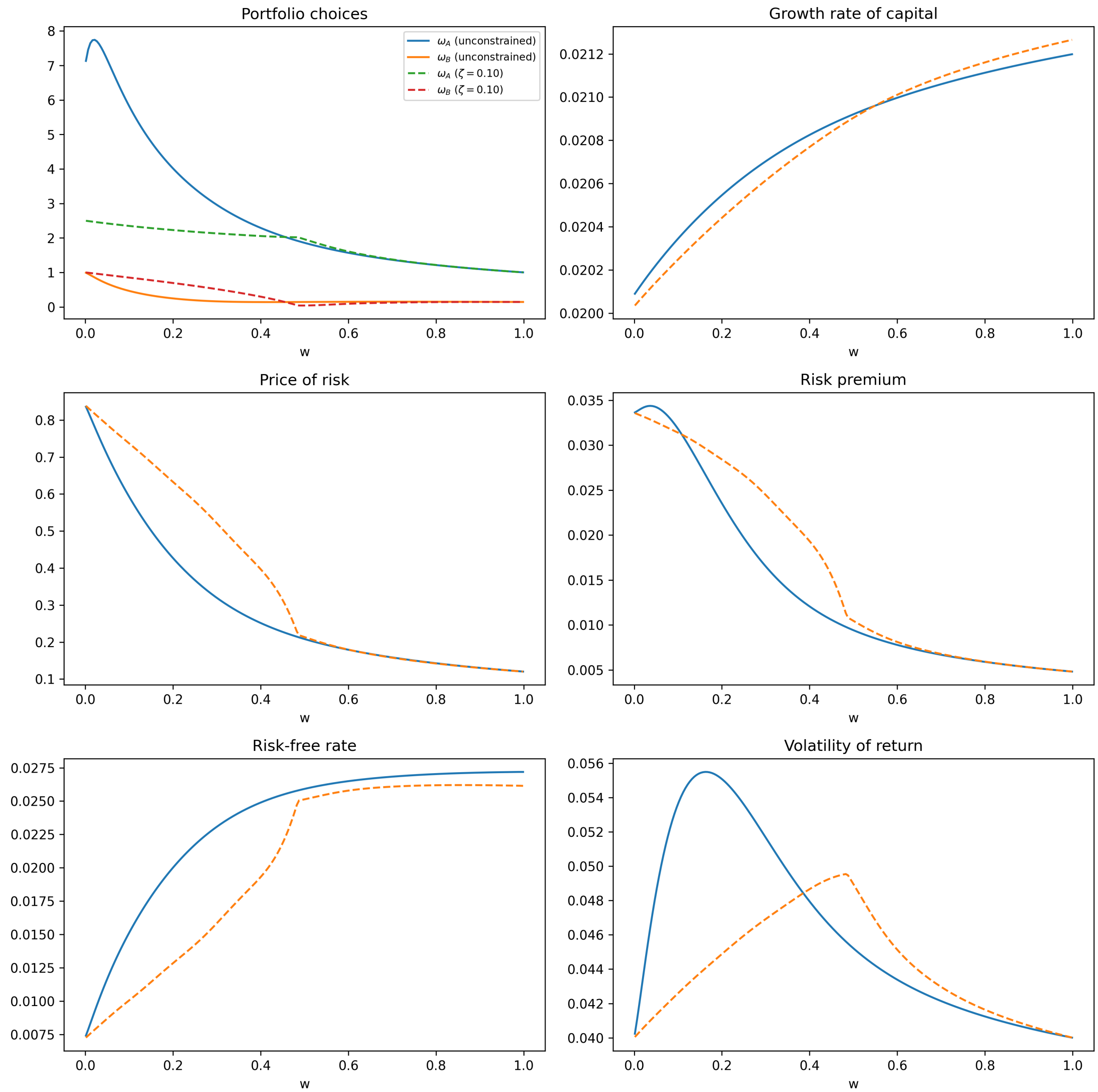
This figure presents the portfolio choices, growth rate of capital, price of risk, risk premium, risk-free rate and volatility of return in constrained ($\hat{\zeta} = 0.15$) and unconstrained economy. Dashed lines represent the constrained economy and solid lines represent the unconstrained economy. Each quantity is plotted against w, the wealth share of intermediaries.

Figure 3: Constrained economy ($\hat{\zeta} = 0.12$) vs. unconstrained economy



This figure presents the portfolio choices, growth rate of capital, price of risk, risk premium, risk-free rate and volatility of return in constrained ($\hat{\zeta} = 0.12$) and unconstrained economy. Dashed lines represent the constrained economy and solid lines represent the unconstrained economy. Each quantity is plotted against w, the wealth share of intermediaries.

Figure 4: Constrained economy ($\hat{\zeta} = 0.10$) vs. unconstrained economy



This figure presents the portfolio choices, growth rate of capital, price of risk, risk premium, risk-free rate and volatility of return in constrained ($\hat{\zeta} = 0.10$) and unconstrained economy. Dashed lines represent the constrained economy and solid lines represent the unconstrained economy. Each quantity is plotted against w, the wealth share of intermediaries.

weight) of intermediaries is weakly procyclical: intermediaries take on more risk when they are wealthier due to the cyclicality of the constraint. The economy behaves like a transition between the unconstrained benchmark and the constraint scenarios. When w is high, the constraint is not binding. As w declines, the constraint starts to bind at a relatively loose level of $\hat{\zeta}$, so the economy first deviates from the benchmark in a manner similar to the $\hat{\zeta} = 0.15$ scenario. As w continues to decline, the constraint tightens further, and the economy behaves more like the $\hat{\zeta} = 0.10$ scenario. In particular, the volatility of the return peaks at a similar level as the unconstrained benchmark.

4.4 Impulse response functions

While the equilibrium functions characterize how the economy behaves at each point in the state space, impulse response functions (IRFs) reveal the dynamic paths following a shock. For each variable of interest, I compute the impulse response following a one standard deviation negative shock. More specifically, the economy is hit contemporaneously by the shock at $t = 0$, and I perform Monte-Carlo simulations of the economy over a horizon of $T = 20$ years to trace the impact of the shock. The expected conditional response at each horizon h is computed by averaging across the simulated trajectories. Such procedure is performed for the economy with no shock at $t = 0$. The relative difference between the two expected paths at each horizon h gives the IRF of the variable of interest. The Monte-Carlo simulations start from the stochastic steady state, defined as the point where the drift of w is zero. Formally, for any equilibrium object V,the IRF at horizon $h \geq 0$ is defined as

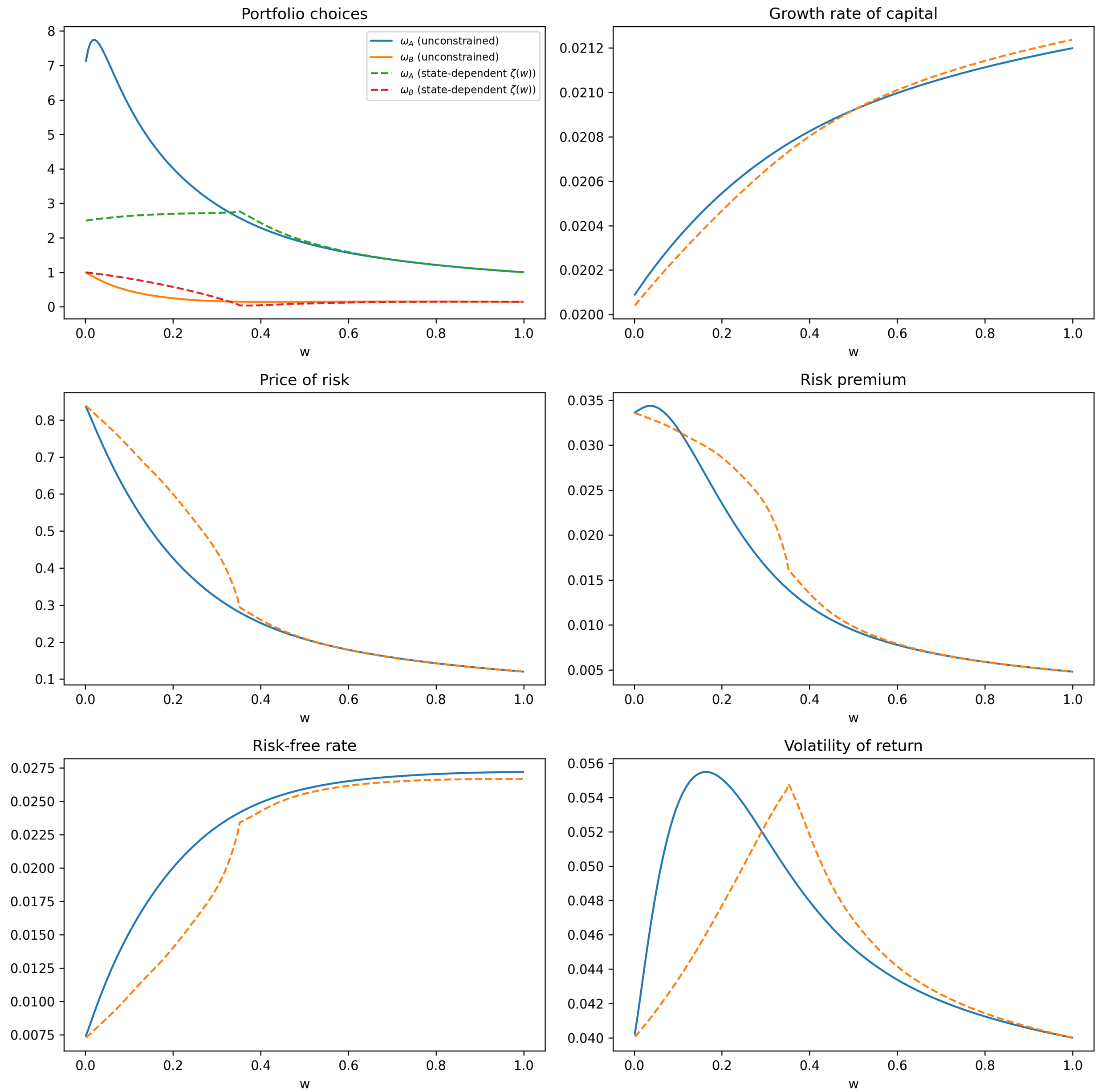
$$\mathrm{IRF}_{t\to T} = \frac{\mathbb{E}[V_T \mid B_t = 1] - \mathbb{E}[V_T \mid B_t = 0]}{\mathbb{E}[V_T \mid B_t = 0]}, \tag{41}$$

where $B_t = 1$ represents the realization of the shock. Note that due to the nonlinearity of the model, the IRFs cannot be calculated by zeroing out future shocks.

Figures 6-8 present the IRFs of key macroeconomic variables for each constrained scenario against the unconstrained benchmark. In each figure, solid lines represent the unconstrained economy and dashed lines represent the constrained economy.

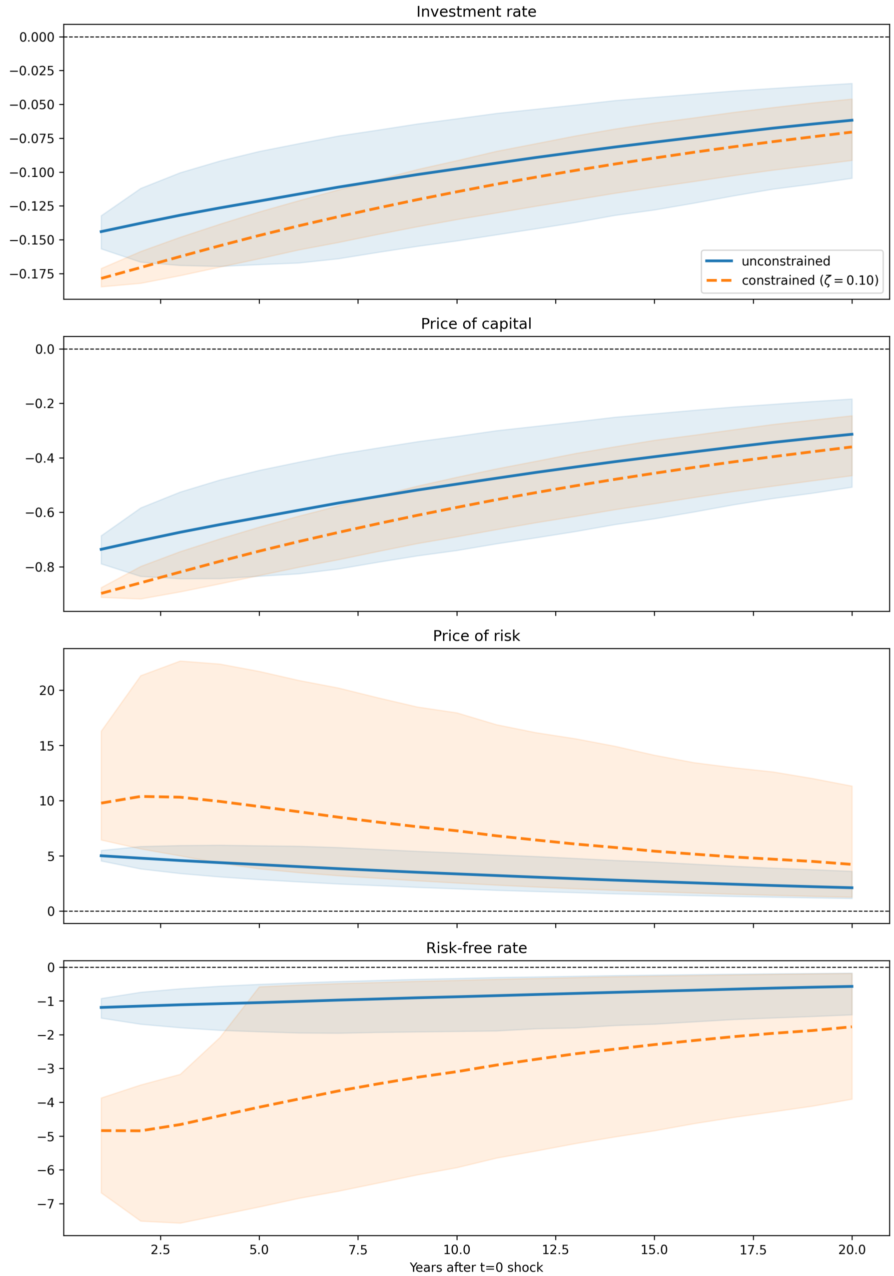
Consider first the constraint scenario of $\hat{\zeta} = 0.10$ (Figure 6). The negative shock erodes intermediaries' net worth, reducing the wealth share w. In the unconstrained economy, intermediaries freely rebalance their portfolios: the investment rate and price of capital q decline, the market price of risk rises, and risk-free rate decreases. After the shock, the economy adjusts smoothly. In the constrained economy, the response is more pronounced on impact and the recovery is more protracted. This is because the VaR constraint occasionally binds as

Figure 5: Constrained economy ($\hat{\zeta}(w)$) vs. unconstrained economy



This figure presents the portfolio choices, growth rate of capital, price of risk, risk premium, risk-free rate and volatility of return in constrained ($\hat{\zeta}(w)$) and unconstrained economy. Dashed lines represent the constrained economy and solid lines represent the unconstrained economy. Each quantity is plotted against w, the wealth share of intermediaries.

Figure 6: Impulse response functions: $\hat{\zeta} = 0.10$ vs. unconstrained economy



This figure presents the impulse response functions of key equilibrium variables in the constrained ($\hat{\zeta} = 0.10$) and the unconstrained economy. Dashed lines represent the constrained economy and solid lines represent the unconstrained economy. Each variable is plotted against time (in years) after a one-standard-deviation negative shock.

intermediaries' risk exposure reaches $\hat{\zeta}$, forcing them to deleverage. Households thus absorb more risk and demand higher compensation, which amplifies the increase in the market price of risk. This amplification, as a form of *financial fire sale*, is particularly persistent as can be seen from the third panel of Figure 6.

The risk-free rate drops more sharply in the constrained economy, reflecting the higher precautionary savings motive: constrained intermediaries reduce their demand for borrowing, pushing the equilibrium interest rate downward. Given the EIS of $\psi > 1$, the decline in the risk-free rate does not fully offset the increase in the price of risk, so the price of capital falls more sharply in the constrained economy.

As the constraint loosens to $\hat{\zeta} = 0.12$ (Figure 7), the qualitative patterns are preserved but the amplification is larger on impact, as the constraint binds over a narrower region of the state space. The point where the constraint binds is more extreme, so the shock has a more severe impact on the price of risk. However, as the economy adjusts and w recovers, the constraint becomes slack sooner, so the amplification is not as persistent as in the $\hat{\zeta} = 0.10$ scenario.

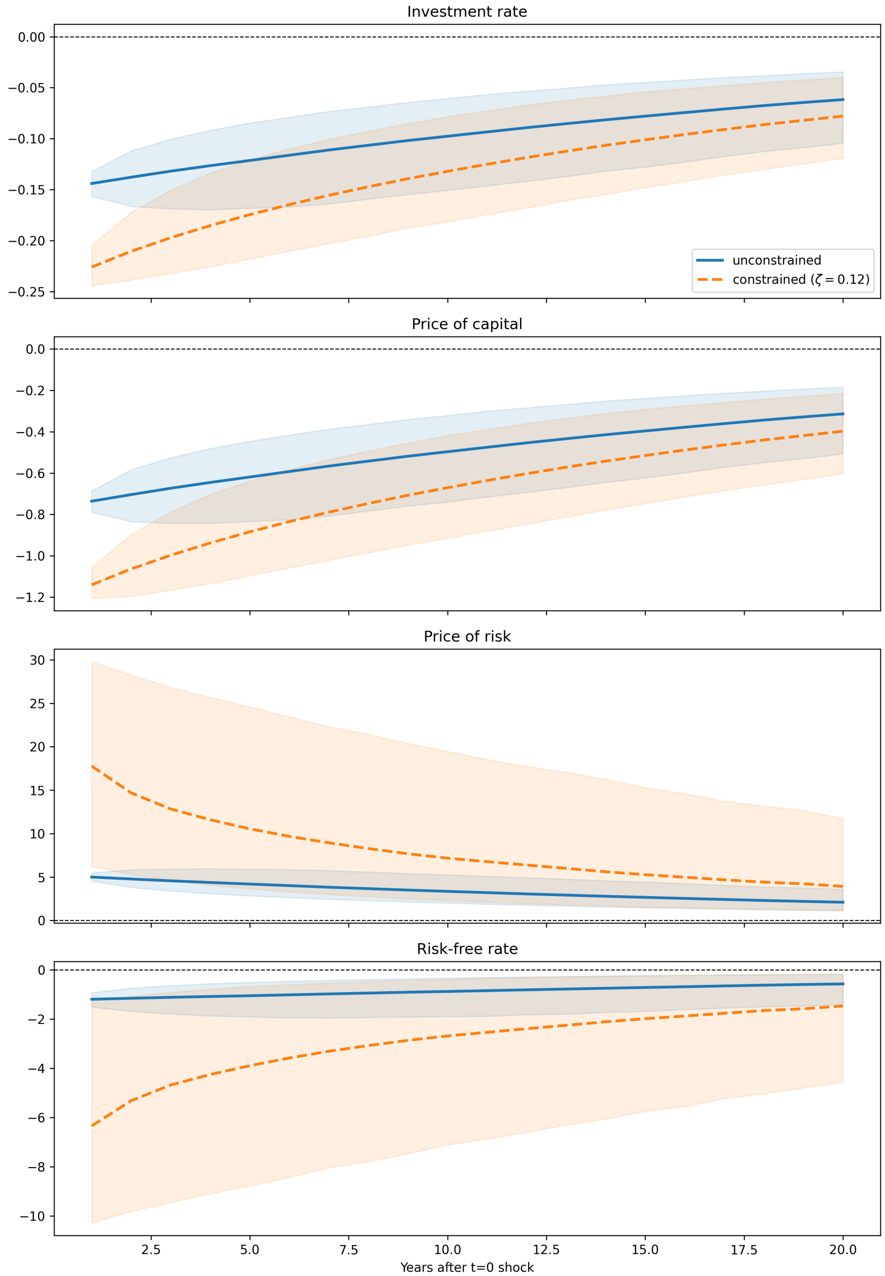
Under the loosest constraint, $\hat{\zeta} = 0.15$ (Figure 8), the IRFs of the constrained economy are much more aligned with the unconstrained benchmark. The constraint only binds briefly after the shock, and the economy quickly transitions to the unconstrained region. The amplification on impact is smaller, and the recovery is faster. The non-monotone pattern of the amplification is particularly evident in the IRF of price of risk.

The state-dependent scenario is illustrated in Figure 9. The amplification and persistence lies in between the previous scenarios. The response on impact is pronounced as the constraint starts to bind at a relatively loose level, but the recovery is faster as the constraint becomes tighter. Taken together, the IRFs highlight an important asymmetry in how macroprudential regulation affects the economy. The VaR constraint has negligible effects during expansions when intermediaries are well-capitalized and the constraint is slack. During downturns, however, the constraint amplifies and prolongs the transition process: the forced deleveraging depresses asset prices, and reduces investment further. The severity of this amplification is not monotonic in the tightness of the constraint.

5 Financial crises

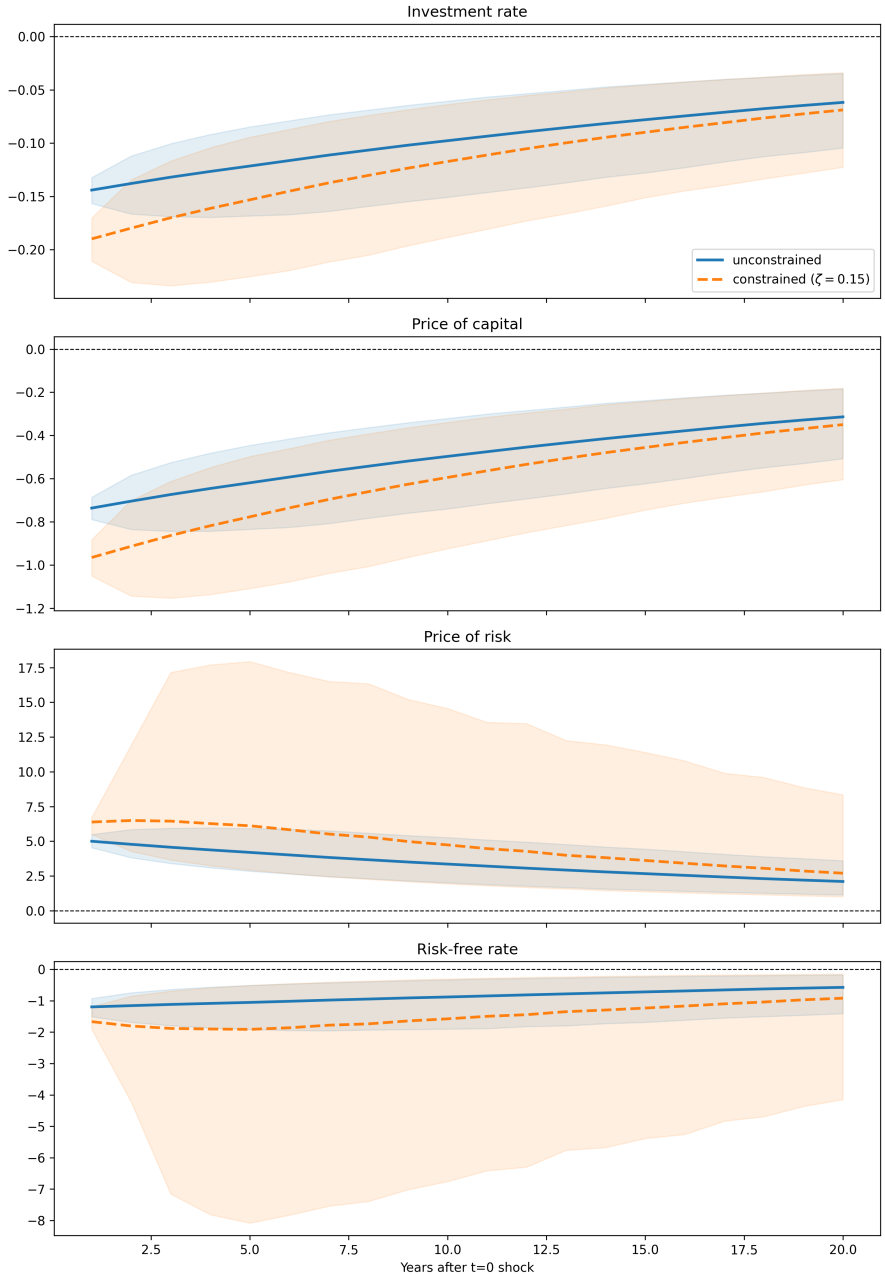
In this section, I explore the model's implications for crises and macro-stability. I first examine the stationary distribution of growth rates across VaR constraint scenarios, characterizing how these constraints reshape tail risk. Then, I analyze the frequency and duration of crises under different VaR constraints, comparing them to historical data.

Figure 7: Impulse response functions: $\hat{\zeta} = 0.12$ vs. unconstrained economy



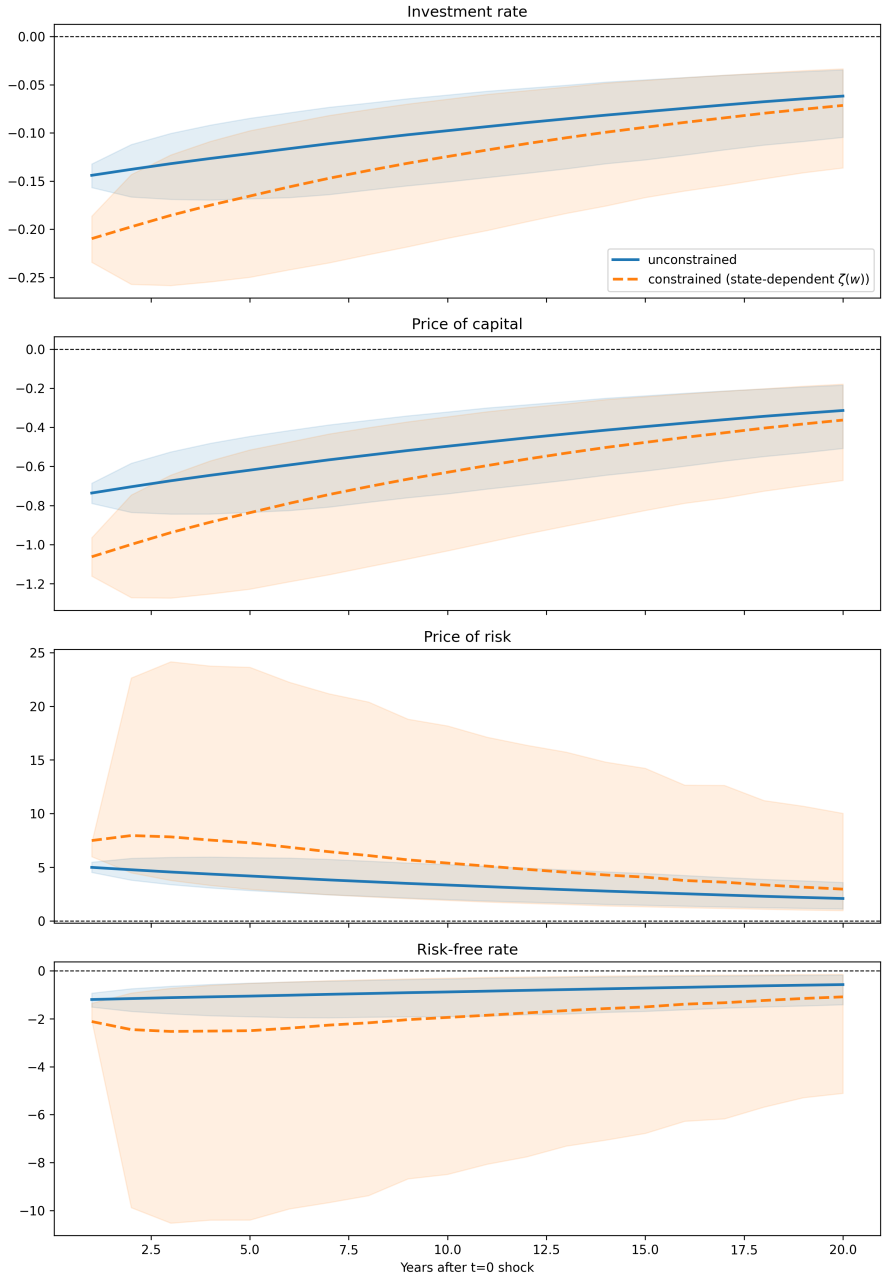
This figure presents the impulse response functions of key equilibrium variables in the constrained ($\hat{\zeta} = 0.12$) and the unconstrained economy. Dashed lines represent the constrained economy and solid lines represent the unconstrained economy. Each variable is plotted against time (in years) after a one-standard-deviation negative shock.

Figure 8: Impulse response functions: $\hat{\zeta} = 0.15$ vs. unconstrained economy



This figure presents the impulse response functions of key equilibrium variables in the constrained ($\hat{\zeta} = 0.15$) and the unconstrained economy. Dashed lines represent the constrained economy and solid lines represent the unconstrained economy. Each variable is plotted against time (in years) after a one-standard-deviation negative shock.

Figure 9: Impulse response functions: state dependent $\hat{\zeta}(w)$ vs. unconstrained economy



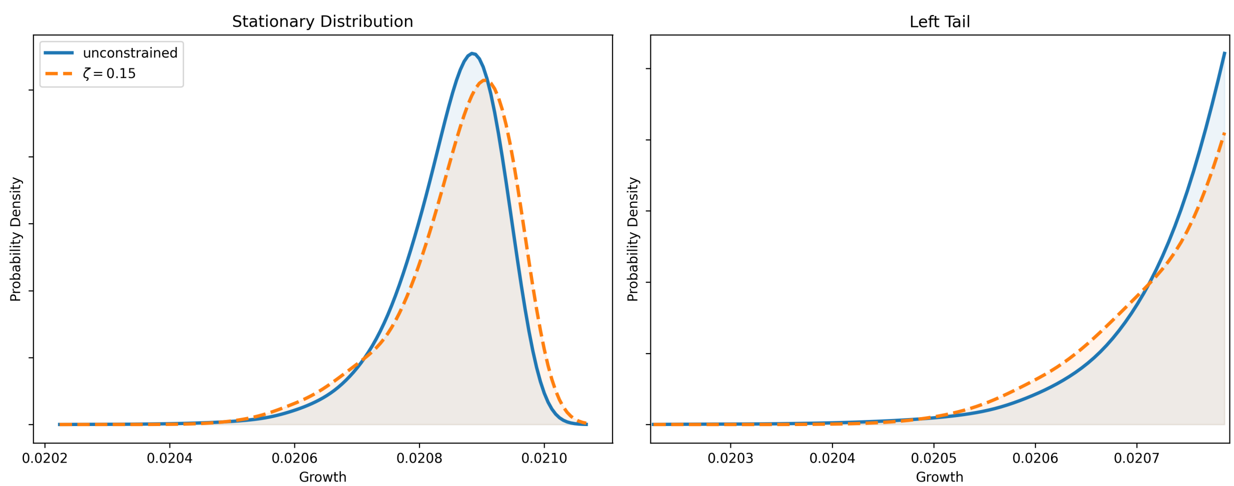
This figure presents the impulse response functions of key equilibrium variables in the constrained (state-dependent$\hat{\zeta}(w)$) and the unconstrained economy. Dashed lines represent the constrained economy and solid lines represent the unconstrained economy. Each variable is plotted against time (in years) after a one-standard-deviation negative shock.

5.1 Stability and tail risk

To assess the implications of VaR constraints for macroeconomic stability, I simulate the economy under each scenario and examine the stationary distribution of growth rates. While the equilibrium dynamics characterize how the economy behaves at each point in the state space, the simulated distributions reveal where the economy spends most of its time and the frequency and severity of adverse outcomes.

Figures 10-13 display the probability density functions of the simulated growth rate for each constrained scenario, compared with the unconstrained benchmark. Under $\hat{\zeta} = 0.15$ (Figure 10), the distribution of the constrained economy is modestly wider and less peaked than the benchmark. The right tail is thicker while the left tail is slightly thinner, reflecting the fact that the constraint limits the extent of leveraging during downturns. As the constraint tightens to $\hat{\zeta} = 0.12$ (Figure 11) and $\hat{\zeta} = 0.10$ (Figure 12), the reshaping becomes more pronounced: the density flattens near the mode, the left tail compresses visibly, and the overall distribution widens. Across all three constant-$\hat{\zeta}$ scenarios, the mode of the growth rate distribution shifts only slightly, consistent with the fact that the constraint is slack during normal times and primarily affects the tails. The state-dependent constraint (Figure 13) produces a density that lies between the $\hat{\zeta} = 0.15$ and $\hat{\zeta} = 0.12$ cases, reflecting the fact that $\hat{\zeta}(w)$ transitions smoothly from a looser to a tighter regime as w declines.

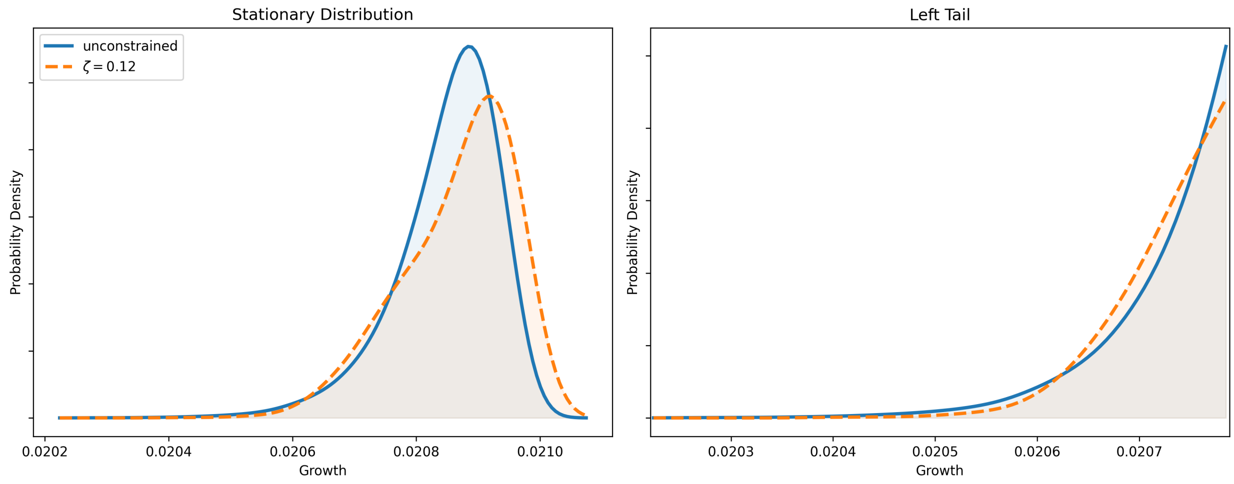
Figure 10: Growth rate distribution: $\hat{\zeta} = 0.15$ vs. unconstrained economy



This figure shows the simulated probability density function of the growth rate of capital under $\hat{\zeta} = 0.15$ (dashed line) compared with the unconstrained benchmark (solid line).

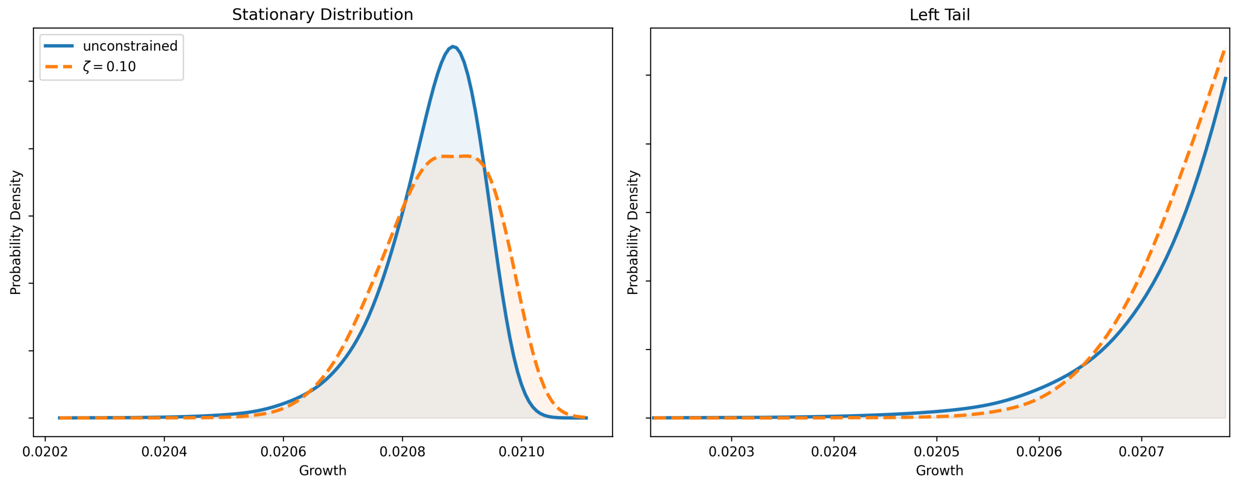
Table 2 quantifies these distributional shifts by reporting the summary statistics of the stationary distribution of $\log(q_t)$, the log price of capital. Since q is the key equilibrium object

Figure 11: Growth rate distribution: $\hat{\zeta} = 0.12$ vs. unconstrained economy



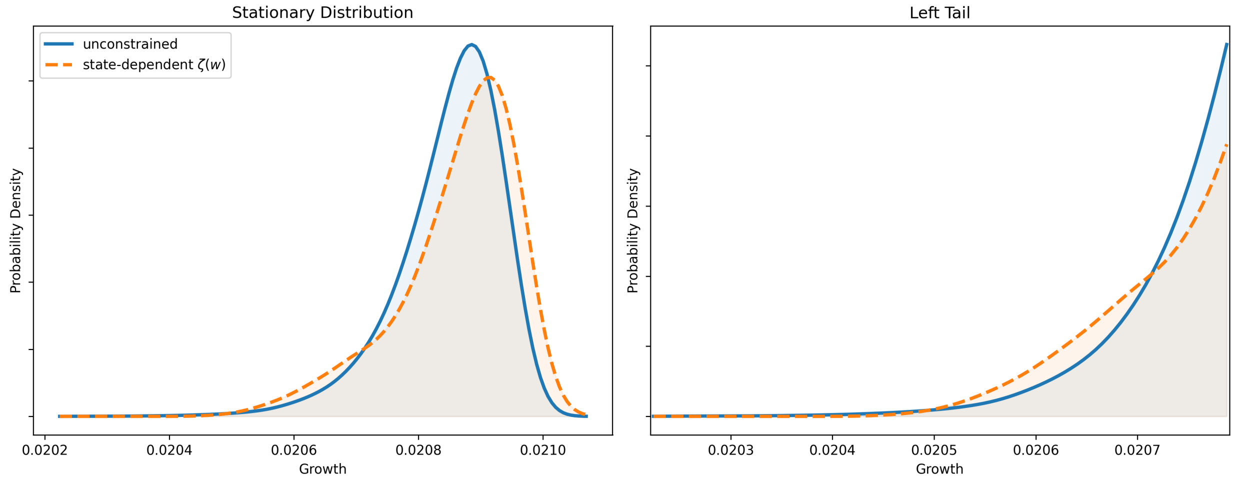
This figure shows the simulated probability density function of the growth rate of capital under $\hat{\zeta} = 0.12$ (dashed line) compared with the unconstrained benchmark (solid line).

Figure 12: Growth rate distribution: $\hat{\zeta} = 0.10$ vs. unconstrained economy



This figure shows the simulated probability density function of the growth rate of capital under $\hat{\zeta} = 0.10$ (dashed line) compared with the unconstrained benchmark (solid line).

Figure 13: Growth rate distribution: state-dependent $\hat{\zeta}(w)$ vs. unconstrained economy



This figure shows the simulated probability density function of the growth rate of capital under the state-dependent constraint $\hat{\zeta}(w)$ (dashed line) compared with the unconstrained benchmark (solid line).

linking financial conditions to investment via the optimality condition (21), shifts in the distribution of $\log(q_t)$ map directly into changes in the distribution of growth rates. Therefore, the moments of $\log(q_t)$ provide a direct measure of financial stability across different policy regimes.

Table 2: Summary statistics of stationary distributions for $\log(q_t)$

	Mean	Std (%)	Skewness	Kurtosis
unconstrained	1.818	2.487	-1.021	4.750
$\zeta = 0.15$	1.821	2.783	-0.911	3.739
$\zeta = 0.12$	1.823	2.794	-0.617	3.009
$\zeta = 0.10$	1.821	2.760	-0.307	2.685
state-dependent	1.822	2.890	-0.885	3.500

This table reports the mean, standard deviation, skewness, and kurtosis of stationary distributions for $\log(q_t)$ across different VaR constraint scenarios. The unconstrained economy corresponds to $\hat{\zeta} \to \infty$.

The moment statistics in Table 2 are consistent with the patterns from the density plots. First, the mean of $\log(q_t)$ is nearly identical across all scenarios (approximately 1.82), indicating that the VaR constraint does not substantially alter the average level of asset valuations. This is because the constraint mainly binds during downturns (low w) and is slack during normal times.

Second, and more importantly, the *shape* of the distribution varies substantially across scenarios, consistent with the density plots above. The skewness and kurtosis are particularly informative. The unconstrained economy exhibits the most negative skewness (-1.021) and the highest kurtosis (4.750), reflecting a distribution of asset valuations with a heavy left tail and sharp peak. This arises because even without binding constraints, the wealth distribution dynamics (Proposition 1) and recursive preferences generate endogenous nonlinearity.

As the VaR constraint becomes tighter, the skewness of the distribution becomes less negative: from -1.021 (unconstrained) to -0.911 ($\hat{\zeta} = 0.15$), -0.617 ($\hat{\zeta} = 0.12$), and -0.307 ($\hat{\zeta} = 0.10$). Simultaneously, kurtosis declines from 4.750 to 3.739, 3.009, and 2.685, respectively. These shifts indicate that tighter constraints reduce the concentration of the distribution near the peak and compress the tails, as visible in Figures 10-12. The state-dependent constraint produces intermediate results: a skewness of -0.885 and kurtosis of 3.50, broadly between the $\hat{\zeta} = 0.15$ and $\hat{\zeta} = 0.12$ scenarios. This reflects the countercyclical nature of the state-dependent constraint.

Hence, the system presents less of an asymmetric response to shocks and the constraint reduces skewness and kurtosis. The intuition is that the VaR constraint limits the extent to which intermediaries can lever up and forces deleveraging during bad times. Therefore, if we measure financial stability by left-tail risk, we can conclude that the VaR constraints enhance financial stability. In the meantime, the average price does not change too much, thus the average growth rate does not change too much either. Hence, the VaR constraint mainly leads to a reallocation of probability mass from the left tail to the center and right tail.

The standard deviation tells a complementary story. The unconstrained economy has the lowest standard deviation (2.487%), while all constrained scenarios exhibit higher dispersion, with the state-dependent constraint producing the highest (2.890%). This seemingly paradoxical finding—that constraints intended to compress risk-taking actually increase the variation of asset valuations—is a direct consequence of the amplification mechanism. When constraints bind, the forced deleveraging depresses asset prices and investment, magnifying the impact of the underlying shock on q. The state-dependent constraint amplifies this effect further by tightening precisely when the economy is most vulnerable. Taken together, the density plots and moment statistics show that macroprudential policies, such as VaR constraint, reshape the distribution of macroeconomic outcomes, especially at the tail. These results highlight a tension in macroprudential regulation. While tighter VaR constraints reduce the left-skewness and fat tails of the distribution of asset valuations and growth, they do so at the cost of higher overall volatility and lower investment and growth during periods of distress.

5.2 Crisis

The analysis above characterizes the unconditional properties of growth rates and asset valuations. I now turn to the dynamics conditional on crisis events and examining their frequency, and duration across VaR constraint scenarios.

I define a crisis as an episode during which the growth rate falls below a threshold g^*, chosen as the 5th percentile of the stationary distribution of g in the unconstrained economy.[12] A crisis episode begins when g_t first crosses below g^* and ends when g_t returns above g^*. Formally, the crisis indicator is

$$\text{Crisis}_t = \mathbf{1}\{g_t < g^*\}. \tag{42}$$

To compute the probability and duration of the crisis, I simulate long sample paths of the economy (100,000 months for 50 parallel paths) under each VaR constraint scenario and identify all crisis episodes. For each scenario, I report the following statistics: (i) the unconditional crisis probability, defined as the fraction of time spent in the crisis region; (ii) the crisis duration, computed as the average length (in months) of each crisis episode.

Table 3: Crisis probability and duration

	Data	Unconstrained	$\hat{\zeta} = 0.15$	$\hat{\zeta} = 0.12$	$\hat{\zeta} = 0.10$	$\hat{\zeta}(w)$
Probability (%)	7.00	5.00	6.22	5.85	4.41	7.88
Duration (months)	17.50	13.86	14.01	14.31	14.39	14.88

This table reports crisis probability and (mean) duration in months across different VaR constraint scenarios. Data for computing the empirical duration of crisis is from NBER website. The model implied probability and duration are obtained from simulating each of the models at monthly frequency.

Table 3 reports the statistics across scenarios. Several patterns emerge. The crisis probability decreases as the constant constraint tightens: from 6.22% under the loosest constraint ($\hat{\zeta} = 0.15$) to 5.85% ($\hat{\zeta} = 0.12$) and 4.41% ($\hat{\zeta} = 0.10$). In each case, the tighter constraint attenuates the volatility of w, making it less likely for the economy to drift into the crisis region. The unconstrained economy, at 5.00%, lies between the loosest and intermediate constrained scenarios. Introducing a loose constraint actually raises crisis probability relative to the unconstrained benchmark: the amplification mechanism from binding constraints during downturns depresses asset prices and growth enough to push the economy into crisis more frequently, while the constraint is not tight enough to substantially reduce the volatility of w in normal times. Only when the constraint is sufficiently tight ($\hat{\zeta} = 0.10$) does the volatility-suppression effect dominate, bringing the crisis probability below the unconstrained level.

[12] This threshold-based definition is standard in the macro-finance literature.

The state-dependent constraint produces the highest crisis probability at 7.88%, consistent with the fact that it generates the largest standard deviation of $\log(q_t)$ (Table 2): by tightening precisely when the economy is most vulnerable, it amplifies adverse shocks enough to push the growth rate below g^* more frequently.

The mean crisis duration increases with constraint tightness: from 13.86 months in the unconstrained economy to 14.01 ($\hat{\zeta} = 0.15$), 14.31 ($\hat{\zeta} = 0.12$), and 14.39 months ($\hat{\zeta} = 0.10$). The state-dependent constraint produces the longest average duration at 14.88 months. These patterns are consistent with the IRF analysis in Section 4: tighter constraints slow the recovery of intermediary wealth by forcing deleveraging at precisely the moments when risk absorption would accelerate the return to normal conditions. Notably, while the differences in duration are modest in absolute terms, they compound over repeated crisis episodes and contribute to the higher unconditional volatility documented in Table 2.

Compared to the data, where the crisis probability is 7.00% and mean duration is 17.50 months, the model broadly captures the right order of magnitude. The unconstrained economy understates both the frequency and persistence of crises, while the state-dependent constraint comes closest to matching the empirical crisis probability. All model scenarios produce durations somewhat shorter than the data, suggesting that additional frictions may be needed to fully account for the observed persistence of downturns. Nevertheless, the cross-scenario variation confirms a key trade-off: tighter constant constraints can reduce crisis frequency at the cost of longer crises, whereas the state-dependent constraint increases both frequency and duration, highlighting the importance of constraint design for macroeconomic stability.

6 Conclusion

This paper develops a macro-finance model to study how VaR constraints on financial intermediaries affect aggregate investment and economic growth through the balance sheet channel. When constraints bind, forced deleveraging amplifies the impact of adverse shocks on asset prices, risk premia, and investment, widening the gap between constrained and unconstrained equilibrium outcomes.

The quantitative analysis reveals a central tension in macroprudential regulation. Tighter VaR constraints reduce the left-skewness and excess kurtosis of the stationary distribution of asset valuations—mitigating tail risk—but simultaneously increase overall volatility, as forced deleveraging magnifies the impact of aggregate shocks on the price of capital. Impulse response functions confirm this asymmetry dynamically: the constraint has negligible effects during expansions but amplifies and prolongs the adjustment following adverse shocks. The crisis analysis sharpens this tension: tighter constant constraints reduce crisis frequency but increase average duration, since forced deleveraging slows the recovery of intermediary balance sheets. The state-dependent constraint, despite producing intermediate distributional outcomes, generates the highest crisis probability by amplifying shocks precisely when the economy is most vulnerable. Compared to NBER-dated recession statistics, the model captures the right order of magnitude for both crisis probability and duration.

These findings speak to the design of post-crisis regulatory frameworks such as Basel III, where risk-weighted capital requirements function similarly to the VaR limit in the model. The results highlight the importance of constraint design—in particular, the potential benefits of state-dependent regulation that loosens during distress to avoid exacerbating the amplification mechanism. The comparison across scenarios also reveals a nonlinear relationship between constraint tightness and outcomes: moderate tightening produces substantial reductions in tail risk with only modest increases in volatility, while further tightening yields diminishing improvements. Formally characterizing the optimal constraint—for example by maximizing a social welfare criterion that penalizes both tail risk and volatility—and incorporating additional frictions to match the observed persistence of downturns are promising directions for future work.

References

Adrian, T. and Shin, H. S. (2014). Procyclical Leverage and Value-at-Risk. *The Review of Financial Studies*, 27(2):373–403.

Basak, S. and Cuoco, D. (1998). An Equilibrium Model with Restricted Stock Market Participation. *The Review of Financial Studies*, 11(2):309–341.

Bernanke, B. and Gertler, M. (1989). Agency Costs, Net Worth, And Business Fluctuations. *The American Economic Review*, 79(1):14.

Bernanke, B. S., Gertler, M., and Gilchrist, S. (1999). Chapter 21 The financial accelerator in a quantitative business cycle framework. In *Handbook of Macroeconomics*, volume 1, pages 1341–1393. Elsevier.

Bhamra, H. S. and Uppal, R. (2009). The Effect of Introducing a Non-Redundant Derivative on the Volatility of Stock-Market Returns When Agents Differ in Risk Aversion. *The Review of Financial Studies*, 22(6):2303–2330.

Bhamra, H. S. and Uppal, R. (2014). Asset Prices with Heterogeneity in Preferences and Beliefs. *Review of Financial Studies*, 27(2):519–580.

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

Brunnermeier, M. K. and Sannikov, Y. (2016). Chapter 18 - Macro, Money, and Finance: A Continuous-Time Approach. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2, pages 1497–1545. Elsevier.

Chabakauri, G. (2013). Dynamic Equilibrium with Two Stocks, Heterogeneous Investors, and Portfolio Constraints. *The Review of Financial Studies*, 26(12):3104–3141.

Chan, Y. L. and Kogan, L. (2002). Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices. *Journal of Political Economy*, 110(1255–1285).

Danielsson, J., Shin, H. S., and Zigrand, J. (2012). Procyclical Leverage and Endogenous Risk. *Working Paper*.

Di Tella, S. (2017). Uncertainty Shocks and Balance Sheet Recessions. *Journal of Political Economy*, 125(6):2038–2081.

Drechsler, I., Savov, A., and Schnabl, P. (2018). A Model of Monetary Policy and Risk Premia. *The Journal of Finance*, 73(1):317–373.

Duffie, D. and Epstein, L. G. (1992a). Asset Pricing with Stochastic Differential Utility. *The Review of Financial Studies*, 5(3):411–436.

Duffie, D. and Epstein, L. G. (1992b). Stochastic Differential Utility. *Econometrica*, 60(2):353–394.

Dumas, B. (1989). Two-Person Dynamic Equilibrium in the Capital Market. *The Review of Financial Studies*, 2(2):157–188.

Dumas, B. (1992). Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World. *The Review of Financial Studies*, 5(2):153–180.

Epstein, L. G. and Zin, S. E. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, 57(4):937–969.

Gârleanu, N. and Panageas, S. (2015). Young, Old, Conservative, and Bold: The Implications of Heterogeneity and Finite Lives for Asset Pricing. *Journal of Political Economy*, 123(3):1979–2010.

Gârleanu, N. and Pedersen, L. H. (2011). Margin-based Asset Pricing and Deviations from the Law of One Price. *The Review of Financial Studies*, 24(6):1980–2022.

Gopalakrishna, G. (2023). A Macro-Finance Model with Realistic Crisis Dynamics. *Working Paper*.

Hayashi, F. (1982). Tobin's Marginal q and Average q: A Neoclassical Interpretation. *Econometrica*, 50(1):213–224.

He, Z. and Krishnamurthy, A. (2012). A Model of Capital and Crises. *The Review of Economic Studies*, 79(2):735–777.

He, Z. and Krishnamurthy, A. (2013). Intermediary Asset Pricing. *American Economic Review*, 103(2):732–770.

Jermann, U. J. (1998). Asset pricing in production economies. *Journal of Monetary Economics*, 41(2):257–275.

Kargar, M. (2021). Heterogeneous intermediary asset pricing. *Journal of Financial Economics*, 141(2):505–532.

Kiyotaki, N. and Moore, J. (1997). Credit cycles. *Journal of Political Economy*, 105(2):211–248.

Kogan, L. (2001). An equilibrium model of irreversible investment. *Journal of Financial Economics*, 62(2):201–245.

Kogan, L. (2004). Asset prices and real investment. *Journal of Financial Economics*, 73(3):411–431.

Kreps, D. M. and Porteus, E. L. (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica*, 46(1):185–200.

Krishnamurthy, A. and Li, W. (2025). Dissecting Mechanisms of Financial Crises: Intermediation and Sentiment. *Journal of Political Economy*, 133(3):935–985.

Longstaff, F. A. and Wang, J. (2012). Asset Pricing and the Credit Market. *The Review of Financial Studies*, 25(11):3169–3215.

Maxted, P. (2024). A Macro-Finance Model with Sentiment. *The Review of Economic Studies*, 91(1):438–475.

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, 41(5):867–887.

Panageas, S. (2020). The Implications of Heterogeneity and Inequality for Asset Pricing. *Foundations and Trends® in Finance*, 12(3):199–275.

Pindyck, R. S. and Wang, N. (2013). The Economic and Policy Consequences of Catastrophes. *American Economic Journal: Economic Policy*, 5(4):306–339.

Rytchkov, O. (2014). Asset Pricing with Dynamic Margin Constraints. *The Journal of Finance*, 69(1):405–452.

Schneider, A. (2022). Risk-Sharing and the Term Structure of Interest Rates. *The Journal of Finance*, 77(4):2331–2374.

Silva, D. H. (2020). The Risk Channel of Unconventional Monetary Policy. *Working Paper*.

Wang, J. (1996). The term structure of interest rates in a pure exchange economy with heterogeneous investors. *Journal of Financial Economics*, 41(1):75–110.

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics*, 24(3):401–421.