# A Macroeconomic Model with a Financial Sector*

Markus K. Brunnermeier and Yuliy Sannikov†

April 8, 2012

**Abstract**

This paper studies the full equilibrium dynamics of an economy with financial frictions. Due to highly non-linear amplification effects, the economy is prone to instability and occasionally enters volatile episodes. Risk is endogenous and asset price correlations are high in down turns. In an environment of low exogenous risk experts assume higher leverage making the system more prone to systemic volatility spikes - a volatility paradox. Securitization and derivatives contracts leads to better sharing of exogenous risk but to higher endogenous systemic risk. Financial experts may impose a negative externality on each other and the economy by not maintaining adequate capital cushion.

*We thank Nobu Kiyotaki, Hyun Shin, Thomas Philippon, Ricardo Reis, Guido Lorenzoni, Huberto Ennis, V. V. Chari, Simon Potter, Emmanuel Farhi, Monika Piazzesi, Simon Gilchrist and seminar participants at Princeton, HKU Theory Conference, FESAMES 2009, Tokyo University, City University of Hong Kong, University of Toulouse, University of Maryland, UPF, UAB, CUFE, Duke, NYU 5-star Conference, Stanford, Berkeley, San Francisco Fed, USC, UCLA, MIT, University of Wisconsin, IMF, Cambridge University, Cowles Foundation, Minneapolis Fed, New York Fed, University of Chicago, the Bank of Portugal Conference, Econometric Society World Congress in Shanghai, Seoul National University, European Central Bank and UT Austin. We also thank Wei Cui, Ji Huang, Andrei Rachkov and Martin Schmalz for excellent research assistance.

†Brunnermeier: Department of Economics, Princeton University, markus@princeton.edu, Sannikov: Department of Economics, Princeton University, sannikov@gmail.com

# 1 Introduction

Economists such as Fisher (1933), Keynes (1936) and Minsky (1986) have attributed the economic downturn of the Great Depression to the failure of financial markets. Kindleberger (1993) documents that financial crises are common in history - having occurred at roughly 10-year intervals in Western Europe over the past four centuries. The current financial crisis has underscored once again the importance of the financial frictions for the business cycles. These facts motivate questions about financial stability. How resilient is the financial system to various shocks? At what point does the system enter a crisis regime, in the sense that market volatility, credit spreads and financing activity change drastically? To what extent is risk exogenous, and to what extent is it generated by the interactions within the system? How does one quantify systemic risk? Does financial innovation really destabilize the financial system? How does the system respond to various policies, and how do policies affect spillovers and welfare?

The seminal contributions of Bernanke and Gertler (1989), Kiyotaki and Moore (1997) (hereafter KM) and Bernanke, Gertler, and Gilchrist (1999) (hereafter BGG) uncover several important channels how financial frictions affect the macroeconomy. First, temporary shocks can have *persistent* effects on economic activity as they affect the net worth of levered agents, and financial constraints. Net worth takes time to rebuild. Second, financial frictions lead to the *amplification* of shocks, directly through leverage and indirectly through prices. Thus, small shocks can have large effects on the economy. The amplification through prices works through adverse feedback loops, as declining net worth of levered agents leads to drop in prices of assets concentrated in their hands, further lowering these agents' net worth.

Both BGG and KM consider on the amplification and propagation of small shocks that hit the system at its deterministic steady state, and focus on linear approximations of system dynamics. In both papers the location of the steady state is pinned down by an exogenous parameter. BGG exogenously fix the rate at which accumulated net worth flows out of the constrained sector. In KM, the wealth distribution at the steady state is determined by the exogenous leverage constraint. In this paper we build upon the work of BGG and KM, but we develop them in two important ways. First, instead of focusing on approximate dynamics near the steady state, we instead use continuous-time methodology to solve for the full dynamics of the system. This allows us to capture the dynamics both near and away from the steady state, and understand the size of shocks that push the system away from the steady state into the crisis regime. Second, instead of pinning down the location of the steady state by an exogenous parameter, we allow the wealth distribution to evolve endogenously through risk-taking and payout decisions. The endogeneity

of the steady state leads an important relationship between the risk environment and equilibrium leverage, which implies that occasional crises are an essential characteristic of system dynamics.

As BGG and KM, the core of our model has two types of agents: productive *experts* and less productive *households.* Because of financial frictions, the wealth of experts is important for their ability to buy physical capital and use it productively. The evolution of the wealth distribution depends on the agent's consumption decisions, as well as macro shocks that affect their balance sheets. Physical capital can be traded in markets, and its equilibrium price is determined endogenously by the agents' net worth and financial constraints. Unlike in BGG and KM, agents in our model anticipate shocks. In *normal* times, the system is near the stochastic steady state: a point at which agents reach their target leverage. The stochastic steady state is defined as the balance point to which the system tends to come back after it is hit by small shocks. It is a point such that, generally, after an adverse shock leads to losses, experts have sufficient time to rebuild net worth before the next shock arrives.

The most important phenomena occur when the system is knocked off balance away from the steady state. The full characterization of system dynamics allows us to derive a number of important implications.

First, the system reaction to shocks is highly nonlinear. While the system is resilient to most shocks near the steady state, unusually large shocks get much more amplified. Once in a crisis regime, even small shocks become amplified, leading to significant endogenous risk. The reason is that while at the steady state, experts can absorb moderate shocks to their net worths easily by adjusting payouts, away from the steady state payouts cannot be further reduced. Hence, near the steady state, shocks have small effect on the experts' demand for physical capital. In the crisis states away from the steady state, experts have to sell capital to cut their risk exposures. Overall, the stability of the system depends on the experts' endogenous choice of capital cushions. As it is costly to retain earnings, excess profits are paid out when experts are comfortable with their capital ratios.

Second, system reaction to shocks is asymmetric. Positive shocks at the steady state lead to larger payouts and little amplification, while large negative shocks are amplified into crises episodes resulting in significant inefficiencies, disinvestment, and slow recovery.

Third, increased volatility in the crisis regime affects the experts' precautionary motive. When changes in asset prices are driven endogenously by the constraints of market participants rather than changes in cash flow fundamentals, incentives to hold cash and wait to pick up assets at the bottom increase. In case prices fall further, the same amount of money can buy a larger quantity of assets, and at a lower price, increasing the expected return. In our equilibrium this phenomenon

leads to price drops in anticipation of the crisis, and higher expected return in times of increased endogenous risk. In equilibrium leverage is determined by experts' responses to everybody else's leverage: higher aggregate leverage increases endogenous risk, increases the precautionary motive and reduces individual incentives to lever up.[1]

Fourth, after moving through a high volatility region, the system can get trapped for some time in a recession with low growth and low market liquidity. The stationary distribution is $\cup$-shaped, implying that while the system spends most of the time around the steady state, it also spends some time in the depressed regime with low growth.

Fifth, the model has important asset pricing implications. In crisis regimes, credit spreads and risk premia increase and asset prices become more correlated due to endogenous risk. In general, asset returns exhibit fat tails and option prices display volatility smirks. In our model these features are implied by equilibrium behavior rather than exogenously assumed rare events.

In addition, a number of comparative statics arise because, unlike the earlier literature, we endogenize the experts' payout policy. Economically, a phenomenon we refer to as the *volatility paradox* arises. Paradoxically, lower exogenous risk can lead to more extreme volatility spikes in the crisis regime. This happens because low fundamental risk leads to higher equilibrium leverage. In sum, whatever the exogenous risk, it is normal for the system to sporadically enter volatile regimes away from the steady state. In fact, our results suggest that low risk environments are conducive to greater buildup of systemic risk.

*Financial innovation* that allows experts to hedge their idiosyncratic risk can be self-defeating as it leads to higher systemic risk. For example, securitization of home loans into mortgage-backed securities allows institutions that originate loans to unload some of the risks to other institutions. Institutions can also share risks through contracts like credit-default swaps, through integration of commercial banks and investment banks, and through more complex intermediation chains (e.g. see Shin (2010)). We find that, when experts can hedge idiosyncratic risks better among each other in our model, they take on more leverage. This makes the system less stable. Thus, while securitization is in principle a good thing - it reduces the costs of idiosyncratic shocks and thus interest rate spreads - it ends up amplifying systemic risks in equilibrium.

All our results extend to a setting, in which intermediaries facilitate lending from households to experts. In this case, the net worth of both intermediaries and end borrowers matter for system dynamics. As in the models of Diamond (1984) and Holmström and Tirole (1997) the role of

[1]The fact that in reality risk taking by leveraged market participants is not observable to others can lead to risk management strategies that are in aggregate mutually inconsistent. Too many of them might be planning to sell their capital in case of an adverse shock, leading to larger than expected price drops. Brunnermeier, Gorton, and Krishnamurthy (2010) argue that this is one contributing factor to systemic risk.

the intermediaries is to monitor end borrowers. In this process intermediaries become exposed to macro risks.

Our model implies important lessons for *financial regulation* when financial crises lead to spillovers into the real economy. Obviously, regulation is subject to time inconsistency. For example, policies intended to ex-post recapitalize the financial sector in crisis times can lead to moral hazard in normal times. In addition, even prophylactic well-intentioned policies can have unintended consequences. For example, capital requirements, if set improperly, can easily harm welfare, as they have little effect on behavior in good times, but bind in downturns. That is, in good times the fear of hitting a capital constraint in the future may be too weak to induce experts to build sufficient net worth buffers to overturn the destabilizing effects in downturns. Overall, our model argues in favor of countercyclical regulation that encourages financial institutions to retain earnings and build up capital buffers in good times and relaxes constraints in downturns.

Our model makes a strong case in favor of *macro-prudential* regulation. For example, regulation that restricts payouts (such as dividends and bonus payments) should depend primarily on aggregate net worth of all intermediaries. That is, even if some of the intermediaries are well capitalized, allowing them to pay out dividends can destabilize the system if others are undercapitalized.

**Literature review.** This paper builds upon several strands of literature. At firm level, microfoundations of financial frictions lie the heart of papers that study capital structure in the presence of informational and agency frictions, as well as papers that look at financial intermediation and bank runs. In the aggregate, papers that study the effects of prices and collateral value, and more generally consider financial frictions in a general equilibrium context, are relevant.

On the firm level, papers such as Townsend (1979), Bolton and Scharfstein (1990) and DeMarzo and Sannikov (2006) explain why violations of Modigliani-Miller assumptions lead to bounds on the agents' borrowing capacity, as well as restrictions on risk sharing. Sannikov (2012) provides a survey of capital structure implications of financial frictions. It follows that in the aggregate, the wealth distribution among agents matters for the allocation of productive resources. Diamond (1984) and Holmström and Tirole (1997) emphasize the monitoring role that intermediaries perform as they channel funds from lenders to borrowers. Diamond and Dybvig (1983) and Allen and Gale (2007) intermediaries are subject to runs. He and Xiong (2009) model runs on non-financial firms, and Shleifer and Vishny (2010) focus on bank stability and investor sentiment. These observations serve as a microfoundation to the balance sheet assumptions made by the literature that analyze financial frictions in the macroeconomy, including our paper.

In the aggregate, a number of papers also build on the idea that adverse price movements affect

the borrowers' net worth, and thus financial constraints. Shleifer and Vishny (1992) emphasize the importance of the *liquidating* price of capital, determined at the time when natural buyers, who are typically in the same industry, are also constrained. Shleifer and Vishny (1997) emphasize the role of the solvency constraint in the fund managers' ability to trade against mispricing. In Geanakoplos (1997, 2003), the identity of the marginal buyer affects prices. Brunnermeier and Pedersen (2009) focus on margin constraints that depend on volatility, and Rampini and Viswanathan (2011) stress that highly productive firms go closer to their debt capacity and hence are harder hit in a downturns.

Important papers that analyze financial frictions in infinite-horizon macro settings include KM, Carlstrom and Fuerst (1997) and BGG. These paper make use of log-linear approximations to study how financial frictions amplify shocks near the steady state of the system. Other papers, such as Christiano, Eichenbaum, and Evans (2005), Christiano, Motto, and Rostagno (2003, 2007), Curdia and Woodford (2009), Gertler and Karadi (2009) and Gertler and Kiyotaki (2011), use these techniques to study related questions, including the impact of monetary policy on financial frictions. See Brunnermeier, Eisenbach, and Sannikov (2012) for a survey of literature on economies with financial frictions.

Several recent papers avoid log-linearization, including Mendoza (2010) and He and Krishnamurthy (2010b,a). Perhaps most closely related to our model is He and Krishnamurthy (2010b). The latter studies an endowment economy to derive a two-factor asset pricing model for assets held exclusively by financial experts. Like in our paper, financial experts face equity issuance constraints. When experts are well capitalized, risk premia are determined by aggregate risk aversion since the outside equity constraint does not bind. However, after a severe adverse shock experts, who cannot sell risky assets to households, become constrained and risk premia rise sharply. He and Krishnamurthy (2010a) calibrate a variant of the model and show that in crisis equity injection is a superior policy compared to interest rate cuts or asset purchasing programs by the central bank.

Several papers identify important externalities that exist due to financial frictions. These include Bhattacharya and Gale (1987), in which externalities arise in the interbank market, Gromb and Vayanos (2002), who provide welfare analysis for a setting with credit constraints, andCaballero and Krishnamurthy (2004), who study externalities an international open economy framework. On a more abstract level these effects can be traced back to inefficiency results within an incomplete markets general equilibrium setting, see e.g. Stiglitz (1982) and Geanakoplos and Polemarchakis (1986). In Lorenzoni (2008) and Jeanne and Korinek (2010) funding constraints depend on prices that each individual investor takes as given. Adrian and Brunnermeier (2010) provide a

systemic risk measure and argue that financial regulation should focus on these externalities.

Our paper is organized as follows. We set up our baseline model in Section 2. In Section 3 we develop methodology to solve the model, and characterize the equilibrium that is Markov in the experts' aggregate net worth and present a computed example. Section 4 discusses equilibrium dynamics and properties of asset prices. Section 5 focuses on comparative statics that depend on the endogeneity of the wealth distribution. In Section 6 we discuss efficiency, spillovers and regulation. Section 7 concludes.

# 2 The Baseline Model

In an economy without financial frictions and complete markets, the flow of funds to the most productive agents is unconstrained, and hence the distribution of wealth is irrelevant. With frictions, the wealth distribution may change with macro shocks and affect aggregate productivity. When the net worth of productive agents become depressed, the allocation of resources (such as capital) in the economy becomes less efficient and asset prices may decline.

In this section we develop a simple baseline model with two types of agents, in which productive agents, experts, can finance their projects only by issuing risk-free debt. The simple capital structure of experts simplifies exposition, but it is not crucial for our results. As long as frictions restrict risk-sharing among the agents, aggregate shocks affect the wealth distribution and thus the agents' risk bearing capacity. In Appendix A, we examine how an agency problem leads to various capital structures, in which experts may issue some issue equity and other payoff-sensitive securities, such as risky debt. We also generalize the model to include intermediaries that facilitate the flow of funds to productive projects, so that the net worth of intermediaries also becomes important.

**Technology.** We consider an economy populated by experts and households. Both types of agents can own capital, but experts are able to manage it more productively.

We denote the aggregate account of efficiency units of capital in the economy by $K_t$ and capital held by an individual agent by $k_t$, where $t \in [0, \infty)$ is time. Physical capital $k_t$ held by an expert produces output at rate

$$y_t = ak_t,$$

per unit of time, where $a$ is a parameter. Output serves as numeraire and its price is normalized to one. New capital can be built through internal investment. When held by an expert, capital

evolves according to

$$dk_t = (\Phi(\iota_t) - \delta)k_t\, dt + \sigma k_t\, dZ_t \tag{1}$$

where $\iota_t$ is the investment rate per unit of capital (i.e. $\iota_t k_t$ is the total investment rate) and $dZ_t$ are exogenous aggregate Brownian shocks. Function $\Phi$, which satisfies $\Phi(0) = 0$, $\Phi'(0) = 1$, $\Phi'(\cdot) > 0$ and $\Phi''(\cdot) < 0$, represents a standard investment technology with adjustment costs. In the absence of investment, capital managed by experts depreciates at rate $\delta$. The concavity of $\Phi(\iota)$ when $\iota$ is negative represents *technological illiquidity*, i.e. large-scale disinvestments are less effective when $\Phi''$ is more negative.

Households are less productive. Capital managed by households produces output of only

$$\underline{y}_t = \underline{a}\, \underline{k}_t$$

with $\underline{a} \leq a$, and evolves according to

$$d\underline{k}_t = (\Phi(\underline{\iota}_t) - \underline{\delta})\, \underline{k}_t\, dt + \sigma \underline{k}_t\, dZ_t,$$

with $\underline{\delta} > \delta$, where $\underline{\iota}_t$ is the household investment rate per unit of capital.

The Brownian shocks $dZ_t$ reflect the fact that one learns over time how "effective" the capital stock is.[2] That is, the shocks $dZ_t$ capture changes in expectations about the future productivity of capital, and $k_t$ reflects the "efficiency units" of capital, measured in expected future output rather than in simple units of physical capital (number of machines). For example, when a company reports current earnings, it not only reveals information about current but also future expected cash flow. In this sense our model is also linked to the literature on connects news to business cycles, see e.g. Jaimovich and Rebelo (2009).

**Preferences.** Experts and less productive households are risk neutral. Households have the discount rate $r$ and they may consume both positive and negative amounts. This assumption ensures that households provide fully elastic lending at the risk-free rate of $r$. Denote by $\underline{c}_t$ the cumulative consumption of an individual household until time $t$, so that $d\underline{c}_t$ is consumption at

[2] Alternatively, one can also assume that the economy experiences aggregate TFP shocks $a_t$ with $da_t = a_t \sigma dZ_t$. Output would be $y_t = a_t \kappa_t$, where capital $\kappa$ is now measured in physical (instead of efficiency) units and evolves according to $d\kappa_t = (\Phi(\iota_t/a_t) - \delta)\kappa_t dt$. To preserve the tractable scale invariance property one has to modify the adjustment cost function to $\Phi(\iota_t/a_t)$. The fact that adjustment costs are higher for high $a_t$ can be justified by the fact that high TFP economies are more specialized.

time $t$. Then the utility of a household is given by[3]

$$E\left[\int_0^\infty e^{-rt}\, d\underline{c}_t\right].$$

In contrast, experts have the discount rate $\rho > r$, and they cannot have negative consumption. That is, cumulative consumption of an individual expert $c_t$ must be a *nondecreasing* process, i.e. $dc_t \geq 0$. Expert utility is

$$E\left[\int_0^\infty e^{-\rho t}\, dc_t\right].$$

**First Best, Financial Frictions and Capital Structure.** In the economy without frictions, experts would manage capital forever. Because they are less patient than households, experts would consume their entire net worths at time 0, and finance their future capital holdings by issuing *equity* to households. The Gordon growth formula implies that price of capital would be

$$\bar{q} = \max_{\iota} \frac{a - \iota}{r - \Phi(\iota) + \delta}, \tag{2}$$

so that capital earns the required return on equity, which equals to the discount rate $r$ of risk-neutral households.

If experts cannot issue 100% of equity to households, they may become constrained, leading to inefficiencies. Experts' ability to absorb risks, in order to manage capital, depends on their wealth. If experts lost their wealth, then the price of capital would permanently drop to

$$\underline{q} = \max_{\underline{\iota}} \frac{\underline{a} - \underline{\iota}}{r - \Phi(\underline{\iota}) + \underline{\delta}},$$

the price that the households would be willing to pay if they had to hold capital forever.

The constraint on expert equity issuance can be justified in many ways, e.g. through the existence of an agency problem between the experts and households. There is extensive literature in corporate finance, which argues that firm insiders must have some "skin in the game" to align their incentives with those of the outside equity holders.[4] Typically, agency models imply that the expert's incentives and effort increase in his equity stake. The productivity is the greatest when the expert owns the entire equity stake and borrows from outside investors exclusively through risk-free debt.

While agency models place a restriction on the risk that expert net worth must absorb, they

[3]Note that we do not denote by $c(t)$ the flow of consumption and write $E\left[\int_0^\infty e^{-\rho t} c(t)\, dt\right]$, because consumption can be lumpy and singular and hence $c(t)$ may be not well defined.

[4]See Jensen and Meckling (1976), Bolton and Scharfstein (1990) and DeMarzo and Sannikov (2006).

imply nothing about how the remaining cash flows are divided among outside investors. That is, the Modigliani-Miller theorem holds with respect to those cash flows. They can be divided among various securities, including risk-free debt, risky debt, equity and hybrid securities. The choice of the securities has no effect on firm value or the equilibrium. Moreover, because the assumptions of Harrison and Kreps (1979) hold in our setting, there exists an analytically convenient capital structure that in which outsiders hold only equity and risk-free debt. Indeed, any other security can be perfectly replicated by continuous trading of equity and risk-free debt. More generally, an equivalent capital structure involving risky long-term debt provides an important framework for studying default in our setting. We propose an agency model and analyze its capital structure implications in Appendix A.

For now, to simplify exposition, we focus on the simplest assumption that delivers the main results of this paper: that experts must retain 100% of their equity and can issue only risk-free debt.

**Market for Capital.** There is a fully liquid market for physical capital, in which experts can trade capital among each other or with households. Denote the equilibrium market price of capital in terms of output by $q_t$ assume that its law of motion is of the form

$$dq_t = \mu_t^q q_t \ dt + \sigma_t^q q_t \, dZ_t. \tag{3}$$

That is, capital $k_t$ is worth $q_t k_t$. In equilibrium $q_t$ is determined *endogenously*, and it is bounded between $\underline{q}$ and $\bar{q}$.

**Return from Holding Capital.** When an expert buys and holds $k_t$ units of capital at price $q_t$, by Ito's lemma the value of this capital evolves according to[5]

$$\frac{d(k_t q_t)}{k_t q_t} = (\Phi(\iota_t) - \delta + \mu_t^q + \sigma\sigma_t^q) \, dt + (\sigma + \sigma_t^q) \, dZ_t. \tag{4}$$

This is the *capital gains rate* that the expert is earning. The total risk of this position consists of *fundamental* risk due to news about the future productivity of capital $\sigma \, dZ_t$, and *endogenous* risk due to financial frictions in the economy, $\sigma_t^q \, dZ_t$. Capital also generates a *dividend yield* of $(a - \iota_t)/q_t$ from output remaining after internal investment. Thus, the total return that experts

[5]We use Ito's product rule. If $dX_t/X_t = \mu_t^X \, dt + \sigma_t^X \, dZ_t$ and $dY_t/Y_t = \mu_t^Y \, dt + \sigma_t^Y \, dZ_t$, then

$$d(X_t Y_t) = Y_t \, dX_t + X_t \, dY_t + (\sigma_t^X \sigma_t^Y)(X_t Y_t) \, dt.$$

earn from capital (per unit of wealth invested) is

$$dr_t^k = \underbrace{\frac{a-\iota_t}{q_t}\,dt}_{\text{dividend yield}} + \underbrace{(\Phi(\iota_t) - \delta + \mu_t^q + \sigma\sigma_t^q)\,dt + (\sigma + \sigma_t^q)\,dZ_t}_{\text{capital gains rate}}. \tag{5}$$

Similarly, less productive households earn the return of

$$d\underline{r}_t^k = \underbrace{\frac{\underline{a}-\underline{\iota}_t}{q_t}\,dt}_{\text{dividend yield}} + \underbrace{(\Phi(\underline{\iota}_t) - \underline{\delta} + \mu_t^q + \sigma\sigma_t^q)\,dt + (\sigma + \sigma_t^q)\,dZ_t}_{\text{capital gains rate}}. \tag{6}$$

**Dynamic Trading and Experts' Problem.** The net worth $n_t$ of an expert who invests fraction $x_t$ of his wealth in capital, $1-x_t$ in the risk-free asset, and consumes at rate $dc_t$, evolves according to[6]

$$\frac{dn_t}{n_t} = x_t\,dr_t^k + (1-x_t)\,r\,dt - \frac{dc_t}{n_t}. \tag{7}$$

We expect $x_t$ to be greater than 1, i.e. experts use leverage. Less productive households provide fully elastic debt funding for the interest rate $r < \rho$ to any expert with positive net worth.[7] Any expert with positive net worth can guarantee to repay any the loan with probability one, because prices change continuously, and individual experts are small and have no price impact.

Formally, each expert solves

$$\max_{x_t \geq 0,\, dc_t \geq 0,\, \iota_t} E\left[\int_0^\infty e^{-\rho t} dc_t\right],$$

subject to the solvency constraint $n_t \geq 0, \forall t$ and the dynamic budget constraint (7).

We refer to $dc_t/n_t$ as the consumption *rate* of an expert. Note that whenever two experts choose the same portfolio weights and consume wealth at the same rate, their expected discounted payoffs will be proportional to their net worth.

**Households' problem.** Similarly, the net worth $\underline{n}_t$ of any household that invests fraction $\underline{x}_t$ of wealth in capital, $1-\underline{x}_t$ in the risk-free asset, and consumes $d\underline{c}_t$, evolves according to

$$\frac{d\underline{n}_t}{\underline{n}_t} = \underline{x}_t\,d\underline{r}_t^k + (1-\underline{x}_t)\,r\,dt - \frac{d\underline{c}_t}{\underline{n}_t}. \tag{8}$$

[6] Chapter 5 of Duffie (2010) is an excellent overview of the mathematics of portfolio returns in continuous time.

[7] In the short run, an individual expert can hold an arbitrarily large amount of capital by borrowing through risk-free debt because prices change continuously in our model, and individual experts are small and have no price impact.

Each household solves

$$\max_{\underline{x}_t \geq 0,\, d\underline{c}_t,\, \underline{\iota}_t} E\left[\int_0^\infty e^{-rt} d\underline{c}_t\right],$$

subject to $\underline{n}_t \geq 0$ and the dynamic budget constraint (8). Note that household consumption $d\underline{c}_t$ can be both positive and negative, unlike that of experts.

In sum, experts and households differ in three ways: First, experts are more productive since $a \geq \underline{a}$ and/or $\delta < \underline{\delta}$. Second, experts are less patient than households, i.e. $\rho > r$. Third, experts' consumption has to be positive while household consumption is allowed to be negative, to ensure that the risk free rate is always $r$.[8]

**Equilibrium.** Informally, an equilibrium is characterized by the processes of the market price of capital $\{q_t\}$, as well as the investment and consumption choices of agents, such that, given prices, agents maximize their expected utilities and markets clear. To define an equilibrium formally, we denote the set of experts to be the interval $\mathbb{I} = [0, 1]$, and index individual experts by $i \in \mathbb{I}$, and similarly denote the set of less productive households by $\mathbb{J} = (1, 2]$ with index $j$.

**Definition 1** *For any initial endowments of capital $\{k_0^i, \underline{k}_0^j; i \in \mathbb{I},\, j \in \mathbb{J}\}$ such that*

$$\int_{\mathbb{I}} k_0^i \, di + \int_{\mathbb{J}} \underline{k}_0^j \, dj = K_0,$$

*an equilibrium is described by a group of stochastic processes on the filtered probability space defined by the Brownian motion $\{Z_t, t \geq 0\}$: the price process of capital $\{q_t\}$, net worths $\{n_t^i, \underline{n}_t^j \geq 0\}$, capital holdings $\{k_t^i, \underline{k}_t^j \geq 0\}$, investment decisions $\{\iota_t^i, \underline{\iota}_t^j \in \mathbb{R}\}$, and consumption choices $\{dc_t^i \geq 0, d\underline{c}_t^j\}$ of individual agents $i \in \mathbb{I}$, $j \in \mathbb{J}$; such that*

*(i) initial net worths satisfy $n_0^i = k_0^i q_0$ and $\underline{n}_0^j = \underline{k}_0^j q_0$, for $i \in \mathbb{I}$ and $j \in \mathbb{J}$,*

*(ii) each expert $i \in \mathbb{I}$ and each household $j \in \mathbb{J}$ solve his problem given prices*

*(iii) markets for consumption goods,[9] and capital clear*

$$\int_{\mathbb{I}} (dc_t^i)\, di + \int_{\mathbb{J}} (d\underline{c}_t^j)\, dj = \left(\int_{\mathbb{I}} (a - \iota_t^i) k_t^i \, di + \int_{\mathbb{J}} (\underline{a} - \underline{\iota}_t^j)\, \underline{k}_t^j \, dj\right) dt, \quad \text{and} \quad \int_{\mathbb{I}} k_t^i \, di + \int_{\mathbb{J}} \underline{k}_t^j \, dj = K_t,$$

$$\text{where} \quad dK_t = \left(\int_{\mathbb{I}} (\Phi(\iota_t^i) - \delta) k_t^i \, di + \int_{\mathbb{J}} (\Phi(\underline{\iota}_t^j) - \underline{\delta})\, \underline{k}_t^j \, dj\right) dt + \sigma K_t \, dZ_t. \tag{9}$$

[8] Negative consumption could be interpreted as the disutility from an additional labor input to produce extra output.

[9] In equilibrium while aggregate consumption is continuous with respect to time, the experts' and households' consumption is not. However, their singular parts cancel out in the aggregate.

Note that if three of the markets clear, then the remaining market for risk-free lending and borrowing at rate $r$ automatically clears by Walras' Law.

Since agents are atomistic perfectly competitive price-takers, the distribution of wealth among experts and among households in this economy does not matter. However, the wealth of experts *relative* to that of households plays a crucial role in our model, as we discuss in the next section.

# 3 Solving for the Equilibrium

We have to determine how the equilibrium allocation of capital and price $q_t$, as well as the agents' consumption decisions, depend on the history of macro shocks $\{Z_s;\ 0 \leq s \leq t\}$. Our procedure to solve for the equilibrium has two major steps. First, we use the equilibrium conditions, agent utility maximization and market clearing, to derive certain properties of the price $q_t$, the expert's value functions and other processes. Second, we show that the equilibrium dynamics can be described by a single state variable and derive a system of equations that determine the price of capital $q_t$ and other variables as functions of this state variable.

Intuitively, we expect the equilibrium prices to fall after negative macro shocks, because those shocks lead to expert losses and make them more constrained. At some point, prices may drop so far that less productive households may find it profitable to buy capital from experts. Less productive households are speculative as they hope to sell capital back to experts at a higher price in the future. In this sense households are liquidity providers as they pick up some of the functions of the traditional financial sector in times of crises.

**Internal Investment.** The returns (5) and (6) that experts and households receive from capital are maximized by choosing the investment rate $\iota$ that solves

$$\max_{\iota} \Phi(\iota) - \iota/q_t.$$

The first-order condition $\Phi'(\iota) = 1/q_t$ (marginal Tobin's q) implies that the optimal investment rate is a functions of the price $q_t$, i.e.

$$\iota_t = \underline{\iota}_t = \iota(q_t),$$

The determination of the optimal investment rate is a completely static problem: it depends only on the current price of capital $q_t$. From now on, we *incorporate the optimal investment rate* in the expressions for the returns $dr_t^k$ and $d\underline{r}_t^k$ that experts and households earn.

**Households' optimal portfolio choice.** Denote the fraction of physical capital held by households by

$$1 - \psi_t = \frac{1}{K_t} \int_{\mathbb{J}} \underline{k}_t^j \, dj.$$

The problem of households is straightforward as they are not financially constrained. In equilibrium they must earn a return of $r$, their discount rate, from risk-free lending to experts and, if $1 - \psi_t > 0$, from holding capital. If households do not hold any physical capital, i.e. $\psi_t = 1$, their expected return on capital must be *less than or equal* to $r$. This leads to the equilibrium condition

$$\underbrace{\frac{\underline{a} - \iota(q_t)}{q_t} + \Phi(\iota(q_t)) - \underline{\delta} + \mu_t^q + \sigma\sigma_t^q}_{E_t[d\underline{r}_t^k]/dt} \leq r, \text{ with equality if } 1 - \psi_t > 0. \tag{H}$$

**Experts' optimal portfolio and consumption choices.** The experts face a significantly more complex problem, because they are financially constrained. Their problem is *dynamic*, that is, their choice of leverage depends not only on the current price levels, but also on the entire future law of motion of prices. Even though experts are risk-neutral with respect to consumption, they exhibit risk-averse behavior in our model (in aggregate) because their marginal utility of *wealth* is stochastic - it depends on the *time-varying* investment opportunities. Greater leverage leads to higher profit and also greater risk. Experts who take on more risk suffer greater losses exactly when they value their funds the most: after negative shocks depress prices and create attractive investment opportunities.

We characterize the experts optimal dynamic strategies through the Bellman equation for their value functions. Consider a feasible strategy $\{x_t, d\zeta_t\}$, which specifies the fraction of wealth invested in capital $x_t$ and the consumption *rate* $d\zeta_t = dc_t/n_t$ of an expert, and denote by

$$\theta_t n_t = E_t \left[ \int_t^\infty e^{-\rho(s-t)} dc_s \right] \tag{10}$$

the expert's future expected payoff under this strategy. Note that the expected payoff is proportional to expert wealth $n_t$ because experts are price-takers, and their consumption under the strategy $\{x_t, d\zeta_t\}$ is proportional to their wealth because $dc_t = d\zeta_t n_t$ by definition. The following proposition provides necessary and sufficient conditions for the strategy $\{x_t, d\zeta_t\}$ to be optimal, given the price process $\{q_t, t \geq 0\}$.

**Lemma 1** *Let $\{q_t, t \geq 0\}$ be a price process for which the maximal payoff that any expert can*

*attain is finite.*[10] *Then the process* $\{\theta_t\}$ *satisfies (10) under the strategy* $\{x_t, d\zeta_t\}$ *if and only if*

$$\rho\theta_t n_t\, dt = n_t\, d\zeta_t + E[d(\theta_t n_t)] \tag{11}$$

*when* $n_t$ *follows (7), and the transversality condition* $E[e^{-\rho t}\theta_t n_t] \to 0$ *holds.*

*Moreover, this strategy is optimal if and only if*

$$\rho\theta_t n_t\, dt = \max_{\hat{x}_t \geq 0, d\hat{\zeta}_t \geq 0} n_t\, d\hat{\zeta}_t + E[d(\theta_t n_t)] \quad s.t. \quad \frac{dn_t}{n_t} = \hat{x}_t\, dr_t^k + (1-\hat{x}_t)\, r\, dt - d\hat{\zeta}_t. \tag{12}$$

Proposition 1 breaks down the Bellman equation (12) into specific conditions that the stochastic laws of motion of $q_t$ and $\theta_t$, together with the experts' optimal strategies, have to satisfy.

**Proposition 1** *Consider a finite process*

$$\frac{d\theta_t}{\theta_t} = \mu_t^\theta\, dt + \sigma_t^\theta\, dZ_t.$$

*Then* $n_t\theta_t$ *represents the maximal future expected payoff that an expert with net worth* $n_t$ *can attain and* $\{x_t, d\zeta_t\}$ *is an optimal strategy if and only if*

*(i)* $\theta_t \geq 1$ *at all times, and* $d\zeta_t > 0$ *only when* $\theta_t = 1$,

*(ii)*
$$\mu_t^\theta = \rho - r, \tag{E}$$

*(iii) either* $x_t > 0$ and
$$\underbrace{\frac{a-\iota(q_t)}{q_t} + \Phi(\iota(q_t)) - \delta + \mu_t^q + \sigma\sigma_t^q - r}_{\text{expected excess return on capital, } E_t[dr_t^k]/dt - r} = \underbrace{-\sigma_t^\theta(\sigma+\sigma_t^q)}_{\text{risk premium}}, \tag{EK}$$

*or* $x_t = 0$ *and* $E[dr_t^k]/dt - r \leq -\sigma_t^\theta(\sigma+\sigma_t^q)$,

*(iv) and the transversality condition* $E[e^{-\rho t}\theta_t n_t] \to 0$ *holds under the strategy* $\{x_t, d\zeta_t\}$.

Equation (EK) is instructive. We will see below that in equilibrium $E[dr_t^k]/dt \geq r$ and $\sigma_t^q \geq 0$, so that experts earn profit and take on more risk by levering up to buy capital. Because also $\sigma_t^\theta \leq 0$, a loss of $(\sigma+\sigma_t^q)\, dZ_t$ per dollar invested in capital happens exactly in the event that better investment opportunities arise as $\theta_t$ goes up by $\sigma_t^\theta\theta_t\, dZ_t$. Thus, while the left hand side of (EK) reflects the experts' incentives to hold more capital, the expression $\sigma_t^\theta(\sigma+\sigma_t^q)$ on the right hand side reflects the experts' *precautionary* motive. If endogenous risk ever made the right hand side

[10] In our setting, because experts are risk-neutral, their value functions under many price processes can be easily infinite (although, of course, in equilibrium they are finite).

of (EK) greater than the left hand side, experts would choose to hold no capital in volatile times, and instead lend to households at the risk-free rate, waiting to pick up assets at low prices at the bottom ("flight to quality").

As becomes clear in further analysis, the precautionary motive increases with aggregate leverage of experts, but disappears completely if experts invest in capital without using leverage. Therefore, the incentives of individual experts to take on risk are decreasing in the risks taken by other experts in the aggregate. This leads to an equilibrium choice of leverage. We conjecture, and later verify, that experts always use positive leverage in equilibrium, so that

$$\psi_t q_t K_t > N_t, \quad \text{where} \quad N_t = \int_{\mathbb{I}} n_t^i \, di.$$

While not directly relevant to our derivation of the equilibrium, it is interesting to note that because $\theta_t$ is the experts' marginal utility of wealth, at any time $t$ they use the *stochastic discount factor* (SDF)

$$e^{-\rho s} \frac{\theta_{t+s}}{\theta_t} \tag{13}$$

to price cash flows at a future time $t+s$. That is, the price of any asset that pays a random cash flow of $\hat{n}_{t+s}$ at time $t+s$ is

$$\hat{n}_t = E_t \left[ \frac{e^{-\rho s} \theta_{t+s}}{\theta_t} \hat{n}_{t+s} \right].$$

**Market Clearing.** Besides household and expert optimization, the equilibrium has to satisfy market-clearing conditions. If the experts always hold a positive amount of capital, i.e. $\psi_t > 0$, then, interestingly, the market-clearing conditions can be automatically satisfied as long as the condition (H) for household optimization as well as conditions (E) and (EK) for experts hold. Indeed, according to Proposition 1, as long as (EK) holds, any nonnegative amount of capital in the experts' portfolios is consistent with the experts' utility maximization, so markets for capital can clear. Markets for consumption goods and risk-free assets clear because the households, whose consumption may be positive or negative, are willing to borrow and lend arbitrary amounts at the risk-free rate $r$.

**Wealth distribution.** Due to financial frictions, the wealth distributions across agents matters in our economy. In aggregate, experts and households have wealth

$$N_t = \int_{\mathbb{I}} n_t^i \, di \qquad \text{and} \qquad q_t K_t - N_t = \int_{\mathbb{J}} \underline{n}_t^j \, dj$$

respectively. The experts' wealth share is

$$\eta_t \equiv \frac{N_t}{q_t K_t} \in [0,1].$$

Experts become constrained when $\eta_t$ falls, leading a low price of capital $q_t$, low investment rate $\iota(q_t)$, and a larger fraction of capital $1-\psi_t$ allocated to households.

Our model has convenient scale-invariance properties, which imply that the price level, as well as inefficiencies with respect to investment and capital allocation, depend on $\eta_t$. That is, for any equilibrium in one economy, there is an equivalent equilibrium with the same laws of motion of $\eta_t$, $q_t$ $\theta_t$ and $\psi_t$ in any scaled version of the economy by any factor $\varsigma \in (0,\infty)$.

We will characterize an equilibrium that is Markov in the state variable $\eta_t$. Before we proceed, Lemma 2 derives the equilibrium law of motion of $\eta_t = N_t/(q_t K_t)$ from the laws of motion of $N_t$, $q_t$ and $K_t$. In Lemma 2, we do not assume that the equilibrium is Markov.[11]

**Lemma 2** *The equilibrium law of motion of $\eta_t$ is*

$$\frac{d\eta_t}{\eta_t} = \frac{\psi_t - \eta_t}{\eta_t}(dr_t^k - r\,dt - (\sigma+\sigma_t^q)^2\,dt) + \frac{a-\iota(q_t)}{q_t}\,dt + (1-\psi_t)(\underline{\delta}-\delta)\,dt - d\zeta_t, \tag{14}$$

*where $d\zeta_t = dC_t/N_t$, with $dC_t = \int_{\mathbb{I}}(dc_t^i)\,di$, is the aggregate expert consumption rate. Moreover, if $\psi_t > 0$, then (EK) implies that we can write*

$$\frac{d\eta_t}{\eta_t} = \mu_t^\eta\,dt + \sigma_t^\eta\,dZ_t - d\zeta_t, \tag{15}$$

$$\textit{where} \quad \sigma_t^\eta = \frac{\psi_t-\eta_t}{\eta_t}(\sigma+\sigma_t^q) \quad \textit{and} \quad \mu_t^\eta = -\sigma_t^\eta(\sigma+\sigma_t^q+\sigma_t^\theta) + \frac{a-\iota(q_t)}{q_t} + (1-\psi_t)(\underline{\delta}-\delta).$$

**Markov Equilibrium.** Because of scale invariance, it is natural to look for an equilibrium that is Markov in the state variable $\eta_t$. In a Markov equilibrium, all processes are functions of $\eta_t$, i.e.

$$q_t = q(\eta_t), \quad \theta_t = \theta(\eta_t) \text{ and } \psi_t = \psi(\eta_t). \tag{16}$$

If these functions are known, then we can use equation (15) to map any path of aggregate shocks $\{Z_s, s \le t\}$ into the current values of $\eta_t$, and subsequently $q_t$, $\theta_t$ and $\psi_t$.

[11] We conjecture that the Markov equilibrium we derive in this paper is unique, i.e. there are no other equilibria in the model (Markov or non-Markov). While the proof of uniqueness is beyond the scope of the paper, a result like Lemma 2 should be helpful for the proof of uniqueness.

Proposition 2 reduces the problem of finding functions (16) to a system differential equations. For a given value of $\eta$, to get the second derivatives $q''(\eta)$ and $\theta''(\eta)$ from $q(\eta)$, $\theta(\eta)$ and their first derivatives, we use equations (E), (EK), (H) and (15) to

1. using Ito's lemma, express the volatilities $\sigma_t^\eta$, $\sigma_t^q$, and $\sigma_t^\theta$ and
2. after calculating the drifts $\mu_t^\eta$, $\mu_t^q$ and $\mu_t^\theta$, use Ito's lemma again to get the second derivatives of $q$ and $\theta$.

In these calculations, we assume that $\psi_t > \eta_t$, i.e. expert leverage is positive.

**Proposition 2** *The equilibrium domain of functions $q(\eta)$, $\theta(\eta)$ and $\psi(\eta)$ is an interval $[0, \eta^*]$. The following formulas can be used to compute $q''(\eta)$, $\theta''(\eta)$ and $\psi(\eta)$ from $(\eta, q(\eta), q'(\eta), \theta(\eta), \theta'(\eta))$*

1. *Find $\psi \in (\eta,\ \eta + q(\eta)/q'(\eta))$ such that*[12]

$$\frac{a-\underline{a}}{q(\eta)} + \underline{\delta} - \delta + (\sigma + \sigma_t^q)\sigma_t^\theta = 0, \tag{17}$$

$$\textit{where} \qquad \sigma_t^\eta \eta = \frac{(\psi-\eta)\sigma}{1-(\psi-\eta)q'(\eta)/q(\eta)}, \quad \sigma_t^q = \frac{q'(\eta)}{q(\eta)}\sigma_t^\eta \eta \quad \text{and} \quad \sigma_t^\theta = \frac{\theta'(\eta)}{\theta(\eta)}\sigma_t^\eta \eta. \tag{18}$$

2. *If $\psi > 1$, set $\psi = 1$ and recalculate (18).*
3. *Compute*

$$\mu_t^\eta = -\sigma_t^\eta(\sigma + \sigma_t^q + \sigma_t^\theta) + \frac{a - \iota(q(\eta))}{q(\eta)} + (1-\psi)(\underline{\delta} - \delta),$$

$$\mu_t^q = r - \frac{a - \iota(q(\eta))}{q(\eta)} - \Phi(q(\eta)) + \delta - \sigma\sigma_t^q - \sigma_t^\theta(\sigma + \sigma_t^q), \quad \mu_t^\theta = \rho - r,$$

$$q''(\eta) = \frac{2(\mu_t^q q(\eta) - q'(\eta)\mu_t^\eta \eta)}{(\sigma_t^\eta)^2 \eta^2} \qquad \text{and} \qquad \theta''(\eta) = \frac{2\left[\mu_t^\theta \theta(\eta) - \theta'(\eta)\mu_t^\eta \eta\right]}{(\sigma_t^\eta)^2 \eta^2}. \tag{19}$$

*Function $q(\eta)$ is increasing, $\theta(\eta)$ is decreasing, and the boundary conditions are*

$$q(0) = \underline{q}, \quad \theta(\eta^*) = 1, \quad q'(\eta^*) = 0, \quad \theta'(\eta^*) = 0 \quad \textit{and} \quad \lim_{\eta \to 0} \theta(\eta) = \infty.$$

*Experts consume only when $\eta_t = \eta^*$, which is a reflecting boundary for the process $\eta_t$ due to the aggregate expert consumption rate $d\zeta_t$.*

Proposition 2 allows us to derive analytical results about equilibrium behavior and asset prices, and to compute equilibria numerically.

[12] The left hand side of (17) decreases from $(a-\underline{a})/q(\eta)+\underline{\delta}-\delta > 0$ to $-\infty$ over the interval $\psi = [\eta, \eta+q(\eta)/q'(\eta)]$.

**Algorithm to Solve the Equations.** The numerical computation of the functions $q(\eta)$, $\theta(\eta)$ and $\psi(\eta)$ poses challenges because of the singularity near $\eta = 0$. In addition, we need to determine the endogenous endpoint $\eta^*$ and match the boundary conditions both at 0 and $\eta^*$. To match the boundary conditions, it is helpful to observe that if a function $\theta(\eta)$ solves the equations of Proposition 2, then so does any function $\varsigma\theta(\eta)$, for any constant $\varsigma > 0$. Therefore, one can always adjust the level of $\theta(\eta)$ *ex post* to match the boundary condition $\theta(\eta^*) = 1$. We use the following algorithm to calculate our numerical examples.

1. Set $q(0) = \underline{q}$, $\theta(0) = 1$ and $\theta'(0) = -10^{10}$.
2. Set $q_L = 0$ and $q_H = 10^{15}$.
3. *Guess* that $q'(0) = (q_L + q_H)/2$. Use the Matlab function *ode45* to solve for $q(\eta)$ and $\theta(\eta)$ until either (a) $q(\eta)$ reaches $\bar{q}$ or (b) $\theta'(\eta)$ reaches 0 or (c) $q'(\eta)$ reaches 0, whichever happens soonest. If $q'(\eta)$ reaches 0 before any of the other events happens, then increase the *guess* by setting $q_L = q'(0)$. Otherwise, let $q_H = q'(0)$. Repeat until convergence (e.g. 50 times).
4. If $q_H$ was chosen to be large enough, then $\theta'(\eta)$ and $q'(\eta)$ will reach 0 at the same point $\eta^*$.
5. Divide the entire function $\theta(\eta)$ by $\theta(\eta^*)$ to match the boundary condition $\theta(\eta^*) = 1$.

The more negative the initial choice of $\theta'(0)$, the better we approximate the boundary condition $\theta(0) = \infty$, that is, the higher the value of $\theta(0)$ becomes after we divide the entire solution by $\theta(\eta^*)$. We provide our Matlab implementation of this algorithm in the Online Appendix.

**Numerical Example.** To compute the example in Figure 1, we took parameter values $\rho = 6\%$, $r = 5\%$, $a = 10\%$, $\underline{a} = 5\%$, $\delta = 3\%$, $\underline{\delta} = 5\%$, $\sigma = 40\%$ and $\Phi(\iota) = \frac{1}{10}(\sqrt{1 + 20\iota} - 1)$.[13] Under these assumptions,

$$\underline{q} = 0.5858 \quad \text{and} \quad \bar{q} = 1.3101.$$

As $\eta$ increases, the price of capital $q(\eta)$ increases, investment rate $\iota(q(\eta_t))$ increases and the value that experts get per unit of wealth $\theta(\eta)$ declines. The allocation of capital to experts $\psi(\eta)$ increases in $\eta$ and reaches 100% on $[\eta^\psi, \eta^*]$.

The map from the history of aggregate shocks $dZ_t$ to the state variable $\eta_t$ is captured by the *drift* $\mu_t^\eta \eta$ and the *volatility* $\sigma_t^\eta \eta$, depicted on the top panels of Figure 2. The drift of $\eta_t$ is positive on the entire interval $[0, \eta^*)$, because levered experts refrain from consumption and earn a positive

[13] The investment technology in this example has quadratic adjustment costs: an investment of $\Phi + 5\Phi^2$ generates new capital at rate $\Phi$.

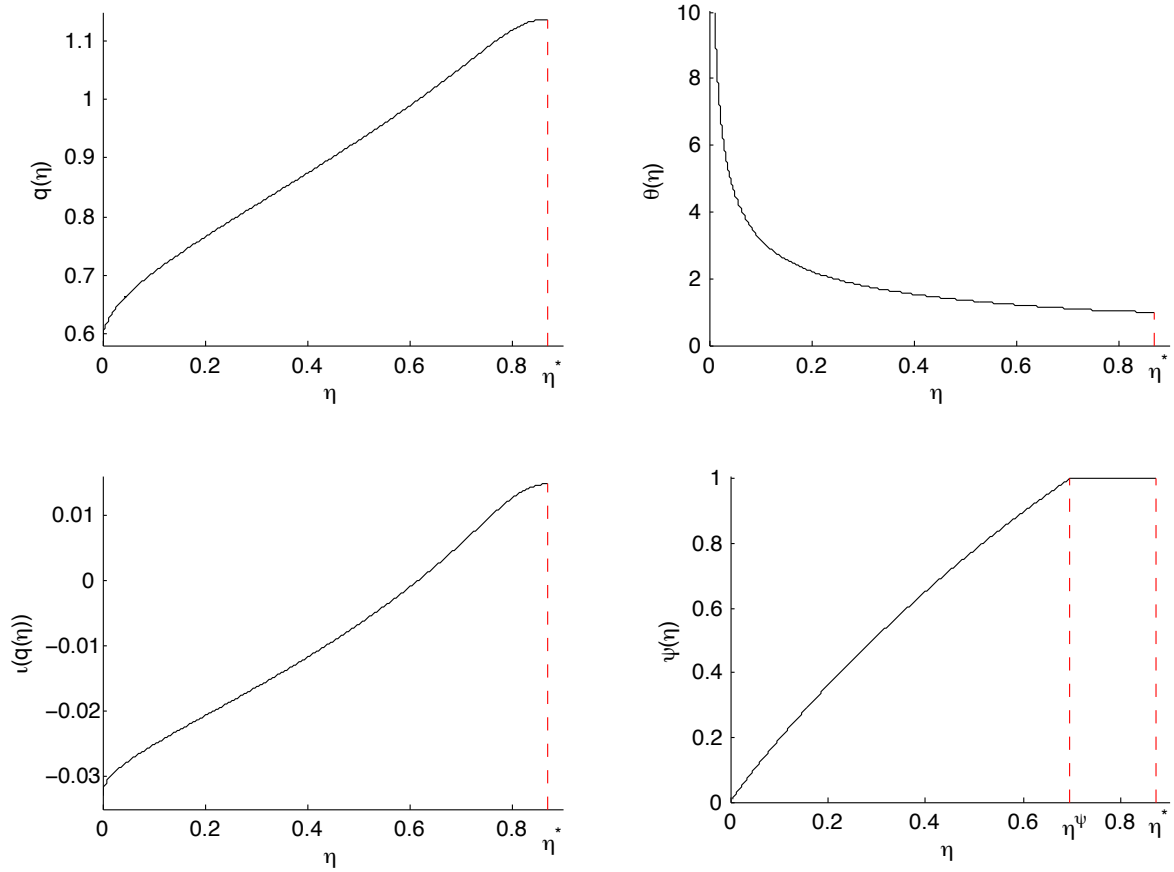


Figure 1: Equilibrium functions $q(\eta)$, $\theta(\eta)$, $\iota(q(\eta))$ and $\psi(\eta)$.

risk premium $(-\sigma_t^\theta)(\sigma+\sigma_t^q) > 0$ from capital, while less productive households are consuming and earning the risk-free return $r$ in expectation. The magnitude of the drift is increasing in expert leverage, shown on the bottom left panel of Figure 2, and the return that experts earn on capital, shown on the bottom right panel. The bottom right panel shows the lower return that households earn from capital, which equals $r$ on the interval $[0, \eta^\psi]$.

The volatility of $\eta_t$ is $\cap$-shaped: from zero, it increases towards its maximum near $\eta^\psi$ and lowers down towards $\eta^*$. We discuss the volatility dynamics in detail in Section 4. Point 0 is an absorbing boundary, but in equilibrium $\eta_t$ never reaches 0 because it evolves like a geometric Brownian motion in the neighborhood of 0, as we show in Proposition 3. Point $\eta^*$ is a reflecting boundary where experts consume excess net worth.

Because in expectation $\eta_t$ gravitates towards the reflecting boundary $\eta^*$, the point $\eta^*$ is the *stochastic steady state* of our system. Point $\eta^*$ in our model is analogous to the *deterministic steady state* in traditional macro models, such as BGG and KM. Just like the steady state in BGG and KM, $\eta^*$ is the point of global attraction of the system and, as we see from Figure 2 and as we discuss below, the volatility near $\eta^*$ is low.

However, point $\eta^*$ is also different from the steady state in BGG and KM in important ways. Unlike in traditional macro models, we do not consider the limit as noise $\sigma$ goes to 0 to identify

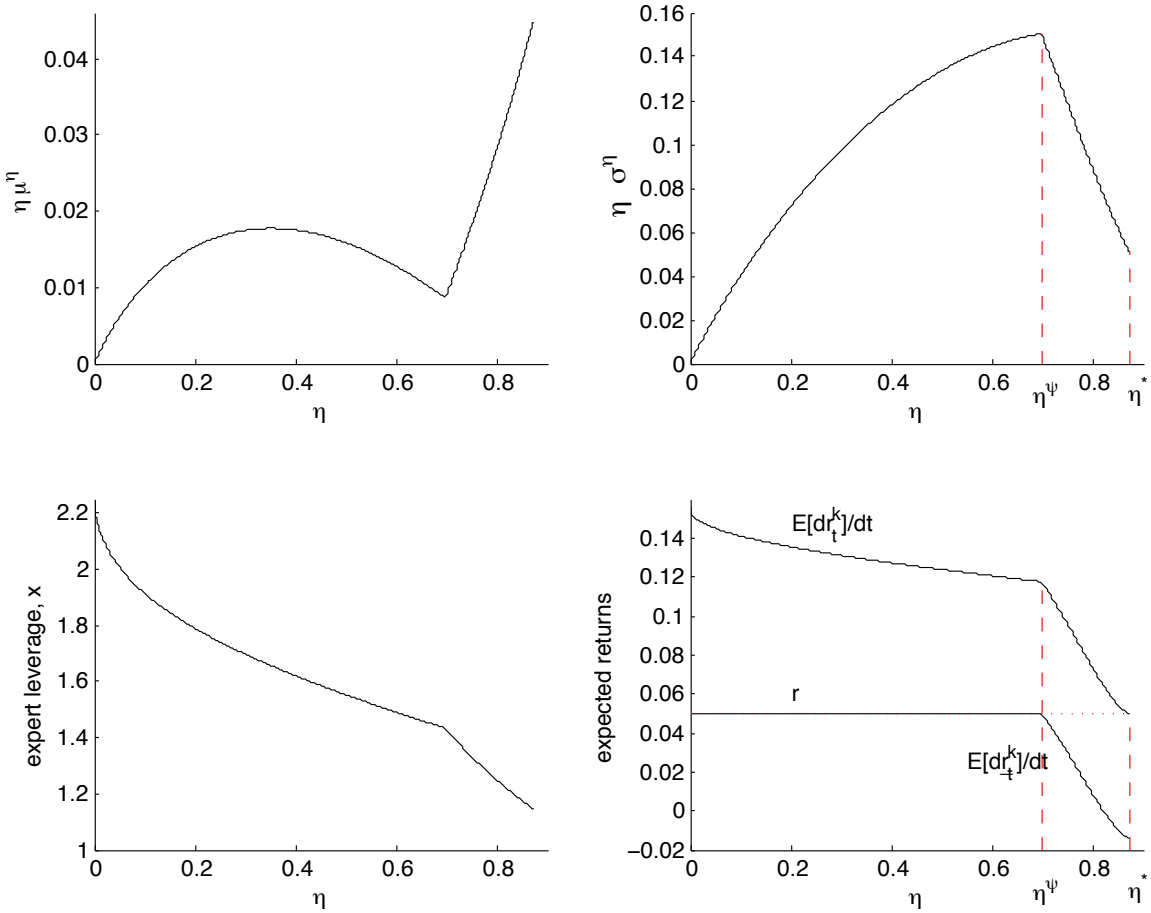


Figure 2: The drift and volatility of $\eta_t$, expert leverage, and expected asset returns.

the steady state, but rather fix the volatility of macro shocks and look for the point where the system remains still in the absence of shocks.[14] Thus, the location of $\eta^*$ endogenously depends on volatility! It is determined indirectly through the consumption and portfolio decisions, which the agents make while taking shocks into account. As we discuss in detail in Sections 4 and 5, the endogeneity of $\eta^*$ leads to a number of important phenomena, including *nonlinearity* - the system responds very differently to small and large shocks near $\eta^*$ - and the *volatility paradox* - that lower values of $\sigma$ may lead higher maximal values of endogenous risk $\sigma_t^q$.

**Inefficiencies in Equilibrium.** Without financial frictions, experts would permanently manage all capital in the economy. Capital would be priced at $\bar{q}$, leading to an investment rate of $\iota(\bar{q})$. Moreover, experts would consume their net worth in lump sum at time 0, so that the sum of utilities of all agents would be $\bar{q}K_0$. With frictions, however, there are three types of inefficiencies in our model,

[14]Strictly speaking, the *deterministic steady state* of our system is $\eta = 0$ : as $\sigma \to 0$, financial frictions go away and experts do not require any net worth to manage capital.

(i) capital misallocation, since less productive households end up managing capital for low $\eta_t$, when $\psi_t < 1$,

(ii) under-investment, since $\iota(q_t) < \iota(\bar{q})$, and

(iii) consumption distortion, since experts postpone some of their consumption into the future.

Note that these inefficiencies vary with the state of the economy: they get worse when $\eta_t$ drops.

Due to these inefficiencies, the sum of utilities of all agents is less than first best utility $\bar{q}K_0$. Even at point $\eta^*$ the sum of the agents' utilities is

$$\underbrace{E\left[\int_0^\infty e^{-\rho t} dC_t\right]}_{\text{expert payoff}} + \underbrace{E\left[\int_0^\infty e^{-rt} d\underline{C}_t\right]}_{\text{household payoff}} = \underbrace{\theta(\eta^*)N_t}_{\text{expert payoff}} + \underbrace{q(\eta^*)K_t - N_t}_{\text{household payoff/wealth}} = q(\eta^*)K_t < \bar{q}K_t, \tag{20}$$

since $\theta(\eta^*) = 1$ and $q(\eta^*) < \bar{q}$.

# 4 Instability, Endogenous Risk, and Asset Pricing

Having solved for the full dynamics, we can address various economic questions like (i) How important is fundamental cash flow risk relative to endogenous risk created by the system? (ii) Does the economy react to large exogenous shocks differently compared to small shocks? (iii) Is the dynamical system unstable and hence the economy is subject to systemic risk?

The equilibrium exhibits *instability*, which distinguishes our analysis from the log-linearized solutions of BGG and KM. Like in those papers, the price of capital in our model is subject to *endogenous risk* $\sigma_t^q$, which leads to *excess volatility*. However, different from BGG and KM, the amount of endogenous risk varies over the cycle: it goes to zero near the steady state $\eta^*$, but it is large below the *stochastic steady state* $\eta^*$. Thus, an unusually long sequence of negative shocks throws the economy into a volatile crisis regime. If more bad shocks arrive, they get amplified, pushing the system into a depressed region, from which the economy takes a long time to recover. These dynamics imply a *bimodal* stationary distribution over the state space. This is in sharp contrast to the linear approximations: they predict a normal stationary distribution around the the steady state, suggesting a much more stable system. Papers such as BGG and KM do not capture the distinction between relatively stable dynamics near the steady state, and much stronger amplification below the steady state. Our analysis highlights the sharp distinction between crisis and normal times.

The nonlinearities of system dynamics are robust to modeling assumptions. For example, a model with logarithmic utility would also generate low (but nonzero) amplification near the steady state, and high amplification below the steady state, especially at the point where experts start selling capital to households.

The differences in system dynamics near the steady state and away have to do with the forces that determine the steady state: the experts' profits and their endogenous payout/consumption decisions. The system tends to gravitate towards a point where these two forces exactly balance each other out: the steady state. Experts accumulate net worth in crisis regimes, where volatility and *risk premia* are high. They start paying out once their aggregate net worth recovers enough that the probability of the next crisis becomes tolerable.

At the end of this section, we briefly discuss the asset pricing implications of our model.

**Amplification due to Endogenous Risk** Endogenous risk refers to changes in asset prices that arise not due to changes in fundamentals, but rather due to adjustments that agents make in response to shocks, which may be driven by constraints or simply the precautionary motive. While exogenous fundamental shocks cause initial losses, endogenous risk is created through feedback loops that arise when agents react to losses. In our model, exogenous risk $\sigma$ is assumed to be constant, but endogenous risk $\sigma_t^q$ varies with the state of the system. The total volatility is the sum of exogenous and endogenous risk, $\sigma + \sigma_t^q$.

The amplification of shocks that creates endogenous risk depends on (i) expert leverage and (ii) feedback loops that arise as prices react to changes in expert net worth, and affect expert net worth further. While the experts finance themselves through fully liquid short-term debt, their assets are subject to aggregate market illiquidity.[15] Figure 3 illustrates the feedback mechanism of amplification, which has been identified by both BGG and KM near the steady state of their models.

The immediate effect of a shock $dZ_t$ that reduces $K_t$ by 1% is a drop of $N_t$ by $\frac{\psi}{\eta}$% and a drop of $\eta_t$ by $(\frac{\psi}{\eta} - 1)$%, where $\frac{\psi}{\eta}$ is the experts' leverage ratio (assets to net worth). This drop in $\eta_t$

[15] Recall that the price impact of a single expert is zero in our setting. However, the price impact due to aggregate shocks can be large. Hence, a "liquidity mismatch index" that tries captures the mismatch between market liquidity of experts' asset and funding liquidity on the liability side has to focus on price impact of assets caused by aggregate shocks rather than idiosyncratic shocks.

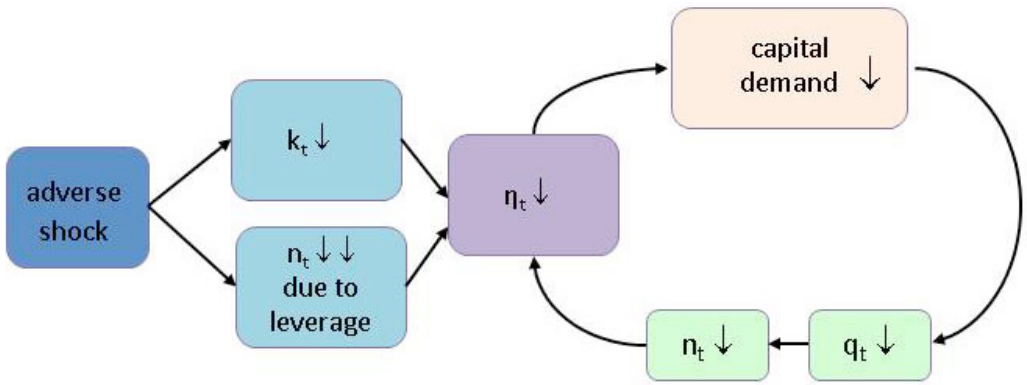


Figure 3: Adverse Feedback Loop.

causes the price $q(\eta_t)$ to drop by

$$\phi\% \equiv \frac{q'(\eta_t)\,\overbrace{(\frac{\psi_t}{\eta_t}-1)\eta_t}^{\text{total drop in } \eta}}{q(\eta)}\%,$$

where the numerator reflects the total drop in $q(\eta)$. That is, this aftershock causes $q_t K_t$ to drop further by $\phi\%$, $N_t$ further by $\frac{\psi}{\eta}\phi\%$ and a $\eta_t$ further by $(\frac{\psi}{\eta}-1)\phi\%$. We see that the initial shock gets amplified by a factor of $\phi$ each time it goes through the feedback loop. If $\phi < 1$ then this loop converges with a total amplification factor of $1/(1-\phi)$ and total impacts on $\eta_t$ and $q(\eta_t)$ of

$$\frac{d\eta_t}{\eta_t} = \frac{\frac{\psi}{\eta}-1}{1-\phi}\% = \frac{1}{\eta}\frac{\psi-\eta}{1-(\psi-\eta)q'(\eta)/q(\eta)}\% \quad \text{and} \quad \frac{dq_t}{q_t} = \frac{q'(\eta)}{q(\eta)}\frac{\psi-\eta}{1-(\psi-\eta)q'(\eta)/q(\eta)}\% \tag{21}$$

respectively. This leads us to formulas (18), provided by Proposition 2, that capture how leverage and feedback loops contribute to endogenous risk.

The amplification effect of $q'(\eta)$ on the endogenous volatility $\sigma_t^q$ is *nonlinear*, since $q'(\eta)$ enters not only the numerator, but also the denominator of (21) and (18). If $q'(\eta)$ is so large that $\phi > 1$, then the feedback effect would be completely unstable, leading to infinite volatility.

**Normal versus Crisis Times and "Ergodic Instability."** The equilibrium in our model has no endogenous risk near the stochastic steady state $\eta^*$, and significant endogenous risk below the steady state. This result strongly resonates what we observe in practice during normal times and crisis episodes.

**Theorem 1** *In equilibrium, at $\eta^*$ the system has no amplification and $\sigma_t^q = 0$, since $q'(\eta^*) = 0$.*

*For $\eta_t < \eta^*$, exogenous shocks spill over into prices, leading to the indirect dynamic amplification factor of* $1/(1-(\psi_t-\eta_t)q'(\eta_t)/q(\eta_t))$.

**Proof.** This result follows directly from Proposition 2. ■

The left panel of Figure 4 shows the volatility of the value of capital $\sigma + \sigma_t^q$, for our computed example. Because *endogenous* risk $\sigma_t^q$ rises sharply below steady state, the system exhibits nonlinearities: large shocks have a very different effect on the system than small shocks. Near the point $\eta^\psi$, increased endogenous risk and leverage lead to a high volatility of $\eta_t$, as seen in Figure 2. This leads to *systemic* risk that the economy occasionally ends up in a depressed regime far below the steady state, where most of the capital is allocated inefficiently to households.

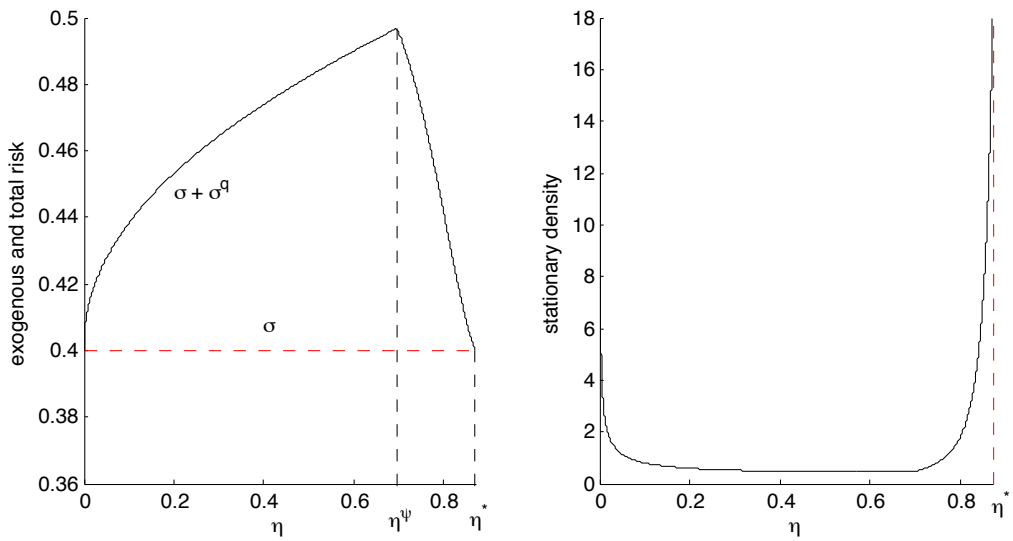


Figure 4: Systemic risk: total volatility of capital and the stationary density of $\eta_t$.

The right panel of Figure 4 shows the stationary distribution of $\eta_t$. Starting from any point $\eta_0 \in (0, \eta^*)$ in the state space, the density of the state variable $\eta_t$ converges to the stationary density in the long run as $t \to \infty$. Stationary density also measures the average amount of time that the variable $\eta_t$ spends in the long run near each point. It can be computed from the drift and volatility of $\eta_t$ using Kolmogorov forward equations (see Appendix B).

The key feature of the stationary distribution in Figure 4 is that it is bimodal with high densities at the extremes. We refer to this characteristic as "ergodic instability." The system exhibits large swings, but it is still ergodic ensuring that a stationary distribution exists.

The stationary density is high near $\eta^*$, which is the attracting point of the system, but very thin in the middle region below $\eta^*$ where the volatility is high. The system moves fast through regions of high volatility, and so the time spent there is very short. These excursions below the steady state are characterized by high uncertainty, and occasionally may take the system very far

below the steady state. In other words, the economy is subject to break-downs – i.e. systemic risk. At the extreme low end of the state space, assets are essentially valued by unproductive households, with $q_t \sim \underline{q}$, and so the volatility is low. The system spends most of the time around the extreme points: either experts are well capitalized and financial system can deal well with small adverse shocks or it drops off quite rapidly to very low $\eta$-values, where prices and experts' net worth drop dramatically. As the economy occasionally implodes, it exhibits systemic risk.

The following proposition formally demonstrates that the stationary density (if it exists) indeed has peaks at $\eta = 0$ and $\eta = \eta^*$. The proof in Appendix D shows that in a neighborhood of 0, variable $\eta_t$ evolves like a geometric Brownian motion, and uses the Kolmogorov forward equation to characterize the stationary density near 0. The stationary distribution may fail to exist if the experts' productive advantage is small relative to the volatility of capital: in that case the system gets trapped near $\eta = 0$ in the long run.

**Proposition 3** *Denote by $\kappa = (a - \underline{a})/\underline{q} + \underline{\delta} - \delta$ the risk premium that experts earn from capital at $\eta = 0$. As long as*

$$2(\rho - r)\sigma^2 < \kappa^2 \tag{22}$$

*the stationary density $d(\eta)$ exists and satisfies $d'(\eta^*) > 0$ and $d(\eta) \to \infty$ as $\eta \to 0$. If $2(\rho - r)\sigma^2 > \kappa^2$ then the stationary density does not exist and in the long run $\eta_t$ ends up in an arbitrarily small neighborhood of 0 with probability close to 1.*

In our numerical examples, $\kappa = 0.1054$.

**Robustness of the Equilibrium Features.** Our model exhibits stability in normal times, and strong amplification in crisis times, because the wealth distribution evolves *endogenously.* The steady state of the wealth distribution is determined by the relative rates of consumption of experts and households on one hand, and the difference in returns that experts and households earn on their portfolios. Variable $\eta_t$ reaches the steady state when experts accumulate enough wealth to absorb most shocks easily. At that point, competition among experts pushes up the price of capital and drives down the risk premia that experts earn. These factors encourage experts to consume their net worth instead of reinvesting it.

Near the stochastic steady state $\eta^*$, where experts become comfortable and risk premia come down, the price of capital reacts less to shocks. Thus, amplification and endogenous risk are significantly lower near the steady state. In fact, in our risk-neutral model, the risk premium $-\sigma_t^\theta(\sigma + \sigma_t^q)$ and endogenous risk $\sigma_t^q$ both drop to *zero* at $\eta^*$.

How robust are these equilibrium features? Do they depend on risk neutrality? To answer these questions, we solved a variation of our model, in which experts and households have *logarithmic* utility with the same discount rates $\rho$ and $r$, instead of being risk-neutral. All other features of the model, including production technologies, financial frictions and asset markets, are the same. Thus, equations (5) and (6) expressing the agent's return on capital are unchanged. The law of motion of $\eta_t$ takes the same form as (14), except that the risk-free return $dr_t$ is no longer constant.

Models with logarithmic utility are easy to solve, because they lead to myopic optimal consumption and portfolio choice decisions. Specifically, (1) given the net worth of $n_t$, the optimal consumption of an expert is given by $\rho n_t\, dt$ (and household, $r\underline{n}_t\, dt$), regardless of investment opportunities, and (2) the agent's optimal portfolio choice always results in the percent volatility of net worth equal to the Sharpe ratio of risky investment. The first property implies that the market-clearing condition for consumption goods is

$$r(q_tK_t - N_t) + \rho N_t = (\psi a + (1-\psi)\underline{a} - \iota(q_t))K_t \quad \Rightarrow \quad \psi_t = \frac{(r(1-\eta_t) + \rho\eta_t)q_t + \iota(q_t) - \underline{a}}{a - \underline{a}}. \tag{23}$$

Since the volatilities of expert and household net worths are $\frac{\psi_t}{\eta_t}(\sigma+\sigma_t^q)$ and $\frac{1-\psi_t}{1-\eta_t}(\sigma+\sigma_t^q)$ respectively, the second property implies that

$$E[dr_t^k - dr_t]/dt = \frac{\psi_t}{\eta_t}(\sigma+\sigma_t^q)^2 \quad \text{and} \quad E[d\underline{r}_t^k - dr_t]/dt \le \frac{1-\psi_t}{1-\eta_t}(\sigma+\sigma_t^q)^2, \tag{24}$$

with equality in the last expression if households hold a positive amount of capital $1-\psi_t > 0$. As the experts' return on capital is higher than that of households by $(a-\underline{a})/q_t + \underline{\delta} - \delta$, it follows that

$$(a-\underline{a})/q_t + \underline{\delta} - \delta = \left(\frac{\psi_t}{\eta_t} - \frac{1-\psi_t}{1-\eta_t}\right)(\sigma+\sigma_t^q)^2 \quad \Rightarrow \quad \sigma+\sigma_t^q = \sqrt{\frac{(a-\underline{a})/q_t + \underline{\delta} - \delta}{\psi_t/\eta_t - (1-\psi_t)/(1-\eta_t)}} \tag{25}$$

when $\psi_t < 1$. Also, as in the risk-neutral model,

$$\sigma+\sigma_t^q = \frac{\sigma}{1-(\psi_t-\eta_t)q'(\eta)/q(\eta)} \quad \Rightarrow \quad q'(\eta) = \frac{q(\eta)}{\psi_t-\eta_t}\left(1 - \frac{\sigma}{\sigma+\sigma_t^q}\right). \tag{26}$$

These equations are sufficient to solve for the equilibrium. Indeed, (23) leads to the boundary condition of $rq(0) + \iota(q(0)) = \underline{a}$ at $\eta = 0$. From that point, we can use equations (23), (25) and (26) to compute $q'(\eta)$ from $\eta$ and $q(\eta)$, and solve a first-order ordinary differential equation for $q(\eta)$ until the point $\eta^\psi$ where $\psi_t$ reaches 1. On $[\eta^\psi, 1]$, we can solve for $q(\eta)$ from (23), using

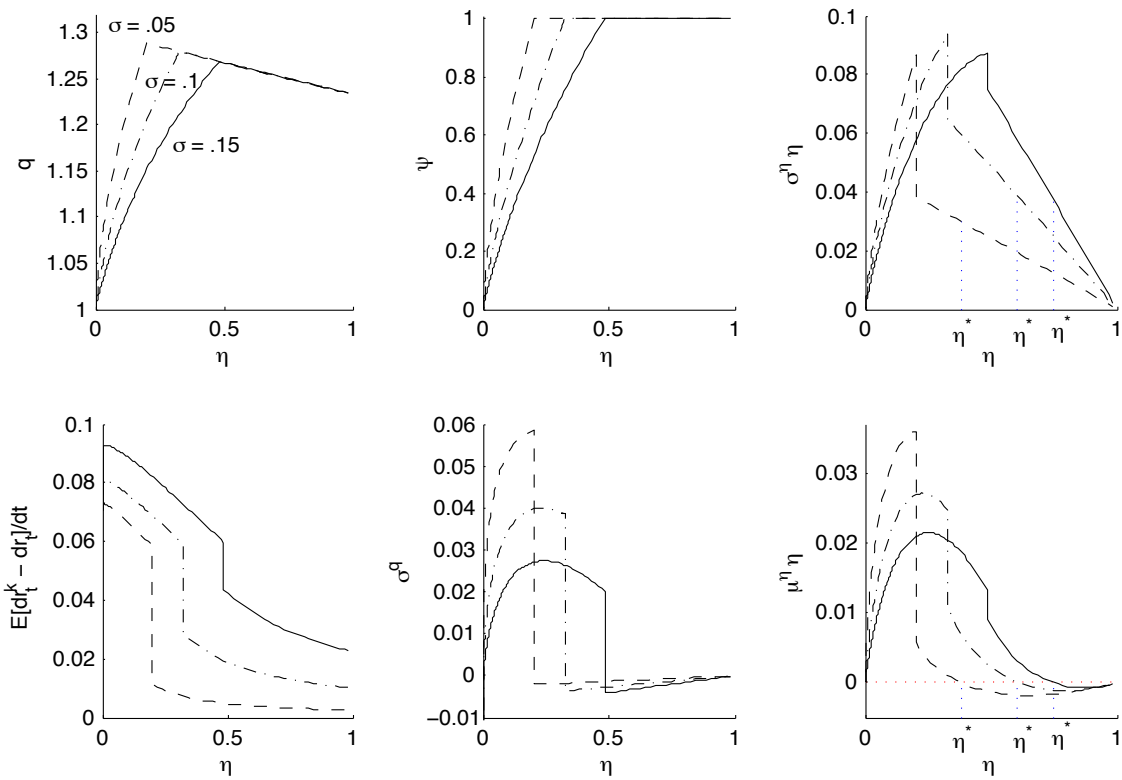


Figure 5: Equilibrium with logarithmic utility.

$\psi_t = 1$. Equations (14) and (24) imply that the volatility and drift of $\eta_t$ are given by

$$\sigma_t^\eta = \frac{\psi_t - \eta_t}{\eta_t}(\sigma + \sigma_t^q) \quad \text{and} \quad \mu_t^\eta = (\sigma_t^\eta)^2 + \frac{a - \iota(q_t)}{q_t} + (1 - \psi_t)(\underline{\delta} - \delta) - \rho, \tag{27}$$

where we used $d\zeta_t = \rho\eta_t\, dt$ for the experts' consumption rate. The stochastic steady state is defined as the point $\eta^*$ where $\mu_t^\eta = 0$.

Figure 5 shows a computed example for the same parameter values of $\rho$, $r$, $a$, $\underline{a}$, $\delta$, $\underline{\delta}$ and $\theta$ as we used in Section 3, and $\sigma = 5\%$, $10\%$ and $15\%$. Certainly, the sharp result of the risk-neutral model that $\sigma_t^q = 0$ at the steady state $\eta^*$ no longer holds exactly. However, it is still true that the volatility of $\eta_t$ is low near $\eta^*$ and it rises below $\eta^*$ as the experts' leverage increases.[16] If $\eta_t$ falls below the point $\eta^\psi$ where households start investing, the volatility of $\eta_t$ jumps sharply.[17]

These features arise because the wealth distribution is endogenous. For example, as $\sigma$ increases, risk premia rise, experts make more profit and the steady state $\eta^*$ shifts to the right into the region where experts are *less levered*. This endogenous force stabilizes the steady state as exogenous risk

[16]The location of the steady state $\eta^*$ above $\eta^\psi$ depends on the assumption that $\rho$ is not much larger than $r$. If so, then the steady state falls in the region where only experts invest.

[17]The jump in volatility at point $\eta^\psi$ occurs because the price $q(\eta)$ has a kink at $\eta^\psi$, which occurs due to the mechanical relationship (23) between $\psi_t$ and the market price in the log utility model. Technically, because of this feature, we have to write the risk-free return in the model in the form $dr_t$ because the risk-free rate is undefined at the kink.

increases. At the same time, point $\eta^\psi$ where households start participating in capital markets also shifts to the right. Interestingly, the spike in volatility at point $\eta^\psi$ is the highest when exogenous risk $\sigma$ is the lowest.

One may wonder how robust the stationary distribution of $\eta_t$ is to the agents' preferences. Lemma 3 in Appendix C shows that hump of the stationary distribution near $\eta = 0$ exists only under some parameter values when agents have logarithmic utility (specifically, if $\rho > r + \kappa$). Intuitively, because risk-averse experts are more cautious than risk-neutral experts, they use less leverage and the economy is less likely to get stuck near $\eta = 0$.

**Correlation in Asset Prices and "Fat Tails."** Excess volatility due to endogenous risk spills over across all assets held by constrained agents, making asset prices in cross-section significantly more correlated in crisis times. Erb, Harvey, and Viskanta (1994) document this increase in correlation within an international context. This phenomenon is important in practice as many risk models have failed to take this correlation effects into account in the recent housing price crash.[18]

Our model, extended to multiple types of capital, generates this result. Specifically, we can reinterpret equation (1),

$$dk_t = (\Phi(\iota_t) - \delta)k_t\, dt + \sigma k_t\, dZ_t,$$

as the law of motion of fully diversified portfolios of capital held by experts, composed of specific types of capital $l \in [0, 1]$ that follow

$$dk_t^l = (\Phi(\iota_t) - \delta)k_t^l\, dt + \sigma k_t^l\, dZ_t + \hat{\sigma} k_t^l\, dZ_t^l.$$

The diversifiable specific Brownian shocks $dZ_t^l$ are uncorrelated with the aggregate shock $dZ_t$. Because of this, the specific shocks carry no risk premium, and so all types of capital are trading at the same price $q_t$.

In equilibrium the laws of motion of $\eta_t$ and $q_t$ are the same as in our basic model, and depend only on the aggregate shocks $dZ_t$. The return of capital $l$ is given by

$$dr_t^{k,l} = \left(\frac{a - \iota(q_t)}{q_t} + \Phi(\iota(q_t)) - \delta + \mu_t^q + \sigma\sigma_t^q\right) dt + (\sigma + \sigma_t^q)\, dZ_t + \hat{\sigma}\, dZ_t^l.$$

[18] See "Efficiency and Beyond" in The Economist, July 16, 2009.

The correlation between assets $l$ and $l'$,

$$\frac{Cov[q_t k_t^l, q_t k_t^{l'}]}{\sqrt{Var[q_t k_t^l]\, Var[q_t k_t^{l'}]}} = \frac{(\sigma + \sigma_t^q)^2}{(\sigma + \sigma_t^q)^2 + \hat{\sigma}^2},$$

increases in the amount of endogenous risk $\sigma_t^q$. Near the steady state $\eta^*$, $\sigma_t^q = 0$ and so the correlation is $\sigma^2/(\sigma^2+\hat{\sigma}^2)$. All the correlation near $\eta^*$ is fundamental. Away from the steady state, some of the correlation becomes *endogenous*: it arises when both assets are held in portfolios of constrained agents.

The equilibrium patterns of volatility and correlation have implications on the pricing of derivatives. First, since the volatility rises in crisis times, option prices exhibit a "volatility smirk" in normal times. This observation is broadly consistent with empirical evidence (see e.g. Bates (2000)). Put options have a higher implied volatility when they are further out of the money. That is, the larger the price drop has to be for an option to ultimately pay off, the higher is the implied volatility or, put differently, far out of the money options are overpriced relative to at the money options. Second, so called "dispersion trades" try to exploit the empirical pattern that the smirk effect is more pronounced for index options than for options written on individual stocks (Driessen, Maenhout, and Vilkov (2009)). Index options are primarily driven by macro shocks, while individual stock options are also affected by idiosyncratic shocks. The observed option price patterns arise quite naturally in our setting as the correlation across stock prices increases in crisis times.[19]

Since data for crisis periods is limited, option prices provide valuable information about the market participants' implicit probability weights of extreme events.

# 5 Volatility Paradox and Endogenous Leverage

In this section, we explore how equilibrium dynamics change with model parameters, such as fundamental aggregate risk $\sigma$, and with modifications of the model that affect the experts' endogenous leverage. Specifically, we add expert-specific shocks to the model, which increase the experts' borrowing costs and make them leverage-dependent. These idiosyncratic shocks, which play an important role in BGG, allows us to make a better comparison with their analysis. Finally, we introduce financial innovation that allows experts to hedge shocks among each other.

System stability depends on the equilibrium leverage, which is endogenous. The model ex-

[19] In our setting options are redundant assets as their payoffs can be replicated by the underlying asset and the bond, since the volatility is a smooth function in $q_t$. This is in contrast to stochastic volatility models, in which volatility is independently drawn and subject to a further stochastic factor, for which no hedging instrument exists.

hibits a *volatility paradox*: as exogenous risk $\sigma$ decreases, the maximal endogenous risk $\sigma_t^q$ may rise. The reason is that the experts' *endogenous* consumption decisions affect the location of the stochastic steady state $\eta^*$. Thus, lower fundamental exogenous volatility, which leads to higher expert leverage, may result in the rise of the self-generated systemic risk.

When we introduce idiosyncratic shocks, higher borrowing costs naturally lead to lower equilibrium leverage. However, leverage rises again with the introduction of new financial products, like derivatives, that allow experts to (better) hedge idiosyncratic risks. Thus, while financial innovation allows for more efficient risk-sharing, it also affects system stability.

**Volatility Paradox.** A reduction in exogenous cash flow risk $\sigma$ reduces financial frictions. Paradoxically, it can make the economy less stable. That is, it can increase the maximum volatility of prices and the experts' net worth. The reason is that a decline in cash flow volatility encourages experts to increase their leverage by reducing their net worth buffer.

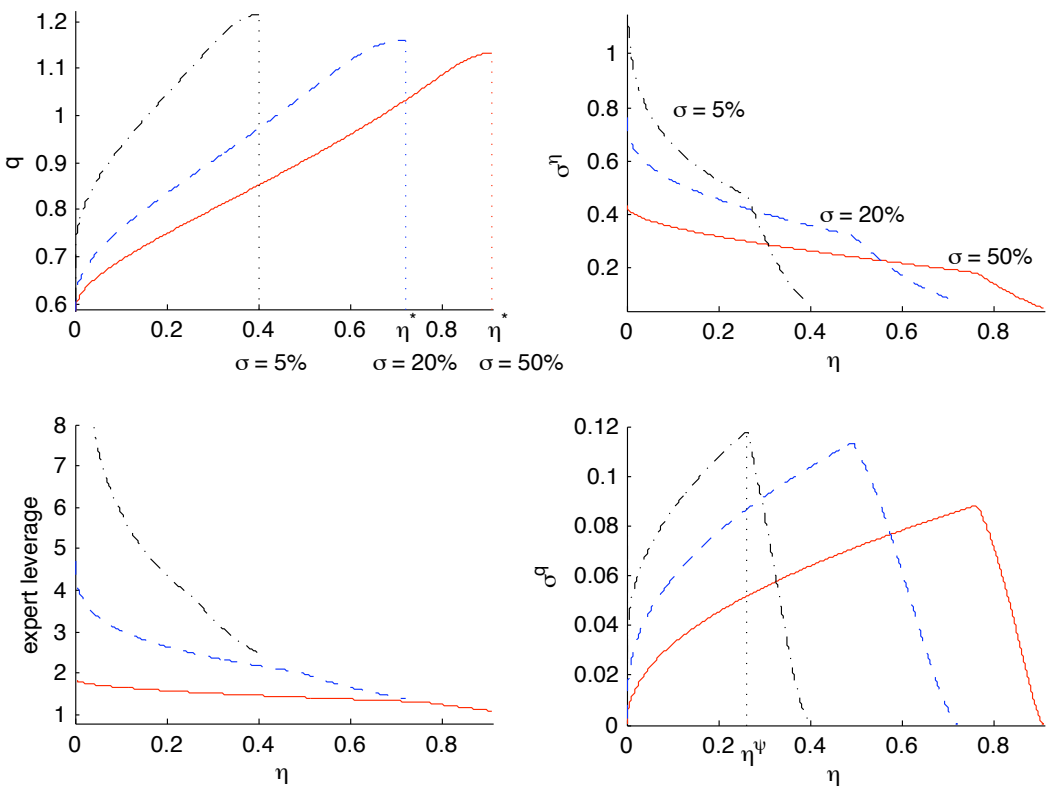


Figure 6: Equilibrium for three levels of exogenous risk, $\sigma = 50\%$, $20\%$ and $5\%$.

Figure 6 illustrates how exogenous risk affects equilibrium prices, volatility and expert leverage in our example of Section 3, for $\sigma = 50\%$, $20\%$ and $5\%$. As exogenous risk *declines*, financial frictions become less severe and the maximal price of capital $q(\eta^*)$ rises. One would expect the system to become more stable. However, to the contrary, experts keep lower net worth buffers and

$\eta^*$ declines as aggregate risk falls. Due to increased leverage, the maximal endogenous risk $\sigma_t^q$ (at point $\eta^\psi$) increases in this example. Likewise, $\sigma_t^\eta$ exhibits the build-up of systemic risk in crises.

The volatility paradox is fairly robust. For example, in Figure 5 in a version of our model with logarithmic utility, the maximal endogenous risk $\sigma_t^q$ also rises as exogenous risk $\sigma$ declines. In fact, the following proposition shows that both with risk-neutrality and logarithmic utility, the equilibrium level of $\sigma_t^\eta$ when $\eta$ is small *increases* as $\sigma$ declines.

**Proposition 4** *Both in the baseline risk-neutral model, and in the variation with logarithmic utility of Section 4,*[20]

$$\sigma_t^\eta \to \frac{\kappa}{\sigma} + O(\sigma), \tag{28}$$

*as* $\eta_t \to 0$, *where* $\kappa = (a - \underline{a})/\underline{q} + \underline{\delta} - \delta$.

This "volatility paradox" is consistent with the fact that the current crisis was preceded by a low volatility environment, referred to as the "great moderation." In other words, in the absence of leverage restrictions, the system is prone to instabilities even and especially when the level of aggregate risk is low.

**Idiosyncratic Risk and Borrowing Costs.** Next, we explore the impact on financial stability of borrowing frictions and financial innovations that facilitate risk management. To study these question we add idiosyncratic jump risk to our baseline setting. With this risk, debt may default and so we can examine credit spreads between risky loans and the risk-free rate. This model variation also allows us to draw a more direct comparison to BGG.[21]

Formally, assume that capital $k_t$ managed by expert $i$ evolves according to

$$dk_t = (\Phi(i_t) - \delta)k_t\, dt + \sigma k_t\, dZ_t + k_t\, dJ_t^i, \tag{29}$$

instead of (1). The new term $dJ_t^i$ is a *compensated* (i.e. mean zero) Poisson process with intensity $\lambda$ and jump distribution $F(y)$, $y \in [-1, 0]$ (if $y = -1$, the expert's entire capital is wiped out). Jumps are independent across experts and cancel out in the aggregate, so that total capital evolves according to the same equation as in the baseline model, (9). As in BGG, jump distribution is *the same* for all experts and does not depend on the balance sheet size.

As BGG, we adopt the costly state verification framework of Townsend (1979) to deal with default. If large enough jump arrives such that the expert's net worth becomes negative, lenders

[20] The term $O(\sigma)$ indicates that the difference $\lim_{\eta\to 0} \sigma_t^\eta - \kappa/\sigma$ converges to 0 at the same rate as $\sigma$ as $\sigma \to 0$.

[21] In BGG the location of the *deterministic* steady state is determined by idiosyncratic jump risk, while aggregate risk is set to zero.

trigger a costly verification procedure to make sure that capital was really destroyed by a shock and not stolen. Verification costs are proportional to the balance sheet size.[22]

We can solve for the equilibrium through the same *two* steps we took in Section 3. First, we extend conditions (H), (E) and (EK) to this setting. Second, we derive the law of motion of the state variable $\eta_t$ that drives the system.

Step 1. Levered experts have to compensate lenders for imperfect recovery and deadweight losses of verification in the event of default. Both of these costs, $L(x_t)$ and $\Lambda(x_t)$ respectively per dollar borrowed, are increasing in leverage $x_t$ (with $L(x) = \Lambda(x) = 0$ if $x \leq 1$). So is the interest $r + L(x_t) + \Lambda(x_t)$ the expert has to pay on debt. Function $L(x_t)$ depends on the intensity and distribution of jumps, and $\Lambda(x_t)$ depends in addition on the verification costs.

The jump term $k_t\, dJ_t^i$ adds $dJ_t^i/q_t$ to the return on capital $dr_t^k$ in (5), but does not affect the expected return. Expert net worth evolves according to

$$dn_t/n_t = x_t\, dr_t^k + (1 - x_t)\,(r + L(x_t) + \Lambda(x_t))\, dt - dc_t/n_t, \tag{30}$$

except that it cannot become negative. If a jump puts $n_t$ into the *negative* territory, it is debt holders who are responsible for the loss so that

$$E[dn_t/n_t] = x_t E[dr_t^k] + (1 - x_t)\,(r + \Lambda(x_t))\, dt - dc_t/n_t. \tag{31}$$

Note the absence of $(1-x_t)L(x_t)$, the expected loss rate of debt holders due to imperfect recovery. That is, because debt holders have to earn the expected return of $r$, the expert's *expected* cost of borrowing is $r+\Lambda(x_t)$, where $\Lambda(x_t)$ reflects the deadweight costs of verification. Thus, the experts' Bellman equation (12) becomes transformed to

$$\rho = \mu_\theta + \max_x \left( xE[dr_t^k]/dt + (1 - x)(r + \Lambda(x)) + x\sigma_\theta(\sigma + \sigma_t^q) \right) \tag{EEK}$$

This equation replaces (E) and (EK) in Proposition 1, and it implies (E) and (EK) if $\Lambda(x) = 0$, i.e. there are no verification costs. In equilibrium, $x_t = \psi_t/\eta_t$ should solve the maximization problem in (EEK). As before, $\theta_t \geq 1$ and experts consume only when $\theta_t = 1$. The household optimal portfolio choice condition (H) remains the same.

[22]The basic costly state verification framework, developed by Townsend (1979) and adopted by BGG is a two-period contracting framework. At date 0, the agent requires investment $I$ from the principal, and at date 1 he receives random output $\tilde{y}$ distributed on the interval $[0, \bar{y}]$. The agent privately observes output $\tilde{y}$, but the principal can verify it at a cost. The optimal contract is a standard debt contract with some face value $D$. If the agent receives $\tilde{y} \geq D$, then he pays the principal $D$ and there is no verification. The principal commits to verify if the agent reports that $\tilde{y} < D$, and receives the entire output.

Step 2. Aggregating the experts' net worth, equation (31) implies that

$$dN_t = \psi_t q_t K_t \, dr_t^k - (\psi_t q_t K_t - N_t)(r + \Lambda(\psi_t/\eta_t)) \, dt - dC_t.$$

With the extra term $\Lambda(\psi_t/\eta_t)$, an analogue of Lemma 2 leads to the formula

$$\frac{d\eta_t}{\eta_t} = \frac{\psi_t - \eta_t}{\eta_t}(dr_t^k - r\,dt - \Lambda(\psi_t/\eta_t)\,dt - (\sigma + \sigma_t^q)^2\,dt) + \frac{a - \iota(q_t)}{q_t}\,dt + (1 - \psi_t)(\underline{\delta} - \delta)\,dt - d\zeta_t. \quad (32)$$

**Borrowing Costs and Equilibrium Dynamics.** Figure 7 illustrates the equilibrium, in which we add borrowing costs, due to costly verification, of $\Lambda(x) = \xi(x-1)$, with $\xi = 0$, 0.01 and 0.02, to the example of Section 3, with $\sigma = 10\%$. The effects of borrowing frictions on equilibrium dynamics may seem surprising at first. One may guess that these frictions, which make it harder for experts to get funding, particularly in downturns, cause amplification effects to become more severe.

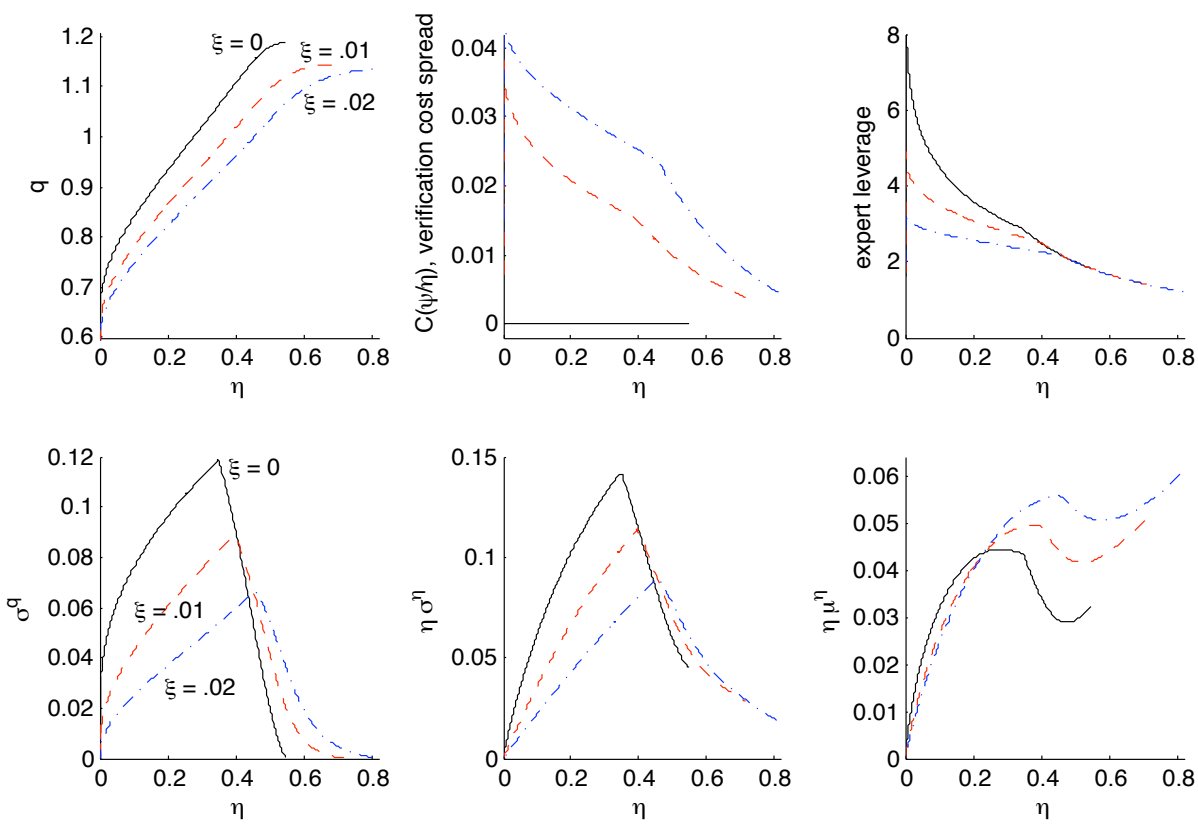


Figure 7: Equilibrium with $\sigma = 10\%$ and $\Lambda(x) = \xi(x-1)$, $\xi = 0$, .01 and .02.

In fact, the opposite is true: while borrowing frictions depress prices and investment, they actually lead to a *more stable* equilibrium. The amount of endogenous risk $\sigma_t^q$ drops significantly because expert leverage decreases and, to a lesser extent, because prices in booms are lower.

Surprisingly, the drift of $\eta$ rises for high $\eta$, despite higher borrowing costs, because the competition among experts becomes less intense due to lower leverage.

If borrowing costs $\Lambda(x)$ become large enough as $x \to \infty$, then even as $\sigma \to 0$, experts cannot hold all capital in the economy when $\eta$ is close to 0. This leads to a non-degenerate *deterministic steady state* $\eta^0 > 0$, with constant price $q^0$ and all capital held by experts, i.e. $\psi = 1$. In the example in Figure 7, $\eta^0 = 0.8165$ when $\xi = 0.02$, very close to the stochastic steady state of $\eta^* = 0.8193$. As $\xi$ decreases towards 0, $\eta^0$ converges to 0, while $\eta^*$ converges to 0.5472.

**Proposition 5** *Leverage $x^0 = 1/\eta^0$ at the non-degenerate* deterministic steady state *of the model with idiosyncratic shocks is characterized by equation*

$$\rho - r = x^0(x^0 - 1)\Lambda'(x^0) + \Lambda(x^0), \tag{33}$$

*and the price of capital is characterized by*

$$\max_{\iota} \frac{a - \iota}{q^0} + \Phi(\iota) - \delta = r + \Lambda(x^0) + (x^0 - 1)\Lambda'(x^0). \tag{34}$$

**Proof.** When $\sigma = 0$, then $\sigma_t^q = \sigma_t^\theta = 0$ and system dynamics is *deterministic*. Taking the first-order condition with respect to $x$ in the Bellman equation (EEK), we get

$$E[dr_t^k]/dt = r + \Lambda(x) + (x - 1)\Lambda'(x), \tag{35}$$

where $x = \psi/\eta$ in equilibrium. Furthermore, at $\eta^0$ we have $\mu_t^\theta = 0$, since it is an absorbing state of the system. Using (35) and $\mu_t^\theta = 0$, the Bellman equation (EEK) at $\eta = \eta^0$ implies (33). Finally, since $\mu^q = 0$ at $\eta^0$, the left hand side of (34) is the expected return on capital, and so (35) implies (34). ■

Proposition 5 allows us to make a more direct comparison of our model with BGG and KM, who explore system dynamics near the deterministic steady state. There are some *cosmetic* differences *near* the steady state, because our system responds differently to unanticipated positive and negative shocks near $\eta^0$. Positive shocks lead to immediate payouts and otherwise do not affect the system, while negative shocks are amplified and lead to gradual recovery. At the end of this section, we discuss adjustments to our model that map more directly to BGG and KM, to replicate their dynamics near the steady state.

There are also some *essential* differences away from the steady state. For example, BGG focus on an environment with *small* aggregate shocks and *large* idiosyncratic shocks, in which the deterministic steady state is nondegenerate. Under these assumptions, leverage stays bounded and the volatility paradox disappears. The system becomes a collection of stand-alone units, in which idiosyncratic risks are primary determinants of leverage. These units occasionally fail individually, in a non-systemic way.

However, in contrast to BGG, the analysis below shows that if the system becomes integrated and individual experts share idiosyncratic risks, then aggregate risk becomes the primary determinant of systemwide leverage and systemic risk reappears.

**Risk Management and Financial Innovation.** Next, we explore the impact of financial innovations that allow experts to share risk better, and in particular hedge idiosyncratic risks. These products can involve *securitization*, including pooling and trenching, credit default swaps, and various options and futures contracts. We find that risk sharing among experts reduces inefficiencies from idiosyncratic risk on one hand, but on the other hand emboldens them to live with smaller net worth buffers and higher leverage. This leads to an increase of systemic risk. Ironically, tools intended for more efficient risk management may make the system less stable.

Assume that all shocks, including idiosyncratic jumps $dJ_t^i$ are observable and contractible among experts, but not between experts and households. Then experts can trade insurance contracts that cover jump losses $dJ_t^i$ on expert $i$'s capital for the risk premium of $\omega_t^i$. We can similarly allow experts to contract on the aggregate risk $dZ_t$.

**Proposition 6** *If experts can contract on all shocks among each other, then the equilibrium in a setting with idiosyncratic shocks equivalent to that in the baseline setting. Experts fully hedge their idiosyncratic risks, which carry the risk premium of zero.*

**Sketch of Proof.** Idiosyncratic risk of any expert $i$ carries the risk premium of zero because it can be fully diversified among other experts. Given that, experts choose to insure their idiosyncratic risks enough to make their debt risk-free. With borrowing frictions eliminated, the laws of motion of $\eta_t$ and functions $q(\eta)$, $\theta(\eta)$ and $\psi(\eta)$ are the same as in the baseline setting with $\Lambda(x) = 0$. Contracts on aggregate risk among expert do not change the equilibrium, as they do not alter the total aggregate risk exposure of the expert sector. ■

When experts can contract on idiosyncratic shocks, then they face the cost of borrowing of only $r$ and equilibrium dynamics ends up being the same as in our baseline model. Thus, in the

example of Figure 7, for any function $\Lambda(x)$ the equilibrium becomes transformed to that described by the blue plot, which corresponds to the parameter $\xi = 0$.

Clearly, instruments that help experts share risks eliminate the deadweight losses due to costly state verification in this model. These instruments also lead to greater systemic risk, because experts endogenously increase leverage by lowering their net worth buffers. If instability harms the economy, e.g. from (not yet modeled) spillovers to other sectors, then the overall effect on welfare is not clear. Whichever setting (with or without risk sharing) is better, certainly risk sharing in combination with a policy that encourages experts to postpone payout leads to even higher welfare (see Section 6).

The link between financial innovation and aggregate leverage has also been illustrated concurrently by Gennaioli, Shleifer and Vishny (2010), who build a two-period model in which agents ignore the possibility of certain bad events. In particular, they interpret securitization as one important form of risk-sharing.

**Comparison with BGG and KM.** One key feature of our model that differs from those of BGG and KM is that experts choose payouts/consumption *endogenously*, taking into account the need for a net worth buffer to absorb future risks. In contrast, in BGG expert payouts are determined by an exogenous parameter, which determines the rate at which experts are forced to retire and consume their net worth. In KM, expert leverage is always given by an exogenous constraint. With these assumptions, BGG and KM do not capture how equilibrium leverage changes with aggregate risk. Thus, even fully solved versions of these models (not just near the steady state) would not be able to generate the volatility paradox.

To clarify driving forces behind our results, below we discuss modifications of our continuous-time model that most closely replicate the discrete-time assumptions of BGG and KM.

In BGG, experts face borrowing costs due to idiosyncratic shocks, and they are equally patient as households i.e. $\rho = r$. These assumptions alone would imply that both the deterministic and stochastic steady states are at $\eta^0 = \eta^* = 1$. Since experts earn higher returns than households, they would eventually overwhelm the economy *unless* they consume at a higher rate than households. To generate a deterministic steady state $\eta^0 < 1$, BGG assume also that with Poisson intensity $\phi > 0$ experts are hit by idiosyncratic Poisson shocks that force them to consume their net worth. We can easily accommodate this feature into our model for *comparison* with BGG.[23]

[23] With these assumptions, the experts' Bellman equation (EEK) becomes transformed to

$$r = \mu_\theta - \phi(\theta_t - 1)/\theta_t + \max_x \left( xE[dr_t^k]/dt - (x-1)(r + \Lambda(x)) + x\sigma^\theta(\sigma + \sigma_t^q) \right), \qquad \text{(EEK}'\text{)}$$

where $(\theta_t - 1)/\theta_t$ is the fraction of value experts lose in the event of an exit shock, while the household equation

Parameter $\phi$ directly affects the location of the deterministic steady state in BGG. This is a convenient feature under log-linearization, which characterizes approximate dynamics only near the steady state. However, focusing on a particular part of the state space through changes of an exogenous parameter $\phi$ also has serious drawbacks, as it completely changes system dynamics. The dynamics of the wealth distribution become exogenously determined, instead of endogenously generated through consumption and investment decisions of individual agents.

In the setting of KM, experts can borrow at rate $r$, but their leverage cannot exceed $\bar{x}$. This can be captured in our model by setting $\Lambda(x) = 0$ on $[0, \bar{x}]$ and $\infty$ on $(\bar{x}, \infty)$. This assumption leads to a deterministic steady state of $\eta^0 = 1/\bar{x}$, at which experts lever up to the constraint.[24]

# 6 Efficiency, Externalities and Macroprudential Policies

The fact that financial frictions lead to systemic instability and excess volatility does not necessarily prescribe strict financial regulation. Making the system more stable might stifle long-run economic growth. To study financial regulation one has to conduct a welfare analysis. This section makes a first small step in this direction.

We start by quantifying welfare. There may be inefficiencies within the financial sector, and spillovers from the financial sector into the real economy. Since these spillovers may be important, we design a simple way to include them into our model (see expression (36)).

Then, we study the effects of policies on the equilibrium outcome. We prove a simple result that a social planner, who faces the same constraints as the market, can attain the first-best efficient outcome by making decisions for the agents and implementing appropriate transfers. However, this policy is not realistic or practical as it interferes with the economy in significant ways. We then consider specific simple policies that are closer to those implemented in reality. These policies impose small constraints on the agents' actions, but otherwise let the economy run on its own.

**Inefficiencies and Externalities.** Incomplete-market settings such as ours are prone to pecuniary externalities. These subtle externalities arise because individual market participants take prices as given, but as a group they affect them. Stiglitz (1982), Geanakoplos and Polemarchakis (1986), and Bhattacharya and Gale (1987) were among the first to highlight the inefficiency of

(H) remains the same. The law of motion of $\eta_t$ (32) also remains the same, except that expert consumption is now exogenously given by $d\zeta = \phi\, dt$.

[24] There are many other interesting modifications of our model, which we do not explore here. For example, one can assume that margins that depend on the value-at-risk (VaR) as in Brunnermeier and Pedersen (2009) and Shin (2010). In the former margins increase with endogenous price volatility. These effects can be captured in our model by assuming that the cost of borrowing is also a function of $\sigma_t^q$.

pecuniary externalities. In particular, the fire-sale externality is a pecuniary externality that arises when in crisis (i) experts are able to sell assets to another sector, e.g. vulture investors, the government or the household sector (in our case) and (ii) the new asset buyers provide a downward-sloping demand function. When levering up ex-ante, financial experts do not take into account that in crisis, their own fire sales will depress prices of assets held by other experts. This effect leads to excess leverage since experts take fire-sale prices as given, i.e. a social planner would lever up less. Recent applications of this inefficiency due to pecuniary externalities within a finance context are Lorenzoni (2008) and Jeanne and Korinek (2010).[25]

As we discussed in Section 3, the sum of utilities of all agents under the first-best outcome is $\bar{q}K$. It is attained if experts consume their net worth at time 0, and manage capital forever thereafter while investing at rate $\iota(\bar{q})$. The maximal *equilibrium* sum of utilities of all agents is lower: it is $q(\eta^*)K$ (see (20)).

In addition to externalities within financial markets, there may be very direct spillovers outside the financial markets into the real economy. This externality is not so subtle, and here we present a crude way to incorporate it into our model, in order to be able to quantify the effects of policies on welfare. Specifically, assume the activities of experts and households who manage capital create a positive spillover on the labor market that is proportional to total output. Formally, we define the spillover as

$$\kappa E\left[\int_0^\infty e^{-rt}(\psi_t a + (1-\psi_t)\underline{a})K_t\, dt\right], \tag{36}$$

where $\kappa \geq 0$ is a constant. Then crises lower the value of this expression, because the reallocation of capital to households lowers output, and depressed capital prices lead to lower investment and growth. Our goal behind expression (36) is to avoid a fully-blown model of spillovers, and instead have a transparent reduced-form expression that is easy to interpret.[26] With the spillovers, the sum of utilities of all agents under the first-best outcome is $\bar{\bar{q}}K$, where

$$\bar{\bar{q}} = \max_{\iota} \frac{(1+\kappa)a - \iota}{r + \delta - \Phi(\iota)}. \tag{37}$$

**Constrained Efficiency.** It turns out that a social planner can achieve the first-best efficient outcome while respecting the same financing frictions with respect to equity issuance that individual experts face. To formalize this result, we define a set of constrained-feasible policies, under which the central planner controls prices and the agents' consumption and investment choices, but

[25] Davila (2011) highlights distinctions between various types of pecuniary externalities.

[26] Such spillover effects can be obtained by explicitly including a labor sector into the model. If so, then equilibrium wages have to depend on capital productivity, and so expression (36) captures the essence of spillovers that would arise in such a model.

treats all experts and households symmetrically.

**Definition 2** *A symmetric* constrained-feasible *policy is described by a group of stochastic processes on the filtered probability space defined by the Brownian motion* $\{Z_t, t \geq 0\}$*: the price process* $q_t$*, investment rates* $\iota_t$ *and* $\underline{\iota}_t$*, capital allocations* $\psi_t K_t$ *and* $(1-\psi_t)K_t$*, consumptions* $dC_t$ *and* $d\underline{C}_t$*, and transfers* $d\tau_t$ *such that*

*(i) representative expert net worth* $dN_t = d\tau_t + \psi_t K_t q_t \, dr_t^k - dC_t$ *stays nonnegative,*

*(ii) representative household net worth is defined by* $\underline{N}_t = q_t K_t - N_t$*,*

*(iii) and the resource constraints are satisfied, i.e.*

$$\frac{dC_t + d\underline{C}_t}{K_t} = (\psi_t(a-\iota_t)+(1-\psi_t)(\underline{a}-\underline{\iota}_t))dt, \quad \frac{dK_t}{K_t} = (\psi_t(\Phi(\iota_t)-\delta)+(1-\psi_t)(\Phi(\underline{\iota}_t)-\underline{\delta}))dt+\sigma dZ_t.$$

Note that since the sum of net worth adds to the total wealth in the economy $q_t K_t$, aggregate transfers across both sectors are zero. Moreover, because of transfers we can set the risk-free rate to zero, without loss of generality. The following proposition characterizes constrained-feasible policies that achieve the first-best allocation.

**Proposition 7** *Constrained-feasible policies that achieve a first-best outcome are those that satisfy* $\psi_t = 1$ *and* $\iota_t = \iota(\bar{q})$ *for all* $t \geq 0$*, and* $dC_t = 0$ *for all* $t > 0$*, although experts may consume positively at time 0, and transfers* $d\tau_t$ *are chosen to keep the net worth of experts nonnegative.*[27]

Proposition 7 generates first-best outcomes only in the setting without spillovers. With spillovers of the form (36), the social planner should choose the investment rate that solves (37).

**Proof.** The policies outlined in the proposition are constrained feasible because the experts' net worth stays nonnegative. They attain first-best outcomes because experts consume *only* at time 0, because all capital is always allocated to experts and because experts are forced to invest at the first-best rate of $\iota_t = \iota(\bar{q})$. Note that after time 0, experts may receive large transfers of wealth to keep their net worth nonnegative, but they are not allowed to consume any of their net worth. ■

[27] One may be wondering whether the suggested policies preserve incentive compatibility. According to our microfoundation of balance sheets in Appendix A, experts must retain full equity stakes in their projects because otherwise they would divert some of the capital and use it in another firm, while original outside equity holders suffer losses. Under any policy of Proposition 7, such a deviation would not enhance the expert's utility because the social planner controls the experts' consumption and sets it to zero for all $t > 0$. Even if experts could secretly *consume* diverted funds, a policy of Proposition 7 would still achieve a first best outcome as long as transfers are chosen to keep the net worth of a representative expert $N_t$ at zero.

The main role of transfers is to ensure that the net worth of experts stays nonnegative. Interestingly, this can also be achieved by an appropriate choice of the price process $q_t$. That is, price stabilization policies that prevents a decline of $q_t K_t$ after an adverse shock to $K_t$ can enhance welfare since they reduce the volatility of experts' net worth. For example, by picking $\sigma^q = -\sigma$, the planner can make the experts' net worth non-random.

Of course, these policies work only if in fact the social planner can completely enforce the experts' consumption and investment decisions. If the planner cannot fully control these decisions, then many well-intentioned policies could backfire. For example, by the logic of our volatility paradox, price stabilization policies could lead to higher endogenous leverage, and thus a less stable equilibrium. Next, we investigate the effects of policies that place only mild restrictions on the equilibrium behavior, such as capital requirements and taxes on dividends.

**Policy Experiments.** Many policies have been proposed or implemented with the goal of improving financial stability. Some, such as equity infusions, asset purchases or funding subsidies by the central bank (see Gertler and Kiyotaki (2011)) are aimed at recapitalizing financial institutions in crises. Others are aimed at controlling the overall risk within the financial system. When considering policies, it is important to understand how they affect the entire equilibrium. For that purpose, our framework is particularly convenient as it captures the nonlinearities that arise when the economy enters a crisis regime, and the endogeneity of leverage at the steady state (i.e. in normal times).

Our policy experiments suggest that good regulation (1) in boom times encourages banks to retain earnings, so that they can absorb losses more easily in the event of a possible downturn and (2) in downturns allows banks to take on more leverage, in order to stabilize the market. In particular, a leverage constraint alone can easily be counterproductive, as it is unlikely to bind in booms and it leads to fire sales in downturns.[28]

We performed numerical experiments, in which we imposed various leverage constraints and restrictions on the experts' consumption. A leverage constraint prohibits experts from taking on leverage greater than $\bar{x}(\eta)$. Generally experts respond to leverage constraints by accumulating more net worth, so $\eta^*$ goes up - an effect that could potentially stabilize the system and increase welfare. However, in our numerical experiments this effect is small, e.g. a flat constraint that binds on 70% of the state space (in downturns) increases $\eta^*$ by only 2%. Leverage constraints also create many inefficiencies, including capital misallocation and depressed prices that cause

[28]Capital requirements restrict leverage by weighing assets according to risk. In our setting, they would restrict leverage the most in times of highest endogenous risk, and have potentially destabilizing effects.

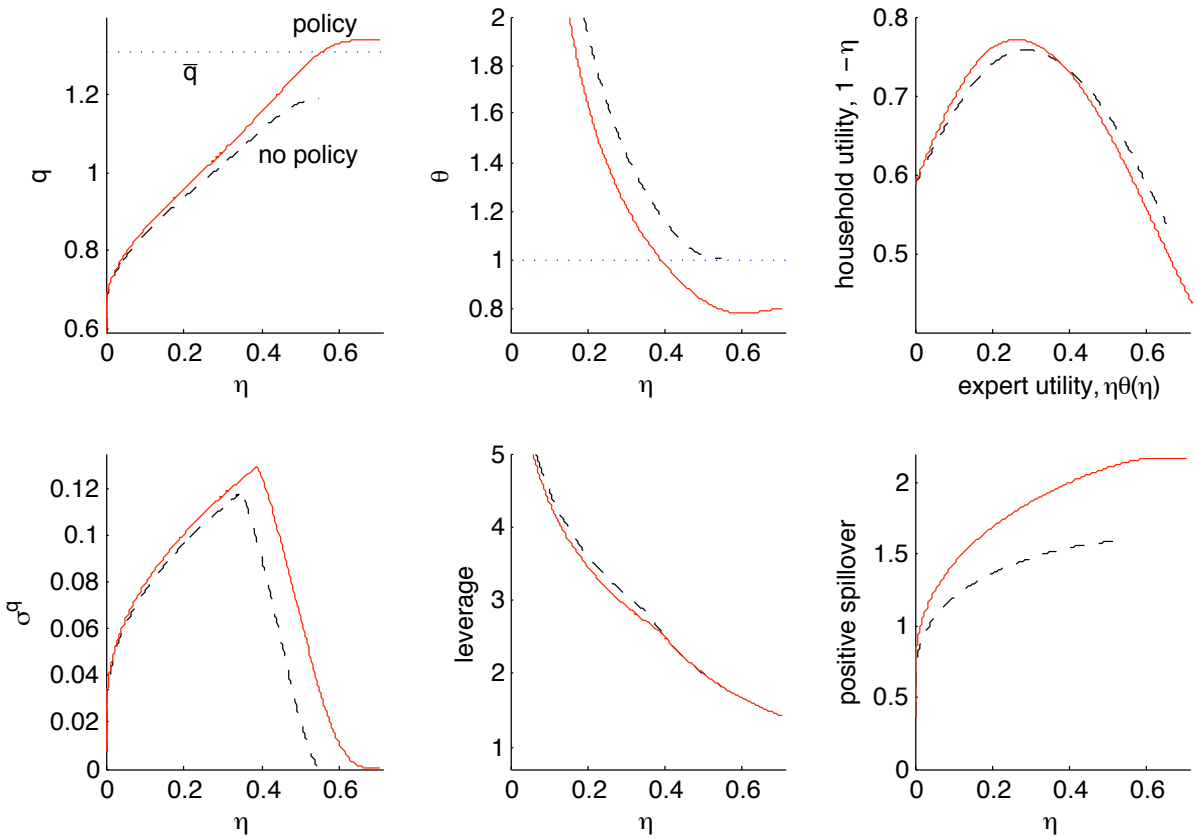


Figure 8: Equilibrium without and with restrictions on payouts.

underinvestment. The overall effect on welfare is mostly negative, although some countercyclical leverage restrictions do appear to mildly improve the spillover measure (36).

We also consider policies that encourage experts to retain earnings, i.e. raise the level $\eta^*$ where experts consume. Figure 8 shows the effects of a policy that forces experts to retain all earnings until $\eta^* = 0.7$ in the example of Section 3 for $\sigma = 10\%$ (in which $\eta^* = 0.5472$ in the absence of policy interventions). As experts accumulate more net worth, crises become less likely. This improves welfare, according to the spillover measure (36) plotted on the bottom right panel.[29] Aside from improved welfare, the policy leads to a number of other interesting consequences as it affects the entire equilibrium. As experts are forced to retain more net worth, the price of capital rises, and even becomes greater than $\bar{q}$ at the steady state. The marginal value of expert net worth $\theta(\eta)$ falls and becomes non-monotonic near $\eta^*$. As a result, risk premia near $\eta^*$ become negative.[30] Crises episodes become less frequent, but more severe. The maximal endogenous risk in crises increases since prices have more room to fall. Moreover, the drift of $\eta_t$, which reflects the speed of

[29] Lower values of $\eta^*$, such as 0.56, even improve the welfare within the financial system, plotted on the top right panel of Figure 8.

[30] As $\eta_t$ gets close to $\eta^*$, $\theta(\eta_t) < 1$ and experts wish they could pay out funds. At $\eta^*$, experts are able to pay out some funds, and so $\theta(\eta_t)$ increases. However, payouts are restricted to be just sufficient for $\eta_t$ to reflect at $\eta^*$.

recovery, is lower as both the price of capital is higher and expert leverage is lower in downturns.

# 7 Conclusions

Events during the great liquidity and credit crunch in 2007-10 have highlighted the importance of financing frictions for macroeconomics. Unlike many existing papers in macroeconomics, our analysis is not restricted to local effects around the steady state. Importantly, we show that endogenous risk due to adverse feedback loops is significantly larger away from the steady state. This leads to non-linearities: small shocks keep the economy near the stable steady state, but large shocks put the economy in the unstable crisis regime characterized by liquidity spirals. The economy is prone to instability regardless of the level of aggregate risk because leverage and risk-taking are endogenous. As aggregate risk goes down, equilibrium leverage goes up, and amplification loops in crisis regimes become more severe - a volatility paradox. In an environment with idiosyncratic and aggregate risks, equilibrium leverage also increases with diversification and with financial instruments that facilitate the hedging of idiosyncratic risks. Thus, paradoxically, tools designed to better manage risks may increase systemic risk.

Policy interventions can make crisis episodes less likely, although many seemingly reasonable policies can harm welfare. Policies for crisis episodes alone, such as those aimed at recapitalizing the financial system, can increase risk-taking incentives ex-ante. More surprisingly, simple restrictions on leverage may do more harm than good, as they bind only in downturns and may have little impact on behavior in booms. Our numerical experiments suggest that policies, which encourage financial institutions to retain earnings longer in booms, appear to be most effective.

# References


Adrian, T., and M. K. Brunnermeier (2010): "CoVaR," *working paper*.

Allen, F., and D. Gale (2007): *Understanding Financial Crisis*. Oxford University Press.

Bates, D. S. (2000): "Post-'87 Crash Fears in the S&P Futures Market," *Journal of Econometrics*, 94, 181–238.

Bernanke, B., M. Gertler, and S. Gilchrist (1999): "The Financial Accelerator in a Quantitative Business Cycle Framework," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1, chap. 21, pp. 1341–1393. Elsevier.

Bernanke, B. S., and M. Gertler (1989): "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, 79(1), 14–31.

Bhattacharya, S., and D. Gale (1987): "Preference Shocks, Liquidity, and Central Bank Policy," in *New Approaches to Monetary Economics*, ed. by W. A. Barnett, and K. J. Singleton, pp. 69–88. Cambridge University Press.

Bolton, P., and D. S. Scharfstein (1990): "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, 80(1), 93–106.

Brunnermeier, M. K., T. Eisenbach, and Y. Sannikov (2012): "Macroeconomics with Financial Frictions: A Survey," in *Advances in Economics and Econometrics, 10th World Congress of the Econometric Society*, ed. by D. Acemoglu, M. Arellano, and E. Dekel. Cambridge University Press, Cambridge, UK.

Brunnermeier, M. K., G. Gorton, and A. Krishnamurthy (2010): "Risk Topography," working paper.

Brunnermeier, M. K., and L. Pedersen (2009): "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22(6), pp.2201–2238.

Caballero, R., and A. Krishnamurthy (2004): "Smoothing Sudden Stops," *Journal of Economic Theory*, 119(4), pp.104–127.

Carlstrom, C. T., and T. S. Fuerst (1997): "Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis," *American Economic Review*, 87(5), 893–910.

Christiano, L., M. Eichenbaum, and C. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Christiano, L., R. Motto, and M. Rostagno (2003): "The Great Depression and the Friedman-Schwartz Hypothesis," *Journal of Money, Credit & Banking*, 35(6), pp.1119–1197.

——— (2007): "Shocks, Structures or Monetary Policies? The Euro Area and US After 2001," *Journal of Economic Dyanmics and Control*, 32(8), 2476–2506.

Curdia, V., and M. Woodford (2009): "Credit Spreads and Monetary Policy," working paper.

DeMarzo, P. M., and Y. Sannikov (2006): "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *The Journal of Finance*, 61(6), 2681–2724.

Diamond, D. W. (1984): "Financial Intermediation and Delegated Monitoring," *The Review of Economic Studies*, 51(3), 393–414.

Diamond, D. W., and P. H. Dybvig (1983): "Bank Runs, Deposit Insurance, and Liquidity," *The Journal of Political Economy*, 91(3), 401–419.

Driessen, J., P. Maenhout, and G. Vilkov (2009): "The Price of Correlation Risk: Evidence from Equity Options," *Journal of Finance*, 64(3), 1377–1406.

Erb, C. B., C. R. Harvey, and T. E. Viskanta (1994): "Forecasting International Equity Correlations," *Financial Analysts Journal*, 50(6), 32–45.

Fisher, I. (1933): "The Debt-Deflation Theory of Great Depressions," *Econometrica*, 1(4), 337–357.

Fudenberg, D., B. Holmström, and P. Milgrom (1990): "Short-term Contracts and Long-term Agency Relationships," *Journal of Economic Theory*, 51(1), 1–31.

Geanakoplos, J. (1997): "Promises Promises," in *The Economy as an Evolving Complex System II*, ed. by W. B. Arthur, S. Durlauf, and D. Lane, pp. 285–320. Addison-Wesley, Reading, MA.

——— (2003): "Liquidity, Defaults, and Crashes," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume 2*, vol. 2 of *Econometric Society Monographs*, pp. 170–205. Cambridge University Press.

Geanakoplos, J., and H. Polemarchakis (1986): "Existence, Regularity, and Constrained Suboptimality of Competitive Allocations when the Asset Market is Incomplete," in *Uncertainty, Information and Communication: Essays in Honor of Kenneth J. Arrow*, ed. by W. P. Heller, R. M. Starr, and D. A. Starrett, vol. 3, pp. 65–95. Cambridge University Press.

Gertler, M., and P. Karadi (2009): "A Model of Unconventional Monetary Policy," working paper.

Gertler, M., and N. Kiyotaki (2011): "Financial Intermediation and Credit Policy in Business Cycle Ana," in *Handbook of Monetary Economics*. North Holland Elsevier.

Gromb, D., and D. Vayanos (2002): "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs," *Journal of Financial Economics*, 66(2-3), 361 – 407.

He, Z., and A. Krishnamurthy (2010a): "Intermediary Asset Prices," working paper.

——— (2010b): "A Model of Capital Crisis," working paper.

He, Z., and W. Xiong (2009): "Dynamic Debt Runs," working paper.

Holmström, B., and J. Tirole (1997): "Financial Intermediation, Loanable Funds, and the Real Sector," *The Quarterly Journal of Economics*, 112(3), pp.663–691.

Jaimovich, N., and S. Rebelo (2009): "Can News about the Future Drive the Business Cycle," *American Economic*, 99(4), 1097–1118.

Jeanne, O., and A. Korinek (2010): "Managing Credit Booms and Busts: A Pigouvian Taxation Approach," working paper.

Keynes, John, M. (1936): *The General Theory of Employment, Interest and Money*. Macmillan, London.

Kindleberger, C. (1993): *A Financial History of Western Europe*. Oxfod University Press, 2nd edn.

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

Kocherlakota, N. R. (2000): "Creating Business Cycles Through Credit Constraints," *Federal Reserve Bank of Minneapolis Quarterly Review*, 24(3), 2–10.

Lorenzoni, G. (2008): "Inefficient Credit Booms," *The Review of Economic Studies*, 75(3), 809–833.

Mendoza, E. G. (2010): "Sudden Stops, Financial Crisis, and Leverage," *American Economic Rev*, 100, 1941–1966.

Minsky, H. (1986): *Stabilizing and Unstable Economy*. Yale University Press.

Myerson, R. B. (2010): "A Model of Moral Hazard Credit Cycles," Working Paper, University of Chicago.

Rampini, A., and S. Viswanathan (2011): "Collateral, Risk Management, and the Distribution of Debt Capacity," *Journal of Finance*, forthcoming.

Sannikov, Y. (2012): "Dynamic Security Design and Corporate Financing," in *Handbook of Economics and Finance*, vol. 2. Elsevier-Northholland.

Shin, H. (2010): *Risk and Liquidity*. Oxford University Press, Oxford, UK, Clarendon Lecture.

Shleifer, A., and R. W. Vishny (1992): "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *The Journal of Finance*, 47(4), 1343–1366.

——— (1997): "The Limits of Arbitrage," *The Journal of Finance*, 52(1), 35–55.

——— (2010): "Unstable Banking," *Journal of Financial Economics*, 97(3), 306 – 318, The 2007-8 financial crisis: Lessons from corporate finance.

Stiglitz, J. E. (1982): "The Inefficiency of the Stock Market Equilibrium," *The Review of Economic Studies*, 49(2), 241–261.

Townsend, R. M. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21(2), 265–293.

# Appendix

## A Microfoundation of Balance Sheets and Intermediation

This section describes the connection between balance sheets in our model and agency problems. Extensive corporate finance literature (see Townsend (1979), Bolton and Scharfstein (1990), DeMarzo and Sannikov (2006), Biais, Mariotti, Plantin and Rochet (2007), or Sannikov (2012) for a survey of these models). suggests that agency frictions increase when the agent's net worth falls. In a macroeconomic setting, this logic points to the aggregate net worth of end borrowers, as well as that of intermediaries.

Incentive provision requires the agent to have some "skin in the game" in the projects he manages. When projects are risky, it follows that the agent must absorb some of project risk through net worth. Some of the risks may be identified and hedged, reducing the agent's risk exposure. However, whenever some aggregate risk exposures of constrained agents cannot be hedged, macroeconomic fluctuations due to financial frictions arise, as these residual risks have aggregate impact on the net worth constrained agents.

Our baseline model assumes the simplest form of balance sheets, in which constrained agents (experts) absorb all risk and issue just risk-free debt. Qualitatively, however, our results still hold if experts can issue some outside equity and even hedge some of their risks, as long as they *cannot* hedge *all* the risks. Quantitatively, the assumption regarding equity issuance matters: if experts can issue more equity or hedge more risks, then they can operate efficiently with much lower net worths. This does not necessarily lead to a more stable system, because, as we saw in Section 5, the steady state in our model is endogenous. Agents who can function with lower wealth accumulate lower net worth buffers. Thus, we expect that our baseline model with simple balance sheets captures many characteristics of equilibria of more general models.

To illustrate the connection between balance sheets and agency models, first, we discuss the agency problem with direct lending from investors to a single agent. Second, we illustrate agency problems that arise with intermediaries. In this case, the net worth of intermediaries matters as well. At the end of this section we discuss contracting with idiosyncratic jump risk, which is relevant for Section 5.

**Agency Frictions between an Expert and Households.** Assume that experts are able to divert capital returns at rate $b_t \in [0, \infty)$. Diversion is inefficient: of the funds $b_t$ diverted, an expert is able to recover only a portion $h(b_t) \in [0, b_t]$, where $h(0) = 0$, $h' \leq 1$, $h'' \leq 0$. Net of diverted

funds, capital generates the return of

$$dr_t^k - b_t\, dt$$

and the expert receives an income flow of $h(b_t)\, dt$.

If capital is partially financed by outside equity held by households, then households receive the return of

$$dr_t^k - b_t\, dt - f_t\, dt,$$

where $f_t$ is the fee paid to the expert. When the expert holds a fraction $\Delta_t$ of equity, then per dollar invested in capital he gets

$$\Delta_t(dr_t^k - b_t\, dt) + (1-\Delta_t) f_t\, dt + h(b_t)\, dt.$$

The incentives with respect to diversion are summarized by the first-order condition $\Delta_t = h'(b_t)$, which leads to a weakly decreasing function $b(\Delta)$ with $b(1) = 0$. In equilibrium, the fee $f_t$ is chosen so that household investors get the expected required return of $r$ on their investment, i.e. $f_t = E[dr_t^k]/\, dt - b_t - r$. As a result, the return on the expert's equity stake in capital (including the benefits of diversion) is

$$\frac{\Delta_t(dr_t^k - b_t\, dt) + (1-\Delta_t) f_t\, dt + h(b_t)\, dt}{\Delta_t} = \frac{E[dr_t^k] - r\, dt}{\Delta_t} + r\, dt + (\sigma + \sigma_t^q) dZ_t - \frac{b(\Delta_t) - h(b(\Delta_t))}{\Delta_t} dt,$$

where $\frac{b(\Delta_t)-h(b(\Delta_t))}{\Delta_t}$ is the deadweight loss rate due to the agency problem. The law of motion of the expert's net worth in this setting is of the form

$$\frac{dn_t}{n_t} = x_t \left( \frac{E[dr_t^k] - r\, dt}{\Delta_t} + r\, dt + (\sigma + \sigma_t^q)\, dZ_t - \frac{b(\Delta_t) - h(b(\Delta_t))}{\Delta_t}\, dt \right) + (1 - x_t) r\, dt - d\zeta_t, \qquad (38)$$

where $x_t$ is the portfolio allocation to inside equity and $d\zeta_t$ is the expert's consumption rate. It is convenient to view equation (38) as capturing the issuance of equity and risk-free debt. However, it is possible to reinterpret this capital structure in many other ways, since securities such as risky debt can be replicated by continuous trading in the firm's stock and risk-free debt.

Equation (EK) generalizes to

$$\max_{\Delta} \frac{E[dr_t^k]/dt - r - (b(\Delta) - h(b(\Delta)))}{\Delta} = -\sigma_t^\theta(\sigma + \sigma_t^q),$$

and determines optimal equity issuance. Our results suggest that, as risk premia $E[dr_t^k]/dt - r$

rise in downturns, the experts' equity retention $\Delta$ decreases and deadweight losses increase.[31]

Our baseline setting is a special case of this formulation, in which there are no costs to the diversion of funds, i.e. $h(b) = b$ for all $b \geq 0$. In this case, the agency problem can be solved only by setting $\Delta = 1$, i.e. experts can finance themselves only through risk-free debt. Our analysis can be generalized easily, but for expositional purposes we keep our baseline model as simple as possible.

We would like to be clear about our assumptions regarding the space of acceptable contracts, which specify how observable cash flows are divided between the expert and household investors. We make the following two restrictions on the contracting space

A The allocation of profit is determined by the total value of capital, and shocks to $k_t$ or $q_t$ separately are not contractible

B Lockups are not allowed - at any moment of time any party can break the contractual relationship. The value of assets is divided among the parties the same way independently of who breaks the relationship

Condition B simplifies analysis, as it allows us to focus only on expert net worth, rather than a summary of the expert's individual past performance history. It subsumes a degree of anonymity, so that once the relationship breaks, parties never meet again and the outcome of the relationship that just ended affects future relationships only through net worth. This condition prevents commitment to long-term contracts, such as in the setting of Myerson (2010). However, in many settings this restriction alone does not rule out optimal contracts: Fudenberg, Holmström, and Milgrom (1990) show that it is possible to implement the optimal long-term contract through short-term contracts with continuous marking-to-market.

Condition A requires that contracts have to be written on the total return of capital, and that innovations in $k_t$, $q_t$ or the aggregate risk $dZ_t$ cannot be hedged separately. This assumption is clearly restrictive, but it creates a convenient and simple way to capture important phenomena that we observe in practice. Specifically, condition A creates an amplification channel, in which market prices affect the agents' net worth, and is consistent with the models of Kiyotaki and Moore (1997) and Bernanke, Gertler, and Gilchrist (1999). Informally, contracting directly on $k_t$ is difficult because we view $k_t$ not as something objective and static like the number of machines, but rather something much more forward looking, like the expected NPV of assets under a particular management strategy. Moreover, even though in our model there is a one-to-one correspondence

[31] One natural way to interpret this is through a capital structure that involves risky debt, as it becomes riskier (more equity-like) after experts suffer losses. Many agency problems become worse when experts are "under water."

between $k_t$ and output, in a more general model this relationship could differ across projects, depend on the expert's private information, and be manipulable, e.g. through underinvestment.

More generally, we could assume that aggregate shocks $dZ_t$ to the experts' balance sheets can be hedged partially. As long as it is impossible to design a perfect hedge, and to perfectly share all aggregate risks with households, the model will generate economic fluctuations driven by the shocks to the net worth of the constrained agents. Thus, to generate economic fluctuations, we make assumptions that would otherwise allow agents to write optimal contracts, but place restrictions on hedging. Experts can still choose their risk exposure $\Delta_t$ but cannot hedge aggregate shocks $dZ_t$.

**Intermediary Sector.** It is possible to reinterpret our model to discuss the capitalization of intermediaries as well as end borrowers.

One natural model of intermediation involves a double moral hazard problem motivated by Holmström and Tirole (1997). Let us separate experts into two classes of agents: entrepreneurs who manage capital under the productive technology, and intermediaries who can channel funds from households to entrepreneurs. Intermediaries are able to, through costly monitoring actions that are unobservable by outside investors, reduce the benefits that entrepreneurs get from the diversion of funds. Specifically, the entrepreneurs' marginal benefit of fund diversion $\frac{\partial}{\partial b}h(b_t|m_t)$ is continuously decreasing with the proportional cost of monitoring $m_t \geq 0$, i.e. $\frac{\partial^2}{\partial b \partial m}h(b_t|m_t) < 0$. Thus, for a fixed equity stake $\Delta_t$ of the entrepreneur, higher monitoring intensity $m_t$ leads to a lower diversion rate $b_t = b(\Delta_t|m_t)$. Assuming that $\frac{\partial^2}{\partial b^2}h(\cdot|m_t) < 0$, the entrepreneur's optimal diversion rate $b_t$ is uniquely determined by the first-order condition $\frac{\partial}{\partial b}h(b_t|m_t) = \Delta_t$, and is continuously decreasing in $\Delta_t$.

Intermediaries have no incentives to exert costly monitoring effort unless they themselves have a stake in the entrepreneur's project. An intermediary who holds a fraction $\Delta_t^I$ of the entrepreneur's equity optimally chooses the monitoring intensity $m_t$ that solves

$$\min_m \ \Delta_t^I b(\Delta_t|m) + m.$$

The solution to this problem determines how the rates of monitoring $m(\Delta_t, \Delta_t^I)$ and cash flow diversion $b(\Delta_t, \Delta_t^I)$ depend on the allocations of equity to the entrepreneur and the intermediary.

By reducing the entrepreneurs' agency problem through monitoring, intermediaries are able to increase the amount of financing available to entrepreneurs. However, intermediation itself requires risk-taking, as the intermediaries need to absorb the risk in their equity stake $\Delta_t^I$ through

their net worth. Thus, the aggregate net worth of intermediaries becomes related to the amount of financing available to entrepreneurs. Figure 9 depicts the interlinked balance sheets of entrepreneurs, intermediaries and households. Fraction $\Delta_t + \Delta_t^I$ of entrepreneur risk gets absorbed by the entrepreneur and intermediary net worths, while fraction $1 - \Delta_t - \Delta_t^I$ is held by households.

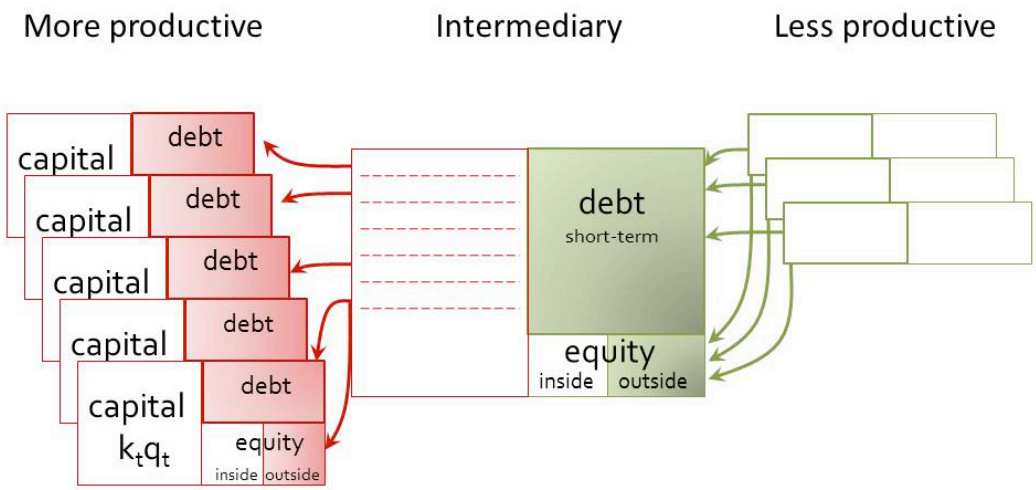


Figure 9: Balance sheets structures of entrepreneurs and financial intermediaries.

The marginal values of entrepreneur and intermediary net worths, $\theta_t$ and $\theta_t^I$, can easily differ in this economy. If so, then the capital-pricing equation (EK) generalizes to

$$\max_{\Delta,\Delta^I} E[dr_t^k]/dt - r - (b(\Delta,\Delta^I) - h(b(\Delta,\Delta^I))) - m(\Delta,\Delta^I) + (\Delta\sigma_t^\theta + \Delta^I\sigma_t^{\theta,I})(\sigma + \sigma_t^q) = 0.$$

Equilibrium dynamics in this economy depends on two state variables, the shares of net worth that belong to the entrepreneurs $\eta_t$ and intermediaries $\eta_t^I$. Generally, these are imperfect substitutes, as intermediaries can reduce the entrepreneurs' required risk exposure by taking on risk and monitoring. However, several special cases can be reduced to a single state variable. For example, if entrepreneurs and intermediaries can write contracts on aggregate shocks among themselves (but not with households), then the two groups of agents have identical risk premia (i.e. $\sigma_t^\theta = \sigma_t^{\theta,I}$) and the sum $\eta_t + \eta_t^I$ determines the equilibrium dynamics.

**Contracting with Idiosyncratic Losses and Costly State Verification.** Next, we discuss contracting in an environment of Section 5, where experts may suffer idiosyncratic loss shocks. For simplicity, we focus on the simplest form of the agency problem without intermediaries, in which $h(b) = b$ for all $b \geq 0$. As we discussed earlier in the Appendix, in our *baseline* model this assumption leads to a simple capital structure, in which experts can borrow only through risk-free

debt.

Assume, as in Section 5, that in the absence of benefit extraction, capital managed by expert $i \in \mathbb{I}$ evolves according to

$$dk_t = (\Phi(\iota_t) - \delta)\, k_t\, dt + \sigma k_t\, dZ_t + k_t\, dJ_t^i,$$

where $dJ_t^i$ is a *compensated* loss process with intensity $\lambda$ and jump distribution $F(y)$, $y \in [-1, 0]$. Then in the absence of jumps, $J_t^i$ has a positive drift of

$$dJ_t^i = \left(\lambda \int_0^1 (-y) dF(y)\right)\, dt,$$

so that $E[dJ_t^i] = 0$.

The entrepreneur can extract benefits continuously or via discrete-jumps. Benefit extraction is described by a non-decreasing process $\{B_t, t \geq 0\}$, which changes the law of motion of capital to

$$dk_t = (\Phi(\iota) - \delta)\, k_t\, dt + \sigma k_t\, dZ_t + k_t\, dJ_t^i - dB_t,$$

and gives entrepreneur benefits at the rate of $dB_t$ units of capital. The jumps in $B_t$ are bounded by $k_{t-}$, the total amount of capital under the entrepreneur's management just before time $t$.

Unlike in our earlier specification of the agency problem, in which the entrepreneur's *rate* of benefit extraction $b_t = \frac{dB_t}{(q_t k_t)\, dt}$ must be finite, now the entrepreneur can also extract benefits discontinuously, including in quantities that reduce the value of capital under management below the value of debt.

We assume that there is a verification technology that can be employed in the event of discrete drops in capital. In particular, if a verification action is triggered by outside investors when capital drops from $k_{t-}$ to $k_t$ at time $t$, then investors

(i) learn whether a drop in capital was caused partially by entrepreneur's benefit extraction at time $t$ and in what amount

(ii) recover all capital that was diverted by the entrepreneur at time $t$

(iii) pay a cost of $(q_t k_{t-}) c(dJ_t^i)$, that is proportional to the value of the investment prior to verification[32]

[32] The assumption that the verification cost depends only on the amount of capital recovered, regardless of the diverted amount, is without loss of generality since on the equilibrium path the entrepreneur does not divert funds.

If verification reveals that the drop in capital at time $t$ was partially caused by benefit extraction, i.e. $k_{t-}(1 + dJ_t^i) > k_t$, then the entrepreneur cannot extract any benefits, as diverted capital $k_{t-}(1 + dJ_t^i) - k_t$ is returned to the investors.

We maintain the same assumptions as before about the form of the contract in the absence of verification, i.e. (A) the contract determines how the total market value of assets is divided between the entrepreneur and outside investors, and (B) at any moment either party can break the relationship and walk away with its share of assets. In particular, contracting on $k_t$ or $q_t$ separately is not possible. In addition, the contract specifies conditions, under which a sudden drop in the market value of the expert's assets $q_t k_t$ triggers a verification action. In this event, the contract specifies how the remaining assets, net of verification costs, are divided among the contracting parties conditional on the amount of capital that was diverted at time $t$. We assume that the monitoring action is not randomized, i.e. it is completely determined by the asset value history.

**Proposition 8** *With idiosyncratic jump risk, it is optimal to trigger verification only in the event that the market value of the expert's assets $q_t k_t$ falls below the value of debt. In the event of verification, it is optimal for debt holders to receive the value of the remaining assets net of verification costs.*

**Proof.** Because jumps are idiosyncratic, they carry no risk premium. Therefore, it is better to deter fund diversion that does not bankrupt the expert by requiring him to absorb jump risk through equity rather than triggering costly state verification (which leads to a deadweight loss). However, verification is required to deter the expert from diverting more funds than his net worth at a single moment of time.

The division of value between debt holders and the expert in the event of verification matters for the expert's incentives only if it is in fact revealed that the expert diverted cash. If no cash was diverted (i.e. it is clear that the loss was caused by an exogenous jump), the division of value between debt holders and the expert can be arbitrary (as long as the expected return of debt holders, net of verification costs, is $r$) since idiosyncratic jump risk carries no risk premium. Without loss of generality we can assume that debt holders receive the entire remaining value in case of verification. ■

Proposition 8 implies that with idiosyncratic jump risk, debt is no longer risk-free.

# B Stationary Distribution

Suppose that $X_t$ is a stochastic process that evolve on the state space $[x_L, x_R]$ according to the equation

$$dX_t = \mu^x(X_t)\,dt + \sigma^x(X_t)\,dZ_t \tag{39}$$

If at time $t = 0$, $X_t$ is distributed according to the density $d(x, 0)$, then the density of $X_t$ at all future dates $t \geq 0$ is described by the forward Kolmogorov equations:

$$\frac{\partial}{\partial t} d\left(x,t\right) = -\frac{\partial}{\partial x}\left(\mu^x\left(x\right) d\left(x,t\right)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^x\left(x\right)^2 d\left(x,t\right)\right).$$

If one of the endpoints is a reflecting barrier, then the boundary condition at that point is

$$-\mu^x(x)d(x,t) + \frac{1}{2}\frac{\partial}{\partial x}(\sigma^x(x)^2 d(x,t)) = 0.$$

A *stationary density* stays fixed over time under the law of motion of the process, so the left-hand side of the Kolmogorov forward equation is $\frac{\partial d(x,t)}{\partial t} = 0$. If one of the endpoints of the interval $[x_L, x_R]$ is reflecting, then integrating with respect to $x$ and using the boundary condition at the reflecting barrier to pin down the integration constant, we find that the stationary density is characterized by the first-order ordinary differential equation

$$-\mu^x(x)d(x) + \frac{1}{2}\frac{\partial}{\partial x}(\sigma^x(x)^2 d(x)) = 0.$$

To compute the stationary density numerically, it is convenient to work with the function $D(x) = \sigma^x(x)^2 d(x)$, which satisfies the ODE

$$D'(x) = 2\frac{\mu^x(x)}{\sigma^x(x)^2}D(x). \tag{40}$$

Then $d(x)$ can be found from $D(x)$ using $d(x) = \frac{D(x)}{\sigma^x(x)^2}$.

With absorbing boundaries, the process eventually ends up absorbed (and so the stationary distribution is degenerate) unless the law of motion prevents (39) it from hitting the boundary with probability one. A non-degenerate stationary density exists with an absorbing boundary at $x_L$ if the boundary condition $D(x_L) = 0$ can be satisfied together with $D(x_0) > 0$ for $x_0 > x_L$. For this to happen, we need

$$\log D(x) = \log D(x_0) - \int_x^{x_0} \frac{2\mu^x(x')}{\sigma^x(x')^2}\,dx' \to -\infty, \text{ as } x \to x_L$$

i.e $\int_{x_L}^{x_0} \frac{2\mu^x(x)}{\sigma^x(x)^2}\, dx = \infty$. This condition is satisfied whenever the drift that pushes $X_t$ away from the boundary $x_L$ (so we need $\mu^x(x) > 0$) is strong enough working against the volatility that may move $X_t$ towards $x_L$. For example, if $X_t$ behaves as a geometric Brownian motion near the boundary $x_L = 0$, i.e. $\mu^x(x) = \mu x$ and $\sigma^x(x) = \sigma x$, with $\mu > 0$, then $\int_0^{x_0} \frac{2\mu^x(x)}{\sigma^x(x)^2}\, dx = \int_0^{x_0} \frac{2\mu}{\sigma^2 x}\, dx = \infty$.

# C Proofs

**Proof of Lemma 1.** Let us show that if the process $\theta_t$ satisfies (11) and the transversality condition holds, then $\theta_t$ represents the expert's continuation payoff, i.e. satisfies (10). Consider the process

$$\Theta_t = \int_0^t e^{-\rho s} n_s\, d\zeta_s + e^{-\rho t}\theta_t n_t.$$

Differentiating $\Theta_t$ with respect to $t$ using Ito's lemma, we find

$$d\Theta_t = e^{-\rho t}(n_t\, d\zeta_t - \rho\theta_t n_t\, dt + d(\theta_t n_t)).$$

If (11) holds, then $E[d\Theta_t] = 0$, so $\Theta_t$ is a martingale under the strategy $\{x_t, d\zeta_t\}$. Therefore,

$$\theta_0 n_0 = \Theta_0 = E[\Theta_t] = E\left[\int_0^t e^{-\rho s} n_s\, d\zeta_s\right] + E\left[e^{-\rho t}\theta_t n_t\right].$$

Taking the limit $t \to \infty$ and using the transversality condition, we find that

$$\theta_0 n_0 = E\left[\int_0^\infty e^{-\rho s} n_s\, d\zeta_s\right],$$

and the same calculation can be done for any other time $t$ instead of 0.

Conversely, if $\theta_t$ satisfies (10), then $\Theta_t$ is a martingale since

$$\Theta_t = E_t\left[\int_0^\infty e^{-\rho s} n_s\, d\zeta_s\right].$$

Therefore, the drift of $\Theta_t$ must be zero, and so (11) holds.

Next, let us show that the strategy $\{x_t, d\zeta_t\}$ is optimal if and only if the Bellman equation (12) holds. Under any alternative strategy $\{\hat{x}_t, d\hat{\zeta}_t\}$, define

$$\hat{\Theta}_t = \int_0^t e^{-\rho s} n_s\, d\hat{\zeta}_s + e^{-\rho t}\theta_t n_t, \quad \text{so that} \quad d\hat{\Theta}_t = e^{-\rho t}(n_t\, d\hat{\zeta}_t - \rho\theta_t n_t\, dt + d(\theta_t n_t)).$$

If the Bellman equation (12) holds, then $\hat{\Theta}_t$ is a supermartingale under an arbitrary alternative

strategy, so

$$\theta_0 n_0 = \hat{\Theta}_0 \geq E[\hat{\Theta}_t] \geq E\left[\int_0^t e^{-\rho s} n_s \, d\hat{\zeta}_s\right].$$

Taking the limit $t \to \infty$, we find that $\theta_0 n_0$ is an upper bound on the expert's payoff from an arbitrary strategy.

Conversely, if the Bellman equation (12) fails then there exists a strategy $\{\hat{x}_t, d\hat{\zeta}_t\}$ such that

$$n_t \, d\hat{\zeta}_t - \rho\theta_t n_t \, dt + E[d(\theta_t n_t)] \geq 0,$$

with a strict inequality on the set of positive measure. Then for large enough $t$,

$$\theta_0 n_0 = \hat{\Theta}_0 < E[\hat{\Theta}_t]$$

and so the expert's expected payoff from following the strategy $\{\hat{x}_t, d\hat{\zeta}_t\}$ until time $t$, and $\{x_t, d\zeta_t\}$ thereafter exceeds that from following $\{x_t, d\zeta_t\}$ throughout. ■

**Proof of Proposition 1.** Using the laws of motion of $\theta_t$ and $n_t$ as well as Ito's lemma, we can transform the Bellman equation (12) to

$$\rho\theta_t n_t \, dt = \max_{\hat{x}_t \geq 0, d\hat{\zeta}_t \geq 0} (1-\theta_t) n_t \, d\hat{\zeta}_t + r\theta_t n_t \, dt + n_t \, E_t[d\theta_t] + \hat{x}_t \theta_t n_t \left(E_t[dr_t^k] - r \, dt + \sigma_t^\theta(\sigma + \sigma_t^q) \, dt\right).$$

Assume that $n_t\theta_t$ represents the expert's maximal expected future payoff, so that by Lemma 1 the Bellman equation holds, and let us justify (i) through (iii). The Bellman equation cannot hold unless $1 \leq \theta_t$ and $E_t[dr_t^k]/dt - r + \sigma_t^\theta(\sigma + \sigma_t^q) \leq 0$, since otherwise the right hand side of the Bellman equation can be made arbitrarily large. If so, then the choices $d\hat{\zeta}_t = 0$ and $\hat{x}_t = 0$ maximize the right hand side, which becomes equal to $r\theta_t n_t \, dt + \theta_t n_t \mu_t^\theta \, dt$. Thus,

$$\rho\theta_t n_t \, dt = r\theta_t n_t \, dt + \theta_t n_t \mu_t^\theta \, dt \quad \Rightarrow \quad \text{(E)}$$

Furthermore, any $d\hat{\zeta}_t > 0$ maximizes the right hand side only if $\theta_t = 1$, and $\hat{x}_t > 0$ does only if $E_t[dr_t^k]/dt - r + \sigma_t^\theta(\sigma + \sigma_t^q) \leq 0$. This proves (i) through (iii). Finally, Lemma 1 implies that the transversality condition must hold for any strategy that attains value $n_t\theta_t$, proving (iv).

Conversely, it is easy to show that if (i) through (iii) hold then the Bellman equation also holds and the strategy $\{x_t, d\zeta_t\}$ satisfies (11). Thus, by Lemma 1, the strategy $\{x_t, d\zeta_t\}$ is optimal and attains value $\theta_t n_t$. ■

**Proof of Lemma 2.** Aggregating over all experts, the law of motion of $N_t$ is

$$dN_t = rN_t\,dt + \psi_t q_t K_t(dr_t^k - r\,dt) - dC_t. \tag{41}$$

where $C_t$ is are aggregate payouts. Furthermore, note that $d(q_tK_t)/(q_tK_t)$ are the capital gains earned by a world portfolio of capital, with weight $\psi_t$ on expert capital and $1-\psi_t$ on household capital. Thus, from (5) and (6),

$$\frac{d(q_tK_t)}{q_tK_t} = \underbrace{dr_t^k - \frac{a-\iota(q_t)}{q_t}\,dt}_{\text{expert capital gains}} - \underbrace{(1-\psi_t)(\underline{\delta}-\delta)\,dt}_{\text{adjustment for household held capital}},$$

since household capital gains are less than those of experts by $\underline{\delta}-\delta$. Using Ito's lemma,

$$\frac{d(1/(q_tK_t))}{1/(q_tK_t)} = -dr_t^k + \frac{a-\iota(q_t)}{q_t}\,dt + (1-\psi_t)(\underline{\delta}-\delta)\,dt + (\sigma+\sigma_t^q)^2\,dt.$$

Combining this equation with (41) and using Ito's lemma, we get

$$d\eta_t = (dN_t)\frac{1}{q_tK_t} + N_t\,d\left(\frac{1}{q_tK_t}\right) + \psi_t q_t K_t(\sigma+\sigma_t^q)\frac{-1}{q_tK_t}(\sigma+\sigma_t^q)\,dt =$$

$$(\psi_t-\eta_t)(dr_t^k - r\,dt - (\sigma+\sigma_t^q)^2\,dt) + \eta_t\left(\frac{a-\iota(q_t)}{q_t} + (1-\psi_t)(\underline{\delta}-\delta)\right)dt - \eta_t d\zeta_t,$$

where $d\zeta_t = dC_t/N_t$. If $\psi_t > 0$, then Proposition 1 implies that $E[dr_t^k] - r\,dt = -\sigma_t^\theta(\sigma+\sigma_t^q)\,dt$, and the law of motion of $\eta_t$ can be written as in (15). ■

**Proof of Proposition 2.** First, we derive expressions for the volatilities of $\eta_t$, $q_t$ and $\theta_t$. Using (15), the law of motion of $\eta_t$, and Ito's lemma, the volatility of $q_t$ is given by

$$\sigma_t^q q(\eta) = q'(\eta)\underbrace{(\psi-\eta)(\sigma+\sigma_t^q)}_{\sigma_t^\eta\eta} \quad\Rightarrow\quad \sigma_t^q = \frac{q'(\eta)}{q(\eta)}\underbrace{\frac{(\psi-\eta)\sigma}{1-\frac{q'(\eta)}{q(\eta)}(\psi-\eta)}}_{\sigma_t^\eta\eta}$$

The expressions for $\sigma_t^\eta$ and $\sigma_t^\theta$ follow immediately from Ito's lemma.

Second, note that from (EK) and (H), it follows that

$$\frac{a-\underline{a}}{q(\eta)} + \underline{\delta} - \delta + (\sigma+\sigma_t^q)\sigma_t^\theta \geq 0, \tag{42}$$

with equality if $\psi < 1$. Moreover, when $q(\eta), q'(\eta), \theta(\eta) > 0$ and $\theta'(\eta) < 0$, then $\sigma_t^q, \sigma_t^\eta > 0$ are

increasing in $\psi$ while $\sigma_t^\theta < 0$ is decreasing in $\psi$. Thus, left hand side of (42) is decreasing from $\frac{a-\underline{a}}{q(\eta)} + \underline{\delta} - \delta$ at $\psi = \eta$ to $-\infty$ at $\psi = \eta + \frac{q(\eta)}{q'(\eta)}$, justifying our procedure for determining $\psi$.

We get $\mu_t^\eta$ from (15), $\mu_t^q$ from (EK) and $\mu_t^\theta$ from (E). The expressions for $q''(\eta)$ and $\theta''(\eta)$ then follow directly from Ito's lemma and (15), the law of motion of $\eta_t$.

Finally, let us justify the five boundary conditions. First, because in the event that $\eta_t$ drops to 0 experts are pushed to the solvency constraint and must liquidate any capital holdings to households, we have $q(0) = \underline{q}$. The households are willing to pay this price for capital if they have to hold it forever. Second, because $\eta^*$ is defined as the point where experts consume, expert optimization implies that $\theta(\eta^*) = 1$ (see Proposition 1). Third and fourth, $q'(\eta^*) = 0$ and $\theta'(\eta^*) = 0$ are the standard boundary conditions at a reflecting boundary. If one of these conditions were violated, e.g. if $q'(\eta^*) < 0$, then any expert holding capital when $\eta_t = \eta^*$ would suffer losses at an infinite expected rate.[33] Likewise, if $\theta'(\eta^*) < 0$, then the drift of $\theta(\eta_t)$ would be infinite at the moment when $\eta_t = \eta^*$, contradicting Proposition 1. Fifth, if $\eta_t$ ever reaches 0, it becomes absorbed there. If any expert had an infinitesimal amount of capital at that point, he would face a permanent price of capital of $\underline{q}$. At this price, he is able to generate the return on capital of

$$\frac{a - \iota(\underline{q})}{\underline{q}} + \Phi(\iota(\underline{q})) - \delta > r$$

without leverage, and arbitrarily high return with leverage. In particular, with high enough leverage this expert can generate a return that exceeds his rate of time preference $\rho$, and since he is risk-neutral, he can attain infinite utility. It follows that $\theta(0) = \infty$.

Note that we have five boundary conditions required to solve a system of two second-order ordinary differential equations with an unknown boundary $\eta^*$. ■

**Proof of Proposition 3.** Since $q'(\eta^*) = \theta'(\eta^*) = 0$, the drift and volatility of $\eta$ at $\eta^*$ are given by

$$\mu_t^\eta(\eta^*)\eta^* = (1-\eta^*)\sigma^2 + \frac{a - \iota(q(\eta^*))}{q(\eta^*)}\eta^* > 0 \quad \text{and} \quad \sigma_t^\eta(\eta^*)\eta^* = (1-\eta^*)\sigma.$$

Hence, $D'(\eta^*) = 2\mu_t^\eta(\eta^*)\eta^*/(\sigma_t^\eta(\eta^*)\eta^*)^2 D(\eta^*) > 0$, where $D(\eta) = d(\eta)(\sigma_t^\eta(\eta)\eta)^2$. Furthermore, because in the neighborhood of $\eta^*$,

$$\sigma_t^\eta(\eta)\eta = \frac{(1-\eta)\sigma}{1 - (1-\eta)q'(\eta)/q(\eta)}.$$

[33] To see intuition behind this result, if $\eta_t = \eta^*$ then $\eta_{t+\epsilon}$ is approximately distributed as $\eta^* - \bar{\omega}$, where $\bar{\omega}$ is the absolute value of a normal random variable with mean 0 and variance $(\sigma_t^\eta)^2\,\epsilon$ As a result, $\eta_{t+\epsilon} \sim \eta^* - \sigma_t^\eta\sqrt{\epsilon}$, so $q(\eta^*) - q'(\eta^*)\sigma_t^\eta\sqrt{\epsilon}$. Thus, the loss per unit of time $\epsilon$ is $q'(\eta^*)\sigma_t^\eta\sqrt{\epsilon}$, and the average rate of loss is $q'(\eta^*)\sigma_t^\eta/\sqrt{\epsilon} \to \infty$ as $\epsilon \to 0$.

is *decreasing* in $\eta$, it follows that the density $d(\eta)$ must be *increasing* in $\eta$.

The dynamics near $\eta = 0$ is more difficult to characterize because of the singularity there. We will do that by conjecturing, and they verifying, that asymptotically as $\eta \to 0$,

$$\mu_t^\eta = \hat{\mu} + o(1) \quad \text{and} \quad \sigma_t^\eta = \hat{\sigma} + o(1),$$

i.e. $\eta_t$ evolves as a geometric Brownian motion, and that

$$\psi(\eta) = C_\psi \eta + o(\eta), \quad q(\eta) = \underline{q} + C_q \eta^\alpha + o(\eta^\alpha) \quad \text{and} \quad \theta(\eta) = C_\theta \eta^{-\beta} + o(\eta^{-\beta})$$

for some constants $C_\psi > 1$, $C_q$, $C_\theta > 0$, $\alpha, \beta \in (0,1)$. We need to verify that the equilibrium equations hold, up to terms of smaller order. Using the equations of Proposition 2, we have

$$\sigma_t^\eta = \frac{(C_\psi - 1)\sigma + o(1)}{1 - O(\eta^\alpha)} \;\Rightarrow\; \hat{\sigma} = (C_\psi - 1)\sigma,$$

$$\sigma_t^q = \frac{\alpha C_q \eta^\alpha}{\underline{q}} \hat{\sigma} + o(\eta^\alpha) = o(1), \quad \sigma_t^\theta = -\beta \hat{\sigma} + o(1),$$

$$(17) \quad \Rightarrow \quad \beta \hat{\sigma} \sigma = \kappa \quad \Rightarrow \quad \hat{\sigma} = (C_\psi - 1)\sigma = \frac{\kappa}{\beta \sigma} \qquad \text{and} \tag{43}$$

$$\hat{\mu} = -\hat{\sigma}(\sigma - \beta \hat{\sigma}) + \frac{a - \iota(\underline{q})}{\underline{q}} + \underline{\delta} - \delta = -\frac{\kappa}{\beta \sigma}\left(\sigma - \frac{\kappa}{\sigma}\right) + \frac{\underline{a} - \iota(\underline{q})}{\underline{q}} + \kappa.$$

We can determine $\mu_t^q$ from the household valuation equation

$$\frac{\underline{a} - \iota(q_t)}{q_t} + \Phi(\iota(q_t)) - \underline{\delta} + \mu_t^q + \sigma \sigma_t^q = r$$

instead of that of experts, because we already took into account (17). By the envelope theorem,

$$\frac{\underline{a} - \iota(q(\eta))}{q(\eta)} + \Phi(q(\eta)) - \underline{\delta} = \underbrace{\frac{\underline{a} - \iota(\underline{q})}{\underline{q}} + \Phi(\underline{q}) - \underline{\delta}}_{r} - \frac{\underline{a} - \iota(\underline{q})}{\underline{q}^2}(C_q \eta^\alpha + o(\eta^\alpha)) + o(\eta^\alpha).$$

Therefore,

$$\mu_t^q = \frac{\underline{a} - \iota(\underline{q})}{\underline{q}^2} C_q \eta^\alpha - \frac{\alpha C_q \eta^\alpha}{\underline{q}} \hat{\sigma} \sigma + o(\eta^\alpha).$$

Our conjecture is valid if equations

$$\mu^q q(\eta) = q'(\eta) \mu_t^\eta \eta + \frac{1}{2} q''(\eta)(\sigma_t^\eta \eta)^2 \quad \text{and} \quad \mu^\theta \theta(\eta) = \theta'(\eta) \mu_t^\eta \eta + \frac{1}{2} \theta''(\eta)(\sigma_t^\eta \eta)^2$$

hold up to higher-order terms of $o(\eta^\alpha)$ and $o(\eta^{-\beta})$, respectively. Ignoring those terms, we need

$$\frac{\underline{a}-\iota(\underline{q})}{\underline{q}}C_q\eta^\alpha - \alpha C_q\eta^\alpha\hat{\sigma}\sigma = \alpha C_q\eta^\alpha\hat{\mu} + \frac{1}{2}\alpha(\alpha-1)C_q\eta^\alpha\hat{\sigma}^2 \quad \text{and}$$

$$(\rho-r)C_\theta\eta^{-\beta} = -\beta C_\theta\eta^{-\beta}\hat{\mu} + \frac{1}{2}\beta(\beta+1)C_\theta\eta^{-\beta}\hat{\sigma}^2. \tag{44}$$

These equations lead to

$$\frac{\underline{a}-\iota(\underline{q})}{\underline{q}} - \alpha\frac{\kappa}{\beta} = \alpha\left(-\frac{\kappa}{\beta\sigma}\left(\sigma-\frac{\kappa}{\sigma}\right) + \frac{\underline{a}-\iota(\underline{q})}{\underline{q}} + \kappa\right) + \frac{1}{2}\alpha(\alpha-1)\frac{\kappa^2}{\beta^2\sigma^2} \quad \text{and}$$

$$\rho - r = -\beta\left(-\frac{\kappa}{\beta\sigma}\left(\sigma-\frac{\kappa}{\sigma}\right) + \frac{\underline{a}-\iota(\underline{q})}{\underline{q}} + \kappa\right) + \frac{1}{2}\beta(\beta+1)\frac{\kappa^2}{\beta^2\sigma^2} \quad \Rightarrow$$

$$\alpha\left(\frac{\kappa^2}{\beta\sigma^2}+\kappa\right) + \frac{\underline{a}-\iota(\underline{q})}{\underline{q}}(\alpha-1) + \frac{1}{2}\alpha(\alpha-1)\frac{\kappa^2}{\beta^2\sigma^2} = 0 \quad \text{and} \tag{45}$$

$$\rho - r = \kappa - \beta\left(\frac{\underline{a}-\iota(\underline{q})}{\underline{q}} + \kappa\right) + \frac{(1-\beta)\kappa^2}{2\beta\sigma^2} \tag{46}$$

We can solve for $\beta$, $C_\psi$ and $\alpha$ in the following order. First, equation (46) has a solution $\beta \in (0,1)$. To see this, note that as $\beta \to 0$ from above, the right hand side of (46) converges to infinity. For $\beta = 1$, the right hand side becomes

$$-\frac{\underline{a}-\iota(\underline{q})}{\underline{q}} < 0.$$

We have $\underline{a}-\iota(\underline{q}) > 0$, since the net rate of output that households receive at $\eta = 0$ must be positive. Second, equation (43) determines the value of $C_\psi > 1$ for any $\beta > 0$. Lastly, equation (45) has a solution $\alpha \in (0,1)$. To see this, note that the left hand side is negative when $\alpha = 0$ and positive when $\alpha = 1$.

This confirms our conjecture about the asymptotic form of the equilibria near $\eta = 0$. Arbitrary values of constants $C_q$ and $C_\theta$ are consistent with these asymptotic dynamics. The value of $C_q$ has to be chosen to ensure that functions $q(\eta)$ and $\theta(\eta)$ reach slope 0 at the same point $\eta^*$, and the $C_\theta$, to ensure that $\theta(\eta^*) = 1$.

We are now ready to characterize the asymptotic form of the stationary distribution bear $\eta = 0$. We have $D'(\eta) = 2\hat{\mu}/\hat{\sigma}^2\, D(\eta)/\eta$, so

$$D(\eta) = C_D\,\eta^{2\hat{\mu}/\hat{\sigma}^2} \quad \text{and} \quad d(\eta) = D(\eta)/(\hat{\sigma}\eta)^2 = C_d\,\eta^{2\hat{\mu}/\hat{\sigma}^2-2}. \tag{47}$$

Equation (44) implies that

$$\frac{2(\rho - r)}{\hat{\sigma}^2} = -\beta \frac{2\hat{\mu}}{\hat{\sigma}^2} + \beta(\beta + 1) \quad \Rightarrow \quad \frac{2\hat{\mu}}{\hat{\sigma}^2} - 2 = \beta - 1 - \frac{2(\rho - r)}{\kappa^2}\sigma^2\beta.$$

We see that $2\hat{\mu}/(\hat{\sigma}^2) - 2 < 0$, and so $d(\eta) = C_d\, \eta^{2\hat{\mu}/\hat{\sigma}^2 - 2} \to \infty$ as $\eta \to 0$. Furthermore, if

$$\frac{2\hat{\mu}}{\hat{\sigma}^2} - 2 > -1 \quad \Leftrightarrow \quad 1 - \frac{2(\rho - r)}{\kappa^2}\sigma^2 > 0 \quad \Leftrightarrow \quad 2(\rho - r)\sigma^2 < \kappa^2,$$

then the stationary density exists and has a hump near $\eta = 0$. Otherwise if $2(\rho - r)\sigma^2 \geq \kappa^2$, then the integral of $d(\eta)$ is infinity, implying that the stationary density does not exist and in the long run $\eta_t$ ends up in an arbitrarily small neighborhood of 0 with probability close to 1. ■

**Lemma 3** *Under the logarithmic utility model, the stationary density exists if*

$$2\sigma^2(\kappa + r - \rho) + \kappa^2 > 0$$

*and has a hump at 0 if also* $\rho > r + \kappa$, *where* $\kappa = (a - \underline{a})/q(0) + \underline{\delta} - \delta$.

**Proof.** Note that asymptotically $\sigma_t^q \to 0$ as $\eta \to 0$. Thus, from equation (25),

$$\psi_t = \eta_t \frac{\kappa}{\sigma^2} + o(\eta_t).$$

Therefore, equation (27) implies that

$$\sigma_t^\eta = \frac{\kappa}{\sigma} + o(1) \quad \text{and} \quad \mu_t^\eta = (\sigma_t^\eta)^2 + \kappa + r - \rho + o(1).$$

The Kolmogorov forward equation (see (47)) implies that asymptotically the stationary density of $\eta_t$ takes the form

$$d(\eta) = C_d \eta^{\beta_d}, \quad \text{where} \quad \beta_d = 2\left(\frac{\mu_t^\eta}{(\sigma_t^\eta)^2} - 1\right) = 2\sigma^2 \frac{\kappa + r - \rho}{\kappa^2}.$$

Thus, unlike in the risk-neutral case, the stationary density is nonsingular if $2\sigma^2(\kappa + r - \rho) + \kappa^2 > 0$ and has a hump at 0 if $\rho > r + \kappa$. ■

**Proof of Proposition 4.** From the proof of Lemma 3,

$$\psi_t = \eta_t \frac{\kappa}{\sigma^2} + o(\eta_t)$$

under logarithmic utility. Under risk neutrality,

$$\psi_t = C_\psi \eta_t + o(\eta_t), \quad \text{where} \quad C_\psi = \frac{\kappa}{\beta\sigma^2} + 1.$$

The variable $\beta$ is determined by equation (46), which implies that $\beta = 1 + O(\sigma^2)$ when $\sigma$ is small. Thus,

$$\psi_t = \eta_t \left( \frac{\kappa}{\sigma^2} + O(1) \right) + o(\eta_t).$$

In both cases,

$$\sigma_t^\eta = \frac{\psi_t - \eta_t}{\eta_t} (\sigma + \sigma_t^q),$$

and $\sigma_t^q \to 0$ as $\eta \to 0$. Thus,

$$\sigma_t^\eta = \frac{\kappa}{\sigma} + O(\sigma)$$

as $\eta \to 0$. ■