

Uncertainty Shocks and Balance Sheet Recessions

Sebastian Di Tella

Stanford University

This paper investigates the origin and propagation of balance sheet recessions. I first show that in standard models driven by TFP shocks, the balance sheet channel disappears when agents can write contracts on the aggregate state of the economy. Optimal contracts sever the link between leverage and aggregate risk sharing, eliminating the concentration of aggregate risk that drives balance sheet recessions. I then show that uncertainty shocks can help explain this concentration of aggregate risk and drive balance sheet recessions, even with contracts on aggregate shocks. The mechanism is quantitatively important, and I explore implications for financial regulation.

I. Introduction

The recent financial crisis has underscored the importance of the financial system in the transmission and amplification of aggregate shocks. During normal times, the financial system helps allocate resources to their most productive use and provides liquidity and risk-sharing services to the economy. During downturns, however, the concentration of aggregate risk on the balance sheets of leveraged agents can lead to balance sheet recessions. Aggregate shocks will be amplified when these agents

I am grateful to my advisors Daron Acemoglu, Guido Lorenzoni, and Ivan Werning for their invaluable guidance. I also thank George-Marios Angeletos, Vladimir Asriyan, Ricardo Caballero, Marco Di Maggio, David Gamarnik, Veronica Guerrieri, Leonid Kogan, Andrey Malenko, William Mullins, Juan Passadore, Mercedes Politi, Felipe Severino, Rob Townsend, Victoria Vanasco, Xiao Yu Wang, Juan Pablo Xandri, Luis Zermelo, seminar participants at the Massachusetts Institute of Technology Macroeconomics Lunch, and two anonymous referees. Data are provided as supplementary material online.

Electronically published November 2, 2017

[*Journal of Political Economy*, 2017, vol. 125, no. 6]

© 2017 by The University of Chicago. All rights reserved. 0022-3808/2017/12506-0004\$10.00

lose net worth and become less willing or able to hold assets, further depressing asset prices and growth. While we have a good understanding of why balance sheets matter in an economy with financial frictions, we do not have a good explanation for why aggregate risk is so concentrated in the first place. In this paper I show that uncertainty shocks can help explain this concentration of aggregate risk and drive balance sheet recessions.

In order to understand the concentration of aggregate risk, I derive financial frictions from a moral hazard problem and allow agents to write contracts on all observable variables. I find that the type of aggregate shock hitting the economy takes on a prominent role. The first contribution of this paper is to show that in standard models of balance sheet recessions driven by Brownian total factor productivity (TFP) shocks, the balance sheet channel completely disappears when agents are allowed to write complete contracts.¹ Optimal contracts break the link between leverage and aggregate risk sharing and eliminate the concentration of aggregate risk that drives balance sheet recessions. As a result, balance sheets play no role in the transmission and amplification of aggregate shocks. Furthermore, these contracts are simple to implement using standard financial instruments such as equity and a market index. In fact, the balance sheet channel vanishes as long as agents can trade a simple market index.

The second contribution is to show that, in contrast to Brownian TFP shocks, uncertainty shocks can create balance sheet recessions. I introduce an aggregate uncertainty shock that increases idiosyncratic risk in the economy. Because of the moral hazard problem, an increase in idiosyncratic risk depresses investment and asset prices. This induces more productive (leveraged) agents to take on aggregate risk *ex ante* in order to hedge endogenously stochastic investment opportunities. As a result, weak balance sheets amplify the effects of the uncertainty shock, further depressing investment and asset prices in a two-way feedback loop. In addition, an increase in idiosyncratic risk leads to an endogenous increase in aggregate risk, low interest rates, and high risk premia.

I use a continuous-time growth model similar to the He and Krishnamurthy (2012) and Brunnermeier and Sannikov (2014) models of financial crises. There are two types of agents: experts who can trade and use capital to produce and households that finance them. Capital is exposed to both aggregate and (expert-specific) idiosyncratic Brownian TFP shocks. Experts want to raise funds from households and share risk with them, but

¹ In standard models of balance sheet recessions such as Bernanke and Gertler (1989) and Kiyotaki and Moore (1997) or, more recently, Gertler, Kiyotaki, and Queralto (2012), He and Krishnamurthy (2012), or Brunnermeier and Sannikov (2014), agents face *ad hoc* constraints on their ability to share aggregate risk.

they face a moral hazard problem that imposes a “skin in the game” constraint: experts must keep a fraction of their equity to deter them from diverting funds to a private account. This limits their ability to share idiosyncratic risk and makes leverage costly. The more capital an expert buys, the more idiosyncratic risk he must carry on his balance sheet. Experts will therefore require a higher excess return on capital when idiosyncratic risk is high and their balance sheets are weak.

When contracts cannot be written on the aggregate state of the economy, experts are mechanically exposed to aggregate risk through the capital they hold, and any aggregate shock that depresses the value of assets will have a large impact on their leveraged balance sheets. In contrast, when contracts can be written on the aggregate state of the economy, the decision of how much capital to buy (leverage) is separated from aggregate risk sharing, and optimal contracts hedge the (endogenously) stochastic investment opportunities provided by the market. In equilibrium, aggregate risk sharing is governed by the hedging motive of experts relative to households. Brownian TFP shocks do not affect the relative investment opportunities of experts and households, so they share this aggregate risk proportionally to their wealth. In equilibrium, TFP shocks have a direct impact only on output but are not amplified through balance sheets and do not affect the price of capital, investment, or the financial market.

In contrast to Brownian TFP shocks, uncertainty shocks create an endogenous hedging motive that induces experts to choose a large exposure to aggregate risk. The intuition is as follows. Downturns are periods of high idiosyncratic risk, with depressed asset prices and high risk premia. Experts who invest in these assets and receive the risk premia have relatively better investment opportunities during downturns and get more utility per dollar relative to households. On the one hand, this creates a *substitution effect*: if agents are close to risk neutral, experts will prefer to have more net worth during downturns in order to get more “bang for the buck.” This effect works against the balance sheet channel, since it induces experts to insure against aggregate risk. On the other hand, experts require more net worth during booms in order to achieve any given level of utility. This creates an *income effect*: risk-averse experts will prefer to have relatively more net worth during booms. The income effect dominates in the empirically relevant case with relative risk aversion greater than one. As a result, after an uncertainty shock financial losses are concentrated on experts’ balance sheets, further depressing asset prices and raising risk premia, and inducing experts to take even more aggregate risk ex ante in a two-way feedback loop.

To evaluate the size of these mechanisms I calibrate the model to US data. I find that uncertainty shocks can generate significant fluctuations in investment and asset prices, with financial losses heavily concentrated on the balance sheets of experts. Empirically, idiosyncratic risk rises

sharply during downturns, as Bloom et al. (2012) or Christiano, Motto, and Rostagno (2014) show.² More generally, the results in this paper suggest that the type of aggregate shock hitting the economy can play an important role in explaining the concentration of aggregate risk that drives balance sheet recessions. When the income effect dominates, experts will choose to face large financial losses after an aggregate shock that improves their investment opportunities relative to households. The same tools presented here can be used to study the effects of other aggregate shocks. The continuous-time setup allows me to characterize the equilibrium as the solution to a system of partial differential equations and analyze the full equilibrium dynamics instead of linearizing around a steady state. It also makes results comparable to the asset pricing literature.

The contracting setup is related to the literature on dynamic contracts in continuous time, such as Sannikov (2008) and especially DeMarzo and Sannikov (2006) (or DeMarzo and Fishman [2007] in discrete time). Here I consider short-term contracts to make results comparable to those of He and Krishnamurthy (2012) and Brunnermeier and Sannikov (2014).³ A possible concern with an optimal contracts approach is that they might require very complex and unrealistic financial arrangements. I show that optimal contracts can be implemented in a complete financial market with a simple equity constraint. In fact, the neutrality result for Brownian TFP shocks does not require the financial market to be complete. It is enough that it spans the aggregate return to capital, and a market index of experts' equity accomplishes precisely this.

Understanding why aggregate risk is concentrated on some agents' balance sheets is important for the design of financial regulation. If markets are incomplete and agents are not able to share aggregate risk, it is optimal to facilitate this risk sharing and eliminate the balance sheet channel. This is the case in the setting in Brunnermeier and Sannikov (2014), for example. In contrast, if contracts are complete and experts are choosing to be highly exposed to aggregate risk, this is no longer true. I show that the competitive equilibrium is not constrained efficient because of the presence of an externality. However, the policy that aims to eliminate the concentration of aggregate risk is not optimal either. I solve a Ramsey problem focusing on a class of simple policy interventions and show that a social planner would like to concentrate aggregate risk on households

² For example, Bloom et al. (2012) report that during the financial crisis in 2008–9, plant-level TFP shocks increased in variance by 76 percent, while output growth dispersion increased by 152 percent. This is also reflected in the idiosyncratic volatility of stock returns (see Campbell et al. 2001). An increase in idiosyncratic risk could also reflect greater disagreement over the value of assets (Simsek 2013) or an increased interest in acquiring information about assets (Gorton and Ordoñez 2014).

³ Di Tella (2014) studies a similar environment with long-term dynamic contracts.

and make experts' balance sheets countercyclical, in order to dampen the effects of uncertainty shocks.

Literature review.—This paper fits within the literature on the balance sheet channel going back to the seminal works of Bernanke and Gertler (1989), Kiyotaki and Moore (1997), and Bernanke, Gertler, and Gilchrist (1999). It is most closely related to the more recent He and Krishnamurthy (2012) and Brunnermeier and Sannikov (2014).⁴ The main difference with these papers is that I allow agents to write contracts on all observable variables, including the aggregate state of the economy.

Krishnamurthy (2003) was the first to explore the concentration of aggregate risk and its role in balance sheet recessions when contracts can be written on the aggregate state of the economy. He finds that when agents are able to trade state-contingent assets, the feedback from asset prices to balance sheets disappears. He then shows that this feedback reappears when limited commitment on households' side is introduced: if households also need collateral to credibly promise to make payments during downturns, they might be constrained in their ability to share aggregate risk with experts. This mechanism also appears in Holmstrom and Tirole (1996). The limited commitment on the households' side is binding, however, only when experts as a whole need fresh cash infusions from households. Typically, debt reductions are enough to provide the necessary aggregate risk sharing and evade households' limited commitment. Rampini and Viswanathan (2010, 2013) also study the concentration of aggregate risk, focusing on the trade-off between financing and risk sharing. They show that firms that are severely collateral constrained might forgo insurance in order to have more funds up-front for investment. Cooley, Quadrini, and Marimon (2004) show how limited contract enforceability can prevent full aggregate risk sharing, while Asriyan (2015) shows how dispersed information can make it costly for agents to share aggregate risk in over-the-counter markets. In contrast to all these papers, I study a setting in which agents are able to leverage and share aggregate risk freely, which highlights their incentives for sharing different types of aggregate shocks. In the same line, Geanakoplos (2009) emphasizes the role of heterogeneous beliefs. More optimistic agents place a higher value on assets and are naturally more exposed to aggregate risk. A similar explanation could be built on heterogeneous preferences for risk. In contrast, the mechanism in this paper does not depend on heterogeneous beliefs or preferences.

Several papers make the empirical case for the importance of balance sheets. Sraer, Chaney, and Thesmar (2012) use local variation in real estate prices to identify the impact of firm collateral on investment. They

⁴ Adrian and Boyarchenko (2012) and Gertler et al. (2012) also study financial crises in settings with incomplete contracts.

find that each extra dollar of collateral increases investment by \$0.06. Gabaix, Krishnamurthy, and Vigneron (2007) provide evidence for balance sheet effects in asset pricing. They show that the marginal investor in mortgage-backed securities is a specialized intermediary instead of a diversified representative agent. Adrian, Etula, and Muir (2014) use shocks to the leverage of securities broker-dealers to construct an “intermediary stochastic discount factor” (SDF) and use it to explain asset returns.

The role of uncertainty shocks in business cycles is explored in Bloom (2009) and, more recently, Bloom et al. (2012). Christiano et al. (2014) introduce shocks to idiosyncratic risk in a model with financial frictions and incomplete contracts, and they report that this shock is the most important factor driving business cycles.⁵ In the asset pricing literature, Campbell et al. (2012) introduce a volatility factor into an intertemporal capital asset pricing model. They find that this volatility factor can help explain the growth-value spread in expected returns. Bansal and Yaron (2004) study aggregate shocks to the growth rate and volatility of the economy. Idiosyncratic risk, in particular, is studied by Campbell et al. (2001). Herskovic et al. (2016) show that idiosyncratic risk is a priced factor in the financial market, consistent with the mechanism here.

Layout.—The rest of the paper is organized as follows. Section II introduces the model. Section III characterizes the equilibrium using a recursive formulation and studies the effects of different types of aggregate shocks. Section IV looks at financial regulation. Section V presents conclusions.

II. The Model

The model purposefully builds on He and Krishnamurthy (2012) and Brunnermeier and Sannikov (2014), adding idiosyncratic risk and general Epstein-Zin preferences to their framework. As in those papers, I derive financial frictions endogenously from a moral hazard problem. In contrast to those papers, however, contracts can be written on all observable variables.

Technology.—Consider an economy populated by two types of agents: “experts” and “households,” identical in every respect except that experts are able to use capital to produce consumption goods.⁶ Denote by k_t the aggregate “efficiency units” of capital in the economy and by $k_{i,t}$ the individual holdings of an expert $i \in [0, 1]$, where $t \in [0, \infty)$ is time. An ex-

⁵ Fernández-Villaverde and Rubio-Ramírez (2010) and Fernández-Villaverde et al. (2011) also study the impact of uncertainty shocks in standard macroeconomic models. On the other hand, Bachmann and Moscarini (2011) argue that causation may run in the opposite direction, with downturns inducing higher risk.

⁶ We could allow households to use capital less productively, as in Kiyotaki and Moore (1997) or Brunnermeier and Sannikov (2014). This does not change the main results.

pert can use capital to produce a flow of consumption goods over a short period of time:

$$y_{i,t} = [a - \iota(g_{i,t})]k_{i,t}.$$

The function ι with $\iota' > 0$, $\iota'' > 0$ represents a standard investment technology with adjustment costs: in order to achieve a growth rate g for his capital stock, the expert must invest a flow of $\iota(g)$ consumption goods. The change in his “effective capital” in a short period of time is

$$\frac{dk_{i,t}}{k_{i,t}} = g_{i,t}dt + \sigma dZ_t + \nu_t dW_{i,t},$$

where $Z = \{Z_t \in \mathbb{R}^d; \mathcal{F}_t, t \geq 0\}$ is an aggregate Brownian motion, and $W_i = \{W_{i,t}; \mathcal{F}_t, t \geq 0\}$ is an idiosyncratic Brownian motion for expert i , in a probability space (Ω, P, \mathcal{F}) equipped with a filtration $\mathbb{F} = \{\mathcal{F}_t; t \geq 0\}$ with the usual conditions. Idiosyncratic shocks W_i represent shocks to the capital held by expert i over a short period, not to the productivity of the expert i . All experts are always equally good at using all capital. The aggregate shock Z can be interpreted as a TFP shock if we let k be “effective” units of capital.⁷

While the exposure of capital to aggregate risk $\sigma \geq 0 \in \mathbb{R}^d$ is constant,⁸ its exposure to idiosyncratic risk $\nu_t > 0$ follows an exogenous stochastic process

$$d\nu_t = \lambda(\bar{\nu} - \nu_t)dt + \sigma_\nu \sqrt{\bar{\nu}_t} dZ_{\nu,t}, \quad (1)$$

where $\bar{\nu}$ is the long-run mean and λ the mean reversion parameter.⁹ The loading of the idiosyncratic volatility of capital on aggregate risk is $\sigma_\nu \leq 0$, so that we may think of Z as a “good” aggregate shock that increases the effective capital stock and reduces idiosyncratic risk. This is just a naming convention. The fact that Z both affects the level of capital as a TFP shock and drives idiosyncratic volatility ν as an uncertainty shock is without loss of generality since it can be multidimensional. We may take some shocks to be pure TFP shocks with $\sigma_\nu^{(i)} = 0$, other pure uncertainty shocks with $\sigma^{(i)} = 0$, and yet other mixed shocks. For most results, however, there is no loss from taking $d = 1$ and focusing on a single aggregate shock.

⁷ If $\kappa_{i,t}$ is physical capital, $k_{i,t} = a_i \kappa_{i,t}$ is “effective capital” in the hands of expert i , so aggregate shocks to $k_{i,t}$ can be interpreted as persistent shocks to TFP a_i , i.e., $da_t = a_t \sigma dZ_t$. To preserve scale invariance we must also have investment costs proportional to a_t .

⁸ I will use the convention that σ is a row vector, while Z_t is a column vector. I will also write σ^2 , e.g., instead of $\sigma\sigma'$ to avoid clutter. Throughout the paper I will not point this out unless it is necessary for clarity.

⁹ If $2\lambda\bar{\nu} \geq \sigma_\nu^2$, this Cox-Ingersoll-Ross process is always strictly positive and has a long-run distribution with mean $\bar{\nu}$. I assume this condition holds.

The law of motion for the aggregate capital stock $k_t = \int_{[0,1]} k_{i,t} di$ is not affected by the idiosyncratic shocks $W_{i,t}$ which wash away in the aggregate:

$$dk_t = \left(\int_{[0,1]} g_{i,t} k_{i,t} di \right) dt + \sigma k_t dZ_t.$$

Preferences.—Both experts and households have Epstein-Zin preferences with the same discount rate ρ , risk aversion γ , and elasticity of intertemporal substitution (EIS) ψ^{-1} . If we let $\gamma = \psi$, we get the standard constant relative risk aversion utility case as a special case. They are defined recursively (see Duffie and Epstein 1992):

$$U_t = \mathbb{E}_t \left[\int_t^\infty f(c_u, U_u) du \right], \tag{2}$$

where

$$f(c, U) = \frac{1}{1 - \psi} \left\{ \frac{\rho c^{1-\psi}}{[(1 - \gamma)U]^{(\gamma-\psi)/(1-\gamma)}} - \rho(1 - \gamma)U \right\}.$$

I will later also introduce turnover among experts in order to obtain a nondegenerate stationary distribution for the economy. Experts will retire with independent Poisson arrival rate τ and become households. There is no loss in intuition from taking $\tau = 0$ for most of the results, however.

Markets.—Experts can trade capital continuously at a competitive price $p > 0$, which we conjecture follows an Ito process:

$$\frac{dp_t}{p_t} = \mu_{p,t} dt + \sigma_{p,t} dZ_t.$$

The total value of the aggregate capital stock is $p k_t$, and it constitutes the total wealth of the economy. There is also a complete financial market with SDF η :

$$\frac{d\eta_t}{\eta_t} = -r_t dt - \pi_t dZ_t.$$

Here r_t is the risk-free interest rate and π_t the price of aggregate risk Z . I am already using the fact that idiosyncratic risks $\{W_i\}_{i \in [0,1]}$ have price zero in equilibrium because they can be aggregated away. Both the price of capital p and the SDF η are determined endogenously in equilibrium and depend only on the history of aggregate shocks Z .

Households' problem.—Households face a standard portfolio problem. They cannot hold capital, but they have access to a complete financial

market. They start with wealth w_0 derived from ownership of a fraction of aggregate capital (which they immediately sell to experts). Taking the aggregate process η as given, they solve the following problem,

$$\begin{aligned} & \max_{(c \geq 0, \sigma_w)} U(c) \\ \text{subject to } & \frac{dw_t}{w_t} = (r_t + \sigma_{w,t}\pi_t - \hat{c}_t)dt + \sigma_{w,t}dZ_t, \end{aligned}$$

and a solvency constraint $w_t \geq 0$, where the hat on \hat{c} denotes that the variable is normalized by wealth. I use w for the wealth of households and reserve n for experts' wealth, which I will call "net worth." Households get the risk-free interest rate on their wealth, plus a premium π_t for the exposure to aggregate risk $\sigma_{w,t}$ they choose to take. Since the price of expert-specific idiosyncratic risks $\{W_i\}$ is zero in equilibrium, they will never hold idiosyncratic risk, so their consumption and wealth depend only on the history of aggregate shocks Z . This is already baked into their budget constraint.

Experts' problem.—Experts face a more complex problem. An expert can continuously trade and use capital for production, as well as participate in the financial market. The cumulative return from investing a dollar in capital for expert i is R_i^k , with

$$dR_{i,t}^k = \underbrace{\left[\frac{a - \iota(g_{i,t})}{p_t} + g_{i,t} + \mu_{p,t} + \sigma\sigma'_{p,t} \right]}_{\mathbb{E}_i[dR_{i,t}^k]} dt + (\sigma + \sigma_{p,t})dZ_t + \nu_i dW_{i,t}.$$

He would like to share risk with the market, but he faces a "skin in the game" constraint. In the online appendix I derive this financial friction from a moral hazard problem, similarly to He and Krishnamurthy (2012) and Brunnermeier and Sannikov (2014). The expert can secretly divert capital to a private account but can keep only a fraction $\phi \in (0, 1)$ of what he steals. I allow experts to write complete short-term contracts on all observable variables, including aggregate shocks. In order to provide incentives to not steal, the expert must keep an exposure ϕ to the return of his capital dR_i^k , so that stealing is not profitable for him. The expert's problem is to choose his consumption and trading strategies to maximize his expected utility

$$\begin{aligned} & \max_{(e \geq 0, g, k \geq 0, \theta)} U(e) \\ \text{subject to } & \frac{dn_{i,t}}{n_{i,t}} = (\mu_{i,n,t} - \hat{e}_{i,t})dt + \sigma_{i,n,t}dZ_t + \tilde{\sigma}_{i,n,t}dW_{i,t}, \quad (3) \end{aligned}$$

where

$$\begin{aligned} \mu_{i,n,t} &= r_t + p_t \hat{k}_{i,t} (\mathbb{E}_t[dR_{i,t}^k] - r_t) - (1 - \phi) p_t \hat{k}_{i,t} (\sigma + \sigma_{p,t}) \pi_t + \theta_{i,t} \pi_t, \\ \sigma_{i,n,t} &= \phi p_t \hat{k}_{i,t} (\sigma + \sigma_{p,t}) + \theta_{i,t}, \\ \tilde{\sigma}_{i,n,t} &= \phi p_t \hat{k}_{i,t} \nu_t, \end{aligned}$$

and a solvency constraint $n_t \geq 0$. As before, the hatted variables denote that they are divided by the net worth $n_{i,t}$. The expert invests $p_t \hat{k}_{i,t}$ in capital and must keep an exposure ϕ to his own return $dR_{i,t}^k$ because of the moral hazard problem. He sells the rest of $1 - \phi$ on the market. The market does not mind the idiosyncratic risk $\nu_t dW_{i,t}$ contained in $dR_{i,t}^k$, but it does demand a price π_t for the aggregate risk $(\sigma + \sigma_{p,t}) dZ_t$ that the expert is off-loading. The skin in the game constraint limits the expert's ability to share the idiosyncratic risk: his exposure $\tilde{\sigma}_{n,t}$ to W_i comes from the fraction ϕ of his return that he keeps. This also exposes him to aggregate risk $\phi p_t \hat{k}_{i,t} (\sigma + \sigma_{p,t})$. Crucially, the moral hazard problem does not limit his ability to share aggregate risk. The term $\theta_{i,t}$ allows him to separate the decision of how much to invest in capital $p_t \hat{k}_{i,t}$ from the decision of how much aggregate risk to hold $\sigma_{n,t}$. This is the main difference with the contractual setup in He and Krishnamurthy (2012) and Brunnermeier and Sannikov (2014), where the additional constraint $\theta_{i,t} = 0$ is imposed: contracts cannot be written on the aggregate state of the economy. In this case investment in capital and exposure to aggregate risk become entangled. The separation between investment in capital (or leverage) and aggregate risk sharing is at the core of the Brownian TFP neutrality result.

The optimal contract is easy to implement. The expert creates a firm with $p_t k_{i,t}$ assets, keeps a fraction ϕ of the equity, and sells the rest and borrows to raise funds (if $n_{i,t} > \phi p_t k_{i,t}$ he does not need to borrow, and he invests $n_{i,t} - \phi p_t k_{i,t}$ outside the firm). In addition, he trades aggregate securities (possibly indices of other firms' equity), and he receives a payment as CEO of the firm, which compensates him for the idiosyncratic risk he takes by keeping a fraction ϕ of his firm's equity. We can think of $\theta_{i,t}$ as the fraction of the expert's wealth invested in a set of aggregate securities that span Z (normalized to have an identity loading on Z). In the special case with only one aggregate shock, $d = 1$, we can think of this security as a normalized market index. More generally, we can consider the intermediate case in which contracts may be written only on a linear combination of aggregate shocks $\tilde{Z}_t = B_t Z_t$ for some full rank matrix $B_t \in \mathbb{R}^{d' \times d}$ with $d' < d$.¹⁰ In this case we will be restricted to choosing $\theta_{i,t} = \hat{\theta}_{i,t} B_t$. In particular, with $B_t = 0$, contracts cannot be written on Z .

¹⁰ In terms of $\theta_{i,t}$ as a set of aggregate securities, this corresponds to an incomplete financial market.

Equilibrium.—Denote the set of experts $\mathbb{I} = [0, 1]$ and the set of households $\mathbb{J} = (1, 2]$. We take the initial capital stock k_0 and its distribution among agents $\{k_i^0\}_{i \in \mathbb{I}}, \{k_j^0\}_{j \in \mathbb{J}}$ as given, with $\int_{\mathbb{I}} k_i^0 di + \int_{\mathbb{J}} k_j^0 dj = k_0$. Let $k_i^0 > 0$ and $k_j^0 > 0$ so that all agents start with strictly positive net worth.

DEFINITION 1. A competitive equilibrium is a set of aggregate stochastic processes: the price of capital p , the state price density η , the aggregate capital stock k , and a set of stochastic processes for each expert $i \in \mathbb{I}$ and each household $j \in \mathbb{J}$: net worth n_i and wealth w_j , consumption e_i and c_j , capital holdings k_i , investment g_i , and aggregate risk sharing $\sigma_{i,n}$ and $\sigma_{j,w}$ such that

- i. initial net worth satisfies $n_{i,0} = p_0 k_i^0$ and wealth $w_{j,0} = p_0 k_j^0$;
- ii. each expert and household solves his problem taking aggregate conditions as given;
- iii. markets clear:

$$\begin{aligned} \int_{\mathbb{I}} e_{i,t} di + \int_{\mathbb{J}} c_{j,t} dj &= \int_{\mathbb{I}} [a - \iota(g_{i,t})] k_{i,t} di, \\ \int_{\mathbb{I}} k_{i,t} di &= k_t, \\ \int_{\mathbb{I}} \sigma_{i,n,t} n_{i,t} di + \int_{\mathbb{J}} \sigma_{j,w,t} w_{j,t} dj &= \int_{\mathbb{I}} p_t k_{i,t} (\sigma + \sigma_{p,t}) di; \end{aligned}$$

- iv. and aggregate capital stock satisfies the law of motion, starting with k_0 :

$$dk_t = \left(\int_{\mathbb{I}} g_{i,t} k_{i,t} di \right) dt + k_t \sigma dZ_t.$$

The market clearing conditions for the consumption goods and capital market are standard. The condition for market clearing in the financial market is derived as follows: we already know each expert keeps a fraction ϕ of his own equity. If we aggregate the equity sold on the market into indices with identity loading on Z , there is a total supply of these indices $(1 - \phi)p_t k_t (\sigma + \sigma_{t,p})$. Households absorb $\int_{\mathbb{J}} \sigma_{j,w,t} w_{j,t} dj$ and experts $\int_{\mathbb{I}} \theta_{i,t} n_{i,t} di$ of these indices. Rearranging, we obtain the expression above. By Walras's law, the market for risk-free debt clears automatically.

III. Solving the Model

Experts and households face a dynamic problem, where their optimal decisions depend on the stochastic investment opportunities they face given by the price of capital p and the SDF η . The equilibrium is driven by the exogenous stochastic process for ν_t and by the endogenous distribution of wealth between experts and households. The recursive Epstein-Zin

preferences generate optimal strategies that are linear in net worth and allow us to simplify the state-space: we need to keep track only of the net worth of experts relative to the total value of assets that they must hold in equilibrium, $x_i = n_i/p_i k_i$. The distribution of net worth across experts, and of wealth across households, is not important. The strategy is to use a recursive formulation of the problem and look for a Markov equilibrium in (v_t, x_t) , taking advantage of the scale invariance property of the economy that allows us to abstract from the level of the capital stock.

The layout of this section is as follows. First I solve a first-best benchmark without moral hazard and show that the economy follows a stable growth path. Then back to the moral hazard case, I recast the equilibrium in recursive form and characterize agents' optimal plans. I study the effect of Brownian TFP shocks under different contractual environments. I then show how uncertainty shocks can create balance sheet recessions as a result of agents' optimal aggregate risk-sharing decisions.

A. Benchmark without Moral Hazard

Without any financial frictions, this is a standard AK growth model in which balance sheets do not play any role. Because there is no moral hazard, experts share all of their idiosyncratic risk, so the dynamics of idiosyncratic shocks v_t are irrelevant. Without financial frictions, the price of capital and the growth rate of the economy do not depend on experts' net worth: balance sheets are relevant only to determine consumption of experts and households. The economy follows a stable growth path.

PROPOSITION 1 (First-best benchmark). If $\rho - (1 - \psi)g^* + (1 - \psi)(\gamma/2)\sigma^2 > 0$ and there are no financial frictions, there is a stable growth equilibrium with price of capital p^* and growth rate g^* given by

$$v'(g^*) = p^*, \tag{4}$$

$$p^* = \frac{a - v(g^*)}{\rho - (1 - \psi)g^* + (1 - \psi)(\gamma/2)\sigma^2}. \tag{5}$$

This is a very clean benchmark. Anything we get away from this balanced growth path can be attributed to the introduction of moral hazard.

B. Back to Moral Hazard

From homothetic preferences we know that the value function for an expert with net worth n takes the following power form:

$$V_i(n) = \frac{(\xi_i n)^{1-\gamma}}{1 - \gamma} \tag{6}$$

for some stochastic process $\xi > 0$ that captures the forward-looking stochastic investment opportunities the expert faces (the price of capital p , the interest rate r , and the price of aggregate risk π). When ξ_t is high, the expert is able to obtain a large amount of utility from a given net worth n_t , as if his actual net worth was $\xi_t n_t$. It depends only on the history of aggregate shocks Z and must be determined in equilibrium. Conjecture that it follows an Ito process

$$\frac{d\xi_t}{\xi_t} = \mu_{\xi,t} dt + \sigma_{\xi,t} dZ_t.$$

For households, the utility function takes the same form, $U_t(n) = (\zeta_t n)^{1-\gamma}/(1-\gamma)$, but instead of ξ_t , we have ζ_t , which follows $d\zeta_t/\zeta_t = \mu_{\zeta,t} dt + \sigma_{\zeta,t} dZ_t$ and captures households' investment opportunities.

The Hamilton-Jacobi-Bellman (HJB) equation associated with experts' problem after some algebra is¹¹

$$\begin{aligned} \frac{\rho}{1-\psi} = \max_{\hat{e}, g, \hat{k}, \theta} & \left\{ \frac{\hat{e}^{1-\psi}}{1-\psi} \rho \xi^{\psi-1} + \mu_n - \hat{e} + \mu_{\xi} \right. \\ & \left. - \frac{\gamma}{2} \left(\sigma_n^2 + \sigma_{\xi}^2 - 2 \frac{1-\gamma}{\gamma} \sigma_n \sigma'_{\xi} + \tilde{\sigma}_n^2 \right) \right\} \end{aligned} \quad (7)$$

subject to the dynamic budget constraint (3) and $\hat{k}, \hat{e} \geq 0$. Households have an analogous HJB equation (but with $\hat{k} = 0$). The first-order condition (FOC) for g ,

$$i'(g) = p, \quad (8)$$

links investment and asset prices: anything that depresses the price of capital will have a real effect through investment and growth. In addition, the combination of homothetic preferences and linear budget constraints implies that policy functions are linear in net worth or wealth: all experts choose the same $\hat{e}_t, g_t, \hat{k}_t$, and θ_t and all households the same \hat{c}_t and $\sigma_{w,t}$. This allows us to abstract from the distribution of wealth. We need to keep track only of the share of aggregate wealth that belongs to experts: $x_t = n_t/p_t k_t \in (0, 1)$. We can therefore look for a Markov equilibrium with two state variables (v_t, x_t) :

$$\begin{aligned} p_t &= p(v_t, x_t), & \xi_t &= \xi(v_t, x_t), & \zeta_t &= \zeta(v_t, x_t), \\ r_t &= r(v_t, x_t), & \pi_t &= \pi(v_t, x_t), \end{aligned}$$

¹¹ We look for a solution ξ to the HJB equation such that $\xi^{1-\gamma}$ is strictly positive and bounded and such that the resulting policy functions $\hat{e}, g, \hat{k}, \theta$ generate a plan that delivers the utility indicated by the value function; likewise for ζ and households' HJB. See App. B for details.

where p , ξ , and ζ are strictly positive and twice continuously differentiable. While idiosyncratic risk ν_t evolves exogenously according to (1), experts' share of aggregate wealth x_t is endogenous, with law of motion $dx_t = \mu_x(\nu, x)dt + \sigma_x(\nu, x)dZ_t$, where

$$\begin{aligned} \mu_x(\nu, x) &= x[\mu_n - \hat{e} - g - \mu_p - \sigma\sigma'_p + (\sigma + \sigma_p)^2 - \sigma_n(\sigma + \sigma_p)], \\ \sigma_x(\nu, x) &= x(\sigma_n - \sigma - \sigma_p). \end{aligned} \tag{9}$$

The endogenous state variable x_t has an interpretation in terms of experts' balance sheets. Since experts must hold all the capital in the economy, the denominator captures their assets while the numerator is the net worth of the expert sector as a whole. We can think of x_t as capturing the strength of experts' balance sheets. We know from proposition 1 that without moral hazard, experts would be able to off-load all of their idiosyncratic risk onto the market, and hence neither ν_t nor x_t would play any role in equilibrium (other than to determine consumption). In an economy with financial frictions $\phi > 0$, both idiosyncratic risk ν_t and experts' balance sheets x_t will affect the equilibrium. If aggregate risk is concentrated on experts' balance sheets, they will face large financial losses after a bad aggregate shock and their share of aggregate wealth x_t will go down ($\sigma_{x,t} > 0$), amplifying the effects of the shock. We can now give a definition for a Markov equilibrium.

DEFINITION 2. A Markov equilibrium in (ν, x) is a set of aggregate functions p, ξ, ζ, r, π and policy functions $\hat{e}, g, \hat{k}, \theta$ for experts and $\hat{c}, \sigma_{w,t}$ for households, and a law of motion for the endogenous aggregate state variable μ_x and σ_x such that

- i. ξ and ζ solve the experts' and households' HJB equations, and $\hat{e}, g, \hat{k}, \theta$ and $\hat{c}, \sigma_{w,t}$ are the corresponding policy functions, taking p, r, π and the laws of motion of ν and x as given;
- ii. markets clear:

$$\begin{aligned} \hat{e}px + \hat{c}p(1 - x) &= a - u(g), \\ p\hat{k}x &= 1, \\ \sigma_n x + \sigma_w(1 - x) &= \sigma + \sigma_p; \end{aligned}$$

- iii. and the law of motion of x satisfies (9).

This recursive definition abstracts from the absolute level of the aggregate capital stock, which we can recover using $dk_t/k_t = g_t dt + \sigma dZ_t$.

Capital holdings.—Experts' demand for capital is pinned down by the FOC from the HJB equation. Using the FOC for \hat{k} and θ , we obtain after some algebra an expression for \hat{k} :

$$\frac{a - \iota_t}{p_t} + g_t + \mu_{p,t} + \sigma\sigma'_{p,t} - r_t \leq (\sigma + \sigma_{p,t})\pi_t + \gamma p_t \hat{k}_t (\phi \nu_t)^2.$$

Idiosyncratic risk is not priced in the financial market because it can be aggregated away. However, because experts face an equity constraint that forces them to keep an exposure ϕ to the return of their capital, they know that the more capital they hold, the more idiosyncratic risk they must bear on their balance sheets, $\tilde{\sigma}_{n,t} = \phi p_t \hat{k}_t \nu_t$. Since they are risk averse, they demand a premium on capital for that idiosyncratic risk. Using the equilibrium condition $p \hat{k} x = 1$, we obtain an equilibrium pricing equation for capital:

$$\underbrace{\frac{a - \iota_t}{p_t} + g_t + \mu_{p,t} + \sigma\sigma'_{p,t} - r_t}_{\text{excess return}} = \underbrace{(\sigma + \sigma_{p,t})\pi_t}_{\text{agg. risk premium}} + \underbrace{\gamma \frac{1}{x_t} (\phi \nu_t)^2}_{\text{id. risk premium}}. \quad (10)$$

The left-hand side is the excess return of capital. The right-hand side is made up of the risk premium corresponding to the aggregate risk capital carries and a risk premium for the idiosyncratic risk it carries. When experts' balance sheets are weak (low x_t) and idiosyncratic risk ν_t high, experts demand a high premium on capital. This is how x_t and ν_t affect the economy, and we can see that without moral hazard ($\phi = 0$) neither x_t nor ν_t would play any role, and experts would be indifferent about how much capital to hold as long as it was properly priced. With moral hazard, instead, they have a well-defined demand for capital, proportional to their net worth.

It is useful to reformulate experts' problem with a fictitious price of idiosyncratic risk

$$\alpha_t = \gamma \frac{\phi \nu_t}{x_t}.$$

Under this formulation, each expert faces a complete financial market without the equity constraint, but where his own idiosyncratic risk W_i pays a premium α_t . Capital is priced as an asset with exposure $\phi \nu_t$ to this idiosyncratic risk and can be abstracted from.¹² We can verify that the ex-

¹² We can use (10) to rewrite experts' dynamic budget constraint

$$\frac{dn_{i,t}}{n_{i,t}} = (r_t + \pi_t \sigma_{n,i,t} + \alpha_t \tilde{\sigma}_{n,i,t} - \hat{e}_{i,t}) dt + \sigma_{n,i,t} dZ_t + \tilde{\sigma}_{n,i,t} dW_{i,t},$$

where the expert can freely choose $\sigma_{n,i,t}$ and $\tilde{\sigma}_{n,i,t}$. Experts' problem then is to maximize their objective function subject to an intertemporal budget constraint

$$\mathbb{E} \left[\int_0^{\infty} \tilde{\eta}_{i,u} e_{i,u} du \right] = n_0,$$

where the fictitious SDF $\tilde{\eta}_i$ follows $d\tilde{\eta}_{i,t}/\tilde{\eta}_{i,t} = -r_t dt - \pi_t dZ_t - \alpha_t dW_{i,t}$ for expert i .

pert will choose an exposure to his own idiosyncratic risk $\tilde{\sigma}_{n,t} = \alpha_t/\gamma = \phi(1/x_t)\nu_t$ as required in equilibrium. An advantage of this formulation is that the only difference between experts' and households' problem is that experts perceive a positive price $\alpha_t > 0$ for their idiosyncratic risk W_t , while households perceive a price of zero.

Aggregate risk sharing.—Optimal contracts allow experts to share aggregate risk freely and separate the decision of how much capital to hold $k_{i,t}$ from the decision of how much aggregate risk to hold $\sigma_{n,t}$. The FOC for θ for experts yields

$$\sigma_{n,t} = \underbrace{\frac{\pi'_t}{\gamma}}_{\text{myopic}} - \underbrace{\frac{\gamma - 1}{\gamma}}_{\text{hedging motive}} \sigma_{\xi,t}. \tag{11}$$

Experts' optimal aggregate risk exposure depends on a myopic risk-taking motive given by the price of risk and a hedging motive driven by the stochastic investment opportunity sets. This hedging motive will play a crucial role in concentrating aggregate risk on experts' balance sheets. It is useful to think about it in terms of income and substitution effects. Recall that ξ_t captures experts' stochastic investment opportunities in the value function (6). If the expert is risk neutral, he will prefer to have more net worth when ξ_t is high, since he can then get more utility out of each unit of net worth (more “bang for the buck”). This is the substitution effect. On the other hand, when ξ_t is low, he requires more net worth to achieve any given level of utility. If the expert is risk averse, he will prefer to have more net worth when ξ_t is low to stabilize his utility across states of the world. This is the income effect. Which effect dominates depends on the risk aversion parameter. I focus on the empirically relevant case with $\gamma > 1$ where the income effect dominates.

Households have analogous FOCs for aggregate risk sharing,

$$\sigma_{w,t} = \frac{\pi'_t}{\gamma} - \frac{\gamma - 1}{\gamma} \sigma_{\zeta,t}, \tag{12}$$

where the only difference is that households' investment opportunity sets are captured by ζ_t instead of ξ_t . Since households cannot buy capital, its price and idiosyncratic risk premium do not affect them, but they still face a stochastic investment opportunity set from interest rates r_t and the price of aggregate risk π_t .

The volatility of balance sheets $\sigma_{x,t}$ arises from the interaction of experts' and households' risk-taking decisions. Using the equilibrium condition $\sigma_n x + \sigma_w(1 - x) = \sigma + \sigma_b$, we obtain the following aggregate risk-sharing equation:

$$\sigma_{x,t} = (1 - x_t)x_t \frac{1 - \gamma}{\gamma} \sigma_{\Omega,t}, \quad (13)$$

where $\Omega_t = \xi_t/\zeta_t$ captures the investment opportunities of experts relative to households and follows the law of motion $d\Omega_t = \Omega_t \mu_{\Omega,t} dt + \Omega_t \sigma_{\Omega,t} dZ_t$. The term $(1 - x_t)x_t$ arises because households must take the other side of experts' position; the $(1 - \gamma)/\gamma$ term captures the substitution and income effects; and $\sigma_{\Omega,t} = \sigma_{\xi,t} - \sigma_{\zeta,t}$ captures how experts' and households' relative investment opportunities depend on aggregate shocks. Since experts and households cannot both hedge in the same direction in equilibrium, it is the difference in their hedging motives, captured by their relative investment opportunities, that can cause aggregate risk to be concentrated on experts' balance sheets $\sigma_{x,t} > 0$.

To understand aggregate risk sharing better, notice that because experts have the option of investing in capital, they always get more utility per dollar of net worth than households, so $\Omega_t = \xi_t/\zeta_t > 1$ always. This ratio is not constant, however: it depends on the aggregate state of the economy. Equation (13) says that if the income effect dominates ($\gamma > 1$), agents will share aggregate risk so that experts have a smaller share of aggregate wealth x_t after an aggregate shock that improves experts' relative investment opportunities Ω_t .

Experts' and households' investment opportunities depend on experts' share of aggregate wealth x_t and so are endogenously determined in equilibrium in a two-way feedback loop: aggregate risk is concentrated on experts' balance sheets to hedge stochastic relative investment opportunities, but the effect of aggregate shocks on experts' relative investment opportunities depends on the concentration of aggregate risk on experts' balance sheets. We can use Ito's lemma to obtain a simple expression for the volatility of Ω_t :

$$\sigma_{\Omega,t} = \underbrace{\frac{\Omega_v}{\Omega} \sigma_v \sqrt{v_t}}_{\text{exogenous}} + \underbrace{\frac{\Omega_x}{\Omega} \sigma_{x,t}}_{\text{endogenous}}, \quad (14)$$

where the function Ω and its derivatives are evaluated at (v_t, x_t) . The locally linear representation allows a neat decomposition into an *exogenous* source, from the uncertainty shock to v_t , and an *endogenous* source from optimal contracts' aggregate risk sharing $\sigma_{x,t}$. We can solve for the fixed point of this two-way feedback:

$$\sigma_{x,t} = \frac{(1 - x_t)x_t \frac{1 - \gamma}{\gamma} \frac{\Omega_v}{\Omega}}{1 - (1 - x_t)x_t \frac{1 - \gamma}{\gamma} \frac{\Omega_x}{\Omega}} \sqrt{v_t} \sigma_v. \quad (15)$$

Notice that even though the presence of moral hazard does not directly restrict experts' ability to share aggregate risk, it introduces hedging motives through the general equilibrium that would not be present without moral hazard, as shown by proposition 1.

C. *Brownian TFP Shocks*

When aggregate shocks come only in the form of Brownian TFP shocks ($\sigma_v = 0$) and we allow agents to write contracts on all observable variables, there is no balance sheet channel. After a negative TFP shock the value of all assets p_k falls and experts and households divide losses proportionally, so $\sigma_{x_i} = 0$. Relative investment opportunities Ω_i are not affected by the aggregate shock, and consequently, there is no concentration of aggregate risk. Balance sheets x_i may still affect the economy because of the presence of financial frictions derived from the moral hazard problem, but they will not be exposed to aggregate risk and hence will not play any role in the amplification of aggregate TFP shocks. In fact, the equilibrium is completely deterministic up to the direct effect of TFP shocks on the aggregate capital stock.

PROPOSITION 2. With only Brownian TFP shocks ($\sigma_v = 0$), if agents can write contracts on the aggregate state of the economy, the balance sheet channel disappears: the state variable x_b , the price of capital p_b , the growth rate of the economy g_b , the interest rate r_b , and the price of risk π_i all follow deterministic paths and are not affected by aggregate shocks.

The neutrality result of proposition 2 has two ingredients: (1) Optimal contracts separate the decision of how much capital to buy (leverage) from the decision of how much aggregate risk to hold (risk sharing). Experts and households will share aggregate risk to hedge their relative investment opportunities Ω_b , as given by expression (13). And (2) aggregate Brownian TFP shocks do not affect the relative investment opportunities Ω_i because the economy is scale invariant with respect to effective capital k . The exogenous source of volatility in Ω_i disappears, so we are left with only the endogenous component in expression (14). With no exogenous source, however, the unique Markov equilibrium has deterministic relative investment opportunities Ω_i and hence no concentration of aggregate risk on experts' balance sheets. Without any source of aggregate volatility, the economy then follows a deterministic path. This neutrality result should be understood as a theoretical benchmark. These TFP shocks are very salient and widely used in the literature. Proposition 2 suggests that in order to understand the concentration of aggregate risk, we should move away from the benchmark setting with Brownian TFP shocks.

Implementation and constrained contracts.—We can compute experts' trading position in the normalized market index θ_i explicitly:

$$\theta_t = (\sigma + \sigma_{p,t}) \underbrace{\frac{x_t - \phi}{x_t}}_{\geq 0} + \underbrace{\frac{\sigma_{x,t}}{x_t}}_{=0} = \underbrace{\frac{x_t - \phi}{x_t}}_{\geq 0} \sigma.$$

Experts are required to hold a fraction ϕ of their equity, which already exposes them to a fraction ϕ of aggregate risk. Since they would like to hold a fraction of aggregate risk proportional to their share of aggregate wealth x_t , they will long or short the normalized market index to hit this target.

In general, the economy may be hit by a large number of orthogonal aggregate shocks, that is, $d > 1$. The neutrality result in proposition 2 does not require complete markets, only that leverage and aggregate risk sharing be separated. In terms of implementation in a financial market, we need the financial market to span the exposure to aggregate risk of the return of capital $(\sigma + \sigma_{p,t})dZ$. In this case, an expert can buy capital and immediately get rid of the aggregate risk using financial instruments. In particular, if experts can short the equity of their competitors who have an exposure to aggregate risk similar to theirs, they can get rid of the aggregate risk in their capital. In a competitive market, there is a large number of competitors, so their idiosyncratic risks can be aggregated away. In other words, an index made up competitors' equity is exactly the instrument required to separate leverage from risk sharing and obtain the neutrality result.

Consider in contrast what happens if we rule out contracts on aggregate shocks, that is, if we constrain experts to $\theta_t = 0$. In this case, experts' leverage $p_t \hat{k}_t$ and aggregate risk sharing $\sigma_{n,t}$ become entangled. We can see this in experts' budget constraint, where we now have $\sigma_{n,t} = \phi p_t \hat{k}_t (\sigma + \sigma_{p,t})$. In the simplest case with $\phi = 1$ as in the baseline setting in Brunnermeier and Sannikov (2014), since experts are leveraged in equilibrium, $p_t k_t > n_t$, aggregate risk is concentrated on their balance sheets and their share of aggregate wealth x_t falls after a bad aggregate shock:

$$\sigma_{x,t} = x_t (\sigma_{n,t} - \sigma - \sigma_{p,t}) = x_t \underbrace{(p_t \hat{k}_t - 1)}_{>0} (\sigma + \sigma_{p,t}) > 0.$$

This reduces their ability to hold capital and lowers asset prices, further hurting their balance sheets, and amplifying and propagating the initial shock. This is precisely the mechanism behind the balance sheet channel in He and Krishnamurthy (2012) and Brunnermeier and Sannikov (2014).¹³

¹³ In He and Krishnamurthy (2012) a similar mechanism underlies the volatility of experts' net worth (specialists in their model), but the price of capital falls because households are more impatient and interest rates must rise for consumption goods markets to clear.

D. *Uncertainty Shocks*

In contrast to TFP shocks, uncertainty shocks that increase idiosyncratic risk create balance sheet recessions. Because of the “skin in the game” constraint, experts must keep a fraction of the idiosyncratic risk in their capital, so after an uncertainty shock, asset prices and investment fall. Even though experts can share aggregate risk freely, they choose to be highly exposed to this aggregate shock ex ante in order to take advantage of ex post investment opportunity sets. As a result, financial losses are disproportionately concentrated on the balance sheets of experts. Weak balance sheets further depress asset prices and investment, which in turn amplifies experts’ incentives to take even more aggregate risk ex ante in a two-way feedback loop. In addition, an increase in *idiosyncratic* risk leads to an endogenous increase in *aggregate* risk and to a low interest rate and high risk premium.

Numerical calibration.—The strategy to solve for the equilibrium with uncertainty shocks is to map it into a system of partial differential equations (PDEs) for $p(v, x)$, $\xi(v, x)$, and $\zeta(v, x)$. For simplicity, I consider a single aggregate shock that affects both effective capital and idiosyncratic risk.

I use the following calibration.¹⁴ *Technology:* I normalize $a = 1$ and set the volatility of TFP shocks $\sigma = 1.25$ percent in order to target an (annualized) volatility of quarterly GDP of 2 percent. For the investment technology I use a quadratic specification $\iota(g) = A(g + \delta)^2 + B(g + \delta)$, where $\delta = 5$ percent. I pick A and B so that the annualized average growth rate of GDP is 2 percent and the average investment to GDP ratio is 20 percent. *Preferences:* I set the discount rate $\rho = 6.65$ percent to target an average risk-free rate of 1 percent. In order to have a long-run stationary distribution, I introduce a Poisson retirement rate for experts τ . When they retire, they become households. I set $\tau = 1.15$ so that average leverage is $l = A/NW = 10$.¹⁵ For the risk aversion and EIS I use reasonable numbers from the literature: $\gamma = 5$ and $EIS = 2$. Below I explore the role that each of these parameters plays in the model in more detail and discuss empirical evidence. *Moral hazard:* Following He and Krishnamurthy (2012), I set $\phi = 0.2$ to match the 20 percent share of profits that hedge funds charge. *Idiosyncratic risk:* I use data from Campbell et al. (2001) on the id-

¹⁴ While numerical results are specific to this calibration, the qualitative properties of the equilibrium are very general as long as $\gamma > 1$ and $\psi^{-1} > 1$. Below I explore the role of these two parameters in detail.

¹⁵ The retirement rate $\tau = 1.15$ implies a very short half-life for experts. It is better to keep in mind that this is just a tool to obtain a stationary distribution. A lower τ would lead to a higher share of aggregate wealth for experts on average and lower average leverage. Alternatively, we could have assumed a higher discount rate for experts, as in Brunnermeier and Sannikov (2014).

idiosyncratic volatility of stock returns to estimate the process for idiosyncratic risk

$$d\nu_t = 1.38(0.25 - \nu_t)dt - 0.17\sqrt{\bar{\nu}_t}dZ_t.$$

The long-run mean is $\bar{\nu} = 25$ percent and the long-run standard deviation $\sigma_{\nu}\sqrt{\bar{\nu}/2\lambda} = 5.1$ percent. The autoregression coefficient $\lambda = 1.38$ implies a half-life of half a year, so we are considering short-lived shocks. Appendix B shows the calibration and numerical solution procedure in detail.

Uncertainty shocks and balance sheet recessions.—An uncertainty shock exogenously increases idiosyncratic risk ν_t . From the pricing equation for capital (10) we see that this raises the premium for idiosyncratic risk and therefore drives down the price of capital, as can be seen in figure 1. Because of the moral hazard problem, experts must keep a fraction of the idiosyncratic risk in the capital they manage, so capital becomes less attractive when ν_t is higher. With EIS greater than one, the price of capital falls. As a result, investment also falls, given by the FOC $\iota'(g_t) = p_t$.¹⁶

The resulting financial losses are concentrated on experts' balance sheets, so their share of aggregate wealth x_t goes down after an uncertainty shock. Figure 1 shows $\sigma_x > 0$ throughout. Weaker balance sheets further depress asset prices, as can also be seen in figure 1. With weaker balance sheets, experts must leverage up more to hold all the capital in the economy, so their exposure to idiosyncratic risk is higher. They will accept this only if capital pays an appropriately higher excess return. As a result, an uncertainty shock produces a balance sheet recession: a downturn with lower investment and asset prices and financial losses concentrated on the balance sheets of experts that amplify the effect of the initial shock.

To understand why financial losses are concentrated on the balance sheets of experts, recall equation (13) for aggregate risk sharing. When risk aversion is greater than one, the income effect dominates and optimal contracts will give experts a smaller share of aggregate wealth x_t after an aggregate shock that improves their investment opportunities $\Omega_t = \xi_t/\zeta_t$ relative to households, in order to stabilize their utility across states of the world. Figure 1 shows that experts' relative investment opportunities Ω_t are better when there is more idiosyncratic risk ν_t and when their share of aggregate wealth x_t is low (weak balance sheets). To understand why this is the case, recall that the only difference between households and experts is that by investing in capital, experts perceive a positive

¹⁶ With $\text{EIS} < 1$ an intertemporal income effect would dominate: even though capital is less attractive, agents feel poorer in certainty equivalent terms and would try to accumulate more, so the price of capital and investment would go up (the interest rate r would fall so that [10] holds). I explore the role of both risk aversion γ and EIS ψ^{-1} below.

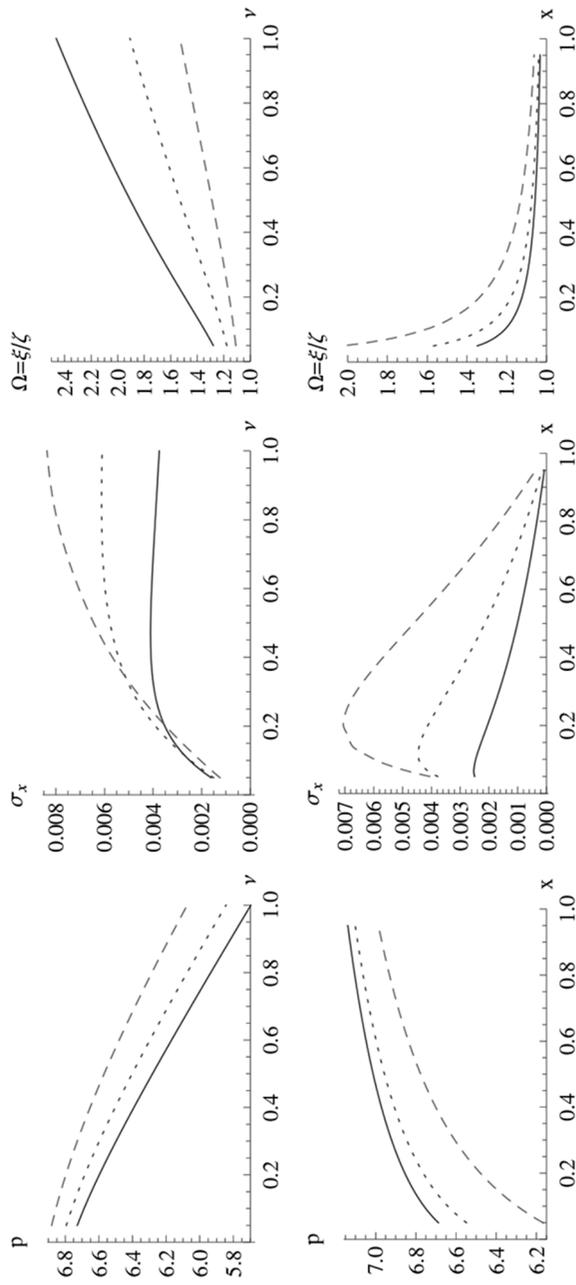


FIG. 1.—The price of capital p , volatility of x , σ_x , and relative investment opportunities $\Omega = \xi/\zeta$, as functions of y (above) for $x = 0.05$ (solid), $x = 0.10$ (dotted), and $x = 0.2$ (dashed), and as a function of x (below) for $y = 0.1$ (solid), $y = 0.25$ (dotted), and $y = 0.6$ (dashed).

price $\alpha_i = \gamma(\phi\nu_i/x_i)$ for their own idiosyncratic risk W_i . This price is higher when idiosyncratic risk is higher, so experts, who in equilibrium go long on their own idiosyncratic risk, benefit from this. Since financial losses are concentrated on experts' balance sheets, their share of aggregate wealth x_i goes down after an uncertainty shock, and this further drives α_i and Ω_i up, providing further incentives for experts to take on aggregate risk ex ante, in a two-way amplification loop.¹⁷

It is worth emphasizing that experts are not necessarily better off during downturns, first, because they (endogenously) face large financial losses. But even conditional on net worth, experts' investment opportunities (captured by ξ_i) may well be worse after an uncertainty shock because interest rates r_t and the risk premia π_t are also affected. What matters for aggregate risk sharing, however, is that the investment opportunities of experts relative to households $\Omega_i = \xi_i/\zeta_i$ improve after an uncertainty shock, because experts at least get higher premiums α_i on idiosyncratic risk. As a result, although experts and households are equally risk averse, for a given price of aggregate risk π_t , experts find taking aggregate risk more attractive than households, and in equilibrium the market concentrates a disproportionate share of aggregate risk on the balance sheets of experts.

Figure 2 shows how an uncertainty shock affects the financial market. The risk-free interest rate r_t falls (it can even become negative) and the price of aggregate risk π_t goes up, both because idiosyncratic risk ν_t goes up and because balance sheets x_t become weaker. In addition, although the exogenous shock increases only idiosyncratic risk ν_t , it endogenously amplifies aggregate risk $\sigma + \sigma_{p,t}$. The model therefore provides an explanation for the observation that idiosyncratic and aggregate volatility seem to move together (see Bloom et al. 2012) and generates stochastic risk premia.¹⁸

What is it about uncertainty shocks that leads to the concentration of aggregate risk? When agents can write contracts on the aggregate state of the economy, they will hedge their relative investment opportunities according to equation (13). Any aggregate shock that improves the forward-looking investment opportunities of experts relative to households can lead to the concentration of aggregate risk on their balance sheets. Notice that although for calibration purposes I assumed that the single

¹⁷ Notice that the endogenous response of asset prices amplifies the effect of the exogenous shock on the balance sheets of experts, as in Kiyotaki and Moore (1997). In that paper, however, this happens ex post because experts cannot hedge this risk. Here, instead, it happens ex ante because they can hedge and choose to increase their exposure to aggregate risk in anticipation of the response of asset prices to the exogenous shock.

¹⁸ There is a large literature on stochastic risk premia. Campbell and Cochrane (1999) introduce habit formation to obtain stochastic risk premia, while He and Krishnamurthy (2012) introduce financial frictions.

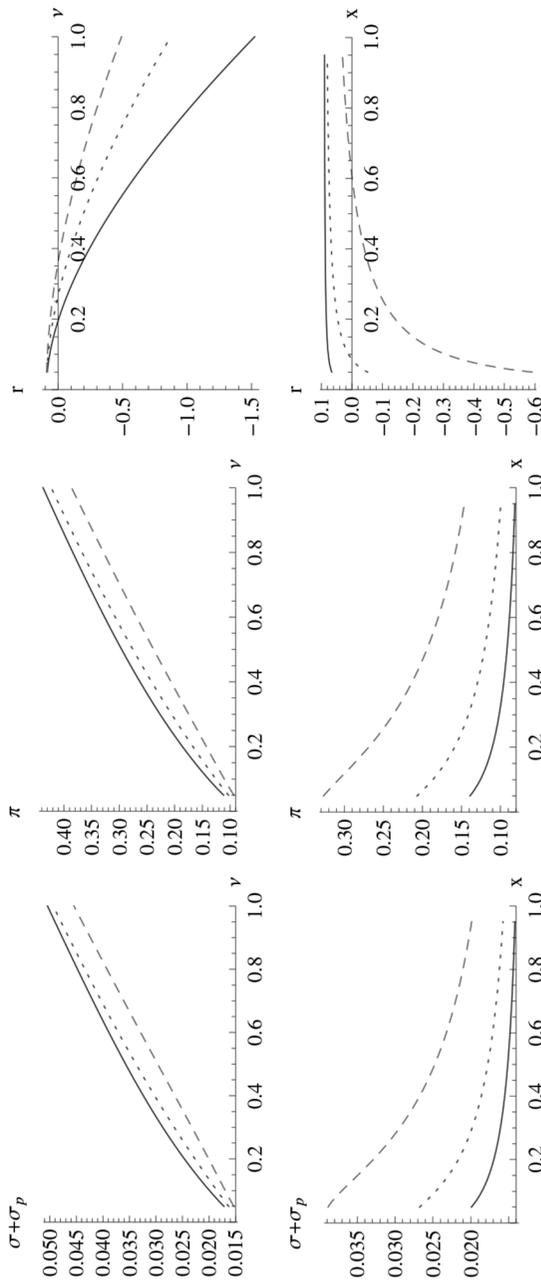


FIG. 2.—Aggregate risk $\sigma + \sigma_p$, the price of risk π , and the risk-free rate r as functions of v (above) for $x = 0.05$ (solid), $x = 0.10$ (dotted), and $x = 0.2$ (dashed), and as a function of x (below) for $v = 0.1$ (solid), $v = 0.25$ (dotted), and $v = 0.6$ (dashed).

aggregate shock Z also affects TFP, that is, $\sigma > 0$, this does not play any role in inducing concentration of aggregate risk. We would still get concentration of aggregate risk even with $\sigma = 0$. On the other hand, shocks to the financial friction ϕ will have the same effect as uncertainty shocks, because $\phi\nu$ enter together in the model: what matters is the idiosyncratic risk that cannot be insured away. Other aggregate shocks can be studied using the tools developed here.

Dynamics and long-run distribution.—To understand the model dynamics, it is useful to look at the phase diagram in figure 3. The state variables (ν, x) live in $(0, \infty) \times (0, 1)$, never reaching any boundary. Uncertainty shocks shift the system as indicated by the diagonal arrows. A bad uncertainty shock raises idiosyncratic risk ν and reduces experts' share of aggregate wealth x . In the absence of shocks, the system would converge to a "steady state." The dashed lines indicate actual equilibrium paths toward the steady state, with dots indicating the progress at quarterly intervals (recall we have assumed that uncertainty shocks are relatively short-lived, with a half-life of two quarters). The system has strong forces that push it to-

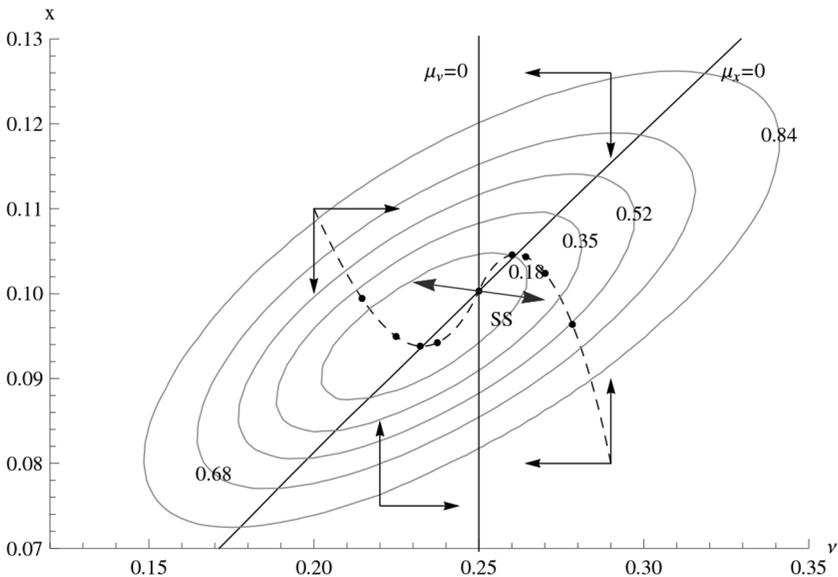


FIG. 3.—Phase diagram showing the loci of $\mu_\nu = 0$ and $\mu_x = 0$. The perpendicular arrows indicate the direction of the drift, and the dashed lines are actual equilibrium paths in the absence of shocks, leading to a "steady state." The dots indicate progress at quarterly intervals. The diagonal arrows indicate the effects of uncertainty shocks (the slope $\sigma_x/\sigma_\nu\sqrt{\nu}$ has the same sign everywhere). The contour lines of the density of the long-run distribution are plotted in the background (for density levels 51, 102, 153, 204, and 255 from the outer line toward the steady state) and the probability contained within each line indicated on the graph.

ward the steady state. When experts' balance sheets are very weak and idiosyncratic risk high (low x and high ν), excess returns are high and experts postpone consumption. As a result, although uncertainty shocks weaken experts' balance sheets on impact, experts subsequently accumulate net worth and rebuild their balance sheets relatively fast, leading to possibly stronger balance sheets in the "medium run." This is reflected in a positive correlation between idiosyncratic risk and experts' share of aggregate wealth in the long-run distribution, also plotted in figure 3. Brunnermeier and Sannikov (2014) refer to this effect as a "volatility paradox": a more volatile economy induces experts to accumulate more net worth and leads to stronger balance sheets.

How big are these effects?—I simulate the model to get an idea of the relevance of the mechanisms described here. As a benchmark, with only TFP shocks (if we shut down uncertainty shocks but keep the calibration) the volatility of growth in GDP, consumption, investment, and asset prices would be 2 percent: we targeted 2 percent volatility of GDP growth, and without uncertainty shocks there is no other amplification mechanism in the model. This would also be the case with uncertainty shocks but no moral hazard, since in that case there are no financial frictions and idiosyncratic risk would be perfectly shared. In contrast, in the full model with financial frictions and uncertainty shocks, while the volatility of output growth remains at 2 percent because we targeted it, the average volatility of investment growth increases to 4.84 percent (the volatility of consumption growth decreases to 1.56 percent to compensate) and the volatility of growth in asset values p_k increases to 4.36 percent.¹⁹ This represents a significant amplification compared to the benchmark with only TFP shocks.

The model also delivers a significant concentration of aggregate risk on the balance sheets of experts. Consider $m = \sigma_n / (\sigma + \sigma_p)$ as a measure of this concentration: if an aggregate shock reduces the value of experts' assets by 1 percent, their net worth will fall by $m \times 1$ percent. The average m in the model is 2.79 (it would be one in the benchmark with only TFP shocks). To get an idea of what this implies, from 2007:II to 2009:I the Case-Shiller index fell by 27.5 percent. If m was constant, experts would lose 77 percent of their net worth in the model.²⁰ For the sake of comparison, He and Krishnamurthy (2014) report that in the same period fi-

¹⁹ In this simple model only TFP shocks to effective capital k can affect output $y = ak$ in the short run. As a result, while consumption remains procyclical, the shock to idiosyncratic risk has a countercyclical effect on consumption.

²⁰ Alternatively, the Federal Reserve Economic Data series Households and Nonprofit Organizations; Net Worth implies total financial losses of 20 percent for the same period, which with a constant $m = 2.79$ would yield a loss of 56 percent of experts' net worth in the model.

nancial intermediaries lost 70 percent of their net worth. However, m is nonlinear, and it is difficult to generate such a large shock to the value of assets in the model, so I explicitly consider a large uncertainty shock below. Weaker balance sheets amplify the direct effect of higher idiosyncratic risk. Consider $f = \sigma_p / (p_v / p) \sigma_v \sqrt{v}$ as a measure of the amplification we get from weaker balance sheets: if the direct effect of an increase of idiosyncratic risk reduces the price of capital by 1 percent, there is an additional amplification through weak balance sheets and the price of capital falls by $f \times 1$ percent (we can also interpret f in terms of investment). The average f in the model is 1.09, meaning that the endogenous response of experts' balance sheets further reduces the price of capital by an extra 9 percent on average. While this is an economically significant amplification channel, the direct effect of higher idiosyncratic risk dominates (recall, however, that idiosyncratic risk matters only because of financial frictions). This is not that surprising in light of the results of Christiano et al. (2014), for example, who show that shocks to idiosyncratic risk can play a preeminent role in business cycles. On the other hand, it is possible that a richer model that allows for fire sales and bankruptcy, for example, may find a larger relative role for balance sheets. Empirically, it is unclear how big we would like f to be, since we observe only the joint effect of shocks. Overall, uncertainty shocks generate significant economic fluctuations that look like balance sheet recessions, with lower investment and asset prices, and financial losses heavily concentrated on the balance sheets of experts.

To better understand the size of these mechanisms, let us consider the effect of a large uncertainty shock. Take an economy with an initial low level of idiosyncratic risk $\nu_0 = 10$ percent and the long-run level of x associated with ν_0 , so $x_0 = 0.04$, which implies a leverage ratio of 24. Suppose that an uncertainty shock hits the economy and drives idiosyncratic risk to $\nu_1 = 60$ percent. As a result, investment falls by 22.5 percent and asset values by 14.74 percent, of which 6.6 percent corresponds to the direct effect of lower effective capital (output also falls by 6.6 percent) and the rest to the uncertainty shock. However, the net worth of experts falls by 47.95 percent, for an implied average m of 3.2. The weak balance sheets amplify the direct effect of the uncertainty shock by roughly 9.5 percent.

In terms of asset pricing, the average price of aggregate risk π is 0.19. A market portfolio has a conditional volatility $\sigma + \sigma_{p,t}$ of 2.45 percent, which implies an average risk premium $\pi(\sigma + \sigma_{p,t})$ of 0.46 percent. In contrast, with only TFP shocks $\pi = \gamma\sigma = 0.0625$, and the excess return of the market portfolio would be 0.08 percent. The uncertainty shock has a positive price in the financial market, consistent with empirical evidence in Herskovic et al. (2016). In addition, uncertainty shocks affect the financial market. After a large uncertainty shock as above, the price of risk π more than doubles from 0.14 to 0.33, while the conditional vol-

atility of the market portfolio goes from 2.02 percent to 3.91 percent. As a result, the risk premium on the market portfolio shoots up from 0.29 percent to 1.31 percent. At the same time, the risk-free rate drops from 6.16 percent before the shock to an unrealistic -105.91 percent.

The role of EIS ψ^{-1} and relative risk aversion γ .—Epstein-Zin preferences separate agents' relative risk aversion γ from their EIS ψ^{-1} . Each plays a different role in the model. A relative risk aversion of $\gamma > 1$ is needed for financial losses to be concentrated on the balance sheet of experts. With $\gamma < 1$ the substitution effect would dominate and financial losses would be concentrated on households, as experts try to preserve more "dry powder" for downturns. The EIS > 1 instead is needed for the price of capital and investment to go down when its risk premium goes up after an uncertainty shock. With EIS > 1 , an intertemporal substitution effect dominates, and agents prefer to consume when capital is unattractive because of high idiosyncratic risk and weak balance sheets. With EIS < 1 instead, an income effect would dominate: agents feel poorer in certainty equivalent terms and try to accumulate more capital. As a result the price of capital and investment would go up (the risk-free rate would drop so that capital still pays a higher excess return). In the special case with EIS = 1, the price of capital is constant.

We therefore need both a risk aversion and an EIS greater than one in order for uncertainty shocks to create downturns with financial losses concentrated on experts' balance sheets, for which Epstein-Zin preferences are required. While the empirical evidence on risk aversion supports $\gamma > 1$, for the EIS the evidence is mixed, but it is a common ingredient of models with stochastic volatility.²¹

A simplified environment.—To better understand why experts' share of aggregate wealth x falls after uncertainty shocks, it is useful to consider a simplified environment with $\psi = 1$ and no retirement, $\tau = 0$. Consider an economy with a constant idiosyncratic risk ν that is realized at time $t = 0$ and can take values with equal probability $\nu \in \{\nu_l, \nu_h\}$ with $\nu_l < \nu_h$. Once ν is realized, we have an economy without uncertainty shocks (so $\sigma_{x,t} = 0$), but we can think about how experts and households share the risk over ν before it is realized.²² Appendix A develops this setting in detail.

Imagine that experts and households start with net worth n_0 and w_0 before ν has realized and can trade Arrow securities to share this risk. No-

²¹ Beeler and Campbell (2012) use an EIS of 1.5, while Bansal et al. (2014) use an EIS of 2. Gruber (2013) estimates an EIS of 2 on the basis of variation across individuals in the capital income tax rate, while Mulligan (2002) finds an EIS considerably higher. On the other hand, Hall (1988) and Vissing-Jørgensen (2002) find an EIS < 1 .

²² Notice that without retirement, $x_t \rightarrow 1$ in the long run for any ν . We are interested in how x responds on impact to the realization of ν at $t = 0$.

tice that as the argument above shows, with $\psi = 1$ the price of capital and the growth rate do not depend on ν or x , so total wealth $n + w = \rho k$ is not affected by the realization of ν . Experts solve

$$\max_{n_h, n_l} \frac{1}{2} \frac{(\xi_h n_h)^{1-\gamma}}{1-\gamma} + \frac{1}{2} \frac{(\xi_l n_l)^{1-\gamma}}{1-\gamma}$$

subject to $q_h n_h + q_l n_l = n_0$,

where q_h and q_l are the prices of Arrow securities. Households have an analogous problem. Using the FOCs we can get

$$\frac{n_h/w_h}{n_l/w_l} = \left(\frac{\Omega_h}{\Omega_l} \right)^{(1-\gamma)/\gamma}.$$

This captures the same intuition as (13) in terms of hedging relative investment opportunities Ω . With $\gamma > 1$, experts' share of aggregate wealth x must be smaller when their relative investment opportunities are better (high Ω). In this setting we can prove that experts' relative investment opportunities $\Omega(x; \nu) = \xi(x; \nu)/\zeta(x; \nu) > 1$ are better if idiosyncratic risk ν is high and their share of aggregate wealth x is low.

PROPOSITION 3. $\Omega(x; \nu)$ is strictly increasing in ν and strictly decreasing in x .

The intuition is that the only difference between experts and households is that experts get a premium for idiosyncratic risk $\alpha = \gamma(\phi\nu/x)$, which is increasing in ν and decreasing in x . Figure 1 shows that this property also holds in the numerical solution of the full model. Using proposition 3 we can show that experts' share of aggregate wealth x falls if high idiosyncratic risk ν_h is realized.

PROPOSITION 4. For $\gamma > 1$, experts' share of aggregate wealth x falls if ν_h is realized and goes up if ν_l is realized, that is, $x_h < x_0 < x_l$.

This is the counterpart of $\sigma_x > 0$ in the full model, also shown in figure 1. The advantage of this simple environment is that it allows a clear analytical characterization that helps us understand the mechanisms in the full model.

IV. Financial Regulation

In the previous section I have shown how agents' privately optimal contracts can lead to the concentration of aggregate risk on experts' balance sheets. It is worth asking if this concentration of aggregate risk is efficient. In standard models of balance sheet recessions driven by TFP shocks, where contracts cannot be written on the aggregate state of the economy, providing aggregate insurance to experts in order to eliminate the con-

centration of aggregate risk is a Pareto improving policy. By doing this the planner is in fact completing the market. Brunnermeier and Sannikov (2014), for example, show how a social planner can achieve first-best allocations in this way. In this section I show that the unregulated competitive equilibrium is not constrained efficient, but that proportional aggregate risk sharing is not optimal either.

A class of policies.—To understand the efficiency of aggregate risk sharing, consider a social planner who can regulate experts’ exposure to aggregate risk $\sigma_{n,t}$. Experts’ problem then is to pick a strategy (e, g, k, θ) to maximize utility $U(e)$ subject to the budget constraint (3) and hitting the $\sigma_{n,t}$ mandated by the government.²³ Households are not regulated, so their problem is unchanged, and prices $p, r,$ and π adjust to clear markets, as in the unregulated competitive equilibrium. In particular, let us consider a class of policies in which the planner picks a constant $\kappa \in \mathbb{R}$ and implements $\sigma_{n,t}/\sigma_{w,t} = \kappa$ in equilibrium by setting the mandate²⁴

$$\sigma_{n,t} = (\sigma + \sigma_{p,t}) \frac{\kappa}{1 + (\kappa - 1)x_t}. \tag{16}$$

For a given $\kappa \in \mathbb{R}$, the *regulated competitive equilibrium* is a competitive equilibrium with a modified experts’ problem that takes as given the mandate (16). The higher κ is, the more aggregate risk is concentrated on experts’ balance sheets. For example, if we set $\kappa = 1$, we get proportional aggregate risk sharing, $\sigma_n = \sigma_w = \sigma + \sigma_p$. Appendix B shows how to characterize the competitive equilibrium under this policy.

Starting from some initial (ν_0, x_0) , the planner picks a policy $\kappa \in \mathbb{R}$. Let

$$U^\kappa(\nu, x) = \frac{[\xi^\kappa(1 - x)p^\kappa]^{1-\gamma}}{1 - \gamma}$$

and

$$U(\nu, x) = \frac{[\xi(1 - x)p]^{1-\gamma}}{1 - \gamma}$$

denote the utility of households under policy κ and in the unregulated competitive equilibrium, respectively, and

$$V^\kappa(\nu, x) = \frac{(\xi^\kappa x p^\kappa)^{1-\gamma}}{1 - \gamma}$$

²³ Notice that since the moral hazard problem does not limit experts’ aggregate risk-sharing decisions and both θ and k are observable and contractible, this policy intervention is not allowing agents to do anything they could not do on their own. It is simply regulating their observable behavior.

²⁴ Notice that the mandate (16) involves equilibrium objects, but each expert takes it as given.

and

$$V(\nu, x) = \frac{(\xi x \phi)^{1-\gamma}}{1-\gamma}$$

the utility of experts.²⁵ To make welfare comparisons straightforward, the planner also carries out a one-time wealth transfer between experts and households in order to keep households indifferent.²⁶ The planner's problem starting at (ν_0, x_0) then is

$$V^*(\nu_0, x_0) = \max_{\kappa, x_1} V^*(\nu_0, x_1)$$

$$\text{subject to } U^*(\nu_0, x_1) = U(\nu_0, x_0).$$

Optimal policy.—I solve for the optimal policy κ^* numerically (see App. B). I use the same calibration as in the previous section and pick the steady state of the unregulated competitive equilibrium as the initial state, $\nu_0 = 25$ percent and $x_0 = 10$ percent. I find an optimal $\kappa^* = 1$ percent, which means that the planner concentrates almost all financial risk on households and almost fully insures experts. The planner prefers countercyclical balance sheets that can dampen the effect of higher idiosyncratic risk. The average concentration of aggregate risk m is 0.011, so if the value of experts' assets falls by 1 percent, their net worth falls by only $m \times 1$ percent = 0.011 percent. As a result, the balance sheet channel reverses: experts' share of aggregate wealth x_t goes up after uncertainty shocks, $\sigma_x < 0$, and this dampens their effect. With higher x_t experts need a lower leverage ratio to hold the capital stock, and this dampens the impact of higher ν_t on their exposure to idiosyncratic risk $\tilde{\sigma}_{n,t} = (1/x_t)\phi\nu_t$. Below I explain why this is welfare improving. The price of capital and investment still go down after uncertainty shocks, but less so than in the unregulated competitive equilibrium. The average amplification from the balance sheet channel f is 0.95, which means that the endogenous response of experts' balance sheets reduces the fall in the price of capital by 5 percent, as opposed to the 9 percent amplification in the unregulated competitive equilibrium. This is reflected in a lower average conditional volatility of the price of capital σ_p , which falls from 1.2 percent to 1.04 percent. The volatility of investment growth falls from 4.84 percent to 4.43 percent, and the volatility of the growth of asset values falls from 4.36 percent to

²⁵ I am normalizing $k_0 = 1$ and ignoring the distribution of wealth within experts and households because it enters only multiplicatively and does not affect the choice of κ .

²⁶ This one-time, unexpected initial transfer also does not affect the contractual setup. It only makes welfare comparisons straightforward. Note that transfers on their own are not Pareto improving. Also, the welfare impact of the policy intervention depends on the state (ν_0, x_0) of the economy when it is implemented, but we can solve it for any initial (ν_0, x_0) .

4.11 percent (recall that in both cases 2 percent comes from the volatility of growth of effective capital). The welfare gains from this policy are equivalent to a 4.33 percent proportional increase in experts' consumption (households are indifferent by construction, so this is considerably less than a 4.33 percent increase in everyone's consumption).

Source of inefficiency.—To understand why the competitive equilibrium is not constrained efficient, notice that there is an externality in this setting because the value of capital relative to the net worth of the expert, $p_t k_{i,t}/n_{i,t}$, appears in the incentive compatibility constraint. To see why, notice that experts do not derive utility from diverting capital directly, only from consumption. Short-term contracts cannot prevent agents from trading the diverted capital to get more consumption—either immediately or in the future through new contracts. Since their continuation utility depends on their net worth, experts must be exposed to idiosyncratic risk (in net worth and therefore in consumption) proportionally to the value of capital relative to their net worth $n_{i,t}$, that is, $\tilde{\sigma}_{i,n,t} = (p_t k_{i,t}/n_{i,t})\phi\nu_t$. Agents, however, do not internalize how their actions affect the incentive compatibility constraints of others through prices. This provides some scope for a social planner to improve over the competitive allocation by affecting the equilibrium behavior of $p_t k_t/n_t$.

Dávila and Korinek (2016) refer to this type of externality as a “binding price constraint” externality. It is important to distinguish it from the externality in Lorenzoni (2008), where incomplete aggregate risk sharing prevents marginal rates of substitution from equalizing across aggregate states. Raising the price of capital in some aggregate state then is a way of transferring wealth to its holder and improving the allocation. Here instead, since aggregate risk sharing is not constrained, it is easy to check that agents equalize the marginal rates of substitution across aggregate states (the ratio of marginal utilities $\partial_e f(e, V)/\partial_c f(c, U)$ has zero aggregate volatility). A policy that distorts the allocation of aggregate risk by reducing $\sigma_{x,t}$ (so experts' share of aggregate wealth x_t is larger after a bad shock than in the competitive equilibrium and smaller after a good one) does not have a first-order benefit from shifting consumption between households and experts. However, it does relax the idiosyncratic risk-sharing constraint after bad shocks, because $x_t = n_t/p_t k_t$ is then larger than without the policy. As a result, the value of capital is smaller relative to experts' net worth, so they can be exposed to less idiosyncratic risk. Of course, it makes the idiosyncratic risk-sharing problem worse after a good shock, because x_t is then lower than without the policy. So this policy will be attractive only if the marginal value of improving idiosyncratic risk sharing is higher after a bad shock than after a good shock. In the case with only TFP shocks, experts' exposure to idiosyncratic risk is the same after good and bad shocks in the unregulated competitive equilibrium, so distorting aggregate risk sharing is not attractive: what we gain in terms

of improved idiosyncratic risk sharing after a bad shock, we lose after a good shock. With uncertainty shocks, however, in equilibrium experts are exposed to more idiosyncratic risk after a bad uncertainty shock that raises ν_t and reduces x_t . Therefore, distorting $\sigma_{x,t}$ so that x_t is higher after a bad uncertainty shock compared to the unregulated competitive equilibrium (and lower after a good shock) is welfare improving. What we gain from improved idiosyncratic risk sharing after a bad uncertainty shock is larger than what we lose after a good shock.

V. Conclusions

In this paper I have shown how the type of aggregate shock hitting the economy can help explain the concentration of aggregate risk and drive balance sheet recessions. While we have a good understanding of why the balance sheets of more productive agents matter in an economy with financial frictions, we do not have a good explanation for why these agents are so exposed to aggregate risk. Even if agents face a moral hazard problem that limits their ability to issue equity, this does not prevent them from sharing aggregate risk, which can be accomplished by trading a simple market index. In fact, I show that in standard models of balance sheet recessions driven by Brownian TFP shocks, the balance sheet channel completely vanishes when agents are allowed to write contracts contingent on the aggregate state of the economy.

In contrast to TFP shocks, uncertainty shocks can create balance sheet recessions with depressed asset prices and investment, and financial losses disproportionately concentrated on the balance sheets of more productive agents. Even though agents can write complete contracts on all observable variables, experts choose to be highly exposed to aggregate risk in order to take advantage of stochastic investment opportunities. In addition, uncertainty shocks also affect financial markets, with higher aggregate volatility, lower risk-free interest rates, and higher risk premia. The model also has lessons for the design of financial regulation. I show how financial regulation may be welfare improving, even when agents are able to write privately optimal complete contracts.

These results suggest two avenues for future research. The first is to think about optimal financial regulation more carefully. While the unregulated competitive equilibrium is not constrained efficient, neither is eliminating the concentration of aggregate risk completely. So how much concentration of aggregate risk is “right,” and what are the appropriate instruments to regulate the economy? The second is to consider alternative aggregate shocks. While I have focused on uncertainty shocks, the same tools developed in this paper can be used to study other kinds of aggregate shocks.

Appendix A

A. Proof of Proposition 1

Without any financial frictions, idiosyncratic risk can be perfectly shared and has zero price in equilibrium. Capital then must be priced by arbitrage

$$g_{i,t} + \mu_{p,t} + \sigma \sigma'_{p,t} + \frac{a - u(g_{i,t})}{p_t} - r_t = \pi_t(\sigma + \sigma_{p,t}), \tag{A1}$$

and experts face the same portfolio problem as households, with the exception of the choice of the growth rate g , pinned down by the static FOC

$$u'(g_t) = p_t.$$

We have, in effect, a standard representative agent model with a stationary growth path with risk-free interest rate

$$r_t = \rho + \psi g_t - \frac{1}{2}(1 + \psi)\gamma\sigma^2$$

and price of aggregate risk

$$\pi_t = \gamma\sigma.$$

In a stationary equilibrium the price of capital is constant, so we have $\mu_{p,t} = \sigma_{p,t} = 0$, and replacing all of this in (A1) gives (5). QED

B. Proof of Proposition 2

From (15) we see that if $\sigma_\nu = 0$, then $\sigma_{x,t} = 0$. Furthermore, the idiosyncratic volatility of capital, ν_b , is then deterministic because it is the solution to an ordinary differential equation (ODE) (1). We can replace ν_t with t in the Markov equilibrium (and obtain a time-dependent equilibrium). The only possibly stochastic state variable is x_t , but we have seen that it can have only a stochastic drift. However, since all equilibrium objects are functions of x and time t , then by (9) we see that x_t is the deterministic solution to a time-dependent ODE. QED

C. A Simplified Environment

Consider the simplified environment in Section III. Once ν is realized, we have an economy without uncertainty shocks. We already know that with only TFP shocks, $\sigma_x = 0$ and therefore also $\sigma_\xi = \sigma_\zeta = \sigma_p = 0$, and $\mu_x > 0$ with $x_t \rightarrow 1$ (but never reaches $x = 1$). Value functions still take the form $(\xi n)^{1-\gamma}/(1-\gamma)$ and $(\zeta n)^{1-\gamma}/(1-\gamma)$. Experts' HJB equation is

$$\rho \log \xi = \max_{\hat{c}, \hat{g}, \hat{h}, \hat{\theta}} \rho \log \hat{c} + \mu_n - \hat{c} - \frac{\gamma}{2}\sigma_n^2 - \frac{\gamma}{2}\tilde{\sigma}_n^2 + \frac{\xi'_x}{\xi} \mu_x$$

subject to the budget constraint (3).²⁷ Households have an analogous one with $\hat{k} = 0$. The FOCs give us $\hat{e} = \hat{c} = \rho$ and $\sigma_n = \sigma_w = \pi/\gamma = \sigma$ (using the equilibrium conditions). From the market clearing condition for consumption goods we obtain that the price of capital is constant and equal to

$$p = \frac{a - \iota(g)}{\rho}$$

with $\iota'(g) = p$ (this is a result of assuming an EIS $\psi = 1$). Of course, we also have the pricing equation for capital, which simplifies to

$$\begin{aligned} \frac{a - \iota(g)}{p} + g - r &= \gamma\sigma^2 + \gamma\frac{(\phi\nu)^2}{x}, \\ r &= \rho + g - \gamma\sigma^2 - \gamma\frac{(\phi\nu)^2}{x}. \end{aligned}$$

If holding capital requires a large exposure to idiosyncratic risk, the risk-free rate must be low to keep the price of capital high enough for market clearing in the consumption goods market. As x grows, experts must leverage less to hold the capital stock, and therefore their exposure to idiosyncratic risk is smaller. As a result, the risk-free interest rate can go up. Notice that this is the same pattern we observe in the numerical solution of the full model in figure 2.

The law of motion of x is also simplified, because $\hat{e} = \hat{c} = \rho$ and $\sigma_x = 0$. We get

$$\mu_x = x(1-x)(\mu_n - \mu_w) = x(1-x)\gamma\left(\frac{\phi\nu}{x}\right)^2 > 0.$$

Plugging in these equilibrium conditions into experts' and households' HJB, we obtain an ODE for the ratio of investment opportunities $\Omega = \xi/\zeta$:

$$\rho \log \Omega = \frac{\gamma}{2} \left(\frac{\phi\nu}{x}\right)^2 + \partial_x(\log \Omega)x(1-x)\gamma\left(\frac{\phi\nu}{x}\right)^2. \quad (\text{A2})$$

Notice that for $x \rightarrow 1$ we get $\log \Omega(x) \rightarrow (\gamma/2)(\phi\nu)^2(1/\rho)$, which corresponds to an economy in which $x = 1$ always.²⁸ In the long run, experts' investment opportunities are relatively better, compared to households', when idiosyncratic risk ν is high, because capital pays a higher excess return, that is, $\alpha = \gamma(\phi\nu/x)$ is higher. In fact, this is true for any $x \in (0, 1)$. In addition, as x grows, $\alpha = \gamma(\phi\nu/x)$ falls, so experts' investment opportunities worsen relative to households'.

PROPOSITION 3. $\Omega(x; \nu)$ is strictly increasing in ν and strictly decreasing in x .

Proof. Take without loss of generality $\nu_2 > \nu_1$. First, we already know that at $x = 1$, $\log \Omega(1; \nu_2) > \log \Omega(1; \nu_1)$. We want to establish that this is true for any

²⁷ The Epstein-Zin aggregator in this case takes the form

$$f(c, U) = \rho(1 - \gamma)U \left(\log(c) - \frac{1}{1 - \gamma} \log[(1 - \gamma)U] \right).$$

²⁸ If $x = 1$ always, there are no households, but we can still compute the value function ζ of a single (measure zero) household in that environment.

$x \in (0, 1)$. Suppose toward contradiction that this is not true. Because $\log \Omega(x; \nu)$ are continuous, there must be some x^* where they intersect for the last time, that is, $x^* = \max\{x \in (0, 1) : \log \Omega(x; \nu_2) = \log \Omega(x; \nu_1)\}$, so that $\log \Omega(x; \nu_2) > \log \Omega(x; \nu_1)$ for all $x > x^*$. Using (A2) we compute

$$\begin{aligned} \partial_x[\log \Omega(x^*, \nu_2)] &= \frac{1}{2} \frac{1}{x^*(1-x^*)} \left\{ \log \Omega(x^*; \nu_2) \left[\frac{\gamma}{2} \left(\frac{\phi \nu_2}{x^*} \right)^2 \frac{1}{\rho} \right]^{-1} - 1 \right\} \\ &< \frac{1}{2} \frac{1}{x^*(1-x^*)} \left\{ \log \Omega(x^*; \nu_1) \left[\frac{\gamma}{2} \left(\frac{\phi \nu_1}{x^*} \right)^2 \frac{1}{\rho} \right]^{-1} - 1 \right\} = \partial_x[\log \Omega(x^*, \nu_1)], \end{aligned}$$

where the inequality uses $\Omega(x; \nu) > 1$. But this means that $\log \Omega(x^* + \epsilon; \nu_2) < \log \Omega(x^* + \epsilon; \nu_1)$. This is a contradiction and shows that $\Omega(x; \nu)$ is strictly increasing in ν .

To show that $\Omega(x; \nu)$ is strictly decreasing in x , notice that $\partial_x[\log \Omega(x, \nu)]$ is non-negative iff

$$\log \Omega(x; \nu) \geq h(x; \nu) \equiv \frac{\gamma}{2} \left(\frac{\phi \nu}{x} \right)^2 \frac{1}{\rho}$$

(and strictly positive if the inequality is strict). Since $h(x)$ is strictly decreasing in x , if $\partial_x[\log \Omega(x^*, \nu)] \geq 0$ for some $x^* \in (0, 1)$, then $\partial_x[\log \Omega(x, \nu)] > 0$ for all $x > x^*$, and therefore $\log \Omega(x; \nu) > h(x^*; \nu)$ for all $x > x^*$. This violates the boundary condition

$$\log \Omega(x; \nu) \rightarrow \frac{\gamma}{2} (\phi \nu)^2 \frac{1}{\rho} = h(1; \nu) < h(x^*; \nu)$$

as $x \rightarrow 1$. So we must have $\partial_x[\log \Omega(x, \nu)] < 0$ for all x . QED

Let us consider now the problem of experts and households at time $t = 0$ before $\nu \in \{\nu_h, \nu_l\}$ has realized. They start with net worth n_0 and wealth w_0 and can trade Arrow securities to share the risk over ν . Experts solve

$$\begin{aligned} \max_{n_h, n_l} \frac{1}{2} \frac{(\xi_h n_h)^{1-\gamma}}{1-\gamma} + \frac{1}{2} \frac{(\xi_l n_l)^{1-\gamma}}{1-\gamma} \\ \text{subject to } q_h n_h + q_l n_l = n_0, \end{aligned}$$

where q_h and q_l are the prices of the Arrow securities that pay one dollar in each state, and $n_h, n_l, \xi_h,$ and ξ_l are the net worth and values of ξ at time $t = 0$ in each state after idiosyncratic risk is realized. The FOCs yield

$$\left(\frac{n_h}{n_l} \right)^{-\gamma} = \frac{q_h}{q_l} \left(\frac{\xi_h}{\xi_l} \right)^{\gamma-1}.$$

Households have an analogous problem, with ζ instead of ξ and w instead of n . Notice that since the price of capital does not depend on ν , the total wealth

is constant: $n_l + w_l = n_h + w_h = n_0 + w_0$. It is easy to see that in equilibrium we must have

$$\frac{n_h/w_h}{n_l/w_l} = \left(\frac{\Omega_h}{\Omega_l}\right)^{(1-\gamma)/\gamma}. \quad (\text{A3})$$

For $\gamma > 1$, experts should have a relatively smaller share of aggregate wealth in the state in which their investment opportunities are relatively better, that is, where Ω is higher. Since $\Omega(x; \nu)$ is strictly increasing in ν , experts' share of aggregate wealth x should be lower when idiosyncratic risk is high, that is, $x_h < x_l$. In addition, $\Omega(x; \nu)$ is strictly decreasing in x , so this aggregate risk-sharing arrangement makes experts' investment opportunities even better in the state with high idiosyncratic risk, amplifying incentives to reduce experts' share of aggregate wealth.

PROPOSITION 4. For $\gamma > 1$, experts' share of aggregate wealth x falls if ν_h is realized and goes up if ν_l is realized, that is, $x_h < x_0 < x_l$.

Proof. First use $x = n/(n + w)$ to write (A3) as

$$\frac{\frac{x_h}{1-x_h}}{\frac{x_l}{1-x_l}} = \left(\frac{\Omega_h}{\Omega_l}\right)^{(1-\gamma)/\gamma}.$$

Taking logs,

$$\begin{aligned} \log \frac{x_h}{1-x_h} - \log \frac{x_l}{1-x_l} &= \frac{1-\gamma}{\gamma} (\log \Omega_h - \log \Omega_l), \\ \log \frac{x_h}{1-x_h} - \frac{1-\gamma}{\gamma} \log \Omega(x_h; \nu_h) &= \log \frac{x_l}{1-x_l} - \frac{1-\gamma}{\gamma} \log \Omega(x_l; \nu_l). \end{aligned} \quad (\text{A4})$$

Now let us compute the derivative of the left-hand side with respect to x :

$$\begin{aligned} \partial_x \left[\log \frac{x}{1-x} - \frac{1-\gamma}{\gamma} \log \Omega(x; \nu_h) \right] &= \frac{1-x}{x} \left[\frac{1}{1-x} + \frac{x}{(1-x)^2} \right] - \frac{1-\gamma}{\gamma} \partial_x \log \Omega(x; \nu_h), \\ \partial_x \left[\log \frac{x}{1-x} - \frac{1-\gamma}{\gamma} \log \Omega(x; \nu_h) \right] &= \frac{1}{x(1-x)} - \frac{1-\gamma}{\gamma} \partial_x \log \Omega(x; \nu_h) \\ &= \frac{1}{x(1-x)} \left\{ 1 - \frac{1-\gamma}{\gamma} \frac{1}{2} \log \Omega(x; \nu_h) \left[\frac{\gamma}{2} \left(\frac{\phi \nu_h}{x} \right)^2 \frac{1}{\rho} \right]^{-1} + \frac{11-\gamma}{2} \frac{1-\gamma}{\gamma} \right\} \\ &= \frac{1}{x(1-x)} \left\{ -\frac{1-\gamma}{\gamma} \frac{1}{2} \log \Omega(x; \nu_h) \left[\frac{\gamma}{2} \left(\frac{\phi \nu_h}{x} \right)^2 \frac{1}{\rho} \right]^{-1} + \frac{11+\gamma}{2} \frac{1+\gamma}{\gamma} \right\} > 0, \end{aligned}$$

where the last inequality uses $\Omega(x; \nu) > 1$. If we pick $x_h = x_l$, since $\Omega(x; \nu)$ is strictly increasing in ν , the left-hand side is larger than the right-hand side. It follows that $x_h < x_l$.

To see that x_0 is between x_h and x_l , use the fact that total wealth is constant to divide the budget constraints by $n + w$ and write $q_h x_h + q_l x_l = x_0$ and $q_h(1-x_h) + q_l(1-x_l) = 1-x_0$. Putting them together we get $q_h + q_l = 1$. Then $q_h x_h + q_l x_l = x_0$ implies they cannot both be greater or lower than x_0 , and this yields the desired result. QED

Appendix B

The strategy to solve for the equilibrium when uncertainty shocks hit the economy is to first use optimality and market clearing conditions to obtain expressions for equilibrium objects in terms of the stochastic processes for p , ξ , and ζ and then use Ito's lemma to map the problem into a system of PDEs. In order to obtain a nondegenerate stationary long-run distribution for x , I also introduce turnover among experts: they retire with independent Poisson arrival rate $\tau > 0$. When they retire they do not consume their wealth right away; they simply become households. Without turnover, experts want to postpone consumption and approach $x_t \rightarrow 1$ as $t \rightarrow \infty$. Turnover modifies experts' HJB slightly:

$$\begin{aligned} \frac{\rho}{1 - \psi} = \max_{i,g,h,\theta} & \frac{\hat{e}^{1-\psi}}{1 - \psi} \rho \xi_t^{\psi-1} + \frac{\tau}{1 - \gamma} \left[\left(\frac{\zeta}{\xi} \right)^{1-\gamma} - 1 \right] + \mu_n \\ & - \hat{e} + \mu_\xi - \frac{\gamma}{2} \left(\sigma_n^2 + \sigma_\xi^2 - 2 \frac{1 - \gamma}{\gamma} \sigma_n \sigma_\xi + \tilde{\sigma}_n^2 \right). \end{aligned} \tag{B1}$$

With Poisson intensity τ the expert retires and becomes a household, losing the continuation utility of an expert but gaining that of a household. For this reason, households' wealth multiplier ζ appears in experts' HJB equation. Households have the same HJB equation as before. The FOCs for consumption for experts and households are

$$\begin{aligned} \hat{e} &= \rho^{1/\psi} \xi^{(\psi-1)/\psi}, \\ \hat{c} &= \rho^{1/\psi} \zeta^{(\psi-1)/\psi}. \end{aligned}$$

So market clearing in the consumption goods market requires

$$\rho^{1/\psi} \left[\xi^{(\psi-1)/\psi} x + \zeta^{(\psi-1)/\psi} (1 - x) \right] = \frac{a - \iota}{p}. \tag{B2}$$

Equation (15) provides a formula for σ_x using

$$\frac{\Omega_p}{\Omega} = \frac{\xi_p}{\xi} - \frac{\zeta_p}{\zeta}$$

and

$$\frac{\Omega_x}{\Omega} = \frac{\xi_x}{\xi} - \frac{\zeta_x}{\zeta}.$$

$$\sigma_x = \frac{(1 - x)x \frac{1 - \gamma}{\gamma} \left(\frac{\xi_p}{\xi} - \frac{\zeta_p}{\zeta} \right)}{1 - (1 - x)x \frac{1 - \gamma}{\gamma} \left(\frac{\xi_x}{\xi} - \frac{\zeta_x}{\zeta} \right)} \sigma_p \sqrt{v_t}.$$

We can use Ito's lemma to obtain expressions for

$$\begin{aligned}\sigma_p &= \frac{h_p}{p} \sigma_v \sqrt{v_t} + \frac{h_x}{p} \sigma_x, \\ \sigma_\xi &= \frac{\xi_v}{\xi} \sigma_v \sqrt{v_t} + \frac{\xi_x}{\xi} \sigma_x, \\ \sigma_\zeta &= \frac{\zeta_v}{\zeta} \sigma_v \sqrt{v_t} + \frac{\zeta_x}{\zeta} \sigma_x,\end{aligned}$$

and the definition of σ_x from (9) to obtain an expression for

$$\sigma_n = \sigma + \sigma_p + \frac{\sigma_x}{x}.$$

Then we use experts' FOC for aggregate risk sharing (11) to obtain an expression for the price of aggregate risk

$$\pi = \gamma \sigma_n + (\gamma - 1) \sigma_\xi.$$

Households' exposure to aggregate risk is taken from (12):

$$\sigma_w = \frac{\pi}{\gamma} - \frac{\gamma - 1}{\gamma} \sigma_\xi.$$

Experts' exposure to idiosyncratic risk is given by $\tilde{\sigma}_n = (\phi/x)v$. We can now use households' budget constraint to obtain the drift of their wealth (before consumption)

$$\mu_w = r + \pi \sigma_w,$$

and plugging into their HJB equation, we obtain an expression for the risk-free interest rate

$$r = \frac{\rho}{1 - \psi} - \frac{\psi}{1 - \psi} \rho^{1/\psi} \zeta^{(\psi-1)/\psi} - \pi \sigma_w - \mu_\zeta + \frac{\gamma}{2} \left(\sigma_w^2 + \sigma_\zeta^2 - 2 \frac{1 - \gamma}{\gamma} \sigma_w \sigma_\zeta \right),$$

where the only term that has not been solved for yet is μ_ζ . We use the FOC for capital (10) and the expression for the risk-free interest rate and plug into the formula for μ_n from equation (3) to get

$$\mu_n = r + \underbrace{\gamma \frac{1}{x^2} (\phi v)^2}_{\alpha, \beta h(\phi v)} + \pi \sigma_n.$$

In equilibrium experts receive the risk-free interest on their net worth, plus a premium for the idiosyncratic risk they carry through capital, $\gamma(1/x^2)(\phi v)^2$, and a risk premium for the aggregate risk they carry, $\pi \sigma_n$. This allows us to com-

pute the drift of the endogenous state variable x in terms of known objects, from (9) (appropriately modified for turnover) and (10):

$$\mu_x = x_i \left[\mu_n - \hat{e} - \tau + \frac{a - t}{p} - r - \pi(\sigma + \sigma_p) - \frac{\gamma}{x} (\phi\nu)^2 + (\sigma + \sigma_p)^2 - \sigma_n(\sigma + \sigma_p) \right].$$

Turnover works to reduce the fraction of aggregate wealth that belongs to experts through the term $-\tau$. Using Ito’s lemma we get expressions for the drift of p , ξ , and ζ :

$$\begin{aligned} \mu_p &= \frac{p_r}{p} \lambda(\bar{\nu} - \nu) + \frac{p_x}{p} \mu_x + \frac{1}{2} \left(\frac{p_{vv}}{p} \sigma_v^2 \nu + 2 \frac{p_{vx}}{p} \sigma_v \sqrt{\nu} \sigma_x + \frac{p_{xx}}{p} \sigma_x^2 \right), \\ \mu_\xi &= \frac{\xi_r}{\xi} \lambda(\bar{\nu} - \nu) + \frac{\xi_x}{\xi} \mu_x + \frac{1}{2} \left(\frac{\xi_{vv}}{\xi} \sigma_v^2 \nu + 2 \frac{\xi_{vx}}{\xi} \sigma_v \sqrt{\nu} \sigma_x + \frac{\xi_{xx}}{\xi} \sigma_x^2 \right), \\ \mu_\zeta &= \frac{\zeta_r}{\zeta} \lambda(\bar{\nu} - \nu) + \frac{\zeta_x}{\zeta} \mu_x + \frac{1}{2} \left(\frac{\zeta_{vv}}{\zeta} \sigma_v^2 \nu + 2 \frac{\zeta_{vx}}{\zeta} \sigma_v \sqrt{\nu} \sigma_x + \frac{\zeta_{xx}}{\zeta} \sigma_x^2 \right). \end{aligned}$$

Finally, experts’ HJB (B1) and their FOC for capital (10) provide two second-order PDEs in p , ξ , and ζ . Together with the market clearing condition for consumption (B2), they characterize the Markov equilibrium.

Since we assume $2\lambda\bar{\nu} \geq \sigma_v^2$, we know that $\nu_t \in (0, \infty)$. From the way we constructed the system of PDEs, the market clearing conditions, the law of motion of x , and the FOCs are satisfied, so we just need to make sure that agents’ plans really are optimal. Check that (1) p is C^2 and strictly positive, and the process x generated by μ_x and σ_x has $x_t \in (0, 1)$ always; (2) ξ and ζ are C^2 and strictly positive, and $\xi^{1-\gamma}$ and $\zeta^{1-\gamma}$ are bounded above; and (3) the resulting policy functions \hat{e} , \hat{g} , \hat{k} , $\hat{\theta}$ and \hat{c} , σ_w generate plans that actually deliver utility $(\xi n)^{1-\gamma}/(1-\gamma)$ and $(\zeta n)^{1-\gamma}/(1-\gamma)$ for experts and households, respectively. Indeed, strictly positive p and $x_t \in (0, 1)$ make sure experts’ and households’ wealth are always strictly positive, and therefore, $c_{i,t} = \hat{c}_i w_{i,t} > 0$ and $e_{i,t} = \hat{e}_i n_{i,t} > 0$ (using ξ and ζ strictly positive) and $k_{i,t} = \hat{k}_i n_{i,t} > 0$. Conditions 2 and 3 ensure that $(\xi n)^{1-\gamma}/(1-\gamma)$ and $(\zeta n)^{1-\gamma}/(1-\gamma)$ really are experts’ and households’ value functions, respectively, and the resulting plans optimal. These conditions can be checked numerically once we obtain a solution to the system of PDEs. In particular, since we are dealing with Markovian consumption streams, we can use the setup in Duffie and Lions (1992) to check condition 3. It is worth pointing out, however, that if there are multiple Markov equilibria, then it is in principle possible that there is more than one solution to the system of PDEs satisfying conditions 1–3, each one corresponding to one equilibrium. However, although I do not provide a uniqueness theorem, numerically the equilibrium seems to be unique.

Numerical algorithm.—The system of PDEs can be solved in several ways. I use a finite difference scheme with a false transient. The idea is that instead of solving the infinite-horizon PDEs directly, we add a time dimension and solve the system as if there was a finite horizon T . Now we must look for p , ξ , and ζ as functions of

(v, x, t) . This requires modifying the HJB equations and FOC for capital by adding a time derivative when computing the drifts using Ito's lemma. We also need to come up with terminal values for ξ , ζ , and p at time T and then solve backward in time. The important insight, however, is that regardless of the terminal values chosen, if somehow we find a stationary point of this system such that the time derivatives $\xi'_i(v, x, t)$, $\zeta'_i(v, x, t)$, and $p'_i(v, x, t)$ vanish, we have found a solution for the system of PDEs characterizing the infinite-horizon problem we are interested in. Because the market clearing condition for consumption is an algebraic constraint (does not involve derivatives), it is easier to totally differentiate it with respect to time and obtain a differential equation

$$\partial_t \{ \rho^{1/\psi} [\xi(v, x, t)^{(\psi-1)/\psi} x + \zeta(v, x, t)^{(\psi-1)/\psi} (1-x)] p(v, x, t) - [a - u(t^{-1}(p(v, x, t)))] \} = 0.$$

Terminal values at $t = T$ are not particularly important as long as they satisfy the market clearing condition. As we solve backward in time the market clearing condition will be preserved, and we can check again at the end. Terminal values play a role analogous to the initial guess when solving nonlinear equations.

Start with a grid for v and x , and pick terminal values for p , ξ , and ζ at each point in the grid (i.e., pick a terminal $p(v, x, T)$, etc.). We can use any finite difference scheme (or a collocation method) to compute the first and second derivatives of ξ , ζ , and p with respect to v and x . Then we can solve the system of PDEs (including the differential version of the market clearing condition for consumption) for the time derivatives $\xi'_i(v, x, t)$, $\zeta'_i(v, x, t)$, and $p'_i(v, x, t)$ at each point in the (v, x) grid. We can use these to take a "step back in time" and update the value of ξ , ζ , and p for each point in the grid. Notice that what we have is a system of first-order differential equations in time, which we can solve backward using any standard integrator such as Runge-Kutta 4, for example.

As we move backward in time from T , we are letting the finite horizon go to infinity, so it is a good guess that the solution should approach the solution for the time-homogeneous system that characterizes the infinite-horizon equilibrium. This suggests that the algorithm will converge to the desired solution for a wide variety of terminal conditions. However, we do not need to guess: we can verify this *ex post*. Solve the system backward in time until we find a stationary point such that the time derivatives vanish. Then we have found a solution to the system of PDEs that we were interested in. We can in fact forget about how we found the functions $\xi(v, x)$, $\zeta(v, x)$, and $p(v, x)$ and check that we satisfy the two original PDEs and the algebraic constraint. We can then check conditions 1–3 to make sure we indeed have a Markov equilibrium.

Calibration and simulation.—Most of the calibration has already been explained. For the stochastic process for idiosyncratic risk v I use data from Campbell et al. (2001) about the monthly standard deviation of stock returns msd_t , and I map it to $v_t = \sqrt{12} \times \text{msd}_t$ (here I use a monthly discretization).²⁹ I then use the follow-

²⁹ I use the series `disg2nt.tab.tsv`, which can be downloaded from <http://scholar.harvard.edu/campbell/data>.

ing moments from the stationary distribution of a square root process. The mean is just $\mathbb{E}[\nu_t] = \bar{\nu} = 25$ percent. The covariance between ν_{t+h} and ν_t is

$$\begin{aligned} \text{cov}(\nu_{t+h}, \nu_t) &= \text{var}(\nu_t) \exp(-\lambda h) \\ \Rightarrow \lambda &= -\frac{\ln \left[\frac{\text{cov}(\nu_{t+h}, \nu_t)}{\text{var}(\nu_t)} \right]}{h}. \end{aligned}$$

The data for ν_t do not fit a square root process perfectly. Depending on the covariance lag h I target, I get a different λ ranging from 0.76 for $h = 1$ (1 year) implying a half-life of around 11 months to 3.44 for $h = 1/12$ (1 month) implying a half-life of 2.5 months. I settle for $\lambda = 1.38$, which implies a half-life of half a year. We are dealing with pretty short-lived shocks, consistent with Bloom (2009), for example. Finally, I use $\text{var}(\nu_t) = \sigma_v^2(\bar{\nu}/2\lambda)$ to solve for $\sigma_v = 0.17$.

I simulate the model for 100,000 years, starting from the “steady state.” For this purpose the model is discretized with a time step $dt = 1/360$, so each time step is a day, and a unit of time is a year. The aggregate shock dZ is approximated with a binomial distribution with a size \sqrt{dt} , as usual. For flow variables such as GDP and investment, I split them into quarters and integrate to obtain quarterly GDP and investment, for example, $y_q = \int_{q \times 1/4}^{(q+1) \times 1/4} y_t dt$, where $q = 1, 2, \dots$ represents the quarter. For stock variables such as the value of assets p_k , I use the value at the beginning of each quarter. I then compute the quarterly growth rate and annualize it, for example, $g_y = [\ln(y_{q+1}) - \ln(y_q)]400$ in percentage terms and compute the average and standard deviation. Notice that the time aggregation explains why $\sigma = 1.25$ percent translates into a 2 percent annualized volatility of growth rate for quarterly GDP instead of 2.5 percent, which we would get if we used beginning of quarter measurements for y . For the investment to GDP ratio I just compute i_q/y_q for each quarter. For financial variables I just take the average of r_t, π_t , and so forth as well as for m_t and f_t .

To simulate the large uncertainty shock, I shut down the terms of order dt and I hit the economy with a sequence of bad aggregate shocks: set $dZ_t = -\sqrt{dt}$ until ν_t reaches the target. This means we are considering a large negative shock that comes very fast, but I am integrating the variables along the path, for example, $x_{i+1} = -\sigma_x(\nu_i, x_i)\sqrt{dt}$.

With financial regulation.—For the regulated equilibrium in which the planner controls experts’ exposure to aggregate risk, we just drop the FOC for θ_t . Notice that the FOCs for consumption, investment, and capital holdings are unchanged. We now have $\sigma_n/\sigma_w = \kappa$. Using the market clearing condition $\sigma_n x + \sigma_w(1 - x) = \sigma + \sigma_p$, we obtain $\sigma_n = (\sigma + \sigma_p)\kappa/[1 + (\kappa - 1)x]$ and $\sigma_w = (\sigma + \sigma_p)/[1 + (\kappa - 1)x]$. As a result, using the definition $\sigma_x = x(\sigma_n - \sigma - \sigma_p)$ and Ito’s lemma for $\sigma_p = (p_v/p)\sigma_v\sqrt{\nu} + (p_x/p)\sigma_x$, we replace the formula for σ_x by

$$\sigma_x = x(1 - x) \frac{\left(\sigma + \frac{p_v}{p} \sigma_v \sqrt{\nu} \right) \frac{\kappa - 1}{1 + (\kappa - 1)x}}{1 - x(1 - x) \frac{p_x}{p} \frac{\kappa - 1}{1 + (\kappa - 1)x}}.$$

Since households are not regulated, their FOC for σ_w still pins down the price of aggregate risk $\pi = \gamma\sigma_w - (1 - \gamma)\sigma_{\tilde{r}}$. Everything else is the same as in the case of the unregulated competitive equilibrium and can be solved in the same way.

For the optimal policy, we just solve the regulated competitive equilibrium for different values of κ and find the optimal one. Each time the initial redistribution from x_0 to x_1 that keeps households indifferent is found numerically using households' value functions under the unregulated and regulated competitive equilibrium.

References

- Adrian, T., and N. Boyarchenko. 2012. "Intermediary Leverage Cycles and Financial Stability." Staff Report no. 567, Fed. Reserve Bank New York.
- Adrian, T., E. Etula, and T. Muir. 2014. "Financial Intermediaries and the Cross-Section of Asset Returns." *J. Finance* 69 (6): 2557–96.
- Asriyan, V. 2015. "Balance Sheet Recessions with Informational and Trading Frictions." Manuscript no. 2565603, Soc. Sci. Res. Network.
- Bachmann, R., and G. Moscarini. 2011. "Business Cycles and Endogenous Uncertainty." Paper no. 36 presented at the annual meeting of Soc. Econ. Dynamics, Ghent.
- Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron. 2014. "Volatility, the Macroeconomy, and Asset Prices." *J. Finance* 69 (6): 2471–2511.
- Bansal, R., and A. Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *J. Finance* 59 (4): 1481–1509.
- Beeler, J., and J. Y. Campbell. 2012. "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment." *Critical Finance Rev.* 1 (January): 141–82.
- Bernanke, B., and M. Gertler. 1989. "Agency Costs, Net Worth, and Business Fluctuations." *A.E.R.* 79 (1): 14–31.
- Bernanke, B., M. Gertler, and S. Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics*, vol. 1, edited by J. Taylor and M. Woodford, 1341–93. Amsterdam: Elsevier.
- Bloom, N. 2009. "The Impact of Uncertainty Shocks." *Econometrica* 77 (3): 623–85.
- Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. Terry. 2012. "Really Uncertain Business Cycles." Working Paper no. 18245, NBER, Cambridge, MA.
- Brunnermeier, M., and Y. Sannikov. 2014. "A Macroeconomic Model with a Financial Sector." *A.E.R.* 104 (2): 379–421.
- Campbell, J., and J. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *J.P.E.* 107 (April): 205–51.
- Campbell, J., S. Giglio, C. Polk, and R. Turley. 2012. "An Intertemporal CAPM with Stochastic Volatility." Working paper, NBER, Cambridge, MA.
- Campbell, J., M. Lettau, B. G. Malkiel, and Y. Xu. 2001. "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk." *J. Finance* 56 (1): 1–43.
- Christiano, L., R. Motto, and M. Rostagno. 2014. "Risk Shocks." *A.E.R.* 104 (1): 27–65.
- Cooley, T., V. Quadrini, and R. Marimon. 2004. "Aggregate Consequences of Limited Contract Enforceability." *J.P.E.* 112 (4): 817–47.
- Dávila, E., and A. Korinek. 2016. "Pecuniary Externalities in Economies with Financial Frictions." Working Paper no. w22444, NBER, Cambridge, MA.
- DeMarzo, P., and M. Fishman. 2007. "Optimal Long Term Financial Contracting." *Rev. Financial Studies* 20:2079–2127.

- DeMarzo, P., and Y. Sannikov. 2006. "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model." *J. Finance* 61 (6): 2681–2724.
- Di Tella, S. 2014. "Optimal Financial Regulation and the Concentration of Aggregate Risk." Working paper, Stanford Univ.
- Duffie, D., and L. Epstein. 1992. "Stochastic Differential Utility." *Econometrica* 60 (2): 353–94.
- Duffie, D., and P. Lions. 1992. "PDE Solutions of Stochastic Differential Utility." *J. Math. Econ.* 21:577–606.
- Fernández-Villaverde, J., and J. Rubio-Ramírez. 2010. "Macroeconomics and Volatility: Data, Models, and Estimation." Working Paper no. 16618, NBER, Cambridge, MA.
- Fernández-Villaverde, J., J. Rubio-Ramírez, A. Guerron-Quintana, and M. Uribe. 2011. "Risk Matters: The Real Effects of Volatility Shocks." *A.E.R.* 101 (6): 2530–61.
- Gabaix, X., A. Krishnamurthy, and O. Vigneron. 2007. "Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market." *J. Finance* 62 (2): 557–95.
- Geanakoplos, J. 2009. "The Leverage Cycle." In *NBER Macroeconomics Annual*, vol. 24, edited by D. Acemoglu, K. Rogoff, and M. Woodford, 1–65. Chicago: Univ. Chicago Press.
- Gertler, M., N. Kiyotaki, and A. Queralto. 2012. "Financial Crises, Bank Risk Exposure and Government Financial Policy." *J. Monetary Econ.* 59 (suppl.): S17–S34.
- Gorton, G., and G. Ordoñez. 2014. "Collateral Crises." *A.E.R.* 104 (2): 343–78.
- Gruber, J. 2013. "A Tax-Based Estimate of the Elasticity of Intertemporal Substitution." *Q. J. Finance* 3 (1).
- Hall, R. E. 1988. "Intertemporal Substitution in Consumption." *J.P.E.* 96 (2): 339–57.
- He, Z., and A. Krishnamurthy. 2012. "A Model of Capital and Crises." *Rev. Econ. Studies* 79 (2): 735–77.
- . 2014. "A Macroeconomic Framework for Quantifying Systemic Risk." Working Paper no. w19885, NBER, Cambridge, MA.
- Herskovic, B., B. Kelly, H. Lustig, and S. Van Nieuwerburgh. 2016. "The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications." *J. Financial Econ.* 119 (2): 249–83.
- Holmstrom, B., and J. Tirole. 1996. "Modeling Aggregate Liquidity." *A.E.R.* 86 (2): 187–91.
- Kiyotaki, N., and J. Moore. 1997. "Credit Cycles." *J.P.E.* 105:211–48.
- Krishnamurthy, A. 2003. "Collateral Constraints and the Amplification Mechanism." *J. Econ. Theory* 111 (2): 277–92.
- Lorenzoni, G. 2008. "Inefficient Credit Booms." *Rev. Econ. Studies* 75 (3): 809–33.
- Mulligan, C. 2002. "Capital, Interest, and Aggregate Intertemporal Substitution." Working Paper no. 9373 (December), NBER, Cambridge, MA.
- Rampini, A., and S. Viswanathan. 2010. "Collateral, Risk Management, and the Distribution of Debt Capacity." *J. Finance* 65:2293–2322.
- . 2013. "Collateral and Capital Structure." *J. Financial Econ.* 109:466–92.
- Sannikov, Y. 2008. "A Continuous-Time Version of the Principal-Agent Problem." *Rev. Econ. Studies* 75 (3): 957–84.
- Simsek, A. 2013. "Speculation and Risk Sharing with New Financial Assets." *Q.J.E.* 128 (3): 1365–96.
- Sraer, D., T. Chaney, and D. Thesmar. 2012. "The Collateral Channel." *A.E.R.* 102 (6): 2381–2409.
- Vissing-Jørgensen, A. 2002. "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution." *J.P.E.* 110 (4): 825–53.