For Online Publication

Online Appendix for

"A Static Capital Buffer is Hard To Beat"

A The Bank's Problem

A.1 Baseline: First-Order Conditions

Substituting $d_t = l_t - e_t$ into equation (??) and writing $\mathrm{d}G(\varepsilon_{t+1})$ explicitly turn the objective into:

$$\max_{l_t,e_t,\sigma_t} E_t \left\{ \psi_{t,t+1} \left[\int_{\varepsilon^*_{t+1}}^{\infty} \left(\left(R^s_{t+1} + \sigma_t \frac{\varepsilon_{t+1}}{Q_t} \right) l_t - R^d_t \left(l_t - e_t \right) \right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \mathrm{d}\varepsilon_{t+1} \right] - e_t \right\},$$

subject to

$$\begin{aligned} e_t &\geq \gamma_t l_t, \\ l_t &\geq 0, \\ \underline{\sigma} \leq \sigma_t &\leq \bar{\sigma}, \end{aligned}$$

where $\psi_{t,t+1} = \beta \frac{\lambda_{ct+1}}{\lambda_{ct}}$ is the stochastic discount factor and $\varepsilon^*_{t+1} = \left(\frac{R^d_t - R^s_{t+1}}{\sigma_t} - \frac{R^d_t e_t}{\sigma_t l_t} \right) Q_t$ is the shield of limited liability. Note that we expressed ε^*_{t+1} from $\left(R^s_{t+1} + \sigma_t \frac{\varepsilon^*_{t+1}}{Q_t} \right) l_t - R^d_t \left(l_t - e_t \right) = 0$ to get the lower limit of the integral.

Append the Lagrangian multiplier χ_{1t} to the constraint $e_t \geq \gamma l_t$ and χ_{2t} to the constraint $l_t \geq 0$. Conditional on the optimal choice of σ_t, the first-order conditions are:

$$\frac{\partial \mathcal{L}}{\partial l_t} = E_t \left[\psi_{t,t+1} \overbrace{\left(\left(R^s_{t+1} + \sigma_t \left(\frac{R^d_t \left(l_t - e_t \right)}{\sigma_t l_t} - \frac{R^s_{t+1}}{\sigma_t} \right) \right) l_t - R^d_t \left(l_t - e_t \right) \right)}^{=0} \cdot \frac{\partial \varepsilon^*_{t+1}}{\partial l_t} \right] + \chi_{2t} +$$

$$E_t \left[\int_{\varepsilon^*_{t+1}}^{\infty} \psi_{t,t+1} \frac{\partial}{\partial l_t} \left(\left(R^s_{t+1} + \sigma_t \frac{\varepsilon_{t+1}}{Q_t} \right) l_t - R^d_t \left(l_t - e_t \right) \right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \mathrm{d}\varepsilon_{t+1} \right] - \gamma \chi_{1t} = 0,$$

$$\frac{\partial \mathcal{L}}{\partial e_t} = -E_t \left[\psi_{t,t+1} \overbrace{\left(\left(R^s_{t+1} + \sigma_t \left(\frac{R^d_t \left(l_t - e_t \right)}{\sigma_t l_t} - \frac{R^s_{t+1}}{\sigma_t} \right) \right) l_t - R^d_t \left(l_t - e_t \right) \right)}^{=0} \cdot \frac{\partial \varepsilon^*_{t+1}}{\partial e_t} \right] + \chi_{1t} +$$

$$E_t \left[\int_{\varepsilon^*_{t+1}}^{\infty} \psi_{t,t+1} \frac{\partial}{\partial e_t} \left(\left(R^s_{t+1} + \sigma_t \frac{\varepsilon_{t+1}}{Q_t} \right) l_t - R^d_t \left(l_t - e_t \right) \right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \mathrm{d}\varepsilon_{t+1} \right] - 1 = 0,$$

$$\chi_{1t}\left(e_t - \gamma_t l_t\right) = 0,$$
$$\chi_{2t} l_t = 0,$$
$$e_t - \gamma_t l_t \geq 0,$$
$$l_t \geq 0,$$
$$\chi_{1t} \geq 0,$$
$$\chi_{2t} \geq 0,$$

We are using the Leibniz integral rule above to find the partial derivatives of the profit function. Note that the first term is zero in the differentiation because the upper limit of the integral does not depend on any of the choice variables.

Next, express the integrals in the first-order conditions above using the erf function, wherever possible. Note that in order to make the next expressions more neat we omit the stochastic discount factor and the expectation operator from consideration. We include them in the final exposition.

Work on $\frac{\partial}{\partial l_t}$:

$$\int\limits_{\left(\frac{R_t^d - R_{t+1}^s}{\sigma_t} - \frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} \frac{\partial}{\partial l_t}\left(\left(R_{t+1}^s + \sigma_t \frac{\varepsilon_{t+1}}{Q_t}\right) l_t - R_t^d\left(l_t - e_t\right)\right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \,\mathrm{d}\varepsilon_{t+1} =$$

$$\int\limits_{\left(\frac{R_t^d - R_{t+1}^s}{\sigma_t} - \frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} \left(R_{t+1}^s + \sigma_t \frac{\varepsilon_{t+1}}{Q_t} - R_t^d\right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \,\mathrm{d}\varepsilon_{t+1} =$$

$$\frac{\sigma_t}{Q_t} \int\limits_{\left(\frac{R_t^d - R_{t+1}^s}{\sigma_t} - \frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} \varepsilon_{t+1} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \,\mathrm{d}\varepsilon_{t+1} +$$

$$\left(R_{t+1}^s - R_t^d\right) \int\limits_{\left(\frac{R_t^d - R_{t+1}^s}{\sigma_t} - \frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \,\mathrm{d}\varepsilon_{t+1}.$$

Break the calculation of the integral into two parts.

$$\int\limits_{\left(\frac{R_t^d - R_{t+1}^s}{\sigma_t} - \frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} \varepsilon_{t+1} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \,\mathrm{d}\varepsilon_{t+1} =$$

Introduce a change in variables to recast the integral in terms of the Standard Normal distribution. Use $v = \frac{\varepsilon_{t+1}+\xi}{\sqrt{2}\tau}$, or equivalently $\varepsilon_{t+1} = v\sqrt{2}\tau - \xi$, and remember that for the change $x = \varphi(t)$, the integral$\int_{\varphi(a)}^{\varphi(b)} f(x)dx$ becomes $\int_a^b f(\varphi(t))\varphi'(t)dt$. Here we use that $dv = \frac{d\varepsilon_{t+1}}{\sqrt{2}\tau}$, so we need to multiply dv by $\sqrt{2}\tau$ to express $d\varepsilon_{t+1}$ in terms of dv. Moreover, we need to transform the lower limit using v. So we need to add ξ to the lower limit of the integral and divide the result by $\sqrt{2}\tau$.

$$\int_{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} \left(v\sqrt{2}\tau - \xi\right)\frac{\sqrt{2}\tau}{\sqrt{2\pi\tau^2}}e^{-v^2}\,\mathrm{d}v =$$

$$\frac{\sqrt{2}\tau}{\sqrt{\pi}}\int_{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} ve^{-v^2}\,\mathrm{d}v - \frac{\xi}{\sqrt{\pi}}\int_{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} e^{-v^2}\,\mathrm{d}v =$$

$$-\frac{\sqrt{2}\tau}{2\sqrt{\pi}}e^{-v^2}\Bigg|_{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} - \frac{\xi}{\sqrt{\pi}}\left[\int_0^{\infty} e^{-v^2}\,\mathrm{d}v - \int_0^{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}} e^{-v^2}\,\mathrm{d}v\right] =$$

$$0 + l_t\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2} -$$

$$\frac{\xi}{\sqrt{\pi}}\left[\frac{\sqrt{\pi}}{2}\mathrm{erf}(\infty) - \frac{\sqrt{\pi}}{2}\mathrm{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right] =$$

$$\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2} - \frac{\xi}{2}\left[1-\mathrm{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right],$$

where we used that $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-v^2}$.

Let's express $\int_{\left(\frac{R_t^d-R_{t+1}^s}{\sigma_t}-\frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty}\left(\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\right)\mathrm{d}\varepsilon_{t+1}$ in terms of the error function.

Again, use the transformation $v = \frac{\varepsilon_{t+1}+\xi}{\sqrt{2}\tau}$ or $\varepsilon_{t+1} = v\sqrt{2}\tau - \xi$

$$\int_{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} \frac{\sqrt{2}\tau}{\sqrt{2\pi\tau^2}} e^{-v^2}\,\mathrm{d}v = \frac{1}{\sqrt{\pi}} \int_{\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}}^{\infty} e^{-v^2}\,\mathrm{d}v =$$

$$\frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right).$$

Therefore,

$$E_t\left[\int_{\left(\frac{R_t^d-R_{t+1}^s}{\sigma_t}-\frac{R_t^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} \frac{\partial}{\partial l_t}\left(\left(R_{t+1}^s+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t - R_t^d\left(l_t-e_t\right)\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right] =$$

$$E_t\left[\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2} - \frac{\sigma_t\xi}{2Q_t}\left[1-\operatorname{erf}\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right]\right] +$$

$$E_t\left[\left(R_{t+1}^s - R_t^d\right)\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right)\right] =$$

$$E_t\left[\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2} + \right.$$

$$\left.\left(\frac{R_{t+1}^s-\frac{\sigma_t\xi}{Q_t}-R_t^d}{2}\right)\left[1-\operatorname{erf}\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right]\right].$$

Similarly, work on $\frac{\partial}{\partial e_t}$

$$\int_{\left(\frac{R_t^d-R_{t+1}^s}{\sigma_t}-\frac{R_{t+1}^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} \frac{\partial}{\partial e_t}\left(\left(R_{t+1}^s+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t - R_t^d\left(l_t-e_t\right)\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1} =$$

$$\int_{\left(\frac{R_t^d-R_{t+1}^s}{\sigma_t}-\frac{R_{t+1}^d e_t}{\sigma_t l_t}\right)Q_t}^{\infty} R_t^d\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1} = R_t^d\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{R_t^d\left(l_t-e_t\right)-R_{t+1}^l l_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right).$$

In sum, the FOCs can be written as follows:

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R_t^d\left(1-\frac{e_t}{l_t}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2}+\right.\right.$$

$$\left.\left.\left(\frac{R_{t+1}^s-\frac{\sigma_t\xi}{Q_t}-R_t^d}{2}\right)\left[1-\operatorname{erf}\left(\frac{\left(R_t^d\left(1-\frac{e_t}{l_t}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right]\right\}+\chi_{2t}=\gamma\chi_{1t},$$

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\left(R_t^d\left(1-\frac{e_t}{l_t}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right)\right]\right\}-1+\chi_{1t}=0.$$

There are complementary slackness conditions which can be described by:

$$\begin{aligned}(e_t-\gamma l_t)\,\chi_{1t} &= 0,\\ l_t\chi_{2t} &= 0.\end{aligned}$$

A.2 Proof of Proposition ??

Equations (??) and (??) can be expressed as

$$\beta E_t\frac{\lambda_{ct+1}}{\lambda_{ct}}R_{t+1}^{e,i}=1-\frac{\zeta_t^i}{\lambda_{ct}},$$

where $i\in\{s,r\}$ denotes the type of equity. Using the expression, substitute for 1 in the bank's FOC with respect to e_t. Therefore,

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\left(R_t^d\left(1-\frac{e_t^i}{l_t^i}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t^i}{\sigma_t^i\sqrt{2}\tau}\right)\right)\right]-R_{t+1}^{e,i}\right\}-\frac{\zeta_t^i}{\lambda_{ct}}+\chi_{1t}^i=0.$$

Since the range of the erf function is between -1 and 1, i.e.$-1\le \operatorname{erf}(x)\le 1$, we know that the following expression is between Ψ_1^* and Ψ_2^*:

$$\Psi_1^*\le E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\left(R_t^d\left(1-\frac{e_t^i}{l_t^i}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t^i}{\sigma_t^i\sqrt{2}\tau}\right)\right)-R_{t+1}^{e,i}\right]\right\}\le\Psi_2^*,$$

where

$$\begin{aligned}\Psi_1^* &= E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[0-R_{t+1}^{e,i}\right]\right\},\\ \Psi_2^* &= E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d-R_{t+1}^{e,i}\right]\right\}.\end{aligned}$$

$$\frac{\partial}{\partial D_t} = \varsigma_0 D_t^{-\varsigma_d} - \lambda_{ct} + E_t \beta \lambda_{ct+1} R_t^d = 0,$$

Use $E_t \beta \lambda_{ct+1} R_{t+1}^{e,i} + \zeta_t^i = \lambda_{ct}$ (that comes from the household's FOCs with respect to e_t^i for each $i \in \{s, r\}$) to substitute for λ_{ct} in equation (**??**) . We get:

$$E_t \left\{ \beta \lambda_{ct+1} \left[R_t^d - R_{t+1}^{e,i} \right] \right\} = -\varsigma_0 D_t^{-\varsigma_d} + \zeta_t^i.$$

Note that $\varsigma_0 D_t^{-\varsigma_d} > 0$ under the usual (and mild) assumptions on the preferences for liquidity. Moreover, the Lagrangian multiplier on the households budget constraint, λ_{ct}, is positive. It reflects the fact that the budget constraint always binds given the standard assumptions on the preferences (Inada conditions). The latest expression is transformed into the following after dividing it by λ_{ct}:

$$\underbrace{E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[R_t^d - R_{t+1}^{e,i} \right] \right\}}_{=\Psi_2^*} - \frac{\zeta_t^i}{\lambda_{ct}} = -\frac{\varsigma_0 D_t^{-\varsigma_d}}{\lambda_{ct}} < 0.$$

Thus, $\Psi_2^* < \frac{\zeta_t^i}{\lambda_{ct}}$. Therefore,

$$E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[R_t^d \frac{1}{2} \left(1 - \text{erf} \left(\frac{\left(R_t^d \left(1 - \frac{e_t^i}{l_t^i} \right) - R_{t+1}^s \right) Q_t + \xi \sigma_t^i}{\sigma_t^i \sqrt{2} \tau} \right) \right) \right] - R_{t+1}^{e,i} \right\} - \frac{\zeta_t^i}{\lambda_{ct}} + \chi_{1t}^i =$$

$$0 < \Psi_2^* - \frac{\zeta_t^i}{\lambda_{ct}} + \chi_{1t} < \frac{\zeta_t^i}{\lambda_{ct}} - \frac{\zeta_t^i}{\lambda_{ct}} + \chi_{1t}^i = \chi_{1t}^i.$$

Hence, $\chi_{1t}^i > 0$. □

A.3 Combined First-Order Conditions

$$E_t \left\{ \beta \tfrac{\lambda_{ct+1}}{\lambda_{ct}} \left[\tfrac{\sigma_t}{Q_t} \tfrac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d \left(1 - \frac{e_t}{l_t} \right) - R_{t+1}^s \right) Q_t + \xi \sigma_t}{\sigma_t \sqrt{2} \tau} \right)^2} + \right. \right.$$

$$\left. \left. \left(\tfrac{R_{t+1}^s - \frac{\sigma_t \xi}{Q_t} - R_t^d}{2} \right) \left[1 - \text{erf} \left(\tfrac{\left(R_t^d \left(1 - \frac{e_t}{l_t} \right) - R_{t+1}^s \right) Q_t + \xi \sigma_t}{\sigma_t \sqrt{2} \tau} \right) \right] \right] \right\} + \chi_{2t} = \gamma \chi_{1t},$$

$$E_t \left\{ \beta \tfrac{\lambda_{ct+1}}{\lambda_{ct}} \left[R_t^d \tfrac{1}{2} \left(1 - \text{erf} \left(\tfrac{\left(R_t^d \left(1 - \frac{e_t}{l_t} \right) - R_{t+1}^s \right) Q_t + \xi \sigma_t}{\sigma_t \sqrt{2} \tau} \right) \right) \right] \right\} - 1 + \chi_{1t} = 0.$$

Since $\chi_{1t} > 0$, multiply the second equation by γ_t and add it to the first equation using $\frac{e_t}{l_t} = \gamma_t$. Therefore, the FOCs can be combined into:

$$E_t \left\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \left[\frac{\sigma_t}{Q_t} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d(1-\gamma_t) - R_{t+1}^s\right) Q_t + \xi\sigma_t}{\sigma_t \sqrt{2}\tau} \right)^2} + \right.\right.$$

$$\left.\left. \frac{1}{2}\left(R_{t+1}^s - \frac{\sigma_t \xi}{Q_t} - R_t^d \right) \left[1 - \operatorname{erf}\left(\frac{\left(R_t^d(1-\gamma_t) - R_{t+1}^s\right) Q_t + \xi\sigma_t}{\sigma_t \sqrt{2}\tau} \right) \right] \right] \right\} = \gamma_t - \chi_{2t},$$

$$\chi_{2t} l_t = 0.$$

A.4 Zero-Profit Condition

Consider the zero-profit condition under all states of nature. Since there is no agency problem between banks and households, this condition captures the fact that all the profits (or losses) are distributed to equity holders after realization of shocks at the beginning of each period. In each aggregate state, banks whose investments in risky firms pan out will have returns that satisfy on average (over the realizations of the idiosyncratic shock) $\left[\left(R_{t+1}^s + \frac{\sigma_t}{Q_t}\right) l_t - R_t^d \left(l_t - e_t\right)\right] - \int R_{t+1,b}^e(b) \cdot e_t = 0$, where the bounds of the integral are chosen such that we integrate over banks for which the profit is non-negative, while banks whose risky investments earn low (negative) returns will have $R_{t+1,b}^e = 0$. Therefore,

$$R_{t+1}^e = \int_{\left(\frac{R_t^d(1-\gamma_t) - R_{t+1}^s}{\sigma_t}\right) Q_t}^{\infty} \frac{\left(\left(R_{t+1}^s + \sigma_t \frac{\varepsilon_{t+1}}{Q_t}\right) l_t - R_t^d d_t\right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \mathrm{d}\varepsilon_{t+1}}{e_t} +$$

$$\int_{-\infty}^{\left(\frac{R_t^d(1-\gamma_t) - R_{t+1}^s}{\sigma_t}\right) Q_t} 0 \cdot \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \mathrm{d}\varepsilon_{t+1} =$$

$$\frac{1}{e_t} \int_{\left(\frac{R_t^d(1-\gamma_t) - R_{t+1}^s}{\sigma_t}\right) Q_t}^{\infty} \left(R_{t+1}^s l_t - R_t^d d_t\right) \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \mathrm{d}\varepsilon_{t+1} +$$

$$\frac{1}{e_t} \int_{\left(\frac{R_t^d(1-\gamma_t) - R_{t+1}^s}{\sigma_t}\right) Q_t}^{\infty} \sigma_t \frac{\varepsilon_{t+1}}{Q_t} l_t \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}} \mathrm{d}\varepsilon_{t+1} =$$

$$\frac{1}{e_t}\left[\left(R^s_{t+1}l_t - R^d_t d_t\right)\frac{1}{2}\left(1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right)+\right.$$
$$\left.\frac{\sigma_t l_t}{Q_t}\left(\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2}-\frac{\xi}{2}\left[1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right)\right]=$$

$$\frac{l_t}{e_t}\left\{\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2}+\right.$$
$$\left.\frac{1}{2}\left(R^s_{t+1}-\frac{\sigma_t\xi}{Q_t}-R^d_t\left(1-\gamma_t\right)\right)\left[1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right\}.$$

Since $\frac{l_t}{e_t}=\frac{1}{\gamma_t}$, we can rewrite the latter condition as (using that it holds for each $i\in\{s,r\}$):

$$R^{e,i}_{t+1}=\frac{\frac{\sigma^i_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma^i_t}{\sigma^i_t\sqrt{2}\tau}\right)^2}+\frac{1}{2}\left(R^s_{t+1}-\frac{\sigma^i_t\xi}{Q_t}-R^d_t(1-\gamma_t)\right)\left[1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma^i_t}{\sigma^i_t\sqrt{2}\tau}\right)\right]}{\gamma_t}.$$

Note that the combined FOC from Appendix A.3 can be expressed as:

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\frac{\sigma^i_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma^i_t}{\sigma^i_t\sqrt{2}\tau}\right)^2}+\right.\right.$$
$$\left.\left.\frac{1}{2}\left(R^s_{t+1}-\frac{\sigma^i_t\xi}{Q_t}-R^d_t\right)\left[1-\text{erf}\left(\frac{\left(R^d_t(1-\gamma_t)-R^s_{t+1}\right)Q_t+\xi\sigma^i_t}{\sigma^i_t\sqrt{2}\tau}\right)\right]\right]\right\}=$$
$$\gamma_t-\chi^i_{2t}=\gamma_t\left(E_t\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}R^{e,i}_{t+1}+\frac{\zeta^i_t}{\lambda_{ct}}\right)-\chi^i_{2t},$$

where we substitute for 1 from Household's FOC with respect to two types of equity: $\beta E_t\frac{\lambda_{ct+1}}{\lambda_{ct}}R^{e,i}_{t+1}=1-\frac{\zeta^i_t}{\lambda_{ct}}$.

Notice that $l^i_t>0$ implies both $\chi^i_{2t}=0$ and $\zeta^i_t=0$ which say that the zero-profit condition implies the FOC.

A.5 Expression of Expected Dividends

Expected dividends (valued on date t) are defined as

$$\Omega\left(\mu_t,\sigma_t;\, l_t,\, d_t,\, e_t\right) =$$

$$E_t\left[\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\int\limits_{\left(\frac{R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)Q_t}^{\infty}\left(\left(R_{t+1}^l+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R_t^d\left(l_t-e_t\right)\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right]=$$

We have already calculated all the necessary integrals in Appendix A.1. Therefore,

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\frac{\sigma_t l_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(R_t^d(l_t-e_t)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)^2}+\right.\right.$$
$$\left.\left.\frac{\left(R_{t+1}^s l_t-R_t^d\left(l_t-e_t\right)-\frac{\sigma_t\xi}{Q_t}l_t\right)}{2}\left[1-\operatorname{erf}\left(\frac{\left(R_t^d\left(l_t-e_t\right)-R_{t+1}^s l_t\right)Q_t+\xi\sigma_t l_t}{\sigma_t l_t\sqrt{2}\tau}\right)\right]\right]\right\}.$$

A.6 Linear Cost of Banking: FOCs of Banks

We use $\left(R_{t+1}^s+\sigma_t\frac{\varepsilon_{t+1}^*}{Q_t}\right)l_t-R_t^d d_t-fl_t=0$ to get $\varepsilon_{t+1}^*=\left(\frac{fl_t+R_t^d(l_t-e_t)}{\sigma_t l_t}-\frac{R_{t+1}^l}{\sigma_t}\right)Q_t$. Conditional on the optimal choice of σ_t, the first-order conditions are:

$$E_t\left[\int\limits_{\varepsilon_{t+1}^*}^{\infty}\psi_{t,t+1}\frac{\partial}{\partial l_t}\left(\left(R_{t+1}^s+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R_t^d\left(l_t-e_t\right)-fl_t\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right]+$$
$$\chi_{2t}-\gamma\chi_{1t}=0,$$
$$E_t\left[\int\limits_{\varepsilon_{t+1}^*}^{\infty}\psi_{t,t+1}\frac{\partial}{\partial e_t}\left(\left(R_{t+1}^s+\sigma_t\frac{\varepsilon_{t+1}}{Q_t}\right)l_t-R_t^d\left(l_t-e_t\right)-fl_t\right)\frac{1}{\sqrt{2\pi\tau^2}}e^{-\frac{(\varepsilon_{t+1}+\xi)^2}{2\tau^2}}\,\mathrm{d}\varepsilon_{t+1}\right]-$$
$$1+\chi_{1t}=0.$$

The derivations are similar to the ones described in Appendix A.1. The only difference is that the lower bound of the integral now contains the additional term fl_t. Hence, adding ξ to the lower limit of the integral and dividing the result by $\sqrt{2}\tau$ make the terms in the final expressions. Moreover, note that we should carry f in the expressions of the FOC with respect to l_t. In sum, the FOCs can be written as follows:

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[\frac{\sigma_t}{Q_t}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(f+R_t^d\left(1-\frac{e_t}{l_t}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)^2}+\right.\right.$$

$$\left.\left.\left(\frac{R_{t+1}^s-\frac{\sigma_t\xi}{Q_t}-R_t^d-f}{2}\right)\left[1-\operatorname{erf}\left(\frac{\left(f+R_t^d\left(1-\frac{e_t}{l_t}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right]\right]\right\}+\chi_{2t}=\gamma\chi_{1t},$$

$$E_t\left\{\beta\frac{\lambda_{ct+1}}{\lambda_{ct}}\left[R_t^d\frac{1}{2}\left(1-\operatorname{erf}\left(\frac{\left(f+R_t^d\left(1-\frac{e_t}{l_t}\right)-R_{t+1}^s\right)Q_t+\xi\sigma_t}{\sigma_t\sqrt{2}\tau}\right)\right)\right]\right\}-1+\chi_{1t}=0.$$

B The Non-Financial Firm's Problem

B.1 Safe firms

Let π_{t+1}^s denote the revenue of a safe firm in period $t+1$ net of expenses:

$$\pi_{t+1}^s=y_{t+1}^s+(1-\delta)Q_tk_{t+1}^s-W_{t+1}h_{t+1}^s-R_{t+1}^sl_t^{f,s}.$$

In this notation, the problem of the safe firm is to

$$\max_{l_t^{f,s},k_{t+1}^s} E_t\left\{\max_{h_{t+1}^s}\pi_{t+1}^s\right\}.$$

The first-order condition for $\max_{h_{t+1}^s}\pi_{t+1}^s$ is $\frac{\partial\pi_{t+1}^s}{\partial h_{t+1}^s}=0$. It implies that

$$W_{t+1} = \frac{\partial y_{t+1}^s}{\partial h_{t+1}^s}=(1-\alpha)\frac{y_{t+1}^s}{h_{t+1}^s}=(1-\alpha)A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^\alpha, \tag{B.1}$$

$$h_{t+1}^s = (1-\alpha)\frac{y_{t+1}^s}{W_{t+1}}=(1-\alpha)\frac{A_{t+1}\left(k_{t+1}^s\right)^\alpha\left(h_{t+1}^s\right)^{1-\alpha}}{W_{t+1}}. \tag{B.2}$$

Accordingly, the safe firm's Lagrangian is:

$$\mathcal{L}^{\text{safe}}=E_t\left\{A_{t+1}\left(k_{t+1}^s\right)^\alpha\left(h_{t+1}^s\right)^{1-\alpha}+(1-\delta)Q_{t+1}k_{t+1}^s-W_{t+1}h_{t+1}^s-R_{t+1}^sl_t^{f,s}\right\}+$$

$$\lambda_{ht}^sE_t\left\{(1-\alpha)\frac{A_{t+1}\left(k_{t+1}^s\right)^\alpha\left(h_{t+1}^s\right)^{1-\alpha}}{W_{t+1}}-h_{t+1}^s\right\}+\lambda_{lt}^s\left(l_t^{f,s}-Q_tk_{t+1}^s\right).$$

Notice that there is no expectation operator on the Lagrangian multipliers because those constraints hold under every state of nature. The problem implies the following first-order conditions

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial l_t^{f,s}} = -E_t\left[R_{t+1}^s\right] + \lambda_{lt}^s = 0,$$

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial k_{t+1}^s} = E_t\left[\alpha \frac{y_{t+1}^s}{k_{t+1}^s} + (1-\delta)Q_{t+1}\right] + \lambda_{ht}^s(1-\alpha)\alpha E_t\left[\frac{A_{t+1}}{W_{t+1}}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1}\right] - \lambda_{lt}^s Q_t = 0,$$

$$\frac{\partial \mathcal{L}^{\text{safe}}}{\partial h_{t+1}^s} = (1-\alpha)\frac{A_{t+1}\left(k_{t+1}^s\right)^\alpha\left(h_{t+1}^s\right)^{1-\alpha}}{W_{t+1}} - W_{t+1} + \lambda_{ht}^s\left[(1-\alpha)^2\frac{A_{t+1}}{W_{t+1}}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^\alpha - 1\right] = 0.$$

Combining $\frac{\partial \mathcal{L}^{\text{safe}}}{\partial h_{t+1}^s} = 0$ with equation (B.2) yields $\lambda_{ht}^s = 0$. Then, plugging $\frac{\partial \mathcal{L}^{\text{safe}}}{\partial l_t^{f,s}} = 0$ into $\frac{\partial \mathcal{L}^{\text{safe}}}{\partial k_{t+1}^s}$ for λ_{lt}^s, we get

$$E_t\left[R_{t+1}^s\right]Q_t = E_t\left[\alpha \frac{y_{t+1}^s}{k_{t+1}^s} + (1-\delta)Q_{t+1}\right].$$

Consider the zero-profit condition of the safe firm under all states of nature. Since output function has constant returns to scale,

$$y_{t+1}^s = \frac{\partial y_{t+1}^s}{\partial k_{t+1}^s}k_{t+1}^s + \frac{\partial y_{t+1}^s}{\partial h_{t+1}^s}h_{t+1}^s = \alpha A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1}k_{t+1}^s + W_{t+1}h_{t+1}^s,$$

where we use equation (B.2) to substitute for W_{t+1} in the last equality. Plugging the expression of y_{t+1}^s into $\pi_{t+1}^s = 0$ and using $Q_t k_{t+1}^s = l_t^{f,s}$, we find that:

$$\alpha A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1}k_{t+1}^s + (1-\delta)Q_{t+1}k_{t+1}^s - R_{t+1}^s Q_t k_{t+1}^s = 0.$$

Since $k_{t+1}^s > 0$, we can divide by k_{t+1}^s to get

$$R_{t+1}^s Q_t = \alpha A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1} + (1-\delta)Q_{t+1} \tag{B.3}$$

under all states of nature. This condition implies the first-order condition

$$E_t\left[R_{t+1}^s\right]Q_t = E_t\left[\alpha A_{t+1}\left(\frac{k_{t+1}^s}{h_{t+1}^s}\right)^{\alpha-1} + (1-\delta)Q_{t+1}\right].$$

B.2 Risky Firms

Let π_{t+1}^r denote the revenue of a risky firm in period $t+1$ net of expenses:

$$\pi_{t+1}^r = y_{t+1}^r + (1-\delta)Q_t k_{t+1}^r - W_{t+1}h_{t+1}^r - R_{t+1}^r l_t^{f,r}.$$

In this notation, the problem of the risky firm is to

$$\max_{l_t^{f,r},k_{t+1}^r} E_t\left\{\max_{h_{t+1}^r}\pi_{t+1}^r\right\}.$$

The first-order condition for $\max_{h_{t+1}^r}\pi_{t+1}^r$ is $\frac{\partial\pi_{t+1}^r}{\partial h_{t+1}^r}=0$. It implies that

$$W_{t+1} = \frac{\partial y_{t+1}^r}{\partial h_{t+1}^r} = (1-\alpha)A_{t+1}\left(\frac{k_{t+1}^r}{h_{t+1}^r}\right)^{\alpha}, \tag{B.4}$$

$$h_{t+1}^r = (1-\alpha)\frac{A_{t+1}\left(k_{t+1}^r\right)^{\alpha}\left(h_{t+1}^r\right)^{1-\alpha}}{W_{t+1}}. \tag{B.5}$$

Accordingly, the risky firm's Lagrangian is:

$$\mathcal{L}^{\text{risky}} = E_t\left[A_{t+1}\left(k_{t+1}^r\right)^{\alpha}\left(h_{t+1}^r\right)^{1-\alpha}+\varepsilon_{t+1}k_{t+1}^r+(1-\delta)Q_{t+1}k_{t+1}^r-W_{t+1}h_{t+1}^r-R_{t+1}^r l_t^{f,r}\right]+$$
$$\lambda_{ht}^r E_t\left[(1-\alpha)\frac{A_{t+1}\left(k_{t+1}^r\right)^{\alpha}\left(h_{t+1}^r\right)^{1-\alpha}}{W_{t+1}}-h_{t+1}^r\right]+\lambda_{lt}^r\left(l_t^{f,r}-Q_t k_{t+1}^r\right).$$

Notice that there is no expectation operator on the Lagrangian multipliers because those constraints hold under every state of nature. The problem implies the following first-order conditions

$$\frac{\partial\mathcal{L}^{\text{risky}}}{\partial l_t^{f,r}} = -E_t\left[R_{t+1}^r\right]+\lambda_{lt}^r=0,$$

$$\frac{\partial\mathcal{L}^{\text{risky}}}{\partial k_{t+1}^r} = E_t\left[\alpha A_{t+1}\left(\frac{k_{t+1}^r}{h_{t+1}^r}\right)^{\alpha-1}+\varepsilon_{t+1}+(1-\delta)Q_{t+1}\right]+$$
$$\lambda_{ht}^r E_t\left[\alpha(1-\alpha)\frac{A_{t+1}}{W_{t+1}}\left(\frac{k_{t+1}^r}{h_{t+1}^r}\right)^{\alpha-1}\right]-\lambda_{lt}^r Q_t=0,$$

$$\frac{\partial\mathcal{L}^{\text{risky}}}{\partial h_{t+1}^r} = (1-\alpha)A_{t+1}\left(\frac{k_{t+1}^r}{h_{t+1}^r}\right)^{\alpha}-W_{t+1}+\lambda_{ht}^r\left[(1-\alpha)^2\frac{A_{t+1}}{W_{t+1}}\left(\frac{k_{t+1}^r}{h_{t+1}^r}\right)^{\alpha}-1\right]=0.$$

Equation (B.4) together with $\frac{\partial\mathcal{L}^{\text{risky}}}{\partial h_{t+1}^r}=0$ yield $\lambda_{ht}^r=0$. Plugging $\frac{\partial\mathcal{L}^{\text{risky}}}{\partial l_t^{f,r}}=0$ into $\frac{\partial\mathcal{L}^{\text{risky}}}{\partial k_{t+1}^r}$ for λ_{lt}^r, we get

$$E_t\left[R_{t+1}^r\right]Q_t = E_t\left[\alpha A_{t+1}\left(\frac{k_{t+1}^r}{h_{t+1}^r}\right)^{\alpha-1}+(1-\delta)Q_{t+1}+\varepsilon_{t+1}\right].$$

Combining equation (B.1) with equation (B.4):

$$\frac{k^s_{t+1}}{h^s_{t+1}} = \frac{k^r_{t+1}}{h^r_{t+1}} \tag{B.6}$$

under all states of nature. But remember that the first-order condition of the safe firm implies

$$E_t\left[R^s_{t+1}\right]Q_t = E_t\left[\alpha A_{t+1}\left(\frac{k^s_{t+1}}{h^s_{t+1}}\right)^{\alpha-1} + (1-\delta)Q_{t+1}\right].$$

Therefore

$$E_t\left[R^s_{t+1}\right]Q_t = E_t\left[R^s_{t+1}Q_t + \varepsilon_{t+1}\right].$$

Consider the zero-profit condition of the risky firm under all states of nature.

$$\begin{aligned}\pi^r_{t+1} = y^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - W_{t+1}h^r_{t+1} - R^r_{t+1}l^{f,r}_t = \\ y^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - (1-\alpha)A_{t+1}\left(k^r_{t+1}\right)^{\alpha}\left(h^r_{t+1}\right)^{1-\alpha} - R^r_{t+1}l^{f,r}_t = \\ \alpha A_{t+1}\left(k^r_{t+1}\right)^{\alpha}\left(h^r_{t+1}\right)^{1-\alpha} + \varepsilon_{t+1}k^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - R^r_{t+1}l^{f,r}_t = \\ \alpha A_{t+1}\left(\frac{k^r_{t+1}}{h^r_{t+1}}\right)^{\alpha-1} k^r_{t+1} + \varepsilon_{t+1}k^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - R^r_{t+1}l^{f,r}_t = 0,\end{aligned}$$

where we use equation (B.5) to substitute for $W_{t+1}h^r_{t+1}$. Using equation (B.3) together with equation (B.6), we can express

$$\alpha A_{t+1}\left(\frac{k^r_{t+1}}{h^r_{t+1}}\right)^{\alpha-1} = R^s_{t+1}Q_t - (1-\delta)Q_{t+1},$$

that holds under all states of nature. Plugging it into the zero-profit condition and using $Q_t k^r_{t+1} = l^{f,r}_t$, we find that:

$$R^s_{t+1}Q_t k^r_{t+1} - (1-\delta)Q_{t+1}k^r_{t+1} + \varepsilon_{t+1}k^r_{t+1} + (1-\delta)Q_t k^r_{t+1} - R^r_{t+1}Q_t k^r_{t+1} = 0.$$

Since $k^r_{t+1} > 0$, we can divide by k^r_{t+1} to get

$$R^r_{t+1}Q_t = R^s_{t+1}Q_t + \varepsilon_{t+1}$$

under all states of nature. This condition implies

$$E_t\left[R^r_{t+1}\right]Q_t = E_t\left[R^s_{t+1}Q_t + \varepsilon_{t+1}\right].$$

B.3 Aggregating across firms

Here we show that we can aggregate individual firms into two representative firms. Let denote $k_{j,t}^i$ the capital chosen by firm i that is financed by borrowing from bank j. Both i and j lie within the continuum of measure 1 of banks and firms, respectively. In this notation, the equation (B.6) is written as

$$\frac{k_{j,t+1}^i}{h_{j,t+1}^i} = \frac{k_{t+1}}{h_{t+1}}, \tag{B.7}$$

for all $j \in [0,1]$ and $i \in [0,1]$. Each firm chooses the same capital-to-labor ratio independently of the type of bank it borrows from.

Notice is that σ_t is the fraction of risky firms at date t; the remaining fraction $1-\sigma_t$ of firms are safe firms. Let's index firms as follows: firm j_1, with $j_1 \in [0, \sigma_t]$, can only access a risky technology subject to both aggregate and idiosyncratic shocks; firm j_2, with $j_2 \in [\sigma_t, 1]$ has access to a safe production technology subject to aggregate shocks only. Since there are no equilibria with $\underline{\sigma} < \sigma_t < \bar{\sigma}$, the fraction of risky firms is linked to the fraction of banks with risky portfolios as follows:

$$\sigma_t = (1-\mu_t)\,\underline{\sigma} + \mu_t \bar{\sigma}.$$

Define the following objects: Let $K_{s,t+1}^s = \int_{\sigma_t}^1 \int_{\mu_t}^1 k_{j,t+1}^i dj di$ be the total capital allocated to the safe technology and financed by borrowing from the banks that choose a fraction $\underline{\sigma}$ of risky projects. Let $K_{r,t+1}^s = \int_{\sigma_t}^1 \int_0^{\mu_t} k_{j,t+1}^i dj di$ be the total capital allocated to the safe technology and financed by borrowing from the banks that choose a fraction $\bar{\sigma}$ of risky projects. We let K_{t+1}^s denote the total capital allocated to the safe technology. Thus,

$$K_{t+1}^s = \int_{\sigma_t}^{1} \int_{0}^{1} k_{j,t+1}^i dj di = K_{s,t+1}^s + K_{r,t+1}^s,$$

Let $K_{s,t+1}^r = \int_0^{\sigma_t} \int_{\mu_t}^1 k_{j,t+1}^i dj di$ be the total capital allocated to the risky technology and financed by borrowing from the banks that choose a fraction $\underline{\sigma}$ of risky projects. Let $K_{r,t+1}^r = \int_0^{\sigma_t} \int_0^{\mu_t} k_{j,t+1}^i dj di$ be the total capital allocated to the safe technology and financed by borrowing from the banks that choose a fraction $\bar{\sigma}$ of risky projects. We let K_{t+1}^r denote the total capital allocated to the risky technology. Thus,

$$K_{t+1}^r = \int_{0}^{\sigma_t} \int_{0}^{1} k_{j,t+1}^i dj di = K_{s,t+1}^r + K_{r,t+1}^r,$$

The same upper and lower case notation applies to labor, i.e. $H^s_{s,t+1} = \int_{\sigma_t}^1 \int_{\mu_t}^1 h^i_{j,t+1} dj di$; $H^s_{r,t+1} = \int_{\sigma_t}^1 \int_0^{\mu_t} h^i_{j,t+1} dj di$; $H^r_{s,t+1} = \int_0^{\sigma_t} \int_{\mu_t}^1 h^i_{j,t+1} dj di$; $H^r_{r,t+1} = \int_0^{\sigma_t} \int_0^{\mu_t} h^i_{j,t+1} dj di$.

Safe representative firm produces:

$$Y^s_t = \int_{\sigma_{t-1}}^1 \int_0^1 A_t \left(k^i_{j,t}\right)^\alpha \left(h^i_{j,t}\right)^{1-\alpha} dj di = \int_{\sigma_{t-1}}^1 \int_0^1 F\left(k^i_{j,t}, h^i_{j,t}\right) dj di =$$

Using that the technology has Constant Returns to Scale:

$$= \int_{\sigma_{t-1}}^1 \int_0^1 \left[F_{k^i_{j,t}}\left(k^i_{j,t}, h^i_{j,t}\right) k^i_{j,t} + F_{h^i_{j,t}}\left(k^i_{j,t}, h^i_{j,t}\right) h^i_{j,t}\right] dj di =$$

where $F_{k^i_{j,t}}\left(k^i_{j,t}, h^i_{j,t}\right)$ and $F_{h^i_{j,t}}\left(k^i_{j,t}, h^i_{j,t}\right)$ denote the partial derivative of $F\left(k^i_{j,t}, h^i_{j,t}\right)$ with respect to $k^i_{j,t}$ and $h^i_{j,t}$, respectively. Since these partial derivatives are homogeneous of degree zero, we can express them in term of capital-labor ratio, i.e.

$$= \int_{\sigma_{t-1}}^1 \int_0^1 \left[f_{k^i_{j,t}}\left(\frac{k^i_{j,t}}{h^i_{j,t}}\right) k^i_{j,t} + f_{h^i_{j,t}}\left(\frac{k^i_{j,t}}{h^i_{j,t}}\right) h^i_{j,t}\right] dj di = \text{Plugging equation (B.7)} =$$

$$= \int_{\sigma_{t-1}}^1 \int_0^1 \left[f_{k_t}\left(\frac{k_t}{h_t}\right) k^i_{j,t} + f_{h_t}\left(\frac{k_t}{h_t}\right) h^i_{j,t}\right] dj di =$$

$$f_{k_t}\left(\frac{k_t}{h_t}\right)\left[\int_{\sigma_t}^1 \int_0^1 k^i_{j,t} dj di\right] + f_{h_t}\left(\frac{k_t}{h_t}\right)\left[\int_{\sigma_t}^1 \int_0^1 h^i_{j,t} dj di\right] = f_{k_t}\left(\frac{k_t}{h_t}\right) K^s_t + f_{h_t}\left(\frac{k_t}{h_t}\right) H^s_t =$$

Since $\frac{K^s_{s,t}}{H^s_{s,t}} = \frac{K^s_{r,t}}{H^s_{r,t}} = \frac{k_t}{h_t}$,then $\frac{K^s_t}{H^s_t}\frac{h_t}{k_t} = \left(\frac{K^s_{s,t}+K^s_{r,t}}{H^s_{s,t}+H^s_{r,t}}\right)\frac{H^s_{r,t}}{K^s_{r,t}} = 1$. Therefore $\frac{K^s_t}{H^s_t} = \frac{k_t}{h_t}$.

$$= f_{K^s_t}\left(\frac{K^s_t}{H^s_t}\right) K^s_t + f_{H^s_t}\left(\frac{K^s_t}{H^s_t}\right) H^s_t = A_t \left(K^s_t\right)^\alpha \left(H^s_t\right)^{1-\alpha}.$$

Risky representative firm:

$$Y^r_t = \int_0^{\sigma_{t-1}} \int_0^1 \left[A_t \left(k^i_{j,t}\right)^\alpha \left(h^i_{j,t}\right)^{1-\alpha} + \varepsilon^i_{j,t} k^i_{j,t}\right] dj di = \int_0^{\sigma_{t-1}} \int_0^1 F\left(k^i_{j,t}, h^i_{j,t}\right) dj di + \int_0^{\sigma_{t-1}} \int_0^1 \varepsilon^i_{j,t} k^i_{j,t} dj di$$

Note that the similar steps described above apply to the first term in the summation, so that $\int_0^{\sigma_{t-1}} \int_0^1 F\left(k^i_{j,t}, h^i_{j,t}\right) dj di = A_t \left(K^r_t\right)^\alpha \left(H^r_t\right)^{1-\alpha}$. To express the second term, notice that $\int_0^{\sigma_{t-1}}\int_0^1 \varepsilon_{j,t}^i k_{j,t}^i dj di = -\xi$. Moreover since each risky firm solves the same maximization problem, it chooses the same amount of capital independently of the type of bank it borrows from. Therefore, $\int_0^{\sigma_{t-1}}\int_0^1 \varepsilon_{j,t}^i k_{j,t}^i dj di = -\xi K_t^r$. Hence,

$$Y_t^r = A_t \left(K_t^r\right)^{\alpha} \left(H_t^r\right)^{1-\alpha} - \xi K_t^r.$$

C The Government

The government levies the tax to fully compensate for the loss to the deposit insurance fund due to rescue of defaulted banks.

C.1 Baseline: No linear cost of banking

$$T_t = -\int_{-\infty}^{\left(\frac{R_{t-1}^d D_{t-1}}{\sigma_{t-1} L_{t-1}} - \frac{R_t^s}{\sigma_{t-1}}\right) Q_{t-1}} \left(\left(R_t^l + \frac{\sigma_{t-1}\varepsilon_t}{Q_{t-1}}\right) L_{t-1} - R_{t-1}^d D_{t-1}\right) \mathrm{d}G(\varepsilon_t) =$$

$$-\left[\int_{-\infty}^{\infty} \left(\left(R_t^l + \frac{\sigma_{t-1}\varepsilon_t}{Q_{t-1}}\right) L_{t-1} - R_{t-1}^d D_{t-1}\right) \mathrm{d}G(\varepsilon_t) - \right.$$

$$\left. \int_{\left(\frac{R_{t-1}^d D_{t-1}}{\sigma_{t-1} L_{t-1}} - \frac{R_t^s}{\sigma_{t-1}}\right) Q_{t-1}}^{\infty} \left(\left(R_t^s + \frac{\sigma_{t-1}\varepsilon_t}{Q_{t-1}}\right) L_{t-1} - R_{t-1}^d D_{t-1}\right) \mathrm{d}G(\varepsilon_t)\right] =$$

Note that the first term equals $\left(R_t^s - \frac{\sigma_{t-1}\xi}{Q_{t-1}}\right) L_{t-1} + R_{t-1}^d D_{t-1}$ in the square bracket. We have already calculated the second term. Therefore,

$$= \frac{\sigma_{t-1} L_{t-1}}{Q_{t-1}} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{R_{t-1}^d (1-\gamma_{t-1}) Q_{t-1} - R_t^s Q_{t-1} + \xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)^2} - \left(R_t^s - \frac{\sigma_{t-1}\xi}{Q_{t-1}}\right) L_{t-1} + R_{t-1}^d D_{t-1} +$$

$$\frac{1}{2} L_{t-1} \left(R_t^s - \frac{\sigma_{t-1}\xi}{Q_{t-1}} - (1-\gamma_{t-1}) R_{t-1}^d\right) \left[1 - \mathrm{erf}\left(\frac{R_{t-1}^d (1-\gamma_{t-1}) Q_{t-1} - R_t^s Q_{t-1} + \xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right] =$$

$$\frac{\sigma_{t-1} L_{t-1}}{Q_{t-1}} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{R_{t-1}^d (1-\gamma_{t-1}) Q_{t-1} - R_t^s Q_{t-1} + \xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)^2} -$$

$$\frac{1}{2}\left(R_t^s L_{t-1} - \frac{\sigma_{t-1}\xi}{Q_{t-1}} L_{t-1} - R_{t-1}^d D_{t-1}\right) \left[1 + \mathrm{erf}\left(\frac{R_{t-1}^d (1-\gamma_{t-1}) Q_{t-1} - R_t^s Q_{t-1} + \xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right].$$

C.2 Linear Cost of Banking: Tax

The tax that accounts for the cost of banking is described as follows:

$$T_t = -\int_{-\infty}^{\left(\frac{R^d_{t-1}d_{t-1}}{\sigma_{t-1}l_{t-1}}-\frac{R^s_t-f}{\sigma_{t-1}}\right)Q_{t-1}} \left(\left(R^s_t+\frac{\sigma_{t-1}\varepsilon_t}{Q_{t-1}}-f\right)l_{t-1}-R^d_{t-1}d_{t-1}\right)\,\mathrm{d}G(\varepsilon_t) =$$

$$\frac{\sigma_{t-1}l_{t-1}}{Q_{t-1}}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(f+R^d_{t-1}(1-\gamma_{t-1})-R^s_t\right)Q_{t-1}+\xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)^2} - \left(R^l_t-\frac{\sigma_{t-1}\xi}{Q_{t-1}}-f\right)l_{t-1}+R^d_{t-1}d_{t-1}+$$

$$\frac{1}{2}l_{t-1}\left(R^s_t-\frac{\sigma_{t-1}\xi}{Q_{t-1}}-(1-\gamma_{t-1})R^d_{t-1}-f\right)\left[1-\operatorname{erf}\left(\frac{\left(f+R^d_{t-1}(1-\gamma_{t-1})-R^s_t\right)Q_{t-1}+\xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right]=$$

$$\frac{\sigma_{t-1}l_{t-1}}{Q_{t-1}}\frac{\tau}{\sqrt{2\pi}}e^{-\left(\frac{\left(f+R^d_{t-1}(1-\gamma_{t-1})-R^s_t\right)Q_{t-1}+\xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)^2} -$$

$$\frac{1}{2}\left(R^s_t l_{t-1}-\frac{\sigma_{t-1}\xi}{Q_{t-1}}l_{t-1}-R^d_{t-1}d_{t-1}-fl_{t-1}\right)\left[1+\operatorname{erf}\left(\frac{\left(f+R^d_{t-1}(1-\gamma_{t-1})-R^s_t\right)Q_{t-1}+\xi\sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right].$$

D Choice of Risk

This appendix shows a proof that the expected dividends function of banks is convex in the risk parameter σ_t. This result guarantees that banks choose either the maximum risk, $\bar{\sigma}$, or the minimum risk, $\underline{\sigma}$, to maximize their profits, so all the intermediate values of σ_t, which may result from the first-order conditions with respect to σ_t, are not optimal.

We generalize the proof taken from Van den Heuvel (2008) to the case with aggregate uncertainty. The proof applies to an arbitrary distribution of the idiosyncratic shock, ε_{t+1}, with non-positive mean, so our example of a Normal distribution considered in the analysis is not a special case which can drive our results. It is used for expositional reasons and quantitative work.

Assumption. *ε has a cumulative distribution function G_ε with support $[\underline{\varepsilon}, \bar{\varepsilon}]$, with $\underline{\varepsilon} < 0 < \bar{\varepsilon}$. The mean of ε is equal to $-\xi$ ($\xi > 0$). ε is independent of the aggregate shock. The aggregate shock does not depend on the choice of σ_t.*

Note that we do not restrict the analysis to the bounded support[1], so $\underline{\varepsilon}$ and $\bar{\varepsilon}$ can take $-\infty$ and $+\infty$, respectively. Note that G_ε need not be continuous.

Let $\hat{\varepsilon}(\sigma_t, R^s_{t+1}) \equiv \left(\frac{R^d_t d_t}{\sigma_t l_t}-\frac{R^l_{t+1}}{\sigma_t}\right)Q_t = \frac{R^d_t(1-\gamma_t)-R^s_{t+1}}{\sigma_t}Q_t$, where the latter equation uses the result that the capital requirement constraint always binds. Therefore, $\left(R^s_{t+1}+\sigma_t\frac{\hat{\varepsilon}(\sigma_t)}{Q_t}\right)l_t -$

[1] Unbounded support is more relevant if we consider aggregate risk

$R_t^d d_t = 0$. Let $\pi(\sigma_t,\ R_{t+1}^s) = E_\varepsilon\left[\left(\left(R_{t+1}^s + \frac{\sigma_t\varepsilon}{Q_t}\right)l_t - R_t^d d_t\right)^+\right]$ be a function of expected dividends (taken over the idiosyncratic shock only) under some realization of R_{t+1}^s which is considered to be fixed in this function. To account for the aggregate uncertainty, R_{t+1}^s needs to be a random variable. Therefore, expected dividends taken into account both idiosyncratic and aggregate uncertainty are

$$\Pi(\sigma_t) = \int\limits_\Omega \pi\left(\sigma_t,\ R_{t+1}^s(\omega)\right) P(d\omega) = E_t\left[\int\limits_{\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s)}^{\bar{\varepsilon}} \left(\left(R_{t+1}^s + \frac{\sigma_t\varepsilon}{Q_t}\right)l_t - R_t^d d_t\right) dG_\varepsilon\right] =$$

$$E_t\left[\int\limits_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left(\left(R_{t+1}^s + \frac{\sigma_t\varepsilon}{Q_t}\right)l_t - R_t^d d_t\right) dG_\varepsilon\right] - E_t\left[\int\limits_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s)} \left(\left(R_{t+1}^s + \frac{\sigma_t\varepsilon}{Q_t}\right)l_t - R_t^d d_t\right) dG_\varepsilon\right] =$$

$$E_t R_{t+1}^s l_t - R_t^d d_t - \frac{\sigma_t\xi}{Q_t}l_t - \frac{\sigma_t l_t}{Q_t}E_t\left[\int\limits_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s)} \left(\varepsilon - \hat{\varepsilon}(\sigma_t,\ R_{t+1}^s)\right) dG_\varepsilon\right] =$$

$$E_t R_{t+1}^s l_t - R_t^d d_t + \frac{l_t}{Q_t}\left(\sigma_t E_t\left[\int\limits_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s)} \left(\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s) - \varepsilon\right) dG_\varepsilon\right] - \sigma_t\xi\right).$$

Note that in the derivations above we express $\left(R_{t+1}^s + \frac{\sigma_t\varepsilon}{Q_t}\right)l_t - R_t^d d_t$ in terms of $\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s)$ and ε using the definition of $\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s)$.

The proof below shows that $\Pi(\sigma_t)$ is convex in σ_t. Since the expression of $\Pi(\sigma_t)$ involves the term which is linear in σ_t and $\frac{l_t}{Q_t} \geq 0$, the sufficient condition for $\Pi(\sigma_t)$ to be convex in σ_t is that

$$H(\sigma_t) \equiv E_t\left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t)} \left(\hat{\varepsilon}(\sigma_t) - \varepsilon\right) dG_\varepsilon\right]\sigma_t$$

is convex in σ_t.

Claim. $H(\sigma_t) \equiv l_t E_t\left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t)} \left(\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s) - \varepsilon\right) dG_\varepsilon\right]\sigma_t$ *is convex in* σ_t*:*

Proof. Steps of the proof: □

1. Define $h(\sigma_t,\ R_{t+1}^s) \equiv \sigma_t\left[\int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s)} \left(\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s) - \varepsilon\right) dG_\varepsilon\right]$ in which the aggregate uncertainty is taken off. Consider 3 cases:

    (a) Realization of R_{t+1}^s is such that $\hat{\varepsilon}(\sigma_t,\ R_{t+1}^s) = \frac{R_t^d(1-\gamma_t) - R_{t+1}^s}{\sigma_t} > 0$, so $R_{t+1}^s < R_t^d\left(1-\gamma_t\right)$,

(b) Realization of R^s_{t+1} is such that $\hat{\varepsilon}(\sigma_t,\, R^l_{t+1}) = \frac{R^d_t(1-\gamma_t)-R^s_{t+1}}{\sigma_t} < 0$, so $R^s_{t+1} > R^d_t\,(1-\gamma_t)\,,$

(c) Realization of R^s_{t+1} is such that $\hat{\varepsilon}(\sigma_t,\, R^l_{t+1}) = \frac{R^d_t(1-\gamma_t)-R^s_{t+1}}{\sigma_t} = 0$, so $R^s_{t+1} = R^d_t\,(1-\gamma_t)\,,$

Show that $h(\sigma_t,\, R^s_{t+1})$ is convex in σ_t in cases 1a and 1b and $h(\sigma_t,\, R^s_{t+1})$ is linear in σ_t in case 1c.

2. Employ the argument that convexity is preserved under non-negative scaling and addition (guaranteed by the expectation operator over the aggregate uncertainty) to find that $H(\sigma_t)$ is convex.

Let's show each step of the proof formally

1. Let $\sigma_{1t} < \sigma_{2t}$ and, for $\lambda \in (0,\, 1)$, define $\sigma_{\lambda t} = \lambda\sigma_{1t} + (1-\lambda)\sigma_{2t}$. Let $\hat{\varepsilon}_i = \hat{\varepsilon}(\sigma_{it},\, R^s_{t+1}) \equiv \frac{R^d_t(1-\gamma_t)-R^s_{t+1}}{\sigma_{it}} Q_t$, for $i = 1,\ 2,\ \lambda$.

   (a) $R^s_{t+1} < R^d_t\,(1-\gamma_t)$: it implies that $\hat{\varepsilon}_2 < \hat{\varepsilon}_\lambda < \hat{\varepsilon}_1$,

$$
\begin{aligned}
h(\sigma_{\lambda t}) = (\lambda\sigma_{1t} + (1-\lambda)\sigma_{2t}) &\left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_{\lambda t})} (\hat{\varepsilon}(\sigma_{\lambda t}) - \varepsilon)\, dG_\varepsilon \right\} = \\
\lambda\sigma_{1t} &\left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon - \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} + \\
(1-\lambda)\sigma_{2t} &\left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} = \\
\lambda\sigma_{1t} &\left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \right\} + \\
(1-\lambda)\sigma_{2t} &\left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} \leq \\
\lambda\sigma_{1t} &\left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \right\} + \\
(1-\lambda)\sigma_{2t} &\left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, dG_\varepsilon \right\},
\end{aligned}
$$

where the inequality sign comes from $\int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon$ and $\int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_2}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, dG_\varepsilon$. Substituting for the definitions of $h(\sigma_{1t}) =$

$\sigma_{1t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon$ and $h(\sigma_{2t}) = \sigma_{2t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon$, we get:

$$h(\sigma_{\lambda t}) \leq \lambda h(\sigma_{1t}) + (1-\lambda)h(\sigma_{2t}) + \lambda\sigma_{1t} \left\{ (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_\lambda) \right\} + (1-\lambda)\sigma_{2t} \left\{ (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_\lambda) \right\} = \lambda h(\sigma_{1t}) + (1-\lambda)h(\sigma_{2t}) + G_\varepsilon(\hat{\varepsilon}_\lambda) \left(\lambda\sigma_{1t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) + (1-\lambda)\sigma_{2t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) \right) = \lambda h(\sigma_{1t}) + (1-\lambda)h(\sigma_{2t}),$$

where we use that $\sigma_{1t} = l_t \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) = \sigma_{2t}\hat{\varepsilon}_2 = \sigma_{\lambda t}\hat{\varepsilon}_\lambda$ in the last equality. So,

$$\lambda\sigma_{1t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) + (1-\lambda)\sigma_{2t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) = \hat{\varepsilon}_\lambda (\lambda\sigma_{1t} + (1-\lambda)\sigma_{2t}) - \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) (\lambda + (1-\lambda)) = \sigma_{\lambda t}\hat{\varepsilon}_\lambda - \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) = \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) - \left(R_t^d (1-\gamma_t) - R_{t+1}^s \right) = 0.$$

Therefore, $h(\sigma_t)$ is convex in σ_t for $R_{t+1}^s < R_t^d (1-\gamma_t)$.

(b) $R_{t+1}^s > R_t^d (1-\gamma_t)$: it implies that $\hat{\varepsilon}_1 < \hat{\varepsilon}_\lambda < \hat{\varepsilon}_2$

$$h(\sigma_{\lambda t}) = (\lambda\sigma_{1t} + (1-\lambda)\sigma_{2t}) \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\sigma_{\lambda t})} (\hat{\varepsilon}(\sigma_{\lambda t}) - \varepsilon)\, dG_\varepsilon \right\} = \lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} + (1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon - \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} = \lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \right\} + (1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \right\} \leq \lambda\sigma_{1t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_1) + \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, dG_\varepsilon \right\} + (1-\lambda)\sigma_{2t} \left\{ \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon + (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_2) + \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \right\},$$

where the inequality sign comes from $\int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \varepsilon)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_1}^{\hat{\varepsilon}_\lambda} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, dG_\varepsilon$ and $\int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\varepsilon - \hat{\varepsilon}_\lambda)\, dG_\varepsilon \leq \int_{\hat{\varepsilon}_\lambda}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \hat{\varepsilon}_\lambda)\, dG_\varepsilon$. Substituting for the definitions of $h(\sigma_{1t}) =$

$\sigma_{1t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_1} (\hat{\varepsilon}_1 - \varepsilon)\, dG_\varepsilon$ and $h(\sigma_{2t}) = \sigma_{2t} \int_{\underline{\varepsilon}}^{\hat{\varepsilon}_2} (\hat{\varepsilon}_2 - \varepsilon)\, dG_\varepsilon$, we get:

$$h(\sigma_{\lambda t}) \leq \lambda h(\sigma_{1t}) + (1-\lambda) h(\sigma_{2t}) + \lambda \sigma_{1t} \left\{ (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1)\, G_\varepsilon(\hat{\varepsilon}_\lambda) \right\} + \\ (1-\lambda)\sigma_{2t} \left\{ (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2)\, G_\varepsilon(\hat{\varepsilon}_\lambda) \right\} = \lambda h(\sigma_{1t}) + (1-\lambda) h(\sigma_{2t}) + \\ G_\varepsilon(\hat{\varepsilon}_\lambda) \left(\lambda \sigma_{1t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_1) + (1-\lambda)\sigma_{2t} (\hat{\varepsilon}_\lambda - \hat{\varepsilon}_2) \right) = \lambda h(\sigma_{1t}) + (1-\lambda) h(\sigma_{2t}),$$

where the last equality follows from the same reasoning employed in the previous case. Therefore, $h(\sigma_t)$ is convex in σ_t for $R^s_{t+1} > R^d_t (1 - \gamma_t)$.

(c) $R^s_{t+1} = R^d_t (1 - \gamma_t)$. Hence, $\hat{\varepsilon}(\sigma_t) = 0$ and

$$h(\sigma_t) = \sigma_t \left[\int_{\underline{\varepsilon}}^{0} (0 - \varepsilon)\, dG_\varepsilon \right],$$

which is linear in σ_t

2. We found in 1 that $h(\sigma_t,\ R^s_{t+1})$ is convex in σ_t for each $R^s_{t+1} \in \mathbb{R}$. Consider $P(\omega) \geq 0$ for each $R^l_{t+1}(\omega) \in \mathbb{R}$. Then the following function[2]:

$$\int_\Omega h\left(\sigma_t,\ R^s_{t+1}(\omega)\right) P(d\omega) = E_t\, h(\sigma_t,\ R^s_{t+1}) \equiv H(\sigma_t)$$

is convex in σ_t. It follows directly from the linearity of the expectation operator which puts a non-negative weight on every realization of R^s_{t+1} and the fact that the sum of convex functions is a convex function. Therefore, $\Pi(\sigma_t)$ is convex in σ_t. □

[2]Linearity in σ_t for one particular value of R^s_{t+1} can be considered as a weakly convex function, so it does not change the nature of the argument

E Equilibrium Conditions

For $\forall i \in [s, r]$:

$$(C_t - \kappa C_{t-1})^{-\varsigma_c} - \beta\kappa E_t \left(C_{t+1} - \kappa C_t\right)^{-\varsigma_c} - \lambda_{ct} = 0 \tag{E.1}$$

$$\varsigma_0 D_t^{-\varsigma_d} - \lambda_{ct} + E_t \beta \lambda_{ct+1} R_t^d = 0, \tag{E.2}$$

$$-\lambda_{ct} + E_t \beta \lambda_{ct+1} R_{t+1}^{e,s} + \zeta_t^s = 0, \tag{E.3}$$

$$-\lambda_{ct} + E_t \beta \lambda_{ct+1} R_{t+1}^{e,r} + \zeta_t^r = 0, \tag{E.4}$$

$$\zeta_t^s E_t^s = 0, \tag{E.5}$$

$$\zeta_t^r E_t^r = 0 \tag{E.6}$$

$$\begin{aligned} \gamma_t - \chi_{2t}^i = E_t \Bigg\{ \beta \frac{\lambda_{ct+1}}{\lambda_{ct}} \Bigg[\frac{\sigma_t^i}{Q_t} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d(1-\gamma_t) - R_{t+1}^s\right) Q_t + \xi \sigma_t^i}{\sigma_t^i \sqrt{2}\tau}\right)^2} + \\ \frac{1}{2}\left(R_{t+1}^s - \frac{\sigma_t^i \xi}{Q_t} - R_t^d\right) \left[1 - \text{erf}\left(\frac{\left(R_t^d\left(1-\gamma_t\right) - R_{t+1}^s\right) Q_t + \xi\sigma_t^i}{\sigma_t^i \sqrt{2}\tau}\right)\right]\Bigg]\Bigg\}, \end{aligned} \tag{E.7}$$

$$\begin{aligned} R_{t+1}^{e,i} = \frac{1}{\gamma_t} \Bigg\{ \frac{\sigma_t^i}{Q_t} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_t^d(1-\gamma_t) - R_{t+1}^s\right) Q_t + \xi \sigma_t^i}{\sigma_t^i \sqrt{2}\tau}\right)^2} + \\ \frac{1}{2}\left(R_{t+1}^s - \frac{\sigma_t^i \xi}{Q_t} - R_t^d\right) \left[1 - \text{erf}\left(\frac{\left(R_t^d\left(1-\gamma_t\right) - R_{t+1}^s\right) Q_t + \xi\sigma_t^i}{\sigma_t^i \sqrt{2}\tau}\right)\right]\Bigg\}, \end{aligned} \tag{E.8}$$

$$\chi_{2t}^i l_t^i = 0, \tag{E.9}$$

$$\sigma^s = \underline{\sigma}, \tag{E.10}$$

$$\sigma^r = \bar{\sigma}, \tag{E.11}$$

$$l_t^i = d_t^i + e_t^i, \tag{E.12}$$

$$e_t^i = \gamma_t l_t^i, \tag{E.13}$$

$$\mathit{\Omega}(\sigma_t^i;\ l_t^i,\ d_t^i,\ e_t^i) = E_t \left[\beta \tfrac{\lambda_{ct+1}}{\lambda_{ct}} R_{t+1}^{e,i} e_t^i\right], \tag{E.14}$$

$$\mu_t = \tfrac{E_t^r}{E_t^s + E_t^r}, \tag{E.15}$$

$$L_t^s = (1-\mu_t)\, l_t^s, \tag{E.16}$$

$$L_t^r = \mu_t l_t^r, \tag{E.17}$$

$$E_t^i = \gamma_t L_t^i, \tag{E.18}$$

$$L_t^i = D_t^i + E_t^i, \tag{E.19}$$

$$D_t = D_t^s + D_t^r, \tag{E.20}$$

$$Y_t^s = A_t \left(K_t^s\right)^{\alpha} \left(H_t^s\right)^{1-\alpha}, \tag{E.21}$$

$$Y_t^r = A_t \left(K_t^r\right)^{\alpha} \left(H_t^r\right)^{1-\alpha} - \xi K_t^r, \tag{E.22}$$

$$Q_t K_{t+1}^s = (1-\underline{\sigma})\, L_t^s + (1-\bar{\sigma})\, L_t^r, \tag{E.23}$$

$$Q_t K_{t+1}^r = \underline{\sigma} L_t^s + \bar{\sigma} L_t^r, \tag{E.24}$$

$$W_t = (1-\alpha)\, \tfrac{Y_t^s}{H_t^s}, \tag{E.25}$$

$$R_t^s = \tfrac{\alpha A_t}{Q_t} \left(\tfrac{K_t^s}{H_t^s}\right)^{\alpha-1} + (1-\delta)\, \tfrac{Q_{t+1}}{Q_t}, \tag{E.26}$$

$$R_t^r = R_t^s + \tfrac{\varepsilon_t}{Q_{t-1}}, \tag{E.27}$$

$$\tfrac{K_t^s}{H_t^s} = \tfrac{K_t^r}{H_t^r}, \tag{E.28}$$

$$H_t^s + H_t^r = 1, \tag{E.29}$$

$$K_t = K_t^s + K_t^r, \tag{E.30}$$

$$K_{t+1} = I_t + (1-\delta) K_t, \tag{E.31}$$

$$I_t = \eta_t \left[1 - \tfrac{\phi}{2}\left(\tfrac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] I_t^g, \tag{E.32}$$

$$\begin{aligned} \eta_t Q_t \left[1 - \frac{\phi}{2}\left(\frac{I_t^g}{I_{t-1}^g} - 1\right)^2\right] - \eta_t Q_t \phi \left(\frac{I_t^g}{I_{t-1}^g} - 1\right) \frac{I_t^g}{I_{t-1}^g} - 1 + \\ \eta_{t+1} \psi_{t,t+1} Q_{t+1} \phi \left(\frac{I_{t+1}^g}{I_t^g} - 1\right) \frac{I_{t+1}^g}{\left(I_t^g\right)^2} I_{t+1}^g = 0, \end{aligned} \tag{E.33}$$

$$Y_t^s + Y_t^r = C_t + I_t^g, \tag{E.34}$$

$$\begin{aligned} T_t = L_{t-1} \Bigg\{ \frac{\sigma_{t-1}}{Q_{t-1}} \frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\left(R_{t-1}^d(1-\gamma_{t-1}) - R_t^s\right) Q_{t-1} + \xi \sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)^2} - \\ \frac{1}{2}\left(R_t^s - R_{t-1}^d (1-\gamma_{t-1}) - \frac{\xi \sigma_{t-1}}{Q_{t-1}}\right) \left[1 + \operatorname{erf}\left(\frac{\left(R_{t-1}^d (1-\gamma_{t-1}) - R_t^s\right) Q_{t-1} + \xi \sigma_{t-1}}{\sigma_{t-1}\sqrt{2}\tau}\right)\right] \Bigg\}. \end{aligned} \tag{E.35}$$

F Discussion of the Excessive Risk-Taking Mechanism

Following our the result derived earlier, we can express the erf function in terms of the share of non-defaulted deposits of the representative bank and then decompose the expected dividend into two components:

$$\Omega\left(\mu_t, \sigma_t;\ l_t\right) = E_t\left\{\Lambda_{t,t+1} l_t\left[\omega_1 + \omega_2 - \left(1 - \gamma_t\right)\right]\right\},$$

where

$$\left[\omega_1 + \omega_2\right] = \left[\underbrace{\left(R^s_{t+1} - R^d_t\left(1 - \gamma_t\right) - \frac{\xi\sigma_t}{Q_t}\right)\underbrace{\left(1 - G(\varepsilon^*_{t+1})\right)}_{\text{non-defaulted}}}_{\substack{\omega_1 \equiv \text{ returns from a loan} \\ \text{portfolio with riskiness } \sigma_t}} + \underbrace{\left(\frac{\sigma_t}{Q_t}\right)\frac{\tau}{\sqrt{2\pi}} e^{-\left(\frac{\varepsilon^*_{t+1}+\xi}{\tau\sqrt{2}}\right)^2}}_{\substack{\omega_2 \equiv \text{ bonus from} \\ \text{projects volatility}}}\right],$$

and the cutoff point ε^*_{t+1} is defined by $R^d_t\left(1 - \gamma_t\right)Q_t - R^s_{t+1}Q_t = \sigma_t\varepsilon^*_{t+1}$.

The first component, ω_1, distinguishes loan returns of riskiness σ_t controlling for the variance of idiosyncratic shock (when τ is taken as given). The bank trades off the benefits from limited liability and deposit insurance with a smaller profitability of riskier projects. The term $\frac{\xi\sigma_t}{Q_t}$ reflects, in expectation, the reduction of loan returns for the bank holding σ_t share of risky projects. The bank receives net income on loans, $R^s_{t+1} - R^d_t\left(1 - \gamma_t\right) - \frac{\xi\sigma_t}{Q_t}$, if it does not default on deposits which happens with probability $1 - G(\varepsilon^*_{t+1})$. If the bank defaults, it gets zero, i.e. $0 \cdot G(\varepsilon^*_{t+1})$ which is not shown in the expression explicitly.

The second counterpart of the above decomposition, ω_2, comprises the extra effect of σ_t on expected dividends that comes from more dispersed returns from projects. In fact, ω_2 is strictly increasing in τ: the bank views projects as a call option the value of which rises with volatility associated with higher upside. Limited liability bounds the payoff to zero in the worst case scenario.

Risk-taking incentives depend on the difference between returns on safe loans and returns on deposits. Table 1 illustrates the effects of greater risk-taking on two components of dividends for each realization of the aggregate returns. We map aggregate returns into states of nature and consider two cases depending on the sign of ε^*_{t+1}. The aggregate returns influence the value of the shield of limited liability. Risk amplifies the effect of the idiosyncratic shock. So, in every state of nature, the bank's choice of risk is determined by the expected effect of the idiosyncratic shock on the value of the shield of limited liability and returns on loans. The up-turn arrow, $\Uparrow$, indicates that greater risk-taking increases the corresponding component of bank's dividends. The down-turn arrow, $\Downarrow$, means that the corresponding component of bank's dividends decreases with greater risk-taking. Two arrows turned in the opposite directions, $\Uparrow\Downarrow$, signify that the effect of greater risk-taking is undetermined and depends the parameterization.

Table 1: Illustrating the Effects of Higher Risk on Dividends.

States of nature where	Effects on ω_1		Effects on ω_2
	$R^l_{t+1} - R^d_t\,(1-\gamma_t) - \frac{\xi\sigma_t}{Q_t}$	$1 - G(\varepsilon^*_{t+1})$	
$R^l_{t+1} < R^d_t\,(1-\gamma_t) \quad \Leftrightarrow \quad \varepsilon^*_{t+1} > 0$	$\Downarrow$	$\Uparrow$	$\Uparrow$
$R^l_{t+1} > R^d_t\,(1-\gamma_t) \quad \Leftrightarrow \quad \varepsilon^*_{t+1} < 0$	$\Downarrow$	$\Downarrow$	if $\varepsilon^*_{t+1} > -\xi$, then $\Uparrow\Downarrow$ if $\varepsilon^*_{t+1} \leqslant -\xi$, then $\Uparrow$

First, $\varepsilon^*_{t+1} > 0$ indicates that the bank makes losses on safe loans. It happens in those states of nature where the net income from the zero-risk portfolio is negative, so the bank is behind the shield of limited liability. By accepting more risk, the bank is more likely to get a positive net return under a favorable realization of the idiosyncratic shock as risk acts like a leverage on the size of the shock. Therefore, $1 - G(\varepsilon^*_{t+1})$ rises. This balances with smaller returns on a portfolio with more risky loans, i.e. $R^s_{t+1} - R^d_t\,(1-\gamma_t) - \frac{\xi\sigma_t}{Q_t}$ goes down. Similarly, gambling on more dispersed returns allows the bank to move away from a zero return that comes from the limited liability to some positive return that is accompanied by less frequent defaults. So, the effect of σ_t on expected dividends from ω_2 is positive.

Second, $\varepsilon^*_{t+1} < 0$ shows that the bank makes positive profits on safe loans. The bank is more likely to default when it takes on more risk because any negative idiosyncratic shock would be amplified by risk. The bank internalizes that riskier projects are less profitable. Therefore, the overall effect of greater risk on ω_1 is negative when $\varepsilon^*_{t+1} < 0$.

Then consider the bonus from projects volatility. If $-\xi < \varepsilon^*_{t+1} < 0$, there are two contrasting forces. On the one hand, the bank always benefits from limited liability that makes the variance of projects returns attractive. On the other hand, the bank is more concerned about (and more vulnerable to) the variability of returns in the situation when taking on more risk would result in zero payoff instead of some positive payoff achieved by smaller risk. It occurs when $-\xi < \varepsilon^*_{t+1} < 0$. In these states of nature, the bank requires greater than average realization of the idiosyncratic shock in order to get a positive return. Call this type of shock a good idiosyncratic shock. This shock happens with probability smaller than 0.5. Define a bad idiosyncratic shock as a complement to a good idiosyncratic shock. An increase in risk increases the profits under a good shock. It captures the benefits from greater upside. At the same time, an increase in risk makes it more likely to get a bad shock. The bank trades off marginal profits coming from a good shock with marginal losses coming from the reduction of profits due to more defaults. Since the probability of the latter is greater than the probability of the former, the losses from defaults can dominate the benefits from greater volatility. This force goes in the opposite direction when $\varepsilon^*_{t+1} \leqslant -\xi$. The difference is that here the bank is more likely to get a good shock than a bad shock. Therefore, the bank puts more weight on the benefits from risk-taking than on its costs. It is verified mathematically that the effects of σ_t on ω_2 is unambiguously positive when $\varepsilon^*_{t+1} \leqslant -\xi$.

In sum, we find that net returns on safe loans, $R^s_{t+1} - R^d_t (1 - \gamma_t)$, is the main driver for the bank's choice of risk. In the partial-equilibrium setting, we differentiate between three cases that characterize incentives for risk-taking.

First, $R^s_{t+1} < R^d_t (1 - \gamma_t)$ applies to the states of nature where a relatively large negative aggregate shock is realized. Two forces against the one that seems to be of lesser relevance make the bank benefit most from taking risk. Second, $-\xi < R^d_t (1 - \gamma_t) - R^s_{t+1} < 0$ applies to the states of nature where intermediate values (not too large and not too small) of either negative or positive aggregate shock are realized. There are more forces that lower incentives for risk. Third, $R^d_t (1 - \gamma_t) - R^s_{t+1} < -\xi$ applies to the states of nature where a positive aggregate shock of a larger size is realized. Interestingly, there is a force associated with the bonus from projects volatility that makes it possible for the bank to increase risk. The choice of risk depends on the strength of that force, ω_2, relative to the negative exposure of returns from a loan portfolio to risk, ω_1. It still remains a quantitative question to find out how risk-taking is determined in the general equilibrium set-up.

Capital requirements affect risk-taking through a change in ε^*_{t+1}. When γ_t increases, ε^*_{t+1} falls. It means that the bank will be more likely to find itself in the states of nature where ε^*_{t+1} is negative. It forces the bank to keep more skin in the game, make the shield of limited liability less attractive and prevent the switch into financing risky projects.

G Calibration of τ

To calibrate the variance of the idiosyncratic shock τ, we link the production function of the risky firm to the production function of the safe firm that has a preexisting debt.

Remember that the next period returns to safe and risky loans are given by

$$\begin{aligned} R^s_{t+1} &= \frac{\alpha A_{t+1}}{Q_t}\left(\frac{K_{t+1}}{H_{t+1}}\right)^{\alpha-1} + (1-\delta)\frac{Q_{t+1}}{Q_t}, \\ R^r_{t+1} &= R^s_{t+1} + \sigma_{\mathrm{RF}}\frac{\varepsilon_{t+1}}{Q_t}, \end{aligned}$$

respectively. The parameter σ_{RF} is needed to distill the exposure of banks (versus other financial intermediaries) to the risk arising in the leveraged loan market. It captures the fact that a certain fraction of leveraged loans is held by the nonbank sector which we do not model here. The risky bank that finances the maximum share of risky projects earns

$$\Omega^{risky}_{t+1} = R^r_{t+1} Q_t K^r_{t+1}.$$

It comprises EBITDA and what the bank makes or loses by selling capital to capital producers. The safe bank with preexisting debt earns

$$\Omega^{safe}_{t+1} = R^s_{t+1} Q_t \left(K_{t+1} + B_t\right) - Q_t B_t R^B_t = \left(R^s_{t+1}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R^B_t\right) Q_t K_{t+1},$$

where B_t is a predetermined debt, measured in units of capital, and R^B_t is a predetermined interest rate. We equate the conditional variances of the returns to loans

$$Var_t\left(R^r_{t+1}\right) = Var_t\left(R^s_{t+1}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R^B_t\right)$$

to find the variance of the idiosyncratic shock that matches $\frac{\text{Debt}}{\text{EBITDA}} = 6$. Note that

$$\begin{aligned} Var_t\left(R^r_{t+1}\right) &= Var_t\left(R^s_{t+1}\right) + \left(\frac{\sigma_{\mathrm{RF}}}{Q_t}\right)^2 \tau^2, \\ Var_t\left(R^s_{t+1}\left(1 + \frac{B_t}{K_{t+1}}\right) - \frac{B_t}{K_{t+1}} R^B_t\right) &= \left(1 + \frac{B_t}{K_{t+1}}\right)^2 Var_t\left(R^s_{t+1}\right), \end{aligned}$$

where K_{t+1} is the steady-state level of capital of the safe firms that are financed by commercial banks and $Q_t = 1$ in the steady state.

The conditional variance of the returns on safe loans is given by

$$\begin{aligned} Var_t\left(R^s_{t+1}\right) = \alpha^2\left(\frac{K_{t+1}}{H_{t+1}}\right)^{2\alpha-2} Var_t\left(A_{t+1}\right) + (1-\delta)^2 Var_t\left(Q_{t+1}\right) + \\ 2\alpha\left(\frac{K_{t+1}}{H_{t+1}}\right)^{\alpha-1}(1-\delta) Cov_t\left(A_{t+1}, Q_{t+1}\right). \end{aligned}$$

We can calculate the conditional variance of Q_{t+1} by picking up its process from the optimization problem of capital producers. However, our approach is meant to be suggestive, and we equate the conditional variances of Q_{t+1} and the aggregate shock. The covariance term is expected to be positive, but we drop it in our calculation because the terms that multiply the covariance are small. The model's counterpart for EBITDA is a total output net of compensation for labor. Thus

$$\frac{\text{Debt}}{\text{EBITDA}} = \frac{B_t}{Y_t^{safe} - W_t H_t^{safe}} = \frac{B_t}{\alpha Y_t^{safe}}.$$

The data analog of σ_{RF} is the share of leveraged loans held by banks (where the remaining fraction is held by nonbanks). We choose $\sigma_{\text{RF}} = 45\%$ from the Shared National Credit Report issued by the Fed, OCC, and FDIC.

H Robustness Checks

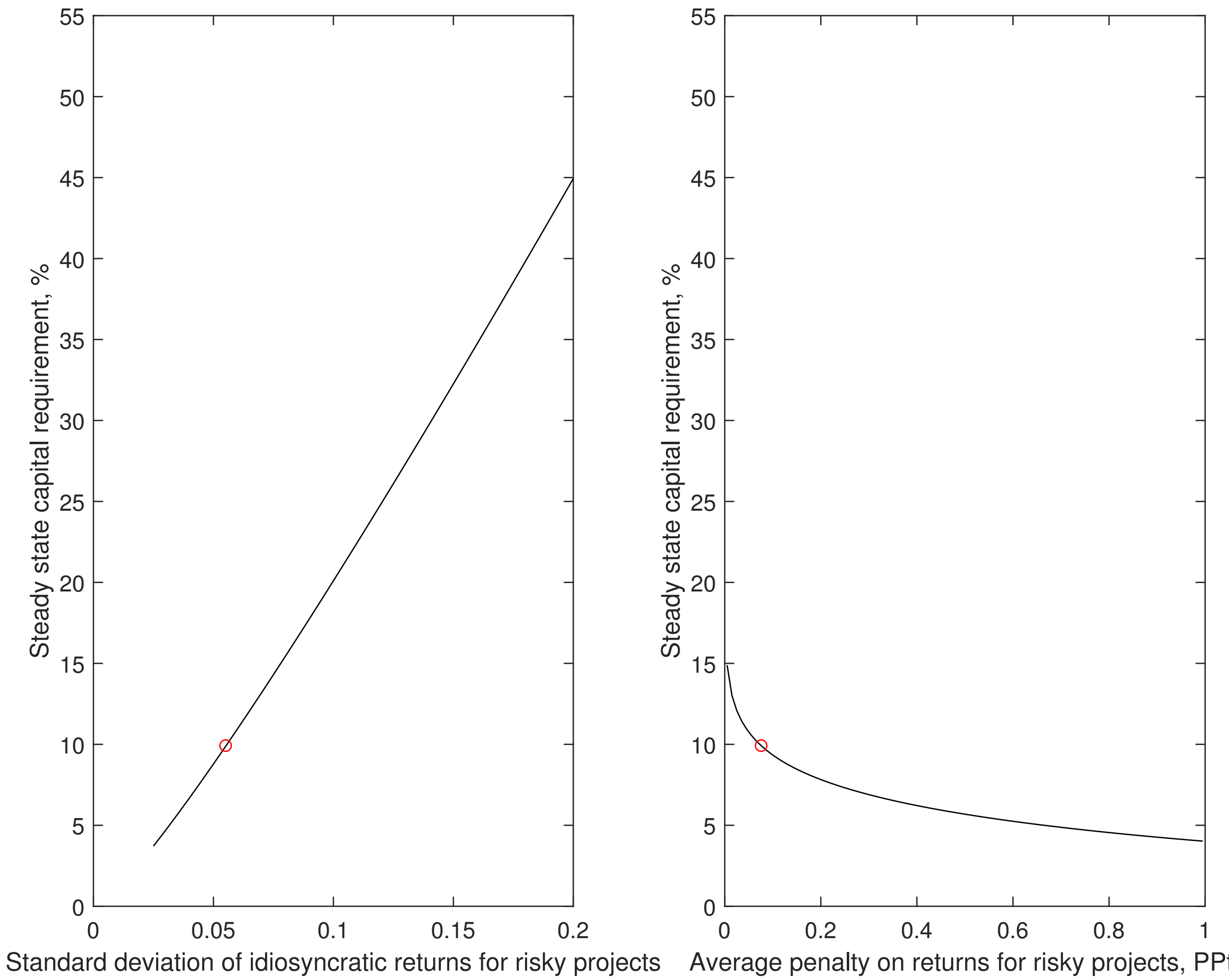

Figure 1: Robustness Checks.