Internet Appendix for "Information Aggregation with Asymmetric Asset Payoffs"

Elias Albagli, Central Bank of Chile
Christian Hellwig, Toulouse School of Economics
Aleh Tsyvinski, Yale University

April 10, 2023

The structure of the appendix is as follows. Section 1 fully characterizes the CARA-normal and CARA-binary models with informed and uninformed traders and provides the proof of proposition 1 in the paper. Section 2 provides further details about the calibration of excess weight on tail risks to forecast dispersion and forecast accuracy. Section 3 discusses a model of belief dispersion in which a noisy sample of individual forecasts is publicly observed, to discuss conditions under which measured dispersion of analyst forecasts remains a valid proxy for the dispersion of private information. Section 4 discusses additional results for section 4 and 5 in the paper: the proof of proposition 5 in the paper, the numerical solution methods for an example with CRRA preferences and binary dividends, and additional limit cases which highlight the amplification of skewness premia through noisy information aggregation. Section 5 discusses additional results for multi-asset extensions with independent or common fundamentals, including an application of our risk-neutral, normal model to security design or departures from the Modigliani-Miller theorem.

1 CARA-normal and CARA-binary models with uninformed traders

Here we analyze the model with CARA preferences and normal or binary dividends, with informed and uninformed traders, the latter slightly generalizing the set-up relative to the main text.

We assume that $\pi(\theta) = \theta$, and that there is a fraction $\kappa > 0$ of traders are informed, the remaining $1 - \kappa$ are uninformed. In the CARA-normal model, we assume that $\theta \sim \mathcal{N}\left(0, \sigma_\theta^2\right)$, in the CARA-binary that $\theta \in \{0, 1\}$, with ex-ante probability $\Pr(\theta = 1) = \lambda \in (0, 1)$.

The informed investors' private signals are normally distributed and centered at θ, $x_i \sim \mathcal{N}(\theta, 1/\beta)$. All other elements are as in the general model (as described in the text).

1.1 CARA-normal Model

Assume that P is informationally equivalent to $z \sim \mathcal{N}\left(\theta, \tau^{-1}\right)$. In the CARA-normal model, the informed traders' demand satisfies $d^I\left(x, P\left(z\right)\right) = \frac{1}{\chi Var(\theta|x,z)}\left(\mathbb{E}\left(\theta|x,z\right) - P\left(z\right)\right)$, where $\mathbb{E}\left(\theta|x,z\right) = \frac{\beta x + \tau z}{\beta+\tau+1/\sigma_\theta^2}$ and $Var\left(\theta|x,z\right) = \left(\beta+\tau+1/\sigma_\theta^2\right)^{-1}$. The uninformed traders' demand satisfies $d^U\left(P\left(z\right)\right) = \frac{1}{\chi Var(\theta|z)}\left(\mathbb{E}\left(\theta|z\right) - P\left(z\right)\right)$, where $\mathbb{E}\left(\theta|z\right) = \gamma z$ and $Var\left(\theta|z\right) = \left(\tau+1/\sigma_\theta^2\right)^{-1}$. Hence, market-clearing requires

$$s = \kappa \int d^I\left(x, P\right) dF\left(x-\theta\right) + \left(1-\kappa\right) d^U\left(P\right) = \frac{\kappa\beta\theta + \tau z}{\chi} - \frac{\kappa\left(\beta+\tau+1/\sigma_\theta^2\right) + \left(1-\kappa\right)\left(\tau+1/\sigma_\theta^2\right)}{\chi}\sigma_\theta^2 P\left(z\right)$$

Hence, the conjecture is satisfied with $z = \theta - \frac{\chi}{\kappa\beta}\left(s-\overline{s}\right)$ and $\tau = \left(\beta\kappa/\chi\right)^2 \cdot \sigma_s^{-2}$. The equilibrium price is

$$P\left(z\right) = \frac{\kappa\beta+\tau}{\kappa\beta+\tau+1/\sigma_\theta^2} z - \frac{\chi}{\kappa\beta+\tau+1/\sigma_\theta^2}\overline{s}$$

or $P\left(z\right) = \hat{\gamma} z - \chi\left(1-\hat{\gamma}\right)\sigma_\theta^2\overline{s}$, where $\hat{\gamma} = \frac{\kappa\beta+\tau}{\kappa\beta+\tau+1/\sigma_\theta^2}$. Hence including uninformed traders in the CARA-normal model mutes the updating wedge by rescaling β by the fraction of informed traders κ; the equilibrium price then responds to z as if it had precision $\kappa\beta+\tau$ instead of τ. All other arguments then carry through replacing β in the model without informed traders by $\beta\kappa$ in the model combining informed and uninformed traders.

1.2 CARA-binary Model

In the CARA normal model, the first-order condition for asset demand yields

$$d\left(\mu, P\right) = \frac{1}{\chi}\left(\log\left(\frac{\mu}{1-\mu}\right) - \log\left(\frac{P}{1-P}\right)\right)$$

where $\mu \in (0,1)$ denotes the investor's posterior belief that $\theta = 1$.

As before, we conjecture that the equilibrium is characterized by a sufficient statistic $z\left(P\right)$ which is distributed according to $z \sim \mathcal{N}\left(\theta, \tau^{-1}\right)$. The uninformed investors' posterior $\mu_U\left(z\right)$ then satisfies

$$\log\left(\frac{\mu_U\left(z\right)}{1-\mu_U\left(z\right)}\right) = \log\left(\frac{\lambda}{1-\lambda}\right) + \tau\left(z - \frac{1}{2}\right),$$

resulting in a demand by the uninformed agents which is given by

$$d_U\left(P\right) = \frac{1}{\chi}\left(\log\left(\frac{\lambda}{1-\lambda}\right) + \tau\left(z-\frac{1}{2}\right) - \log\left(\frac{P}{1-P}\right)\right).$$

The informed investors' posterior $\mu_I\left(x,z\right)$ satisfies

$$\log\left(\frac{\mu_I\left(x,z\right)}{1-\mu_I\left(x,z\right)}\right) = \log\left(\frac{\lambda}{1-\lambda}\right) + \beta\left(x-\frac{1}{2}\right) + \tau\left(z-\frac{1}{2}\right),$$

resulting in a demand by the informed agents which is given by

$$d_I(x,P)=\frac{1}{\chi}\left(\log\left(\frac{\lambda}{1-\lambda}\right)+\tau\left(z-\frac{1}{2}\right)+\beta\left(x-\frac{1}{2}\right)-\log\left(\frac{P}{1-P}\right)\right).$$

Aggregating across agents, we obtain the market-clearing condition

$$s=\frac{1}{\chi}\left(\log\left(\frac{\lambda}{1-\lambda}\right)-\log\left(\frac{P}{1-P}\right)+\tau\left(z-\frac{1}{2}\right)\right)+\frac{\beta\kappa}{\chi}\left(\pi-\frac{1}{2}\right),$$

for $\pi\in\{0,1\}$. Solving for P, we obtain the market-clearing price

$$P(z)=\frac{\lambda e^{(\beta\kappa+\tau)\left(z-\frac{1}{2}\right)-\chi\bar{s}}}{\lambda e^{(\beta\kappa+\tau)\left(z-\frac{1}{2}\right)-\chi\bar{s}}+1-\lambda},$$

where $z=\pi-\frac{\chi}{\kappa\beta}\cdot(s-\bar{s})$ and $\tau=(\beta\kappa/\chi)^2\cdot\sigma_s^{-2}$, confirming our initial conjecture. The log-odds ratio implied by the price attributes a weight $\tau+\beta\kappa$ to the market signal z and includes a risk adjustment $-\chi\bar{s}$ to compensate investors for their expected exposure. As in the CARA-normal model, this equilibrium price corresponds to the risk-adjusted expectation of dividends of a hypothetical investor who treats the market signal z as if it had precision $\beta\kappa+\tau$ and takes on a position equal to the average supply $\bar{s}$. These equations generalize the equilibrium characterization to the case with informed and uninformed traders.

1.3 Proof of proposition 1 in the main text

Let $\hat{\lambda}=\lambda e^{-\chi\bar{s}}/\left(\lambda e^{-\chi\bar{s}}+1-\lambda\right)$ denote a risk-adjusted prior that $\pi=1$. The following proposition generalizes proposition 1 from the paper to arbitrary values of $\bar{s}$ and $\kappa\in[0,1]$:

Proposition 1 : *There exist positive numbers $\Delta\in(0,1)$ and $R\in(0,1)$ such that the expected price takes the form*

$$\mathbb{E}(P(z))=\lambda+\left(\frac{1}{2}-\lambda\right)\Delta-\left(\lambda-\hat{\lambda}\right)R.$$

Hence we can decompose the difference between expected price and expected dividend into a risk premium R and an adjustment Δ that captures the role of excess weight on tail risks.

Proof of Proposition 1

Define $\hat{\lambda}=\lambda/\left(\lambda+(1-\lambda)e^{\chi\bar{s}}\right)$, $x=(\beta\kappa+\tau)\left(\frac{1}{2}+\frac{\chi}{\beta\kappa}(s-\bar{s})\right)$, $\hat{x}=\frac{1}{2}(\beta\kappa+\tau)$, $\sigma_x^2=(\beta\kappa+\tau)^2/\tau$, and $\psi=\tau/(\beta\kappa+\tau)$. We show that $\mathbb{E}(P(z))=\lambda+(1-2\lambda)\Delta-\left(\lambda-\hat{\lambda}\right)R$, where

$$\Delta=\int_{-\infty}^{\infty}\left\{\frac{\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{-x}-1\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}\right\}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)$$

$$R=\int_{-\infty}^{\infty}\frac{1}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right).$$

To prove this result, write $P(z)$ as

$$P(\pi,x)=\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x-2\hat{x}\pi}}$$

where $x\sim\mathcal{N}\left(\hat{x},\sigma_x^2\right)$ and $\pi=1$ w.p. λ. Taking expectations, we obtain

$$\begin{aligned}\mathbb{E}\left(P\left(z\right)\right) &= \lambda\int_{-\infty}^{\infty}\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x-2\hat{x}}}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)+\left(1-\lambda\right)\int_{-\infty}^{\infty}\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)\\ &= \hat{\lambda}\int_{-\infty}^{\infty}\left\{\frac{\lambda}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}+\frac{1-\lambda}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}\right\}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)\\ &= \hat{\lambda}\int_{-\infty}^{\infty}\left\{\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}+\frac{1-\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}\right\}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)\\ &\quad+\int_{-\infty}^{\infty}\left\{\frac{\hat{\lambda}\left(\lambda-\hat{\lambda}\right)}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}-\frac{\hat{\lambda}\left(\lambda-\hat{\lambda}\right)}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}\right\}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right).\end{aligned}$$

Here, the first equality exploits the symmetry of $\Phi(\cdot)$ around 0 before changing variables $x'=2\hat{x}-x$:

$$\int_{-\infty}^{\infty}\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x-2\hat{x}}}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)=\int_{-\infty}^{\infty}\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x-2\hat{x}}}d\Phi\left(\frac{\hat{x}-x}{\sigma_x}\right)=\int_{-\infty}^{\infty}\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x'}}d\Phi\left(\frac{x'-\hat{x}}{\sigma_x}\right)$$

Since $\phi\left(\frac{x-\hat{x}}{\sigma_x}\right)/\phi\left(\frac{-x-\hat{x}}{\sigma_x}\right)=e^{2x\hat{x}/\sigma_x^2}=e^{\psi x}$, the first line in the expression for $\mathbb{E}\left(P\left(z\right)\right)$ can be written as

$$\begin{aligned}&\hat{\lambda}\int_{-\infty}^{\infty}\left\{\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}+\frac{1-\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}\right\}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)\\ =\;&\hat{\lambda}\int_{-\infty}^{\infty}\frac{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-\psi x}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)=\hat{\lambda}+\hat{\lambda}\left(1-\hat{\lambda}\right)\int_{-\infty}^{\infty}\frac{e^{-\psi x}-e^{-x}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right).\end{aligned}$$

Splitting the integral at 0, we have

$$\begin{aligned}\int_{-\infty}^{\infty}\frac{e^{-\psi x}-e^{-x}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) &= \int_{0}^{\infty}\left\{\frac{e^{-\psi x}-e^{-x}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}+\frac{\left(e^{\psi x}-e^{x}\right)e^{-\psi x}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}\right\}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)\\ &= \int_{0}^{\infty}\left\{\frac{\left(e^{(1-\psi)x}-1\right)e^{-x}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}-\frac{e^{(1-\psi)x}-1}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}\right\}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)\end{aligned}$$

Since

$$\frac{e^{-x}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}-\frac{1}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}=\frac{\left(1-e^{-x}\right)\left(1-2\hat{\lambda}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)},$$

the first integral then takes the form $\hat{\lambda}+\left(\frac{1}{2}-\hat{\lambda}\right)\Delta$, where

$$\Delta=2\int_{0}^{\infty}\frac{\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{(1-\psi)x}-1\right)\left(1-e^{-x}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)=2\int_{-\infty}^{\infty}\frac{\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{-x}-1\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right).$$

Since $\left(e^{(1-\psi)x}-1\right)\left(1-e^{-x}\right)>0$ for $x>0$, it follows that $\Delta>0$. In addition, since $1>\hat{\lambda}\left(1-\hat{\lambda}\right)>\hat{\lambda}\left(1-\hat{\lambda}\right)\left(1-e^{x}\right)$ for all $x<0$, it follows that

$$\Delta\leq 2\int_{-\infty}^{0}\frac{\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{-x}-1\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)\leq 2\int_{-\infty}^{0}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)=2\Phi\left(\frac{-\hat{x}}{\sigma_x}\right)\leq 1.$$

To compute the second integral, note that

$$g\left(x\right)\equiv\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{-x}}-\frac{\hat{\lambda}}{\hat{\lambda}+\left(1-\hat{\lambda}\right)e^{x}}=\frac{\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}-e^{-x}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}.$$

Therefore we obtain $\mathbb{E}\left(P\left(z\right)\right)=\lambda+\left(\frac{1}{2}-\lambda\right)\Delta-\left(\lambda-\hat{\lambda}\right)R$, where

$$R=1-\Delta-\int_{-\infty}^{\infty}g\left(x\right)d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)=\int_{-\infty}^{\infty}\frac{1}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)<1.$$

1.4 Convergence as $\beta\kappa\to 0$

We also compare the binary model with dispersed information ($\beta\kappa>0$) to its counterpart where $\beta\kappa=0$, where all investors are uninformed and take identical positions to absorb supply shocks. Taking the limit as $\beta\kappa\to 0$ corresponds to the equilibrium price in the benchmark model where all investors are uninformed:

$$\lim_{\beta\kappa\to 0}P\left(z\right)=\frac{\lambda e^{-\chi s}}{\lambda e^{-\chi s}+1-\lambda}.$$

We contrast this limit with the alternative in which both $\beta\kappa\to 0$ and $\chi\sigma_s\to 0$, holding $\tau=\sqrt{\frac{\beta\kappa}{\chi\sigma_s}}$ constant, in which case

$$\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}}P\left(z\right)=\frac{\lambda e^{\tau\left(z-\frac{1}{2}\right)-\chi\bar{s}}}{\lambda e^{\tau\left(z-\frac{1}{2}\right)-\chi\bar{s}}+1-\lambda}.$$

Proposition 2 *: (i) Limit as $\beta\kappa \to 0$, holding $\chi\sigma_s$ constant:* $\lim_{\beta\kappa\to 0}\Delta = \hat{\lambda}\left(1-\hat{\lambda}\right)(\chi\sigma_s)^2 + o\left((\chi\sigma_s)^4\right)$ *and*

$$\lim_{\beta\kappa\to 0}\mathbb{E}\left(P\left(z\right)\right) = \hat{\lambda} + \left(\frac{1}{2}-\hat{\lambda}\right)\hat{\lambda}\left(1-\hat{\lambda}\right)(\chi\sigma_s)^2 + o\left((\chi\sigma_s)^4\right).$$

(ii) Limit as $\beta\kappa \to 0$ and $\chi\sigma_s \to 0$, holding $\tau = \sqrt{\frac{\beta\kappa}{\chi\sigma_s}}$ constant: There exist functions $\Delta(\tau)$ and $R(\tau)$, such that $\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}}\frac{\Delta}{\beta\kappa} = \Delta(\tau)$ and $\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}} R = R(\tau)$. Moreover, for small τ, $R(\tau) = 1-\hat{\lambda}\left(1-\hat{\lambda}\right)\tau + o\left(\tau^2\right)$ and $\Delta(\tau) = \hat{\lambda}\left(1-\hat{\lambda}\right) + o(\tau)$.

Proposition 2 shows that when supply shocks are small, skewness has much larger effects on average prices with dispersed information (part ii) than at the common information benchmark (part i). In the common information benchmark, the skewness premium is proportional to $\chi^2\sigma_s^2$. In the dispersed information economy the skewness premium is proportional to $\chi\sigma_s$. Dispersed information therefore amplifies the effects of skewness on expected prices and returns. Below we argue this result applies very generally and is the consequence of two different sources of skewness premia, with the former driven by investor preferences (downside risk aversion) and the latter driven by information, or co-movement between supply shocks and investor expectations. The risk adjustment reflects the information aggregated through the share price, hence an increase in $\overline{s}$ is passed through to prices by less than the full adjustment to the risk-adjusted prior $\hat{\lambda}$, except when prices become fully uninformative.

Proof of Proposition 2

Part (i): Taking the limit as $\beta\kappa \to 0$ while holding σ_s constant yields $\lim_{\beta\kappa\to 0}\psi = \lim_{\beta\kappa\to 0}\tau = \lim_{\beta\kappa\to 0}\hat{x} = 0$ and $\lim_{\beta\kappa\to 0}\sigma_x = \chi\sigma_s$. Therefore

$$\lim_{\beta\kappa\to 0}\Delta = 2\hat{\lambda}\left(1-\hat{\lambda}\right)\int_{-\infty}^{\infty}\frac{e^{-x}-1}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x}{\chi\sigma_s}\right)$$

and

$$\lim_{\beta\kappa\to 0}R = \int_{-\infty}^{\infty}\frac{1}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x}{\chi\sigma_s}\right)$$

Using a second order expansion for $\frac{e^{-x}-1}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}$ and $\frac{1}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}$ around $x=0$ yields

$$\lim_{\beta\kappa\to 0}\Delta = \hat{\lambda}\left(1-\hat{\lambda}\right)(\chi\sigma_s)^2 + o\left((\chi\sigma_s)^4\right)$$

and

$$\lim_{\beta\kappa\to 0} R = 1-\hat{\lambda}\left(1-\hat{\lambda}\right)(\chi\sigma_s)^2+o\left((\chi\sigma_s)^4\right)$$

where $o(\cdot)$ denotes terms of higher orders. Substituting these two into the expression for $\mathbb{E}(P(z))$ yields

$$\lim_{\beta\kappa\to 0}\mathbb{E}(P(z)) = \hat{\lambda}+\left(\frac{1}{2}-\hat{\lambda}\right)\hat{\lambda}\left(1-\hat{\lambda}\right)(\chi\sigma_s)^2+o\left((\chi\sigma_s)^4\right).$$

Part (*ii*): After substituting for $\hat{x}$ and σ_x, we write

$$\begin{aligned}
\Delta &= 2\int_{-\infty}^{\infty}\frac{\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{-x}-1\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right)\\
&= 2\hat{\lambda}\left(1-\hat{\lambda}\right)e^{-\frac{\tau}{8}}\int_{-\infty}^{\infty}\frac{e^{\frac{\psi}{2}x}\left(e^{-x}-1\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x}{\sigma_x}\right)\\
&= 2\;\hat{\lambda}\left(1-\hat{\lambda}\right)e^{-\frac{\tau}{8}}\int_{-\infty}^{\infty}\frac{e^{\frac{\psi-1}{2}x}\left(e^{-\frac{1}{2}x}-e^{\frac{1}{2}x}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x}{\sigma_x}\right)\\
&= 2\hat{\lambda}\left(1-\hat{\lambda}\right)e^{-\frac{\tau}{8}}\int_{-\infty}^{\infty}\frac{\left(e^{\frac{\psi-1}{2}x}-1\right)\left(e^{-\frac{1}{2}x}-e^{\frac{1}{2}x}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x}{\sigma_x}\right)\\
&= (1-\psi)\,\hat{\lambda}\left(1-\hat{\lambda}\right)e^{-\frac{\tau}{8}}\int_{-\infty}^{\infty}\frac{\frac{2}{1-\psi}\left(e^{\frac{\psi-1}{2}x}-1\right)\left(e^{-\frac{1}{2}x}-e^{\frac{1}{2}x}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x}{\sigma_x}\right)
\end{aligned}$$

Here the second line uses the fact that $\phi\left(\frac{x-\hat{x}}{\sigma_x}\right)=\phi\left(\frac{x}{\sigma_x}\right)e^{\frac{\psi}{2}x-\frac{\tau}{8}}$, while the fourth line uses the fact that $\int_{-\infty}^{\infty}\frac{e^{-\frac{1}{2}x}-e^{\frac{1}{2}x}}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x}{\sqrt{\tau}}\right)=0$ since $\frac{e^{-\frac{1}{2}x}-e^{\frac{1}{2}x}}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}$ is symmetric around $x=0$.

Taking the limit as $\beta\kappa\to 0$ while holding $\tau=\frac{\beta\kappa}{\chi\sigma_s}$ constant yields $\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}}\sigma_x^2=\tau$, $\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}}\psi=1$, and $\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}}\frac{1-\psi}{\beta\kappa}=\frac{1}{\tau}$. By L'Hôpital's Rule, $lim_{\psi\to 1}\frac{2}{1-\psi}\left(e^{\frac{\psi-1}{2}x}-1\right)=-x$. It then follows that

$$\begin{aligned}
\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}}\frac{\Delta}{\beta\kappa} &= \frac{1}{\tau}\hat{\lambda}\left(1-\hat{\lambda}\right)e^{-\frac{\tau}{8}}\int_{-\infty}^{\infty}\frac{x\left(e^{\frac{1}{2}x}-e^{-\frac{1}{2}x}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)}d\Phi\left(\frac{x}{\sqrt{\tau}}\right)\\
&= \frac{1}{\tau}\hat{\lambda}\left(1-\hat{\lambda}\right)e^{-\frac{\tau}{8}}\int_{-\infty}^{\infty}\frac{\sqrt{\tau}v\left(e^{\frac{1}{2}\sqrt{\tau}v}-e^{-\frac{1}{2}\sqrt{\tau}v}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{\sqrt{\tau}v}+e^{-\sqrt{\tau}v}-2\right)}d\Phi(v)
\end{aligned}$$

Which in turn implies that $\Delta \sim \beta\kappa$. A second-order Taylor expansion around $x = 0$ yields $\frac{x\left(e^{\frac{1}{2}x}-e^{-\frac{1}{2}x}\right)}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)(e^{x}+e^{-x}-2)} = x^2 + o\left(x^4\right)$, and therefore

$$\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}} \frac{\Delta}{\beta\kappa} = \hat{\lambda}\left(1-\hat{\lambda}\right) + o\left(\tau\right)$$

for τ sufficiently small.

Following the same steps for R yields

$$R = \int_{-\infty}^{\infty} \frac{1}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)} d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) = e^{-\frac{\tau}{8}} \int_{-\infty}^{\infty} \frac{e^{\frac{\psi}{2}x}}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)} d\Phi\left(\frac{x}{\sigma_x}\right)$$

and

$$\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}} R = e^{-\frac{\tau}{8}} \int_{-\infty}^{\infty} \frac{e^{\frac{1}{2}x}}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)\left(e^{x}+e^{-x}-2\right)} d\Phi\left(\frac{x}{\sqrt{\tau}}\right)$$

A second-order Taylor expansion around $x = 0$ yields $\frac{e^{\frac{1}{2}x}}{1+\hat{\lambda}\left(1-\hat{\lambda}\right)(e^{x}+e^{-x}-2)} = 1+\frac{1}{2}x+\left(\frac{1}{8}-\hat{\lambda}\left(1-\hat{\lambda}\right)\right)x^2+ o\left(x^3\right)$, and therefore $\lim_{\beta\kappa\to 0,\chi\sigma_s=\frac{\beta\kappa}{\sqrt{\tau}}} R = 1-\hat{\lambda}\left(1-\hat{\lambda}\right)\tau + o\left(\tau^2\right)$.

2 Calibration of excess weight on tail risks based on analyst earning forecasts

As a first step towards quantifying the potential impact of noisy information aggregation on asset returns, we need to construct empirical measures for forecast dispersion and excess weight on tail risks. We infer excess weight on tail risks from measures of forecast dispersion $\tilde{D}$ and forecast accuracy $\widehat{\gamma}$ for a cross-section of listed firms, using representation

$$\frac{\widehat{\sigma}_{\theta}}{\sigma_{\theta}} = \sqrt{1+\tilde{D}^2 \frac{\widehat{\gamma}\left(1-\widehat{\gamma}\right)}{\widehat{\gamma}\left(1-\widehat{\gamma}\right)-\tilde{D}^2}} \qquad (1)$$

that we derived in the main text.

The I.B.E.S. data base of analyst forecasts of earnings per share reports a consensus or average earnings forecast $\bar{f}_{it}$, realized earnings per share θ_{it} and the cross-sectional standard deviation of individual earnings forecasts D_{it}, for each firm-year in the sample, $\left\{D_{it}, \bar{f}_{it}, \theta_{it}\right\}$. Using these measures, we obtain a forecast error for the consensus forecast in each year, $e_{it} = \theta_{it} - \bar{f}_{it}$, and if the sample of individual forecasters is sufficiently large, we can estimate the individual forecast error variance as $Var\left(e_{it}^2\right) + \mathbb{E}\left(D_{it}^2\right)$. We can also estimate the variance of realized earnings per share, $Var\left(\theta_{it}\right)$ along with the time-series average forecast dispersion, $\mathbb{E}\left(D_{it}^2\right)$. Together these estimates allows to construct firm-level estimates of forecast accuracy $\widehat{\gamma} = 1 - \left(Var\left(e_{it}^2\right) + \mathbb{E}\left(D_{it}^2\right)\right)/Var\left(\theta_{it}\right)$ and normalized forecast dispersion $\tilde{D} = \sqrt{\mathbb{E}\left(D_{it}^2\right)/Var\left(\theta_{it}\right)}$.

Our main sample focuses on 6820 firms used in the empirical study by Guntay and Hackbarth (2013, henceforth GH), which uses forecasts over relatively short horizons (within quarter) from 1987-1998. As a robustness check we replicate our estimates on a second sample of earnings forecast data from Straub and Ulbricht (2023, henceforth SU), who use the entire I.B.E.S. sample (1976-2016) and forecasts over a longer 8 month horizon. We further restrict our sample to a subset of 2101 firms for which we have at least 10 years of forecast data.

One concern with our approach is that the forecast accuracy and dispersion measures we construct are noisy estimates of the true underlying dispersion and accuracy at the level of each firm. If the number of analysts is small, or the time series of forecasts is short, then our approach may over-estimate the true extent of cross-sectional variation in forecast dispersion and information frictions due to estimation noise. For example, for 13.5% of the sample of firms in GH, the estimated forecast accuracy is negative, meaning that the forecasters' posteriors are systematically more noisy than the unconditional volatility of earnings, a feature inconsistent with *any* model of Bayes-consistent belief updating. A similar fraction of firms has $\tilde{D} > 0.5$, which is in violation of the common prior assumption, and more generally, 23.5% of firms in the sample violate the parameter restriction $\widehat{\gamma}\left(1 - \widehat{\gamma}\right) > \tilde{D}^2$ which is implied by the common prior assumption in the model. This occurs mostly for firms that have high estimated levels of forecast dispersion, and low levels of forecast accuracy. For these 23.5% of firms, $\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$ is not well defined.[1] These violations of the relevant parameter conditions could either be interpreted as evidence of measurement error or point us to a rejection of the common prior assumption underlying our model.

The alternative SU sample, which we consider as a robustness check, gives some support to the hypothesis that measurement error is a concern. We restrict ourselves to firms that have been followed by earnings analysts for at least 10 years, which gives us a sufficiently long time series to estimate forecast accuracy with a higher level of precision. These firms are also likely to be followed by a larger number of analysts., which ceteris paribus increases the precision of the dispersion estimate. The fraction of firms that violate the common prior restriction drops from 23.5 to 13%, the fraction of firms with negative forecast accuracy from 13.5 to 9.6%, and the fraction

[1]The initial GH sample contained 6820 firms. 970 of these firms have $\tilde{D} > 0.5$, 905 of these firms have $\widehat{\gamma} < 0$, and for a total of 1577 of these firms, we find $\widehat{\gamma}\left(1 - \widehat{\gamma}\right) < \tilde{D}^2$. In addition, 92 firms in our sample have a reported forecast dispersion of $\tilde{D} = 0$, suggesting that they were followed by only one forecaster, and forecast dispersion is thus not identified.

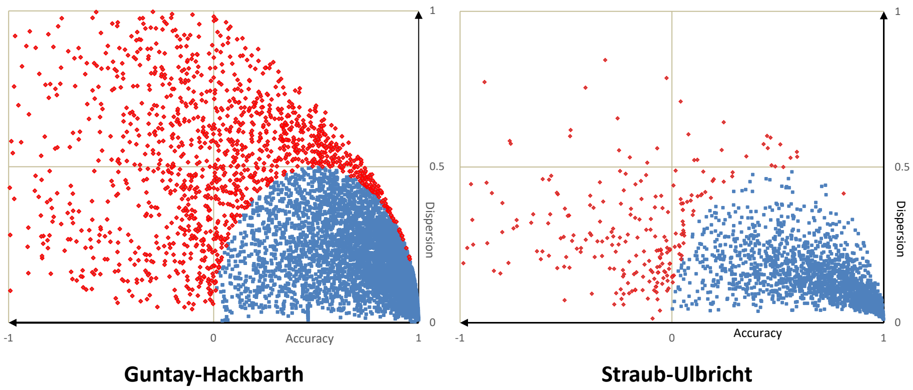

Figure 1: This figure displays scatter plots of Forecast Accuracy (horizontal axis) against Forecast Accuracy (vertical axis) for the GH and SU data sets. The blue ▪-marks correspond to firms that satisfy the parameter restriction $\hat{\gamma}(1-\hat{\gamma}) > \tilde{D}^2$. Red ◊-marks correspond to firms that do not satisfy this parameter restriction. Domains are restricted to [-1,1] and [0,1] respectively.

Figure 1: Forecast Dispersion vs. Accuracy (Scatter Plots)

of firms with $\tilde{D} > 0.5$ from 14.2% to 2.2%.

Figure 1 shows scatter plots of forecast accuracy (horizontal axis) and forecast dispersion (vertical axis) in the two data sets, with negative forecast accuracy ($\widehat{\gamma} < 0$) corresponding to a case where the forecast error variance exceeds the volatility of earnings. The blue ■-marks correspond to firms that satisfy $\widehat{\gamma}(1-\widehat{\gamma}) > \tilde{D}^2$, the red ♦-marks correspond to firms that do not satisfy this restriction. In both samples, the majority of firms lie in an area with high forecast accuracy and low forecast dispersion, which corresponds to low levels of information friction in the market. But there is a non-negligible subset of firms with less accurate earnings forecasts and higher levels of forecast dispersion, even among those firms for which $\widehat{\gamma}(1-\widehat{\gamma}) > \tilde{D}^2$.

We wish to construct measures of EWTR $\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$ for as large as possible a sample of firms, which can be done exactly only for the firms for whom $\widehat{\gamma}(1-\widehat{\gamma}) > \tilde{D}^2$. There are then two ways to incorporate this parameter restriction. The first is to simply exclude all the firms for whom $\widehat{\gamma}(1-\widehat{\gamma}) \leq \tilde{D}^2$ and obtain estimates of $\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$ for the remainder. A second, more conservative approach consists in excluding only the firms for whom $\tilde{D} \geq 0.5$ which are not consistent with our model for any value of $\widehat{\gamma}$. For the remaining firms we can then estimate a lower bound on $\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$ using only the data on forecast dispersion by setting $\widehat{\gamma} = 0.5$. This will yield a lower bound on the implied EWTR and

Table 1a	Guntay-Hackbarth	N	Mean	St. Dev.	10%	30%	Median	70%	90%
Forecast Dispersion ($\tilde{D}$)	Sample I	6820	**0.289**	**0.275**	0.062	0.134	**0.214**	0.324	0.598
	Sample II	6728	**0.293**	**0.275**	0.067	0.137	**0.217**	0.327	0.603
	Sample III	5758	**0.206**	**0.120**	0.061	0.123	**0.186**	0.264	0.388
	Sample IV	5153	**0.192**	**0.111**	0.058	0.116	**0.173**	0.246	0.359
Forecast Accuracy ($\hat{\gamma}$)	Sample I	6820	**0.466**	**1.068**	-0.154	0.460	**0.723**	0.880	0.969
	Sample II	6728	**0.462**	**1.074**	-0.161	0.454	**0.720**	0.879	0.968
	Sample III	5758	**0.651**	**0.485**	0.172	0.599	**0.784**	0.902	0.972
	Sample IV	5153	**0.747**	**0.226**	0.402	0.667	**0.814**	0.915	0.975
Table 1b	**Straub-Ulbricht**	**N**	**Mean**	**St. Dev.**	**10%**	**30%**	**Median**	**70%**	**90%**
Forecast Dispersion ($\tilde{D}$)	Sample I	2101	**0.167**	**0.126**	0.053	0.094	**0.137**	0.194	0.315
	Sample III	2054	**0.156**	**0.096**	0.053	0.093	**0.134**	0.189	0.290
	Sample IV	1839	**0.143**	**0.084**	0.051	0.088	**0.126**	0.175	0.258
Forecast Accuracy ($\hat{\gamma}$)	Sample I	2101	**0.559**	**0.463**	0.014	0.482	**0.695**	0.843	0.954
	Sample III	2054	**0.586**	**0.402**	0.053	0.497	**0.703**	0.845	0.955
	Sample IV	1839	**0.689**	**0.240**	0.322	0.577	**0.745**	0.865	0.959

Table 1: Forecast dispersion and forecast accuracy (summary statistics)

return premia for a larger sample than the first approach. Below we explore both alternatives, which yield fairly similar results.

Tables 2a and b report summary statistics (mean, standard deviation, median, 10th, 30th, 70th and 90th percentiles) for forecast dispersion and forecast accuracy in the two data sets with different sample restrictions. The GH sample came in two versions, one with raw accuracy and dispersion measures, and one based on windsorized data. We report here the moments for the windsorized data, as the procedure appears to mitigate (slightly) the concerns about measurement error in forecast accuracy. Tables for the raw data moments are very similar and available in the online repository.[2] Sample I includes all 6820 firms. In Sample II (6728 firms), we exclude firms for which the reported forecast dispersion $\tilde{D}$ is 0: these firms appear to be followed by a single analyst, which makes it impossible to properly define and measure forecast dispersion in the data. As shown in

[2]The biggest difference between raw and windsorized data moments is in the accuracy measures for the lower end of the distribution in the full sample, where concerns about measurement are the largest. All other moments are virtually identical in the two samples.

the table, the correction induced by removing firms with 0 forecast dispersion is very minor. In Sample III (5758 firms), we remove firms with $\tilde{D} = 0$ and with $\tilde{D} > 0$. This is the sample we will use to construct the lower bound measure on EWTR setting $\widehat{\gamma} = 0$. Forecast dispersion is lower and forecast accuracy higher than in the full sample. Finally, in Sample IV (5153 firms), we remove all firms with $\tilde{D} = 0$ and with $\widehat{\gamma}(1 - \widehat{\gamma}) < \tilde{D}^2$ to use both measures of forecast dispersion and forecast accuracy to estimate EWTR. This sample further reduces forecast dispersion and increases forecast accuracy, consistent with the additional moment condition. Table 1b reports the same statistics for the SU data, but here we only consider samples I (2101 firms), III (2054 firms) and IV (1839 firms), since the original sample did not include any firm for which $\tilde{D} = 0$. The shifts in moments induced by the sample selection criteria are less pronounced in the SU sample, which is consistent with our interpretation that the measures in SU are less noisy. Overall, forecast dispersion and forecast accuracy are lower in SU than in GH, which may be due to the different samples of firms, but also the longer forecast horizons that are considered.

The distributions of forecast dispersion and forecast accuracy are highly skewed in all the different version of the two samples: for most firms, forecast dispersion is low and forecast accuracy is fairly high. However, forecast dispersion can be significant in the top quintile of the distribution: in all the different samples, forecast dispersion at the 90th percentile is more than twice as high as at the median. The picture is reversed for forecast accuracy, which displays a fat lower tail of firms for which survey forecasts have low and even negative accuracy. Figures 2 and 3 confirm these observations visually by showing the CDFs of forecast dispersion and forecast accuracy in the two samples (forsample II, i.e. without $\tilde{D} = 0$ firms in the case of GH, and for sample I in the case of SU).

Table 2a and 2b report summary statistics for firm-level measures of excess weight on tail risks constructed using the data on forecast accuracy and forecast dispersion and equation (1). First, using Sample III which excludes all firms for which $\tilde{D} \geq 0.5$, we construct a "lower bound" measure of EWTR $\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$ by setting $\widehat{\gamma} = 0.5$. Second, using Sample IV which excludes all firms for which $\widehat{\gamma}(1 - \widehat{\gamma}) \leq \tilde{D}^2$, we impute $\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$ using the reported distributions of $\widehat{\gamma}$ and $\tilde{D}$. For the latter sample, we also report statistics on the lower bound to provide some insight into the respective strengths of sample selection going from Sample III to Sample IV, and the role of forecast accuracy in generating higher levels of EWTR.

Our two measures of EWTR are even more skewed than the measures of forecast dispersion: EWTR is small for the vast majority of firms, i.e. less than 5% for roughly 70% of the GH sample and less than 2% for roughly 70% of the SU sample. On the other hand, it can become very

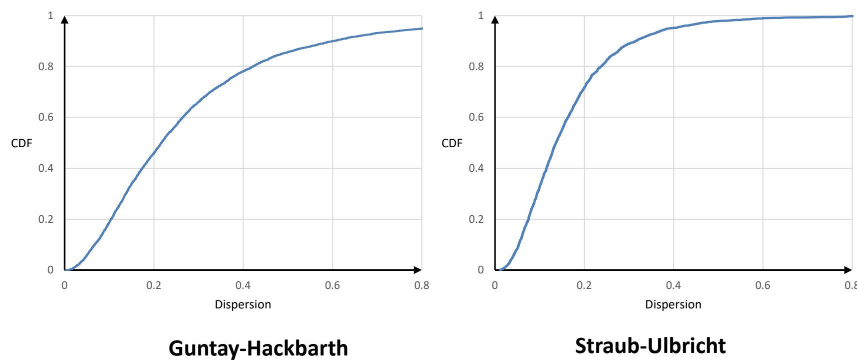

Figure 2: This figure displays cdf of forecast dispersion (horizontal axis), in the GH (sample II, without $\widetilde{D} = 0$) and SU (sample I).

Figure 2: Forecast Dispersion (Cumulative Distribution)

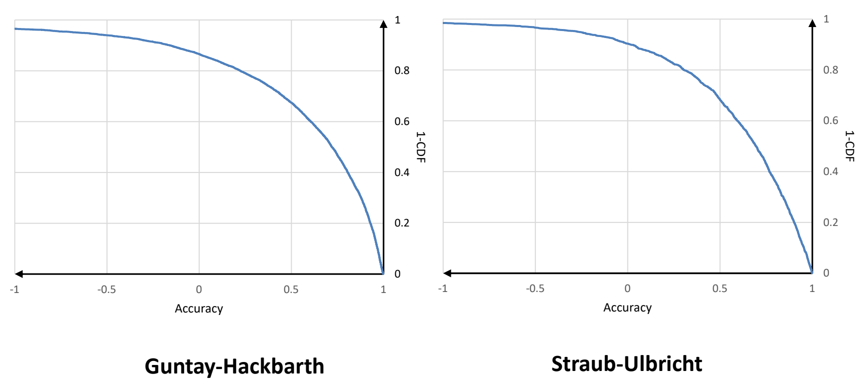

Figure 3: This figure displays 1-cdf of forecast accuracy (horizontal axis), in the GH (sample II, without $\widetilde{D} = 0$) and SU (sample I).

Figure 3: Forecast Accuracy (Cumulative Distribution)

Table 2a	Guntay-Hackbarth	N	Mean	St. Dev.	10%	30%	Median	70%	90%
Lower bound ($\frac{\widehat{\sigma}_\theta}{\sigma_\theta}\Big\vert_{\widehat{\gamma}=0.5}$)	**Sample III**	5758	**1.087**	**0.331**	1.002	1.008	**1.02**	1.047	1.175
	Sample IV	5153	**1.055**	**0.153**	1.002	1.007	**1.017**	1.039	1.126
EWTR ($\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$)	**Sample IV**	5153	**1.102**	**0.493**	1.002	1.009	**1.023**	1.055	1.195
Table 2b	**Straub-Ulbricht**	**N**	**Mean**	**St. Dev.**	**10%**	**30%**	**Median**	**70%**	**90%**
Lower bound ($\frac{\widehat{\sigma}_\theta}{\sigma_\theta}\Big\vert_{\widehat{\gamma}=0.5}$)	**Sample III**	2054	**1.033**	**0.105**	1.001	1.004	**1.010**	1.021	1.061
	Sample IV	1839	**1.021**	**0.054**	1.001	1.004	**1.008**	1.017	1.044
EWTR ($\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$)	**Sample IV**	1839	**1.027**	**0.097**	1.001	1.004	**1.009**	1.018	1.051

Table 2: Summary statistics for excess weight on tail risks

substantial in the top quintile, with EWTR at the 90th percentile of the distribution roughly 8 to 10 times as large as at the median in GH, and around 5 times as large in SU. These results support the qualitative conclusion that EWTR is concentrated in the top quintile of firms with the highest forecast disagreement. Moreover, these ratios do not vary much across the different measures and samples that we considered (Sample III and Sample IV, EWTR and the lower bound implied by setting $\widehat{\gamma} = 0.5$), and the distribution of EWTR estimates from the more restrictive sample IV appears to be very similar to the distribution of the lower bound estimate from sample III.

Figures 4 and 5 illustrate these observations by plotting the cdf for $\frac{\widehat{\sigma}_\theta}{\sigma_\theta}$ and the lower bound measure. These figures confirms that the distribution of excess weight on tail risks is highly right-skewed in both samples: frictions are small for most firms, but the mean level of frictions is much higher than the median and driven by an upper tail of firms for which measured information frictions can be fairly substantial.

Table 2 reports sample correlations between forecast dispersion, forecast accuracy, excess weight on tail risks and the lower bound for EWTR in the different samples.Forecast dispersion and accuracy are negatively correlated in the cross-section. Excess weight on tail risks is weakly negatively correlated with forecast accuracy, but more strongly positively correlated with forecast dispersion. Moreover, this correlation is higher in the less noisy SU sample, and it increases if we use the lower bound for excess weight on tail risks computed by setting $\widehat{\gamma} = 1/2$ rather than dispersion directly. These correlations illustrate that most variation in excess weight on tail risks can be attributed to forecast dispersion, especially in the less noisy SU sample.[3]

[3] $Corr\left(\frac{\widehat{\sigma}_\theta}{\sigma_\theta}, \tilde{D}\right)$ and $Corr\left(\frac{\widehat{\sigma}_\theta}{\sigma_\theta}, lower\ bound\right)$ increase to 0.47 and 0.70, respectively, in the restricted GH sample if one removes a single outlier.

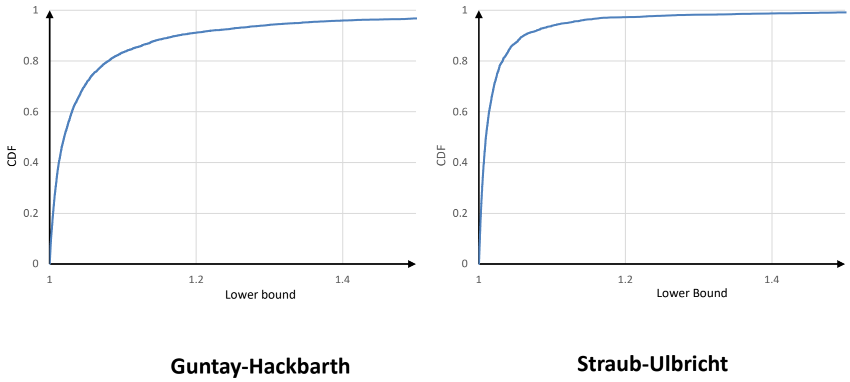

Figure 4: This figure displays cdf of the lower bound on EWTR (horizontal axis), in the GH (sample II, without $\widetilde{D} = 0$) and SU (sample I).

Figure 4: Lower bound (Cumulative Distribution)

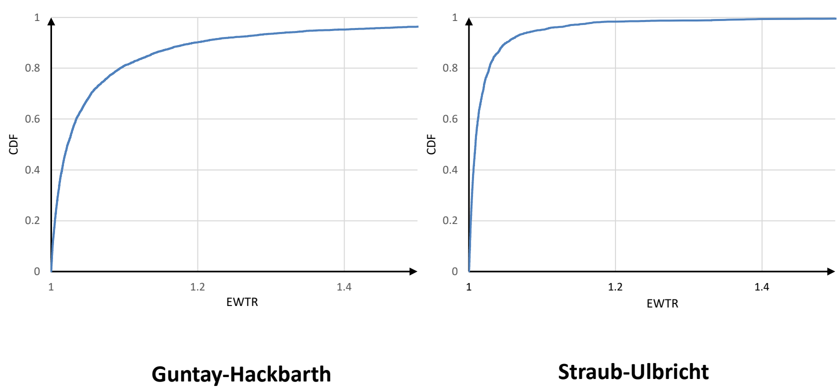

Figure 5: This figure displays cdf of EWTR (horizontal axis), in the GH (sample II, without $\widetilde{D} = 0$) and SU (sample I).

Figure 5: Excess weight on tail risks (Cumulative Distribution)

Sample	$Corr\left(\tilde{D},\widehat{\gamma}\right)$	$Corr\left(\tilde{D}, lower\ bd\right)$	$Corr\left(\widehat{\gamma}, lower\ bd\right)$	$Corr\left(\frac{\widehat{\sigma}_\theta}{\sigma_\theta},\tilde{D}\right)$	$Corr\left(\frac{\widehat{\sigma}_\theta}{\sigma_\theta},\widehat{\gamma}\right)$	$Corr\left(\frac{\widehat{\sigma}_\theta}{\sigma_\theta}, lower\ bd\right)$
GH - Sample II	-0.59					
GH - Sample III	-0.32	0.43	-0.15			
GH - Sample IV	-0.41	0.57	-0.23	0.32	-0.08	0.44
SU - Sample I	-0.65					
SU - Sample III	-0.56	0.62	-0.33			
SU - Sample IV	-0.52	0.66	-0.24	0.55	-0.20	0.86

Table 3: Sample Correlations

The distribution of forecast dispersion, forecast accuracy and excess weight on tail risks in the Straub-Ulbricht sample is qualitatively very similar to the GH sample, with a highly skewed distribution where most variation is concentrated in the top quintile of firms, but average values of forecast dispersion, accuracy and excess weight on tail risks are lower. We can attribute these differences to the different sample length, time period and forecast horizon, but also the fact that firms with a longer track record of being followed by analysts are more likely to garner wider investor interests, have more liquid markets and less overall belief dispersion - in other words, the more restrictive sample is likely to focus on firms that are less subject to noisy information aggregation frictions. However, the fact that two substantially different data sets deliver qualitatively very similar distributions of forecast dispersion and excess weight on tail risks, gives us some confidence in the qualitative robustness of our numerical examples. We have chosen to use GH as our primary sample because it matches most closely with empirical studies on equity and bond returns, which strikes us a better benchmark for assessing the quantitative implications of our model.

To summarize, our data suggest that excess weight on tail risks is likely to be small for most firms, but significantly larger in the top quintile of the distribution. As a ballpark estimate, the data suggest that an average excess weight on tail risks of about 10%, but most of this average is driven by the top quintile where excess weight on tail risks increases to close to 20%, while for a large majority of firms excess weight on tail risks remains very small. Most of the variation in excess weight on tail risks comes from forecast dispersion.

3 Analyst forecasts as a proxy for private information dispersion

On concern with our use of analyst forecast dispersion to measure excess weight on tail risks is that analyst forecasts are in the public domain, but we are using them, in the present case, to proxy for dispersed private information among investors. While this concern is not unique to our paper - it arises whenever publicly available measures of investor disagreement are used to proxy for dispersion of beliefs - it is of special concern to us since the model of noisy information aggregation explicitly posits a common prior assumption, under which publicly available information can on its own not be a source of belief dispersion.

Here, we discuss a simple setting of how public analyst forecasts may be used to construct a proxy for private information dispersion. Formally, we suppose that there is a continuum of investors whose earnings expectations are given by $\mathbb{E}\left(\theta|x,z;y\right)$, where $x \sim \mathcal{N}\left(\theta, \beta^{-1}\right)$ denotes their idiosyncratic private signal, $z \sim \mathcal{N}\left(\theta, \tau^{-1}\right)$ the information content of the price, and $y \sim \mathcal{N}\left(\theta, \tau_y^{-1}\right)$ any other additional public information. This public information may include the finite sample of investor (or analyst) forecasts. We consider a static model where implicitly the information that is aggregated through the published forecasts or the price is already implicitly taken into consideration by the forecasters' announcements - we can think of this as a fixed point of a "tatonnement" process between price formation and belief updating that may otherwise require a more complete dynamic model.

Suppose that the public information y includes a finite sample of N forecasts reported and published as a random sample of investor expectations $\{\mathbb{E}\left(\theta|x_n,z;y\right)\}_{n=1}^{N}$. With the linear-normal model, these forecasts take the form

$$\mathbb{E}\left(\theta|x_n,z;y\right) = \frac{\beta x_n + \tau z + \tau_y y}{\beta + \tau + \tau_y + \sigma_\theta^{-2}}$$

We can then construct the mean or consensus forecast

$$f = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left(\theta|x_n,z;y\right) = \frac{\beta\frac{1}{N}\sum_{n=1}^{N}x_n + \tau z + \tau_y y}{\beta + \tau + \tau_y + \sigma_\theta^{-2}}$$

and a sample estimate of forecast dispersion $\hat{D}$

$$\hat{D}^2 = \frac{1}{N-1}\sum_{n=1}^{N}\left(x_n - f\right)^2 = \frac{1}{N-1}\gamma_x^2\sum_{n=1}^{N}\left(x_n - \frac{1}{N}\sum_{n=1}^{N}x_n\right)^2$$

where $\gamma_x = \frac{\beta}{\beta+\tau+\tau_y+\sigma_\theta^{-2}}$. Upon taking expectations, it is straight-forward to check that

$$\mathbb{E}\left(\hat{D}^2\right) = \gamma_x^2\beta^{-1}$$

which corresponds to the true investor forecast dispersion. Hence $\hat{D}$ provides an unbiased, noisy estimate of the true underlying forecast dispersion.

We also define forecast accuracy as $A = 1 - \frac{Var(\theta|x_n,z;y)}{Var(\theta)} = \frac{\sigma_\theta^{-2}}{\beta+\tau+\tau_y+\sigma_\theta^{-2}}$, i.e. as the ratio of conditional forecast error variance to the unconditional variance of fundamentals. We can then define excess weight on tail risks as follows:

$$\widehat{\sigma}_\theta^2 = \mathbb{E}\left(Var\left(\theta|x=z,z;y\right)\right) + Var\left(\mathbb{E}\left(\theta|x=z,z;y\right)\right),$$

applying the Law of total variance for the risk-neutral measure. It is straight-forward to check that

$$\widehat{\sigma}_\theta^2 = \sigma_\theta^2 A + \sigma_\theta^2 \left(1-A\right)^2 + \frac{\tau_y+\beta+\tau}{\left(\beta+\tau+\tau_y+\sigma_\theta^{-2}\right)^2} + \frac{\left(\beta+\tau\right)\left(\frac{\beta+\tau}{\tau}-1\right)}{\left(\beta+\tau+\tau_y+\sigma_\theta^{-2}\right)^2}$$

$$\widehat{\sigma}_\theta^2 = \sigma_\theta^2\left\{A + \left(1-A\right)^2 + A\left(1-A\right) + \tilde{D}\frac{\beta+\tau}{\tau}\right\} = \sigma_\theta^2\left\{1+\tilde{D}\frac{\beta+\tau}{\tau}\right\}$$

where $\tilde{D}^2 = \mathbb{E}\left(\hat{D}^2\right)/\sigma_\theta^2$, as in the main text. Finally, we check that

$$\frac{\beta+\tau}{\tau} = \frac{\left(1-A\right)\left(A-\tau_y\left(1-A\right)\sigma_\theta^2\right)}{\left(1-A\right)\left(A-\tau_y\left(1-A\right)\sigma_\theta^2\right)-\tilde{D}},$$

which yields

$$\frac{\widehat{\sigma}_\theta^2}{\sigma_\theta^2} = 1 + \tilde{D}^2\frac{\left(1-A\right)\left(A-\tau_y\left(1-A\right)\sigma_\theta^2\right)}{\left(1-A\right)\left(A-\tau_y\left(1-A\right)\sigma_\theta^2\right)-\tilde{D}^2}.$$

If the precision of the additional public information τ_y is known, then this expression generalizes the one given in the paper (with $A = \widehat{\gamma}$). If τ_y is unknown, then the expression in the main text is a lower bound on the true estimate of EWTR. Finally in the special case where y corresponds to the mean or consensus forecast, we have $\tau_y = \beta N$ (assuming the latter is observed without noise).

4 Additional results for Section 4 and 5

4.1 Proof of proposition 5 in the main text

Let $\varphi(z) = \mathbb{E}\left(f\left(z-\theta'\right)|z\right) = \int f\left(z-\theta\right)h\left(\theta|z\right)d\theta$. Since marginalization preserves log-concavity (proposition 3.3 in Saumard and Wellner, 2014), $\varphi(z)$ is log-concave, whenever $f\left(z-\theta\right)h\left(\theta|z\right)$ is log-concave in (z,θ).[4] If $\psi\left(z|\theta\right) = \psi\left(z-\theta\right)$, where ψ is log-concave, we have $\varphi(z) = \frac{\int f(z-\theta)\psi(z-\theta)h(\theta)d\theta}{\int \psi(z-\theta)h(\theta)d\theta} = \frac{\int f(u)\psi(u)h(z-u)du}{\int \psi(u)h(z-u)du}$ and

$$\frac{\varphi'(z)}{\varphi(z)} = \hat{\mathbb{E}}\left(\frac{h'(\theta)}{h(\theta)}|z\right) - \mathbb{E}\left(\frac{h'(\theta)}{h(\theta)}|z\right).$$

[4]Since $h\left(\theta|z\right)$ is generically not log-concave even if h and ψ are, this requirement imposes a lower bound on the log-concavity of f, or equivalently the informativeness of the private signal.

Now,

$$\frac{d}{dz}\hat{\mathbb{E}}\left(\frac{h'(\theta)}{h(\theta)}|z\right)=\hat{\mathbb{E}}\left(\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}|z\right)+\widehat{Var}\left(\frac{h'(\theta)}{h(\theta)}|z\right)$$

Applying Brascamp and Lieb (1976), Theorem 4.1 (see proposition 10.1(a) in Saumard and Wellner, 2014) yields

$$\widehat{Var}\left(\frac{h'(\theta)}{h(\theta)}|z\right)\leq\hat{\mathbb{E}}\left(\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}\left\{-\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}-\frac{d}{dz}\left(\frac{f'(z-\theta)}{f(z-\theta)}+\frac{\psi'(z-\theta)}{\psi(z-\theta)}\right)\right\}^{-1}\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}|z\right)$$

and therefore

$$-\frac{d}{dz}\hat{\mathbb{E}}\left(\frac{h'(\theta)}{h(\theta)}|z\right)$$

$$\geq\hat{\mathbb{E}}\left(\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}\left\{-\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}-\frac{d}{dz}\left(\frac{f'(z-\theta)}{f(z-\theta)}+\frac{\psi'(z-\theta)}{\psi(z-\theta)}\right)\right\}^{-1}\left\{\frac{d}{dz}\left(\frac{f'(z-\theta)}{f(z-\theta)}+\frac{\psi'(z-\theta)}{\psi(z-\theta)}\right)\right\}|z\right)$$

$$=\hat{\mathbb{E}}\left(\left\{\left(-\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}\right)^{-1}+\left(-\frac{d}{dz}\left(\frac{f'(z-\theta)}{f(z-\theta)}+\frac{\psi'(z-\theta)}{\psi(z-\theta)}\right)\right)^{-1}\right\}^{-1}|z\right)\geq\left\{\underline{\tau}_h^{-1}+\left(\underline{\tau}_f+\underline{\tau}_\psi\right)^{-1}\right\}^{-1}=\underline{\widehat{\gamma}}\underline{\tau}_h,$$

where $\underline{\widehat{\gamma}}=\frac{\underline{\tau}_f+\underline{\tau}_\psi}{\underline{\tau}_f+\underline{\tau}_\psi+\underline{\tau}_h}$. In addition,

$$\mathbb{E}\left(\frac{h'(\theta)}{h(\theta)}|z\right)=\mathbb{E}\left(\frac{\psi'(z-\theta)}{\psi(z-\theta)}|z\right)=(1-\gamma)\,\mathbb{E}\left(\frac{\psi'((1-\gamma)z-u)}{\psi((1-\gamma)z-u)}|z\right)+\gamma\mathbb{E}\left(\frac{h'(\gamma z+u)}{h(\gamma z+u)}|z\right)$$

where $\theta=\gamma z+u$, and therefore

$$\frac{d}{dz}\mathbb{E}\left(\frac{h'(\theta)}{h(\theta)}|z\right)=(1-\gamma)^2\,\mathbb{E}\left(\frac{d}{dz}\frac{\psi'(z-\theta)}{\psi(z-\theta)}|z\right)+\gamma^2\mathbb{E}\left(\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}|z\right)+Var\left((1-\gamma)\frac{\psi'(z-\theta)}{\psi(z-\theta)}+\gamma\frac{h'(\theta)}{h(\theta)}|z\right).$$

Setting $\frac{d}{dz}\frac{\psi'(z-\theta)}{\psi(z-\theta)}\geq-\bar{\tau}_\psi$ and $\frac{d}{d\theta}\frac{h'(\theta)}{h(\theta)}\geq-\bar{\tau}_h$and minimizing the RHS w.r.t. γ yields $\frac{d}{dz}\mathbb{E}\left(\frac{h'(\theta)}{h(\theta)}|z\right)\geq-\bar{\gamma}\bar{\tau}_h$, where $\bar{\gamma}=\frac{\bar{\tau}_\psi}{\bar{\tau}_\psi+\bar{\tau}_h}$. Combining the two yields

$$-\frac{d}{dz}\frac{\varphi'(z)}{\varphi(z)}\geq\underline{\widehat{\gamma}}\underline{\tau}_h-\bar{\gamma}\bar{\tau}_h.$$

Next, we show that log-concavity of $\varphi(z)$ implies log-convexity of $\hat{m}(\theta)=\int\frac{f(z-\theta)}{\varphi(z)}\psi(z|\theta)\,dz$. If $\psi(z|\theta)=\psi(z-\theta)$, where ψ is log-concave, we have $\hat{m}(\theta)=\int\frac{f(u)}{\varphi(\theta+u)}\psi(u)\,du$ and $\frac{\hat{m}'(\theta)}{\hat{m}(\theta)}=-\hat{\mathbb{E}}\left(\frac{\varphi'(z)}{\varphi(z)}|\theta\right)$, where $\hat{\mathbb{E}}(\cdot|\theta)$ represents expectations w.r.t. the probability measure $\frac{\frac{f((\rho-1)\theta+v)}{\varphi(\rho\theta+v)}\psi((\rho-1)\theta+v)}{\int\frac{f((\rho-1)\theta+v)}{\varphi(\rho\theta+v)}\psi((\rho-1)\theta+v)dv}$. It follows that

$$\frac{d}{d\theta}\frac{\hat{m}'(\theta)}{\hat{m}(\theta)}=\widehat{Cov}\left(\frac{\varphi'(z)}{\varphi(z)},\frac{f'(z-\theta)}{f(z-\theta)}+\frac{\psi'(z-\theta)}{\psi(z-\theta)}|\theta\right)$$

and since $\frac{f'(z-\theta)}{f(z-\theta)}+\frac{\psi'(z-\theta)}{\psi(z-\theta)}$ is strictly decreasing in θ, $\frac{d}{d\theta}\frac{\hat{m}'(\theta)}{\hat{m}(\theta)}>0$, whenever $\frac{d}{dz}\frac{\varphi'(z)}{\varphi(z)}<0$.

Moreover, since $\frac{\hat{m}'(\theta)}{\hat{m}(\theta)} = -\hat{\mathbb{E}}\left(\frac{\varphi'(z)}{\varphi(z)}|\theta\right) = -\hat{\mathbb{E}}\left(\frac{f'(z-\theta)}{f(z-\theta)} + \frac{\psi'(z-\theta)}{\psi(z-\theta)}|\theta\right)$, changing variables to $v \equiv z - \rho\theta$ yields $\frac{\hat{m}'(\theta)}{\hat{m}(\theta)} = \hat{\mathbb{E}}\left(K\left(\theta, v\right)|\theta\right)$ and

$$\frac{d}{d\theta}\frac{\hat{m}'(\theta)}{\hat{m}(\theta)} = \hat{\mathbb{E}}\left(\frac{d}{d\theta}K(\theta,v)|\theta\right) + \widehat{Var}\left(K(\theta,v)|\theta\right)$$

$$\text{where } K(\theta,v) = (\rho-1)\left(\frac{f'((\rho-1)\theta+v)}{f((\rho-1)\theta+v)} + \frac{\psi'((\rho-1)\theta+v)}{\psi((\rho-1)\theta+v)}\right) - \rho\frac{\varphi'(\rho\theta+v)}{\varphi(\rho\theta+v)}.$$

It follows that $\frac{d}{d\theta}\frac{\hat{m}'(\theta)}{\hat{m}(\theta)} > 0$ for all θ whenever $\frac{d}{d\theta}K(\theta,v) > 0$ for all (θ, v), or

$$(\rho-1)^2\left(-\tau_f(z-\theta) - \tau_\psi(z-\theta)\right) + \rho^2\tau_\varphi(z) > 0$$

for all (θ, z), where $\tau_f(z-\theta) = -\frac{d}{dz}\left(\frac{f'(z-\theta)}{f(z-\theta)}\right)$, $\tau_\varphi(z) = -\frac{d}{dz}\frac{\varphi'(z)}{\varphi(z)}$ and $\tau_\psi = -\frac{d}{dz}\left(\frac{\psi'(z-\theta)}{\psi(z-\theta)}\right)$. Since ρ is a free parameter, we can set $\rho(\theta, v)$ to maximize the LHS of above inequality. From $0 \le \tau_\varphi(z) \le \underline{\widehat{\gamma}}\underline{\tau}_h = \left(1-\underline{\widehat{\gamma}}\right)\left(\underline{\tau}_f + \underline{\tau}_\psi\right)$, it follows that this expression is maximized when $\rho = \frac{\tau_f+\tau_\psi}{\tau_f+\tau_\psi-\tau_\varphi}$, which yields

$$\frac{d}{d\theta}\frac{\hat{m}'(\theta)}{\hat{m}(\theta)} \ge \hat{\mathbb{E}}\left(\frac{\tau_f+\tau_\psi}{\tau_f+\tau_\psi-\tau_\varphi}\tau_\varphi|\theta\right) \ge \frac{\underline{\widehat{\gamma}}}{\underline{\widehat{\gamma}} - \left(1-\underline{\widehat{\gamma}}\right)\frac{\tau_\varphi}{\underline{\tau}_h}}\frac{\tau_\varphi}{\underline{\tau}_h}\underline{\tau}_h,$$

which is strictly positive and bounded away from 0 whenever $\tau_\varphi(z) > 0$. Substituting $\tau_\varphi(z) \ge \underline{\widehat{\gamma}}\underline{\tau}_h - \bar{\gamma}\bar{\tau}_h$ then yields

$$\frac{d}{d\theta}\frac{\hat{m}'(\theta)}{\hat{m}(\theta)} \ge \frac{\underline{\widehat{\gamma}}\left(\underline{\widehat{\gamma}} - \frac{\bar{\tau}_h}{\underline{\tau}_h}\bar{\gamma}\right)}{\underline{\widehat{\gamma}} - \left(1-\underline{\widehat{\gamma}}\right)\left(\underline{\widehat{\gamma}} - \frac{\bar{\tau}_h}{\underline{\tau}_h}\bar{\gamma}\right)}\underline{\tau}_h > 0.$$

Finally, we consider the limit as f, ψ, and h converge to normal distributions, in which case $\tau_f \to \beta$, $\tau_\psi \to \tau$, and $\tau_h \to \sigma_\theta^{-2}$, and the upper and lower bounds on these functions also converge to β, τ, and σ_θ^{-2}, respectively. In this case, $\underline{\widehat{\gamma}} \to \widehat{\gamma}$,$\bar{\gamma} \to \gamma$, and $\tau_\varphi = -\frac{d}{dz}\left(\frac{\varphi'(z)}{\varphi(z)}\right) \to \left(\widehat{\gamma} - \gamma\right)\frac{1}{\sigma_\theta^2}$. This in turn implies that

$$K(\theta,v) \to -(\rho-1)(\beta+\tau)((\rho-1)\theta+v) + \rho\tau_\varphi(\rho\theta+v) = \theta\rho\tau_\varphi,$$

which implies that $\widehat{Var}\left(K(\theta,v)|\theta\right) \to 0$ and $\frac{d}{d\theta}K(\theta,v) \to \rho\tau_\varphi = \rho\left(\widehat{\gamma}-\gamma\right)\frac{1}{\sigma_\theta^2}$ as f, ψ, and h converge to normal distributions. But then,

$$\frac{d}{d\theta}\frac{\hat{m}'(\theta)}{\hat{m}(\theta)} \to \rho\tau_\varphi = \frac{\beta+\tau}{\beta+\tau-\tau_\varphi}\tau_\varphi = \frac{1}{\sigma_\theta^2}\frac{\widehat{\gamma}\left(\widehat{\gamma}-\gamma\right)}{\widehat{\gamma} - \left(1-\widehat{\gamma}\right)\left(\widehat{\gamma}-\gamma\right)} = \frac{1}{\sigma_\theta^2} - \frac{1}{\widehat{\sigma}_\theta^2}.$$

4.2 Numerical solution methods

Here we present the results from the iteration procedure described in section 4. Fix a support of the fundamental θ and prior $H(\cdot)$. We start by conjecturing a distribution of prices conditional on a given value of θ: $\Psi^{(0)}(P'|\theta) \equiv Pr(P \le P'|\theta)$, along with a conditional density

$\psi^{(0)}(P|\theta)$. From $\psi^{(0)}(P|\theta)$, we calculate the posterior distribution for each investor using Bayes rule: $Pr(\theta|x_i, P) = \psi^{(0)}(P|\theta) \cdot Pr(\theta|x_i) / \sum_{\theta'} \psi^{(0)}(P|\theta') \cdot Pr(\theta'|x_i)$, where $Pr(\theta|x_i)$ corresponds to the posterior conditional on observing x_i only. Using the posterior distribution, we find the demand functions $d(x_i, P)$ that maximize $E[u(w)|x_i, P]$ for each agent i, and then determine aggregate demand $D(\theta, P)$ numerically by integrating over x. Using the market-clearing condition, we then characterize the resulting informational content of prices $\Psi^{(1)}(P'|\theta) \equiv 1 - G(D(\theta, P'))$. This new conditional price distribution $\Psi^{(1)}$ is used then as the starting guess in place of $\Psi^{(0)}$, and the exercise is iterated until convergence. Finally, we calculate the price function $P(\theta, s)$ by inverting the function $D(\theta, P) = s$ to obtain $P = P(\theta, s = D)$.[5]

With CRRA preferences and binary asset payoffs, this task is simplified by the fact that the posterior odds ratio and asset demand can be written in closed form. Suppose that all primitives are as in section 3.1, except that preferences are of the CRRA form: $U(w) = w^{1-\chi}/(1-\chi)$, where $w = w_0 + d(\pi - P)$ is an investor's the terminal wealth when purchasing d units of the asset, and w_0 is an initial endowment, identical across agents. An investor's posterior belief $\mu(x, P) = Pr(\pi = 1|x, P)$ satisfies

$$log\frac{\mu(x,P)}{1-\mu(x,P)} = log\frac{\lambda}{1-\lambda} + log\left(\frac{\psi(P|\pi=1)}{\psi(P|\pi=0)}\right) + \beta(x-0.5),$$

and optimal asset demand is

$$d(x,P) = w_0 \frac{\left(\frac{\mu(x,P)}{1-\mu(x,P)}\right)^{\frac{1}{\chi}} - \left(\frac{P}{1-P}\right)^{\frac{1}{\chi}}}{(1-P)\left(\frac{P}{1-P}\right)^{\frac{1}{\chi}} + P\left(\frac{\mu(x,P)}{1-\mu(x,P)}\right)^{\frac{1}{\chi}}}.$$

Aggregating across agents, we obtain the aggregate demand $D(\pi, P)$ and the price likelihood ratio $\psi(P|\pi = 1)/\psi(P|\pi = 0) = g(D(1, P))/g(D(0, P))$. We use these likelihoods as inputs for $\mu(x, P)$, and iterate until convergence. Parameters are set to $w_0 = 5$, $\mathbb{E}(s) = 0.5$, $\sigma_s = 1$, $\beta = 1$, and $\gamma = 3$. [6]

Figure 6 plots the relation between the random supply and three objects of interest, for the dividend realization $\pi = 1$:[7] the price $P(1, s)$, the posterior expectation of dividends conditional

[5] We make our matlab code available for the CRRA, binary payoff case in the online appendix. Under generic preferences and payoff structures, aggregate demand monotonicity w.r.t. prices is not guaranteed (see for example, Barlevy and Veronesi (2003). Without strict monotonicity, then it is no longer true that $Pr(P \leq P'|\theta) = Pr(u \geq D(\theta, P))$ for any price level, and the solution method proposed here would not work. All the examples presented in this section are under parameters which satisfy monotonicity of demand in prices.

[6] For futher details on how to implement the algorithm, see matlab code attached.

[7] The corresponding figures for $\pi = 0$ are qualitatively similar, but horizontally shifted to the left. The average price and average dividend expectations included in the figure of course consider both states of nature.

on the price P, $log\frac{\lambda}{1-\lambda}+log\left(\frac{\psi(P|\pi=1)}{\psi(P|\pi=0)}\right)$, and the posterior expectation $\mu_i(z,P)$ of an investor observing $x_i=z$, where z is set so that the investor purchases $d(z,P)=\mathbb{E}(s)$ exactly the expected asset supply. Thus, the difference between $P(1,s)$ and $\mu_i(z,P)$ equals the risk premium required for holding the average net supply $\mathbb{E}(s)$ of the asset. Panel a) of Figure 6 considers a security with downside risk, $\lambda=0.9$ (i.e., a bond), while panel b) plots the same relations ships for a security with upside risk, $\lambda=0.1$ (i.e., a stock, or option).

In both figures, the price, the dividend expectation conditional on the sufficient statistic z (defined as the signal of the trader which holds the average supply), and the expectations of the marginal trader (which weighs z both as public and private information), are increasing functions of the sufficient statistic, as it indicates a higher probability that $\pi=1$. However, the expectations of the marginal investor are more sensitive to z, which delivers a lower expectations for the marginal investor relative to those that condition only on the price signal, for relatively low values of z,while the converse is true for high realization. This qualitative feature is true regardless of the shape of cash flows, as can be seen by comparing panels a) and b) of Figure 6.

On average, due to the concavity of the cash flow under downside risks in panel a), the average expectation of the marginal investor (over realizations of z) falls below the probability of $\pi=1$ (0.9 in the downside risk case studies). The difference between these objects is the skewness premium (negative in this case). The average price then compounds the skewness premium with the risk premium for holding a positive expected supply of the asset, with both of these premia pushing the average asset price below the unconditional probability that $\pi=1$ which is equal to 0.9.

For the upside risk case in panel b), the convex nature of payoffs implies that now the average expectation of the marginal investor lies above the average asset payoff, representing a positive skewness premium. From these expectations, a positive risk premium is also discounted, resulting in a price that lies below the average expectation of the marginal investor. For the parameters chosen, the skewness premium just dominates the downward price influence of risk aversion, and the resulting average price is slightly above the average dividend.

In summary, the qualitative implications described in the text under CARA-binary preferences survive when we use CRRA preferences instead, as one would expect given the generalizations of the model described in section 4.

4.3 Limit cases of the general model

Here we discuss additional limit cases of the general model that give rise to excess weight on tail risks, returns to skewness and disagreement as mentioned in the discussion in section 4 and 5 of

CRRA-binary model

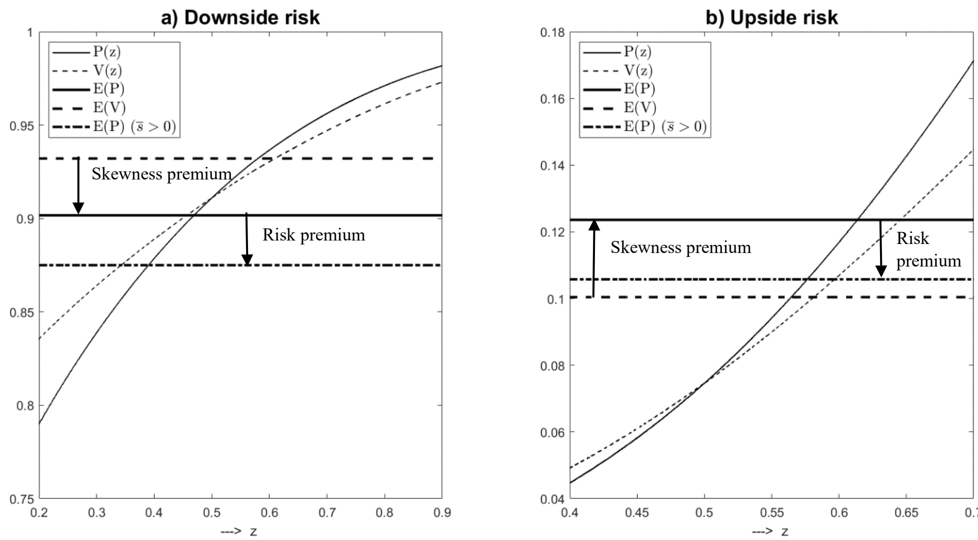

Figure 6: CRRA-binary simulation

the paper. The additional structure provided by these limits allows us to directly characterize the risk-neutral probability or expected equilibrium prices and skewness premia.

It will be convenient to directly consider a market with informed and uninformed traders. Formally, we consider sequences of markets indexed by $n = 1, 2, ...$ and characterized by information structures $\{\kappa_n, F_n, G_n\}$ defined by a fraction of informed traders κ_n, a distribution of private signals F_n, and a distribution of supply shocks G_n, which converge to a limit $\{\kappa_\infty, F_\infty, G_\infty\}$. When F_∞ and G_∞ are degenerate, it will be convenient to fix baseline distributions F and G centered around 0 and simply vary the precision of private signals β_n and the variance of supply shocks $Var_n(s)$ along the sequence. We assume throughout that the corresponding equilibrium sequence $\{d_n^I(x, P); d_n^U(P); P_n(z); \Psi_n(z|\theta)\}$ has a well-defined limit, and that $\left\{d_n^I(x, P)\right\}$ and $\left\{d_n^U(P)\right\}$ are both decreasing in P, for all n.

Finally, we assume that $U'(\cdot)$ is an analytic function, which can be represented by its infinite order Taylor expansion. This allows us to represent the informed trader's first-order condition as follows:

$$\sum_{n=0}^{\infty} \frac{1}{n!} U^{(n+1)}(0)\, d^I(x, P(z))^n \sum_{k=0}^{n+1} \binom{n+1}{k} \mu_k(x, z) \left(\mathbb{E}(\pi(\theta)|x, z) - P(z)\right)^{n+1-k} = 0 \qquad (2)$$

where $\mu_k(x, z) \equiv \mathbb{E}\left(\left(\pi(\theta) - \mathbb{E}(\pi(\theta)|x, z)\right)^k |x, z\right)$. Equation (2) implies that we can represent

$d^I(x, P(z))$ as a function of $\mathbb{E}(\pi(\theta)|x,z) - P(z)$ and the conditional moments $\mu_k(x,z)$:

$$d^I(x, P(z)) = D(\mathbb{E}(\pi(\theta)|x,z) - P(z); \{\mu_k(x,z)\}). \tag{3}$$

Likewise, the uninformed traders' first-order condition can be represented as

$$\sum_{n=0}^{\infty} \frac{1}{n!} U^{(n+1)}(0)\, d^U(P(z))^n \sum_{k=0}^{n+1} \binom{n+1}{k} \mu_k(z) (\mathbb{E}(\pi(\theta)|z) - P(z))^{n+1-k} = 0$$

where $\mu_k(z) \equiv \mathbb{E}\left((\pi(\theta) - \mathbb{E}(\pi(\theta)|z))^k |z\right)$, or equivalently $d^U(P(z)) = D(\mathbb{E}(\pi(\theta)|z) - P(z); \{\mu_k(z)\})$. It follows that $\mathbb{E}(\pi(\theta)|z) - P(z)$ is of order $d^U(P(z))$, and focusing on the first three orders in the Taylor expansion, we obtain

$$0 = \mathbb{E}(\pi(\theta)|z) - P(z) - \chi d^U(P(z)) \left((\mathbb{E}(\pi(\theta)|z) - P(z))^2 + \mu_2(z)\right)$$

$$+\frac{\alpha}{2} d^U(P(z))^2 \left((\mathbb{E}(\pi(\theta)|z) - P(z))^3 + \mu_3(z) + 3\mu_2(z)(\mathbb{E}(\pi(\theta)|z) - P(z))\right) + o\left(d^U(P(z))^3\right)$$

where $\chi = -U^{(2)}(0)/U^{(1)}(0)$ and $\alpha = U^{(3)}(0)/U^{(1)}(0)$, or

$$P(z) = \mathbb{E}(\pi(\theta)|z) - \chi d^U(P(z))\, \mu_2(z) + \frac{\alpha}{2} d^U(P(z))^2 \mu_3(z) + o\left(d^U(P(z))^3\right)$$

Taking prior expectations then yields equation (19) in the paper.

The market-clearing condition can be written as

$$(1-\kappa)\, d^U(P(z)) + \kappa \int d^I(x, P(z))\, dF(x-\theta) = s.$$

Finally, define the *no-information price* $V^{NI}(s)$, which obtains when $\beta\kappa = 0$, i.e. there are no informed traders (or their private signals are infinitely noisy). This no-information price satisfies

$$\sum_{n=0}^{\infty} \frac{1}{n!} U^{(n+1)}(0)\, s^n \sum_{k=0}^{n+1} \binom{n+1}{k} \mu_k \left(\mathbb{E}(\pi(\theta)) - V^{NI}(s)\right)^{n+1-k} = 0 \tag{4}$$

where $\mu_k \equiv \mathbb{E}\left((\pi(\theta) - \mathbb{E}(\pi(\theta)))^k\right)$, or equivalently, $s = D\left(\mathbb{E}(\pi(\theta)) - V_\pi^{NI}(s); \{\mu_k\}\right)$.

We first generalize the discussion of limits when private information vanishes, and supply shocks are small, from section 5 in the paper (with informed and uninformed traders). We then consider two additional limits, in which supply shocks become large. In the first, noisy information aggregation can lead to arbitrarily large, unbounded variation in prices, even though supply has bounded support. In the second, we consider a limit case where supply shocks become large, but private information becomes very price, so that prices remain informative in the limit. Once again the equilibrium price is strictly more variable than the underlying fundamental, and it results in a risk-neutral measure with excess weight on tail risks.

4.3.1 Supply shocks vs. dispersed information

Here we generalize the observation, discussed in section 5 in the paper (as well as in proposition 2 above in the online appendix) that dispersed information amplifies skewness premia that are generated by supply shocks. As in the paper, we distinguish between information-based and preference-based returns to skewness, and show that when supply shocks are small, the information-based returns to skewness are an order of magnitude larger than the preference-based ones.

The no information limit: Consider a sequence of economies $\{\kappa_n\}$, such that $\lim_{n\to 0}\kappa_n = 0$, i.e. the fraction of informed traders converges to 0, while the distribution of private signals F and the distribution of supply shocks G is kept constant. In this case, it follows from the market-clearing condition that $d^U(P(z)) \to s$, which in turn implies that the price must co-move with s but be independent of θ. Therefore $\mu_k(z)$ converges to μ_k, and it follows immediately that $P(z) \to V^{NI}(s)$. It follows that $V^{NI}(s) - \mathbb{E}(\pi(\theta)) = o(s)$, and therefore (from the Taylor expansion for uninformed traders):

$$V^{NI}(s) - \mathbb{E}(\pi(\theta)) = -\chi\mu_2 s + \frac{\alpha}{2}\mu_3 s^2 + o\left(s^3\right)$$

and

$$\mathbb{E}\left(V^{NI}(s)\right) = \mathbb{E}(\pi(\theta)) + \alpha Var(s)\mu_3 + o\left(Var(s)^{3/2}\right).$$

It is straight-forward to check that the same limit is reached for a sequence of economies $\{\beta_n\}$, such that $\lim_{n\to 0}\beta_n = 0$, i.e. the precision of private signals converges to 0, while holding fixed the fraction of informed traders. In this case, $\mu_k(x,z)$ converges to $\mu_k(z)$ and $d^I(x,P)$ converges to $d^U(P)$, which by market-clearing must then both converge to s. But then the price becomes uninformative, and $\mu_k(x,z)$ and $\mu_k(z)$ converge to μ_k, which in turn implies that $P(z) \to V^{NI}(s)$.

This characterization generalizes the observation in section 5 of the *preference-based* skewness premium: At the no-information limit, the expected price premium scales with the skewness of returns μ_3, the variance of supply shocks $Var(s)$, and the degree of downside risk aversion α.

The noisy information limit: Consider next the limit in which $\kappa_n \to 0$ and $\{G_n\}$ is such that $s_n = \kappa_n \tilde{s}$, where $\tilde{s}$ is centered at 0 and distributed according to $\tilde{G}$. In this case, the market-clearing condition implies

$$\frac{1-\kappa_n}{\kappa_n} d_n^U(P(z)) = \tilde{s} - \int d_n^I(x, P(z))\, dF(x-\theta)$$

We first show that prices must remain informative in the limit: suppose to the contrary that $\mu_k(z)$ converges to μ_k and $\mathbb{E}(\pi(\theta)|z)$ to $\mathbb{E}(\pi(\theta))$ in which case the uninformed trader's demand takes the form $d^U(P) = D(\mathbb{E}(\pi(\theta)) - P; \{\mu_k\})$, hence the left-hand-side of the market-clearing condition is independent of θ. On the other hand, we have shown in the proof of Theorem 1 that $d^I(x, P)$ must be strictly increasing in x, even at a limit at which z is uninformative. But then the right-hand side of the market-clearing condition must be strictly decreasing in θ, which yields a contradiction.

It follows that there exists a limit distribution $\Psi^\infty(z|\theta)$ for the sufficient statistic. The market-clearing condition implies that $d^U(P(z)) - s \to 0$ w.p. 1, and therefore at the limit

$$\lim_{n\to\infty} P(z) = \mathbb{E}(\pi(\theta)|z) - \chi s \mu_2(z) + \frac{\alpha}{2} s^2 \mu_3(z) + o\left(s^3\right)$$

and

$$\lim_{n\to\infty} \mathbb{E}(P(z)) = \mathbb{E}(\pi(\theta)) - \chi \kappa_n Cov(\tilde{s}, \mu_2(z)) + o\left(\kappa_n^2\right).$$

This characterization generalizes the observation in section 5 of the information-based skewness premium that depends on the covariance between supply shocks, or uninformed traders' exposures $\tilde{s}$, and posterior uncertainty $\mu_2(z)$ and which scales with the standard deviation of supply shocks, or κ_n.

The preference- and information-based skewness premia both result from negative co-movement of the price with exposure of uninformed agents. Skewness or asymmetry in returns makes this co-movement asymmetric: with upside risk, the co-movement is stronger on the upside, resulting in a convex relation between exposure and prices and a positive price premium, while with downside risk this co-movement is stronger on the downside, resulting in a concave relation and a negative price premium.

However preference- and information-based skewness premia differ in the reasons for asymmetric co-movement. The preference-based skewness premium results from downside risk aversion ($U'''(0) > 0$), which induces investors to seek extra compensation for downside risks. This increases the risk premium for negatively skewed securities and decreases it for positively skewed securities.

The information-based skewness premium instead results from endogenous co-movement between equilibrium supply and uncertainty: since informed investors do not perfectly arbitrage supply shocks, P and z are decreasing in s in equilibrium. But the price P conveys information z, which in turn induces investors to update their expectations and posterior variances. When $\pi(\cdot)$ is dominated by downside risk (convex), an increase in P increases dividend expectations, but also the investors' conditional uncertainty $\mu_2(z)$. $\mu_2(z)$ and s thus co-move negatively, resulting in a positive price premium. When instead $\pi(\cdot)$ is dominated by downside risk (concave), $\mu_2(z)$ and s co-move positively, resulting in a negative price premium.

The same argument extends to the limit of a sequence, in which $\beta_n \to 0$ and $Var_n(s) \to 0$ in such a manner that z remains informative, while holding κ_n constant. In this limit, the posterior $H(\cdot|x,z)$ converges to $H(\cdot|z)$, and moments $\mu_k(x,z)$ to $\mu_k(z)$, as $\beta \to 0$. From the market-clearing condition,

$$d^U(P) = s - \kappa\left(\int d^I(x,P)\,dF(x-\theta) - d^U(P)\right),$$

and therefore the price remains informative, only if $Var_n(s)$ is proportional to the variance of $\int d^I(x,P)\,dF(x-\theta) - d^U(P)$. The term $d^I(x,P) - d^U(P)$ in turn scales with $\mathbb{E}(\pi(\theta)|x,z) - \mathbb{E}(\pi(\theta)|z)$ when β is close to 0, and $\mathbb{E}(\pi(\theta)|x,z) - \mathbb{E}(\pi(\theta)|z)$ scales with β. It follows that $\int d^I(x,P)\,dF(x-\theta) - d^U(P) = o(\beta)$, and the price remains informative in the limit, if and only if $Var_n(s)/\beta_n^2$ converges to a positive finite constant. From the uninformed trader's first-order condition, we then obtain that the expected price premium must satisfy

$$\lim_{n\to\infty}\mathbb{E}(P(z)) = \mathbb{E}(\pi(\theta)) - \chi Cov\left(d^U(P(z)),\mu_2(z)\right) + o(Var_n(s)),$$

where $d^U(P(z)) = o\left(\sqrt{Var_n(s)}\right)$, i.e. to a first order the expected price premium is governed by the co-movement between the uninformed traders' exposure $d^U(P(z))$ and posterior uncertainty $\mu_2(z)$, and the former scales with the standard deviation of supply shocks, when the latter is stationary at the noisy information limit.

4.3.2 The large noise limit

A similar amplification result also holds for positive levels of private signal precision when supply shocks make prices completely uninformative. Formally consider a sequence $\{G_n\}$ of dispersed information economies such that $\lim_{n\to\infty}\left\|G_n(s) - \bar{G}\right\| = 0$ for all $s \in (d_L, d_H)$, and some $\bar{G} \in (0,1)$, i.e. supply realizations are shifted more and more to the boundaries of the support of s. We compare the dispersed information price when private signals are informative about θ with its no information counterpart $V^{NI}(s)$.

At the no-information limit, we obtain $V^{NI}(d_L) > V^{NI}(s) > V^{NI}(d_H)$ and $Var\left(V^{NI}(s)\right) \leq 1/4\left(V^{NI}(d_L) - V^{NI}(d_H)\right)^2$. Supply shocks are bounded by (d_L, d_H) and their variance is bounded by $Var(s) \leq (d_H - d_L)^2\,\bar{G}\left(1-\bar{G}\right)$. This in turn translates into a uniform bound on prices, price volatility, excess weight on tail risks and the skewness premium.

With dispersed information, we first show that the asset price indeed becomes perfectly uninformative as n grows large. In the limit, arbitrarily large fluctuations in marginal investor expectations are then needed to absorb supply shocks. Demand and price functions are well-defined in the limit, but z becomes infinitely volatile, resulting in an arbitrarily large excess weight on tail risks i.e. the risk-neutral measure converges to an improper prior. With convex $\pi(\cdot)$, this in turn implies that the average price and hence the skewness premium must grow arbitrarily large. This limit result is particularly striking in how pure noise in the realization of z introduces large swings in prices because these shocks to z generate large swings in the marginal investors dividend expectations - recall that the market price always treats the sufficient statistic as if it has at least a precision of $\beta > 0$, even when in fact it is completely uninformative.

More specifically, we first show that the asset price becomes perfectly uninformative as n grows large: Suppose to the contrary that $H_n(\theta|P)$ converges to some non-degenerate limit distribution $H(\theta|P)$. Using $H(\theta|P)$, we can derive a limit demand function $d^I_\infty(x,P)$ and a limit price function $P_\infty(z)$. Substituting these into $\Psi_n(z|\theta) = 1 - G_n(D_n(\theta, P_n(z)))$, it follows that $\Psi_n(z|\theta) \to 1 - \bar{G}$ for any finite θ and z. From this, it follows that $\Psi_n(z) = \int \Psi_n(z|\theta)\, dH(\theta) \to 1 - \bar{G}$, and

$$\hat{H}_n(\theta) = \int H(\theta|x = z, z) d\Psi(z) = -\int \Psi(z) \frac{\partial H_n(\theta|x = z, z)}{\partial z} dz \to \left(1 - \bar{G}\right).$$

In addition, note that

$$Pr_n\left(\theta \le \theta'; P \le P'\right) = \int_{-\infty}^{\theta'} \left(1 - G_n\left(D_n(\theta, P)\right)\right) dH(\theta) \to \left(1 - \bar{G}\right) H\left(\theta'\right)$$

On the other hand, since $Pr_n\left(\theta \le \theta'; P \le P'\right) = Pr_n\left(\theta \le \theta' | P \le P'\right) Pr_n\left(P \le P'\right)$. Since $Pr_n\left(P \le P'\right) \to 1 - \bar{G}$ for any interior P', we conclude that $Pr_n\left(\theta \le \theta' | P \le P'\right) \to H\left(\theta'\right)$, and hence $H(\theta|P) \to H(\theta)$.

Next we show that with dispersed information and large noise, price fluctuations and expected prices become arbitrarily large if $\pi(\cdot)$ is convex and hence unbounded on the upside. Given the convergence of beliefs, the uninformed traders demand converges to $d^U(P) = D\left(\mathbb{E}(\pi(\theta)) - P; \{\mu_k\}\right)$ and the informed traders to $\bar{d}^I(x, P) = D\left(\mathbb{E}(\pi(\theta)|x) - P; \{\mu_k(x)\}\right)$ where $\{\mu_k(x)\}$ and $\mathbb{E}(\pi(\theta)|x)$ only condition on the private signal realization x. These limits are uniquely determined by the prior and the distribution of private signals. Furthermore they imply that there exists a unique sufficient statistic $z(P)$ defined by setting $\bar{d}^I(z, P) = 0$, and correspondingly a unique limit price function $P(z) = \mathbb{E}(\pi(\theta)|x = z)$. What's more, if $\pi(\cdot)$ is unbounded, then $lim_{n\to\infty} Var(P(z)) = \infty$. If $\pi(\cdot)$ is strictly convex (more generally, $\pi'(\theta) + \pi'(-\theta)$ is strictly positive and bounded away from 0 for $\theta > 0$ sufficiently large, then $lim_{n\to\infty}\mathbb{E}(P(z)) = \infty$.

Hence, to summarize, skewness premia are again amplified by dispersed information: at the no information benchmark, the expected price premium scales with the variance of supply, which remains bounded. With dispersed information instead, excess weight on tail risks scales with the variance of the marginal trader's private signal realization, which can be unbounded, even if supply shocks are bounded. This explains why the amplification effect from dispersed information can become arbitrarily large.

4.3.3 The precise information limit

Along similar lines, we consider the limit in which $\beta \to \infty$ (private signals become arbitrarily precise), while supply realizations become more and more noisy, i.e. $\{G_n\}$ is such that $\lim_{n\to\infty} \left\|G_n(s) - \bar{G}\right\| = 0$ for all $s \in (d_L, d_H)$, and some $\bar{G} \in (0,1)$.

In this scenario, $\{\mu_k(x,z)\} \to \{0\}$, i.e. informed traders face no uncertainty, and in the limit, they follow a threshold strategy and set $d^I(x,P) = d_L$ if $x < \hat{x}$ and $d^I(x,P) = d_H$ if $x > \hat{x}$, where $\hat{x}$ satisfies $\pi(\hat{x}) = P$. Setting $z = \hat{x}$ yields $P(z) = \pi(z)$, i.e. in the limit the informed traders act as if their private signal conveyed θ perfectly. The uninformed traders in turn update based on the information conveyed by the price.

The demand by informed traders is given by $D^I(\theta, P(z)) = d_L + (d_H - d_L)\left(1 - F\left(\sqrt{\beta}(z-\theta)\right)\right)$, which yields the market-clearing condition

$$\kappa\left(d_L + (d_H - d_L)\left(1 - F\left(\sqrt{\beta}(z-\theta)\right)\right)\right) = s - (1-\kappa)\, d^U(P(z))$$

When $\kappa = 1$ (no uninformed traders), it follows that $z \to \theta + \frac{1}{\sqrt{\beta}} F^{-1}\left(\frac{d_H - s}{d_H - d_L}\right)$, and z decomposes into a shift in the mean and a mean-preserving spread over the fundamental θ. Since $P(z) = \pi(z)$ the distribution of z also defines the risk-neutral probability measure, which thus has strictly fatter tails than the prior H. The same is true by continuity when κ is sufficiently close to 1.

5 Additional results for multi-asset extensions

Here we expand on the characterization of pricing kernels with multiple securities. Recall that

$$P(z) = \mathbb{E}(\pi(\theta)\, m(\theta,z)\,|z) = \mathbb{E}\left(\pi(\theta)\, m^U(\theta,z)\,|z\right)$$

where

$$m(\theta,z) = \frac{U'\left((\pi(\theta) - P(z))' d_0\right)}{\mathbb{E}\left(U'\left((\pi(\theta) - P(z))' d_0\right)|x = z, z\right)} \frac{h(\theta|x = z, z)}{h(\theta|z)}$$

$$m^U(\theta,z) = \frac{U'\left((\pi(\theta) - P(z))' d^U(P(z))\right)}{\mathbb{E}\left(U'\left((\pi(\theta) - P(z))' d^U(P(z))\right)|z\right)}.$$

5.1 Generalization of information- and preference-based skewness premia

A second-order Taylor approximation for the uninformed traders yields

$$P_n(z) = \mathbb{E}(\pi_n(\theta)|z) - \chi \cdot e_n' \Sigma(z) d^U(P(z)) + \frac{\alpha}{2} \cdot d^U(P(z))' \Psi^n(z) d^U(P(z)) + o\left(\left\| d^U(P(z))^3 \right\|\right),$$

where $\chi = -\frac{U''(0)}{U'(0)}$, $\alpha = \frac{U'''(0)}{U'(0)}$, $e_n' = (0, ..., 1, 0, ...0)$ represents the n-th dimension unit vector, $\Sigma(z)$ the $N \times N$ variance-covariance matrix of expected returns with nm-th entry $Cov(\pi_n(\theta), \pi_m(\theta)|z)$, and $\Psi^n(z)$ the n-th third-moment matrix with k, l-th entries

$$\psi_{(k,l)}^n(z) = \mathbb{E}((\pi_n(\theta) - \mathbb{E}(\pi_n(\theta)|z))(\pi_k(\theta) - \mathbb{E}(\pi_k(\theta)|z))(\pi_l(\theta) - \mathbb{E}(\pi_l(\theta)|z))|z).$$

The difference between price and expected dividend thus decomposes into a risk adjustment $-\chi \cdot e_n' \Sigma(z) d^U(P(z))$ that scales with the uninformed traders' exposure $d^U(P(z))$ and a third-moment adjustment term $\frac{\alpha}{2} \cdot d^U(P(z))' \Psi^n(z) d^U(P(z))$ that depends on the squares of exposures. The risk adjustment can be rewritten as $e_n' \Sigma(z) d^U(P(z)) = Cov(\pi_n(\theta), \pi(\theta) \cdot d^U(P(z))|z)$, where $\pi(\theta) \cdot d^U(P(z))$ represents the uninformed traders' total portfolio return. Abstracting from the second-order (third-moment) terms, the model-implied risk premium thus recovers a standard "CAPM" representation from the perspective of uninformed investors.[8]

If $\alpha = 0$ (quadratic preferences, no downside risk aversion), or equivalently, ignoring the third-moment terms, the expected price premium satisfies

$$\begin{aligned}\mathbb{E}(P_n(z)) - \mathbb{E}(\pi_n(\theta)) &\approx -\chi \cdot e_n' \mathbb{E}\left(\Sigma(z) \cdot d^U(P(z))\right) = -\chi \cdot \mathbb{E}\left(Cov\left(\pi_n(\theta), \pi(\theta) \cdot d^U(P(z))|z\right)\right) \\ &= -\chi \cdot e_n' \mathbb{E}(\Sigma(z)) D - \chi \cdot \mathbb{E}\left(e_n' \Sigma(z) \cdot \left(d^U(P(z)) - D\right)\right)\end{aligned}$$

where $D = \mathbb{E}\left(d^U(P(z))\right)$. The expected price premium thus decomposes into an average risk premium that scales with risk aversion χ, expected posterior uncertainty $\mathbb{E}(\Sigma(z))$, expected exposure D, and an adjustment due to the co-movement between the exposure and uncertainty. When assets are conditionally independent ($Cov(\pi_n(\theta), \pi_m(\theta)|z) = 0$), this co-movement term reduces to $\mathbb{E}\left(e_n' \Sigma(z) \cdot \left(d^U(P(z)) - D\right)\right) = Cov\left(Var(\pi_n(\theta)|z), d_n^U(P(z))\right)$. If exposure $d_n^U(P(z))$ is everywhere decreasing in z, then the co-movement term is positive (negative) if uncertainty $Var(\pi_n(\theta)|z)$ is decreasing (increasing) in z. Therefore, controlling for the average exposure, co-movement generates a positive expected price premium if uncertainty and exposure are both

[8] See Andrei et al. (2022) for implications of noisy information aggregation for empirical properties of the CAPM with linear/normal asset returns.

counter-cyclical, and a negative premium if uncertainty is pro-cyclical. This is exactly what return asymmetry generates: for downside risks, a deterioration of z increases the likelihood of adverse tail risks, hence uncertainty is countercyclical. For an upside risk the same deterioration of reduces uncertainty as the positive tail event is less likely to materialize.

This observation leads to an alternative interpretation of the negative relation between skewness and returns as resulting from the combination of (i) counter-cyclical exposure of uninformed traders, and (ii) pro-cyclical (counter-cyclical) uncertainty of upside (downside) risks. The counter-cyclical exposure of uninformed traders emerges naturally from the informed traders' demand and the market-clearing condition: since

$$\kappa_I \int d_I\left(x, P\left(z\right)\right) dF(x-\theta) + \kappa_U d^U\left(P\left(z\right)\right) = s$$

an increase in the fundamental vector θ that raises demand $\int d_I\left(x, P_\pi\left(z\right)\right) dF(x-\theta)$ by the informed traders for all securities must be offset by a reduction in the demand by uninformed traders, resulting in lower exposures for uninformed traders when the fundamental is high, or asset supply is low.[9]

Consider next the case where $\alpha > 0$. With downside risk aversion, the second-order term $\frac{\alpha}{2} \cdot d^U\left(P\left(z\right)\right)' \Psi^n\left(z\right) d^U\left(P\left(z\right)\right)$ multiplies the investors' attitudes towards downside risk with the squared exposures and asymmetries in returns that are summarized by $d^U\left(P\left(z\right)\right)' \Psi^n\left(z\right) d^U\left(P\left(z\right)\right)$. When assets are conditional independent, the latter term reduces to $Skew\left(\pi_n\left(\theta\right)|z\right) \cdot d_n^U\left(P_n\left(z\right)\right)^2$, where $Skew\left(\pi_n\left(\theta\right)|z\right) = \psi^n_{(n,n)}\left(z\right)$ denotes the conditional skewness of asset payoffs. Taking expectations, this term thus generalizes the observation discussed in section 2 that attitudes towards downside risk introduce a *preference-based* skewness premium in asset prices: because of downside risk aversion, traders require additional compensation for accepting the market-clearing exposure level s on negatively skewed securities, while willing to reduce the risk premium for positively skewed securities.

Finally, we note that this representation also yields a generalization of the amplification of price

[9]Similar results obtain in the risk-neutral model with a noise trader demand of the form $s = \Phi\left(u + \omega\left(P - \mathbb{E}\left(\pi\left(\theta\right)|P\right)\right)\right)$, which captures the notion that the residual supply available to informed traders increases in the expected price premium $P - \mathbb{E}\left(\pi\left(\theta\right)|P\right)$. In this formulation, the higher is $\omega > 0$ the more actively the uninformed traders arbitrage the perceived price premium, with $P \to \mathbb{E}\left(\pi\left(\theta\right)|P\right)$ as $\omega \to \infty$, akin to free entry by uninformed risk-neutral arbitrageurs. A micro-foundation for this functional form assumption about asset supply can be obtained by assuming that *(i)* the asset supply is normalized to 1, and *(ii)* there is a unit measure of risk-neutral uninformed arbitrageurs, who each have a stochastic cost of $c_i = c + u_i$ of holding the one unit of the asset, where $c \sim \mathcal{N}\left(\bar{c}, \sigma_c^2\right)$ and $u_i \sim \mathcal{N}\left(0, \gamma^{-1}\right)$. In this case, uninformed arbitrageurs buy the security if and only if $c_i + P \leq \mathbb{E}\left(\pi\left(\theta\right)|P\right)$, resulting in a residual supply schedule of $\Phi\left(\sqrt{\gamma}\left(c + P - \mathbb{E}\left(\pi\left(\theta\right)|P\right)\right)\right)$, which confirms the above representation with $\omega = \sqrt{\gamma}$ and $u = \sqrt{\gamma}c \sim \mathcal{N}\left(\sqrt{\gamma}\bar{c}, \gamma\sigma_c^2\right)$.

premia with dispersed information. Suppose that $\mathbb{E}(s) = 0$ and consider first the limit without informed traders, as κ_I goes to 0. In this case, $d^U(P(z))$ must converge to s/κ_U, P must become completely uninformative, and price fluctuations are exclusively due to supply shocks. Therefore, $\Sigma(z)$ converges to the prior variance-covariance matrix Σ which is independent of $d^U(P(z)) = s/\kappa_U$, and hence the information-based premium vanishes: In the limit, $\mathbb{E}\left(e_n'\Sigma(z)\cdot\left(d^U(P(z)) - D\right)\right) = e_n'\Sigma\cdot\mathbb{E}\left(\left(d^U(P(z)) - D\right)\right) = 0$. However $\mathbb{E}\left(d^U(P(z))'\,\Psi^n(z)\,d^U(P(z))\right)$ converges to $\frac{1}{\kappa_U^2}\mathbb{E}(s'\Psi^n s)$, where Ψ^n is the unconditional third-moment matrix. This last expression is positive (negative) whenever Ψ^n is positive (negative)-definite; with independent assets, this limit is $\sum_{n=1}^N Skew(\pi_n(\theta))\,\mathbb{E}\left(s_n^2\right)$. This limit thus highlights that the preference-based skewness premium scales with the variance of supply shocks.

Alternatively consider the limit, in which $\kappa_I \to 0$ and the distribution of supply shocks is also scaled by κ_I, i.e. $s = \kappa_I\tilde{s}$, where $\tilde{s}$ is distributed according to some fixed distribution $\tilde{G}$, with $\mathbb{E}(\tilde{s}) = 0$. In this limit, z remains informative about $\int d_I(x, P(z))\,dF(x-\theta) - \tilde{s}$, i.e. $\Sigma(z)$ converges to a finite limit, and $d^U(P(z))$ must then scale with κ_I to satisfy market-clearing. The expected price premium satisfies

$$\mathbb{E}(P_n(z)) - \mathbb{E}(\pi_n(\theta)) = -\chi\mathbb{E}\left(Cov\left(\pi_n(\theta), \pi(\theta)\cdot d^U(P(z))\,|z\right)\right) + o\left(\mathbb{E}\left(\left\|s^2\right\|\right)\right)$$

$$= -\chi\cdot e_n'\mathbb{E}\left(\Sigma(z)\cdot d^U(P(z))\right) + o\left(\mathbb{E}\left(\left\|s^2\right\|\right)\right).$$

This last expression, and hence the information-based skewness premium, scales with κ_I, or the standard deviation of supply shocks, and conditional second moments of returns (the conditional variance-covariance matrix, or the conditional covariance of $\pi_n(\theta)$ with the uninformed traders' portfolio return $\pi(\theta)\cdot d^U(P(z))$), rather than third moments.

5.2 Independent Securities

Consider now the case in which trading in N markets is independent. Suppose that fundamentals θ, supply s and signal noise ε^i are all component-wise independent, $\frac{\partial\pi_n}{\partial\theta_{n'}} = 0$ for $n \neq n'$, and traders have CARA preferences. We prove the following proposition:

Proposition 3 : *Suppose that each component market considered in isolation admits a price-monotonic equilibrium* $\{P_n(\theta_n, s_n); d_n^I(x_n, P_n); d_n^U(P_n); H_n(\cdot|P_n)\}$ *with pricing kernels* $m_n(\theta_n, z_n)$ *for informed and* $m_n^U(\theta_n, z_n)$ *for uninformed traders. Then there exists a price-monotonic equilibrium with simultaneous trading in* N *markets characterized by* $\{P_n(\theta_n, s_n); d_n^I(x_n, P_n); d_n^U(P_n); H_n(\cdot|P_n)\}_{n=1}^N$ *and pricing kernels* $m(\theta, z) = \prod_{n=1}^N m_n(\theta_n, z_n)$ *and* $m^U(\theta, z) = \prod_{n=1}^N m_n^U(\theta_n, z_n)$.

This proposition gives sufficient conditions under which trading behavior is independent across markets and the pricing kernel is decomposable asset-by-asset. The key assumptions are CARA preferences and component-wise independence. Under CARA preferences, trading gains or losses do not have spill-overs across markets, which in turn implies that demand for any security n is not affected by risk-taking in other markets. With component-wise independence, prices also do not have informational spill-overs, i.e. information aggregation about asset n is independent of markets $n' \neq n$.

Proof: We first guess and verify that the posterior density $h(\theta|P)$ satisfies $h(\theta|P) = \prod_{n=1}^{N} h_n(\theta_n|P_n)$, i.e. the posterior density is component-wise independent. It then also follows that $h(\theta|x, P) = \prod_{n=1}^{N} h_n(\theta_n|x_n, P_n)$ since private signals are component-wise independent.

If the conjecture is correct, with CARA preferences and component-wise independence the uninformed traders' FOC can be rewritten as

$$0 = \mathbb{E}\left(\prod_{n=1}^{N} \exp\left(-\chi\pi_n(\theta_n) d_n^U(P)\right)(\pi_n(\theta_n) - P_n)|P\right)$$

$$= \mathbb{E}\left(\exp\left(-\chi\pi_n(\theta_n) d_n^U(P)\right)(\pi_n(\theta_n) - P_n)|P_n\right)\mathbb{E}\left(\prod_{k\neq n} \exp\left(-\chi\pi_k(\theta_k) d_k^U(P)\right)|P\right)$$

$$= \mathbb{E}\left(\exp\left(-\chi\pi_n(\theta_n) d_n^U(P)\right)(\pi_n(\theta_n) - P_n)|P_n\right)$$

which in turn implies that

$$P_n = \frac{\mathbb{E}\left(\exp\left(-\chi\pi_n(\theta_n) d_n^U(P)\right)\pi_n(\theta_n)|P_n\right)}{\mathbb{E}\left(\exp\left(-\chi\pi_n(\theta_n) d_n^U(P)\right)|P_n\right)}.$$

Likewise, the informed trader's FOC can be rewritten as

$$P_n = \frac{\mathbb{E}\left(\exp\left(-\chi\pi_n(\theta_n) d_n^I(x, P)\right)\pi_n(\theta_n)|x_n, P_n\right)}{\mathbb{E}\left(\exp\left(-\chi\pi_n(\theta_n) d_n^I(x, P)\right)|x_n, P_n\right)}$$

It follows that under our guess, the informed and uninformed traders' demands only depend on the price in market n and their private signal x_n, and is independent of prices and private signals for other securities. Hence we can write $d_n^U(P) \equiv d_n^U(P_n)$ and $d_n^I(x, P) \equiv d_n^I(x_n, P_n)$. It then follows that market-clearing in market n is given by

$$\kappa_I \int d_n^I(x_n, P_n)\, dF(x_n - \theta_n) + \kappa_U d_n^U(P_n) = s_n,$$

which in turn implies that P_n is a function of θ_n and s_n only, i.e. prices are component-wise independent. But if prices are component-wise independent, then it must be the case that $h(\theta|P) =$

$\prod_{n=1}^{N} h_n(\theta_n|P_n)$, i.e. traders' posteriors are also component wise independent which confirms our initial guess.

It then follows that we can directly apply the results in section 4 of the paper to show that there exist a component-wise independent sufficient statistic vector z as well as component-wise pricing kernels

$$m_n^U(\theta_n, z_n) = \frac{\exp\left(-\chi\pi_n(\theta_n) d_n^U(P_n(z_n))\right)}{\mathbb{E}\left(\exp\left(-\chi\pi_n(\theta_n) d_n^U(P_n(z_n))\right)|z_n\right)}$$

$$m_n(\theta_n, z_n) = \frac{\exp\left(-\chi\pi_n(\theta_n) d_{n,0}\right)}{\mathbb{E}\left(\exp\left(-\chi\pi_n(\theta_n) d_{n,0}\right)|x_n = z_n, z_n\right)} \frac{h(\theta_n|x_n = z_n, z_n)}{h(\theta_n|z_n)}$$

and $P_n(z_n) = \mathbb{E}(\pi_n(\theta_n) m_n(\theta_n, z_n)|z_n) = \mathbb{E}\left(\pi_n(\theta_n) m_n^U(\theta_n, z_n)|z_n\right)$. Moreover, letting $m^U(\theta, z) = \prod_{n=1}^{N} m_n^U(\theta_n, z_n)$ and $m(\theta, z) = \prod_{n=1}^{N} m_n(\theta_n, z_n)$, we obtain

$$P_n(z) = \mathbb{E}(\pi_n(\theta_n) m_n(\theta_n, z_n)|z_n) = \mathbb{E}(\pi_n(\theta_n) m_n(\theta_n, z_n)|z_n)\, \mathbb{E}\left(\prod_{k \neq n} m_k(\theta_k, z_k)|z\right)$$

$$= \mathbb{E}\left(\pi_n(\theta_n) \prod_{k=1}^{N} m_k(\theta_k, z_k)|z\right) = \mathbb{E}(\pi_n(\theta_n) m(\theta, z)|z)$$

and by the same argument $P_n(z) = \mathbb{E}\left(\pi_n(\theta_n) m^U(\theta, z)|z\right)$.

5.3 Security Design: Splitting Cash-flows to influence market value

Our main theoretical results suggest that dispersed information may plausibly account for observed price premia or discounts in equity and bond markets. Here we use our model to argue that such premia provide a novel and potentially important element shaping security design incentives and firms' funding choices.

With perfectly competitive financial markets, the market value of a given cash flow should not depend on how it is allocated to different investors (Modigliani and Miller, 1958). Our model with noisy information aggregation instead suggests that a seller incurs an expected loss from issuing bonds, while generating a gain from selling options or levered equity claims. Furthermore, when investor pools for different claims have different informational characteristics, the seller can impact expected revenues by tailoring the split to different investor types. We illustrate these points in a two-asset version of the risk-neutral normal model.

Consider a risk-neutral securities originator, or *seller*, who owns claims on a stochastic dividend $\pi(\cdot)$. This cash flow is divided into two parts, π_1 and π_2, both monotone in θ, such that $\pi_1 + \pi_2 = \pi$, and then sold to investors in two separate markets at prices P_1 and P_2, to be determined in equilibrium. We assume without loss of generality that π_2 has more upside risk than π_1. For each claim, there is a unit measure of informed investors who obtain a noisy private signal $x_i \sim \mathcal{N}\left(\theta, \beta_i^{-1}\right)$, and a noisy supply $\Phi\left(u_i\right)$, where

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u,1}^2 & \rho\sigma_{u,1}\sigma_{u,2} \\ \rho\sigma_{u,1}\sigma_{u,2} & \sigma_{u,2}^2 \end{pmatrix}\right)$$

That is, each market is affected by a supply shock u_i with market-specific variance $\sigma_{u,i}^2$. The environment is then characterized by the market-characteristics β_i and $\sigma_{u,i}^2$, and by the correlation of demand shocks across markets, ρ. Investors are are risk-neutral and face position limits as in section 3, and are active only in their respective market. We consider both the possibility that investors observe and condition on prices in the other market (informational linkages), and the possibility that they do not (informational segregation).

The seller is risk-neutral, and hence wishes to maximize the expected revenue net of the dividend, $P_1 + P_2 - \pi(\cdot)$. With both informational linkages and informational segregation, the seller's net expected revenue can be represented as the sum of the expected price wedges for the two securities, $W\left(\pi_1, \widehat{\sigma}_{\theta,1}\right) + W\left(\pi_2, \widehat{\sigma}_{\theta,2}\right)$, where $\widehat{\sigma}_{\theta,i}$ denotes the level of informational frictions in market i. Since our results below take these information frictions parameters as given, they apply identically to the models with informational segregation and informational linkages. The two cases only differ in how $\widehat{\sigma}_{\theta,i}$ is determined.[10]

Proposition 4 ***(Modigliani-Miller):***
$W\left(\pi_1, \widehat{\sigma}_{\theta,1}\right) + W\left(\pi_2, \widehat{\sigma}_{\theta,2}\right) \gtreqless W\left(\pi_1, \widehat{\sigma}_{\theta,2}\right) + W\left(\pi_2, \widehat{\sigma}_{\theta,1}\right)$ *if and only if* $\widehat{\sigma}_{\theta,1} \gtreqless \widehat{\sigma}_{\theta,2}$.

The proof follows directly from additivity and increasing differences: If the two markets have identical characteristics, i.e. $\widehat{\sigma}_{\theta,1} = \widehat{\sigma}_{\theta,2} = \widehat{\sigma}_\theta$, the expected wedge is additive across cash flows: $W\left(\pi_1, \widehat{\sigma}_\theta\right) + W\left(\pi_2, \widehat{\sigma}_\theta\right) = W\left(\pi_1 + \pi_2, \widehat{\sigma}_\theta\right)$ for any π_1 and π_2, and only the combined cash flow

[10] With informational segregation the analysis of the two markets can be completely separated and the equilbrium price and information frictions parameters $\widehat{\sigma}_{\theta,i}$ are determined as in section 2.3. With informational linkages the equilibrium analysis has to be adjusted to incorporate the information contained in price 1 for market 2, and vice versa. Informed investors in market i buy a security if and only if their private signal exceeds a threshold $\hat{x}_i(\cdot)$, where $\hat{x}_i(\cdot)$ is conditioned on both prices. By market-clearing, $\hat{x}_i(\cdot) = z_i \equiv \theta - 1/\sqrt{\beta_i} \cdot u_i$. Equilibrium prices $P_1(z_1, z_2) = \mathbb{E}(\pi_1(\theta)|x = z_1; z_1, z_2)$ and $P_2(z_1, z_2) = \mathbb{E}(\pi_2(\theta)|x = z_2; z_1, z_2)$ are invertible functions of (z_1, z_2) and the expected skewness premium is $W\left(\pi_i, \widehat{\sigma}_{\theta,i}\right) = \int \left(\Phi\left(\frac{\theta}{\sigma_\theta}\right) - \Phi\left(\frac{\theta}{\widehat{\sigma}_{\theta,i}}\right)\right) d\pi_i(\theta)$, where

$$\widehat{\sigma}_{\theta,i}^2 = \sigma_\theta^2 + \left(1 + \sigma_{u_i}^2\right) \cdot \frac{\beta_i}{\left(\beta_i + V\right)^2}, \text{ with } V = 1/\sigma_\theta^2 + \frac{1}{1-\rho^2}\left(\frac{\beta_1}{\sigma_{u,1}^2} + \frac{\beta_2}{\sigma_{u,2}^2} - 2\rho\frac{\sqrt{\beta_1\beta_2}}{\sigma_{u,1}\sigma_{u,2}}\right).$$

matters for the total wedge – i.e., the Modigliani-Miller theorem applies. If instead the two markets have different informational characteristics the seller maximizes expected revenue by matching the security with more upside risk to the market that has more severe information frictions (a higher value of $\widehat{\sigma}_\theta$). This maximizes the gains from the positive wedge resulting on the upside, while minimizing the losses from the negative wedge on the downside. The next proposition advances this logic further by considering how the seller can exploit the heterogeneity in investor pools if she gets to design the split of π into π_1 and π_2.

Proposition 5 ***(Designing cash flows):*** *The seller maximizes her expected revenues by splitting cash flows according to $\pi_1^*(\theta) = \min\{\pi(\theta), \pi(0)\}$ and $\pi_2^*(\theta) = \max\{\pi(\theta) - \pi(0), 0\}$, and then assigning π_1^* to the investor pool with the lower $\widehat{\sigma}_{\theta,i}$.*

Proof: *For any alternative split (π_1, π_2), the monotonicity requirements imply that $0 \leq \pi_1(\theta_1) - \pi_1(\theta_2) \leq \pi(\theta_1) - \pi(\theta_2)$ and $0 \geq \pi_1(-\theta_1) - \pi_1(-\theta_2) \geq \pi(-\theta_1) - \pi(-\theta_2)$ for $\theta_1 > \theta_2 \geq 0$. It follows that*

$$\pi_1(\theta_1) - \pi_1(\theta_2) + \pi_1(-\theta_1) - \pi_1(-\theta_2) \geq \pi(-\theta_1) - \pi(-\theta_2) = \pi_1^*(\theta_1) - \pi_1^*(\theta_2) + \pi_1^*(-\theta_1) - \pi_1^*(-\theta_2)$$

i.e. that π_1 has less downside risk and more upside risk than π_1^*. Likewise π_2 has more downside risk and less upside risk than π_2^*. But then, the expected revenue of selling π_1 to the investor pool with $\widehat{\sigma}_{\theta,1}$ and π_2 to the investor pool with $\widehat{\sigma}_{\theta,2}$ is $W(\pi_1, \widehat{\sigma}_{\theta,1}) + W(\pi_2, \widehat{\sigma}_{\theta,2}) = W(\pi, \widehat{\sigma}_{\theta,2}) + W(\pi_1, \widehat{\sigma}_{\theta,1}) - W(\pi_1, \widehat{\sigma}_{\theta,2})$, while the expected revenue from selling π_1^* to the investor pool with $\widehat{\sigma}_{\theta,1}$ and π_2^* to the investor pool with $\widehat{\sigma}_{\theta,2}$ is $W(\pi_1^*, \widehat{\sigma}_{\theta,1}) + W(\pi_2^*, \widehat{\sigma}_{\theta,2}) = W(\pi, \widehat{\sigma}_{\theta,2}) + W(\pi_1^*, \widehat{\sigma}_{\theta,1}) - W(\pi_1^*, \widehat{\sigma}_{\theta,2})$. The difference in revenues is therefore $W(\pi_1^*, \widehat{\sigma}_{\theta,1}) - W(\pi_1^*, \widehat{\sigma}_{\theta,2}) - (W(\pi_1, \widehat{\sigma}_{\theta,1}) - W(\pi_1, \widehat{\sigma}_{\theta,2}))$, which is positive, since π_1^* contains more downside risk than π_1, and $\widehat{\sigma}_{\theta,2} > \widehat{\sigma}_{\theta,1}$.

The seller maximizes the total proceeds by assigning all the cash flow below the line defined by $\pi(.) = \pi(0)$ to the investor group with the lowest information friction parameter, $\widehat{\sigma}_{\theta,1}$, and the complement to the investor group with the highest friction; $\widehat{\sigma}_{\theta,2}$. When $\pi(\cdot) > 0$, this split has a straightforward interpretation in terms of debt and equity, with a default point on debt that is set at the prior median $\pi(0)$. For any other arbitrary division of cash flows $\{\pi_1(\cdot), \pi_2(\cdot)\}$, π_1 has less downside risk than π_1^*, and π_2 has less upside risk than π_2^*. This raises the expected price premium on π_1 and lowers the expected price premium on π_2, but due to increasing differences, the lower expected price premium on π_2 dominates, resulting in strictly lower expected revenue for the seller.

Propositions 4 and 5 are direct applications of additivity and the increasing difference property of W (Proposition 3(ii) in the paper).[11] They highlight how systematic departures from the Modigliani-Miller Theorem arise as a result of information frictions in asset markets, and how such asset market frictions may affect financial structure and security design. Much of the existing security design literature instead focuses on the role of "information-insensitive" debt contracts to mitigate asymmetric information between insiders who sell claims to raise funds and uninformed outsiders (Myers and Majluf, 1984; DeMarzo and Duffie, 1999; DeMarzo 2005), or limit liquidity traders' losses when trading with agents who hold superior information about the claims' quality (Gorton and Pennacchi, 1990; Boot and Thakor, 1993). Boot and Thakor (1993) and Fulghieri and Lukin (2001) emphasize the role of information-sensitive junior securities or equity claims to incentivize information acquisition. In all these models, securities are priced at their expected fundamental value, as prices are either set by a risk-neutral competitive market-maker or by a set of homogeneous outside investors, taking into consideration the signaling effects of the insiders' security design decisions. This rules out mispricing of securities as a force shaping security design incentives.

Closest to our work, Axelson (2007) studies the role of mis-pricing for security design when securities are sold through an auction and investors have private information about a firm's cash flow.[12] The auction mechanism generically results in under-pricing of securities due to a winner's curse, and the optimal security design seeks to limit the losses associated with the winner's curse resulting in either debt or call option contracts - depending on whether the issuer has more to gain from limiting the winner's curse through information insensitivity (debt) or from aligning the cash raised through information sensitive securities (equity or options) with the firm's fundamentals. As the market becomes more and more competitive, the underpricing disappears and the optimal security design approximates an equity claim. Like Axelson (2007), we also emphasize the role of market frictions in optimal security design, but we abstract from the objective of raising cash for investment, and instead allow for liquidity shocks along with informed trading, and investor pools with different informational characteristics. This opens up the possibility that securities can be over- as well as under-priced, resulting in the separation of upside vs. downside risk as the key force driving optimal security design. The debt-equity split then emerges as the optimal way of

[11] We could also consider a simpler security design problem, in which the seller designs a security $\pi_2(\cdot)$ for a single investor pool, keeping the residual $\pi_1 = \pi - \pi_2$ to himself. Using the same logic as above, the seller's optimal design consists in a call option contract $\pi_2^*(\theta) = \max\{\pi(\theta) - \pi(0), 0\}$ that maximizes the price premium on the upside risk, while keeping the downside risk (or debt claim) that would be under-priced to himself.

[12] See Yang (2020) for related results when a single buyer strategically acquires information about the seller's cash flows.

catering specific securities to the different investor pools.[13]

References

[1] Andrei, Daniel, Julien Cujean and Mungo Wilson (2022). "The Lost Capital Asset Pricing Model," Review of Economic Studies, forthcoming.

[2] Axelson, Ulf (2007). "Security design with investor private information," Journal of Finance, 62 (6), 2587-2632.

[3] Barlevy, Gadi and Pietro Veronesi (2003). "Rational Panics and Stock Market Crashes," Journal of Economic Theory, 110 (2), 234-263.

[4] Boot, Arnoud WA, and Anjan V. Thakor (1993). "Security design." The Journal of Finance 48(4), 1349-1378.

[5] Brascamp, Herm Jan, and Elliott H. Lieb (1976). "On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, with an application to the diffusion equation," Journal of Functional Analysis, 22, 366-389 (1976).

[6] DeMarzo, Peter M. (2005) "The pooling and tranching of securities: A model of informed intermediation." Review of Financial Studies 18(1), 1-35.

[7] DeMarzo, Peter, and Darrell Duffie (1999). "A liquidity-based model of security design." Econometrica 67(1), 65-99.

[8] Fulghieri, Paolo, and Dmitry Lukin (2001). "Information production, dilution costs, and optimal security design." Journal of Financial Economics, 61(1), 3-42.

[9] Gorton, Gary, and George Pennacchi (1990). "Financial intermediaries and liquidity creation." Journal of finance, 45 (1), 49-71.

[10] Güntay, Levent, and Dirk Hackbarth (2010). "Corporate Bond Credit Spreads and Forecast Dispersion." Journal of Banking and Finance, 34 (10), 2328-2345.

[13] By taking as given the differences in market characteristics and assuming that the seller can freely assign the cash-flows to these two pools, we omit the possibility that market characteristics themselves respond to how the seller designs the securities. Analyzing this interplay between investor's information choices and the resulting market characteristics, along with the seller's security design question is an important avenue for further work, but clearly beyond the scope of this paper.

[11] Modigliani, Franco and Merton H. Miller (1958). "The Cost of Capital, Corporate Finance and the Theory of Investment," American Economic Review, 48 (3), 261-297.

[12] Myers, Stewart C., and Nicholas S. Majluf (1984). "Corporate financing and investment decisions when firms have information that investors do not have." Journal of Financial Economics 13(2), 187-221.

[13] Saumard, Adrien and Jon Wellner (2014). "Log-Concavity and Strong Log-Concavity: a Review," working paper (arXiv:1404.5886).

[14] Straub, Ludwig, and Robert Ulbricht (2023). "Endogenous Uncertainty and Credit Crunches," working paper, MIT and TSE.

[15] Yang, Ming (2020). "Optimality of Debt under Flexible Information Acquisition," Review of Economic Studies 87(1), 487-536.