WORKING PAPERS

N° 1172

January 2021

“Dispersed Information and Asset Prices”

Elias Albagli, Christian Hellwig and Aleh Tsyvinski

Toulouse School of Economics

Dispersed Information and Asset Prices*

Elias Albagli — Central Bank of Chile

Christian Hellwig — Toulouse School of Economics

Aleh Tsyvinski — Yale University

January 2, 2021

Abstract

We argue that noisy aggregation of dispersed information provides a unified explanation for several prominent cross-sectional return anomalies such as returns to skewness, returns to disagreement and corporate credit spreads. We characterize asset returns with noisy information aggregation by means of a risk-neutral probability measure that features excess weight on tail risks, and link the latter to observable moments of earnings forecasts, in particular forecast dispersion and accuracy. We calibrate our model to match these moments and show that it accounts for a large fraction of the empirical return premia. We further develop asset pricing tools for noisy information aggregation models that do not impose strong parametric restrictions on economic primitives such as preferences, information, or return distributions.

*We thank Bruno Biais, John Geanakoplos, Narayana Kocherlakota, Felix Kubler, Stephen Morris, Guillaume Plantin, Jeremy Stein, Jean Tirole, Xavier Vives, Olivier Wang, Eric Young, and audiences at numerous conferences and seminars for helpful comments. This paper builds on and supersedes our earlier working paper "A theory of asset prices based on heterogeneous information" (Albagli, Hellwig, Tsyvinski 2011, 2015). Hellwig gratefully acknowledges financial support from the European Research Council (starting grant agreement 263790), the Institut Universitaire de France as well as funding from the French National Research Agency (ANR) under grant ANR-17-EURE-0010 (Investissements d'Avenir).Tsyvinski is grateful to NSF (grant 0551175) for support and EIEF for hospitality.

1 Introduction

Dispersed information and disagreement among investors about the expected cash-flows of different securities is a common feature of many, if not most financial markets, and markets are often viewed as playing a central role of aggregating such information through prices. We develop a parsimonious, flexible theory of asset pricing in which aggregation of dispersed information emerges as the core force determining asset prices and expected returns. Our theory links the asset's predicted prices and returns to the distribution of the underlying cash-flow risk and features of the market environment such as market liquidity and investor disagreement. It provides a unified explanation for several widely discussed empirical regularities. We show that our theory is consistent with negative excess return to skewness, with seemingly contradictory empirical evidence on the impact of investor disagreement on returns in equity and bond markets, and that it can account for a significant part of the high level and variation in corporate credit spreads - the credit spread puzzle.

We consider an asset market along the lines of Grossman and Stiglitz (1980), Hellwig (1980), and Diamond and Verrecchia (1981).[1] An investor pool consists of informed investors who observe a noisy private signal about the value of an underlying cash flow, and uninformed noise traders whose random trades determine the net supply of the asset. The asset price equates the demand by informed investors to the available asset supply and serves as an endogenous, noisy public signal of the asset's cash-flow. In contrast to the existing literature we do not impose any parametric restrictions on the distribution of asset dividends, which allows us to derive return implications for a wide range of assets traded in financial markets, and to confront the model-implied returns with their empirical counterparts.

The textbook no-arbitrage paradigm interprets systematic return differences across securities as compensation for risk. We build on no-arbitrage theory by constructing a risk-neutral probability measure for asset prices with dispersed information, and we identify information frictions and limits to arbitrage as a novel source of systematic return differences. In contrast to its no-arbitrage counterpart, the risk-neutral probability measure with noisy information aggregation is asset-specific since it factors in information and noise-trading frictions that may be specific to each individual security. Nevertheless, the equilibrium imposes restrictions on the shape of the risk-neutral probability measure. Specifically, noisy information aggregation generates *excess weight on tail risks*: the risk-neutral measure overweighs the probabilities of both very favorable and very unfavorable

[1] See Brunnermeier (2001), Vives (2008), and Veldkamp (2011) for textbook discussions.

outcomes. This contrasts with a risk adjustment that overweighs unfavorable outcomes relative to favorable ones. What's more, the extent by which the market overweighs tail risks scales with the dispersion of investor expectations about returns. Negative returns to skewness, negative (positive) returns to investor disagreement for positively (negatively) skewed securities, and positive interaction between skewness and investor disagreement are then a natural consequence of interpreting the expected price premium as the value of a mean-preserving spread.

We introduce our model in section 2 and develop the main theoretical results in section 3. The main technical challenge comes from dealing with the endogeneity of information contained in the asset price. By assuming that investors are risk-neutral but face position limits, we are able to fully characterize the information content of the price without any further restrictions on return distributions. Moreover by assuming risk-neutrality, our model abstracts from "standard" risk premia to focus on the role of noisy information aggregation for asset returns.

The theoretical contribution proceeds in three steps.

First, we represent the equilibrium price by means of a *sufficient statistic* variable that summarizes the information aggregated through the market. This sufficient statistics representation reveals the presence of an *updating wedge*: noisy information aggregation makes the asset price more sensitive to fundamental and liquidity shocks than the corresponding risk-adjusted dividend expectations. Hence the prices is higher (lower) than expected dividends whenever the information aggregated through the price is sufficiently (un-)favorable.

Second, we use this equilibrium characterization to represent the expected return of the asset by means of a risk-neutral probability measure. Applying the Law of Total Variance, we show that the updating wedge naturally leads to *excess weight on tail risks*: the difference between the risk-neutral and the objective prior distribution of fundamentals decomposes into a shift in means that accounts for average supply effects akin to risk premia in a model with risk averse investors, and a mean-preserving spread that captures the additional effects of dispersed information and limits to arbitrage: securities characterized by upside (downside) risk are priced above (below) their fundamental value. Moreover, the over-pricing of upside or under-pricing of downside risks scales with excess weight on tail risks.

Third, we relate excess weight on tail risks to two observable moments of return forecasts, forecast dispersion and forecast accuracy. In the model, almost all the variation in excess weight on tail risks is linked to forecast dispersion, which in turn allows us to interpret forecast dispersion as a natural proxy for excess weight on tail risks. This allows us to discipline excess weight on tail risks with earnings forecast data, test the main predictions of our model, and offer a novel interpretation of existing empirical results.

In section 4, we thus show that our theory is consistent with several, seemingly unrelated, asset pricing puzzles:

1. Returns to skewness: Consistent with the first core prediction of our model, a large empirical literature documents a negative relation between skewness of the return distribution and expected returns in equity markets.[2]

2. Credit spread puzzle: Corporate credit spreads are very large relative to the underlying default risks, and difficult to reconcile with compensation for default risk, especially for high investment grade bonds.[3] In our model, credit spreads overweigh default risk, and the resulting ratio of spreads to expected default losses can become arbitrarily large for highly rated bonds.

3. Returns to disagreement: The empirical evidence on returns to investor disagreement is divided. Several studies find negative returns to disagreement in equity markets, which are typically interpreted in support of heterogeneous priors models with short-sales constraints in which securities are over-priced due to an implicit re-sale option whose value is always increasing with forecast dispersion (Miller, 1977). Others find positive returns to disagreement in bonds markets, and interpret disagreement as a proxy for risk.[4] Our model explains why higher disagreement can lead to lower equity returns but higher bond returns, consistent with empirical evidence. It offers a unified explanation for these seemingly contradictory empirical results by identifying upside vs. downside risk as the key determinant for signing the returns to disagreement.

4. Interaction effects: several studies suggest that returns to disagreement interact with proxies for return asymmetry, such as the value premium for equity, or measures of leverage and default risk for bonds.[5] Our theory provides a unified explanation for such interaction effects.

We then calibrate the informational parameters of our model to match dispersion of analysts' earnings forecasts and argue that even moderate degrees of information frictions can account for a quantitatively significant and empirically plausible fraction of the observed returns.

[2]See Conrad, Dittmar and Ghysels (2013), Boyer, Mitton and Vorkink (2010) and Green and Hwang (2012).

[3]See Huang and Huang (2012), Feldhütter and Schaefer (2018) and Bai, Goldstein and Yang (2020) for recent contributions.

[4]See Diether, Malloy, and Sherbina (2002), and Gebhardt, Lee, and Swaminathan (2001), and Yu (2011) for returns to disagreement in equity markets, Guntay and Hackbarth (2010) for bond markets and Carlin, Longstaff and Matoba (2014) for mortgage-backed securities.

[5]Yu (2011) reports that the value premium increases by 6.9% p.a. between the lowest and highest disagreement terciles, and that the returns to disagreement decrease by 7% p.a. from the highest to the lowest book-to-market ratio equity quintiles. Guntay and Hackbarth (2010) report that disagreement has far larger impact on bond spreads and bond returns for firms with high leverage or low credit ratings.

We use I.B.E.S. data on analysts' earning forecasts to measure analyst forecast dispersion and accuracy and impute excess weight on tail risks at the firm level for a cross-section of equity returns. These measures suggest that information frictions and excess weight on tail risks are highly skewed: for most firms, excess weight on tail risks appears to be negligible, but for firms in the top quintile, our measure of excess weight on tail risks can be very significant. For the median firm, excess weight on tail risks corresponds to less than 2% fo the standard deviation of earnings, but the sample average is respectively at 2.7% and 8.4% in two different forecast samples that we study, and 2-3 times as large in the top quintile of the distribution.

We calibrate the precision of private signals and the variance of supply shocks in the model to match the range of excess weight on tail risks suggested by the I.B.E.S. data, and parametrize the asset return distribution to match idiosyncratic skewness and volatility reported for different portfolios in Conrad, Dittmar and Ghysels (2013) and Boyer, Mitton and Vorkink (2010). Our model generates returns to skewness ranging from 34% to 100% of the empirical returns to skewness for empirically plausible levels of information frictions across the two studies. In addition, we account for about 60% of the returns to disagreement in equity markets and a similar fraction of the observed interaction effects.

We also calibrate our model to match observed default risks on corporate bonds and show that plausible levels of information frictions account for up to 35% of the levels and variations in credit spreads for high quality corporate bonds. This is in line with reduced form evidence suggesting that investor disagreement accounts for roughly 20-30% of this variation in the data.[6] Taken together, our quantitative results suggests that information frictions may play a significant role in accounting for a number of seemingly unrelated asset pricing anomalies.

Finally, in section 5, we generalize the key steps of our theoretical argument. We completely dispense with parametric assumptions about economic primitives, allowing for generic distributions of asset fundamentals, supply shocks and investor preferences. We provide a representation of asset prices by means of a risk-neutral measure that decomposes returns into a component due to risk premia and a component due to noisy information aggregation. The latter naturally generalizes the notion of excess weight on tail risks described above. We then identify two sufficient conditions on the equilibrium under which the model gives rise to excess weight on tail risks. These sufficient conditions are satisfied by all canonical applications of noisy information aggregation models in the existing literature.

[6] We also provide a numerical example which shows that our model is able to fully account for the observed spreads with extreme levels of information frictions.

Our paper contributes to the literature on noisy information aggregation in asset markets by offering a variant of the canonical noisy rational expectations model that dispenses with strong parametric assumptions about asset returns. This enables us to derive a characterization of asset prices and returns with noisy information aggregation by means of a risk-neutral probability measure that displays excess weight on tail risks. These results offer a tractable alternative to existing workhorse models that typically rely on restrictive parametric assumptions about preferences, information and asset returns (such as CARA preferences, or normally distributed signals and dividends), and they provide novel quantifiable implications of noisy information aggregation for asset returns. Furthermore, the results in section 5 show that these properties are a general consequence of noisy information aggregation in asset returns, and not due to specific functional form assumptions about information and risk preferences.

Breon-Drish (2015) analyzes non-linear and non-normal variants of the noisy REE framework in the broad exponential family of distributions and CARA preferences. He further derives powerful results on the incentives for information acquisition, whereas we take the information structure as given. Barlevy and Veronesi (2003), Peress (2004) and Yuan (2005) also study non-linear models of noisy information aggregation with a single asset market. Malamud (2015) and Chabakauri, Yuan and Zachariadis (2020) study information aggregation in non-linear, multi-asset noisy REE models with a rich set of state-contingent securities, exploiting spanning properties of state prices with complete or incomplete markets. In contrast to our work, these papers all impose parametric assumptions on the underlying asset returns, probability, information and preference structure to fully characterize the information content of asset prices, rather than identifying properties of asset prices that apply beyond the specifics of their environment.

Our equilibrium characterization with noisy information aggregation by means of a sufficient statistic variable shares similarities with common value auctions; these similarities are even more pronounced in the case with risk-neutral agents and position bounds.[7] Yet whereas the auctions literature seeks to explore under what conditions prices converge to the true fundamental values when the number of bidders grows large, we focus instead on the departures from this competitive limit that arise with noise and information frictions. In other words, rather than emphasizing perfect information aggregation at the competitive limit, we emphasize the impact of frictions and noise on prices away from this limit.

[7]See Wilson (1977), Milgrom (1979, 1981b), Pesendorfer and Swinkels (1997), Kremer (2002) and Perry and Reny (2006) for important contributions to this literature. However our results do not require the restrictions of risk-neutrality or unit demand that typically characterize the auction-theoretic literature.

More generally, any theory of mispricing must rely on some source of noise affecting the market, coupled with limits to investors' ability or willingness to exploit arbitrage (see Gromb and Vayanos, 2010, for an overview and numerous references). We show that noise trading with dispersed information leads not just to random price fluctuations, but to systematic, predictable departures of the price from the asset's fundamental value. This result is independent of the exact nature of the limits to arbitrage imposed by the model.

2 Agents, assets, information structure and financial market

There is a single risky asset with dividends given by a strictly increasing function $\pi(\cdot)$ of a stochastic fundamental θ, which is normally distributed according to $\theta \sim \mathcal{N}\left(0, \sigma_\theta^2\right)$, where σ_θ^2 denotes the variance of fundamentals. The asset supply is stochastic and equal to $s = \Phi(u)$, where $\Phi(\cdot)$ is the cdf of a standard normal distribution, and $u \sim \mathcal{N}\left(\bar{u}, \sigma_u^2\right)$. This supply assumption is adapted from Hellwig, Mukherji and Tsyvinski (2006) and preserves the normality of updating from private signals and market price without imposing any restrictions on the shape of $\pi(\cdot)$.

There is a unit measure of risk-neutral, informed investors who each observe a noisy private signal $x_i = \theta + \varepsilon_i$, where ε_i is i.i.d across agents, and distributed according to $\varepsilon_i \sim \mathcal{N}\left(0, \beta^{-1}\right)$, where β denotes the precision of private signals. Upon observing their signal, these investors submit price-contingent demand schedules $d(x,P)$, corresponding to a combination of limit orders to bid for the available supply. We assume that their positions are restricted by position limits that restrict each investor's demand to the unit interval $[0,1]$. In other words, investors cannot short-sell and can buy at most one unit of the asset.

Once investors submit their orders, trades are executed at a price P that is selected to clear the market. A price P clears the market if and only if

$$s = D(\theta, P) \equiv \int d(x,P)\, d\Phi(\sqrt{\beta}(x-\theta)).$$

Let $H(\cdot|P)$ denote the posterior cdf of θ, conditional on observing the market price P, and $H(\cdot|x,P)$ the investors' posterior conditional on x and P. Given $H(\cdot|x,P)$, a demand function $d(x,P)$ is optimal, if it solves the investors' decision problem

$$\max_{d\in[0,1]} d\left(\int \pi(\theta) dH(\theta|x,P) - P\right).$$

A *Perfect Bayesian Equilibrium* consists of a demand function $d(x,P)$, a price function $P(\theta,s)$, and posterior beliefs $H(\cdot|P)$ such that (i) $d(x,P)$ is optimal given $H(\cdot|x,P)$; (ii) $P(\theta,s)$ clears the market for all $(\theta, s) \in \mathbb{R} \times [0,1]$; and (iii) $H(\cdot|P)$ satisfies Bayes' rule whenever applicable, i.e., for all P such that $\{(\theta, s) : P(\theta, s) = P\}$ is non-empty.

We conclude the model description with two remarks about modeling choices.

First, in our model, investors do not observe signals directly about asset payoffs but rather about a fundamental variable θ, and the asset payoff is some monotone function of θ.[8] This formulation separates the distribution of asset returns from the investors' updating of beliefs, which strikes us as a reasonable approximation of many real world financial markets. For example, equity analysts gather information about a firm's earnings, investment opportunities, returns to capital etc. all of which eventually affect the dividend payouts to shareholders. A bond analyst will assess the issuer's solvency which depends on revenues, but also debt service, leverage etc., often summarized in a single "distance to default" variable. An option trader will forecast where the underlying is heading. In all these cases, the fundamental about which a trader gathers information is distinct from the security's payoffs and the mapping from fundamentals to payoffs is typically non-linear. In some cases, such as different options on the same underlying asset, bonds of different maturity or seniority, or firms that issue both equity and debt, the same fundamentals affect different securities differently. Our model is flexible enough to accommodate any of these possibilities.

Second, assuming risk neutrality and imposing position limits is of special interest because risk preferences disappear from the equilibrium characterization, and asset pricing implications are driven exclusively by noisy information aggregation. The risk-neutral model with position limits strikes us as a natural laboratory for studying cross-sectional return predictions with noisy information aggregation, since investors should be able to diversify asset-specific risks by investing across a wide range of assets. Such diversification can be achieved by limiting the wealth that is invested in any given security, akin to position limits in our model.[9]

3 Equilibrium characterization

Our analysis proceeds in three steps: (i) a representation of the equilibrium price in terms of a sufficient statistic z which highlights that the price responds more strongly to shocks than would be warranted purely on informational grounds (Section 3.1), (ii) a risk-neutral representation of expected prices and returns that displays excess weight on tail risks (Section 3.2), along with comparative statics that derive from the interaction between excess weight on tail risks and return

[8] If $\pi(\theta) = \theta$, our model reduces to the canonical formulation in which investors observe noisy signals of dividends.

[9] See Albagli, Hellwig and Tsyvinski (2017) for further discussion.

asymmetries (Section 3.3), and (iii) a mapping of excess weight on tail risks to observable moments of earnings forecast accuracy and forecast dispersion (Section 3.4).

3.1 Sufficient statistic representation of the equilibrium price

Standard arguments imply that $H(\cdot|x,P)$ is first-order stochastically increasing in the investor's signal x. Given risk-neutrality, there then exists a unique signal threshold $\hat{x}(P)$ such that

$$\int \pi(\theta)\, dH(\theta|x,P) \gtreqless P \quad \text{if and only if} \quad x \gtreqless \hat{x}(P),$$

and investors demand $d(x,P) = 0$ if $x < \hat{x}(P)$ and $d(x,P) = 1$ if $x > \hat{x}(P)$, and $d(\hat{x}(P), P) \in [0,1]$. Therefore, aggregate demand by informed investors equals $D(\theta,P) = 1 - \Phi(\sqrt{\beta}(\hat{x}(P) - \theta))$. Setting $z \equiv \hat{x}(P)$, market-clearing then implies $1 - \Phi(\sqrt{\beta}(z-\theta)) = s = \Phi(u)$, or equivalently $z = \theta - 1/\sqrt{\beta} \cdot u$, and $z|\theta \sim \mathcal{N}\left(\theta - \bar{u}/\sqrt{\beta}, \tau^{-1}\right)$, where $\tau \equiv \beta/\sigma_u^2$. The variable z represents the information conveyed by the price, and τ the precision of the price signal.

Our first proposition provides a closed form characterization of the unique equilibrium in which demand is non-increasing in P, consistent with executing trades through limit orders.

Proposition 1 *: The unique equilibrium in which the investors' demand is non-increasing in P is characterized by the equilibrium price function*

$$P_\pi(z) = \mathbb{E}(\pi(\theta)|x=z,z) = \int \pi(\theta) d\Phi\left(\frac{\theta - \gamma_P \cdot \left(z + \frac{\tau}{\beta+\tau}\bar{u}/\sqrt{\beta}\right)}{\sigma_\theta/\sqrt{1-\gamma_P}}\right) \tag{1}$$

$$\textit{where } \gamma_P \equiv \frac{\beta+\tau}{1/\sigma_\theta^2 + \beta + \tau}.$$

Proposition 1 represents the equilibrium price in terms of an endogenous state variable or sufficient statistic z for the information conveyed through the asset price. This sufficient statistic corresponds to the private signal of the marginal investor who is just indifferent between buying and not buying the asset.

The expected dividend conditional on z instead takes the form

$$\mathbb{E}(\pi(\theta)|z) = \int \pi(\theta) d\Phi\left(\frac{\theta - \gamma_V \cdot (z + \bar{u}/\sqrt{\beta})}{\sigma_\theta/\sqrt{1-\gamma_V}}\right) \quad \text{where} \quad \gamma_V \equiv \frac{\tau}{1/\sigma_\theta^2 + \tau}. \tag{2}$$

Hence the equilibrium price differs from the expected dividend by responding more strongly to the market signal z than would be justified purely based on the information conveyed through z. We term this excess price sensitivity the *updating wedge*.[10]

[10] To our knowledge, the only prior discussion of this updating wedge is by Vives (2008), in the context of the CARA-normal model.

This updating wedge results from market clearing with dispersed information and is perfectly consistent with Bayesian rationality. The equilibrium price is represented as the marginal investor's dividend expectation. An increase from z to z' (due to an increase in θ or a decrease in s) conveys positive news about θ and thus raises dividend expectations for all traders in the market. In addition, an increase in θ or a decrease in s shift the asset demand and supply. To clear the market, the private signal of the marginal investor pricing the asset then has to increase from z to z', hence the marginal investor becomes even more optimistic relative to the other investors in the market. The expression for the equilibrium price incorporates these two effects through the sufficient statistic z appearing once as the price signal and once as the marginal investor's private signal. In contrast, the Bayesian posterior of θ given P only includes the first effect.

3.2 Risk-neutral measure and excess weight on tail risks

We use the equilibrium characterization of Proposition 1 to represent the expected price $\mathbb{E}(P(z)) = \int \pi(\theta)\, d\hat{H}(\theta)$ as an expectation of dividends under a risk-neutral measure $\hat{H}$.[11] We then derive empirical predictions from the comparison between the risk-neutral and the physical prior.

Compounding the risk-neutral posterior $H(\cdot|x = z, z)$ with the prior $z \sim \mathcal{N}\left(-\bar{u}/\sqrt{\beta}, \sigma_\theta^2/\gamma_V\right)$, the risk-neutral measure $\hat{H}(\cdot)$ is normal with mean $\bar{\theta} \equiv -\frac{\sqrt{\beta}}{1/\sigma_\theta^2+\beta+\tau}\bar{u}$ and variance

$$\sigma_P^2 = \left(1 - \gamma_P + \gamma_P^2/\gamma_V\right)\sigma_\theta^2 > \sigma_\theta^2.$$

The risk-neutral prior differs from the physical prior through a shift in means and a mean-preserving spread, a property that we refer to as *excess weight on tail risks*: controlling for the mean $\bar{\theta}$, $\hat{H}(\cdot)$ places higher weight on realizations of θ in both upper and lower tails. This property distinguishes the risk-neutral measure under dispersed information from a counterpart that incorporates pure risk premia or aggregate supply effects, which merely shift probability mass from the upper to the lower tail realizations, analogous to the shift $\bar{\theta}$ in the mean of the distribution.

The following representation of σ_P^2 provides some intuition for excess weight on tail risks:

$$\begin{aligned}
\sigma_P^2 &= Var\left(\mathbb{E}(\theta|x=z,z)\right) + \mathbb{E}\left(Var(\theta|x=z,z)\right) = Var\left(\mathbb{E}(\theta|x=z,z)\right) + \mathbb{E}\left(Var(\theta|x,z)\right) \\
&= Var\left(\theta + \mathbb{E}(\theta|x=z,z) - \mathbb{E}(\theta|x,z)\right) > \sigma_\theta^2.
\end{aligned}$$

Here, the first equality applies Blackwell's variance decomposition to the risk-neutral measure. The second equality uses the fact that $Var(\theta|x,z)$ is independent of x. The third equality follows from

[11] In contrast to the risk-neutral measure with no-arbitrage, the measure $\hat{H}$ includes asset-specific factors relating to the severity of information aggregation frictions in a specific market.

$cov\left(\theta,\mathbb{E}\left(\theta|x,z\right)|x,z\right)=0$. The last inequality follows from the fact that $\mathbb{E}\left(\theta|x=z,z\right)-\mathbb{E}\left(\theta|x,z\right)$ is linear in $x-z$ and independent of θ.

Excess weight on tail risks arises if and only if $Var\left(\mathbb{E}\left(\theta|x=z,z\right)\right)>Var\left(\mathbb{E}\left(\theta|x,z\right)\right)$ and is thus equivalent to another equilibrium property of noisy information aggregation models: namely, that conditional expectations under the risk-neutral measure are strictly more variable than the expectations of an arbitrary informed trader in the market. But this result emerges because supply shocks introduce fluctuations in risk-neutral expectations that are orthogonal to the private signals of fundamentals that are aggregated through the price.

3.3 Asset pricing implications of excess weight on tail risks

Taking expectations and using integration by parts, we represent the expected price premium as $\mathbb{E}\left(P_{\pi}\left(z\right)\right)-\mathbb{E}\left(\pi\left(\theta\right)\right)=R\left(\pi;\bar{\theta},\sigma_P\right)+W\left(\pi,\sigma_P\right)$, where

$$R\left(\pi;\bar{\theta},\sigma_P\right) \equiv \int_{-\infty}^{\infty}\left(\pi\left(\theta+\bar{\theta}\right)-\pi\left(\theta\right)\right)d\Phi\left(\frac{\theta}{\sigma_P}\right) \tag{3}$$

$$W\left(\pi;\sigma_P\right) \equiv \int_{-\infty}^{\infty}\left(\pi\left(\frac{\sigma_P}{\sigma_\theta}\theta\right)-\pi\left(\theta\right)\right)d\Phi\left(\frac{\theta}{\sigma_\theta}\right). \tag{4}$$

The term $R\left(\pi;\bar{\theta},\sigma_P\right)$ summarizes the impact of the shift in means which is governed by the average asset supply.[12] The term $W\left(\pi;\sigma_P\right)$ summarizes the impact of the mean-preserving spread on the expected price premium. For reasons that will quickly become apparent, we will refer to $W\left(\pi,\sigma_P\right)$ as the skewness premium. Our next definition provides a partial order on returns that we use for the comparative statics of $W\left(\pi,\sigma_P\right)$.

Definition 1 ***(Cash flow risks)*** *:*
(i) Dividend function π has symmetric risk if $\pi\left(\theta_1\right)-\pi\left(\theta_2\right)=\pi\left(-\theta_2\right)-\pi\left(-\theta_1\right)$ for all $\theta_1>\theta_2\geq 0$.
(ii) Dividend function π is dominated by upside risk if $\pi\left(\theta_1\right)-\pi\left(\theta_2\right)\geq\pi\left(-\theta_2\right)-\pi\left(-\theta_1\right)$, and dominated by downside risk if $\pi\left(\theta_1\right)-\pi\left(\theta_2\right)\leq\pi\left(-\theta_2\right)-\pi\left(-\theta_1\right)$, for all $\theta_1>\theta_2\geq 0$.
(iii) Dividend function π_1 has more upside risk than π_2 if $\pi_1-\pi_2$ is dominated by upside risk.

This definition classifies payoff functions by comparing marginal gains and losses at fixed distances from the prior mean to determine whether the payoff exposes its owner to bigger payoff fluctuations on the upside or the downside. Any linear dividend function has symmetric risks, any

[12]With position limits, the average supply effect captures the role of "cash-in-the-market" or short-sales constraints, by which average prices are a decreasing function of the expected asset supply. With risk-averse investors and no position limits, the risk premium generates a similar average supply effect.

convex function is dominated by upside risks, and any concave dividend function is dominated by downside risks, but the classification also extends to non-linear functions with symmetric gains and losses, non-convex functions with upside risk or non-concave functions with downside risk.

Securities are easy to classify according to this definition when the fundamental and the return are both observable (for example in the case of defaultable bonds or options). But even without observing fundamentals directly, upside and downside risk directly translate into the distribution of returns being more spread out above or below the median of the return distribution, if the fundamental distribution is symmetric. Intuitively, this means that a security that is dominated by upside (downside) risk has positive (negative) skewness.

Proposition 2 ***(Sign and comparative statics of*** $W(\pi, \sigma_P)$***):***

(i) If π *has symmetric risk, then* $W(\pi; \sigma_P) = 0$*. If* π *is dominated by upside risk, then* $W(\pi; \sigma_P)$ *is positive and increasing in* σ_P*. If* π *is dominated by downside risk, then* $W(\pi; \sigma_P)$ *is negative and decreasing in* σ_P*. Moreover,* $lim_{\sigma_P \to \infty} |W(\pi; \sigma_P)| = \infty$*, whenever* $lim_{\theta \to \infty} |\pi(\theta) + \pi(-\theta)| = \infty$*.*

(ii) If π_1 *has more upside risk than* π_2*, then* $W(\pi_1; \sigma_P) - W(\pi_2; \sigma_P)$ *is non-negative and increasing in* σ_P*.*

Proposition 2 shows how the skewness premium arises as a combination of asymmetry in the dividend profile π and noisy information aggregation, or excess weight on tail risks ($\sigma_P > \sigma_\theta$). The skewness premium is positive for upside risks, negative for downside risks and larger for assets with larger return asymmetries. Furthermore, the skewness premium increases and can grow arbitrarily large as excess weight on tail risks becomes more important, provided that the return asymmetry does not vanish from the tails. Finally, price impact of information aggregation frictions and return asymmetry are mutually reinforcing, since the skewness premium has increasing differences in upside risk and excess weight on tail risks.

Mathematically, these results follow directly from our interpretation of the skewness premium as a mean-preserving spread. This mean-preserving spread becomes more valuable when the payoff function shifts towards more upside risk.

The intuition for proposition 2 comes from the interaction of $\pi(\cdot)$ with excess weight on tail risks. With symmetric $\pi(\cdot)$, excess weight on tail risks does not affect the average price: over-pricing when z is high is just offset by under-pricing when z is low. When instead $\pi(\cdot)$ is dominated by upside risk, the upper tail risk of dividends is more important than the lower tail. Hence over-pricing on the upside is larger than under-pricing on the downside, which results in a positive price premium. When instead $\pi(\cdot)$ is dominated by downside risk, the price premium is negative.

Below, we map excess weight on tail risks to forecast dispersion. Our model thus predicts a negative relation between skewness or upside risk and returns, whose strength is reinforced by forecast dispersion. Moreover, it derives a positive relation between forecast dispersion and returns for securities with downside risks. These predictions clearly distinguish our theory from heterogeneous priors models in which forecast dispersion proxies for investor disagreement, which generates a positive option value of resale regardless of a securities return structure. We will return to these predictions below when relating our model to the empirical evidence.

3.4 Quantifying excess weight on tail risks

We now map the ratio σ_P/σ_θ into two statistics that can be estimated using data on investor's forecasts of fundamentals. This will allow us to translate the comparative statics of Proposition 2 into empirically testable predictions. We can interpret $\gamma_P = 1 - \left(1/\sigma_\theta^2\right)/\left(1/\sigma_\theta^2 + \beta + \tau\right)$ as a measure of *forecast accuracy*, and construct an empirical counterpart by computing the forecast error variance, relative to the variance of earnings. In addition, we define *forecast dispersion* $\tilde{D} \equiv \frac{\sqrt{\beta}}{\left(1/\sigma_\theta^2+\beta+\tau\right)}\frac{1}{\sigma_\theta}$ as the cross-sectional standard deviation of forecast errors normalized by the standard deviation of fundamentals. This yields the following representation of excess weight on tail risks in terms of forecast dispersion and forecast accuracy:

$$\frac{\sigma_P}{\sigma_\theta} = \sqrt{1 + \tilde{D}^2 \frac{\gamma_P\left(1-\gamma_P\right)}{\gamma_P\left(1-\gamma_P\right) - \tilde{D}^2}} \tag{5}$$

Representation (5) uses only fundamentals and forecast data to discipline excess weight on tail risks, without relying on any market data (asset returns and prices). The representations of $\tilde{D}$ and γ_P in turn allow us to back out the primitive parameters β and σ_u^2. We can therefore calibrate these informational parameters using data on earnings and earnings forecasts, and then test the model asset pricing implications on market data.[13]

Furthermore, equation (5) implies that forecast dispersion drives almost all the variation in σ_P/σ_θ, unless forecast accuracy takes on extreme values. Conditional on $\tilde{D}$, σ_P/σ_θ displays a symmetric U-shape in forecast accuracy, reaching a minimum at $\gamma_P = 1/2$, and diverging to infinity when $\gamma_P\left(1-\gamma_P\right) \to \tilde{D}^2$. But for most of its range, the increase in σ_P/σ_θ is very mild as we move

[13]The chosen representation in terms of forecast accuracy and forecast dispersion is by no means the only representation possible. For example initially we defined σ_P/σ_θ in terms of the response coefficients γ_P and γ_V, where γ_P is interpreted as forecast accuracy, and γ_V represents a simple measure of *market accuracy* or price informativeness, which can be computed as the R^2 obtained by regressing fundamentals (earnings) on prices. The same parameters also matter for price volatility or return predictability.

away from $\gamma_P = 0.5$. For example, consider a value of $\tilde{D} = 0.2$, which is close to the median value of forecast dispersion in the sample of Guntay and Hackbarth (2010). This implies a minimum value of $\sigma_P/\sigma_\theta = 1.0235$. But σ_P/σ_θ does not exceed 1.0244 for values of forecast accuracy in the range of $[0.25, 0.75]$. Therefore, the lower bound on σ_P/σ_θ that we obtain by setting $\gamma_P = 0.5$ is actually a very good approximation for a remarkably wide range of forecast accuracies, and forecast dispersion $\tilde{D}$ emerges as a natural proxy variable for excess weight on tail risks.[14]

The key steps in this section were the characterization of a risk-neutral measure that displays excess weight on tail risks and the characterization of asset pricing implications of excess weight on tail risks by Proposition 2. Hence, any theory that rationalizes excess weight on tail risks of the risk neutral measure will yield similar asset pricing predictions. Noisy information aggregation allows us to tie excess weight on tail risks to observable moments of forecast dispersion and accuracy, which leads to an immediate test of the theory. In Section 5, we generalize these steps beyond the present model with risk-neutral investors and position limits.

4 Dispersed information and asset pricing puzzles

Proposition 2 offers three qualitative predictions: returns to skewness, returns to disagreement, and positive interaction effects. In this section, we first discuss empirical evidence supporting these predictions qualitatively. We then calibrate excess weight on tail risks to match forecast dispersion and forecast accuracy, and explore to what extent our model is also able to quantitatively account for empirical facts.[15]

Prediction 1 (Returns to skewness): *Price premia are positive (negative) and return premia negative (positive) for securities dominated by upside (downside) risk. Price premia are increasing and returns decreasing with skewness or upside risk.*

A sizable empirical literature documents a negative relationship between expected skewness and equity returns. For example, Conrad, Dittmar and Ghysels (2013; CDG henceforth) estimate skewness of equity returns from option prices, Boyer, Mitton and Vorkink (2010) from forecasting

[14] In principle, if we represent excess weight on tail risks as a function of two empirical proxies, empirical tests w.r.t. one proxy are valid only if they control for the other - otherwise there is likely to be omitted variable bias if the two proxies are correlated with each other. However, since forecast accuracy has so little effect over much of its range, omitted variable bias is much less of a concern when we test predictions regarding forecast dispersion. This offers further validation for using forecast dispersion as a proxy for excess weight on tail risks.

[15] Our model may potentially account also for options prices, announcement effects or excess price volatility, but addressing those would take us too far afield from the main predictions of Proposition 2.

regressions. Both studies then sort stocks by expected skewness and find that securities with higher skewness earn about 0.7% lower average returns per month, equivalent to more than 8% of yearly excess returns for the strategy of going long/short on low/high skewness stocks. Green and Hwang (2012) find that IPOs with high expected skewness earn significantly more negative abnormal returns in the following one to five years. Zhang (2013) shows that skewness correlates positively with the book-to-market factor and thus helps account for the value premium.

Existing explanations for these empirical findings rely on investor preferences for positive skewness.[16] Proposition 2 offers an alternative explanation for a negative returns to skewness as the result of dispersed information and excess weight on tail risks. In contrast to preference-based theories, this explanation also links returns to skewness and disagreement.

Returns to skewness also manifest themselves in bond markets through the *credit spread puzzle*, i.e. the difficulty to reconcile high levels of corporate bonds spreads with historical default data in standard models pricing credit risk. Huang and Huang (2012) calibrate a number of structural models to historical default data and show that they all produce spreads relative to treasuries that fall significantly short of their empirical counterparts. This shortfall is most severe for short maturity, high investment grade securities, which are almost as safe as treasuries of similar maturity, yet pay significantly larger return premia.[17] Proposition 2 offers an intuitive explanation for this

[16]In Brunnermeier and Parker (2005) and Brunnermeier, Gollier and Parker (2007), overinvestment in highly skewed securities, along with under-diversification, results from a model of optimal expectations. Barberis and Huang (2008) show that cumulative prospect theory results in a demand for skewness or a preference for stocks with lottery-like features. Mitton and Vorkink (2007) develop a model in which investors have heterogeneous preference for skewness.

[17]Among recent contributions to this literature, Chen (2010) considers a dynamic model with time variation in macroeconomic conditions in a context of long-run risks, endogenous default decisions by firms, and recursive preferences. He and Milbradt (2014) study the role of interacting market liquidity with endogenous default decisions. Chen et al. (2018) show that the interaction of all these mechanisms can deliever credit spreads that come closer to the empirical counterparts. However, most purely risk- and liquidity-based models account for at most a small fraction of the level and volatility of spreads that are observed in practice, especially so for short-horizon investment grade bonds.

Our predictions for credit spreads also speak to the empirical asset pricing literature linking credit spreads and equity returns (See Bhamra, Kuehn and Strebulaev, 2010 and citations therein). Bhamra, Kuehn and Strebulaev (2010) reconcile co-movement of equity returns and credit spreads through a structural model of capital structure with Epstein-Zin preferences and time-varying counter-cyclical default risks. McQuade (2018) obtains similar results by introducing stochastic volatility as a second risk factor that is priced by the representative investor. Our analysis instead links credit spreads and (levered) equity returns through dispersed investor information and excess weight on tail risks, which allows us to connect with a different set of stylized facts regarding returns to investor disagreement. We leave a multi-asset version of our noisy information aggregation model to study return co-movement and capital structure choices for future work.

puzzle by arguing that with dispersed information the market has a tendency to be overly concerned with the tail risk of default, relative to actual default probabilities.

Prediction 2 (Returns to disagreement): *Returns to disagreement are negative (positive) with upside (downside) risks.*

A growing literature uses forecast dispersion as a proxy for disagreement. Diether, Malloy and Sherbina (2002; DMS henceforth) sort stocks by the dispersion of earnings forecasts across analysts covering each security. They find that stocks in the highest dispersion quintile have monthly returns which are about 0.62% lower than those in the lowest dispersion quintile, amounting to a yearly excess return over 7% for the strategy of going long/short on low/high dispersion stocks. They interpret this result as evidence consistent with the hypothesis of Miller (1977) of investor disagreement interacting with short-selling constraints.[18] Yu (2011) reports similar results and Gebhardt, Lee and Swaminathan (2001) document that an alternative measure of stock risk premia, the cost of capital, is also negatively related to analyst forecast dispersion.

Guntay and Hackbarth (2010; GH henceforth) perform a similar analysis for bond yields but reach the opposite conclusion as DMS: yield spreads and bond returns are increasing with forecast dispersion, and spreads are 0.14% higher and returns 0.08% higher in the top dispersion quintile, which amounts to a yearly excess return of about 1% for the strategy long/short on high/low dispersion bonds. GH replicate DMS' result of negative returns to disagreement in equity returns in their sample (though the measured excess returns are smaller), which suggests a systematic difference in returns to disagreement for equity and bond markets. GH interpret returns to disagreement as a proxy for risk premia. Carlin, Longstaff and Matoba (2014) confirm GH's results for mortgage-backed securities.

Proposition 2 reconciles these seemingly contradictory empirical results by noting that studies that find negative returns to disagreement focus on securities with upside risk, while studies that find positive returns to disagreement focus on securities where downside risk is dominant.

Prediction 3 (Interaction effects): *There is positive interaction between returns to disagreement and returns to skewness.*

Evidence on interaction effects between skewness and disagreement is more limited. Yu (2011) sorts stocks by book-to-market ratio and disagreement and reports that the value premium increases from 4.3% annual return with the lowest tercile disagreement to 11.3% with the highest tercile, and the returns to disagreement range from -0.26% annual for the highest quintile of book-to-market ratios to -7.2% for the lowest quintile. Following Zhang (2013) who interprets book-to-market

[18] They rule out a risk-based explanation for the anomaly by controling for stocks exposure to standard risk factors

ratios as a proxy for skewness, these results suggest substantial interaction between returns to skewness and disagreement in equity returns.[19]

For bond markets, GH report that the effect of disagreement on spreads and yields doubles in high leverage or low-rated rated firms, two plausible proxies for downside risks. In a regression of credit spreads on leverage, disagreement and their interaction, the interaction term turns out to be highly significant, but disagreement and leverage are insignificant on their own. These empirical results suggest that returns to skewness and disagreement interact in the data along the lines suggested by our theoretical results.[20]

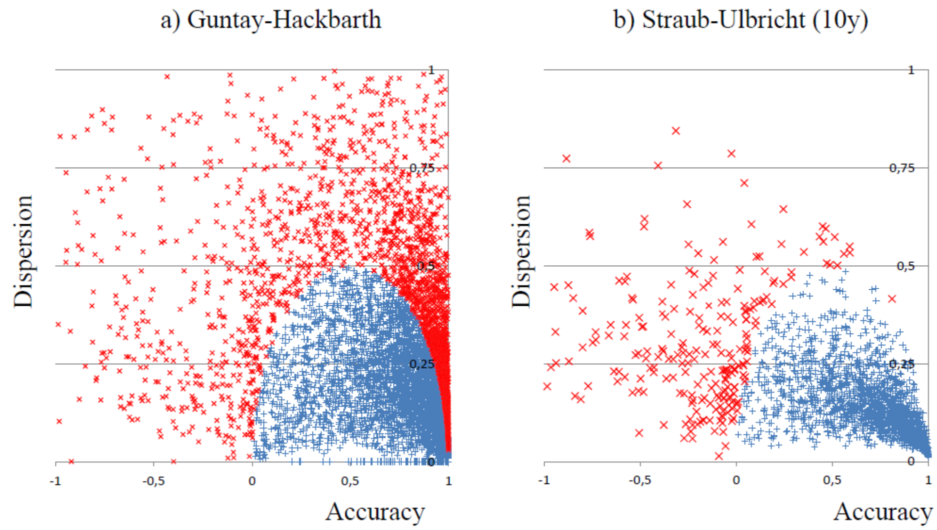
The ability to account for returns to skewness and disagreement in both equity and bond markets distinguishes our theory from heterogeneous prior models with short-sales constraints following Miller (1977): in those models prices incorporate a resale option value that lowers future returns irrespective of the asset characteristics. They are therefore unable to explain why these comparative statics would be different for different security classes such as stocks and bonds.

Indeed, the ambiguous empirical relationship between disagreement and asset returns remains one of the major unresolved puzzles in asset pricing. Perhaps Carlin, Longstaff and Matoba (2014) put it best: "Understanding how disagreement affects security prices in financial markets is one of the most important issues in finance. ...Despite the fundamental nature of this issue, though, there still remains significant controversy in the literature about how disagreement risk affects expected returns and asset prices." To our knowledge, ours is the first explanation that can reconcile the seemingly contradictory empirical results as direct predictions of a unified theory, tractable enough to encompass assets with different underlying cash-flow risks.

[19] More specifically, Zhang (2013) first documents strong positive correlation between book-to-market ratios and skewness of returns, and then shows that the explanatory power of book-to-market ratios for returns is significantly lower when controling for skewness.

[20] Hou, Xue, and Zhang (2020) question the statistical robustness of various return anomalies in equity markets including some of the studies we mention here. They replicate existing studies on a uniform sample and emphasize the importance of small capitalizations and equal vs. value weighting in estimating return anomalies. In their study, returns to skewness remain small and insignificant, but by focusing on the impact of skewness in realized returns, they do not directly replicate CDG or Boyer Mitton and Vorkink (2010), which use measures of expected future skewness. They replicate DMS and show that value-weighting leads to much lower and statistically insignificant returns to disagreement. This suggests that returns to disagreement are concentrated in markets with small capitalizations, a finding which is consistent with the replication of DMS by GH for a subsample of firms that are active also in bond markets. Finally, their estimates of the value premium are similar to the ones reported in Yu (2011). Our predictions are broadly consistent with the results of Hou, Xue and Zhang (2020), if one assumes that larger markets are more liquid and less subject to noisy information aggregation frictions.

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



4.1 Measuring information frictions

Here we construct simple measures of excess weight on tail risks from analysts' earnings forecasts. Using the representation (5), we infer excess weight on tail risks from measures of forecast dispersion $\tilde{D}$ and forecast accuracy γ_P for a cross-section of listed firms. These statistics are derived from the I.B.E.S. data base of analyst forecasts of earnings per share, but based on the data treatments in GH and Straub and Ulbricht (2015; SU henceforth).[21] The database reports a measure of forecast dispersion along with a consensus or average earnings forecast for each fim in the sample. Forecast accuracy is computed from average forecasts and realized earnings as the variance of forecast errors relative to the unconditional variance of earnings. In the GH sample, these measures are computed for all firms at a relatively short horizon (within quarter), using data from 1987-1998. In the SU sample, these measures are computed using the whole I.B.E.S. sample (1976-2016) for a subset of firms for which we have at least 10 years of forecast data, and forecasts are made over a slightly longer 8 month horizon.

Figure 1 shows scatter plots of forecast accuracy (horizontal axis) and forecast dispersion (vertical axis) in the two data sets, with negative forecast accuracy ($\gamma_P < 0$) corresponding to a case where the forecast error variance exceeds the volatility of earnings. The blue + marks correspond to firms that satisfy $\gamma_P (1 - \gamma_P) > \tilde{D}^2$, the red x-marks correspond to firms that do not satisfy this restriction. If this condition is not satisfied then the measured forecast dispersion and forecast accuracy cannot be consistent with our model (or more generally with any common prior, linear

[21]We thank these four authors for sharing their treatment of the raw data with us.

	Source	SU	restricted SU	GH	restricted GH
	N	**2103**	**1839 (87.4%)**	**6820**	**4335 (63.6%)**
Relative	**Mean**	**0.172**	**0.143**	**0.293**	**0.177**
Forecast	**St. Dev.**	**0.2**	**0.084**	**0.305**	**0.108**
Dispersion	10%	0.053	0.051	0.061	0.048
$\tilde{D}$	25%	0.085	0.079	0.117	0.092
	Median	**0.137**	**0.126**	**0.214**	**0.16**
	75%	0.213	0.189	0.368	0.251
	90%	0.315	0.257	0.601	0.342
Forecast	**Mean**	**0.53**	**0.689**	**0.611**	**0.748**
Accuracy	**St. Dev.**	**1.14**	**0.24**	**0.988**	**0.236**
γ_P	10%	0.012	0.325	0.124	0.378
	25%	0.397	0.526	0.54	0.627
	Median	**0.695**	**0.744**	**0.818**	**0.822**
	75%	0.873	0.89	0.947	0.937
	90%	0.954	0.959	0.984	0.978
Correlation	$Corr\left(\tilde{D}, \gamma_P\right)$	**-0.84**	**-0.519**	**-0.282**	**-0.34**

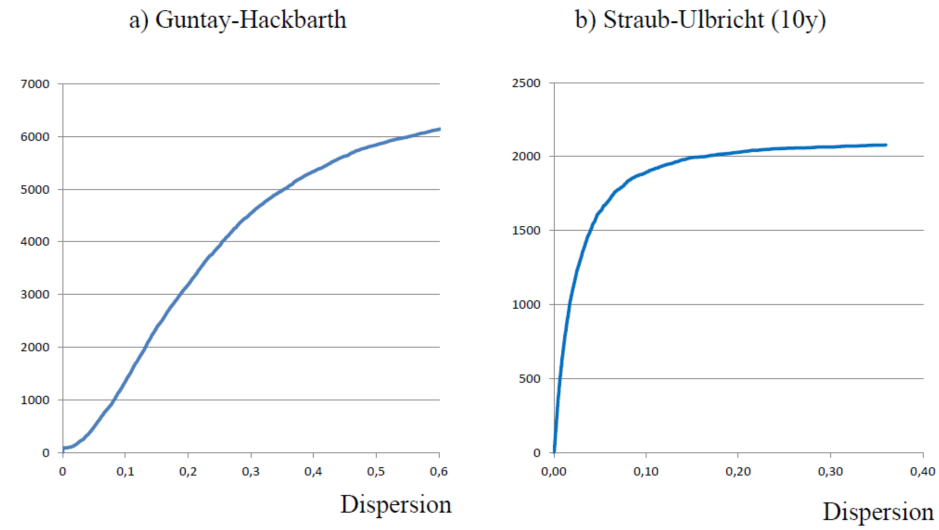
Table 1: Forecast Dispersion and Forecast Accuracy

projection model), and the friction parameter is not well defined.

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 $\gamma_P (1 - \gamma_P) > \tilde{D}^2$.

Table 1 reports summary statistics (mean, median, 10th, 25th, 75th and 90th percentiles) for forecast dispersion and forecast accuracy in the two datasets, as well as the raw correlation between these two measures. We report these statistics both for the full sample, and for a restricted sample of firms in which $\gamma_P (1 - \gamma_P) > \tilde{D}^2$. The proportion of firms that are excluded in the restricted sample drops from 36% in the GH sample to less than 13% in the SU sample, which only considers firms with at least 10 years of forecast data. This suggests that violations of the $\gamma_P (1 - \gamma_P) > \tilde{D}^2$ condition may stem from errors in the estimates of earnings volatility or forecast accuracy that are based on limited sample sizes. In particular, in contrast to GH, SU has no firms with high

Figure 2: Forecast Dispersion (Cumulative Distribution)



forecast dispersion and high forecast accuracy - the restricted SU sample only excludes firms in which measured forecast accuracy is implausibly low, given forecast dispersion.

Figure 2 displays the cross-sectional distribution of forecast dispersion in the two samples and shows that, consistent with the moments reported in Table 1, the distribution of forecast dispersion is highly skewed with a fat upper tail, and the mean of the distribution far higher than the median, in both samples. The distribution of forecast accuracy on the other hand is more uniformly spread with a significant lower tail. In addition, forecast dispersion and forecast accuracy are negatively correlated in the cross-section, but the correlation is far from perfect.

Table 2 reports summary statistics for the cross-sectional distributions of the excess weight on tail risks σ_P/σ_θ that is implied by the joint distribution of forecast dispersion and forecast accuracy in the restricted sample for which σ_P/σ_θ is well-defined. Figure 3 displays the corresponding cross-sectional distribution. The table and figure both confirm that the distribution of excess weight on tail risks is highly right-skewed in both samples: frictions are small for most firms (the median firm in the GH sample has less than 2% of excess weight on tail risks, $\sigma_P/\sigma_\theta < 1.02$, while the median firm in the SU sample has $\sigma_P/\sigma_\theta < 1.01$), but the mean level of frictions is much higher than the median (1.027 in SU and 1.084 in GH) and driven by an upper tail of firms for which measured information frictions can be fairly substantial.

Table 2 also provides summary statistics for the lower bound on σ_P/σ_θ constructed from forecast dispersion by setting forecast accuracy equal to 1/2, and reports the correlation between σ_P/σ_θ, forecast dispersion and accuracy, and this lower bound. Excess weight on tail risks is only weakly

Figure 3: Excess Weight on Tail Risks (Cumulative Distribution)

a) Guntay-Hackbarth b) Straub-Ulbricht (10y)

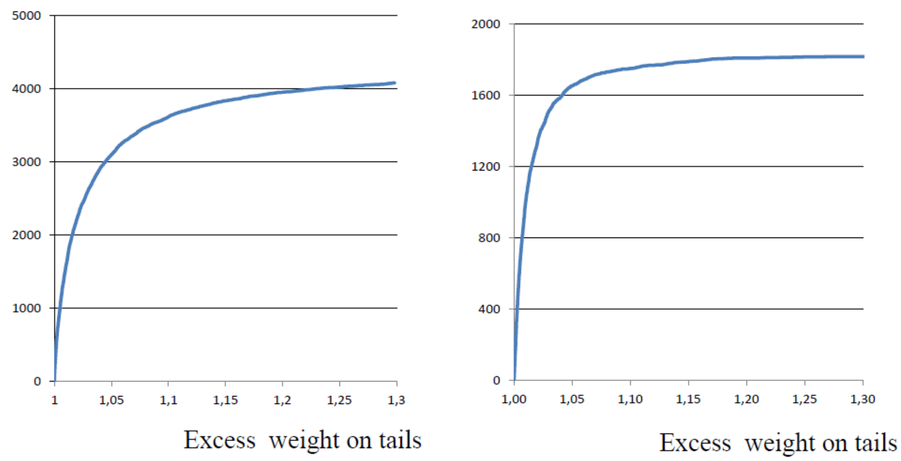
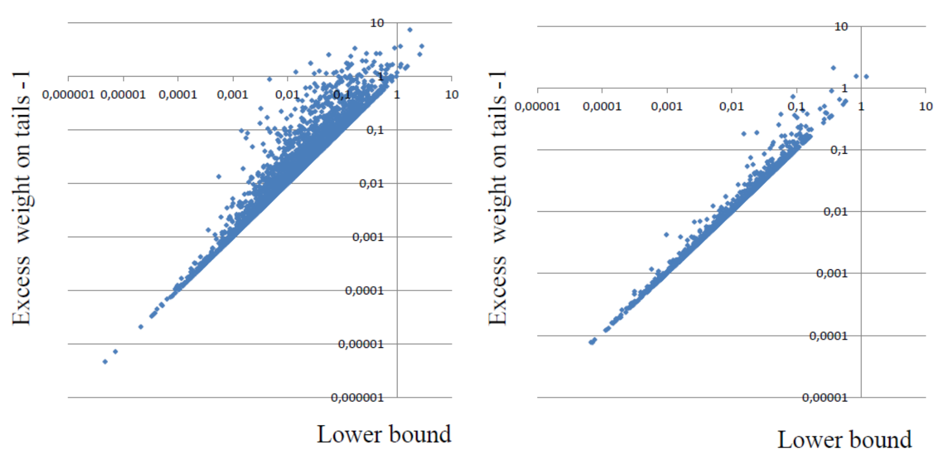


Figure 4: Excess weight on tail risks vs. Lower Bound

a) Guntay-Hackbarth b) Straub-Ulbricht (10y)



Source	restricted SU (N=1839)		restricted GH (N=4335)	
	$\frac{\sigma_P}{\sigma_\theta}$	Lower bound	$\frac{\sigma_P}{\sigma_\theta}$	Lower bound
Mean	**1.027**	**1.021**	**1.084**	**1.044**
St. Dev.	**0.091**	**0.054**	**0.260**	**0.113**
10%	1.001	1.001	1.001	1.001
25%	1.003	1.003	1.005	1.004
Median	**1.009**	**1.008**	**1.019**	**1.014**
75%	1.022	1.021	1.059	1.041
90%	1.05	1.044	1.177	1.104
$Corr\left(\frac{\sigma_P}{\sigma_\theta}, \gamma_P\right)$	-0.200		-0.109	
$Corr\left(\frac{\sigma_P}{\sigma_\theta}, \tilde{D}\right)$	0.550		0.496	
$Corr\left(\frac{\sigma_P}{\sigma_\theta}, Lower\ Bound\right)$	0.860		0.737	

Table 2: Excess Weight on Tail Risks

negatively correlated with forecast accuracy, but significantly positively correlated with forecast dispersion, and this correlation becomes even stronger if we replace forecast dispersion with the lower bound on frictions. Figure 4 displays a scatter plot of firm-level frictions $\sigma_P/\sigma_\theta - 1$ on the vertical axis against its lower bound imputed from forecast dispersion on the horizontal axis, in a log-log scale. Except for very few outliers, our measure of frictions lines up very closely with the lower bound. This confirms that the variation in excess weight on tail risks is mostly driven by forecast dispersion.

To summarize, our data suggest that excess weight on tail risks is likely to be small for most firms, i.e. less than 1.02 in the SU sample and less than 1.05 in the GH sample for 75% of firms, but quite large in the top quartile of the distribution. We show next that the impact on asset returns can nevertheless be fairly significant.

4.2 Returns to skewness

We now argue that our model can generate quantitatively significant skewness premia with realistic levels of information frictions, and that it can account for the observed negative relation between skewness and returns.

We focus on the empirical results of CDG. Define $\pi(\theta) = e^{kx(\theta)}$ such that $x(\cdot)$ is distributed according to a beta distribution. We select the distributional parameters of the beta distribution

	Model					**Data**		
$\tilde{D}$	**0.10**	**0.206**	**0.228**	**0.301**	**0.35**			
σ_P/σ_θ	**1.005**	**1.027**	**1.035**	**1.084**	**1.163**	***Return***	***Skew***	***Vol***
Low	0.11	0.59	0.77	1.87	3.64	***0.57***	*-2.814*	*31.51*
Medium	0.05	0.27	0.36	0.87	1.70	***0.21***	*-0.98*	*32.26*
High	-0.00	-0.01	-0.01	-0.02	-0.05	***-0.22***	*0.026*	*31.14*
H-L	-0.11	-0.59	-0.78	-1.89	-3.69	***-0.79***		
M-L	-0.06	-0.31	-0.41	-1.00	-1.94	***-0.37***		

Table 3: Returns to skewness: model vs data

and k so that $\pi(\cdot)$ matches the skewness and volatility of the different portfolio bins in CDG (Table II). We perform this exercise for skewness inferred from options with three months to maturity, and the respective low, medium and high-skewness portfolios.[22]

We then vary the information frictions parameters β and σ_u so that σ_P/σ_θ falls within the range suggested in section 4.1. We set forecast accuracy γ_P to 0.75, which falls between the sample means and medians of the restricted SU and GH samples, and then vary $\tilde{D}$ between values of 0.1 and 0.35. With these numbers σ_P/σ_θ ranges from 1.005 to 1.163, β from 0.16 to 1.96, and σ_u from 0.24 to 1.37. In all our calibrations, we set $\bar{\theta} = 0$ to focus exclusively on the role of excess weight on tail risks.[23]

Table 3 compares model-implied returns to their empirical counterparts. Its rows represent the portfolios in CDG constructed from options of 3-months to maturity and sorted by skewness. The last three columns report empirical returns, skewness and volatility of these portfolios. The first five columns report the model-implied monthly expected returns (in percentage points) for different levels of forecast dispersion and excess weight on tail risks (reported in the top two lines). The values of $\tilde{D} = 0.206$ and $\tilde{D} = 0.301$ match, respectively, the mean excess weight on tail risks in the SU and GH samples, the values of $\tilde{D} = 0.1$ and $\tilde{D} = 0.35$ have been added to illustrate the quantitative results of the model for the lower half and the top quintiles of the distribution of

[22] We ran the same simulations for skewness measures constructed from 12 months options and for idiosyncratic skewness in CDG, and found the same quantitative results.

[23] We checked the robustness of our quantitative results with respect to variations in $\bar{\theta}$ that bring back average supply effects. Using a second-order Taylor expansion, $R\left(\pi;\bar{\theta},\sigma_P\right) \approx \bar{\theta}\int_{-\infty}^{\infty}\pi'(\theta)\,d\Phi\left(\frac{\theta}{\sigma_P}\right) + \bar{\theta}^2\int_{-\infty}^{\infty}\pi''(\theta)\,d\Phi\left(\frac{\theta}{\sigma_P}\right)$. Since $\bar{\theta} = -\tilde{D}\sigma_\theta\bar{u}$, return asymmetry also interacts with forecast dispersion through average supply effects, but these interactions scale with $\tilde{D}^2$ and are thus an order of magnitude smaller than the effects of excess weight on tail risks. These results are available upon request.

σ_P/σ_θ, using the GH-sample as a benchmark.

Model-implied return premia are negligible for small- to medium dispersion levels, but then grow exponentially for levels of forecast dispersion in the upper quartile of the distribution reported in Table 1, and they far exceed their empirical counterparts for dispersion levels in the top decile.

In addition, return premia appear to grow approximately linearly with the implied excess weight on tail risks, σ_P/σ_θ. We can therefore assess the model's predictive power for average returns by focusing on a calibration that matches the average level of σ_P/σ_θ. Excess weight on tail risks of $\sigma_P/\sigma_\theta = 1.035$ in the middle column match the observed returns to skewness almost exactly. This value is somewhat above the sample mean from SU, but well below the one from GH. Using the sample mean from SU as a conservative estimate, our model accounts for at least 75% of the observed return premia, while using the sample mean from GH, the model generates returns to skewness more than twice as large as the ones in the data.

Our model thus accounts very well for the return differences between high and low skewness portfolios, even if it doesn't match the level of returns.

4.3 The credit spread puzzle

Here we argue that our theory provides a natural explanation for the credit spread puzzle based on excess weight on tail risks. For a given value of information frictions $\sigma_P/\sigma_\theta > 1$, the risk-neutral measure treats a x-standard deviation event as if it was a $x\sigma_\theta/\sigma_P$-standard deviation event. Therefore, if the ex ante probability of default is not too high, the risk-neutral measure will systematically over-estimate default risks, resulting in under-pricing of defaultable bonds. Moreover, this under-pricing becomes more important in relative terms when bonds are deemed safer.

Consider a special case of our model in which the asset payoff is binary: $\pi(\theta) = 1$ for $\theta > \underline{\theta}$ and $\pi(\theta) = 1 - c < 1$ for $\theta \le \underline{\theta}$. With these assumptions, the equilibrium bond price is $P_{\underline{\theta}}(z) = 1 - c\Phi\left(\frac{\theta - \gamma_P \cdot \left(z + \frac{\tau}{\beta+\tau}\bar{u}/\sqrt{\beta}\right)}{\sigma_\theta/\sqrt{1-\gamma_P}}\right)$, while the expected default rate is $Pr(\theta \le \underline{\theta}|z) = \Phi\left(\frac{\underline{\theta} - \gamma_V \cdot \left(z + \bar{u}/\sqrt{\beta}\right)}{\sigma_\theta/\sqrt{1-\gamma_V}}\right)$. The spread is $s_{\underline{\theta}}(z) = -ln\left(P_{\underline{\theta}}(z)\right) \approx c\Phi\left(\frac{\theta - \gamma_P \cdot \left(z + \frac{\tau}{\beta+\tau}\bar{u}/\sqrt{\beta}\right)}{\sigma_\theta/\sqrt{1-\gamma_P}}\right)$, and the expected spread is $\mathbb{E}\left(s_{\underline{\theta}}(z)\right) \approx c\Phi\left(\left(\underline{\theta} - \bar{\theta}\right)/\sigma_P\right)$, while the unconditional default rate is $Pr(\theta \le \underline{\theta}) = \Phi(\underline{\theta}/\sigma_\theta)$. The expected return is

$$\mathbb{E}\left(s_{\underline{\theta}}(z)\right) - cPr(\theta \le \underline{\theta}) \approx c\left(\Phi\left(\left(\underline{\theta} - \bar{\theta}\right)/\sigma_P\right) - \Phi\left(\underline{\theta}/\sigma_\theta\right)\right).$$

This characterization delivers the following two comparative statics results, for $\bar{\theta} \le 0$:[24]

[24] $\bar{\theta} \le 0$ allows for positive average supply, akin to a risk premium for default risk.

- The expected spread $c\Phi\left(\left(\underline{\theta}-\bar{\theta}\right)/\sigma_P\right)$ and return $c\left(\Phi\left(\left(\underline{\theta}-\bar{\theta}\right)/\sigma_P\right)-\Phi\left(\underline{\theta}/\sigma_\theta\right)\right)$ are increasing in σ_P, for $\underline{\theta}<\bar{\theta}$.

- The ratio $\Phi\left(\underline{\theta}/\sigma_\theta\right)/\Phi\left(\left(\underline{\theta}-\bar{\theta}\right)/\sigma_P\right)$ is increasing in $\underline{\theta}$ and converges to 0 as $\underline{\theta}\to-\infty$.

Therefore, we confirm analytically that excess weight on tail risks increases bond spreads and expected returns, and that these effects are more important for safer bonds, i.e. as $\underline{\theta}\to-\infty$.

In Table 4, we illustrate these points numerically. The first four columns of Table 4 report empirical moments for bond maturities of 4 years, taken from Huang and Huang (2012):[25] the cumulative probability of default over the holding period, transformed into an annualized "loss rate" (in basis points) which corresponds to the yield spread at which the bondholder just breaks even, the average spread, which corresponds to $\mathbb{E}\left(s_{\underline{\theta}}\left(z\right)\right)$ in our model, and the L/S ratio between the loss rate and the spread, which corresponds to $\Phi\left(\underline{\theta}/\sigma_\theta\right)/\Phi\left(\left(\underline{\theta}-\bar{\theta}\right)/\sigma_P\right)$ in the model.

The next two columns report model-implied spreads in levels (and as percentage of the data) and the L/S ratio of the structural model of bond pricing in Huang and Huang (2012), which we use as a reference point for our calibration. These columns summarize well the credit spread puzzle: Model-implied spreads are far lower than their data counterparts, especially for investment grade (Baa and above) bonds. This failure can be traced to lack of sufficient variation in the model-implied L/S ratio: In the data, default risk premia vanish at a much slower rate than the corresponding default risks in higher ratings categories. This is difficult to reconcile with standard risk preferences which typically display bounded aversion to small pay-off risks.

We set $c=0.5$ to match empirical recovery rate estimates in case of default, set $\underline{\theta}$ to match observed default frequencies for bonds with different ratings. We set $\bar{\theta}=0$ and vary σ_P/σ_θ: we set forecast accuracy to γ_P to 0.75, corresponding to the median of the restricted G-H sample, and then vary $\tilde{D}$ to match friction levels of $\frac{\sigma_P}{\sigma_\theta}=1.084$ ($\tilde{D}=0.301$), $\frac{\sigma_P}{\sigma_\theta}=1.163$ ($\tilde{D}=0.35$), and $\frac{\sigma_P}{\sigma_\theta}=1.74$ ($\tilde{D}=0.414$). These correspond, respectively, to the average level of frictions in the GH sample, the top quartile of forecast dispersion in GH, and the level of frictions under which the model matches the spread for Baa bonds exactly.

The last six columns report the model-implied average spread and L/S ratio for each ratings category at the three levels of information frictions.

Like the structural bond pricing model, our dispersed information model falls well short of accounting for the observed level of credit spreads, especially for investment grade bonds. For realistic levels of excess weight on tail risks, the model accounts for anywhere from roughly 10%

[25] We have performed the same exercise with bonds of 10 years to maturity. The results are very similar.

	Data				Structural Model		$\frac{\sigma_P}{\sigma_\theta} = 1.084$		$\frac{\sigma_P}{\sigma_\theta} = 1.163$		$\frac{\sigma_P}{\sigma_\theta} = 1.74$	
	Pr(default)	***Loss Rate***	***Spread***	***L/S***	**Spread**	**L/S**	**Spread**	**L/S**	**Spread**	**L/S**	**Spread**	**L/S**
Aaa	*0.04%*	*0.5*	***55***	***0.9%***	1.1 (2%)	46%	1.4 (2.6%)	35%	2.9 (5.2%)	18%	42 (77%)	1.2%
Aa	*0.23%*	*2.9*	***65***	***4.4%***	6 (9.2%)	48%	6.2 (9.6%)	46%	11 (16%)	27%	80 (124%)	3.6%
A	*0.35%*	*4.4*	***96***	***4.6%***	10 (10%)	44%	8.2 (8.5%)	53%	13 (14%)	33%	91 (94%)	5.1%
Baa	*1.24%*	*15.5*	***158***	***9.8%***	32 (20%)	49%	29 (18%)	54%	41 (26%)	38%	158 (100%)	9.8%
Ba	*8.51%*	*109*	***320***	***34%***	172 (54%)	63%	142 (44%)	77%	168 (53%)	65%	333 (104%)	33%
B	*23.3%*	*301*	***470***	***66%***	446 (95%)	70%	397 (84%)	78%	424 (90%)	73%	558 (119%)	56%

Table 4: Credit spreads: model vs data

to 35% of spreads on high quality (Baa and above) bonds. This ratio is consistent with empirical results in GH and in Buraschi, Trojani and Vedolin (2014) who suggest that forecast dispersion accounts for about 20%-30% of the observed levels and variation in credit spreads in the data.

Furthermore, our model generates substantially more variation in the L/S ratio than the structural bond pricing model, and is able to account for a much lower ratio of expected losses relative to spreads for investment grade bonds. In the data, this ratio varies from 66% for the lowest ratings category to 5% for investment grade bonds and less than 1% for Aaa bonds - a variation by a factor of 70 between the highest and the lowest ratings category. The structural model matches the L/S ratio for the lowest ratings categories, but is unable to account for its steep decline in the highest categories, since the L/S ratio stays above 45%, it accounts for a variation of at most a factor 1.5. With $\frac{\sigma_P}{\sigma_\theta} = 1.163$, the dispersed information model can account for L/S ratios as low as 18% for 4-year Aaa bonds, equivalent to a variation by a factor 4.[26]

[26] Feldhütter and Schaefer (2018) question the existence of a credit spread puzzle by arguing that historical default data by ratings category offer only a very noisy estimate of underlying default risks. They offer a model-based alternative that exploits information on leverage and defaults across maturities and ratings categories, to argue that default probabilities do not vary nearly as much across ratings as suggested by Huang and Huang (2012). Bai, Goldstein and Yang (2020) on the other hand question the estimates by Feldhütter and Schaefer (2018) by arguing that model-consistent estimates should use market rather than book values of debt to measure leverage. The model-based estimates extrapolate default boundaries estimated from low-rated firms for which defaults are frequent to higher ratings categories for which defaults are far more infrequent. The gap between market- and book-value of debt for low-rated firms then generates a substantial bias in the estimated default boundary and the variation in the estimated default probability across ratings categories.

Table 3 of Bai, Goldstein and Yang (2020) summarizes the difference between the historical default frequencies and the respective estimates using market value and book value of debt. Using their numbers, we can compute L/S-ratios for the different estimates. For bonds with 5 years to maturity, the L/S-ratios vary by a factor of 20 across ratings

In the last two columns we have set forecast dispersion to match the return on Baa bonds, and we have then calculated the model-implied spreads for all other ratings categories and maturities. While the required excess weight on tail risks may seem implausibly large, this calibration matches the overall level and variation of spreads across categories remarkably well. This result suggests that severe information frictions have the potential to fully account for the credit spread puzzle by generating sufficient variation in the L/S-ratio to match the observed variation in the data.

Finally, we explore the robustness of our results w.r.t. $\bar{\theta}$. By setting $\bar{\theta} = -0.2$ and $\frac{\sigma_P}{\sigma_\theta} = 1$, we replicate the structural model of Huang and Huang (2012) almost exactly with Spreads varying from 1.1 to 446 basis points and the L/S ratio varying from 44% to 66% as we vary credit ratings from Aaa to B. Allowing a similar shift in the mean of the risk-neutral distribution uniformly increases spreads for bonds of all ratings categories, which helps in accounting for the relative magnitude of spreads on investment grade bonds but also results in over-shooting for spreads on speculative grades.

In summary, our model offers a simple explanation for the credit spread puzzle. To account for the puzzle, one must explain why return premia that vanish at a lower rate than the underlying default risks for highly rated bonds, as illustrated by the declining L/S ratio. Our explanation is based on the observation that with excess weight on tail risks the market-implied default probability $\Phi\left(\left(\underline{\theta} - \bar{\theta}\right)/\sigma_P\right)$ converges to 0 at a much slower rate than the actual default rate $\Phi\left(\underline{\theta}/\sigma_\theta\right)$.

Our model abstracts from many additional factors to generate quantitatively realistic credit spreads. Nevertheless, the fact that this simple model does as well or better than substantially more sophisticated structural models of credit spreads along one key dimension, the L/S-ratio, suggests that dispersed information may play a significant role in isolation or in combination with other factors determining credit spreads. In Albagli, Hellwig, and Tsyvinski (2014) we develop a fully dynamic dispersed information model of bond pricing and discuss its quantitative fit for explaining the credit spread puzzle.

---

categories in the estimates derived from historical default frequencies, by a factor of 2.5 in the estimates derived from book value of debt, and by a factor of 8 in the estimates derived from market value of debt. The structural model estimates reported by Bai, Goldstein and Yang (2020) in turn suggest a variation in the L/S ratio of at most a factor of 1.3 for models using either book or market value of debt, which is even lower than the factor of 1.5 derived from the estimates of Huang and Huang (see Table 4).

As a robustness check, we also calibrated our model to match the three series of default probabilities reported in Bai, Goldstein and Yang (2020). Our model generates a variation in the L/S-ratio by a factor of 2 to 5 for the different estimates of default risk with plausible degrees of information frictions. Our model with dispersed information thus accounts for a much larger fraction of the variation in the L/S-ratio than the structural model estimates reported in Bai, Goldstein and Yang (2020), regardless of the estimated default probabilities one chooses to match.

4.4 Returns to disagreement

DMS (table VI) report that stocks in the highest quintile of analyst forecast dispersion have monthly returns that are about 0.62% lower than those in the lowest dispersion quintile, equivalent to an annual return premium of -7.7%. GH replicate these results with a subset of firms that are also active in bonds markets and confirm their finding, though with a lower level of excess returns, of 0.26% per month, or -3.2% per year. The firms in the GH sample correspond to the highest size and lowest dispersion quintiles of the DMS sample. This accounts for most of the difference in returns to disagreement between DMS and GH.[27]

We now use the risk-neutral normal model with the same calibration strategy as above to explore returns to disagreement generated by our model. We follow the same approach as above and calibrate informational parameters to match the variation in forecast dispersion and informational frictions reported for the GH sample, and asset returns to match skewness and volatility of firm-level equity returns. As before we set γ_P to 0.75. We then vary $\tilde{D}$ to match the mean friction, as well as the 10th, 25th, 50th, 75th and 90th percentile of the distribution of forecast dispersion quintiles in GH.

The challenge is then to identify realistic targets for expected skewness and volatility of firm-level equity returns: Unfortunately, DMS or GH do not report these moments when identifying returns to disagreement. CDG report predicted skewness and volatility but their predicted skewness measures are negative at the firm-level, which contradicts other available evidence and cannot be consistent with a model generating negative returns to disagreement in equity markets.

We therefore calibrate our returns to match the realized firm-level skewness and volatility of equity returns that are reported in Boyer et al. (2010), rather than their expected counterparts. In the first row of Table 5, we report the monthly returns to a security that displays the mean level of skewness and volatility reported in Table 1 (Panel A) of Boyer, Mitton and Vorkink (2010), for different levels of forecast dispersion, as well as the excess return of the 90th to 10th dispersion percentile. Our model generates a negative excess return of 0.16% per month, corresponding to roughly 60% of the returns reported in GH,[28] with most of the return differential concentrated in the top quintile, which is qualitatively consistent with the empirical findings.

[27] Similarly, Hou, Xue, and Zhang (2020) report that returns to disagreement are much lower and statistically insignificant at the 5% level, when using value-weighted portfolios. This is consistent with the view that returns to disagreement are largest in markets with low capitalization.

[28] We focus on the returns to disagreement reported by GH, since these are the firms that correspond to our sample of forecast dispersion.

	Data			Model	$\tilde{D}$					
	Ret	*Skew*	*Vol*	**Mean GH (0.301)**	**10% (0.048)**	**25% (0.092)**	**50% (0.16)**	**75% (0.251)**	**90% (0.342)**	*90%-10%*
Mean		*0.851*	*3.6*	-0.09	-0.00	-0.00	-0.01	-0.05	-0.16	-0.16
Q1	*1.19*	*0.167*	*1.888*	-0.01	-0.00	-0.00	-0.00	-0.01	-0.02	-0.02
Q2	*1.12*	*0.375*	*2.087*	-0.02	-0.00	-0.00	-0.00	-0.01	-0.04	-0.04
Q3	*1.10*	*0.565*	*2.672*	-0.04	-0.00	-0.00	-0.01	-0.02	-0.08	-0.08
Q4	*1.06*	*0.809*	*3.406*	-0.08	-0.00	-0.00	-0.01	-0.04	-0.14	-0.14
Q5	*0.52*	*1.629*	*5.342*	-0.24	-0.00	-0.01	-0.04	-0.13	-0.42	-0.42
Q5-Q1	*-0.67*			-0.23	-0.00	-0.01	-0.04	-0.12	-0.41	

Table 5: Returns to disagreement: model vs data

The subsequent rows report model-implied monthly returns for securities that match the different portfolios sorted on idiosyncratic skewness as in Boyer, Mitton and Vorkink (2010), Table 3. These model implied returns to disagreement are decreasing with skewness, as predicted by the theory, and concentrated in the top skewness and top dispersion quintiles. The variation in returns to disagreement from the bottom to top skewness quintile (0.40%), as well as the variation in returns to skewness from the bottom to the top disagreement quintiles (0.41%) are significant and correspond to about 5% annual returns or about two thirds of the variation in returns to disagreement and two thirds of the value premium reported in Yu (2011). In other words, our theory accounts for a sizeable fraction of the level and variation in the observed returns to disagreement.[29]

To summarize, these three quantitative applications suggest that our theory can generate quantitatively significant levels and variation in returns to skewness, disagreement, and the interaction between the two, for plausible levels of information frictions or forecast dispersion in equity and

[29] Compared to the CDG calibration, our calibration accounts for a much smaller fraction of the returns to skewness in Boyer, Mitton and Vorkink (2010), 0.23% in the model vs. 0.67% in the data, or 34% of the observed return premium. This may be due to the fact that we cannot perfectly calibrate the model to match expected skewness as we did in the CDG calibration. A second reason is that CDG report much higher levels of firm-level volatility - raising the volatility targets in the present model to the levels reported in CDG would substantially increase, once again, the ability of our model to account for observed returns to skewness and disagreement. In addition, firm-level volatility turns out to be a strong predictor of idiosyncratic skewness, which may affect returns through a risk channel that our model does not account for.We also calibrated our model using the moments for conditionally sorted portfolios in Table 10 of Boyer, Mitton and Vorkink (2010). In this alternative calibration, controlling for idiosyncratic volatility, the share of the returns to skewness explained by the model increases to 67%, primarily because the empirical return premium is significantly lower.

bond markets. Future work will have to refine these predictions, corroborate the empirical measures used to calibrate information frictions, and evaluate the respective roles of information friction vs. risk premia, but as a first pass these results illustrate that even relatively small levels of information frictions may potentially account for significant cross-sectional return premia. Our quantitative results also replicate the empirical finding that most of these return premia are concentrated in the highest dispersion quintile, in line with our observation that information frictions are small for most firms, but can be very significant in the upper tail.

5 General Results

In this section, we generalize the equilibrium characterization and comparative statics results from section 3. We show that the sufficient statistics representation of the equilibrium price with an updating wedge, the construction of the risk-neutral measure, and the asset pricing implications of excess weight on tail risks generalize immediately without further assumptions on preferences, signal distributions and returns. The only step that is more involved is showing that the updating wedge generates excess weight on tail risks in the risk-neutral measure. We show that excess weight on tail risks arises whenever (i) the gap between the asset price and average investor expectations is orthogonal to fundamentals, i.e. exclusively driven by supply shocks and (ii) investor expectations are approximately linear, i.e. linear projections provide a sufficiently close approximation of Bayesian posteriors. These two sufficient conditions are satisfied by all canonical examples of noisy REE studied in the literature. The core insights of our model thus extend far beyond the model with risk-neutral investors and position limits.

We generalize the model set-up of section 2 by making the following assumptions:

(i) *Generic prior distributions:* the asset fundamental θ is distributed according to a generic cdf. $H(\cdot)$, with variance $Var(\theta) \equiv \sigma_\theta^2$. The specification of asset dividends $\pi(\cdot)$ is unchanged. The stochastic asset supply s is distributed according to a generic cdf. $G(\cdot)$ with support $[d_L, d_H]$, with $d_L \leq 0 \leq d_H$, and $d_L < d_H$.[30]

(ii) *Generic signal distributions:* Informed investors' private signals are $x_i = \theta + \varepsilon_i$, where ε_i is i.i.d across agents, and distributed according to generic cdf. $F(\cdot)$ and smooth, symmetric density function $f(\cdot)$ with unbounded support. We assume $f'(\cdot)/f(\cdot)$ is strictly decreasing and unbounded above and below, and let $\beta \equiv 1/Var(\varepsilon)$ denote the precision of private signals.[31]

[30] The model introduced in section 2 specified $H(\cdot)$ to be normal with mean 0 and variance σ_θ^2, and asset supply had support $[d_L, d_H] = [0, 1]$ and $G(\cdot)$ s.t. $s = \Phi(u)$, where u was normally distributed.

[31] Monotonicity of $f'(\cdot)/f(\cdot)$ implies signals have log-concave density and satisfy the monotone likelihood ratio

(*iii*) *Generic investor preferences:* Investors' preferences are characterized by a strictly increasing, concave utility function $U(\cdot)$ defined on the investors' realized gains or losses $d_i \cdot (\pi(\theta) - P)$. Their asset demand is restricted by position limits, i.e. $d_i \in [d_L, d_H]$.[32] Given the investors' posterior $H(\cdot|x, P)$, a demand function $d(x, P)$ is optimal, if it solves the investors' decision problem $\max_{d \in [d_L, d_H]} \int U(d(\pi(\theta) - P))\, dH(\theta|x, P)$. This leads to the corresponding first-order condition

$$\int (\pi(\theta) - P) \cdot U'(d(\pi(\theta) - P)) dH(\theta|x, P) = 0. \tag{6}$$

As before, a *Perfect Bayesian Equilibrium* $\{d(x, P); P(\theta, s); H(\cdot|P)\}$ consists of a demand function $d(x, P)$, a price function $P(\theta, s)$, and posterior beliefs $H(\cdot|P)$ that jointly satisfy the investors' optimality condition, the market-clearing condition and Bayes' Rule whenever the latter is applicable. We assume that there exists a *price-monotone equilibrium* $\{P(\theta, s); d(x, P); H(\cdot|P)\}$ in which $d(x, P)$ is strictly decreasing in P for $d(x, P) \in (d_L, d_H)$. Price monotonicity of demand arises automatically if trade takes place through a limit-order book.[33]

5.1 Sufficient Statistic Representation

Fix any $\bar{D} \in (d_L, d_H)$ and define $z \equiv z(P)$ as the private signal of an investor who finds it optimal to hold exactly $\bar{D}$ units at price P. $z(P)$ is implicitly defined by $d(z, P) = \bar{D}$. Since $d(x, P)$ is strictly increasing in x, $z(P)$ is strictly increasing in P, and therefore serves as a sufficient statistic for the information conveyed through the equilibrium price. By inverting $z(P)$, we can represent the equilibrium price as a function of z only. Evaluating the investor's first-order condition (6) at $x = z(P)$, we obtain

$$P_\pi(z) = \frac{\mathbb{E}\left(U'\left(\bar{D}\left(\pi(\theta) - P_\pi(z)\right)\right) \cdot \pi(\theta)\,|x = z, z\right)}{\mathbb{E}\left(U'\left(\bar{D}\left(\pi(\theta) - P_\pi(z)\right)\right)|x = z, z\right)}. \tag{7}$$

In addition, we can construct posterior beliefs directly from the market-clearing condition. Since aggregate demand $D(\theta, P) = \int d(x, P)\, dF(x - \theta)$ is strictly decreasing in P, we have $Pr(P \leq P'|\theta) = Pr(D(\theta, P) \geq D(\theta, P')) = Pr(s \geq D(\theta, P')) = 1 - G(D(\theta, P'))$. Therefore conditional on θ, z is distributed according to

$$\Psi(z|\theta) = 1 - G(D(\theta, P_\pi(z))). \tag{8}$$

property. Unboundedness implies extreme signal realizations induce large updates in posterior beliefs, (almost) regardless of the information contained in other signals. In section 2 we assumed $F(\cdot)$ to be iid. normal, which satisfies all these properties.

[32] In section 2 we assumed $U(\cdot)$ to be linear and positions to be bounded by $[0, 1]$.

[33] To our knowledge, no general existence results are available for this class of models.

Together with the prior $H(\cdot)$, this defines the joint distribution of P and θ, from which we derive the posterior $H(\cdot|P)$ using Bayes' Rule whenever applicable. Hence, for given $\bar{D}$, the equilibrium is fully characterized by the fixed point between equations (7) and (8). These observations are summarized in the following theorem:

Theorem 1 *: For any Perfect Bayesian Equilibrium $\{P(\theta, s); d(x, P); H(\cdot|P)\}$ with $d(x, P)$ strictly decreasing in P at $\bar{D} \in (d_L, d_H)$, there exists a sufficient statistic $z = z(\theta, s)$ such that $P(\theta, s) = P_\pi(z(\theta, s))$, where $P_\pi(\cdot)$ satisfies (7) and the cdf of z satisfies (8).*

Theorem 1 generalizes the sufficient statistic representation of Proposition 1 for any price-monotone equilibrium. For each $\bar{D} \in (d_L, d_H)$, there exists a state variable z, function of θ and s only, such that we can represent the price as the risk-adjusted expectation of dividends of an investor who finds it optimal to hold exactly $\bar{D}$ units of the asset when the state is z.[34]

Equation (7) generalizes the result that the risk-neutral expectation of dividends differs from the Bayesian posterior conditional on the same public information z. The risk-neutral expectation of dividends processes the price signal twice, once as a public price signal, and once as the private signal of the threshold investor who finds it optimal to purchase exactly $\bar{D}$ units of the asset. The intuition for this characterization is the same as the one given in section 3, i.e. shifts in fundamentals or noise trader demand result in price adjustment, due to market-clearing, over and above the mere information content of the price. In the expression for the equilibrium price, these two effects are represented by the sufficient statistic z appearing twice in the conditioning set, once through the price signal, and once through the marginal investor's private information. This wedge between the market expectation of dividends and the Bayesion posterior is thus a necessary characteristic of *any* model with noisy information aggregation through asset prices.

Two ingredients are crucial for the appearance of this updating wedge. First the equilibrium demand is not perfectly price-elastic, due to limits to the trader's willingness (i.e. risk aversion) or ability (i.e. position limits) to arbitrage any degree of perceived mispricing. Without such limits to arbitrage, risk-neutral investors would be willing to take unlimited positions, prices would become perfectly revealing and converge to the true dividend values (i.e. z would converge to θ, almost surely), while supply shocks get seamlessly absorbed by the market.

[34] The representation in theorem 1 depends on the initial choice of $\bar{D}$, but the representations for different values of $\bar{D}$ are all monotonic transformations of each other. They correspond to different decompositions of the price into expected dividend and risk premium: the higher is $\bar{D}$, the higher is the required risk premium, and hence also the dividend expectation of the investor who holds $\bar{D}$ in equilibrium. As $\bar{D} \to 0$ the risk premium disappears.

Second, information must be dispersed. To see this, we compare our characterization to an otherwise identical economy in which all investors share the same information. Taking as given supply s and an exogenous public signal z, the asset price with *common information* is

$$V_{\pi}(z,s)=\frac{\mathbb{E}\left(U'\left(s\left(\pi\left(\theta\right)-V_{\pi}\left(z,s\right)\right)\right)\cdot\pi\left(\theta\right)|z\right)}{\mathbb{E}\left(U'\left(s\left(\pi\left(\theta\right)-V_{\pi}\left(z,s\right)\right)\right)|z\right)}. \tag{9}$$

By comparing the common information price (9) to its dispersed information counterpart (7) with $\bar{D}$ set equal to s, we see exactly how dispersed information affects the response of prices to the information aggregated through the market.

Theorem 1 only offers a partial equilibrium characterization: to fully characterize asset valuations, we still need to compute, for some $\bar{D}$, the distribution of the associated sufficient statistic z. This distribution, however, derives from the market clearing condition $D(\theta,P)=s$, which still requires information about the entire demand schedule. Nevertheless, Theorem 1 allows us to develop implications for asset prices through a risk-neutral representation of the equilibrium price.

5.2 A risk-neutral probability measure for markets with dispersed information

Generalizing the construction of the risk-neutral measure is straight-forward: Let $\hat{H}(\theta|z)=H(\theta|x=z,z)$ denote the risk-neutral posterior distribution over θ. Theorem 1 with $\bar{D}=0$ gives us a risk-neutral representation of the price as the conditional dividend expectation under $\hat{H}(\theta|z)$, $P_{\pi}(z)=\hat{\mathbb{E}}\left(\pi\left(\theta\right)|z\right)=\int\pi\left(\theta\right)d\hat{H}(\theta|z)$.[35] The risk-neutral prior distribution $\hat{H}(\theta)$ over θ is then derived by compounding $\hat{H}(\theta|z)$ with the prior over z: $\hat{H}(\theta)=\int H(\theta|x=z,z)d\Psi\left(z\right)$, where $\Psi\left(z\right)=\int\left(1-G\left(D\left(\theta,P_{\pi}\left(z\right)\right)\right)\right)dH\left(\theta\right)$ denotes the prior cdf of z. This leads to the canonical representation of asset prices as expected dividends under the risk-neutral measure $\hat{H}$: $\mathbb{E}\left(P_{\pi}\left(z\right)\right)=\int\hat{\mathbb{E}}\left(\pi\left(\theta\right)|z\right)d\Psi\left(z\right)=\int\pi\left(\theta\right)d\hat{H}(\theta)\equiv\hat{\mathbb{E}}\left(\pi\left(\theta\right)\right).$

We say that $\hat{H}\left(\cdot\right)$ displays *excess weight on tail risks*, if we can decompose the difference between $H\left(\cdot\right)$ and $\hat{H}\left(\cdot\right)$ into a shift in means and a mean-preserving spread, i.e. if the distribution with cdf $\hat{H}\left(\theta+\delta\right)$ with $\delta=\int\theta d\hat{H}\left(\theta\right)-\int\theta dH\left(\theta\right)$ represents a mean-preserving spread over H. We then decompose $\mathbb{E}\left(P_{\pi}\left(z\right)\right)-\mathbb{E}\left(\pi\left(\theta\right)\right)$ as follows:

$$\mathbb{E}\left(P_{\pi}\left(z\right)\right)-\mathbb{E}\left(\pi\left(\theta\right)\right)=\int_{-\infty}^{\infty}\left(\pi\left(\theta\right)-\pi\left(\theta-\delta\right)\right)d\hat{H}\left(\theta\right)+\int_{-\infty}^{\infty}\left(H\left(\theta\right)-\hat{H}\left(\theta+\delta\right)\right)d\pi\left(\theta\right)$$

As before, the term $R\left(\pi;\delta,\hat{H}\left(\cdot\right)\right)\equiv\int_{-\infty}^{\infty}\left(\pi\left(\theta\right)-\pi\left(\theta-\delta\right)\right)d\hat{H}\left(\theta\right)$ accounts for the shift in means. This term varies with the expected asset supply: for a given distribution of dividends, a

[35] With $\bar{D}\neq 0$, a similar representation applies, but in that case, we need to account for a risk adjustment in the prior, which changes the interpretation of the comparative statics results.

first-order stochastic increase in the supply distribution $G(\cdot)$ requires that informed investors buy more shares in equilibrium, which lowers the marginal investor's z. This downwards shift in the price distribution is captured by a decrease in δ.

The second term, denoted $W\left(\pi, \hat{H}\right) \equiv \int_{-\infty}^{\infty} \left(H(\theta) - \hat{H}(\theta + \delta)\right) d\pi(\theta)$, instead corresponds to a second-order shift in the distribution. This term takes expectations of $\pi(\theta)$ w.r.t. two distributions with identical means but where $\hat{H}(\theta + \delta)$ is second-order stochastically dominated by $H(\theta)$, capturing the excess weight on tail risks implied by the market expectations. Using second-order stochastic dominance as a partial order on the severity of information aggregation frictions, we say that $\hat{H}_1(\cdot)$ has more excess weight on tail risk than $\hat{H}_2(\cdot)$ iff $\hat{H}_1(\cdot)$ is second-order stochastically dominated by $\hat{H}_2(\cdot)$. Theorem 2 generalizes the testable implications of excess weight on tail risks for asset prices from Proposition 2.

Theorem 2 *Suppose that the risk-neutral measure $\hat{H}_1$ has more excess weight on tail risk than $\hat{H}_2$.*

(i) ***Comparative Statics w.r.t.*** $\hat{H}$***:*** *If π is convex (concave), then $W\left(\pi_1, \hat{H}_1\right) \geq W\left(\pi_1, \hat{H}_2\right) \geq 0$ ($W\left(\pi_1, \hat{H}_1\right) \leq W\left(\pi_1, \hat{H}_2\right) \leq 0$). Moreover, $\lim_{\hat{H} \to H} W\left(\pi, \hat{H}\right) = 0$, and $|W\left(\pi, \hat{H}\right)|$ grows arbitrarily large if $\hat{H}$ has arbitrarily large excess weight on tail risk.*

(ii) ***Comparative Statics w.r.t.*** π ***and Increasing differences:*** *If $\pi_1 - \pi_2$ is convex, then $W\left(\pi_1, \hat{H}_1\right) - W\left(\pi_2, \hat{H}_1\right) \geq W\left(\pi_1, \hat{H}_2\right) - W\left(\pi_2, \hat{H}_2\right) \geq 0$.*

Generalizing predictions from Proposition 2 to Theorem 2 leads to two additional complications.

First, while the mapping from model parameters to observable forecast statistics was particularly transparent when expectations were linear, a direct empirical proxy for excess weight on tail risks is more difficult to define for the general model, where this property is summarized not by a single parameter but by a second-order stochastic dominance ranking. Nevertheless, if the underlying information frictions are summarized by a small set of parameters, such as noise trader variance and private signal precision, we can still calibrate these parameters to observed moments of forecast accuracy and forecast dispersion.

Second, without additional assumptions of symmetry on the prior and risk-neutral measures, the comparative statics w.r.t. return asymmetry require a more restrictive partial order in terms of convexity and therefore apply only to returns that are continuous and unbounded on at least one side. While this may not restrict the analysis of upside risks (such as options) too much, it seems impractical to assume unboundedly negative dividend payoffs for downside risks.

However, if $H(\theta)$ and $\hat{H}(\theta+\delta)$ are symmetric around 0, $W\left(\pi,\hat{H}\right)$ can be written as

$$W\left(\pi,\hat{H}\right)=\int_{-\infty}^{\infty}\left(H(\theta)-\hat{H}(\theta+\delta)\right)d\hat{\pi}(\theta)=2\int_{0}^{\infty}\left(H(\theta)-\hat{H}(\theta+\delta)\right)d\hat{\pi}(\theta), \tag{10}$$

where $\hat{\pi}(\theta)=1/2\cdot(\pi(\theta)+\pi(-\theta))$. Hence, symmetry allows us to partially order upside and downside risk in terms of the concavity and convexity of $\hat{\pi}$ rather than π. If in addition to being symmetric $H(\theta)$ and $\hat{H}(\theta+\delta)$ only cross at $\theta=0$ (i.e. $H(\theta)\gtreqless\hat{H}(\theta+\delta)$ for $\theta\gtreqless 0$), then upside and downside risks can be ordered according to whether $\hat{\pi}$ is increasing, constant or decreasing. This allows us to generalize the partial order on returns to Definition 1 and recover the comparative statics predictions of Proposition 2.

5.3 Sufficient conditions for excess weight on tail risk

The last remaining step in the argument is to show that the risk-neutral measure with dispersed information indeed displays excess weight on tail risks. Our next proposition provides sufficient conditions for excess weight on tail risks in the general model. We then show that these sufficient conditions have a natural interpretation in terms of model properties and are satisfied in all leading applications of noisy information aggregation models.

Proposition 3 *: Suppose that $\hat{\mathbb{E}}(\theta|z)-\mathbb{E}(\theta|x,z)$ is independent of θ and $\theta-\mathbb{E}(\theta|x,z)$ is independent of x. Then the risk-neutral probability measure $\hat{H}(\cdot)$ displays excess weight on tail risks.*

The first independence condition in proposition 3 requires that the gap between the average informed trader's expectation $\int\mathbb{E}(\theta|x,z)\,dF(x-\theta)$ and the marginal informed trader's expectation $\hat{\mathbb{E}}(\theta|z)$ is *exclusively* driven by random supply shocks. The second independence condition imposes that residual uncertainty is independent of x. These two conditions correspond to the two properties of the normal distribution that we used to link excess weight on tail risks to excess volatility of risk-neutral market expectations, and to show that $\hat{Var}(\theta)=Var\left(\theta+\hat{\mathbb{E}}(\theta|z)-\mathbb{E}(\theta|x,z)\right)>Var(\theta)$. To complete the proof of proposition 3, we show that these conditions not only imply that $\hat{Var}(\theta)>Var(\theta)$, but that the difference between the risk-neutral and objective priors decomposes into a shift in means and a mean-preserving spread.

Next, we argue that these sufficient conditions are satisfied in important benchmark cases. Differentiating the market-clearing condition $D(\theta,P)=s$ with respect to θ, we obtain

$$D_\theta(\theta,P_\pi(z))+D_P(\theta,P_\pi(z))\frac{\partial P_\pi(z)}{\partial z}\frac{\partial z(\theta,s)}{\partial\theta}=0.$$

We then say that equilibrium demand has a *simple structure*, if for all θ and z,

$$D_{\theta}\left(\theta, P_{\pi}\left(z\right)\right) + D_{P}\left(\theta, P_{\pi}\left(z\right)\right)\frac{\partial P_{\pi}\left(z\right)}{\partial z} = 0, \tag{11}$$

or equivalently, $\frac{\partial z(\theta,s)}{\partial \theta} = 1$. Therefore demand has a simple structure, if and only if the sufficient statistic $z\left(\theta, s\right)$ takes a simple additive form "fundamentals plus noise", $z\left(\theta, s\right) = \theta + u\left(s\right)$ for some monotone decreasing function $u\left(\cdot\right)$. As before, we denote the precision of z by $\tau \equiv 1/Var\left(u\left(s\right)\right)$.

Simple demand structures encompass many existing models, such as the CARA normal model and its non-normal extensions (e.g. Breon-Drish, 2015) in which demand is affine in the private signal, or the risk-neutral model studied above. In general, $\frac{\partial P_{\pi}(z)}{\partial z} = -d_{x}\left(z, P\right)/d_{P}\left(z, P\right)$ measures the rate at which the marginal investor trades off between higher price and higher dividend expectation, while $\left(-D_{\theta}/D_{P}\right)^{-1}$ represents the same marginal rate of substitution for aggregate demand, or investors on average. A simple demand structure imposes that the marginal and average investors' marginal rates of substitution between higher price and higher dividend expectation coincide. Departures from simple demand structures require that $-d_{x}\left(x, P\right)/d_{P}\left(x, P\right)$ varies with x, and that this variation does not wash out through aggregation.

Simple demand structures offer a useful benchmark for excess weight on tail risks. If demand is simple and $\mathbb{E}\left(\theta|x, z\right)$ is linear in x, then $\hat{\mathbb{E}}\left(\theta|z\right) - \mathbb{E}\left(\theta|x, z\right)$ is independent of θ, and $Var\left(\theta|x, z\right)$ is independent of x, which implies that $\hat{Var}\left(\theta\right) = \sigma_{P}^{2} > Var\left(\theta\right)$, where σ_{P}^{2} is defined as in section 3.2.[36] The characterization of excess weight on tail risks thus fully generalizes to any linear projection model. These models also satisfy the characterization of excess weight on tail risks in terms of forecast dispersion and accuracy given by equation (5).

6 Concluding Remarks

We have presented a theory of asset price formation based on dispersed information and its aggregation in asset markets. This theory ties expected asset returns to properties of their return distribution and the market's information structure. The theory imposes no restrictions on asset payoffs, investor information and asset supply and therefore speaks to much wider and less stylized asset classes than most of the prior literature on noisy information aggregation. Finally, our theory is tractable and easily lends itself to applications as well as quantitative evaluation of asset pricing puzzles by calibrating model parameters to moments of forecast dispersion and forecast accuracy.

[36] The stronger independence condition in Proposition 3 is required only to show that this excess variance property implies a decomposition into a mean-preserving spread.

In particular we show that our theory can account for a rich set of empirical facts regarding returns to skewness and forecast dispersion in equity and bond markets.

Future work will have to explore the quantitative implications of dispersed information for excess price volatility and return predictability, as well as other asset pricing puzzles. In Albagli, Hellwig, and Tsyvinski (2014) we use our framework to develop a dynamic model of corporate credit spreads. A second direction is to explore the effects of public news and information disclosures. A third direction consists in exploring how market frictions influence real decision-making by firms, households or policy makers. Bassetto and Galli (2019) use a variant of our model to compare information sensitivity of domestic and foreign debt and provide a theory of "original sin". In Albagli, Hellwig, and Tsyvinski (2017), we study the interplay between noisy information aggregation and risk-taking incentives. In an earlier version of this paper, we applied our model to security design and capital structure questions. These applications already suggest that our model may be useful to shed light on other economic phenomena well beyond empirical asset pricing puzzles.

References

[1] Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski (2011). "A Theory of Asset Prices based on Heterogeneous Information," NBER Working paper No. 17548, revised 2015.

[2] Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski (2014). "Dynamic Dispersed Information and the Credit Spread Puzzle," NBER Working paper No. 19788.

[3] Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski (2017). "Imperfect Financial Markets and Shareholder Incentives in Partial and General Equilibrium," NBER Working paper No. 23419.

[4] Bai, Jennie, Robert S. Goldstein, and Fan Yang (2020). "Is the Credit Spread Puzzle a Myth?," Journal of Financial Economics, 137 (2), 297-319.

[5] Barberis, Nicholas and Ming Huang (2008). "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," American Economic Review, 98 (5), 2066 - 2100.

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

[7] Bassetto, Marco and Carlo Galli (2019). "Is Inflation Default? The Role of Information in Debt Crises," American Economic Review, 109 (10), 3556 - 3584.

[8] Bhamra, Harjoat S., Lars-A. Kuehn, and Ilya A. Stebulaev (2010). "The Levered Equity Risk Premium and Credit Spreads: A Unified Framework," Review of Financial Studies, 23 (2), 645-703.

[9] Boyer, Brian, Todd Mitton, and Keith Vorkink (2010). "Expected Idiosyncratic Skewness," Review of Financial Studies, 23 (1), 169-202.

[10] Breon-Drish, Bradyn (2015). "On Existence and Uniqueness of Equilibrium in a Class of Noisy Rational Expectations Models." Review of Economic Studies, 82(3), 868-921.

[11] Brunnermeier, Markus. (2001). *Asset Pricing under Asymmetric Information - Bubbles, Crashes, Technical Analysis and Herding.* Oxford University Press.

[12] Brunnermeier, Markus and Jonathan Parker (2005). "Optimal Expectations," American Economic Review, 95 (4), 1092-1118.

[13] Brunnermeier, Markus, Christian Gollier, and Jonathan Parker (2007). "Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns," American Economic Review, 97 (2), 159-165.

[14] Buraschi, Andrea, Fabio Trojani, and Andrea Vedolin (2014). "Economic Uncertainty, Disagreement, and Credit Markets," Management Science, 60 (5), 1281-1296.

[15] Carlin, Bruce I., Francis A. Longstaff, and Kyle Matoba (2014). "Disagreement and Asset Prices," Journal of Financial Economics, 114 (2), 226-238.

[16] Chabakauri, Georgy, Kathy Yuan, and Konstantinos Zachariadis (2020). "Multi-Asset Noisy Rational Expectations Equilibrium with Contingent Claims," working paper, LSE and QMUL

[17] Chen, Long, Pierre Collin-Dufresne, and Robert S. Goldstein (2009). "On the Relation Between the Credit Spread Puzzle and the Equity Premium Puzzle." Review of Financial Studies, 22 (9), 3367-3409.

[18] Chen, Hui, 2010, "Macroeconomic Conditions and the Puzzles of Credit Spreads and Capital Structure," Journal of Finance, 65 (6), 2171–2212.

[19] Chen, Hui, Rui Cui, Zhiguo He and Konstantin Milbradt (2018). "Quantifying Liquidity and Default Risks of Corporate Bonds over the Business Cycle." Review of Financial Studies, 31 (3), 852-897.

[20] Conrad, Jennifer, Robert Dittmar, and Eric Ghysels (2013). “Ex Ante Skewness and Expected Stock Returns,” Journal of Finance, 68 (1), 85-124.

[21] Diamond, Douglas and Robert Verrecchia (1981). “Information Aggregation in a Noisy Rational Expectations Economy,” Journal of Financial Economics, 9 (3), 221-235.

[22] Diether, Karl, Christopher Malloy, and Anna Scherbina (2002). “Differences of Opinion and the Cross Section of Stock Returns,” Journal of Finance, 57 (5), 2113-2141.

[23] Feldhütter, Peter, and Stephen Schaefer (2018). “The Myth of the Credit Spread Puzzle.” Review of Financial Studies, 31 (8), 2897-2942.

[24] Gebhardt, William R., Charles M. C. Lee, and Bhaskaran Swaminathan (2001). “Toward an Implied Cost of Capital,” Journal of Accounting Research, 39(1), 135-76.

[25] Green, Clifton and Byoung-Hyoun Hwang (2012). “Initial Public Offerings as Lotteries: Skewness Preference and First-Day Returns,” Management Science, 58 (2), 432-444.

[26] Gromb, Denis, and Dimitri Vayanos (2010), “Limits of Arbitrage,” Annual Review of Financial Economics, 2, 251-275.

[27] Grossman, Sanford J. and Joseph E. Stiglitz (1980). “On the Impossibility of Informationally Efficient Markets,” American Economic Review, 70 (3), 393-408.

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

[29] He, Zhiguo and Konstantin Milbradt (2014). “Endogenous Liquidity and Defaultable Bonds.” Econometrica, 82 (4), 1443-1508.

[30] Hellwig, Christian, Arijit Mukherji, and Aleh Tsyvinski (2006). “Self-Fulfilling Currency Crises: The Role of Interest Rates,” American Economic Review, 96 (5), 1769-1787.

[31] Hellwig, Martin (1980). “On the aggregation of Information in Competitive Markets,” Journal of Economic Theory, 22 (3), 477-498.

[32] Hou, Kewei, Chen Xue and Lu Zhang (2020). “Replicating Anomalies,” Review of Financial Studies, 33 (5), 2019-2133.

[33] Huang, Jing-zhi, and Ming Huang (2012). “How much of the Corporate-Treasury Yield Spread is due to Credit Risk?” Review of Asset Pricing Studies, 2 (2), 153-202.

[34] Kremer, Ilan (2002). "Information aggregation in common value auctions." Econometrica, 70 (4), 1675-1682.

[35] Malamud, Semyon (2015). "Noisy Arrow-Debreu Equilibria". Working paper, EPFL.

[36] McQuade, Timothy (2018). "Stochastic Volatility and Asset Pricing Puzzles". Working paper, Stanford GSB.

[37] Milgrom, Paul (1979). "A convergence theorem for competitive bidding with differential information." Econometrica, 47 (3), 679-688.

[38] Milgrom, Paul (1981). "Rational Expectations, Information Acquisition, and Competitive Bidding," Econometrica, 49 (4), 921-943.

[39] Miller, Edward (1977). "Risk, uncertainty, and divergence of opinion." Journal of Finance 32 (4), 1151-1168.

[40] Mitton, Todd and Keith Vorkink (2007). "Equilibrium Underdiversification and the Preference for Skewness", Review of Financial Studies, 20 (4), 1255-1288.

[41] Pálvölgyi,Domotor, and Gyuri Venter (2014). "Multiple Equilibria in Noisy Rational Expectations Economies," Working paper.

[42] Peress, Joel (2004). "Wealth, information acquisition, and portfolio choice." Review of Financial Studies, 17 (3), 879-914.

[43] Perry, Motty, and Philip J. Reny (2006). "Toward a strategic foundation for rational expectations equilibrium." Econometrica, 74 (5), 1231-1269.

[44] Pesendorfer, Wolfgang, and Jeroen M. Swinkels (1997). "The loser's curse and information aggregation in common value auctions." Econometrica, 65 (6), 1247-1281.

[45] Straub, Ludwig, and Robert Ulbricht (2015). "Endogenous Uncertainty and Credit Crunches," working paper, MIT and TSE.

[46] Veldkamp, Laura (2011). *Information Choice in Macroeconomics and Finance*. Princeton University Press.

[47] Vives, Xavier (2008). *Information and Learning in Markets*. Princeton University Press.

[48] Wilson, Robert (1977). "A bidding model of perfect competition," Review of Economic Studies, 44 (3), 511-518.

[49] Yuan, Kathy (2005). "Asymmetric Price Movements and Borrowing Constraints: a Rational Expectations Equilibrium Model of Crises, Contagion and Confusion," Journal of Finance, 60(1), 379- 411.

[50] Yu, Jialin (2011). "Disagreement and Return Predictability of Stock Portfolios," Journal of Financial Economics, 99 (1), 162-183.

[51] Zhang, Xiao-Jun (2013). "Book-to-Market Ratio and Skewness of Stock Returns." The Accounting Review, 88 (6), 2213-2240.

7 Appendix: Proofs

Proof of Proposition 1:

The price function $P_\pi(z) = \mathbb{E}(\pi(\theta)|x = z, z)$ given by (1) is continuous and strictly increasing in z. It then follows from arguments given in the text that when coupled with the threshold $\hat{x}(P) = z$ and the associated posterior beliefs, $P_\pi(z)$ constitutes an equilibrium in which $d(x, P)$ is non-increasing in P. Moreover, by market-clearing, $z = \hat{x}(P_\pi(z))$ and $z' = \hat{x}(P_\pi(z'))$, and therefore $z = z'$ if and only if $P_\pi(z) = P_\pi(z')$. Therefore, the equilibrium conjectured above is the only equilibrium, in which P is informationally equivalent to z.

It remains to be shown that there exists no other equilibrium in which demand is non-increasing in P. In any equilibrium, in which $d(x, P)$ is non-increasing in P, $\hat{x}(P)$ must be non-decreasing in P. Moreover, $\hat{x}(P)$ must be continuous – otherwise, if there were jumps, then there would be certain realizations for z, for which there is no P, such that $\hat{x}(P) = z$, implying that the market cannot clear at these realizations of z. Now, if $\hat{x}(P)$ is strictly increasing in P, it is invertible, and we are therefore back to the equilibrium that we have already characterized. Suppose therefore that $\hat{x}(P)$ $\hat{x}(P) = \hat{x}(P') = \hat{x}(P'')$ for $P \in (P', P'')$ and $P'' > P'$. Suppose further that for sufficiently low $\varepsilon > 0$, $\hat{x}(P)$ is strictly increasing over $(P' - \varepsilon, P')$ and $(P'', P'' + \varepsilon)$, and hence uniquely invertible.[37] But then for $z \in (\hat{x}(P' - \varepsilon), \hat{x}(P')$ and $z \in (\hat{x}(P''), \hat{x}(P'' + \varepsilon))$, $P(z)$ is uniquely defined, so we have

[37] It cannot be flat everywhere, because then informed demand would be completely inelastic, and there would be no way to absorb supply shocks.

$P' \geq lim_{z\uparrow\hat{x}} P(z) = lim_{z\uparrow\hat{x}} \mathbb{E}(\pi(\theta)|x = z, z)$ and $P'' \leq lim_{z\downarrow\hat{x}} P(z) = lim_{z\downarrow\hat{x}} \mathbb{E}(\pi(\theta)|x = z, z)$. But since $\mathbb{E}(\pi(\theta)|x = z, z)$ is continuous, it must be that

$$P'' \leq lim_{z\downarrow\hat{x}} \mathbb{E}(\pi(\theta)|x = z, z) = lim_{z\uparrow\hat{x}} \mathbb{E}(\pi(\theta)|x = z, z) \leq P',$$

which yields a contradiction.

Derivation of equation 4:

Simple algebra shows that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{1-\gamma_P}\sigma_\theta} \phi\left(\frac{\theta - \gamma_P z}{\sqrt{1-\gamma_P}\sigma_\theta}\right) d\Phi\left(\frac{\sqrt{\gamma_V} z}{\sigma_\theta}\right) = \frac{1}{\sigma_P}\phi\left(\frac{\theta}{\sigma_P}\right),$$

where $\sigma_P^2 = (1-\gamma_P)\sigma_\theta^2 + \left(\gamma_P^2/\gamma_V\right)\sigma_\theta^2 = (1 + (\gamma_P/\gamma_V - 1)\gamma_P)\sigma_\theta^2$. Therefore,

$$\mathbb{E}(P(z)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \pi(\theta) \frac{1}{\sqrt{1-\gamma_P}\sigma_\theta} \phi\left(\frac{\theta - \gamma_P z}{\sqrt{1-\gamma_P}\sigma_\theta}\right) d\Phi\left(\frac{\sqrt{\gamma_V} z}{\sigma_\theta}\right) d\theta = \int_{-\infty}^{\infty} \pi(\theta)\frac{1}{\sigma_P}\phi\left(\frac{\theta}{\sigma_P}\right) d\theta$$

and $W(\pi, \sigma_P) = \int_{-\infty}^{\infty} \pi(\theta)\left(\frac{1}{\sigma_P}\phi\left(\frac{\theta}{\sigma_P}\right) - \frac{1}{\sigma_\theta}\phi\left(\frac{\theta}{\sigma_\theta}\right)\right) d\theta = \int_{-\infty}^{\infty}\left(\pi\left(\frac{\sigma_P}{\sigma_\theta}\theta\right) - \pi(\theta)\right) d\Phi\left(\frac{\theta}{\sigma_\theta}\right)$.

Proof of Proposition 2:

Part (i) follows directly from applying Definition 1 in equation (4) and from taking the derivative w.r.t. σ_P. Part (ii) follows from additivity (for given σ_P, $W(\pi_1, \sigma_P) - W(\pi_2, \sigma_P) = W(\pi_1 - \pi_2, \sigma_P)$) and applying part (i) to $\pi_1 - \pi_2$.

For the limit as $\sigma_P \to \infty$, note that $lim_{\sigma_P\to\infty} \int_{-\infty}^{\infty}\left(\pi\left(\frac{\sigma_P}{\sigma_\theta}\theta\right)\right) d\Phi\left(\frac{\theta}{\sigma_\theta}\right) = lim_{\theta\to\infty}\frac{1}{2}(\pi(\theta) + \pi(-\theta))$, and therefore $lim_{\sigma_P\to\infty}|W(\pi, \sigma_P)| = lim_{\theta\to\infty}\frac{1}{2}|\pi(\theta) + \pi(-\theta)|$.

Derivation of equation 5:

Simple algebra shows that

$$\frac{\sigma_P^2}{\sigma_\theta^2} = 1 + \gamma_P\frac{\gamma_P - \gamma_V}{\gamma_V} = 1 + \gamma_P\frac{\beta/\sigma_\theta^2}{\left(1/\sigma_\theta^2 + \beta + \tau\right)\tau} = 1 + \frac{\beta/\sigma_\theta^2}{\left(1/\sigma_\theta^2 + \beta + \tau\right)^2}\frac{\beta + \tau}{\tau}$$

Since $\tilde{D}^2 = \frac{\beta/\sigma_\theta^2}{\left(1/\sigma_\theta^2 + \beta + \tau\right)^2}$ and $\gamma_P(1-\gamma_P) = \frac{(\beta+\tau)/\sigma_\theta^2}{\left(1/\sigma_\theta^2 + \beta + \tau\right)^2}$, it follows that

$$\frac{\sigma_P^2}{\sigma_\theta^2} = 1 + \tilde{D}^2\frac{\gamma_P(1-\gamma_P)}{\gamma_P(1-\gamma_P) - \tilde{D}^2}.$$

Comparative Statics of Credit spreads:

Write $\Phi\left(\frac{\theta}{\sigma_\theta}\right)/\Phi\left(\frac{\theta-\bar{\theta}}{\sigma_P}\right) = \Phi(v)/\Phi(\xi(v-\bar{v}))$, where $v = \underline{\theta}/\sigma_\theta$, $\bar{v} = \bar{\theta}/\sigma_\theta$, and $\xi = \sigma_\theta/\sigma_P$. Comparative statics of expected spreads and returns w.r.t. σ_P^2 are immediate. Taking derivatives w.r.t. v, we obtain

$$\frac{\partial \frac{\Phi(v)}{\Phi(\xi(v-\bar{v}))}}{\partial v} = \frac{\Phi(v)}{\Phi(\xi(v-\bar{v}))}\left(\frac{\phi(v)}{\Phi(v)} - \xi\frac{\phi(\xi(v-\bar{v}))}{\Phi(\xi(v-\bar{v}))}\right).$$

Since $\xi \leq 1$ and $\xi(v-\bar{v}) \geq u$, it follows that $\frac{\phi(v)}{\Phi(v)} \geq \xi\frac{\phi(\xi(v-\bar{v}))}{\Phi(\xi(v-\bar{v}))}$ and $\frac{\partial \frac{\Phi(v)}{\Phi(\xi(v-\bar{v}))}}{\partial v} > 0$. Now, using L'Hôpital's Rule,

$$\lim_{u\to-\infty}\frac{\Phi(v)}{\Phi(\xi(v-\bar{v}))} = \lim_{u\to-\infty}\frac{\phi(v)}{\xi\phi(\xi(v-\bar{v}))} = \xi^{-1}\lim_{u\to-\infty} e^{-\frac{1}{2}v^2+\frac{\xi^2}{2}(v-\bar{v})^2}.$$

which equals 0, whenever $\xi < 1$ ($\sigma_P > \sigma_\theta$) or $\bar{v} < 0$.

Proof of Theorem 1:

We begin with two useful lemmas:

Lemma 1 *Suppose that θ is distributed according to cdf. $H(\cdot)$ and that $f(\cdot)$ is log-concave and $f'(\cdot)/f(\cdot)$ unbounded. Then $H(\theta|x) \equiv \int_{-\infty}^{\theta} f(x-\theta')\,dH(\theta')/\int_{-\infty}^{\infty} f(x-\theta')\,dH(\theta')$ is decreasing in x, with $\lim_{x\to-\infty} H(\theta|x) = 1$ and $\lim_{x\to\infty} H(\theta|x) = 0$.*

Proof. Notice that

$$\frac{H(\theta|x)}{1-H(\theta|x)} = \frac{\int_{-\infty}^{\theta} f(x-\theta')\,dH(\theta')}{\int_{\theta}^{\infty} f(x-\theta')\,dH(\theta')} = \frac{\int_{-\infty}^{\theta}\frac{f(x-\theta')}{f(x-\theta)}dH(\theta')}{\int_{\theta}^{\infty}\frac{f(x-\theta')}{f(x-\theta)}dH(\theta')} = \frac{H(\theta)}{1-H(\theta)}\frac{\mathbb{E}\left(\frac{f(x-\theta')}{f(x-\theta)}|x,\theta'\leq\theta\right)}{\mathbb{E}\left(\frac{f(x-\theta')}{f(x-\theta)}|x,\theta'>\theta\right)}$$

Log-concavity and MLRP of f imply that whenever $\theta' < \theta$, $f(x-\theta')/f(x-\theta)$ is decreasing in x with $lim_{x\to-\infty} f(x-\theta')/f(x-\theta) = \infty$ and $lim_{x\to\infty} f(x-\theta')/f(x-\theta) = 0$. It follows that the second ratio is strictly decreasing in x and converges to 0 as $x\to\infty$ and ∞ as $x\to-\infty$. ■

Lemma 2 *In any equilibrium, and for any P on the interior of the support of $\pi(\theta)$, there exist $x_L(P)$ and $x_H(P)$, such that $d(x,P) = d_L$ for all $x \leq x_L(P)$, $d(x,P) = d_H$ for all $x \geq x_H(P)$, and $d(x,P)$ is strictly increasing in x for $x \in (x_L(P), x_H(P))$*

Proof. For any D, consider the risk-adjusted cdf

$$H(\cdot|P;D) = \frac{\int_{-\infty}^{\theta} U'(D(\pi(\theta)-P))\,dH(\theta|P)}{\int_{-\infty}^{\infty} U'(D(\pi(\theta)-P))\,dH(\theta|P)},$$

and let $H(\cdot|x,P;D)$ and $\mathbb{E}(\pi(\theta)|x,P;D) \equiv \int \pi(\theta)\,dH(\theta|x,P;D)$ denote the cdf and conditional expectations after updating conditional on a private signal x. By lemma 2, $H(\cdot|x,P;D)$ is strictly decreasing in x, $\mathbb{E}(\pi(\theta)|x,P;D)$ is strictly increasing in x and $\lim_{x\to-\infty}\mathbb{E}(\pi(\theta)|x,P;D) < P < \lim_{x\to\infty}\mathbb{E}(\pi(\theta)|x,P;D)$ for any P on the interior of the support of $\pi(\cdot)$. But then there exist $x_L(P)$ s.t. $\mathbb{E}(\pi(\theta)|x_L(P),P;d_L) = P$, which implies that $d(x,P) = d_L$ for all $x \leq x_L(P)$, and $x_H(P)$s.t. $\mathbb{E}(\pi(\theta)|x_H(P),P;d_H(P)) = P$, which implies that $d(x,P) = d_H$ for all $x \geq x_H(P)$.

For $x \in (x_L(P), x_H(P))$ and $x' > x$, lemma 2 implies that $P = \mathbb{E}(\pi(\theta)|x,P;d(x;P)) < \mathbb{E}(\pi(\theta)|x',P;d(x;P))$, or equivalently $\mathbb{E}((\pi(\theta)-P)|x',P;d(x,P)) > 0$. Since the LHS of this condition is strictly decreasing in d, it follows that $d(x',P) > d(x,P)$. ■

Lemmas 1 and 2, and $d(x,P)$ strictly decreasing in P imply that there exists a unique $z(P) \in (x_L(P), x_H(P))$; s.t. $d(z(P),P) = \bar{D}$, or equivalently, $P = P_\pi(z) = \mathbb{E}\left(\pi(\theta)|x = z(P), P; \bar{D}\right)$. Combining with the equilibrium price function, we then define a candidate sufficient statistic function $z(\theta,u) = z(P(\theta,u))$, and since $z(P)$ is invertible, z must be a sufficient statistic for the information contained in P. Therefore we obtain the representation (7), along with representation (8) of equilibrium beliefs.

Proof of Theorem 2:

Parts (i) is well-known. Part (ii) follows from additivity, $W\left(\pi_1,\hat{H}\right) - W\left(\pi_2,\hat{H}\right) = W\left(\pi_1 - \pi_2,\hat{H}\right)$, and applying part (i) to $\pi_1 - \pi_2$.

To complete the proof, we show that $\left|W\left(\pi,\hat{H}\right)\right|$ may become arbitrarily large if $\hat{H}(\cdot)$ converges to an improper distribution characterized by $\hat{H}(\theta) \to \bar{H} \in (0,1)$. Suppose that π is convex (the proof for concave π is analogous), and note that we can write $\pi(\theta) = \tilde{\pi}(\theta) + \psi\theta$, where $\psi = lim_{\theta\to-\infty}\pi'(\theta)$, and $\tilde{\pi}(\theta)$ is bounded below, non-decreasing, and convex. Furthermore, $W\left(\pi,\hat{H}\right) = W\left(\tilde{\pi},\hat{H}\right)$. Now taking the limit as $\hat{H}(\theta) \to \bar{H} \in (0,1)$,

$$lim_{\hat{H}(\theta)\to\bar{H}} W\left(\tilde{\pi},\hat{H}\right) = \left(1-\bar{H}\right) lim_{K\to\infty}\tilde{\pi}(K) + \bar{H} lim_{K\to\infty}\tilde{\pi}(-K) - \mathbb{E}(\pi(\theta)) = \infty,$$

since $\tilde{\pi}(\theta)$ is bounded below but unbounded above for convex π.

Derivation of Equation 10:

To prove the first equality, rewrite $W\left(\pi, \hat{H}\right)$ as

$$\begin{aligned} W\left(\pi, \hat{H}\right) &= \int_{-\infty}^{\infty} \left(H(\theta) - \hat{H}(\theta+\delta)\right) d\pi(\theta) = -\int_{-\infty}^{\infty} \left(H(-\theta) - \hat{H}(-\theta+\delta)\right) d\pi(\theta) \\ &= -\int_{-\infty}^{\infty} \left(H(\theta) - \hat{H}(\theta+\delta)\right) d\pi(-\theta) \end{aligned}$$

and therefore

$$\int_{-\infty}^{\infty} \left(H(\theta) - \hat{H}(\theta+\delta)\right) d\psi(\theta) = \frac{1}{2}\left\{W\left(\pi, \hat{H}\right) - W\left(-\pi, \hat{H}\right)\right\} = W\left(\hat{\pi}, \hat{H}\right).$$

The second equality then follows from symmetry, $\hat{\pi}(\theta) = \hat{\pi}(-\theta)$ and $H(\theta) - \hat{H}(\theta+\delta) = \hat{H}(-\theta+\delta) - H(-\theta)$.

Proof of Proposition 3:

We represent θ under the risk-neutral measure as $\mathbb{E}(\theta|x=z,z)+v$, where $v = \theta - \mathbb{E}(\theta|x=z,z)$. If $\theta - \mathbb{E}(\theta|x,z)$ is independent of x, v has the same probability distribution as $\theta - \mathbb{E}(\theta|x,z)$, and $\mathbb{E}(\theta|x=z,z)+v$ has the same distribution as $\theta - \mathbb{E}(\theta|x,z) + \mathbb{E}(\theta|x=z,z)$. But if $\mathbb{E}(\theta|x=z,z) - \mathbb{E}(\theta|x,z)$ is independent of θ, then $\theta - \mathbb{E}(\theta|x,z) + \mathbb{E}(\theta|x=z,z)$ is a mean-preserving spread over $\theta+\delta$,where $\delta = \mathbb{E}(\mathbb{E}(\theta|x=z,z)) - \mathbb{E}(\theta)$. Hence the risk-neutral distribution decomposes into a shift in means and a mean-preserving spread over the objective distribution of dividends.