A Theory of Asset Prices based on Heterogeneous Information

Elias Albagli
Central Bank of Chile

Christian Hellwig
Toulouse School of Economics

Aleh Tsyvinski
Yale University

July 8, 2015

Abstract

With only minimal restrictions on security payoffs and trader preferences, noisy aggregation of heterogeneous information drives a systematic wedge between the impact of fundamentals on the price of a security, and the corresponding impact on risk-adjusted cash flow expectations. From an ex ante perspective, this wedge leads to a systematic gap between an asset’s expected price and its expected dividend, whose sign and magnitude depend on the asymmetry between upside and downside payoff risks and on the importance of information heterogeneity. We consider applications of our theory to the negative relationship between returns and skewness and optimal security design.

*We thank Dimitri Vayanos, the editor, and three anonymous referees. We also thank Bruno Biais, John Geanakoplos, Narayana Kocherlakota, Felix Kubler, Stephen Morris, Guillaume Plantin, Jeremy Stein, Jean Tirole, Dimitri Vayanos, Xavier Vives, Olivier Wang, Martin Weber, Eric Young, and audiences at numerous conferences and seminars for helpful comments. Hellwig gratefully acknowledges financial support from the European Research Council (starting grant agreement 263790). Tsyvinski is grateful to NSF for support and EIEF for hospitality.
1 Introduction

We develop a parsimonious, flexible theory of asset pricing in which heterogeneity of information and its aggregation in the market emerges as the core force determining asset prices and expected returns. Under mild restrictions on security payoffs, trader preferences and the information structure, we show that noisy aggregation of dispersed information results in market prices that are more sensitive to fundamental and liquidity shocks than the corresponding risk-adjusted dividend expectations. Prices are then higher (lower) than dividend expectations whenever the information aggregated through the price is sufficiently (un-)favorable. Moreover, when assets payoffs are asymmetric, the sensitivity of prices to shocks ex post also affects expectations of prices ex ante: on average, securities characterized by upside risk are thus priced above their fundamental value, while the opposite is true for securities characterized by downside risk. As applications of our theory, we illustrate that our model is consistent with the empirical regularities on the negative return to skewness, and we use our model to discuss how information aggregation frictions affect optimal security design by offering a new rationale for tranching a given cash flow into debt and equity.

We consider an asset market along the lines of Grossman and Stiglitz (1980), Hellwig (1980), and Diamond and Verrecchia (1981). An investor pool consists of informed traders who observe a noisy private signal about the value of an underlying cash flow, and uninformed noise traders. The price is set to equate the demand by informed and noise traders to the available asset supply. The price then serves as an endogenous, noisy public signal of the asset’s cash flow.

As our first main result, in section 2, we provide a very general representation of the equilibrium security price. Consider a trader who, at equilibrium, finds it optimal to hold a given, arbitrary, exposure to the security. Under mild regularity conditions, this trader’s private signal varies monotonically with the price, and can thus be used as a sufficient statistic for the information conveyed through the price. The equilibrium price can then be represented as a function of just this sufficient statistic, corresponding to the traders’ risk-adjusted expectation of dividends. This representation allows us to decompose the equilibrium price into the trader’s dividend expectation and risk compensation required for the given exposure level. What’s more, since the initial exposure level was arbitrary, we obtain a whole class of equivalent representations of the equilibrium price in terms of dividend expectations and risk compensation, each tailored to a different exposure level.

As an immediate consequence of this representation, we obtain that the price is generically more sensitive to fundamental and liquidity shocks than the risk-adjusted expectation of dividends

---

1 See Brunnermeier (2001), Vives (2008), and Veldkamp (2011) for textbook discussions.
conditional on the information conveyed through the price. With dispersed information, the asset price serves to aggregate information, but must also adjust to clear the market. An increase in demand resulting from a more favorable realization of the payoff fundamental, or an increase in the noise traders’ demand, must be met by a price increase to clear the market, even without considering the inference drawn from the price. This market-clearing effect is then compounded by the information conveyed by the price increase, as traders use Bayes’ rule to update their beliefs about fundamentals and noise-trading shocks from prices. The risk-adjusted expectation of dividends also incorporates this update from the information contained in the price but does not incorporate the extra adjustment in prices that is due to market-clearing.

In section 3, we specialize our model to the case in which traders are risk-neutral and face limits on their asset positions. Absent risk premia, noisy information aggregation is then the only force shaping asset prices and returns. Using a market structure first introduced in Hellwig, Mukherji and Tsyvinski (2006), we characterize the equilibrium price and the expected fundamental value in closed form, without imposing any restrictions on the asset’s payoff risk. We then characterize the average price and ex ante expected dividends, as a function of the cash flow distribution and a parameter summarizing information frictions, which depends on the accuracy of informed traders’ private signals and the variance of noise trading shocks.

Our second main result uses this equilibrium characterization to link the asset’s expected price to payoff asymmetry and information frictions. Formally, we establish a negative relation between skewness and returns that is consistent with the data: assets that are dominated by upside risk and thus positively skewed trade at a price premium, while assets that are dominated by downside risk or negatively skewed trade at a price discount. The intuition for this result is as follows: With asymmetric returns, the adjustment of equilibrium prices to shocks is larger in the direction that exposes investors to more uncertainty in payoffs. For securities that are characterized by upside risk, the positive price gaps then dominate, and on average the security trades at a premium. The opposite is true for securities that are characterized by downside risk. In absolute value the unconditional wedge is larger for more asymmetric payoff risks, or for more severe information aggregation frictions, and due to an increasing difference property, payoff asymmetry and information frictions are mutually reinforcing.

In section 4, we study models with risk-averse informed traders. To clearly identify the respective roles of risk aversion and information aggregation, we compare our model with risk aversion to a counterpart in which traders have common information, and prices only incorporate risk premia. In section 4.1 we show how the same over-reaction to the information contained in the price is
present in the standard model with CARA preferences and normally distributed dividends. However due to the symmetry of payoffs, information aggregation does not affect average prices, which differ from expected dividends only through risk premia.

In section 4.2 we analyze a model with CARA preferences and binary asset payoffs, which allows us to consider asymmetric payoff risks. As in the case with risk-neutral traders, noisy information aggregation has a positive impact on average prices if and only if the asset payoffs are characterized by upside risk. Moreover, the impact of asymmetry on average prices is of a similar order of magnitude as the standard risk premium term. In comparison, payoff asymmetry has the same qualitative effect on risk premia and asset prices in the corresponding benchmark economy with common information, but the effect is an order of magnitude smaller. Dispersed information is thus key for amplifying the link between payoff asymmetry and asset prices.\(^2\)

In section 4.3 we numerically solve the binary model with CRRA preferences and show that the results parallel those from the CARA-binary case. Payoff asymmetry has an impact on asset prices that is of similar magnitude as “standard” risk premia, when information is dispersed, but much smaller when traders only have access to common information.

In section 5, we develop two applications of the model. First, we show the model is consistent with the empirically documented negative relationship between returns and skewness.\(^3\) This numerical illustration suggests that even moderate information frictions can generate sizable and empirically plausible excess returns from skewness.

Second, we consider how a seller of a cash-flow may influence its market value by tranching the cash-flow and selling it to different investor pools. The seller’s expected revenue is not affected by the split, if and only if the different investor pools have identical informational characteristics. When investor pools differ, the seller can increase expected revenue by selling downside risks in the market with smaller, and upside risks in the market with larger information aggregation frictions. The optimal security design completely separates upside and downside risks, splitting the cash flow into a debt claim for the downside, and an equity claim for the upside, with a default point for debt at the prior median. In contrast to existing theories of capital structure or optimal security design, our results are not driven by asymmetric information or incentive problems between insiders and outsiders or shareholders and managers. Instead, we focus entirely on how market frictions shape security design incentives - in particular, how the relation between risk asymmetry and prices shapes the incentives to issue debt, which represents a downside risk, and equity, which is tilted

\(^2\)We thank the editor, Dimitri Vayanos, for pushing us to clarify this point.

towards the upside.

Our paper contributes to the literature on noisy information aggregation in asset markets in two ways. First, whereas most of the existing literature imposes strong parametric assumptions for tractability (such as CARA preferences, or normally distributed signals and dividends), our main equilibrium characterization is almost completely free of assumptions about the underlying primitives, which enables us to offer general insights into how noisy information aggregation affects asset prices and returns.\textsuperscript{4} Second, the variant of our model with risk-neutral traders offers closed form solutions without any restrictions on cash-flow distributions, and therefore may serve as a tractable alternative to the existing workhorse models. Breon-Drish (2011, forthcoming) analyzes non-linear and non-normal variants of the noisy REE framework in the broad exponential family of distributions. He further derives powerful results on the incentives for information acquisition in this environment, whereas we take the information structure as given.\textsuperscript{5}

Our equilibrium characterization with noisy information aggregation by means of a “sufficient statistic” variable (Theorem 1) shares similarities with common value auctions; these similarities are even more pronounced in the case with risk-neutral agents and position bounds (Section 3, Theorem 2).\textsuperscript{6} 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.

More generally, any theory of mispricing must rely on some source of noise affecting the market, coupled with limits to the traders’ ability or willingness to exploit arbitrage (see Gromb and Vayanos, 2010, for an overview and numerous references). We show that noise trading under heterogeneous 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.

\textsuperscript{4}To our knowledge, Vives (2008) is the only written statement of the observation that information aggregation drives a wedge between asset prices and expected dividends in the CARA-normal model. Moreover, under the standard assumptions of normally (i.e. symmetrically) distributed dividends necessary to solve such models, any unconditional excess return is attributable to a risk premium alone.

\textsuperscript{5}Barlevy and Veronesi (2003) and Yuan (2005) also study non-linear models of noisy information aggregation, but in each case restricting themselves to specific parametric examples.

\textsuperscript{6}See Wilson (1977), Milgrom (1979, 1981b), Pesendorfer and Swinkels (1997), Kremer (2002) and Perry and Reny (2006) for important contributions to this literature. However note that our THM 1 doesn’t require the restrictions of risk-neutrality or unit demand that typically characterize the auction-theoretic literature.
2 The Model

2.1 Agents, assets, information structure and financial market

The market is set as a Bayesian trading game with a unit measure of informed traders. The dividend of the risky asset is given by a strictly increasing and twice continuously differentiable function \( \pi(\cdot) \) of a stochastic fundamental, \( \theta \). Nature draws \( \theta \in \mathbb{R} \) according to a distribution with a smooth density function \( h(\cdot) \). Each informed trader \( i \) then receives a noisy private signal \( x_i = \theta + \epsilon_i \), where \( \epsilon_i \) is i.i.d across traders, and distributed according to cdf. \( F: \mathbb{R} \rightarrow [0,1] \) and smooth density function \( f \).\(^7\) We assume \( f'(\cdot)/f(\cdot) \) is strictly decreasing and unbounded above and below.\(^8\)

Traders’ preferences are characterized by a strictly increasing, concave utility function \( U: \mathbb{R} \rightarrow \mathbb{R} \) defined on the traders’ realized gains or losses \( d_i \cdot (\pi(\theta) - P) \), where \( d_i(\cdot) \) is a price-contingent demand schedule restricted to lie on the interval \([d_L(P), d_H(P)]\). Here, \( d_L(P) < 0 < d_H(P) \) are arbitrary, continuous, price-contingent limits. Individual trading strategies are a mapping \( d: \mathbb{R}^2 \rightarrow [d_L(P), d_H(P)] \) from signal-price pairs \((x_i, P)\) into asset holdings. Aggregating across traders leads to the aggregate informed demand, \( D: \mathbb{R}^2 \rightarrow [0,1], D(\theta, P) = \int d(x, P)dF(x - \theta), \) where \( F(x - \theta) \) is the cross-sectional distribution of private signals \( x_i \), conditional on \( \theta \). The supply of securities is stochastic, given by a function \( S(u, P) \in [d_L(P), d_H(P)] \) that is increasing in both the price \( P \) and a supply shock \( u \). The shock \( u \) is distributed according to cdf \( G(\cdot) \). Once traders submit their orders, a price \( P \) is selected to clear the market. Formally, let \( \hat{P}: \mathbb{R}^2 \rightarrow \mathbb{R}, \hat{P}(\theta, u) = \{ P \in R : D(\theta, P) = S(u, P) \} \), denote the correspondence of market-clearing prices. A price function \( P: \mathbb{R}^2 \rightarrow \mathbb{R} \) clears the market if and only if \( P(\theta, u) \in \hat{P}(\theta, u) \), for all \((\theta, u) \in \mathbb{R}^2 \). Let \( H(\cdot|P) : \mathbb{R} \rightarrow [0,1] \) denote the posterior cdf of \( \theta \), conditional on observing the market price \( P \). Then the informed traders’ posterior is defined from Bayes’ Rule as \( H(\theta|x, P) = \int_0^\theta f(x - \theta')dH(\theta'|P)/ \left( \int_\theta^\infty f(x - \theta')dH(\theta'|P) \right) \), and their decision problem is

\(^7\)The conventional assumption according to which traders observe noisy signals of dividends is nested by setting \( \pi(\theta) = \theta \). Our generalization separates the updating from prices from the distribution of underlying asset returns, which in turn allows us to derive pricing implications without imposing strong assumptions on asset payoffs. Our formulation has a natural interpretation in the context of firm-specific uncertainty, where investors process information about a firm’s fundamentals (e.g. its future earnings \( \theta \)) to forecast the payoffs to securities the firm has issued, e.g. debt or equity claims, represented by different functions \( \pi(\cdot) \), and the same information about a firm’s earnings will result in different updates for different securities.

\(^8\)Monotonicity of \( f'(\cdot)/f(\cdot) \) implies signals have log-concave density and satisfy the monotone likelihood ratio property. Unboundedness implies extreme signal realizations induce large updates in posterior beliefs, (almost) regardless of the information contained in other signals.
Theorem 1: Let $P(\theta, u)$, and posterior beliefs $H(\cdot | P)$ such that (i) $d(x, P)$ is optimal given $H(\cdot | x, P)$; (ii) the asset market clears for all $(\theta, u)$; and (iii) $H(\cdot | P)$ satisfies Bayes’ rule whenever applicable, i.e., for all $p$ such that $\{(\theta, u) : P(\theta, u) = p\}$ is non-empty.

2.2 A General Characterization Result

We begin our analysis with a general characterization result about the equilibrium structure, assuming that such an equilibrium exists.\(^9\)

**Theorem 1**: Let $\{P(\theta, u); d(x, P); H(\cdot | P)\}$ be a Perfect Bayesian Equilibrium. Assume that $H(\cdot | P)$ admits a continuous density function $h(\cdot | P)$, which is everywhere positive. Let $\bar{S}(P)$ denote an arbitrary continuous function of $P$, such that $\bar{S}(P) \in (d_L(P), d_H(P))$. Then, the following two conditions are equivalent:

1. $\bar{S}(P) - d(x, P)$ is strictly increasing in $P$, for $x$ s.t. $d(x, P) = \bar{S}(P)$.
2. There exists a sufficient statistic function $z(\theta, u)$, with cdf $\Psi(z'|\theta) = \Pr(z(\theta, u) \leq z'|\theta)$ and density $\psi(z'|\theta)$ such that $P(\theta, u) = P_\pi(z(\theta, u))$, where $P_\pi(z)$ is invertible and satisfies\(^10\)

$$P_\pi(z) = \frac{\mathbb{E} (U'(\bar{S}(P_\pi(z)) (\pi(\theta) - P_\pi(z))) \cdot \pi(\theta) | x = z, z)}{\mathbb{E} (U'(\bar{S}(P_\pi(z)) (\pi(\theta) - P_\pi(z))) | x = z, z)}$$

(2)

Theorem 1 introduces a class of sufficient statistic representations for any equilibrium that satisfies mild regularity conditions on posterior beliefs. Formally, there exists a random variable $z$, function of $\theta$ and $u$ only, which is informationally equivalent to the price. Moreover, the price can be represented as a function of this sufficient statistic, as in equation (2).

The theorem derives its interest not only from the existence of such a sufficient statistic but also from the characterization of equilibrium prices that it entails. Formally, the sufficient statistic representation (2) characterizes the price as the risk-adjusted expectation of dividends of a trader.

---

\(^9\)Existence is guaranteed once we impose more structure later on for preferences and distributions. To our knowledge, no general existence results are available for this class of models.

\(^10\)We index an equilibrium function or variable by $\pi$ to make explicit that it is derived from a specific dividend function $\pi(\cdot)$, i.e. $P_\pi(\cdot)$ is the equilibrium price function that is derived from dividend function $\pi(\cdot)$ by equation (2).
who finds it optimal to hold exactly $\bar{S}(P)$ units of the asset, and provides a decomposition into
two terms: the expected dividend, and a compensation for risk. However, the decomposition into
these two components remains indeterminate as it depends on the choice of the exposure function
$\bar{S}(\cdot)$.\textsuperscript{11} The intuition for why the decomposition into expected dividends and risk premium not
unique is as follows. In equilibrium, a trader with a larger demand of the asset (measured by $\bar{S}(\cdot)$)
requires a higher risk premium. For a given asset price, this is possible only if the trader is also
more optimistic about the asset’s expected dividend. At one extreme, a choice of $\bar{S}(P) = 0$ results
in a representation of the price purely as an expected dividend. The larger is $\bar{S}(P)$, the higher is
the required risk compensation, and hence also the dividend expectation of the trader who holds
$\bar{S}(P)$. A reasonable decomposition of prices into expected dividends and compensation for risk
then requires appropriate choice of $\bar{S}(P)$, for example by setting it equal to the expected asset
holdings by informed traders.

The key result of the theorem is that at the interim stage – when the price is observed but
before dividends are realized – the market-implied posterior over $\theta$ conditional on $P$ differs from
the Bayesian posterior conditional on the same public information. The market acts as if the price
signal (or equivalently, the sufficient statistic $z$) enters twice into the updating, once as a public price
signal, and once as the private signal of the threshold trader who finds it optimal to purchase exactly
$\bar{S}(P)$ units of the asset. In the proof of the theorem, we show that for any equilibrium, any $\bar{S}(P)$,
and any $P$, we can construct such a threshold signal $z(P)$, which in turn is sufficient to establish
the wedge between market-implied and objective posterior distributions. Condition (1) is then
required only to show that $z(P)$ is invertible, validating its use as a sufficient statistic. The wedge
between market-implied and objective posterior distributions is thus a necessary characteristic for
any model with noisy information aggregation through share prices.

**Intuition:** The wedge between market-implied and objective posterior expectations is a direct
consequence of market-clearing forces. In equilibrium, demand is increasing in $x$. Suppose that
it is also decreasing in $P$, and with risk aversion or position bounds, it is not perfectly price-
elastic. An increase in the fundamental $\theta$ has two effects (the same logic applies for a decrease in
$u$, corresponding to a supply contraction). First, holding posteriors $H(\cdot|P)$ fixed, an increase in $\theta$
increases demand because traders receive more optimistic private signals. For a given realization
of $u$, this increases the market-clearing price. The reasoning so far follows solely from market
clearing, even without considering agents learning from $P$. Second, the resulting change in $P$\textsuperscript{11} Obviously all these sufficient statistics representations are equivalent, since they amount to nothing other than
monotonic transformations of the price, and hence of each other.
shifts the posterior $H(\cdot|P)$ as traders update beliefs about dividends from observing a higher price. This update uses Bayes’ Rule, and reinforces the first (market-clearing) effect of the price. In the expression for the price (2) the two effects are represented by the sufficient statistics $z$ appearing twice. In contrast, the objective posterior of $\theta$ given $P$ only includes the second effect.

The equilibrium price $P_\pi(z)$ can be interpreted as the risk-adjusted dividend expectation of the trader whose demand is equal to $\bar{S}(P)$. This trader conditions on the information in the price $z$, as well as a private signal. The market clearing condition imposes a restriction on the identity of this trader: By construction, the realization of this trader’s private signal is $z(\theta,u)$ and shifts with the underlying shocks $\theta$ and $u$. Hence, despite its appearance, the wedge is not a consequence of irrational trading decisions, but results from market clearing with investor heterogeneity and is perfectly consistent with Bayesian rationality.

**Common Information Benchmark:** We now compare our model to an otherwise identical economy in which all traders shared the same information. Such a comparison is not straightforward because the information conveyed through prices is endogenously linked to the model’s primitives. Nevertheless, taking as given a supply or exposure $S$ and an exogenous public signal $z$, we obtain a homogeneous information formula for the asset price:

$$V_\pi(z) = \frac{E(U'(S(\pi(\theta) - V_\pi(z))) \cdot \pi(\theta)|z)}{E(U'(S(\pi(\theta) - V_\pi(z)))|z)} = E(\pi(\theta)|z) + \frac{cov(U'(S(\pi(\theta) - V_\pi(z)))\cdot\pi(\theta)|z)}{E(U'(S(\pi(\theta) - V_\pi(z)))|z)}.$$  

Compared to the heterogeneous information representation with $\bar{S}(P)$ set equal to $S$, this characterization highlights the wedge in posterior beliefs that is due to noisy information aggregation. This wedge shifts posteriors in both the expected dividend and the risk premium term.

This comparison also highlights the two crucial ingredients for the impact of noisy information aggregation on asset prices. The first are 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 traders 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 $\theta$, almost

---

12 For example, if we were to simply remove the existence of private signals from the economy but kept other primitives the same, then the price would no longer convey any information about fundamentals. If instead we consider an economy with fixed supply and replicate the information aggregated through prices by a public signal with the same statistical properties, we ignore the fact that supply shocks themselves contain an element of risk for which risk-averse traders require compensation. Finally, if we consider an economy with supply shocks and such a public signal, we leave out the fact that the model of noisy information aggregation introduces an endogenous correlation between supply shocks and information. These examples illustrate the challenge of defining an otherwise identical economy in which information is homogeneous. We return to this issue below in section 4.
surely), while noise trader shocks get seamlessly absorbed by the market. The second crucial ingredient is heterogeneous information. Indeed the difference between $P_\pi(z)$ and $V_\pi(z)$ disappears in the limit where private signals become infinitely noisy, since then the market-implied and objective posterior converge to each other and to the prior if $P$ ceases to carry any information about $\theta$.

In summary, the wedge between market-implied and objective posteriors requires heterogeneity of beliefs in equilibrium, once the informational content of prices has been accounted for.

**Implications for Expected Returns:** We can also characterize how dispersed information influences expected returns. The asset’s realized return is $R(\theta, z) = \pi(\theta) / P_\pi(z)$. Therefore, for given exposure $\bar{S}(\cdot)$, the market-implied return expectations only result from the existence of a risk premium:

$$
E(R(\theta, z) | x = z, z) = 1 - \frac{\text{cov}(U' (\bar{S}(P_\pi(z)) (\pi(\theta) - P_\pi(z))) ; \pi(\theta) | x = z, z)}{E(U' (\bar{S}(P_\pi(z)) (\pi(\theta) - P_\pi(z))) \cdot \pi(\theta) | x = z, z)}.
$$

The expected returns then correct the market-implied expected returns by applying the correct expectations of dividends:

$$
E(R(\theta, z) | z) = \frac{E(\pi(\theta) | z)}{E(\pi(\theta) | x = z, z)} E(R(\theta, z) | x = z, z).
$$

Hence, the expected return consists of two factors: the risk premium that is embedded in the market-implied expected return, and the ratio between the objective and the market-implied expectation of dividends. The latter factor is a novel implication of the dispersed information economy.\(^\text{13}\)

**Limitation of Partial Characterization:** Unfortunately, Theorem 1 offers only a partial characterization of the equilibrium asset price: it characterizes asset valuations only once we have been able to compute, for some $\bar{S}(P)$, the distribution of the associated sufficient statistic $z$. This distribution, however, derives from the market clearing condition $D(\theta, P) = S(u, P)$, which still requires information about the entire demand schedule. We are thus still left with a fairly complex fixed point problem. We nevertheless view this characterization as an important result. First, as we discussed above, the theorem shows exactly how dispersed information models differ from their common information counterparts. This property is underlying the structure of many well-known dispersed information models such as the canonical CARA-normal model of Hellwig (1980) and Diamond and Verrecchia (1981), or the recent non-linear generalizations in Breon-Drish (2012). Second, the sufficient statistics characterization is a key intermediate step for the full equilibrium solution in dispersed information models, either for numerical solution procedures, or for cases where we are able to obtain closed form solutions with additional assumptions about preferences and shocks. In the next sections we derive such solutions for special cases.

---

\(^\text{13}\)Similarly, expected log-returns decompose additively into a risk premium and an information aggregation term.
3 The Risk-neutral, Normal Model

In this section, we focus on a special case of our general model in which traders are risk-neutral, and position limits \( d_L (P) \) and \( d_H (P) \) bind. This case is of special interest because, since traders are risk-neutral, the asset pricing implications result exclusively from the wedge between objective and market-implied dividend expectations. We further introduce functional form assumptions so that the updating from prices preserves normality, and therefore the model becomes tractable for very general dividend distributions. We refer to this as the risk-neutral normal model.\(^{14}\)

3.1 Model and Equilibrium Characterization

Formally, we introduce the following two additional assumptions:

(1) Traders are risk-neutral, and their positions are bounded by \( d_L (P) = 0 \) and \( d_H (P) = 1 \).

(2) Asset supply is inelastic and given by \( S(u, P) = S(u) \).

Monotonicity and unboundedness of \( f' / f \) imply that \( \int \pi (\theta) dH\left(\theta | x, P\right) \) is an increasing function of \( x \) (Milgrom, 1981a), and there exists a unique \( \hat{x} \), such that \( \int \pi (\theta) dH\left(\theta | \hat{x}, P\right) = P \). When traders are risk-neutral, any trader demands \( d(\hat{x}, P) = 0 \) if \( x < \hat{x} (P) \) and \( d(\hat{x}, P) = 1 \) if \( x > \hat{x} (P) \), and \( \hat{x}(P), P \in [0, 1] \). Furthermore, the aggregate demand for the asset, given \( P \) and \( \theta \), is \( 1 - F(z - \theta) \), and market-clearing requires that \( 1 - F(z - \theta) = S(u) \), or equivalently \( z \equiv \hat{x}(P) = \theta + F^{-1} (1 - S(u)) \). This sufficient statistic \( z \) is centered at \( \theta \) and distributed according to cdf \( \Psi (z - \theta) = 1 - G\left(S^{-1} (1 - F(z - \theta))\right) \), with the corresponding pdf \( \psi(z - \theta) \). This characterization of the distribution \( \Psi (z - \theta) \) of the sufficient statistic is uniquely defined from the primitives \( G(\cdot), F(\cdot), \) and \( S(\cdot) \), and independent of the dividend function \( \pi(\cdot) \). We then arrive at the following proposition showing that the class of sufficient statistic characterizations in Theorem 1 reduces to a single sufficient statistic characterization \( z = \hat{x}(P) \), for which \( P_{\pi}(z) = \mathbb{E}(\pi(\theta) | x = z, z) \).\(^{15}\)

Proposition 1: Let \( \psi(\cdot) \) be as defined above, and suppose that the function

\[
P_{\pi}(z) = \mathbb{E}(\pi(\theta) | x = z, z) = \frac{\int \pi(\theta) \psi(z - \theta) f(z - \theta) h(\theta) d\theta}{\int \psi(z - \theta) f(z - \theta) h(\theta) d\theta}\]

is strictly increasing in \( z \). Then the price function \( P_{\pi}(z) \) characterizes the unique equilibrium in which the traders’ demand is non-increasing in \( P \).

\(^{14}\)The risk-neutral model with position limits also strikes us as a natural laboratory for analyzing aggregation of information about firm-specific risks, which investors should be able to diversify by investing across a wide range of assets. In practice such diversification can be achieved by limiting ex ante the wealth that is invested in any given security, akin to the position limits in our model. See Albagli, Hellwig and Tsyvinski (2014b) for further discussion.

\(^{15}\)In other words, the risk premium term disappears from the equilibrium characterization.
Thus with inelastic asset supply and position limits, we obtain a closed form characterization of the unique equilibrium, in which demand is non-increasing in $P$. This monotonicity restriction can be justified for instance by assuming that trade takes place through a limit-order book. The characterization requires $\mathbb{E}(\pi(\theta) | x = z, z)$ to be strictly increasing in $z$, otherwise this equilibrium fails to exist.\(^\text{16}\) Monotonicity w.r.t. $z$ holds for example whenever the density $\psi(\cdot)$ is log-concave, which imposes restrictions on $G(\cdot)$, $F(\cdot)$, and $S(\cdot)$.

Assumptions (1) and (2) mainly serve to simplify the characterization of the sufficient statistic $z$ in closed-form. As long as traders are risk-neutral, there exists a threshold signal $\hat{x}(P)$ that serves as indifference point for the traders and fully pins down the asset demand. For general position bounds and supply functions, the market-clearing condition is then given by

$$1 - F(z - \theta) = \frac{S(u, P) - d_L(P)}{d_H(P) - d_L(P)} \equiv \hat{S}(u, P).$$

Thus changes in position bounds are isomorphic to changes in the supply schedule $S(u, P)$. For a given price function $P_\pi(z)$, the cdf $\Psi(z|\theta)$ satisfies $\Psi(z|\theta) = 1 - G(\hat{u}(z, \theta))$, where $\hat{u}(z, \theta)$ is implicitly defined by the market-clearing condition $1 - F(z - \theta) = \hat{S}(\hat{u}, P_\pi(z))$. The equilibrium is then implicitly defined as a fixed point between the equation characterizing the equilibrium price $P_\pi(z)$, and the characterization of the cdf. $\Psi(z|\theta)$ from the market-clearing condition.

To specialize the characterization even further, suppose that $\theta$ is distributed according to $\theta \sim \mathcal{N}(0, \sigma^2_\theta)$, private signals are distributed according to $x_i \sim \mathcal{N}(\theta, \beta^{-1})$, so $F(x - \theta) = \Phi(\sqrt{\beta}(x - \theta))$, and the asset supply is $S(u) = \Phi(u)$, where $\Phi(\cdot)$ is the cdf of a standard normal distribution, and $u \sim \mathcal{N}(0, \sigma^2_u)$. This supply assumption is adapted from Hellwig, Mukherji and Tsyvinski (2006). Market clearing then implies $1 - \Phi(\sqrt{\beta}(\hat{x}(P) - \theta)) = \Phi(u)$, and therefore $z = \hat{x}(P) = \theta - 1/\sqrt{\beta} \cdot u$. It follows that $z|\theta \sim \mathcal{N}(\theta, \sigma^2_u/\beta)$, and the equilibrium price and expected dividends are given by:\(^\text{17}\)

$$P_\pi(z) = \int \pi(\theta) d\Phi\left(\frac{1}{\sigma^2_\theta} + \frac{\beta}{\sigma^2_u}\left(\theta - \frac{\beta + \beta/\sigma^2_u}{1/\sigma^2_\theta + \beta/\sigma^2_u} z\right)\right),$$  \(3\)

$$V_\pi(z) = \int \pi(\theta) d\Phi\left(\frac{1}{\sigma^2_\theta} + \frac{\beta}{\sigma^2_u}\left(\theta - \frac{\beta/\sigma^2_u}{1/\sigma^2_\theta + \beta/\sigma^2_u} z\right)\right).$$  \(4\)

\(^\text{16}\)If $\mathbb{E}(\pi(\theta) | x = z, z)$ is non-monotone, then the market price no longer suffices to infer $z$, resulting in violations of market-clearing for some realizations of shocks. The only alternative then is that (i) the equilibrium demand is non-monotone in $P$, and (ii) the equilibrium price function is discontinuous and conveys finer information than the partition on the space of $(u, \theta)$ that is induced by the sufficient statistic $z$. Such equilibria cannot easily be ruled out; see e.g. Pálvölgyi and Venter (2014) for an example of such a construction in the Grossman-Stiglitz model.

\(^\text{17}\) $V_\pi(z)$ also corresponds to the equilibrium price with common information and risk-neutral traders.
This risk-neutral, normal model allows us to isolate non-linearities in the dividend function $\pi(\cdot)$, or equivalently, asymmetries in the distribution of dividend risks, while preserving the normality of updating from private signals and market price. The market treats the information contained in $z$ as if it had a precision $\beta + \beta/\sigma^2_u$, when the true precision of $z$ is only $\beta/\sigma^2_u$.

### 3.2 Unconditional prices, dividends, and wedge

We now analyze how noisy information aggregation and payoff asymmetries impact average prices and returns in the risk-neutral, normal model. Our next lemma derives a closed-form solution for the unconditional wedge between expected price and expected dividend.

**Lemma 1 (Unconditional wedge):** Define $\gamma_P \equiv \frac{\beta + \beta/\sigma^2_u}{1/\sigma^2_\theta + \beta + \beta/\sigma^2_u}$ and $\gamma_V \equiv \frac{\beta/\sigma^2_u}{1/\sigma^2_\theta + \beta/\sigma^2_u}$.

Then the difference $\Delta_\pi$ between expected price and expected dividend is

$$\Delta_\pi \equiv \mathbb{E}(P_\pi(z)) - \mathbb{E}(V_\pi(z)) = \int_0^\infty \left( \pi'(\theta) - \pi'(-\theta) \right) \left( \Phi \left( \frac{\theta}{\sigma_\theta} \right) - \Phi \left( \frac{\theta}{\sigma_P} \right) \right) d\theta. \quad (5)$$

This characterization shows how the difference between expected price and expected dividend depends on both the shape of the payoff function, $\pi(\theta)$, and the parameters of the informational environment. By taking ex ante expectations over $z$, the expected price and dividend $\mathbb{E}(P_\pi(z))$ and $\mathbb{E}(V_\pi(z))$ are determined as expectations of $\pi(\theta)$, with respect to, respectively, the market-implied and objective priors over $\theta$. Both priors are normal and centered at 0, but characterized respectively by variances $\sigma^2_P$ and $\sigma^2_\theta$. The ratio $\sigma_P/\sigma_\theta > 1$ summarizes the importance of informational frictions in the market. The market-implied prior thus amounts to a mean-preserving spread relative to the objective prior distribution over $\theta$.

The parameter $\gamma_P$ represents the weight that the market price attaches to $z$, while $\gamma_V$ represents the weight on $z$ in expected dividends. Since the market attaches larger weight to the realization of $z$ ($\gamma_P > \gamma_V$), conditional market expectations of dividends are more volatile relative to the information aggregated in the market. This in turn raises the volatility, or equivalently, the ex ante uncertainty about the asset price, which translates into a higher market-implied uncertainty about fundamentals (i.e., $\sigma_P > \sigma_\theta$). In the formula for expected dividends, the posterior of $\theta$ conditional on $z$ is normal with mean $\gamma_V z$ and variance $(1 - \gamma_V) \sigma^2_\theta$, while the prior of $z$ is normal with mean 0 and variance $\sigma^2_\theta/\gamma_V$; compounding these two distributions yields (by the law of iterated
expectations) the objective prior distribution of \( \theta \), with a prior mean 0 and variance \( \sigma^2_\theta \). The market-implied posterior of \( \theta \) conditional on \( z \) is normal with mean \( \gamma_P z \) and variance \((1 - \gamma_P) \sigma^2_\theta \), and is characterized by a strictly larger responsiveness to \( z \) and a lower posterior variance \((\gamma_P > \gamma_V)\). Compounding this posterior with the prior of \( z \) yields a market-implied prior with mean 0 and variance \( \sigma^2_P \), strictly higher than \( \sigma^2_\theta \) because the increased responsiveness to \( z \) increases ex ante uncertainty about the market’s posterior expectation.

Our next definition provides a partial order on payoff functions that we use for the comparative statics of the expression derived in Lemma 1 with respect to the shape of the dividend function and the informational parameters.

**Definition 1 (Cash flow risks):**

(i) A dividend function \( \pi \) has symmetric risks if \( \pi'(\theta) = \pi'(-\theta) \) for all \( \theta > 0 \).

(ii) A dividend function \( \pi \) is dominated by upside risks if \( \pi'(\theta) \geq \pi'(-\theta) \), and dominated by downside risks if \( \pi'(\theta) \leq \pi'(-\theta) \), for all \( \theta > 0 \).

(iii) A dividend function \( \pi_1 \) has more upside (less downside) risk than \( \pi_2 \) if \( \pi'_1(\theta) - \pi'_1(-\theta) \geq \pi'_2(\theta) - \pi'_2(-\theta) \) for all \( \theta > 0 \).

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 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.

The next Theorem describes the sign of the unconditional wedge and its comparative statics. The results follow directly from this partial order, and the characterization in Lemma 1.

**Theorem 2 (Average price and dividend value):**

(i) **Sign:** If \( \pi \) has symmetric risk, then \( \Delta_\pi = 0 \). If \( \pi \) is dominated by upside risk, then \( \Delta_\pi \geq 0 \). If \( \pi \) is dominated by downside risk, then \( \Delta_\pi \leq 0 \).

(ii) **Comparative Statics w.r.t. \( \sigma^2_P \):** If \( \pi \) is dominated by upside or downside risk, then \( |\Delta_\pi| \) is increasing in \( \sigma_P \). Moreover, \( \lim_{\sigma_P \to \sigma_0} |\Delta_\pi| = 0 \), and \( \lim_{\sigma_P \to \infty} |\Delta_\pi| = \infty \), whenever there exists \( \varepsilon > 0 \), such that \( |\pi'(\theta) - \pi'(-\theta)| > \varepsilon \) for all \( \theta \geq 1/\varepsilon \).

(iii) **Comparative Statics w.r.t. \( \pi \):** If \( \pi_1 \) has more upside risk than \( \pi_2 \), then \( \Delta_{\pi_1} \geq \Delta_{\pi_2} \).

(iv) **Increasing differences:** If \( \pi_1 \) has more upside risk than \( \pi_2 \), then \( \Delta_{\pi_1} - \Delta_{\pi_2} \) is increasing in \( \sigma_P \).
Theorem 2 shows how the shape of the dividend function and the informational parameters combine to determine the difference between expected price and expected dividend in the risk-neutral, normal model. Part (i) shows that unconditional price premia or discounts arise as a combination of two elements: upside or downside risks in the dividend profile π, and the information aggregation friction (σ_P > σ_θ). Mathematically, the result follows directly from our interpretation of the market-implied prior distribution as a symmetric, mean-preserving spread of the true underlying fundamental distribution.

To gain some intuition for this result, recall from the previous section that the equilibrium price attaches too much weight to \( z = \theta - 1/\sqrt{\beta} \cdot u \), or in other words, the price response exceeds the response of dividend expectations to shocks to fundamentals and noise trading. Consider for example an increase in supply \( u \). In equilibrium, more traders purchase the asset, but with dispersed information, this can be accomplished only be a drop in prices that induces less optimistic traders to enter the market. This in turn implies that the difference between price and expected dividends, \( P(z) - V(z) \), is an increasing function of \( z \). Dispersed information thus induces a negative co-movement between supply \( u \) and the wedge \( P(z) - V(z) \), and a positive co-movement between fundamentals \( \theta \) (or equivalently dividends \( \pi(\theta) \)) and the wedge.

What’s more, the response of the wedge to these shocks depends on the shape of the dividend distribution: if fundamentals have a stronger impact on dividends in one direction, then the marginal trader’s expectations will respond more strongly and the co-movement of the wedge with the underlying shocks will be stronger in that direction. When taking averages with respect to the underlying fundamental and supply shocks, the asymmetry in responses then generates the bias in the average price.

A simple way to see this is to apply a first-order Taylor expansion of \( \pi(\cdot) \) around \( \theta_0 = \mathbb{E}(\theta|z) \) to approximate \( P(z) - V(z) \) by

\[
P(z) - V(z) \approx \pi'(\mathbb{E}(\theta|z)) (\mathbb{E}(\theta|x = z, z) - \mathbb{E}(\theta|z)) = \pi'(\mathbb{E}(\theta|z)) (\gamma_P - \gamma_V) z. \tag{6}
\]

Thus, the wedge is approximated by the product of the sensitivity of dividends to expected fundamentals \( \pi'(\mathbb{E}(\theta|z)) \), and the gap between the market implied and objective fundamental expectations. This gap is an increasing function of \( z \). What’s more, in the risk-neutral model this gap is linear in \( z \) and has ex ante expectation of 0, so the unconditional wedge is approximated by \( \Delta_{\pi} \approx (\gamma_P - \gamma_V) \text{cov}(\pi'(\mathbb{E}(\theta|z)), z) \). The sensitivity of dividends to expected fundamentals is increasing or decreasing in \( z \), depending on whether \( \pi(\cdot) \) is convex or concave, and therefore \( \Delta_{\pi} > 0 \)

---

18 A second-order Taylor expansion would yield additional terms that only reinforce the co-movement.
when $\pi(\cdot)$ is convex, and $\Delta_\pi < 0$ when $\pi(\cdot)$ is concave. The characterization of the wedge in terms of upside and downside risks in Theorem 2 then generalizes this observation.

This decomposition can be extended beyond the risk-neutral model. Applying the same approximation to the general sufficient statistic characterization of Theorem 1, and considering the case with an exposure function of $S(\cdot) = 0$, we obtain that the same approximation of the conditional wedge. The unconditional wedge is then approximated by

$$
\Delta_\pi \approx \text{cov}\left( \pi' (E(\theta|z)) , E(\theta|x = z, z) - E(\theta|z) \right) + E\left( \pi'(E(\theta|z)) \right) E\left( E(\theta|x = z, z) - E(\theta|z) \right) - E(\theta) .
$$

(7)

Thus, in the general model, the unconditional wedge is decomposed into two terms. The first term captures the effects of asymmetries through the covariance of $\pi'(E(\theta|z))$ with $E(\theta|x = z, z) - E(\theta|z)$, just like in the risk-neutral model. The second term results from the fact that the average market-implied expectation may differ from the prior i.e. $E(E(\theta|x = z, z)) \neq E(\theta)$. This ex ante difference in expectations could emerge for example from a risk premium, when the security is on average in positive net supply, or equivalently, the marginal trader pricing the security tends to be on average less optimistic than the average trader in the market.\(^{19}\) We’ll defer a full discussion of the relation between these two terms to the next section, which analyzes information wedges in models with risk aversion.

Parts (ii), (iii), and (iv) of Theorem 2 complement the first result with specific predictions on how the magnitude of the unconditional wedge depends on information frictions and cash flow characteristics. Part (ii) shows that the unconditional wedge increases in absolute value as we increase information aggregation frictions (higher $\sigma_P$). Moreover, the expected price wedge disappears when information frictions disappear ($\sigma_P \rightarrow \sigma_\theta$), or becomes arbitrarily large as frictions become more important, as long as the payoff asymmetry does not disappear in the tails. Part (iii) shows that an asset with more upside or less downside risk trades at a higher expected price. Simply put, the mean-preserving spread becomes more valuable when the payoff function shifts towards more upside risk. Part (iv) shows that the unconditional wedge has increasing differences between the dominance of upside risk and the level of market noise. This implies that the effects of market noise and asymmetry in dividend risk on the magnitude of the wedge are mutually reinforcing.

We conclude by describing how $\sigma_P/\sigma_\theta$ depends on the informational parameters $\beta$ and $\sigma_u$. First notice that $\sigma_P/\sigma_\theta$ is increasing in $\gamma_P$ and decreasing in $\gamma_V$, and the information friction is therefore

\(^{19}\)Miller (1977) argues that securities can be overpriced when traders have heterogeneous beliefs and there are limits to short-selling. The underlying mechanism in our setting is quite different since over-pricing results from the interaction between dispersed information and asymmetric cash-flow risks, and as this discussion shows our result does not require short-sales constraints.
largest when $\gamma_P$ is close to 1 and $\gamma_V$ close to 0, resulting in a large discrepancy between market-implied and objective posterior beliefs. This corresponds to a situation where the market signal $z$ is much noisier than private signals $x$, and mainly driven by noise trading shocks $u$. These noise trading shocks however must be absorbed by prices, resulting in large movements of the threshold private signal required to clear the market, and large shifts in the corresponding market-implied expectation of dividends, if private signals are very informative.

Now, $\gamma_P$ and $\gamma_V$ are both decreasing in $\sigma^2_u$ and increasing in $\beta$, which suggests that the overall impact of noise trading and private information on $\sigma_P/\sigma_\theta$ is ambiguous. For the precision of private signals ($\beta$), this is indeed the case. For very small $\beta$, traders disregard private information and beliefs are determined by the common prior, which reduces the ratios $\gamma_P/\gamma_V$ and $\sigma_P/\sigma_\theta$ towards 1. As $\beta$ increases, beliefs become more heterogeneous, raising $\gamma_P/\gamma_V$ and $\sigma_P/\sigma_\theta$. At the other extreme when $\beta \to \infty$, the market signal becomes arbitrarily precise as well, and the ratios $\gamma_P/\gamma_V$ and $\sigma_P/\sigma_\theta$ converge again to 1. The information frictions are therefore largest for intermediate levels of signal precision. In both of these limits, information frictions disappear because investor heterogeneity also vanishes, and informed traders are able to perfectly absorb noise trader shocks without impacting prices (in one case this is because in the limit everyone is perfectly informed through the price, in the other it is because no investor has any superior information). For $\sigma_u$ however, one can show that $\sigma_P/\sigma_\theta$ is monotonically increasing in $\sigma_u$. More noise trading thus unambiguously worsens the information aggregation friction.

4 Risk Aversion

In this section, we study the interaction of noisy information aggregation with risk aversion. We begin with two examples which we solve in closed form. First, we analyze the well-known model with CARA preferences and normally distributed dividends. This allows us to connect our analysis to the existing literature, but remains of limited interest for the discussion of ex ante returns because of the symmetry assumptions built into the model. In the second example, dividends have a binary distribution, which allows us to incorporate payoff asymmetries and extend the insights of Theorem 2 to a model with risk aversion.

We also use these examples to compare our model to two settings with common information and stochastic supply. First, in the no-information case, we remove the assumption that traders obtain private signals. At this benchmark, prices no longer contain any information about dividends, which increases traders’ uncertainty and affects the required compensation for risk. At the second
benchmark, which we term the public information case, we consider a common information economy in which all traders have access to a common signal $z$ which has the same distribution conditional on $\theta$ as in the dispersed information economy (hence the two economies are equivalent in terms of public information).

We complete the section with a numerical example that extends our findings to a noisy information aggregation model with CRRA preferences where closed-form solutions are not available.

4.1 The CARA-normal model

The canonical CARA-normal model with dispersed information is characterized by the following assumptions: (i) the dividend is normally distributed, $\pi(\theta) = \theta$, with $\theta \sim N(0, \sigma^2_\theta)$; (ii) the supply of shares is normally distributed, $u \sim N(\mu, \sigma^2_u)$; (iii) traders have CARA preferences over terminal wealth, $U(w) = -\exp(-\gamma w)$, (iv) traders do not face limits on their portfolio holdings, and (v) each trader observes a private signal that is normally distributed according to $x_i \sim N(\theta, \beta^{-1})$.

The demand of traders with CARA preferences and posteriors over dividends that are normal with mean $\mu$ and variance $\Sigma$ is given by $d(\mu, \Sigma; P) = (\mu - P)/\gamma \Sigma$. In the dispersed information economy, we conjecture and verify that the equilibrium is characterized by a sufficient statistic $z(P)$ which is distributed according to $z(P) \sim N(\theta, \tau^{-1}_P)$, where $\tau_P$ denotes the endogenous informativeness of the price signal. With this conjecture, the traders’ expectation of $\theta$ given $x$ and $z$ is $\mu(x, z) = (\beta x + \tau_P z)/(\sigma^{-2}_\theta + \beta + \tau_P)$, and the posterior variance is $\Sigma = 1/(\sigma^{-2}_\theta + \beta + \tau_P)$. Aggregating demand across traders and imposing market-clearing gives

$$P = \frac{\beta \theta + \tau_P z}{\sigma^{-2}_\theta + \beta + \tau_P} - \frac{\gamma}{\sigma^{-2}_\theta + \beta + \tau_P} u.$$ 

Hence, we can define the sufficient statistic $z$ as $z(P) = \theta - \gamma/\beta \cdot (u - \bar{u})$, which confirms the conjecture with $\tau_P = (\beta/\gamma)^2 \cdot \sigma^{-2}_u$. The equilibrium price is thus represented as

$$P(z) = \frac{\beta + \tau_P}{\sigma^{-2}_\theta + \beta + \tau_P} z - \frac{\gamma}{\sigma^{-2}_\theta + \beta + \tau_P} \bar{u},$$

and the expected price is $\mathbb{E}(P(z)) = -\gamma/(\sigma^{-2}_\theta + \beta + \tau_P) \cdot \bar{u}$.

In a common information economy with dividend expectation $\mu$ and supply realization $u$, the market-clearing condition implies that $\gamma \Sigma u = \mu - P$, or equivalently $P = \mu - \gamma \Sigma u$. The asset price thus decomposes unambiguously into a dividend expectation $\mu$, and a risk premium $-\gamma \Sigma u$. At the no information benchmark, $\mu = 0$ and $\Sigma = \sigma^2_\theta$, so the no information price is $V^{NI}(u) = -\gamma \sigma^2_\theta u$. 


and its expectation is $E(V^{NI}(u)) = -\gamma \sigma^2 \bar{u}$. At the public information benchmark with signal $z \sim \mathcal{N}(\theta, \tau_P^{-1})$, we have

$$
\mu = E(\theta | z) = \tau_P / (\sigma^2 + \tau_P) \cdot z \quad \text{and} \quad \Sigma = (\sigma^2 + \tau_P)^{-1},
$$

so the price is

$$
V^{PI}(z, u) = z \cdot \tau_P / (\sigma^2 + \tau_P) - u \cdot \gamma / (\sigma^2 + \tau_P), \quad \text{and} \quad E(V^{PI}(z, u)) = -\gamma \bar{u} / (\sigma^2 + \tau_P).
$$

Prices and dividend expectations respond to the signal $z$ only in proportion to its information content $\tau_P / (\sigma^2 + \tau_P)$, while fluctuations in supply only affect the risk adjustment.

The equilibrium price with dispersed information in equation (8) thus differs from its common information counterparts by (i) responding more strongly to the market signal $z$, and (ii) adjusting the average risk premium to the level consistent with the informed traders’ posterior uncertainty.\(^{20}\) This exactly mirrors our results from sections 2 and 3. Consider either an increase in $\theta$, or a decrease in $u$, both of which result in a higher price. With dispersed information, an increase in $\theta$ raises the expectation of traders through the distribution of private signals, while a decrease in $u$ reduces the informed traders’ equilibrium exposure and therefore the risk premium. These two effects are captured by the weight $\beta / (\sigma^2 + \beta + \tau_P)$ in equation (8), and correspond to the market-clearing channel we described in section 2. In addition, traders update their expectations about dividends from the observation of $z$, assigning it an additional weight of $\tau_P / (\sigma^2 + \beta + \tau_P)$, the informational channel.\(^{21}\) From an ex ante perspective, however, the differential responses to supply shocks and signals cancel out. In other words, because of the linearity built into the CARA-normal model, expected prices do not feature the effects of asymmetric payoffs described in section 3, and depend only on the average risk premium.

### 4.2 The CARA-binary model

We now analyze a model with CARA preferences and binary dividends to derive a closed-form characterization in an environment with both risk aversion and payoff asymmetries. Specifically, we assume dividends can take only two values: $\pi \in \{0, 1\}$, with ex-ante probability $Pr(\pi = 1) = \lambda > 0$. The parameter $\lambda$ measures the degree of upside versus downside risk: if $\lambda > 1/2$, the security is a downside risk; if $\lambda = 1/2$, the risk is symmetric; if $\lambda < 1/2$, the security is an upside risk. The traders’ private signals are normally distributed and centered at $\pi$, $x_i \sim \mathcal{N}(\pi, 1/\beta)$. All other elements are kept the same as in the CARA-normal model.

\(^{20}\)As discussed in section 2, the decomposition of the price into dividend expectations and risk premium depends on the choice of an exposure function $\bar{S}(\cdot)$ for the dispersed information economy. Here, we are basing the discussion on a choice of $\bar{S}(\cdot) = \bar{\pi}$, but the same discussion about average risk premia apply for any exposure function $\bar{S}(\cdot)$, for which average exposure $E(\bar{S}(\cdot))$ equals $\bar{\pi}$.

\(^{21}\)Remarkably, the only prior discussion of this observation that we are aware of is by Vives (2008).
With binary dividends, it is easy to check from first-order conditions that a trader who assigns probability \( \mu \in (0, 1) \) to \( \pi = 1 \) demands 
\[
d(\mu, P) = \frac{1}{\gamma} \left( \log \left( \frac{\mu}{1-\mu} \right) - \log \left( \frac{P}{1-P} \right) \right).
\]
As before, we conjecture and verify that the equilibrium is characterized by a sufficient statistic \( z(P) \) which is distributed according to \( z(P) \sim N(\pi, \tau_P) \), where the precision of the price signal \( \tau_P \) remains to be determined. With this conjecture, the posterior \( \mu(x, z) \) after observing two independent, normally distributed signals \( x \) and \( z \) satisfies 
\[
\log \left( \frac{\mu(x, z)}{1-\mu(x, z)} \right) = \log \left( \frac{\lambda}{1-\lambda} \right) + \beta \left( x - \frac{1}{2} \right) + \tau_P \left( z - \frac{1}{2} \right).
\]
Hence, with normally distributed signals, the log-odds ratio, and therefore the asset demand, is linear in the signal realizations. Aggregating across traders, we obtain that the aggregate demand
\[
D(\pi, P) = \frac{1}{\gamma} \left( \log \left( \frac{\lambda}{1-\lambda} \right) - \log \left( \frac{P}{1-P} \right) + \tau_P \left( z - \frac{1}{2} \right) \right) + \frac{\beta}{\gamma} \left( \pi - \frac{1}{2} \right),
\]
for \( \pi \in \{0, 1\} \). Therefore, the market-clearing price function \( P(z) \) satisfies
\[
\log \left( \frac{P(z)}{1-P(z)} \right) = \log \left( \frac{\lambda}{1-\lambda} \right) - \gamma \bar{u} + (\tau_P + \beta) \left( z - \frac{1}{2} \right),
\]
or 
\[
P(z) = \lambda \left\{ \lambda + (1-\lambda) e^{\gamma \bar{u} - (\beta+\tau_P)(z-\frac{1}{2})} \right\}^{-1},
\]
where \( z = \pi - \gamma/\beta \cdot (u - \bar{u}) \) and \( \tau_P = (\beta/\gamma)^2 \cdot \sigma_u^{-2} \), confirming our initial conjecture. The price (or equivalently, the log-odds ratio implied by the price) attributes a weight \( \tau_P + \beta \) to the market signal \( z \) due to the market clearing effect, while its true information content is only \( \tau_P \). In addition, the log-odds ratio implied by the price includes a risk adjustment \(-\gamma \bar{u}\) to compensate traders for their expected exposure.

In the homogeneous information market, we invert the demand schedule for given supply realization \( u \) and posterior \( \mu \) to obtain the equilibrium price; \( V(\mu, u) = \mu \{\mu + (1-\mu) e^{\gamma u}\}^{-1} \). In the no information case we have \( \mu = \lambda \), so \( V^{NI}(u) = \lambda \{\lambda + (1-\lambda) e^{\gamma u}\}^{-1} \), while in the public information case, \( V^{PI}(u, z) = \lambda \{\lambda + (1-\lambda) e^{\gamma u - \tau_P(z-1/2)}\}^{-1} \). The weight assigned to \( z \) thus corresponds to its true information content, \( \tau_P \).

Our next proposition compares expected prices and dividends in the heterogeneous and common information models, in the case in which the asset is on average in zero net supply: \( \bar{u} = 0 \).

**Proposition 2 (Wedge with binary cash-flows I):** Suppose that \( \bar{u} = 0 \). There exist positive numbers \( \Delta^P \), \( \Delta^{NI} \), and \( \Delta^{PI} \) with the following properties:
1. The average price in the private information model and the common information benchmarks take the form 

\[ \mathbb{E}(P(z)) = \lambda + (1 - 2\lambda) \Delta^P, \quad \mathbb{E}(V^{NI}(u)) = \lambda + (1 - 2\lambda) \Delta^{NI}, \quad \text{and} \quad \mathbb{E}(V^{PI}(u, z)) = \lambda + (1 - 2\lambda) \Delta^{PI}. \]

2. It is always the case that \( \Delta^{NI} > \Delta^{PI} \), but \( \Delta^P \) may be larger or smaller than \( \Delta^{PI} \) and \( \Delta^{NI} \), depending on parameters.

3. Holding \( \tau_P \) constant, and for small \( \gamma \sigma_u \), we have

\[ \Delta^{NI} \approx \frac{1}{2} \lambda (1 - \lambda) \gamma^2 \sigma_u^2, \quad \Delta^{PI} \approx e^{-\frac{1}{2} \gamma^2 \sigma_u^2 \tau^2} \frac{1}{2} \lambda (1 - \lambda) \gamma^2 \sigma_u^2, \quad \Delta^P \approx e^{-\frac{1}{2} \gamma^2 \sigma_u^2 \tau^2} \frac{1}{2} \lambda (1 - \lambda) \gamma \sigma_u \sqrt{\tau}. \]

This proposition shows that with dispersed information, the expected price exceeds the asset’s expected value, whenever \( \lambda < 1/2 \), i.e. whenever the asset is characterized by upside risk (part 1). The same result also arises at the common information benchmarks with stochastic supply. Parts 2-3 of the proposition compare the extent to which payoff asymmetry determines prices across the different models, with the numbers \( \Delta^P, \Delta^{NI}, \) and \( \Delta^{PI} \) summarizing the impact of upside or downside risk on prices. The central observation from this comparison is that when supply shocks are small, payoff asymmetry has a much larger effect on prices with dispersed information than at the common information benchmarks.

The impact of payoff asymmetry on average prices follows from a similar logic as in section 3: in all three models, there is a negative co-movement between the stochastic asset supply, and the equilibrium price (or more precisely, the wedge between the price and the conditional dividend expectation). In the common information model with or without public signals, the price is uninformative, and this negative relation is entirely due to the shifts in risk premia that are required to induce traders to absorb the supply fluctuations: If \( u > 0 \), in equilibrium traders must take a long position to clear the market, and the price is lower than expected dividend to compensate for risk. For \( u < 0 \), traders must short the asset, and the price is higher than the expected dividend. This negative co-movement of the wedge between price and expected dividend with supply then interacts with the asymmetry in payoffs in much the same way as in the risk-neutral model: For upside risks (\( \lambda < 1/2 \)), the potential losses are larger for short than for long positions, so the compensation for risk must be larger in absolute terms for the short positions. This results in an average price that exceeds that unconditional dividend expectation \( \lambda \). By the reverse argument, with downside risk (\( \lambda > 1/2 \)) the expected price is lower than \( \lambda \), that is, a positive average risk premium.

At the no information or common information benchmarks, these risk premia scale with \( \gamma u \) conditional on the realization of supply \( u \). The expected risk premium then scales with \( \gamma^2 \sigma_u^2 \) (when \( \lambda \neq 1/2 \)). Adding a noisy public signal to the no information benchmark reduces risk premia
and brings the price closer in line with expected dividends, but otherwise doesn’t alter the impact of payoff asymmetry. This implies that \( \Delta NI > \Delta PI \), i.e. the gap between prices and expected dividends is unambiguously lower in the public information economy.

With dispersed information, the asset price fluctuates with supply shocks and fundamentals through two channels. First, just as in the common information environment, traders must be compensated for the risk associated with exposure to asymmetric payoffs. Second, these shifts are reinforced by the additional wedge between market-implied and objective dividend expectations: a lower supply of the asset, or a higher fundamental realization increases the expectation of the marginal trader whose indifference is required to clear the market, and does so in excess of the impact of these shocks on objective dividend expectations. In other words, dispersed information generates an amplification of the impact of noisy asset supply on the wedge between price and expected dividend.

The impact of dispersed information on average prices, relative to the common and no information benchmarks, is then ambiguous: On the one hand, the access to better information reduces the risk associated with the asset; on the other the dispersed information amplifies the gap between price and expected dividend. The risk premium continues to scale in expectation with \( \gamma^2 \sigma_u^2 \), but for given price informativeness \( \tau_p \), the impact of supply and fundamentals on this expectations wedge scales with \( \gamma \sigma_u \), and therefore, when \( \gamma \sigma_u \) is small, \( \Delta P \) is of the order \( \gamma \sigma_u \). Hence, when \( \gamma \sigma_u \) is small, the wedge between market-implied and objective expectations becomes far more important in accounting for the impact of payoff asymmetry than risk premia. Moreover this amplification can be arbitrarily large, since \( \Delta NI \) and \( \Delta PI \) are of the order \( \gamma^2 \sigma_u^2 \), while \( \Delta P \) is of the order \( \gamma \sigma_u \). This observation however does not hold universally: if the market signal is sufficiently precise or the supply shocks become very large, then the impact of information on risk premia dominates, and eventually \( \Delta P \) is strictly less than \( \Delta NI \) and \( \Delta PI \).

We conclude this part with the case where \( \bar{u} \neq 0 \), i.e. supply is not centered at 0. We can replace the prior \( \lambda \) in the equilibrium characterization by its risk-adjusted counterpart \( \hat{\lambda} = \lambda / (\lambda + (1 - \lambda) e^{\gamma \bar{u}}) \). With this adjustment, we can separate the gap between average price and average dividend into a component that is due to noisy information aggregation, and an average risk premium. The information aggregation component takes the same form as in proposition 2.

**Proposition 3 (Wedge with binary cash-flows II):** Suppose that \( \bar{u} \neq 0 \). There exist positive numbers, \( \Delta P, \Delta NI, \Delta PI, R^P, R^{PI} \) and \( R^P \) with the following properties

1. The average price in the private information model and the common information benchmarks take the form \( \mathbb{E}(V^{NI}(u)) = \hat{\lambda} + (1 - 2\hat{\lambda}) \Delta NI \), \( \mathbb{E}(P(z)) = \lambda + (1 - 2\lambda) \Delta P - (\lambda - \hat{\lambda}) R^P \), and
\[ E(V^{PI}(u,z)) = \lambda + \left(1 - 2\hat{\lambda}\right) \Delta^{PI} - \left(\lambda - \hat{\lambda}\right) R^{PI}. \]

2. The numbers \( \Delta^P, \Delta^{NI}, \Delta^{PI} \) take the same form as in Prop. 2, but with \( \hat{\lambda} \) in place of \( \lambda \).

3. The numbers \( R^{PI} \) and \( R^P \) satisfy \( 1 > R^{PI} > R^P \) and \( \lim_{\gamma \sigma_u \to 0} R^P = \lim_{\gamma \sigma_u \to 0} R^{PI} = R \) for some \( R < 1 \).

Hence, in all three models, the average price can be decomposed into an average risk premium component and an asymmetry component, mirroring the decomposition that we had obtained from the Taylor expansion in section 3 (equation (7) for the risk-averse case). The asymmetry component retains all the properties that were highlighted in proposition 2, including its amplification in the model with noisy information aggregation. The two effects reinforce each other in the case of downside risk \((\hat{\lambda} < 1/2)\), but partially offset each other in the case of upside risk \((\hat{\lambda} > 1/2)\). Which one ends up dominating then depends on parameters such as the magnitude of average supply or the informativeness of the private information and prices. The risk premium scales with \( \lambda - \hat{\lambda} = \hat{\lambda}(1 - \lambda)(e^{\gamma \pi} - 1) \approx \gamma \pi \hat{\lambda}(1 - \lambda) \), is bounded away from zero, and is rankable across models: highest in the model with no information, and lowest in the model with noisy information aggregation, in which traders face the least uncertainty ex post about dividends. The asymmetry term instead scales with \( \gamma^2 \) in the common information benchmarks, and is thus a second-order concern relative to average risk premia. In the model with dispersed information, both the risk component and the asymmetry component scale with \( \gamma \), which illustrates that dispersed information is key for payoff asymmetry to have effects on average prices that are of similar magnitude as standard risk premia.

4.3 Numerical solutions for CRRA preferences

In the model with general preferences and distributions, closed form solutions are no longer readily available. To solve such models, we propose a procedure which solves for the informational content of prices.\(^{22}\) We start by conjecturing a distribution of prices conditional on a given value of \( \theta \): \( F(P'|\theta) \equiv Pr(P \leq P'|\theta) \), for each \( \theta \) in the grid considered, along with conditional density \( f(P'|\theta) \). From \( f(P'|\theta) \), we calculate the posterior distribution of \( \theta \) for each trader using Bayes rule: \( Pr(\theta|x_i, P) = f(P|\theta) \cdot Pr(\theta|x_i) / \sum_\theta f(P|\theta') \cdot Pr(\theta'|x_i) \), where \( Pr(\theta|x_i) \) corre-

\(^{22}\)Bernardo and Judd (2000) and Peress (2004) numerically solve a RE equilibrium under asymmetric information and CRRA preferences by “guessing” price and demand functions using hermite polynomials under the structural moment conditions implied by demand optimality and market clearing. The central difference of our approach is that here we explicitly solve for the price likelihood function, which allows a clean characterization of the informational content of prices, for different price realizations. To our knowledge, this methodology is new in the REE literature.
sponds to the posterior conditional on observing $x_i$. Using the posterior distribution, we find the demand function $d(x, P)$ that maximizes $E[u(w)|x, P]$, and then determine aggregate demand $D(\theta, P)$ numerically. From this, we re-compute the conditional distribution of prices $F(P|\theta)$ as $F(P|\theta) = Pr(u \geq D(\theta, P)) = 1 - G(D(\theta, P))$, making use of the market-clearing condition. We then iterate until the resulting $Pr(P|\theta')$ converges. Finally, we calculate the price function $P(\theta, u)$ by inverting the function $D(\theta, P) = u$ to obtain $P = P(\theta, u)$.

In the case with CRRA preferences and binary asset payoffs, this task is simplified by the fact that the posterior odds ratio and asset demand can be written in closed form. Suppose that all primitives are as in the model in section 4.2, except that preferences are of the CRRA form: $U(w) = w^{1-\gamma}/(1-\gamma)$. Here, $w = w_0 + d(\pi - P)$ is the terminal wealth of a trader purchasing $d$ units of the asset, and $w_0$ is an initial endowment, identical across traders. From Bayes rule, the posterior belief of a trader observing his private signal and the price is given by

$$ \frac{\mu_i}{1-\mu_i} = \frac{Pr(\pi = 1|x_i, P)}{Pr(\pi = 1|x_i, P')} = \frac{\lambda}{1-\lambda} \cdot \frac{Pr(P|\pi = 1)}{Pr(P|\pi = 0)} \cdot e^{\beta(x_i - 0.5)}. $$

For a trader with posterior $\mu_i$, the optimal asset demand solves $E[w^{-\gamma} \cdot (\pi - P) | x_i, P] = 0$, or

$$ d_i = w_0 \frac{\left( \frac{\mu_i}{1-\mu_i} \right)^{\frac{1}{\gamma}} - \left( \frac{P}{1-P} \right)^{\frac{1}{\gamma}}}{(1-P) \left( \frac{P}{1-P} \right)^{\frac{1}{\gamma}} + P \left( \frac{\mu_i}{1-\mu_i} \right)^{\frac{1}{\gamma}}}. \quad (9) $$

Aggregating across traders gives the aggregate demand, $D(\pi, P; w_0)$, and using the market-clearing condition $D(\pi, P; w_0) = u$, we then update on the price likelihood ratio $Pr(P|\pi = 1)/Pr(P|\pi = 0)$. We use these likelihoods as inputs in the expression for $\mu_i$, and iterate until convergence.

We contrast this economy with the common information benchmarks. For the no information economy, the price is obtained from equation (9) using the prior $\lambda$ instead of $\mu_i$, and imposing the market-clearing condition $d(P; w) = u$. This gives $V^{NI}(u; w_0)$. For the common information economy where traders observe an exogenous public signal $z$, the common posterior belief corresponds to $\mu = Pr(\pi = 1|z)$. We then solve for demand using (9) replacing $\mu_i$ with $\mu$, and impose the market-clearing condition $d(z, P, w_0) = u$ to find the equilibrium price function $V^{PI}(z, u; w_0)$.

Figure 1 computes numerically the total unconditional wedge – the difference between the unconditional price and $\lambda$, as a fraction of $\lambda$ – for the heterogeneous and common information economies. We set parameters equal to $\sigma_u = \beta = 1$, and $\gamma = 3$. In all panels, the solid line represents

\[\text{To simplify the computation of the common information counterpart, we consider a distribution of the exogenous public signal that is normal and has a precision } \tau_P, \text{ such that traders have on average the same posterior variance about } \pi \text{ than they would by observing only the price in the heterogeneous information economy.}\]
the unconditional wedge in the model with dispersed information, while the dashed line represents its common information counterpart. Consistent with theorem 2, dispersed information results in higher prices relative to the common information benchmark, when the asset is characterized by upside risk, and in lower prices when it is characterized by downside risk.

Panel a) then studies the comparative statics with respect to initial wealth, \( w_0 \), for an average asset supply of \( \pi = 0 \). When the asset is dominated by upside risk, the unconditional wedge is positive, but as wealth increases, the wedge becomes smaller in absolute magnitude and unconditional prices converge to \( \lambda \). Dispersed information considerably amplifies this wedge, to the point that when wealth levels are sufficiently high (small local risk aversion), the wedge with dispersed information remains an order of magnitude larger than under common information, consistent with Proposition 2. The same picture emerges for downside risks, but unconditional risk premia and information wedges in this case are negative.

Panel b) presents results for an intermediate level of aggregate risk; \( \pi = 0.5 \). The aggregate exposure \( \pi = 0.5 \) adds a compensation for risk to the price, which under both common and dispersed information results in lower average prices. For upside risks, this leads to a positive average risk premium (i.e., prices below payoffs) under common information, but with dispersed information, the positive information wedge arising with upside risk is large enough to offset the additional risk premium at low levels of wealth. As wealth increases and the informational content of prices improves, the information wedge is reduced and the risk premium component dominates. For downside risks, the negative information wedge now adds to the the risk premium component, inducing a more negative wedge under heterogeneous information than under common information.

Panel c) presents results for an even larger level of aggregate risk, \( \pi = 1 \), with qualitative effects similar as in previous cases. The main difference is that for upside risks, the positive information wedge under heterogeneous information is not sufficient to undo the large positive risk premium, resulting in a negative overall wedge for all wealth levels.

For completeness, panel d) presents the comparative statics with respect to \( \pi \), with a fixed level of wealth, \( w_0 = 10 \). For upside risks, the positive information wedge implies a larger overall wedge under heterogeneous information. Both curves are decreasing in \( \pi \), and eventually become negative, as the risk premium component becomes dominant. For downside risks, the negative information wedge adds to the risk premium component under heterogeneous information, increasing the magnitude of the unconditional wedge.

In summary, the results of numerical simulations for the model with CRRA utility parallel the theoretical results of section 4.2.
Figure 1: Unconditional wedge in the CRRA-binomial model: \( \frac{\mathbb{E}[P] - \lambda}{\lambda} \) (in bps)

- **Upside risk**: \( \lambda = 0.1 \)
- **Downside risk**: \( \lambda = 0.9 \)

### a) \( \overline{\mu} = 0 \)

### b) \( \overline{\mu} = 0.5 \)

### c) \( \overline{\mu} = 1 \)

### d) \( w_0 = 10 \)
5 Applications

In this section, we develop two applications in the context of our baseline risk-neutral, normal model (Section 3). First, we develop a numerical illustration of the effects of skewness on expected returns, which shows how our model is qualitatively consistent with the empirical predictions on return to skewness. Second, we reconsider the Modigliani-Miller Theorem and show how the owner of a cash flow can influence its market value through appropriate security design.

5.1 Numerical illustration, returns and skewness

A sizable empirical literature documents a negative relationship between skewness and asset returns. For example, Conrad, Dittmar and Ghysels (2013 - henceforth CDG) use option prices to recover moments of the risk neutral probability distribution and find that ex ante more negatively (positively) skewed returns are associated with subsequent higher (lower) returns. For different stocks, they infer the distribution of cash flows from option prices using options with maturities of 3 and 12 months. They sort the stocks according to the degree of skewness into high-skewed (top 30 percent, skewness $\approx 0$), medium-skewed, and low-skewed securities (bottom 30 percent, skewness $\approx -2.8$). For each maturity, they compute the return difference between the highest and lowest skewness portfolios, and find that securities with higher skewness earn lower average returns: -79 and -67 basis points per month, for 3- and 12-month maturities. Boyer, Mitton and Vorkink (2010) estimate a cross-sectional model and find that expected idiosyncratic skewness and returns are negatively correlated. Green and Hwang (2012) find that IPOs with high expected skewness earn significantly more negative abnormal returns in the following one to five years.

Existing theories account for these empirical findings through over-investment in securities with high skewness, resulting in over-pricing and negative expected returns. 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.

Our model can account for these empirical findings without relying on specific features of demand or preferences. Specifically, Theorem 2 establishes that securities that are dominated by upside risk, or in other words, positively skewed, are also the ones with higher prices and hence lower returns; our model thus predicts a negative relation between skewness and expected returns.
The following proposition formalizes this connection for log returns.

**Proposition 4 (Skewness and Returns):** In the risk-neutral, normal model, the unconditional expectation of log returns is

$$
\mathbb{E} (\log R_{\pi}(\theta, z)) = \int_0^\infty \left( \frac{\pi'(\theta)}{\pi(\theta)} - \frac{\pi'(-\theta)}{\pi(-\theta)} \right) \left( \Phi \left( \frac{\theta}{\sigma_P} \right) - \Phi \left( \frac{\theta}{\sigma_{\theta}} \right) \right) d\theta + \mathbb{E} (\Gamma(z)),
$$

where

$$
\Gamma(z) = \mathbb{E} (\log R_{\pi}(\theta, z) | x = z, z) \approx -\frac{1}{2} \text{Var} (R_{\pi}(\theta, z) | x = z, z),
$$

and \( \text{Var} (R_{\pi}(\theta, z) | x = z, z) \) is the residual variance of returns conditional on the price.

Thus, the expected log-return decomposes into a term that measures the impact of upside vs. downside risk after re-defining these in terms of percentage rather than absolute variation in returns, and a second term that results from an adjustment due to Jensen’s inequality which vanishes with the informed traders’ residual uncertainty (i.e. as \( \gamma_P \) converges to 1). Therefore, conditional on holding the market-implied variance of returns fixed, the predictions of Theorem 2 directly extend to expected log-returns: upside risks trade at a premium and offer lower expected returns, while downside risks trade at a discount and offer higher returns, and the magnitude of these premia and discounts increases in the degree of information frictions.

We now provide a numerical illustration, which suggests that even a small degree of informational friction could generate quantitatively significant negative returns to skewness. After normalizing (without loss of generality) \( \sigma^2_{\theta} = 1 \), our risk-neutral, normal model is fully specified by the choice of a dividend function \( \pi(\cdot) \) and the informational parameters \( \beta \) and \( \sigma^2_u \). We thus conduct the following thought experiment: after parametrizing the dividend function \( \pi(\cdot) \) so that it matches the volatility and skewness data of the different portfolios in CDG, we vary the informational parameters \( \beta \) and \( \sigma^2_u \) to target different levels of informational frictions \( \sigma^2_P \), and then compare the associated model-implied returns to their empirical counterparts reported by CDG.

Specifically, we consider a class of dividend functions of the form \( \pi(\theta) = e^{k \cdot x(\theta)} \), where \( k > 0 \) is a scaling parameter, \( x(\cdot) \equiv B^{-1}(\Phi(\cdot)) \) and \( B(\cdot) \) is the cdf of a beta distribution with distributional parameters \( A \) and \( B \). Up to a scalar \( k \), log-dividends are thus distributed according to a beta distribution with expectation \( A/(A+B) \). The skewness of log returns is then independent of \( k \) and inversely related to its expectation \( A/(A+B) \), while the volatility of log returns, for given \( A \) and \( B \) scales with \( k \). For given \( A+B \), the mean \( A/(A+B) \) is thus chosen to match the skewness, and \( k \) is then chosen to match the volatility of returns. We chose this parametrization because it allows us to flexibly match virtually arbitrary skewness/volatility configurations with just two
Table 1: Information frictions, skewness, and expected returns

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Skew</th>
<th>Return</th>
<th>Vol</th>
<th>Skew</th>
<th>Return</th>
<th>ImpVol</th>
<th>Return</th>
<th>ImpVol</th>
<th>Return</th>
<th>ImpVol</th>
<th>Return</th>
<th>ImpVol</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data</td>
<td></td>
<td></td>
<td>Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>σ&lt;sub&gt;P&lt;/sub&gt; = 1.03</td>
<td></td>
<td></td>
<td>σ&lt;sub&gt;P&lt;/sub&gt; = 1.06</td>
<td></td>
<td></td>
<td>σ&lt;sub&gt;P&lt;/sub&gt; = 1.09</td>
<td></td>
<td></td>
<td>σ&lt;sub&gt;P&lt;/sub&gt; = 1.12</td>
<td></td>
</tr>
<tr>
<td>3 m.</td>
<td>Low</td>
<td>0.57</td>
<td>31.51</td>
<td>-2.814</td>
<td>0.61</td>
<td>1.049</td>
<td>1.31</td>
<td>1.094</td>
<td>1.98</td>
<td>1.141</td>
<td>2.64</td>
<td>1.189</td>
</tr>
<tr>
<td></td>
<td>Med</td>
<td>0.21</td>
<td>32.26</td>
<td>-0.980</td>
<td>0.28</td>
<td>1.027</td>
<td>0.57</td>
<td>1.052</td>
<td>0.85</td>
<td>1.079</td>
<td>1.13</td>
<td>1.104</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>-0.22</td>
<td>31.14</td>
<td>0.026</td>
<td>-0.09</td>
<td>1.022</td>
<td>-0.18</td>
<td>1.045</td>
<td>-0.27</td>
<td>1.066</td>
<td>-0.37</td>
<td>1.087</td>
</tr>
<tr>
<td></td>
<td>H-L</td>
<td>-0.79</td>
<td>-0.70</td>
<td>-1.49</td>
<td>-2.25</td>
<td>-3.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M-L</td>
<td>-0.37</td>
<td>-0.32</td>
<td>-0.74</td>
<td>-1.13</td>
<td>-1.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 m.</td>
<td>Low</td>
<td>0.55</td>
<td>30.16</td>
<td>-2.743</td>
<td>0.63</td>
<td>1.046</td>
<td>1.26</td>
<td>1.092</td>
<td>1.89</td>
<td>1.138</td>
<td>2.52</td>
<td>1.185</td>
</tr>
<tr>
<td></td>
<td>Med</td>
<td>0.15</td>
<td>31.31</td>
<td>-0.974</td>
<td>0.27</td>
<td>1.026</td>
<td>0.55</td>
<td>1.053</td>
<td>0.83</td>
<td>1.078</td>
<td>1.10</td>
<td>1.104</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>-0.12</td>
<td>30.69</td>
<td>0.019</td>
<td>-0.09</td>
<td>1.023</td>
<td>-0.17</td>
<td>1.045</td>
<td>-0.26</td>
<td>1.066</td>
<td>-0.35</td>
<td>1.087</td>
</tr>
<tr>
<td></td>
<td>H-L</td>
<td>-0.67</td>
<td>-0.71</td>
<td>-1.43</td>
<td>-2.15</td>
<td>-2.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M-L</td>
<td>-0.40</td>
<td>-0.35</td>
<td>-0.71</td>
<td>-1.06</td>
<td>-1.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Finally, we set σ<sup>2</sup>θ = 1 and σ<sup>2</sup>u = 2, and then vary β to target a specific information friction parameter σ<sub>P</sub>/σ<sub>θ</sub>, for values of σ<sub>P</sub>/σ<sub>θ</sub> ranging from 1.03 to 1.12.25

Table 1 presents our numerical results. Its rows represent the portfolios in CDG sorted by skewness constructed from options of 3-month and 12-month to maturity, respectively. The first three columns (“data”) report empirical moments from CDG (Table II), i.e. average returns, excess return of high-to-low or medium-to-low skewness, and the empirical volatility and skewness of these portfolios. The subsequent columns report our model-implied moments for a security that matches the same log skewness and volatility as in the data. For each such security, we report the asset return from the model, as well as the ratio of the market-implied to the actual volatility.

The table shows that even with a small level of frictions, the premium for skewness falls in the range suggested by the data. For example, a value of σ<sub>P</sub>/σ<sub>θ</sub> of 1.06 (meaning that the market-implied standard deviation of θ exceeds the objective standard deviation of θ by only 6 percent) is already sufficient to generate excess returns that are higher than those observed in the data. Given the stylized nature of our exercise we do not want to over-interpret these results - a full-fledged calibration would require taking a more explicit stand on the key informational parameters, such as

24We report results for A + B = 5, but results for other values are similar.
25Proposition 4 states that the impact of skewness on returns depends primarily on the ratio σ<sub>P</sub>/σ<sub>θ</sub>, and that conditional on this ratio, our results are not affected by the particular choice of σ<sub>u</sub> and β.
the precision of private information, $\beta$, and the volatility of supply shocks, $\sigma_u$.\textsuperscript{26} Nevertheless our results suggest that the quantitative impact of skewness on returns in the presence of dispersed information can be quite significant.

We briefly mention other testable implications of the theory, for completeness. First, positively skewed risks earn a negative excess return, while negatively skewed risks (downside risks) earn positive excess returns. To test this prediction, one needs to define a benchmark return against which the positive and negative excess returns of upside and downside risks may be measured (the benchmark return needs to be equal to the return on a symmetric risk). Second, the impact of skewness on returns is larger when information frictions are more pronounced. To test this prediction, one would require empirical proxies or measures of information aggregation frictions, and one would need to interact those with the measures of skewness to forecast returns.

### 5.2 Splitting Cash-flows to influence market value

In complete and perfectly competitive financial markets, the market value of a given cash flow should not depend on how it is allocated to different investors (Modigliani and Miller, 1958). Here we show that with noisy information aggregation, the relationship between skewness or upside vs downside risk and returns emerges as a new element shaping security design incentives. Furthermore, when the investor pools for different claims have different informational characteristics, the seller can influence expected revenues by tailoring the split to the different investor types.

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

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

\textsuperscript{26}In Albagli, Hellwig, and Tsyvinski (2014 a), we argue that our theory can provide an explanation for the credit spread puzzle. In that paper, we also propose a strategy to infer some of these parameters from data linking analyst forecast dispersion to expected bond returns. Our results suggest that values of $\sigma_P/\sigma_\theta$ of 1.2 are not unreasonable. Such values are also consistent with the calibration to 3 months ahead earnings forecasts by Straub and Ulbricht (2015). David, Hopenhayn and Venkatesvaran (2014) calibrate our model to the co-movement of prices and realized returns over 3 year horizons and arrive at parameter values that imply $\sigma_P/\sigma_\theta \approx 1.8$. 

29
That is, each market is affected by a noisy supply shock $u_i$ with market-specific noise parameter $\sigma_{u,i}^2$. The environment is then characterized by the market-characteristics $\beta_i$ and $\sigma_{u,i}^2$, and by the correlation of demand shocks across markets, $\rho$. Traders are active only in their respective market, have risk-neutral preferences, and face position limits as in Section 3. We consider both the possibility that traders observe and condition on prices in the other market (informational linkages), and the possibility that they do not (informational segregation).

The seller is risk-neutral, and hence wishes to maximize the expected revenue net of the dividend, $P_1 + P_2 - \pi(\cdot)$. With both informational linkages and informational segregation, the seller’s net expected revenue can be represented as the sum of the expected price wedges for the two securities, $\Delta_{\pi_1}(\sigma_{P,1}) + \Delta_{\pi_2}(\sigma_{P,2})$, where $\sigma_{P,i}$ denotes the level of informational frictions in market $i$. The two cases only differ in how the information frictions parameters are determined.

Under informational segregation, the analysis of the two markets can be completely separated. The equilibrium characterization from the single-asset model applies separately in each market, and the information frictions parameters $\sigma_{P,i}$ are determined as in lemma 1. With informational linkages, the equilibrium analysis has to be adjusted to incorporate the information contained in price 1 for the traders in market 2, and vice versa, but otherwise proceeds along the same lines. Informed traders in market $i$ buy a security if and only if their private signal exceeds a threshold $\hat{x}_i(\cdot)$, where $\hat{x}_i(\cdot)$ is conditioned on both prices. By market-clearing, $\hat{x}_i(\cdot) = z_i \equiv \theta - 1/\sqrt{\beta_i} \cdot u_i$. Observing $P_i$ is then isomorphic to observing $z_i$, and observing both prices is isomorphic to observing $(z_1, z_2)$. We then characterize posterior beliefs over $\theta$, market prices, and expected dividends, as functions of $(z_1, z_2)$:

\[
\begin{align*}
P_1(z_1, z_2) &= \mathbb{E}(\pi_1(\theta) | x = z_1; z_1, z_2) \quad \text{and} \quad V_1(z_1, z_2) = \mathbb{E}(\pi_1(\theta) | z_1, z_2), \\
P_2(z_1, z_2) &= \mathbb{E}(\pi_2(\theta) | x = z_2; z_1, z_2) \quad \text{and} \quad V_2(z_1, z_2) = \mathbb{E}(\pi_2(\theta) | z_1, z_2).
\end{align*}
\]

The expected wedge $\Delta_{\pi_i}$ is characterized by

\[
\Delta_{\pi_i}(\sigma_{P,i}) = \int_0^\infty (\pi'_i(\theta) - \pi'_i(-\theta)) \left( \Phi \left( \frac{\theta}{\sigma_{\theta}} \right) - \Phi \left( \frac{\theta}{\sigma_{P,i}} \right) \right) d\theta,
\]

where

\[
\sigma_{P,i}^2 = \sigma_{\theta}^2 + (1 + \sigma_{u,i}^2) \cdot \frac{\beta_i}{(\beta_i + V)^2}, \quad \text{with} \quad V = 1/\sigma_{\theta}^2 + \frac{1}{1 - \rho^2} \left( \frac{\beta_1}{\sigma_{u,1}^2} + \frac{\beta_2}{\sigma_{u,2}^2} - 2 \rho \sqrt{\beta_1 \beta_2} \right).
\]

This characterization of $\Delta_{\pi_i}$ is the same as in lemma 1, but the definition of informational frictions parameter $\sigma_{P,i}$ must be adjusted for the change in the underlying information structure. Since our results below take the information frictions parameters as given, they apply identically to the models with informational segregation and informational linkages.
Proposition 5 (Modigliani-Miller):

(i) The cash-flow split does not affect the seller’s expected revenue, if and only if the market characteristics are identical: $\sigma_{P,1} = \sigma_{P,2}$.

(ii) If $\sigma_{P,1} > \sigma_{P,2}$, $\Delta_{\pi_1} (\sigma_{P,1}) + \Delta_{\pi_2} (\sigma_{P,2}) > \Delta_{\pi_1} (\sigma_{P,2}) + \Delta_{\pi_2} (\sigma_{P,1})$, while if $\sigma_{P,1} < \sigma_{P,2}$, $\Delta_{\pi_1} (\sigma_{P,1}) + \Delta_{\pi_2} (\sigma_{P,2}) < \Delta_{\pi_1} (\sigma_{P,2}) + \Delta_{\pi_2} (\sigma_{P,1})$.

For given values of $\sigma_{P}$, the expected wedge is additive across cash flows: $\Delta_{\pi_1} (\sigma_{P}) + \Delta_{\pi_2} (\sigma_{P}) = \Delta_{\pi_1+\pi_2} (\sigma_{P})$, for any $\pi_1$ and $\pi_2$. If the two markets have identical characteristics, i.e. $\sigma_{P,1} = \sigma_{P,2}$, only the combined cash flow matters for the total wedge – i.e., the Modigliani-Miller theorem applies. If on the other hand the two markets have different informational characteristics, the increasing difference property of $\Delta_{\pi} (\cdot)$ implies the seller’s expected revenue is higher when the security with more upside risk is matched with the market that has more severe information frictions (a higher value of $\sigma_{P}$). This maximizes the gains from the positive wedge resulting on the upside, while minimizing the losses from the negative wedge on the downside.\textsuperscript{27} This logic is advanced further in the next proposition, which considers how the seller can exploit the heterogeneity in investor pools if she gets to design the split of $\pi$ into $\pi_1$ and $\pi_2$.

Proposition 6 (Designing cash flows): The seller maximizes her expected revenues by splitting cash flows according to $\pi_1^*(\theta) = \min \{\pi (\theta), \pi (0)\}$ and $\pi_2^*(\theta) = \max \{\pi (\theta) - \pi (0), 0\}$, and then assigning $\pi_1^*$ to the investor pool with the lower value of $\sigma_{P}$.

Figure 1 sketches the optimal dividend split for an arbitrary dividend function. The seller maximizes the total proceeds by assigning all the cash flow below the line defined by $\pi(\cdot) = \pi(0)$ to the investor group with the lowest information friction parameter; $\sigma_{P,1}$, and the complement to the investor group with the highest friction; $\sigma_{P,2}$. Intuitively, the optimal split loads the entire downside risk on the investor group that discounts the price of the claim the least with respect to its expected payoff (because of the low friction parameter; $\sigma_{P,1}$), while loading the entire upside

\textsuperscript{27}With no arbitrage, the Modigliani-Miller Theorem applies not just to expected prices, but also at an interim stage, i.e. for realized prices and revenues. In our model, this occurs only if $P_{z_1} (z_1, z_2) + P_{z_2} (z_1, z_2)$ is invariant to the dividend split $\{\pi_1(\cdot), \pi_2(\cdot)\}$. This requires that the supply shocks in the two markets are perfectly correlated: $\rho = 1$, and the two markets have identical informational characteristics ($\beta_1 = \beta_2$ and $\sigma_{u,1} = \sigma_{u,2}$). Furthermore, with informational linkages, there exists the additional possibility that when $\rho = 1$ and $\beta_1 \sigma_{u,1}^{-2} \neq \beta_2 \sigma_{u,2}^{-2}$, the observations of the two price signals with different precision but perfectly correlated noise enables every trader to perfectly infer $\theta$ and $u$, implying $P_{\pi_1} (z_1, z_2) = \pi_1 (\theta)$ and $P_{\pi_2} (z_1, z_2) = \pi_2 (\theta)$.

The interim version of the Modigliani-Miller theorem thus requires not just that the informational characteristics of the two markets are the same, but in addition that the noise trading shocks are perfectly correlated so that the second price signal becomes redundant.
risk to the group that overvalues the claim the most with respect to its expected dividend (due to the high information frictions; $\sigma_{P2}$). When $\pi(\cdot) > 0$, this split has a straightforward interpretation in terms of debt and equity, with a default point on debt that is set at the prior median $\pi(0)$. For any other arbitrary division of cash flows $\{\pi_1(\cdot), \pi_2(\cdot)\}$, $\pi_1$ has less downside risk than $\pi_1^*$, and $\pi_2$ has less upside risk than $\pi_2^*$. This results in a higher expected price on $\pi_1$ and a lower expected price on $\pi_2$, but due to the increasing difference property (Theorem 2, part iv), the lower expected price on $\pi_2$ dominates, resulting in strictly lower expected revenue for the seller.\footnote{We could use the same logic to formulate a simpler security design problem, in which the seller designs a security $\pi_2(\cdot)$ for a single investor pool, keeping the residual $\pi_1 = \pi - \pi_2$ to himself. By the same argument as above, the seller’s optimal design consists in an options contract $\pi_2^*(\theta) = \max\{\pi(\theta) - \pi(0), 0\}$ that maximizes the upside sold in the market at a premium, while keeping the downside risk (or debt claim) that would be under-priced to himself.} This proposition illustrates how noisy information aggregation can shape optimal security design through the link between risk asymmetries and prices that emerges in equilibrium.

In this analysis we take as given the differences in market characteristics and assume that the seller can freely assign the cash-flows to these two pools. This is an important limitation, as it omits the possibility that market characteristics themselves respond to how the seller designs the securities – for example because the investor’s incentives to obtain information also depend on the asset risks they face, and their ability or willingness to arbitrage across different markets. Analyzing this interplay between investor’s information choices and the resulting market characteristics, along with the seller’s security design question is an important avenue for further work, but clearly beyond the scope of this paper. The results here are merely intended to highlight the possibility of
systematic departures from the Modigliani-Miller irrelevance result, and to show that information frictions outside the firm give the owner of a cash flow distinct possibility to manipulate its market value through security design.

We now briefly discuss how our application relates to the literature on security design. A key theme in this literature is the role of “information-insensitive” debt contracts (with or without default risk, and possibly in combination with pooling of securities), which mitigate the asymmetric information problem between uninformed outsiders and insiders who sell claims to raise funds (Myers and Majluf, 1984; DeMarzo and Duffie, 1999; DeMarzo 2005), or limit liquidity traders’ exposures to trading losses in a market with buyers or sellers who hold superior information about the claims’ quality (Gorton and Pennacchi, 1990; Boot and Thakor, 1993). Boot and Thakor (1993) and Fulghieri and Lukin (2001) also emphasize the role of information-sensitive junior securities or equity claims to incentivize information acquisition, which benefits high quality firms in equilibrium. In all these models, security prices are either set by a risk-neutral competitive market-maker or by a set of homogeneous outside investors, so securities are priced at their expected fundamental value, taking into consideration the signaling effects of the insiders’ security design decisions. But then, limits to arbitrage or mispricing of securities is not considered as a force shaping security design incentives.

Our results are closer to a second line of papers which abstracts from signaling issues and emphasizes the role of market frictions in security pricing. Closest to our work, Axelson (2007) studies security design when investors have private information about a firm’s cash flow and securities are sold through a K-th price auction. Without liquidity trading, the auction mechanism generically results in under-pricing of securities due to a winners’ curse argument, and the optimal security design seeks to limit the losses associated with the winner’s curse resulting in either debt or call option contract - depending on whether the issuer has more to gain from limiting the winner’s curse through information insensitivity (debt) or from aligning the firm’s fundamentals with amount of cash raised through information sensitive securities (equity or options). As the market becomes more and more competitive, the underpricing disappears and the optimal security design approximates an equity claim. Like Axelson (2007), we also emphasize market frictions as a driving force of optimal security design, but there are several important differences. First, by allowing for liquidity shocks along with informed trading, our model opens up the possibility that securities can be over- as well as under-priced, resulting in the separation of upside vs. downside risk as the key

---

force driving optimal security design. Second, we abstract from the objective of raising external resources by taking the cash flow as exogenous, assuming for simplicity that the firms sells the entire cash flow. Third, we allow for investor pools with different informational characteristics. The debt-equity split then emerges as the optimal way of catering specific securities to the different investor pools.

6 Concluding Remarks

In this paper we have presented a theory of asset price formation based on heterogeneous information. This theory ties expected asset returns to properties of their risk profile and the market’s information structure. The theory is parsimonious, in the sense that all its results follow directly from the interplay of asset payoffs and information heterogeneity. The theory is general: the main characterization result in Theorem 1 imposes no restrictions on the distribution of asset payoffs, and only Theorem 2 works with specific assumptions on preferences and information. The results are therefore able to speak to much wider and much less stylized asset structures than most of the prior literature on noisy information aggregation. The theory is consistent with the empirical facts on return to skewness. Our theory is tractable and easily lends itself to applications, such as our discussion of the Modigliani-Miller theorem.

We conclude with remarks on other potential applications and future research. An obvious direction is to explore the implications of information heterogeneity for volatility of prices and returns; the earlier working paper version (Albagli, Hellwig, and Tsyvinski, 2011) already explored noisy information aggregation as a potential source of excess price volatility and low predictability of returns. Another direction is to explore the effects of public news and information disclosures into our asset pricing model. A third direction is to explore other asset pricing puzzles (such as option pricing anomalies, equity and bond returns) through the lens of information heterogeneity. We explore this direction in the specific case of corporate credit spreads in Albagli, Hellwig, and Tsyvinski (2014a). A fourth direction is to extend the analysis of a multi-period, and multi-asset extensions of our market model. A final direction lies in the integration of financial market frictions with real decisions that shape allocation of resources and firm dividends. In Albagli, Hellwig, and Tsyvinski (2014b), we consider one such model to study the interplay between noisy information aggregation and shareholder incentives.
References


Appendix: Proofs

We first state a lemma that will be useful for the first theorem:

**Lemma 2** Suppose that $\theta$ is distributed according to some continuous bounded density $h(\cdot)$ and that the signal density $f(\cdot)$ satisfies log-concavity and unboundedness of $\frac{f'(\cdot)}{f(\cdot)}$. 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.** Monotonicity is shown by Milgrom (1981a). For the characterization at the extremes, let $\hat{\varepsilon} = \max_{\varepsilon} f(\varepsilon)$, and notice that when $x > \theta + \hat{\varepsilon}$,

$$
\frac{H(\theta|x)}{1 - H(\theta|x)} = \frac{\int_{-\infty}^{\theta} f(x - \theta') dH(\theta')}{{\int_{\theta}^{\infty} f(x - \theta') dH(\theta')}} = \frac{\int_{-\infty}^{\theta} f(x - \theta') dH(\theta')}{{\int_{\theta}^{\infty} f'(x - \theta') [H(\theta') - H(\theta)] d\theta'}}
\leq \frac{1 - F(x - \theta) \max_{\theta} h(\theta)}{f(x - \theta) \hat{H}(x - \hat{\varepsilon}) - H(\theta)}.
$$

Since $f(\varepsilon) / (1 - F(\varepsilon)) = \mathbb{E}(-f'(\varepsilon') / f(\varepsilon') | \varepsilon' \geq \varepsilon)$, it follows that $\lim_{\varepsilon \to \infty} f(\varepsilon) / (1 - F(\varepsilon)) = \infty$, and therefore $\lim_{x \to \infty} H(\theta|x) = 0$. An analogous argument shows that when $x < \theta + \hat{\varepsilon}$,

$$
\frac{1 - H(\theta|x)}{H(\theta|x)} \leq \frac{F(x - \theta) \max_{\theta} h(\theta)}{f(x - \theta) \hat{H}(x - \hat{\varepsilon}) - H(\theta)},
$$

and since $\lim_{\varepsilon \to -\infty} f(\varepsilon) / F(\varepsilon) = \infty$, it follows that $\lim_{x \to -\infty} H(\theta|x) = 1$. $
$

**Lemma 3** 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 $d(x, P) = d_L(P)$ for all $x \leq x_L(P)$, $d(x, P) = d_H(P)$ 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 pdf
\[ h(\cdot|P,d) = \frac{h(\theta|P)U'(d(\pi(\theta) - P))}{\int h(\theta|P)U'(d(\pi(\theta) - P)) d\theta}, \]
and let $\tilde{H}(\cdot|x, P, d)$ and $\hat{E}(\pi(\theta)|x, P, d) \equiv \int \pi(\theta) d\tilde{H}(\cdot|x, P, d)$ denote the cdf and conditional expectations after updating conditional on a private signal $x$. From the previous lemma, it follows immediately that $\tilde{H}(\cdot|x, P, d)$ is strictly decreasing in $x$. Therefore, $\hat{E}(\pi(\theta)|x, P, d)$ is strictly increasing in $x$ and if $P$ is on the interior of the support of $\pi(\theta)$, then $\lim_{x \to -\infty} \hat{E}(\pi(\theta)|x, P, d) < P < \lim_{x \to \infty} \hat{E}(\pi(\theta)|x, P, d)$. But then there exist $x_L(P)$ s.t. $\hat{E}(\pi(\theta)|x_L(P), P, d_L(P)) = P$, which implies that $d(x, P) = d_L(P)$ for all $x \leq x_L(P)$, and $x_H(P)$ s.t. $\hat{E}(\pi(\theta)|x_H(P), P, d_H(P)) = P$, which implies that $d(x, P) = d_H(P)$ for all $x \geq x_H(P)$.

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

Proof of Theorem 1:

For any $\tilde{S}(P) \in (d_L(P), d_H(P))$; lemma 2 implies that there exists a unique $z(P) \in (x_L(P), x_H(P))$; s.t. $d(z(P), P) = \tilde{S}(P)$, or equivalently, $P = \hat{E}(\pi(\theta)|z(P), P, \tilde{S}(P))$. Combining with the equilibrium price function, we then define the sufficient statistic function $z(\theta, u) = z(P(\theta, u))$.

Next, $d(x; P)$ is strictly increasing in $x$, so by construction, $z(P)$ is strictly monotone and hence invertible, if and only if condition 1 is satisfied. But if $z(P)$ is invertible, then $z(P)$ and $P$ are informationally equivalent.

Now, if $z(P)$ is invertible, we have $P = \hat{E}(\pi(\theta)|x, P, \tilde{S}(P))$, if and only if
\[ P(z) = \frac{\int \pi(\theta)U'(\tilde{S}(P(z)))(\pi(\theta) - P(z)) dH(\theta|x = z, z)}{\int U'(\tilde{S}(P(z)))(\pi(\theta) - P(z)) dH(\theta|x = z, z)}. \]
from which we immediately obtain that $z(P)$ and $z(\theta, u)$ as constructed above satisfy condition 2.

To prove the converse, suppose that the equilibrium admits a sufficient statistic function $z(\theta, u)$ with $P(z)$ implicitly defined by
\[ P(z) = \frac{\int \pi(\theta)U'(\tilde{S}(P(z)))(\pi(\theta) - P(z)) dH(\theta|x = z, z, \theta, u = z)}{\int U'(\tilde{S}(P(z)))(\pi(\theta) - P(z)) dH(\theta|x = z, z, \theta, u = z)} \]
and assume that $P(z)$ is invertible. Because the equilibrium price function satisfies $P(\theta, u) = P(z(\theta, u))$, we have

39
But then it follows immediately that $P(z)$ is the inverse of the function $z(P)$ that we defined above, and hence $z(P)$ is invertible.

**Proof of Proposition 1:**

Conjecture first that at equilibrium, $P$ is informationally equivalent to $z$. This occurs if and only if $\hat{x}$ is invertible. In this case, substituting $\hat{x} = z$ into the marginal trader’s indifference condition, we obtain that $P(z) = \mathbb{E}(\pi(\theta) \mid x = z, z)$. Therefore, the conjectured price function and threshold signals form an equilibrium, if and only if $\mathbb{E}(\pi(\theta) \mid x = z, z)$ is strictly increasing. Moreover, by market-clearing, $z = \hat{x}(P(z))$ and $z' = \hat{x}(P(z'))$, and therefore $z = z'$ if and only if $P(z) = P(z')$. Therefore, the equilibrium conjectured above is the only equilibrium, in which $P$ is informationally equivalent to $z$.

Next, in any equilibrium, in which $d(x, P)$ is non-increasing in $P$, $\hat{x}(P)$ must be non-decreasing in $P$. Moreover, if $\hat{x}(P)$ must be non-decreasing in $P$, it 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 $\hat{x}(P)$ is flat over some range, i.e. $\hat{x}(P) = \hat{x}$ for $P \in (P', P'')$. 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.\(^{30}\) But then for $z \in (\hat{x}(P' - \varepsilon), \hat{x})$ and $z \in (\hat{x}, \hat{x}(P'' + \varepsilon))$, $P(z)$ is uniquely defined, so we have $P' \geq \lim_{z \uparrow x} P(z) = \lim_{z \uparrow x} \mathbb{E}(\pi(\theta) \mid x = z, z)$ and $P'' \leq \lim_{z \downarrow x} P(z) = \lim_{z \downarrow x} \mathbb{E}(\pi(\theta) \mid x = z, z)$. But since $\mathbb{E}(\pi(\theta) \mid x = z, z)$ is continuous, it must be that $P'' \leq P'$, which yields a contradiction.

**Proof of Lemma 1:**

By the law of iterated expectations, $\mathbb{E}(V(z)) = \mathbb{E}(\pi(\theta)) = \int_{-\infty}^{\infty} \pi(\theta) d\Phi(\theta / \sigma_{\theta})$. To find $\mathbb{E}(P(z))$, define $\sigma_P^2 = \sigma_{\theta}^2 (1 + (\gamma_P / \gamma_V - 1) \gamma_P)$. Simple algebra shows that

$$
\int_{-\infty}^{\infty} \frac{1}{\sqrt{1 - \gamma_P^2 \sigma_{\theta}^2}} \phi \left( \frac{\theta - \gamma_P z}{\sqrt{1 - \gamma_P^2 \sigma_{\theta}^2}} \right) d\Phi \left( \frac{\sqrt{\gamma_V z}}{\sigma_{\theta}} \right) = \frac{1}{\sigma_P} \phi \left( \frac{\theta}{\sigma_P} \right).
$$

\(^{30}\)It cannot be flat everywhere, because then informed demand would be completely inelastic, and there would be no way to absorb noise trader shocks.
With this, we compute $\mathbb{E}(P(z))$:

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

$$
= \int_{-\infty}^{\infty} \pi(\theta) \int_{-\infty}^{\infty} \frac{1}{\sqrt{1 - \gamma_p \theta}} \phi \left( \frac{\theta - \gamma_p z}{\sqrt{1 - \gamma_p \theta}} \right) d\Phi \left( \frac{\sqrt{\gamma}z}{\sigma_\theta} \right) d\theta
$$

$$
= \int_{-\infty}^{\infty} \pi(\theta) \frac{1}{\sigma_p} \phi \left( \frac{\theta}{\sigma_p} \right) d\theta
$$

and

$$
\Delta_p = \int_{-\infty}^{\infty} \pi(\theta) \left( \frac{1}{\sigma_p} \phi \left( \frac{\theta}{\sigma_p} \right) - \frac{1}{\sigma_\theta} \phi \left( \frac{\theta}{\sigma_\theta} \right) \right) d\theta = \int_{-\infty}^{\infty} \pi'(\theta) \left( \Phi \left( \frac{\theta}{\sigma_\theta} \right) - \Phi \left( \frac{\theta}{\sigma_p} \right) \right) d\theta
$$

$$
= \int_0^{\infty} \left( \pi'(\theta) - \pi'(-\theta) \right) \left( \Phi \left( \frac{\theta}{\sigma_\theta} \right) - \Phi \left( \frac{\theta}{\sigma_p} \right) \right) d\theta,
$$

where the first equality proceeds by integration by parts, the second by a change in variables, and the third step uses the symmetry of the normal distribution ($\Phi(-x) = 1 - \Phi(x)$).

**Proof of Theorem 2:**

Parts (i) - (iii) follow immediately from lemma 1, the definition of upside and downside risk, and the fact that $\Phi(\theta/\sigma_\theta) > \Phi(\theta/\sigma_p)$ for all $\theta$ (since $\sigma_p > \sigma_\theta$). For part (iv) notice that

$$
\Delta_{\pi_1}(\sigma_p) - \Delta_{\pi_2}(\sigma_p) = \int_0^{\infty} \left( \pi'_1(\theta) - \pi'_1(-\theta) - (\pi'_2(\theta) - \pi'_2(-\theta)) \right) \left( \Phi \left( \frac{\theta}{\sigma_\theta} \right) - \Phi \left( \frac{\theta}{\sigma_p} \right) \right) d\theta,
$$

and therefore $\Delta_{\pi_1}(\sigma_p) - \Delta_{\pi_2}(\sigma_p)$ is positive and increasing in $\sigma_p$.

**Proof of Proposition 2:**

**Part 1:** Derivation of $\mathbb{E}(P(z))$: Write $\mathbb{E}(P(z))$ as

$$
\mathbb{E}(P(z)) = \lambda \int_{-\infty}^{\infty} \frac{\lambda}{\lambda + (1 - \lambda) e^{-(\beta + \tau_p)\left(\frac{1}{2} + \frac{1}{3}s\right)}} d\Phi \left( \frac{s}{\sigma_u} \right) + \lambda \int_{-\infty}^{\infty} \frac{1 - \lambda}{\lambda + (1 - \lambda) e^{(\beta + \tau_p)\left(\frac{1}{2} + \frac{1}{3}s\right)}} d\Phi \left( \frac{s}{\sigma_u} \right)
$$

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

where $x = (\beta + \tau_p)\left(\frac{1}{2} + \frac{1}{3}s\right)$, $\hat{x} = \frac{1}{2} (\beta + \tau_p)$, and $\sigma_x^2 = (\beta + \tau_p)^2 / \tau_p$. Using the fact that $\phi \left( \frac{x - \hat{x}}{\sigma_x} \right) / \phi \left( \frac{-x - \hat{x}}{\sigma_x} \right) = e^{2x^2/\sigma_x^2} = e^{\psi x}$, where $\psi = \tau_p / (\tau_p + \beta)$, we obtain

$$
\mathbb{E}(P(z)) = \lambda \int_{-\infty}^{\infty} \left\{ \frac{\lambda + (1 - \lambda) e^{-\psi x}}{\lambda + (1 - \lambda) e^{-x}} \right\} d\Phi \left( \frac{x - \hat{x}}{\sigma_x} \right) = \lambda \int_{-\infty}^{\infty} \left\{ \frac{1 - (1 - \lambda) (1 - e^{-\psi x})}{1 - (1 - \lambda) (1 - e^{-x})} \right\} d\Phi \left( \frac{x - \hat{x}}{\sigma_x} \right)
$$

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

41
Splitting the integral at 0, we have

\[
\begin{align*}
\mathbb{E}(P(z)) - \lambda &= \lambda \int_0^\infty \left\{ \frac{(1 - \lambda) \left( e^{-\psi x} - e^{-x} \right)}{1 - (1 - \lambda) (1 - e^{-x})} + \frac{(1 - \lambda) \left( e^{\psi x} - e^{x} \right) e^{-\psi x}}{1 - (1 - \lambda) (1 - e^{x})} \right\} d\Phi \left( \frac{x - \hat{x}}{\sigma_x} \right) \\
&= \lambda \int_0^\infty \left\{ \frac{(1 - \lambda) (e^{(1-\psi)x} - 1) e^{-x}}{1 - (1 - \lambda) (1 - e^{-x})} - \frac{(1 - \lambda) (e^{(1-\psi)x} - 1)}{1 - (1 - \lambda) (1 - e^{x})} \right\} d\Phi \left( \frac{x - \hat{x}}{\sigma_x} \right)
\end{align*}
\]

Now,

\[
\frac{e^{-x}}{1 - (1 - \lambda) (1 - e^{-x})} - \frac{1}{1 - (1 - \lambda) (1 - e^{x})} = \frac{(1 - e^{-x}) (1 - 2\lambda)}{1 + \lambda (1 - \lambda) (e^{x} + e^{-x} - 2)},
\]

and therefore \(\mathbb{E}(P(z))\) takes the form \(\mathbb{E}(P(z)) = \lambda + (1 - 2\lambda) \Delta^P\), where

\[
\Delta^P = \int_0^\infty \left\{ \frac{\lambda (1 - \lambda) (e^{(1-\psi)x} - 1) (1 - e^{-x})}{1 + \lambda (1 - \lambda) (e^{x} + e^{-x} - 2)} \right\} d\Phi \left( \frac{x - \hat{x}}{\sigma_x} \right).
\]

This formulation encompasses the “No-Information” Case as the limit in which \(\beta \to 0, \psi = \hat{x} = 0\) and \(\sigma_x^2 = \gamma^2 \sigma_u^2\), and therefore \(\mathbb{E}(V^{NI}(u)) = \lambda + (1 - 2\lambda) \Delta^{NI}\), where

\[
\Delta^{NI} = \int_0^\infty \left\{ \frac{\lambda (1 - \lambda) (e^x + e^{-x} - 2)}{1 + \lambda (1 - \lambda) (e^{x} + e^{-x} - 2)} \right\} d\Phi \left( \frac{x}{\gamma \sigma_u} \right).
\]

Derivation of \(\mathbb{E}(V^{PI}(u, z))\): Write \(\mathbb{E}(V^{PI}(u, z))\) as

\[
\mathbb{E}(V^{PI}(u, z)) = \lambda \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{\lambda}{\lambda + (1 - \lambda) e^{-\gamma u - \tau_p \left( \frac{1}{2} + \frac{\gamma}{\beta} s \right)}} d\Phi \left( \frac{s}{\sigma_u} \right) d\Phi \left( \frac{u}{\sigma_u} \right) + \lambda \int_{-\infty}^\infty \frac{1 - \lambda}{\lambda + (1 - \lambda) e^{\gamma u + \tau_p \left( \frac{1}{2} + \frac{\gamma}{\beta} s \right)}} d\Phi \left( \frac{s}{\sigma_u} \right) d\Phi \left( \frac{u}{\sigma_u} \right)
\]

\[
= \lambda \int_{-\infty}^\infty \left\{ \frac{\lambda}{\lambda + (1 - \lambda) e^{-x}} + \frac{1 - \lambda}{\lambda + (1 - \lambda) e^{x}} \right\} d\Phi \left( \frac{x - \tilde{x}}{\sigma_x} \right)
\]

where \(x = \gamma u + \tau_p \left( \frac{1}{2} + \frac{\gamma}{\beta} s \right), \tilde{x} = \frac{1}{2} \tau_p\), and \(\tilde{\sigma}_x^2 = \tau_p + \gamma^2 \sigma_u^2\). Since \(\phi \left( \frac{x - \tilde{x}}{\tilde{\sigma}_x} \right) / \phi \left( \frac{-x - \tilde{x}}{\tilde{\sigma}_x} \right) = e^{2x \bar{\psi} / \tilde{\sigma}_x^2} = e^{\bar{\psi} x}\), where \(\bar{\psi} = \tau_p / (\tau_p + \gamma^2 \sigma_u^2)\), we follow exactly the same steps as above to obtain that \(\mathbb{E}(V^{PI}(u, z)) = \lambda + (1 - 2\lambda) \Delta^{PI}\), where

\[
\Delta^{PI} = \int_0^\infty \left\{ \frac{\lambda (1 - \lambda) (e^{(1-\bar{\psi})x} - 1) (1 - e^{-x})}{1 + \lambda (1 - \lambda) (e^x + e^{-x} - 2)} \right\} d\Phi \left( \frac{x - \tilde{x}}{\sigma_x} \right).
\]

**Part 2:** We first show that \(\Delta^{NI} > \Delta^{PI}\). For \(\Delta^{NI}\), we have \(\Delta^{NI} = \int_0^\infty f(x) d\Phi \left( \frac{x}{\gamma \sigma_u} \right),\) where

\[
f(x) = \frac{\lambda (1 - \lambda) (e^x + e^{-x} - 2)}{1 + \lambda (1 - \lambda) (e^x + e^{-x} - 2)}.
\]
For $\Delta^{PI}$, we have

$$
\Delta^{PI} = \int_0^\infty \frac{e^{-(1-\overline{\psi})x} - 1}{e^x - 1} f(x) d\Phi \left( \frac{x - \overline{x}}{\sigma_x} \right) = \int_0^\infty \frac{e^x - e^{\overline{\psi}x}}{e^x - 1} f(x) d\Phi \left( \frac{x + \overline{x}}{\sigma_x} \right),
$$

using the fact that $\phi \left( \frac{x - \overline{x}}{\sigma_x} \right) = e^{\overline{\psi}x} \phi \left( \frac{x + \overline{x}}{\sigma_x} \right)$. It follows that

$$
\frac{\Delta^{PI}}{\Delta^{NI}} = \frac{\int_0^\infty \frac{e^{x - e^{\overline{\psi}x}}}{e^{\overline{\psi}x} - 1} f(x) d\Phi \left( \frac{x + \overline{x}}{\sigma_x} \right)}{\int_0^\infty f(x) d\Phi \left( \frac{x}{\gamma \sigma_u} \right)} < \frac{\frac{\gamma \sigma_u}{\sigma_x} \int_0^\infty \max \left\{ 1, \phi \left( \frac{x + \overline{x}}{\sigma_x} \right) / \phi \left( \frac{x}{\gamma \sigma_u} \right) \right\} f(x) d\Phi \left( \frac{x}{\gamma \sigma_u} \right)}{\int_0^\infty f(x) d\Phi \left( \frac{x}{\gamma \sigma_u} \right)},
$$

where the first inequality uses the fact that $e^{x - e^{\overline{\psi}x}} < 1$. Next, note that $\max \left\{ 1, \phi \left( \frac{x + \overline{x}}{\sigma_x} \right) / \phi \left( \frac{x}{\gamma \sigma_u} \right) \right\}$ is strictly increasing in $x$ for all $x > \overline{x}$, where $\overline{x} = \frac{1}{2} \gamma^2 \sigma_u^2 \left( 1 + 1/\sqrt{1 - \overline{\psi}} \right)$ solves $\overline{x} / \gamma \sigma_u = (\overline{x} + \overline{x}) / \overline{x}$, and $\max \left\{ 1, \phi \left( \frac{x + \overline{x}}{\sigma_x} \right) / \phi \left( \frac{x}{\gamma \sigma_u} \right) \right\} = 1$ for all $x \leq \overline{x}$. Notice also that we can interpret the function $F(x) = \int_0^x f(x') d\Phi \left( \frac{x'}{\gamma \sigma_u} \right) / \int_0^\infty f(x') d\Phi \left( \frac{x'}{\gamma \sigma_u} \right)$ as a cumulative distribution function, and therefore the last expression as an expectation of $\max \left\{ 1, \phi \left( \frac{x + \overline{x}}{\sigma_x} \right) / \phi \left( \frac{x}{\gamma \sigma_u} \right) \right\}$ w.r.t. the distribution $F(\cdot)$. Moreover, it is straightforward to check that

$$
F(x) \geq \hat{F}(x) \equiv \frac{\Phi \left( \frac{x}{\gamma \sigma_u} \right) - \Phi \left( \frac{\overline{x}}{\gamma \sigma_u} \right)}{1 - \Phi \left( \frac{\overline{x}}{\gamma \sigma_u} \right)}
$$

for $x \geq \overline{x}$, and therefore $\hat{F}(\cdot)$ first-order stochastically dominates $F(\cdot)$. But this then implies that

$$
\frac{\Delta^{PI}}{\Delta^{NI}} < \frac{\frac{\gamma \sigma_u}{\sigma_x} \int_0^\infty \frac{\phi \left( \frac{x + \overline{x}}{\sigma_x} \right)}{\phi \left( \frac{x}{\gamma \sigma_u} \right)} d\Phi \left( \frac{x}{\gamma \sigma_u} \right)}{1 - \Phi \left( \frac{\overline{x}}{\gamma \sigma_u} \right)} = \frac{\int_0^\infty d\Phi \left( \frac{x + \overline{x}}{\sigma_x} \right)}{1 - \Phi \left( \frac{\overline{x}}{\gamma \sigma_u} \right)} = \frac{1 - \Phi \left( \frac{\overline{x} + \overline{x}}{\gamma \sigma_u} \right)}{1 - \Phi \left( \frac{\overline{x}}{\gamma \sigma_u} \right)} = 1,
$$

or $\Delta^{PI} < \Delta^{NI}$.

Next, for the comparison with $\Delta^P$, consider the limiting case where $\gamma \sigma_u \to \infty$: Since $\Delta^{NI} = \int_0^\infty f(\gamma \sigma_u v) d\Phi(v)$ and $\lim_{x \to \infty} f(x) = 1$, $\Delta^{NI}$ converges to $\lim_{\gamma \sigma_u \to \infty} \Delta^{NI} = \int_0^\infty d\Phi(v) = 1/2$. For the public information case,

$$
\Delta^{PI} = e^{-\overline{\psi} \sqrt{\gamma \nu}} \int_0^\infty \left\{ f(\gamma \sigma_u v) \frac{e^{(1-\overline{\psi}/2)\gamma \sigma_u v} - e^{\overline{\psi}/2 \gamma \sigma_u v}}{e^{\gamma \sigma_u v} - 1} \right\} d\Phi \left( \frac{\gamma \sigma_u v}{\sigma_x} \right),
$$

where $\overline{\psi} \to 0$. Both terms inside the bracket then converge to 1, so $\lim_{\gamma \sigma_u \to \infty} \Delta^{PI} = \int_0^\infty d\Phi(\gamma \sigma_u v/\sigma_x) = 1/2$. Finally,

$$
\Delta^P = \int_0^\infty \left\{ f(x) \frac{e^{(1-\overline{\psi})x} - 1}{e^x - 1} \right\} d\Phi \left( \frac{x - \overline{x}}{\sigma_x} \right) = \int_0^\infty \left\{ f(\gamma \sigma_u v) \frac{e^{(1-\overline{\psi})\gamma \sigma_u v} - 1}{e^{\gamma \sigma_u v} - 1} \right\} d\Phi \left( (1-\overline{\psi}) v - \frac{1}{2} \sqrt{\gamma \nu} \right).
$$
Here, the first term converges to 1, the second to $e^{-\psi\gamma\sigma u}$. Since $\psi\gamma\sigma u = (1 - \psi) \sqrt{\tau_p}$, we obtain

$$\lim_{\gamma\sigma u \to \infty} \Delta P = \lim_{\psi \to 0} \int_0^\infty d\Phi \left( (1 - \psi) \frac{1}{2} \sqrt{\tau_p} \right) = 1 - \Phi \left( \frac{1}{2} \sqrt{\tau_p} \right) < \frac{1}{2}. $$

Therefore for $\gamma\sigma u$ sufficiently large, we obtain $\Delta^{NI} > \Delta^{PI} > \Delta^P$. In part 3, we show that $\Delta^P > \Delta^{NI} > \Delta^{PI}$ in the opposite case when $\gamma\sigma u \to 0$.

**Part 3:** As $\gamma\sigma u \to 0$ : For the no-information benchmark, notice that

$$\frac{\Delta^{NI}}{\gamma^2 \sigma^2 u} = \int_0^\infty \left\{ \frac{\lambda (1 - \lambda)}{1 + \lambda (1 - \lambda)} \left( e^{\gamma\sigma u \nu - 1} (1 - e^{-\gamma\sigma u \nu}) \right) \right\} d\Phi \left( v \right),$$

and therefore

$$\lim_{\gamma\sigma u \to 0} \frac{\Delta^{NI}}{\gamma^2 \sigma^2 u} = \int_0^\infty \left\{ \frac{\lambda (1 - \lambda)}{1 + \lambda (1 - \lambda)} \left( e^{\gamma\sigma u \nu - 1} (1 - e^{-\gamma\sigma u \nu}) \right) \right\} d\Phi \left( \frac{\gamma\sigma u \nu}{\sigma x} \right).$$

Now, the second fraction converges to 1, while the first one converges to $\lambda (1 - \lambda) v^2$, as $\gamma\sigma u \to 0$.

In addition, $\tilde{\psi} \to 1$. Therefore,

$$\lim_{\gamma\sigma u \to 0} \frac{\Delta^{PI}}{\gamma^2 \sigma^2 u} = \lim_{\gamma\sigma u \to 0} e^{-\frac{1}{8} \sqrt{\tau_p}} \lambda (1 - \lambda) \int_0^\infty v^2 d\Phi \left( \frac{\gamma\sigma u \nu}{\sigma x} \right) = 1 - \frac{1}{2} \lambda (1 - \lambda) e^{-\frac{1}{8} \sqrt{\tau_p}}.$$ 

Finally, for the dispersed information model, we have

$$\frac{\Delta^P}{\gamma\sigma u} = e^{-\frac{1}{8} \sqrt{\tau_p}} \int_0^\infty \left\{ \frac{\lambda (1 - \lambda)}{1 + \lambda (1 - \lambda)} \left( e^{\gamma\sigma u \nu - 1} (1 - e^{-\gamma\sigma u \nu}) \right) \right\} d\Phi \left( \frac{\gamma\sigma u \nu}{\sigma x} \right),$$

following the same manipulations as in the public information case. As $\gamma\sigma u \to 0$, the second fraction converges to 1, and the first one to $\gamma\sigma u \lambda (1 - \lambda) v^2$. Therefore,

$$\lim_{\gamma\sigma u \to 0} \frac{\Delta^P}{\gamma\sigma u} = \lim_{\gamma\sigma u \to 0} e^{-\frac{1}{8} \sqrt{\tau_p}} \lambda (1 - \lambda) \int_0^\infty v^2 d\Phi \left( \frac{\gamma\sigma u \nu}{\sigma x} \right) = 1 - \frac{1}{2} \lambda (1 - \lambda) e^{-\frac{1}{8} \sqrt{\tau_p}} \sqrt{\tau_p}. $$

44
Proof of Proposition 3:

Parts 1 and 2: Derivation of $\mathbb{E}(P(z))$, when $\overline{u} > 0$: Write $\mathbb{E}(P(z))$ as

$$
\mathbb{E}(P(z)) = \hat{\lambda} \int_{-\infty}^{\infty} \left\{ \frac{\lambda}{\hat{\lambda} + (1 - \hat{\lambda}) e^{-x}} + \frac{1 - \lambda}{\hat{\lambda} + (1 - \hat{\lambda}) e^{x}} \right\} d\Phi \left( \frac{x - \hat{x}}{\sigma_x} \right)
$$

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

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

where $x = (\beta + \tau_p) \left( \frac{1}{2} + \frac{1}{2} \sigma_2 \right)$, $\hat{x} = \frac{1}{2} (\beta + \tau_p)$, and $\sigma^2_x = (\beta + \tau_p)^2 / \tau_p$, as before. Following the same derivations as in the proof of proposition 2, the first integral can be rewritten as $\hat{\lambda} + \left( 1 - 2\hat{\lambda} \right) \Delta^P$, with $\Delta^P$ defined as before, but $\lambda$ replaced by $\hat{\lambda}$. Defining

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

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

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

The no-information case is again nested by letting $\beta \to 0$, $\psi = \hat{x} = 0$ and $\sigma^2_x = \gamma^2 \sigma^2_u$, so that $R^{NI} = 1$. Finally, following the same steps as above, the case with public information results in a representation $\mathbb{E}(V^{PI}(z)) = \lambda + \left( 1 - 2\hat{\lambda} \right) \Delta^{PI} - \left( \lambda - \hat{\lambda} \right) R^{PI}$, where

$$
R^{PI} = 1 - \int_{-\infty}^{\infty} \frac{\hat{\lambda} (1 - \hat{\lambda}) (e^x - e^{-x})}{1 + \hat{\lambda} (1 - \hat{\lambda}) (e^x + e^{-x} - 2)} d\Phi \left( \frac{x - \tilde{x}}{\tilde{\sigma}_x} \right) = 1 - \int_{-\infty}^{\infty} g(x) d\Phi \left( \frac{x - \tilde{x}}{\tilde{\sigma}_x} \right),
$$

with $\tilde{x} = \frac{1}{2} \tau_p$ and $\tilde{\sigma}_x^2 = \tau_p + \gamma^2 \sigma^2_u$ defined as in the proof of proposition 2. Furthermore, since

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

and $1 - 2\hat{\lambda} (1 - \hat{\lambda}) (1 - e^{-x}) > 1 - 2\hat{\lambda} (1 - \hat{\lambda}) > 0$ for all $x$ and $\hat{\lambda}$, we have $R^P > 0$ and $R^{PI} > 0$.
**Part 3:** To see that $R^P < 1$ and $R^{PI} < 1$, we need to show that the respective integrals are positive. Splitting the integrals at zero and using the fact that $\phi\left(\frac{x-\hat{x}}{\sigma_x}\right)/\phi\left(\frac{-x-\hat{x}}{\sigma_x}\right) = e^{2x\hat{x}/\sigma_x^2} = e^{\psi x}$, where $\psi = \tau_p/ (\tau_p + \beta)$, we obtain

$$\int_{-\infty}^{\infty} g(x) d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) = \int_{0}^{\infty} g(x) \left(1 - e^{-\psi x}\right) d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) > 0,$$

and

$$\int_{-\infty}^{\infty} g(x) d\Phi\left(\frac{x-\bar{x}}{\sigma_x}\right) = \int_{0}^{\infty} g(x) \left(1 - e^{-\psi x}\right) d\Phi\left(\frac{x-\bar{x}}{\sigma_x}\right) > 0.$$

To show that $R^P < R^{PI}$, we find, after integration by parts, and exploiting the fact that $g'(x) = g'(-x)$:

$$R^P = 1 - \int_{0}^{\infty} g(x) d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) - \int_{-\infty}^{0} g(x) d\Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) = \int_{0}^{\infty} \Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) dg(x) + \int_{-\infty}^{0} \Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) dg(x) = \int_{0}^{\infty} \left\{ \Phi\left(\frac{x-\hat{x}}{\sigma_x}\right) + 1 - \Phi\left(\frac{x+\hat{x}}{\sigma_x}\right) \right\} dg(x).$$

Likewise, since

$$R^{PI} = \int_{0}^{\infty} \left\{ \Phi\left(\frac{x-\bar{x}}{\sigma_x}\right) + 1 - \Phi\left(\frac{x+\bar{x}}{\sigma_x}\right) \right\} dg(x),$$

it suffices to check that

$$\Phi\left(\frac{x+\bar{x}}{\sigma_x}\right) - \Phi\left(\frac{x-\bar{x}}{\sigma_x}\right) < \Phi\left(\frac{x+\hat{x}}{\sigma_x}\right) - \Phi\left(\frac{x-\hat{x}}{\sigma_x}\right).$$

Since $\bar{x}/\sigma_x < \hat{x}/\sigma_x$, it is immediate that

$$\Phi\left(\frac{x+\bar{x}}{\sigma_x}\right) - \Phi\left(\frac{x-\bar{x}}{\sigma_x}\right) < \Phi\left(\frac{x+\hat{x}}{\sigma_x}\right) - \Phi\left(\frac{x-\hat{x}}{\sigma_x}\right).$$

Furthermore, since $1/\sigma_x > 1/\sigma_x$, and $\Phi\left(\frac{x}{\sigma_x} + \frac{\hat{x}}{\sigma_x}\right) - \Phi\left(\frac{x}{\sigma_x} - \frac{\hat{x}}{\sigma_x}\right)$ is decreasing in $1/\sigma_x$ for $x > 0$, it follows that

$$\Phi\left(\frac{x}{\sigma_x} + \frac{\hat{x}}{\sigma_x}\right) - \Phi\left(\frac{x}{\sigma_x} - \frac{x}{\sigma_x}\right) < \Phi\left(\frac{x}{\sigma_x} + \frac{\hat{x}}{\sigma_x}\right) - \Phi\left(\frac{x-\hat{x}}{\sigma_x}\right).$$

Now, for the limiting results, notice that as $\gamma \sigma_u \to 0$, $\hat{x} \to \frac{1}{2} \tau_p = \bar{x}$, $\sigma_x^2 = (\gamma \sigma_u + \sqrt{\tau_p})^2 \to \tau_p$ and $\sigma_x^2 = \tau_p + \gamma^2 \sigma_u^2 \to \tau_p$, and therefore $\lim_{\gamma \sigma_u \to 0} R^P = \lim_{\gamma \sigma_u \to 0} R^{PI} = R$, where

$$R = 1 - \int_{-\infty}^{\infty} g(x) d\Phi\left(\frac{x-\frac{1}{2} \tau_p}{\sqrt{\tau_p}}\right).$$
Proof of Proposition 4:
Since \( \log R_\pi (\theta, z) = \log \pi (\theta) - \log P_\pi (z) \), we have

\[
\mathbb{E} (\log R_\pi (\theta, z)) = \mathbb{E} (\log \pi (\theta)) - \mathbb{E} (\log P_\pi (z))
\]

\[
= \mathbb{E} (\log \pi (\theta)) - \mathbb{E} \left( \mathbb{E} (\log \pi (\theta) \mid x = z, z) \right) + \mathbb{E} \left( \mathbb{E} (\log \pi (\theta) \mid x = z, z) \right) - \mathbb{E} (\log P_\pi (z)).
\]

The term \( \mathbb{E} (\log \pi (\theta)) - \mathbb{E} \left( \mathbb{E} (\log \pi (\theta) \mid x = z, z) \right) \) can be re-written along the same lines as lemma 1 (with \( \log (\pi (\theta)) \) taking the place of \( \log (\pi (\theta)) \)) to arrive at

\[
\mathbb{E} (\log \pi (\theta)) - \mathbb{E} \left( \mathbb{E} (\log \pi (\theta) \mid x = z, z) \right) = \int_0^\infty \left( \frac{\pi'(\theta)}{\pi(\theta)} - \frac{\pi'(-\theta)}{\pi(-\theta)} \right) \left( \Phi \left( \frac{\theta}{\sigma_\pi} \right) - \Phi \left( \frac{\theta}{\sigma_\theta} \right) \right) d\theta.
\]

The second term is the unconditional expectation of \( \mathbb{E} (\log \pi (\theta) \mid x = z, z) - \log P_\pi (z) \), or equivalently the market-implied expectation of \( \log \text{returns} \mathbb{E} (\log R_\pi (\theta, z) \mid x = z, z) \).

Proof of Proposition 5:
If \( \sigma_{P,1} = \sigma_{P,2} = \sigma_P \), then \( \Delta_{\pi_1} (\sigma_{P,1}) + \Delta_{\pi_2} (\sigma_{P,2}) = \Delta_\pi (\sigma_P) \), and hence the total expected revenue is not affected by the split. If instead \( \sigma_{P,1} \neq \sigma_{P,2} \), then by Theorem 2, \( \Delta_{\pi_1} (\sigma_{P,1}) + \Delta_{\pi_2} (\sigma_{P,2}) > \Delta_{\pi_1} (\sigma_{P,2}) + \Delta_{\pi_2} (\sigma_{P,1}) \), whenever \( \sigma_{P,2} > \sigma_{P,1} \), since \( \pi_2 \) has more upside risk than \( \pi_1 \).

Proof of Proposition 6:
For any alternative split \((\pi_1, \pi_2)\), the monotonicity requirements imply that \( 0 \leq \pi'_1 (\theta) = \pi'_2 (\theta) \leq \pi'_1 (\theta) \). This in turn implies that for all \( \theta \geq 0 \), \( \pi'_1 (\theta) - \pi'_1 (-\theta) = -\pi' (-\theta) \leq \pi'_1 (\theta) - \pi'_1 (-\theta) \) and \( \pi'_2 (\theta) - \pi'_2 (-\theta) = \pi' (\theta) \geq \pi'_2 (\theta) - \pi'_2 (-\theta) \), i.e. \( \pi_1 \) has less downside risk and more upside risk than \( \pi_1 \), and \( \pi_2 \) has more downside risk and less upside risk than \( \pi_2 \). Moreover,

\[
(\pi'_1 (\theta) - \pi'_1 (-\theta)) + (\pi'_2 (\theta) - \pi'_2 (-\theta)) = \pi' (\theta) - \pi' (-\theta) = (\pi'_1 (\theta) - \pi'_1 (-\theta)) + (\pi'_2 (\theta) - \pi'_2 (-\theta)).
\]

But then, the expected revenue of selling \( \pi_1 \) to the investor pool with \( \sigma_{P,1} \) and \( \pi_2 \) to the investor pool with \( \sigma_{P,2} \) is \( \Delta_{\pi_1} (\sigma_{P,1}) + \Delta_{\pi_2} (\sigma_{P,2}) = \Delta_\pi (\sigma_{P,1}) + \Delta_{\pi_2} (\sigma_{P,2}) - \Delta_{\pi_2} (\sigma_{P,1}) \), while the expected revenue from selling \( \pi_1^* \) to the investor pool with \( \sigma_{P,1} \) and \( \pi_2^* \) to the investor pool with \( \sigma_{P,2} \) is \( \Delta_{\pi_1} (\sigma_{P,1}) + \Delta_{\pi_2}^* (\sigma_{P,2}) = \Delta_\pi (\sigma_{P,1}) + \Delta_{\pi_2}^* (\sigma_{P,2}) - \Delta_{\pi_2}^* (\sigma_{P,1}) \). The difference in revenues is therefore \( \Delta_{\pi_2} (\sigma_{P,2}) - \Delta_{\pi_2}^* (\sigma_{P,1}) - (\Delta_{\pi_2} (\sigma_{P,2}) - \Delta_{\pi_2} (\sigma_{P,1})) \), which is positive, since \( \pi_2^* \) contains more upside and less downside risk than \( \pi_2 \), and \( \sigma_{P,2} \geq \sigma_{P,1} \) (Theorem 2, part (iv)).