DISCUSSION PAPER SERIES

No. 9291

A THEORY OF ASSET PRICES BASED ON HETEROGENEOUS INFORMATION

Elias Albagli, Christian Hellwig and Aleh Tsyvinski

FINANCIAL ECONOMICS and INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9291.asp**

ISSN 0265-8003

A THEORY OF ASSET PRICES BASED ON HETEROGENEOUS INFORMATION

Elias Albagli, USC Marshall
Christian Hellwig, Toulouse School of Economics and CEPR
Aleh Tsyvinski, Yale University

Discussion Paper No. 9291
January 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9291

January 2013

ABSTRACT

A Theory of Asset Prices based on Heterogeneous Information*

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 cash flow expectations. From an ex ante perspective, this information aggregation wedge leads to a systematic gap between an asset's expected price and its expected dividend. The sign and magnitude of this expected wedge depend on the asymmetry between upside and downside payoff risks and on the importance of information heterogeneity. We consider three applications of our theory. We first show that predictions of our model provide a novel theoretical justification and are quantitatively consistent with documented empirical regularities on negative relationship between returns and skewness. Second, we illustrate how heterogeneous information leads to systematic departures from the Modigliani-Miller theorem and provide a new theory of debt versus equity. Third, we provide conditions under which permanent over- or under-pricing of assets is sustainable in a dynamic version of our model.

JEL Classification: D82, D84, G12 and G14
Keywords: asset prices, information aggregation, Modigliani-Miller theorem and skewness

Elias Albagli
Finance and Business Economics
USC Marshall School of Business
University of Southern California
Los Angeles, CA 90089-0808
USA

Email: albagli@marshall.usc.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=173450

Christian Hellwig
Toulouse School of Economics
IDEI
Office MF 512
Manufactures des Tabacs
21 Allées de Brienne
31000 Toulouse
FRANCE

Email: christian.hellwig@tse-fr.eu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=161456

Aleh Tsyvinski
Department of Economics
Yale University
28 Hillhouse Ave., Rm. 201
New Haven, CT 06520-8268
US

Email: a.tsyvinski@yale.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=166763

*We thank Bruno Biais, John Geanakoplos, Narayana Kocherlakota, Felix Kubler, Stephen Morris, Guillaume Plantin, Jeremy Stein, Jean Tirole, Dimitri Vayanos, Xavier Vives, Martin Weber, Eric Young, and audiences at Berkeley Haas, Chicago Fed, Columbia, EUI, HEC (Paris), IIES (Stockholm), LBS, LSE, Mannheim, Minneapolis Fed, Munich, Northwestern, NYU, Rice, UCL, UCLA, Yale, the 2nd French Macro-Finance Summer workshop (Sciences Po.), the Hydra workshop (Ajaccio), NBER EFG meeting, and ESSET Gerzensee 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.

Submitted 08 January 2013

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. With only minimal restrictions on security payoffs, trader preferences and the information structure, noisy aggregation of heterogeneous information leads to a systematic gap between a security's equilibrium price, and its corresponding fundamental value. This gap, which we term the *information aggregation wedge*, leads to novel implications and sharp predictions that link the asset's predicted returns to features of the market environment and the distribution of the underlying cash-flow risk. We consider three applications of our theory. First, we show how theoretically and quantitatively predictions of a version of our model are consistent with documented regularities on the negative return to skewness. Second, we reconsider the Modigliani-Miller theorem, and show how dispersed information provides a rationale for tranching a cash flow into debt and equity, regardless of other incentive or tax considerations. Third, we discuss sustainability of mispricing in a dynamic environment with limits to arbitrage.

We consider an asset market along the lines of Grossman and Stiglitz (1980), Hellwig (1980), and Diamond and Verrecchia (1981).[1] An investor pool is divided into 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 state.

Our first main result, in section 2, offers a general characterization of the equilibrium security price, while imposing virtually no restrictions on the asset's distribution of cash flows, the market environment, the trader preferences and the information structure. This price is shown to be generically different from the security's expected "fundamental value", which is defined as the dividend expectation conditional on the information conveyed through the price. The price and the expected dividend both incorporate the information that is conveyed through the price in equilibrium, but in addition, the price must adjust to the fundamental shocks to satisfy the market-clearing condition. Even without considering the inference drawn from the price, 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 an increase in the equilibrium price in order to clear the market. This direct market-clearing effect present only in the determination of the price is then compounded by the information that is conveyed by the price increase. This informational effect

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

is reflected both in the price and in the dividend expectations. The market clearing effect is the reason why the price responds more strongly to fundamentals than dividend expectations.

To arrive at our formal characterization, notice that traders in the market update expectations based on their private signal and the information coming from the price. If information is sufficiently dispersed, then for any price, there exists a threshold trader who finds it optimal in equilibrium neither to purchase, nor to short-sell any quantity of the asset at this price. Under mild regularity conditions, this trader's private signal increases monotonically with the price (and hence with the underlying shocks). We can thus define the value of this private signal as a "sufficient statistic" for the information content of the price, and use it to characterize the equilibrium price as a function of this sufficient statistic. Since a risk-averse investor demands zero if and only if his dividend expectation equals the price, the equilibrium price must equal the threshold investor's dividend expectation. In this characterization, the market-clearing effect of an increase in the payoff fundamental or the noise trader's demand appears through a shift in the marginal trader's private signal, while the informational effect appears through the direct effect of prices on trader's posterior expectations. The characterization of price and expected dividend as two different posterior expectations that are conditioned on the same sufficient statistic makes the subsequent analysis particularly tractable.

In section 3, we introduce additional structure on preferences and signal distributions to derive the implications of this general characterization for unconditional asset returns. Using a market structure first introduced in Hellwig, Mukherji and Tsyvinski (2006), we assume that traders are risk neutral but face limits on their asset positions. With these assumptions, we are able to characterize the sufficient statistic variable, the equilibrium price, and the expected fundamental value in closed form, still without imposing any restrictions on the asset's payoff risk. From an ex ante perspective we characterize the unconditional (expected) information aggregation wedge, defined as the difference between the expected price an the expected dividends, as a function of the cash flow distribution and a parameter summarizing information frictions. This parameter depends on the accuracy of informed traders' private signals and the variance of noise trading shocks.

Our second main result uses this characterization to link the asset's expected return to information frictions and the asymmetry between upside and downside risks. This asymmetry is defined as a partial order on payoff risks that compares the marginal gains and losses at fixed distances from the prior mean of the fundamental. The expected information aggregation wedge is characterized as a mean-preserving spread of the underlying distribution of fundamentals. That is, from an ex ante perspective, the price puts a higher weight on the tails than the distribution of the fundamental. This spread is larger the more important the information frictions are. It affects prices positively when the asset's payoff risk is dominated by the upside, and negatively when dominated by the downside. Therefore, regardless of the informational parameters, the expected wedge is zero when payoff risk is symmetric. The expected wedge is positive (meaning that the expected price exceeds expected dividends) for risks that are dominated by the upside, and is negative for risks that are dominated by the downside. Moreover, in absolute value this expected wedge becomes more pronounced for more asymmetric payoff risks, or for a higher degree of information aggregation frictions, and due to an increasing difference property, the impact of payoff asymmetry is more pronounced as information frictions become larger, and vice versa. We end the section deriving implications for the expected returns and with providing a particularly simple closed form example.

In section 4, we develop three applications of the model. As a first application, we show how our model is consistent with the empirical predictions on a negative relationship between returns and skewness theoretically and quantitatively. Conrad, Dittmar and Ghysels (forthcoming), Boyer, Mitton and Vorkink (2010) and Green and Hwang (2012) establish this relationship for different measures of skewness: securities with higher skewness earn lower returns. We use our model to price a security that matches the skewness and volatility properties of the different stock portfolios in Conrad et al. (forthcoming). We then compute the securities' expected log-returns in our model, for different values of the information friction parameters. We establish that even with the smallest level of frictions, the premium for skewness falls in the range suggested by the empirical findings. In other words, even very moderate information frictions can generate sizeable and empirically plausible excess returns from skewness. Several explanations were proposed for the empirical findings: optimal expectations in Brunnermeier and Parker (2005) and Brunnermeier, Gollier and Parker (2007); cumulative prospect theory implying preference for lottery-like features in Barberis and Huang (2008); heterogeneous preferences for skewness in Mitton and Vorkink (2007). Our model provides a novel theoretical justification for the negative relationship between the skewness and returns based on heterogenous information in a rational expectations equilibrium. The explanation does not rely on non-rational expectations, behavioral phenomena, or on heterogeneity in preferences.

As the second application we consider how a seller of a cash-flow may exploit information heterogeneity to influence asset's market value by tranching the cash-flow and selling it to different investor pools. We show that the seller's expected revenue is not affected by the split, if and only if the different investor pools have identical informational characteristics. However, when the investor pools differ, the seller can manipulate her expected revenue by selling downside risks in the market with smaller information aggregation frictions, and upside risks in the market with larger information aggregation frictions. The seller maximizes expected revenue by completely separating 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. These results offer a new perspective on the Modigliani-Miller Theorem, which establishes that under conditions of no arbitrage, the total market value of a firm is not influenced by how it is divided into separate securities, and hence its optimal capital structure is indeterminate, absent any distortions in the generation of cash flows (Modigliani and Miller, 1958). Capital structure theories then focus mostly on trade-offs that affect the generation of cash flows inside the firm, such as agency costs, information frictions or tax distortions, assuming that the market value of the resulting cash flow is not affected by its split into different securities. Here instead we take the view that firm value may also be influenced by heterogeneous information, and show how these frictions affect the optimal design of securities and capital structure.

As the third application, we consider an infinitely repeated version of our trading model and give conditions under which a security may be permanently over- or under-priced, regardless of current market conditions. As is well known, under no-arbitrage conditions securities cannot permanently – and not even temporarily – deviate from their fundamental values (Tirole, 1982; Santos and Woodford, 1997). While the anticipation of higher future prices would, in principle, induce agents to increase the price bid in the current period, the combination of no arbitrage with transversality conditions (or backwards induction, in case of assets with finite horizons) rules out the possibility of any security trading at a price that exceeds the net present value of expected future cash flows. With heterogeneous information, the anticipation of higher future prices for securities that are viewed as upside risks induces traders to always bid for prices in excess of current dividend expectations. The asset then trades at a premium over its expected dividend value regardless of the current state realization. By the same argument, securities that are dominated by downside risk may be permanently underpriced.

Our paper contributes to the large literature on noisy information aggregation in asset markets in two ways in particular. First, whereas most of the existing literature imposes strong parametric assumptions (such as CARA preferences, or normally distributed signals and dividends) to arrive at closed-form solutions, our main equilibrium characterization is almost completely free of assumptions about the underlying primitives, and thereby enables us to identify at a general level how noisy information affects asset prices and returns.[2] Second, the variant of our model with

[2]To our knowledge, Vives (2008) is the only written statement of an observation that an information aggregation

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, 2012) also explores non-linear and non-normal variants of the noisy REE framework, but he still relies on important distributional restrictions of asset payoffs and signals to the "exponential family" of distributions. Moreover, he uses these results to analyze incentives for information acquisition, whereas we take the information structure as given and instead focus on implications of information aggregation for expected asset returns.[3]

Another influential literature emphasizes heterogeneous beliefs and short sales constraints as potential sources of bubbles, mis-pricing, and market anomalies (Harrison and Kreps, 1978; Allen, Morris and Postlewaite, 1993; Chen, Hong and Stein, 2002; Scheinkman and Xiong, 2003; Hong and Stein, 2007; Hong and Sraer, 2011). Mispricing is sustained by the option to resell an over-valued security to an even more optimistic buyer in the future. This option becomes valuable in the presence of (one-sided) short-sales constraints, and implies a channel for over-valuation. Heterogeneity in prior beliefs is taken as exogenous, and with the exception of Allen, Morris and Postlewaite (1993), traders do not update from the observation of prices. We touch on similar themes, but stay within the REE tradition in which traders' beliefs result from exogenous signals, and information aggregation through prices imposes tight restrictions on the heterogeneity in beliefs. Furthermore, we do not invoke asymmetric limits to arbitrage, and our market environment is static, so that the resale option does not play an important role, and our model gives rise to over- as well as under-valuation results.[4] In this sense, our results are closer to the findings of REE models with short sale constraints. Diamond and Verrecchia (1987) show that constraining short sales reduce the informational efficiency but do not cause overpricing. Bai, Chang, and Wang (2006) find that limiting short sales driven by risk-sharing results in a higher price, but if these are driven by private information, short sale limits increase the uncertainty for the less informed and may result in lower prices.

wedge is present in the CARA-normal models, and there it is only mentioned in passing. Moreover, under the standard assumptions of normally distributed dividends necessary to solve such models, any unconditional excess return is attributable to a risk premium.

[3]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.

[4]This distinction becomes even more clear when one compares the predictions of the two theories for bond markets: while the heterogeneous priors literature does not draw a big distinction between equity and debt, and the over-pricing mechanisms are similar in both cases, our information-based theory views equity as upside risks and bonds as downside risks, therefore generating distinct pricing implications. See in particular Hong and Sraer (2011) for a heterogeneous priors model of debt bubbles.

More generally, any theory of mispricing must rely on some source of noise affecting the market, coupled with some limits to the traders' ability or willingness to exploit the resulting arbitrage opportunity (see Gromb and Vayanos, 2010, for an overview and numerous references). We show that noise trading under heterogenous information leads not just to random errors in the price, but to systematic, predictable departures of the price from the asset's fundamental value. This observation is independent of the exact nature of the limits to arbitrage, which result from risk aversion and position limits in our case.

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 and a 'Walrasian auctioneer'. The dividend of the risky asset is given by a strictly increasing and twice continuously differentiable function $\pi(\cdot)$ of a stochastic fundamental θ.

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 $x_i = \theta + \varepsilon_i$, where ε_i is i.i.d across traders, and distributed according to cdf. $F : \mathbb{R} \to [0, 1]$ and smooth density function f. We further assume that $f'(\cdot)/f(\cdot)$ is strictly decreasing and unbounded above and below. Monotonicity of $f'(\cdot)/f(\cdot)$ implies that signals have log-concave density and hence satisfy the monotone likelihood ratio property. Unboundedness implies that extreme signal realizations induce large updates in posterior beliefs, (almost) regardless of the information contained in other signals.

Each trader decides how many shares in the asset to purchase at the prevailing price P, in exchange for cash. Formally, trader i submits a price-contingent demand schedule $d_i(\cdot)$. Traders' preferences are characterized by a strictly increasing, concave utility function $U : \mathbb{R} \to \mathbb{R}$, which is defined on the traders' realized gains or losses $d_i \cdot (\pi(\theta) - P)$, when they purchase d_i units at a price P. Traders' positions are restricted to lie on the interval $[d_L(P), d_H(P)]$, where $d_L(P) < 0 < d_H(P)$ are arbitrary, continuous, price-contingent limits.

Individual trading strategies are a mapping $d : \mathbb{R}^2 \to [d_L(P), d_H(P)]$ from signal-price pairs (x_i, P) into asset holdings. Aggregating traders' decisions leads to the aggregate demand by informed traders, $D : \mathbb{R}^2 \to [0, 1]$, $D(\theta, P) = \int d(x, P) dF(x - \theta)$, where $F(x - \theta)$ represents the cross-sectional distribution of private signals x_i conditional on the realization of θ.[5] The supply

[5] We assume that the Law of Large Numbers applies to the continuum of traders, so that conditional on θ the cross-sectional distribution of signal realizations ex post is the same as the ex ante distribution of traders' signals.

of securities is stochastic, and 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 all traders have submitted their orders, the auctioneer selects a price P to clear the market. Formally, let $\hat{P} : \mathbb{R}^2 \to \mathbb{R}$, $\hat{P}(\theta, u) = \{P \in R : D(\theta, P) = S(u, P)\}$, denote the correspondence of market-clearing prices.[6] Then a price function $P : \mathbb{R}^2 \to \mathbb{R}$ clears the market, if and only if $P(\theta, u) \in \hat{P}(\theta, u)$, for all $(\theta, u) \in \mathbb{R}^2$.

Let $H(\cdot|P) : \mathbb{R} \to [0, 1]$ denote the posterior cdf of θ, conditional on observing the market price P. Then the informed traders' posterior is defined from Bayes' Rule as:

$$H(\theta|x, P) = \frac{\int_{-\infty}^{\theta} f(x - \theta')\, dH(\theta'|P)}{\int_{-\infty}^{\infty} f(x - \theta') dH(\theta'|P)}.$$

The traders' decision problem is stated as $\max_{d \in [d_L(P), d_H(P)]} \int U(d(\pi(\theta) - P))\, dH(\theta|x, P)$. The corresponding first-order condition is $\int (\pi(\theta) - P) \cdot U'(d(\pi(\theta) - P)) dH(\theta|x, P) = 0$.

A *Perfect Bayesian Equilibrium* consists of demand functions $d(x, P)$ for informed traders, a price function $P(\theta, u)$, 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 (θ, 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.[7]

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.[8]

Theorem 1 (Equilibrium Characterization) *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. Then, the following two conditions are equivalent:*

(1) *Demand $d(x, P)$ is strictly decreasing in P, whenever $d(x, P) = 0$.*

(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*[9]

$$P_\pi(z) = \mathbb{E}(\pi(\theta)|x = z, z) = \frac{\int \pi(\theta) f(z - \theta)\, \psi(z|\theta) h(\theta)\, d\theta}{\int f(z - \theta) \psi(z|\theta) h(\theta) d\theta}. \tag{1}$$

[6] We can without loss of generality restrict the range of $P(\cdot)$ to coincide with the range of $\pi(\cdot)$.

[7] Implicitly, this definition assumes that traders agree on their update about θ from the price, $H(\cdot|P)$. While this is obviously true along the equilibrium path, the assumption also applies for prices not observed in equilibrium.

[8] Existence will be 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.

[9] We index an equilibrium function or variable by π 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 (1).

Thus, under a weak regularity condition on posterior beliefs, any equilibrium with a demand function that crosses zero monotonically in P admits a sufficient statistic representation. That is, there exists a random variable z that is only a function of θ and u, and contains the same information as the price. Moreover, by means of this sufficient statistic z, the price can be represented in the form given by (1).

The proof of Theorem 1 proceeds in several steps. First, if $h(\cdot|P)$ is continuous, log concavity of f, and the unboundedness of f'/f implies that $\mathbb{E}(\pi(\theta)|x, P)$ is increasing in x, and $\lim_{x\to-\infty} \mathbb{E}(\pi(\theta)|x, P) < P < \lim_{x\to\infty} \mathbb{E}(\pi(\theta)|x, P)$, for any P that is in the interior of the support of $\pi(\cdot)$. This in turn implies that there exists a unique critical private signal $z(P)$ such that $\mathbb{E}(\pi(\theta)|z, P) = P$. We can now use the equilibrium price function $P(\theta, u)$ to construct the random variable $z(\theta, u) = z(P(\theta, u))$. Second, $d = 0$ is optimal for a trader if and only if $\mathbb{E}(\pi(\theta)|x, P) = P$ or $x = z(P)$, and any trader with signal $x > z(P)$ will have a positive demand, while any trader with signal $x < z(P)$ will have negative demand. Third, observing z is informationally equivalent to observing P if and only if $z(P)$ is strictly monotone. By construction, this is the case, if and only if $\mathbb{E}(\pi(\theta)|x, P) - P$ is strictly decreasing in P at $x = z(P)$, or equivalently $d(x, P)$ is strictly decreasing in P, whenever $d(x, P) = 0$. Substituting $x = z(P)$ into $P = \mathbb{E}(\pi(\theta)|x, P)$ and using the fact that P is informationally equivalent to z then leads to the characterization in (1).

The theorem derives its interest not just from the result that such a sufficient statistic can be constructed, but from the characterization of equilibrium prices that it entails. The main observation from the characterization theorem is that at the interim stage – when the share price is observed but before dividends are realized – the equilibrium price differs from the expected dividend, conditional on the public information. Specifically, the expected dividends conditional on public information are given by $V_\pi(z) = E(\pi(\theta)|P_\pi(z))$, or equivalently by

$$V_\pi(z) = \mathbb{E}(\pi(\theta)|z) = \frac{\int \pi(\theta)\psi(z|\theta)h(\theta)d\theta}{\int \psi(z|\theta)h(\theta)d\theta}.$$

The difference between the price and the expected dividends, conditional on public information z, is $W_\pi(z) \equiv P_\pi(z) - V_\pi(z)$. The unconditional difference is

$$W_\pi = \int \left\{ \int P(z)d\Psi(z|\theta) - \pi(\theta) \right\} dH(\theta).$$

We refer to W_π as the unconditional and $W_\pi(z)$ as the conditional *"information aggregation wedge"*. Since the equilibrium price $P_\pi(\cdot)$ to the expected dividend value $V_\pi(\cdot)$ both have direct empirical counterparts in any set of price-return data, this comparison allows us to focus directly on the empirical, testable implications of our model.[10]

The key reason for the presence of the conditional information aggregation wedge is that an equilibrium has to satisfy the market clearing condition. The intuition is as follows. Suppose that demand is increasing in x and decreasing in P. Consider an increase in the fundamental θ (the same logic applies for a decrease in u, corresponding to supply contraction). There are two effects. The first is, holding posteriors $\hat{H}(\cdot)$ fixed, an increase in θ would lead to an increase in demand. This increase, for a given realization of u, has to be met by an increase in price in order to clear markets. The reasoning so far follows solely from and is a consequence of market clearing, even without considering agents learning from P.

There is also the second effect – the change in the posterior $\hat{H}(\cdot)$. The traders compute in equilibrium how the shifts in θ (or u) translate into shifts in the equilibrium price and therefore form an updated belief about dividends from observing the price. This update uses Bayes' Rule, and reinforces the first (market-clearing) effect of the price.

In the expression for the price (1) the two effects are represented by the sufficient statistics z appearing twice. In contrast, in the expression for the the expected value of the dividends $V_\pi(z)$, the sufficient statistics appears only once to represent only the second effect. The first effect is absent as there is no requirement for the markets to clear in this expression.

The equilibrium price $P_\pi(z)$ can be interpreted as the expectation of the dividends of the trader whose demand is equal to zero. This trader conditions on the information in the price z, as well as a private signal. The equilibrium requires market clearing and imposes a restriction on the identity of this trader. Specifically, the private signal of this trader is equal to $z(P)$, by construction of the sufficient statistic z. Moreover, the identity of this trader, $z(\theta, u)$, shifts with the underlying shocks θ and u.

Despite its appearance, the information aggregation wedge is not the result of non-Bayesian updating or irrational trading decisions, but results from market clearing with investor heterogeneity and is perfectly consistent with Bayesian Rationality at the level of individual investors.

[10] This benchmark of comparison differs from the benchmarks chosen, e.g. by Harrison and Kreps (1978), who compare an asset's value to the dividend expectation of any trader in their market, or to an average of those expectations, as in Allen, Morris and Shin (2006) and Bacchetta and van Wincoop (2006). Like these antecedents, however, our characterization of the wedge as well as its implications for returns can be understood in terms of a failure of the law of iterated expectations, i.e. generically, $\mathbb{E}(\mathbb{E}(\pi(\theta)|x=z,z)) \neq \mathbb{E}(\mathbb{E}(\pi(\theta)|z)) = \mathbb{E}(\pi(\theta))$.

2.3 Discussion: Assumptions, extensions, and generalizations

Monotonicity of demand. The characterization result in terms of a sufficient statistic z is valid if and only if demand is decreasing in the price at the level where demand is zero, or equivalently, that $\mathbb{E}(\pi(\theta)|x, P)$ cannot increase faster in P than P itself. This condition limits how much an increase in P translates into "good news" about the asset's dividend value. When the monotonicity condition does not hold, we can still identify $z(P)$ as the private signal threshold at which $\mathbb{E}(\pi(\theta)|z, P) = P$, or equivalently $d(z, P) = 0$. With this characterization, the equilibrium price satisfies $P = \mathbb{E}(\pi(\theta)|z(P), P)$, which is generically different from the expected dividend value, $\mathbb{E}(\pi(\theta)|P)$. Therefore, monotonicity of demand is not needed to show the existence of the wedge. However, if the function $z(P)$ is no longer invertible, then z no longer fully summarizes the information content of P, and we are no longer able to use it as a sufficient statistic for the information conveyed through the equilibrium price, which is needed to arrive at the characterization in (1).[11]

Sufficient information dispersion. The assumptions of log-concavity of $f(\cdot)$ and unboundedness of $f'(\cdot)/f(\cdot)$ are needed to show the existence of a threshold trader. They imply that posterior beliefs are first-order stochastically increasing in x, and sufficiently diverse so that in equilibrium some informed traders end up buying the asset, while others end up as (short-)sellers.

Limits to arbitrage assumptions. The above characterization extends with appropriate adjustments to the case where $d_L(P) = 0$ ($d_H(P) = 0$) for some P. In that case, condition 1 of the theorem applies only at $\inf_x \{x : d(x, P) > 0\}$ ($\sup_x \{x : d(x, P) < 0\}$), and for any strictly lower x (strictly higher x) the lower (upper) bound constraints are binding.

Likewise, it is important for the (non zero) wedge to have some restrictions on arbitrage. The first, and most immediate one comes from the informed traders' risk aversion and position limits that restrict how much traders are willing to bet on the private information they have.[12] The second form of arbitrage restrictions, is the absence of uninformed, risk-neutral arbitrageurs that are willing to bet on the difference between $P_\pi(z)$ and $V_\pi(z)$. A simple form to introduce the latter is to adjust the supply assumption so that the number of shares available to informed traders is $S(u, P) = S(u; P - \mathbb{E}(\pi(\theta)|P))$, i.e., to let the supply respond positively to the wedge. We could also include in the model a positive measure of risk-averse, uninformed traders, in which case the

[11] The random variable $z(\theta, u) = z(P(\theta, u))$ is still well defined. However, if the same realization of z is consistent with multiple values of P along the equilibrium path, then the observation of a particular realization of P informs the traders not only of the corresponding value of z, but possibly conveys additional information about (θ, u) through the selection among multiple market-clearing prices. Thus, the information conveyed through z may be coarser than the information communicated through P.

[12] Notice that formally the position limits may, but don't have to be binding for our characterization result.

residual supply available to informed traders will be again a function of the noise trader shock u and the price P (through the demand of uninformed traders). We analyzed some special cases with this feature in an earlier working paper version (Albagli, Hellwig, and Tsyvinski, 2011a). While this does not alter the construction of the sufficient statistic, the resulting wedge becomes smaller, since the uninformed traders' positions will partially offset the price impact of the informed traders. The wedge is largest, when supply is completely exogenous and inelastic.

Value of information. Note that traders value having access to the private signal, and earn strictly higher expected returns on average than if they were uninformed. From an ex ante perspective, the value of information depends on the expected utility of an average informed trader. To see this notice simply that an uninformed trader in our model will set $d(P) \in \arg\max_{d\in[d_L(P),d_H(P)]} \mathbb{E}\left(U\left(d(\pi(\theta)-P)\right)|z\right)$, which using the Law of Iterated Expectations could also be written as $d(P) \in \arg\max_{d\in[d_L(P),d_H(P)]} \mathbb{E}\left[\mathbb{E}\left(U\left(d(\pi(\theta)-P)\right)|x,z\right)|z\right]$. The resulting expected utility is always strictly less than what an informed trader obtains by making d contingent on x as well as P: $\mathbb{E}\left[\arg\max_{d\in[d_L(P),d_H(P)]} \mathbb{E}\left(U\left(d(\pi(\theta)-P)\right)|x,z\right)|z\right]$.

The role of payoff non-linearities. Our characterization of equilibrium prices is almost completely free of assumptions about the security's payoffs, and can therefore be used to derive return implications for broad security classes. Furthermore, this general characterization suggests some simple conjectures about the shape of the information aggregation wedge – namely that $P_\pi(z)$ responds more to z than $V_\pi(z)$, due to the role of market clearing, so that $W_\pi(z)$ is monotone in z; or that at least it crosses 0 in a single location, and is negative for low z and positive for high z. Such conjectures however do not hold generally without strong additional restrictions on the shape of $\pi(\cdot)$ or signals and distributions.

The comparison between the price and the expected dividends includes several effects. Using a second-order Taylor expansion of $\pi(\theta)$ around $\theta = \mathbb{E}(\theta|x=z,z)$ for $P_\pi(z)$ and around $\theta = \mathbb{E}(\theta|z)$ for $V_\pi(z)$, we obtain

$$\begin{aligned} W_\pi(z) \approx\ & \pi\left(\mathbb{E}(\theta|x=z,z)\right) - \pi\left(\mathbb{E}(\theta|z)\right) \\ & + \frac{1}{2}\pi''\left(\mathbb{E}(\theta|x=z,z)\right)\mathbb{V}(\theta|x=z,z) - \frac{1}{2}\pi''\left(\mathbb{E}(\theta|z)\right)\mathbb{V}(\theta|z). \end{aligned}$$

The price differs from the expected dividends in two respects. First, the additional conditioning of the posterior on z in the price alters the response of the expectations about θ. This is captured by the first term in this decomposition. Second, the residual uncertainty in the price (after observing z) is (on average) lower. This is captured by the second line of the decomposition.

When $\pi(\cdot)$ is linear, the effect through expectations is the only term influencing the wedge, while residual uncertainty plays no role. For non-linear dividends, however, residual uncertainty shifts the wedge, as the curvature in dividends impacts expected prices and dividends differently. In particular, for z s.t. $\mathbb{E}(\theta|x=z,z)=\mathbb{E}(\theta|z)$, residual uncertainty is the only factor determining the wedge, and if having access to an additional signal reduces residual uncertainty, then the residual uncertainty implies a negative wedge if $\pi''(\cdot)>0$, and a positive wedge if $\pi''(\cdot)<0$. More generally, third- and higher derivatives may modify these comparative statics so that it is impossible to offer precise results on the shape of $W_\pi(\cdot)$ without additional restrictions.[13]

Generalizations. The only property of demand that we have exploited to arrive at the characterization is that a trader's asset demand is zero when the price equals the trader's dividend expectation, and that demand is monotone in P. This suggests that the above method of characterizing the equilibrium is even more general than what is suggested here. For example the same characterization still obtains when there is arbitrary heterogeneity across agents in preferences (provided all are weakly risk averse). The reason is that the point at which their demand is zero only depends on their expected return from holding the asset, and not on the shape of U.

By a similar argument the model can also be extended to allow for background risks. With background risk, an agent's net gains are given by $w^i+d(x,P)(\pi(\theta)-P)$, where w^i is stochastic. Defining the sufficient statistic z in the same way as above by $d(z,P)=0$, and proceeding along the same lines as Theorem 1, the price is characterized as

$$P_\pi(z)=\frac{\mathbb{E}\left(U'\left(w^i\right)\cdot\pi(\theta)\,|x=z,z\right)}{\mathbb{E}\left(U'\left(w^i\right)|x=z,z\right)}=\mathbb{E}\left(\pi(\theta)\,|x=z,z\right)+\frac{cov\left(U'\left(w^i\right);\pi(\theta)\,|x=z,z\right)}{\mathbb{E}\left(U'\left(w^i\right)|x=z,z\right)}.$$

The characterization thus includes a risk-adjustment to account for the background risk. Whenever the background risk is uncorrelated with the asset payoff ($cov(w^i,\theta)=0$), the risk-adjustment cancels and we just recover equation (1). But when the risks are correlated, the co-movement of the asset with the background risk enters as an additional factor in the determination of prices: in particular a positive comovement ($cov(w^i,\theta)>0$) leads to lower prices and higher returns.

The role of risk preferences: alternative characterizations. As in any canonical asset pricing model, risk preferences play an important role in our equilibrium characterization, even if this role may not be immediately apparent in equation (1). Risk preferences enter through market-clearing in the construction of the sufficient statistic. They affect the characterization of the wedge in two ways. First, if traders need to be compensated for the risk of holding a security in positive net supply, the trader who finds it optimal to hold zero assets must be less optimistic than the

[13] In special cases, it is easy to illustrate all these possibilities. The earlier working paper version (Albagli, Hellwig, Tsyvinski, 2011a) provides a formal discussion and examples.

average trader in the market. Second, risk preferences affect the traders' willingness to trade on their private information, which influences the informativeness of the price signal. This appears through the distribution of the sufficient statistic z.[14]

It's possible to link and compare our asset price representation to the canonical risk-based asset pricing models with homogeneous information in a straightforward manner. To this end, suppose that at the equilibrium we are considering $d(x,P)$ is increasing in x and decreasing in P, and reaches the position limits at the extremes. In this case, for any asset supply level $S \in [d_L(P), d_H(P)]$, we can define a different sufficient statistic variable $z_S(P)$ by setting $d(z,P)=S$, each with its own distribution $\Psi_S(z'|\theta)$. The corresponding pricing equation is:

$$\begin{aligned} P_\pi(z) &= \frac{\mathbb{E}\left(U'\left(S\left(\pi(\theta)-P_\pi(z)\right)\right)\cdot\pi(\theta)\,|x=z,z\right)}{\mathbb{E}\left(U'\left(S\left(\pi(\theta)-P_\pi(z)\right)\right)|x=z,z\right)} \\ &= \mathbb{E}\left(\pi(\theta)\,|x=z,z\right)+\frac{cov\left(U'\left(S\left(\pi(\theta)-P_\pi(z)\right)\right);\pi(\theta)\,|x=z,z\right)}{\mathbb{E}\left(U'\left(S\left(\pi(\theta)-P_\pi(z)\right)\right)|x=z,z\right)} \end{aligned}$$

and must be solved explicitly for the price $P_\pi(z)$. Of course all these representations are equivalent, since they are all based on monotone transformations of the equilibrium price, and they are all equivalent to equation (1), which was based on $S=0$. In particular, by setting S instead equal to the prior expectation of the asset supply: $\bar{S}=\mathbb{E}(S)$, we obtain a characterization that offers a decomposition between the informational component of the wedge and a risk premium component. The same characterization can be taken one step further by setting S equal to the expected asset supply, conditional on P, $\bar{S}(P)=\int S(u,P)\,dG(u|P)$, to account for the fact that the informed traders' exposure to the risk varies with the price. Equivalently, we can express $P_\pi(z)$ as

$$\begin{aligned} P_\pi(z) &= \mathbb{E}\left(\pi(\theta)\cdot m_R(\theta,z)\cdot m_I(\theta,z)\,|z\right), \text{ where} \\ m_R(\theta,z) &= \frac{U'\left(\bar{S}\left(P_\pi(z)\right)\left(\pi(\theta)-P_\pi(z)\right)\right)}{\mathbb{E}\left(U'\left(\bar{S}\left(P_\pi(z)\right)\left(\pi(\theta)-P_\pi(z)\right)\right)|x=z,z\right)} \text{ and} \\ m_I(\theta,z) &= \frac{\psi(\theta|x=z,z)}{\psi(\theta|z)}=f(z-\theta)\cdot\frac{\int\psi(z|\theta')h(\theta')d\theta'}{\int f(z-\theta')\,\psi(z|\theta')h(\theta')d\theta'} \end{aligned}$$

factor in the risk adjustment, and the informational aggregation adjustment, respectively. Thus, conditional on z, the informational adjustment reweights realizations of θ according to the private signal density $f(z-\theta)$.

[14] The role of risk preferences can easily be seen in the well-known model with CARA preferences and normally distributed signals and dividends, which is a special case of our model. It is easy to provide a solution to the canonical CARA-Normal REE model using the methodology developed here and decompose the wedge into a level adjustment that comes from a risk premium, and an adjustment of the slope (with respect to the sufficient statistic z), which comes from the information aggregation channel. See Vives (2008), or our working paper version (Albagli, Hellwig, Tsyvinski, 2011a) for details.

Approximating the no-arbitrage model. Our characterization recovers the equilibrium characterization of the no-arbitrage model in the limit where private signals become infinitely noisy, or formally, $F(\varepsilon) \to 1/2$ and $f(\varepsilon)/f(\varepsilon') \to 1$, for all ε and ε'. Consider the formulation above, in which the sufficient statistic was constructed with reference to an endogenous asset supply, $\bar{S}(P)$. It then follows immediately that the information aggregation factor in the above formula, $m_I(\theta, z)$, converges to 1 almost everywhere, and therefore the equilibrium price only includes the usual risk adjustment $P_\pi(z) = \mathbb{E}(\pi(\theta) \cdot m_R(\theta, z)|z)$. In the limit, the traders' posterior beliefs, and hence their demand functions $d(x, P)$, are independent of x, and market demand $D(\theta, P)$ no longer varies with θ, and converges to some limit demand function $\bar{D}(P)$. Several possibilities then arise, depending on how the distribution of supply shocks converges along with the noise in private signals:

(*i*) If S remains stochastic, then P becomes completely uninformative about θ in the limit. P only responds to supply shocks. The equilibrium price then implicitly solves P as a function of u: $P_\pi(u) = \mathbb{E}(\pi(\theta) \cdot m_R(\theta, u))$, where $m_R(\theta, u) = U'(S(u, P)(\pi(\theta) - P)) / \mathbb{E}(U'(S(u, P)(\pi(\theta) - P)))$. The asset price fluctuates because changes in S vary the exposure of traders to dividend risk, but these fluctuations contain no information about θ. Prices and realized dividends are completely uncorrelated.

If $S(u, P)$ converges in probability to a constant function $S(P)$, then the behavior of prices at the limit depends on the order of limits (or equivalently, the relative rates of convergence).

(*ii*) If $S(u, P)$ converges to $S(P)$ at a lower rate than $D(\theta, P)$ converges to $\bar{D}(P)$, then P remains completely uninformative about (θ, u) at the limit. In this case, the price converges to a constant, regardless of (θ, u), and solves $P_\pi = \mathbb{E}(\pi(\theta) \cdot m_R(\theta))$, with $m_R(\theta)$ is defined as the limit of $m_R(\theta, u)$, as $S(u, P) \to S(P)$. In this limit, the price remains disconnected from future dividends, yet at the same time, stochastic exposure as a source of price fluctuations disappears as well.

(*iii*) If $S(u, P)$ converges to $S(P)$ at a faster rate than $D(\theta, P)$ converges to $\bar{D}(P)$, then P becomes perfectly informative of θ in the limit. In this case, $P_\pi(z) \to \pi(\theta)$ in the limit, i.e. all pricing errors disappear and the security becomes perfectly risk-free.

(*iv*) If $S(u, P)$ converges to $S(P)$ at the same rate as $D(\theta, P)$ converges to $\bar{D}(P)$, then z must remain partially informative in the limit. In this case, asset price fluctuations partially respond to the information in z. The characterization includes a risk adjustment to account for the residual uncertainty in dividends. In this case, prices and dividend values are positively but imperfectly correlated even in the limit as noise-trading vanishes.

Therefore, our model embeds the standard common information, no arbitrage model as a limiting case, as private signals become uninformative. In this limit, traders take identical positions regardless of their signal realizations, and the price remains informative of θ only if supply shocks disappear sufficiently fast at the same time as private signals become uninformative. The different limiting cases are all special cases of the no arbitrage model - if supply shocks vanish, then everything from perfect to no learning from the price is possible depending on the speed of convergence. If supply shocks don't vanish then the limiting scenario allows for price fluctuations that are due to stochastic exposure, and that are orthogonal to the asset's fundamental value.

3 The Risk-neutral, Normal Model

We now specialize the general model to further simplify the equilibrium characterization. We first consider the *risk-neutral model* in which agents are risk-neutral, are restricted to hold at most one unit, and are not allowed to short-sell. In this case, all traders with signals $x < z(P)$ will hold 0 in equilibrium, while any trader with signal $x > z(P)$ will hold 1. This specification makes the aggregation of demand and the construction of the sufficient statistic z particularly simple.

We then introduce restrictions that allow us to retain the tractability of normal updating. We term this the *risk-neutral normal model.* In contrast to the CARA-normal model, we do not impose any additional restrictions on the shape of dividends. The risk-neutral, normal model therefore allows us to discuss expected return implications of heterogeneous information, for very general classes of securities. Moreover, by abstracting from risk preferences, this model isolates the information aggregation channel in a particularly clear fashion.

3.1 Characterizing the Risk-Neutral and Risk-Neutral, Normal models

For the model with risk-neutral preferences, we assume that $U(w) = w$, $d_L(P) = 0$, $d_H(P) = 1$, and that supply is inelastic: $S(u, P) = u$. A trader's expected value of holding the asset is $\int \pi(\theta)\, dH(\theta|x, P)$. Under log concavity of f and unboundedness of f'/f, this expectation is an increasing function of x (Milgrom, 1981a), and there exists a unique $\hat{x}$, such that $\int \pi(\theta)\, dH(\theta|\hat{x}, P) = P$. The traders' decisions are therefore characterized by a signal threshold function $\hat{x}: \mathbb{R} \to \mathbb{R}$, such that

$$d(x, P) \begin{cases} = 0 \text{ if } x < \hat{x}(P) \\ \in [0, 1] \text{ if } x = \hat{x}(P) \\ = 1 \text{ if } x > \hat{x}(P) \end{cases} .$$

Aggregating the individual demands, we find the market demand $D(\theta, P) = 1 - F(\hat{x}(P) - \theta)$. Solving the market-clearing condition $D(\theta, P) = u$ then yields $F(\hat{x}(P) - \theta) = 1 - u$, which has to hold for any (θ, u). This allows us to characterize the correspondence of market-clearing prices:

$$\hat{P}(\theta, u) = \left\{P : \hat{x}(P) = \theta + F^{-1}(1 - u)\right\}$$

We focus on equilibria in which the price is conditioned on (θ, u) through the observable state variable $z \equiv \theta + F^{-1}(1 - u)$. This sufficient statistic z is distributed according to cdf. $\Psi(z|\theta) = 1 - G(1 - F(z - \theta))$ and pdf $\psi(z|\theta) = g(1 - F(z - \theta)) f(z - \theta)$. The next lemma characterizes the resulting equilibrium beliefs.

Lemma 1 *In any equilibrium with conditioning on z, the equilibrium price function $P(z)$ is invertible. Posterior beliefs conditional on P are given by*

$$H(\theta|P) = \frac{\int_{-\infty}^{\theta} \psi(\hat{x}(P)|\theta)\, dH(\theta)}{\int_{-\infty}^{\infty} \psi(\hat{x}(P)|\theta)\, dH(\theta)}.$$

Lemma 1 shows that Condition 1 of the Characterization Theorem 1 must be satisfied in any equilibrium. This obtains because the market-clearing condition is fully determined from the realization of the sufficient statistic z: the supply shock u determines how many shares must be bought by informed traders to clear the market. The realization of θ then determines how the traders' private signals are distributed. Hence, θ and u together uniquely pin down the private signal of the trader whose indifference is required to clear the market. If it were now the case that the price function was non-monotone, then the same realization of P would coincide with different thresholds for this trader and would necessarily have to violate the market-clearing condition for some realizations of (θ, u).

From this lemma, we immediately have the following equilibrium characterization:

Proposition 1 *In the Risk-neutral model, for any increasing dividend function $\pi(\cdot)$, an asset market equilibrium exists, is unique, and is characterized by the price function $P_\pi(z)$ and the traders' threshold function $\hat{x}(p) = z = P_\pi^{-1}(p)$, where*

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

and $\psi(z|\theta)$ is as defined above.

The price function $P_\pi(z)$ is uniquely defined and strictly monotone, and therefore defines the unique market equilibrium.[15]

[15] Notice that this only implies the uniqueness of the equilibrium that conditions on the summary statistic z, not overall uniqueness of the equilibrium.

To specialize the characterization to the Risk-neutral, Normal Model, suppose that θ is distributed according to $\theta \sim \mathcal{N}\left(0, \sigma_\theta^2\right)$, private signals are distributed according to $x_i \sim \mathcal{N}\left(\theta, \beta^{-1}\right)$, or equivalently, $F(x-\theta) = \Phi(\sqrt{\beta}(x-\theta))$. The asset supply is $S(u) = \Phi(u)$, where $\Phi(\cdot)$ denotes the cdf of a standard normal distribution, and u is a random variable, distributed according to $u \sim \mathcal{N}\left(0, \sigma_u^2\right)$. This supply assumption is adapted from Hellwig, Mukherji and Tsyvinski (2006). With these assumptions, it follows that $D(\theta, P) = 1 - \Phi(\sqrt{\beta}(\hat{x}(P) - \theta)) = \Phi(u)$, and therefore $z = \hat{x}(P) = \theta - 1/\sqrt{\beta} \cdot u$. It follows that $z|\theta \sim \mathcal{N}\left(\theta, \sigma_u^2/\beta\right)$, and the equilibrium price and expected dividends are given by:

$$
\begin{aligned}
P_\pi(z) &= \int \pi(\theta) d\Phi\left(\sqrt{1/\sigma_\theta^2 + \beta + \beta/\sigma_u^2}\left(\theta - \frac{\beta + \beta/\sigma_u^2}{1/\sigma_\theta^2 + \beta + \beta/\sigma_u^2} \cdot z\right)\right), \qquad (2) \\
V_\pi(z) &= \int \pi(\theta)\, d\Phi\left(\sqrt{1/\sigma_\theta^2 + \beta/\sigma_u^2}\left(\theta - \frac{\beta/\sigma_u^2}{1/\sigma_\theta^2 + \beta/\sigma_u^2} \cdot z\right)\right). \qquad (3)
\end{aligned}
$$

3.2 Remarks and Extensions

The risk-neutral model has the property that the distribution of the sufficient statistic z is derived independently of the assumed payoff function $\pi(\cdot)$, and only depends on the informational primitives, i.e. the distributions of shocks and private signals. This feature is particularly useful for comparative statics discussions with respect to $\pi(\cdot)$.

The risk-neutral model with unit demand is closely related to Milgrom's analysis of $K+1$-th price common value auctions with privately informed bidders (Milgrom, 1981b). In Milgrom's model, bidders submit reservation values up to which they are willing to purchase one unit of the object (when K identical units are on sale, and a finite number of bidders). Our risk-neutral model can immediately be rephrased in terms of a traders' reservation price, in which case the traders' reservation prices, for a private signal z, are given by the equilibrium price function $P(z)$. The two models differ in that Milgrom introduces noise in the price formation through the signal noise among a finite number of agents (and then studies convergence behavior), while we assume that there is an infinite number of traders and the number of units for sale is stochastic.

It is straightforward to allow for supply and for the position limits to depend on the price, i.e., to relax the restrictions on $d_L(P)$, $d_H(P)$, and $S(u, P)$ that we imposed here. Any trader with a signal $x < z(P)$ will then demand $d_L(P)$; any trader with signal $x > z(P)$ demands $d_H(P)$. Straightforward algebra gives the following implicit characterization of $z(P)$:

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

The model with general position bounds is thus equivalent to the model with bounds at 0 and 1, and a supply function that is adjusted correspondingly.[16] In this case, however, we loose the closed form characterization of the distribution of z. While the market-clearing condition implicitly defines a threshold $\hat{u}(z,\theta)$ such that $1-F(z-\theta) = \hat{S}(\hat{u}, P_\pi(z))$, which yields a cdf. $\Psi(z|\theta) = 1-G(\hat{u}(z,\theta))$, notice that the definition of $\hat{u}(z,\theta)$ depends on the function $P_\pi(z)$, which in turn depends on the density $\psi(z|\theta)$. The equilibrium is thus 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.

3.3 Expected Wedge and Prices in the Risk-neutral, Normal model

In this section, we show results about the unconditional expectation of prices and dividends in the risk-neutral, normal model. We define

$$\gamma_P \equiv \frac{\beta + \beta\sigma_u^{-2}}{\sigma_\theta^{-2} + \beta + \beta\sigma_u^{-2}}, \quad \text{and} \quad \gamma_V \equiv \frac{\beta\sigma_u^{-2}}{\sigma_\theta^{-2} + \beta\sigma_u^{-2}}$$

as the response coefficients of the price and the expected dividends to innovations in z. The next lemma derives the closed-form solution for the unconditional wedge.

Lemma 2 (Unconditional Wedge) *Define σ_P as $\sigma_P^2 = \sigma_\theta^2 (1 + (\gamma_P/\gamma_V - 1)\gamma_P)$. The unconditional information aggregation wedge W_π is characterized by*

$$W_\pi = \int_0^\infty \left(\pi'(\theta) - \pi'(-\theta)\right) \left(\Phi\left(\frac{\theta}{\sigma_\theta}\right) - \Phi\left(\frac{\theta}{\sigma_P}\right)\right) d\theta. \tag{4}$$

This characterization shows how the wedge depends on both the shape of the payoff function, $\pi(\theta)$, and the parameters describing the informational environment (the distance of σ_P from σ_θ). The parameter $\sigma_P > \sigma_\theta$ corresponds to the prior variance of θ, as assessed in the price, and summarizes the importance of informational frictions in the market. By taking ex ante expectations over z, the shifts in mean and residual uncertainty combine into a mean-preserving spread between the weights that $\mathbb{E}(P_\pi(z))$ and $\mathbb{E}(V_\pi(z))$ associate with each realization of θ.

The price places more weight on the tails of the fundamental distribution, from an ex ante perspective (i.e., $\sigma_P > \sigma_\theta$). This result can intuitively be understood as follows. In the formula for expected dividends, the posterior of θ conditional on z is normal with mean $\gamma_V z$ and variance $(1-\gamma_V)\sigma_\theta^2$, while the prior of z is normal with mean 0 and variance σ_θ^2/γ_V. The compounded prior distribution then corresponds to the actual prior distribution of θ, with a prior variance

[16]Notice that this observation also applies to the model with inelastic bounds and supply. Widening or tightening the position bounds is thus formally equivalent to rendering the supply less or more noisy.

of $(1-\gamma_V)\,\sigma_\theta^2 + \gamma_V \sigma_\theta^2 = \sigma_\theta^2$. We can do the same compounding in the formula for the expected price, where the posterior of θ, conditional on z, is normal with mean $\gamma_P z$ and variance $(1-\gamma_P)\,\sigma_\theta^2$. Compounding with the normal prior over z, the prior over θ (as assessed in the price) is characterized as a normal distribution with mean 0 and variance $(1-\gamma_P)\,\sigma_\theta^2 + \gamma_P^2 \sigma_\theta^2/\gamma_V = \sigma_P^2$. Hence, the information frictions summarized by the distance of σ_P from σ_θ is large whenever the market signal is noisy relative to private signals, or the ratio γ_P/γ_V is high. The reason is that this leads to a large discrepancy between the posterior beliefs as assessed in the price and the expected dividends.

We use Lemma 2 to sign the unconditional wedge as a function of the shape of the dividend function, and to offer comparative statics with respect to π and the informational parameters γ_P and γ_V. Our next definition provides a partial order on payoff functions that we will use for the comparative statics.

Definition 1 *(i) A dividend function π has symmetric risks if $\pi'(\theta) = \pi'(-\theta)$ for all $\theta > 0$.*

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

(iii) A dividend function π_1 has more upside (less downside) risk than π_2 if $\pi_1'(\theta) - \pi_1'(-\theta) \geq \pi_2'(\theta) - \pi_2'(-\theta)$ for all $\theta > 0$.

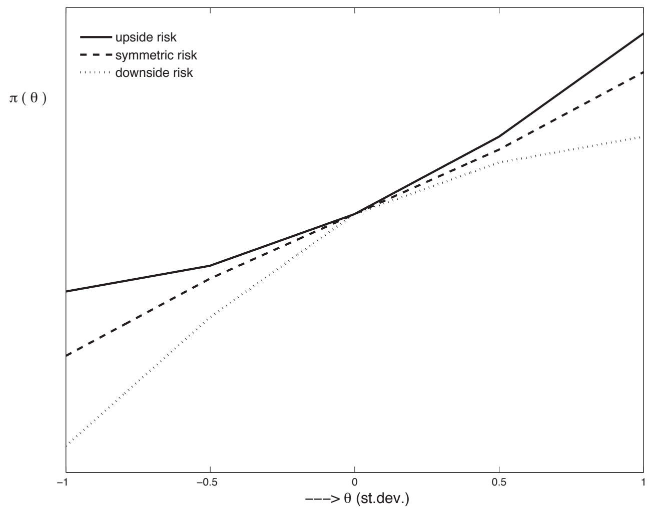
This definition classifies payoff functions by comparing marginal gains and losses at fixed distances from the prior mean to determine whether the payoff exposes its owner to bigger payoff fluctuations on the upside or the downside. Any linear dividend function has symmetric risks, any convex function is dominated by upside risks, and any concave dividend function is dominated by downside risks. The classification however also extends to non-linear functions with symmetric gains and losses, as well as non-convex functions with upside risk or non-concave functions with downside risk[17]. Figure 3 plots examples of payoff functions dominated by different types of risk.

The following Theorem summarizes the sign of the unconditional wedge and the comparative statics implications. The results follow directly from this partial order, and the characterization in lemma 2.

Theorem 2 (Average Price and Dividend Value) *(i)* ***Sign:*** *If π has symmetric risk, then $W_\pi = 0$. If π is dominated by upside risk, then $W_\pi \geq 0$. If π is dominated by downside risk, then $W_\pi \leq 0$.*

[17] In the section on applications, we also show that skewness of the distribution of the payoffs can be interpreted as the upside or downside risk.

Figure 1: Dividend risk types



(*ii*) ***Comparative Statics w.r.t.*** π***:*** *For given* σ_P^2*, if* π_1 *has more downside and less upside risk than* π_2*, then* $W_{\pi_2} \geq W_{\pi_1}$*.*

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

(*iv*) ***Increasing differences:*** *If* π_1 *has more upside risk than* π_2*, then* $W_{\pi_1}(\sigma_P) - W_{\pi_2}(\sigma_P)$ *is increasing in* σ_P*.*

This theorem summarizes how the shape of the dividend function and the informational parameters combine to determine the sign and magnitude of the unconditional information aggregation wedge. It shows that unconditional price premia or discounts arise as a combination of two elements: upside or downside risks in the dividend profile π, and an impact of private information on market prices ($\gamma_P > \gamma_V$). The latter requires that updating from prices is noisy ($\gamma_V < 1$). This theorem shows how noisy information aggregation may influence conditional and unconditional returns of assets through their payoff profile and the informational characteristics of the market.

The result is easily understood from our interpretation of the wedge as the expected value of a symmetric, mean-preserving spread of the true underlying fundamental distribution.

Part (i) shows that the sign of the wedge is determined by whether π is dominated by upside, downside, or symmetric risk. When the dividend function has symmetric risk, the gains from this spread on the upside exactly cancel the expected losses on the downside, and the total effect is 0.

When the dividend is dominated by upside risks, the expected upside gains dominate and the value of the mean-preserving spread is positive, leading to a positive unconditional wedge. Conversely, when the dividend is dominated by downside risks, the expected losses on the downside dominate and the expected value of the spread is negative.

Parts (*ii*), (*iii*), and (*iv*) complement the first result on the possibility of expected price premia or discounts with specific predictions on how its magnitude depends on cash flow and informational characteristics.

Part (*ii*) shows that an asset with more upside or less downside risk on average has a higher price premium or a lower price discount, all else equal. Thus, returns on average are lower (and prices higher) for securities that represent more upside risks. Simply put, the mean-preserving spread becomes more valuable when the payoff function shifts towards more upside risk.

Part (*iii*) shows the role of informational parameters. For a given payoff function, the unconditional wedge increases in absolute value as the information aggregation friction has bigger effects (higher σ_P). For a given set of upside or downside risks, a bigger mean-preserving spread generates bigger gains or losses. Moreover, a wedge obtains only if $\gamma_P > \gamma_V$, i.e., if the heterogeneous beliefs have an impact on price. The wedge is increasing in γ_P and decreasing in γ_V, as the precision of market information and private information move the wedge in opposite directions. If the payoff asymmetry doesn't disappear in the tails, the absolute value of the wedge approaches infinity when $\gamma_V \to 0$. This obtains if for a given value of β, the market noise becomes infinitely large. In this limiting case, the trader with the threshold signal remains responsive to z, even though the z is infinitely noisy.

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.

Importantly, our results on differences between expected prices and dividends are not a consequence of irrational trading strategies, behavioral biases of investors, or agency conflicts. Nor are such differences accounted for by risk premia (since traders are risk neutral). Our model thus offers a theory in which expected prices can differ systematically from expected dividends as a result of the interplay between the dividend structure and the partial aggregation of information into prices, in a context where traders hold heterogeneous beliefs in equilibrium and arbitrage is limited. To our knowledge, this channel is new to the literature.

3.4 Implications for expected returns

We now translate the insights from Theorem 2 into results about expected asset returns. The asset return is given by $R_\pi(\theta, z) = \pi(\theta)/P_\pi(z)$, and the expected return is

$$\begin{aligned}
\mathbb{E}(R_\pi(\theta,z)) &= \mathbb{E}\left(\frac{V_\pi(z)}{P_\pi(z)}\right) = 1 - \mathbb{E}\left(\frac{W_\pi(z)}{P_\pi(z)}\right) \\
&= 1 - W_\pi \mathbb{E}\left(\frac{1}{P_\pi(z)}\right) - cov\left(W_\pi(z), \frac{1}{P_\pi(z)}\right).
\end{aligned}$$

We therefore decompose the expected return into a term that scales with the unconditional wedge, and a term that derives from the covariance between the wedge and the inverse of the asset price. Whenever $W_\pi(\cdot)$ is monotonically increasing, this covariance is negative, and adds a positive return premium. Moreover, both of these terms are increasing in magnitude with the degree of information frictions. Therefore, the conclusions of theorem 2 directly extends to assets which exhibit downside risks, for which both the unconditional wedge and the covariance term move in the same direction. For upside risks, the unconditional wedge and the covariance term move in opposite directions, and the overall effect is therefore ambiguous in general. In the next section we derive an analytical characterization of the case of the exponential dividends, and show how the covariance term matters for returns.

We obtain further characterization for expected log-returns, $\mathbb{E}(\log R_\pi(\theta,z)) = \mathbb{E}(\log \pi(\theta)) - \mathbb{E}(\log P_\pi(z))$, as summarized by the following proposition.

Proposition 2 *The unconditional expectation of log returns is*

$$\begin{aligned}
\mathbb{E}(\log R_\pi(\theta,z)) &= \int_0^\infty \left(\frac{\pi'(\theta)}{\pi(\theta)} - \frac{\pi'(-\theta)}{\pi(-\theta)}\right)\left(\Phi\left(\frac{\theta}{\sigma_P}\right) - \Phi\left(\frac{\theta}{\sigma_\theta}\right)\right) d\theta + \mathbb{E}(\Delta(z)), \textit{ where} \\
\Delta(z) &= \mathbb{E}(\log R_\pi(\theta,z)\,|x=z,z) \approx -\frac{1}{2} Var(R_\pi(\theta,z)\,|x=z,z),
\end{aligned}$$

and $Var(R_\pi(\theta,z)\,|x=z,z)$ denotes the residual variance of returns that is implied by the market price.

The expected log-return is decomposed into two terms. The first inherits all the predictions of Theorem 2, after re-defining upside and downside risks in terms of percentage variation in returns (and noting that a positive wedge translates into a lower expected return). The second term is an adjustment from Jensen's inequality (due to the transformation of returns to logs), and scales with the residual uncertainty of the trader receiving the threshold signal. This second term disappears as γ_P converges to 1, that is, when residual uncertainty disappears. 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 returns, while downside risks trade at a discount and offer higher returns, and these premia and discounts are increasing in the degree of information frictions.

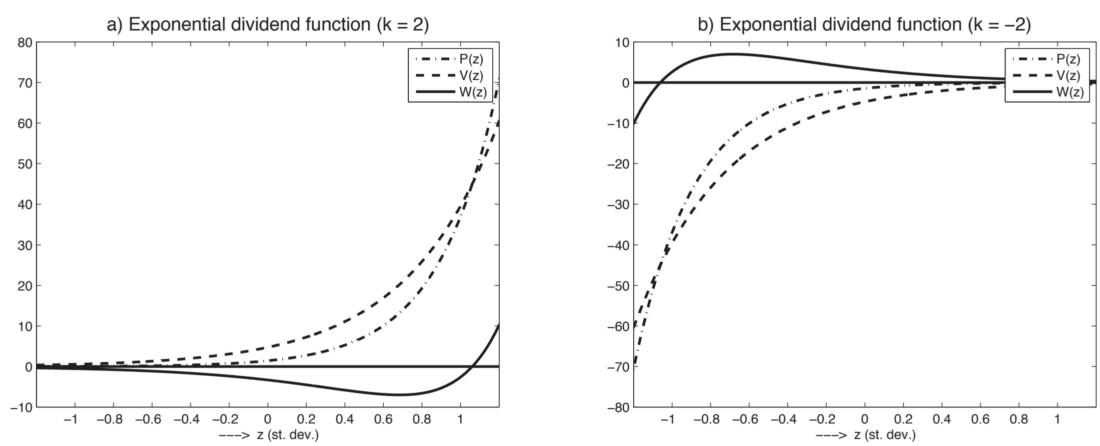
3.5 Example: log-normal returns

Here we illustrate the results of the preceding sections with a brief example that encompasses the different possibilities. Suppose that $\pi(\theta) = \frac{1}{k} e^{k\theta}$, with $k \neq 0$. Expected dividends, prices and the wedge are then characterized by:

$$
\begin{aligned}
V_\pi(z) &= \frac{1}{k} e^{k\gamma_V z + \frac{k^2}{2}\sigma_\theta^2(1-\gamma_V)}, \quad P_\pi(z) = \frac{1}{k} e^{k\gamma_P z + \frac{k^2}{2}\sigma_\theta^2(1-\gamma_P)} \\
W_\pi(z) &= P_\pi(z)\left(1 - e^{-k(\gamma_P-\gamma_V)z + \frac{k^2}{2}(\gamma_P-\gamma_V)\sigma_\theta^2}\right).
\end{aligned}
$$

The price and expected dividend are both exponential functions in z, with a stronger reaction of prices to z. The residual uncertainty affects both $V_\pi(z)$ and $P_\pi(z)$ multiplicatively, but the factor is larger for $V_\pi(z)$.

Figure 2: Exponential dividends



If $k > 0$, the dividend function is increasing, convex, and bounded below by zero (figure 2, panel c). The conditional wedge is negative at $z = 0$ and non-monotone. It decreases at first, reaches its lowest value at some intermediate point, and is increasing and convex from there on, crossing 0 at $z = \frac{k}{2}\sigma_\theta^2 > 0$. The reverse image obtains when $k < 0$, in which case π is increasing, concave, and bounded above by zero (figure 2, panel d). For negative z, the wedge is negative at first and increasing in z, crossing 0 at $z = \frac{k}{2}\sigma_\theta^{-2} < 0$. It reaches its maximum value at a negative z and then monotonically converges towards 0. This example thus confirms the intuitions from the shift in means which makes $P_\pi(z)$ more responsive to a shift in z, and the shift in residual uncertainty that is captured by the multiplicative factors. The curvature parameter k governs the shape of the wedge function, and whether the residual uncertainty increases or decreases the wedge.

The exponential dividend function illustrates the content of Theorem 2: the expected wedge is positive if and only if $k > 0$, and negative if $k < 0$. That is, the security trades at a premium in the case with convex dividends and upside risks, and at a discount in the case with concave dividends and downside risks. Taking expectations, we have

$$\begin{aligned} \mathbb{E}\left(V_\pi(z)\right) &= 1/k \cdot e^{\frac{k^2}{2}\sigma_\theta^2}, \quad \mathbb{E}\left(P_\pi(z)\right) = 1/k \cdot e^{\frac{k^2}{2}\sigma_\theta^2\left[1+\left(\frac{\gamma_P}{\gamma_V}-1\right)\gamma_P\right]} = 1/k \cdot e^{\frac{k^2}{2}\sigma_P^2} \\ \text{and } \mathbb{E}\left(W_\pi(z)\right) &= 1/k \cdot e^{\frac{k^2}{2}\sigma_\theta^2}\left\{e^{\frac{k^2}{2}\sigma_\theta^2(\gamma_P/\gamma_V-1)\gamma_P} - 1\right\} = 1/k \cdot \left\{e^{\frac{k^2}{2}\sigma_P^2} - e^{\frac{k^2}{2}\sigma_\theta^2}\right\}, \end{aligned}$$

which is positive whenever $k > 0$, and negative for $k < 0$ (and can be checked to approach 0 continuously as $k \to 0$). Moreover, expected returns are given by

$$\begin{aligned} \mathbb{E}\left(R_\pi(\theta, z)\right) &= e^{\frac{k^2}{2}\sigma_\theta^2(\gamma_P/\gamma_V-1)\gamma_P} = e^{\frac{k^2}{2}\left(\sigma_P^2-\sigma_\theta^2\right)} > 1, \text{ and} \\ \mathbb{E}\left(\log R_\pi(\theta, z)\right) &= -\frac{k^2}{2}\sigma_\theta^2\left(1-\gamma_P\right) \end{aligned}$$

Here, the covariance term dominates the expected return expression, leading to an expected return premium that is increasing with the magnitude of information frictions. For log-returns, however, the log-normal example represents neither upside nor downside risk, so only the residual uncertainty component remains and lowers the expected return.

4 Applications

In this section, we develop three applications. First, we show how our model is consistent with the empirical predictions on a return to skewness theoretically and quantitatively. Second, we reconsider the Modigliani-Miller Theorem and show how by appropriately splitting a cash flow in two components, and selling them to different clienteles, the owner can influence its market value. Third, we consider a dynamic extension of our model to demonstrate how the information aggregation wedge may give rise to permanent over- or under-valuation of securities in a dynamic context.

4.1 Returns and skewness

A sizeable empirical literature documents a negative relationship between measures of skewness and asset returns. Conrad, Dittmar and Ghysels (forthcoming) (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. Boyer, Mitton and Vorkink (2010) estimate a cross-sectional model of expected skewness that uses a variety of predictive variables 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.

The empirical prediction of these papers are consistent with three types of models. Brunnermeier and Parker (2005) and Brunnermeier, Gollier and Parker (2007) develop a model of optimal expectations. They show that investors are undiversified and overinvest in a security that is more skewed than the average security, hence more positively skewed securities have lower returns. Barberis and Huang (2008) use cumulative prospect theory implying that investors prefer stocks with lottery-like features. A significant position in a positively skewed asset presents a small, but highly valued, chance of making significant amount of money. A positively skewed security can then be overpriced and can earn a negative average excess return. Mitton and Vorkink (2007) develop a model in which investors have heterogeneous preference for skewness.

Our model provides a novel theoretical justification for the negative relationship between the skewness and returns. The explanation does not rely on non-rational expectations, behavioral phenomena, or on heterogeneity in preferences. Moreover, we show quantitatively that even a small degree of informational friction can generate significant negative returns to skewness.

Measuring upside or downside risk of the return distribution by skewness, Theorem 2 and Proposition 2 lead to the following empirical predictions:

Empirical prediction: *Higher skewness leads to lower expected returns.*

As the main comparison for the relation of skewness and returns in our model, we base ourselves on the estimates in CDG. For different stocks, they infer the distribution of cash flows from option prices, using options with maturities of 3 and 12 months. They then sort the stocks according to their degree of skewness in high-skewed (top 30%, skewness ≈ 0), medium-skewed and low-skewed (bottom 30%, skewness ≈ -2.8) securities, and report the subsequent raw returns of the equally weighted skewness-ranked portfolios over the next month. The raw returns differential is negative, at -82 and -73 basis points per month for returns inferred from options with three- and 12-

month maturities, respectively. On average, for each maturity securities with lower skewness earn higher returns in the next month, and securities with less negative or positive skewness earn lower returns. They also subtract the return of the 3 factor Fama-French portfolio from the individual security returns and average the resulting excess or characteristic-adjusted return across firms in the skewness-ranked portfolios. These characteristic-adjusted returns are quite similar: averaging -79 and -67 basis points per month, for three- and 12-month maturities. For the quantitative example below we chose to target these characteristic adjusted returns. These magnitudes are also consistent with Boyer, Mitton, and Vorkink (2010) who find that portfolios sorted on expected skew in the physical distribution generate a return differential of 67 basis points per month and with Chang, Christoffersen, and Jacobs (forthcoming) who find the return to the skewness of -50 to -70 basis points.

We now quantitatively show that even small degrees of information frictions lead to excess returns that are easily in line with the magnitudes reported by CDG. We use our model to price a security that matches the skewness and volatility properties of the different stock sorts in CDG. We then compute the securities' expected log-returns in our model, for different values of the information friction parameter σ_P, and ask at what levels of the model-implied returns are consistent with the actual ones.

Specifically, we assume that $\log \pi(\theta) = k \cdot x(\theta)$ where $x(\cdot)$ is distributed according to a beta distribution, and $k > 0$ is a scaling parameter. We choose k and the two parameters of the beta distribution to match the model-implied volatility and skewness of the return distribution with the market-implied one for the CDG portfolios.[18] For this security we then compute expected returns and market-implied volatilities for different levels of information frictions.

The results are reported in Table 1. The rows represent the portfolios sorted by skewness, with skewness measures constructed from options of 3-month and 12-month to maturity, respectively. That is, the first row, represents the Low-skewness portfolio constructed from options with 3-month maturity, the second row, the corresponding high-skewness portfolio, and the third row reports the return differences between the two.

The first three columns then report empirical moments, and are taken directly from CDG (Table II). Here we report the average returns (and the excess return of high-to-low skewness), which we

[18]Specifically, let A and B denote the distributional parameters. Holding $A+B$ fixed, we vary the mean of the distribution $A/(A+B)$ to match the skewness (higher mean makes the skewness more negative), and we vary the scaling parameter k to match the volatility of returns. Our benchmark results are reported for $A+B=5$, but results for other values are similar. This parametrization was chosen because it allows us to flexibly match the volatility and skewness properties with two parameters, and allows for virtually arbitrary skewness/volatility configurations.

Table 1: Information frictions, skewness, and expected returns

		Data			σp =1.03		σp =1.06		σp =1.09		σp =1.12	
Maturity	Skewness	**Return**	**Vol**	**Skewness**	**Return**	**ImpVol/Vol**	**Return**	**ImpVol/Vol**	**Returns**	**ImpVol/Vol**	**Returns**	**ImpVol/Vol**
	Low	**0.57**	31.51	-2.814	**0.61**	1.049	**1.31**	1.094	**1.98**	1.141	**2.64**	1.189
3 Months	Medium	**0.21**	32.26	-0.980	**0.28**	1.027	**0.57**	1.052	**0.85**	1.079	**1.13**	1.104
	High	**-0.22**	31.14	0.026	**-0.09**	1.022	**-0.18**	1.045	**-0.27**	1.066	**-0.37**	1.087
	H-L	**-0.79**			**-0.70**		**-1.49**		**-2.25**		**-3.01**	
	M-L	**-0.37**			**-0.32**		**-0.74**		**-1.13**		**-1.51**	
	Low	**0.55**	30.16	-2.743	**0.63**	1.046	**1.26**	1.092	**1.89**	1.138	**2.52**	1.185
12 Months	Medium	**0.15**	31.31	-0.974	**0.27**	1.026	**0.55**	1.053	**0.83**	1.078	**1.10**	1.104
	High	**-0.12**	30.69	0.019	**-0.09**	1.023	**-0.17**	1.045	**-0.26**	1.066	**-0.35**	1.087
	H-L	**-0.67**			**-0.71**		**-1.43**		**-2.15**		**-2.87**	
	M-L	**-0.40**			**-0.35**		**-0.71**		**-1.06**		**-1.42**	

seek to explain, and the empirical volatility and skewness of these portfolios.

The subsequent columns then report theoretical moments from our model. For each of the empirical portfolios, we create a theoretical counterpart security by setting the parameters of the beta distributions so that the distribution of $\pi(\cdot)$ matches the reported skewness and volatility from the data. We set $\sigma_\theta^2 = 1$ and $\sigma_u^2 = 2$ and then vary β to set the information friction parameter σ_P/σ_θ equal to values ranging from 1.03 to 1.12. 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.[19]

The table shows that even with the smallest level of frictions, the premium for skewness falls in the range suggested by the data, and matches the excess returns observed in the data remarkably well. A value of σ_P/σ_θ of 1.06 (meaning that the market-implied standard deviation of θ exceeds the true standard deviation of θ by 6%) is already sufficient to generate excess returns that are well above the level observed in the data. In other words, even very moderate information frictions can generate sizeable and empirically plausible excess returns from skewness. This finding is also robust across the different option maturities. What is more for each comparison, the volatilities of high and low-skewed securities are sufficiently close to each other that we can rule out the possibility that the results are driven by volatility differences.

We now 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 a positive excess return. To test this prediction, one needs to define a benchmark

[19] Since the volatility and skewness numbers in the data are imputed from option prices, it might make sense to match these up with the market-implied skewness and volatility measures from the model. The corresponding calibration is more involved (since now the distributional parameters have to be varied as a function of market frictions), but when frictions are small, this is unlikely to have big effects, since the targeted volatility and skewness remain approximately the same.

return, against which the positive and negative excess returns of upside and downside risks may be measured (to be consistent with the prediction, 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 aggregation frictions and limits to arbitrage 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. While there is no such systematic analysis for the measures of the informational friction, Ruf (2012) finds a positive relationship between skewness premium and measures of limits to arbitrage in options.

4.2 Splitting Cash-flows to influence market value

The Modigliani-Miller theorem states that in perfect and complete financial markets, splitting a cash flow into two different securities, and selling these claims separately to investors does not influence its total market value (Modigliani and Miller, 1958). Here we show that the Modigliani-Miller theorem remains valid with noisy information aggregation, only if the different claims are sold to investor pools with *identical* informational characteristics. When the investor pools for different claims have different characteristics, then the nature of the split influences the seller's revenue. The seller in turn can increase her revenues by tailoring the split to the different investor types.

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

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

That is, each market is affected by a noise trader shock u_i with market-specific noise parameter $\sigma_{u,i}^2$. The environment is then characterized by the market-characteristics β_i and $\sigma_{u,i}^2$, and by the correlation of demand shocks across markets, ρ. Traders are active only in their respective market. 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).

Under informational segregation, the analysis of the two markets can be completely separated; any correlation between the two in prices is the result of correlation in demand shocks, as well as the common underlying fundamental, but this does not influence expected revenues. The equilibrium characterization from the single-asset model applies separately in each market: $P_i(z_i) = \mathbb{E}\left(\pi_i(\theta)|x = z_i, z_i; \beta_i, \sigma_{u,i}^2\right)$ and $V_i(z_i) = \mathbb{E}\left(\pi_i(\theta)|z_i; \beta_i, \sigma_{u,i}^2\right)$. The seller's total expected revenue in excess of the cash flow's expected dividend value is then given by $W_{\pi_1}(\sigma_{P,1}) + W_{\pi_2}(\sigma_{P,2})$, where $\sigma_{P,i}$ is determined as in lemma 2, and denotes the level of informational frictions in each market.

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. Since expected dividends are monotone, informed traders in market i will 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, it must be the case that $\hat{x}_i(\cdot) = z_i \equiv \theta + 1/\sqrt{\beta_i} \cdot u_i$. 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 θ, market prices, and expected dividends, as functions of (z_1, z_2):

$$\begin{aligned}
P_1(z_1, z_2) &= \mathbb{E}(\pi_1(\theta)|x = z_1; z_1, z_2) \text{ and } V_1(z_1, z_2) = \mathbb{E}(\pi_1(\theta)|z_1, z_2), \\
P_2(z_1, z_2) &= \mathbb{E}(\pi_2(\theta)|x = z_2; z_1, z_2) \text{ and } V_2(z_1, z_2) = \mathbb{E}(\pi_2(\theta)|z_1, z_2)
\end{aligned}$$

From an ex ante perspective, the unconditional information aggregation wedge W_{π_i} is characterized by

$$\begin{aligned}
W_{\pi_i}(\sigma_{P_i}) &= \int_0^\infty \left(\pi_i'(\theta) - \pi_i'(-\theta)\right)\left(\Phi\left(\frac{\theta}{\sigma_\theta}\right) - \Phi\left(\frac{\theta}{\sigma_{P_i}}\right)\right) d\theta, \text{ where} \\
\sigma_{P_i}^2 &= \sigma_\theta^2 + \left(1 + \sigma_{u_i}^2\right) \cdot \frac{\beta_i}{(\beta_i + V)^2}, \text{ with } V = 1/\sigma_\theta^2 + \frac{1}{1-\rho^2}\left(\frac{\beta_1}{\sigma_{u,1}^2} + \frac{\beta_2}{\sigma_{u,2}^2} - 2\rho\frac{\sqrt{\beta_1\beta_2}}{\sigma_{u,1}\sigma_{u,2}}\right).
\end{aligned}$$

The characterization of lemma 2 continues to apply separately for each market, after adjusting the definition of informational frictions $\sigma_{P,i}$ for the change in the underlying information structure. For given values of π_i and $\sigma_{P,i}$, the seller's expected revenue net of expected dividends in both cases is $W_{\pi_1}(\sigma_{P,1}) + W_{\pi_2}(\sigma_{P,2})$. We can now state a first version of the Modigliani-Miller theorem for expected revenues in our model.

Proposition 3 (Modigliani-Miller I) *(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 $\sigma_{P,1} > \sigma_{P,2}$, $W_{\pi_1}(\sigma_{P,1}) + W_{\pi_2}(\sigma_{P,2}) > W_{\pi_1}(\sigma_{P,2}) + W_{\pi_2}(\sigma_{P,1})$, while if $\sigma_{P,1} < \sigma_{P,2}$, $W_{\pi_1}(\sigma_{P,1}) + W_{\pi_2}(\sigma_{P,2}) < W_{\pi_1}(\sigma_{P,2}) + W_{\pi_2}(\sigma_{P,1})$.

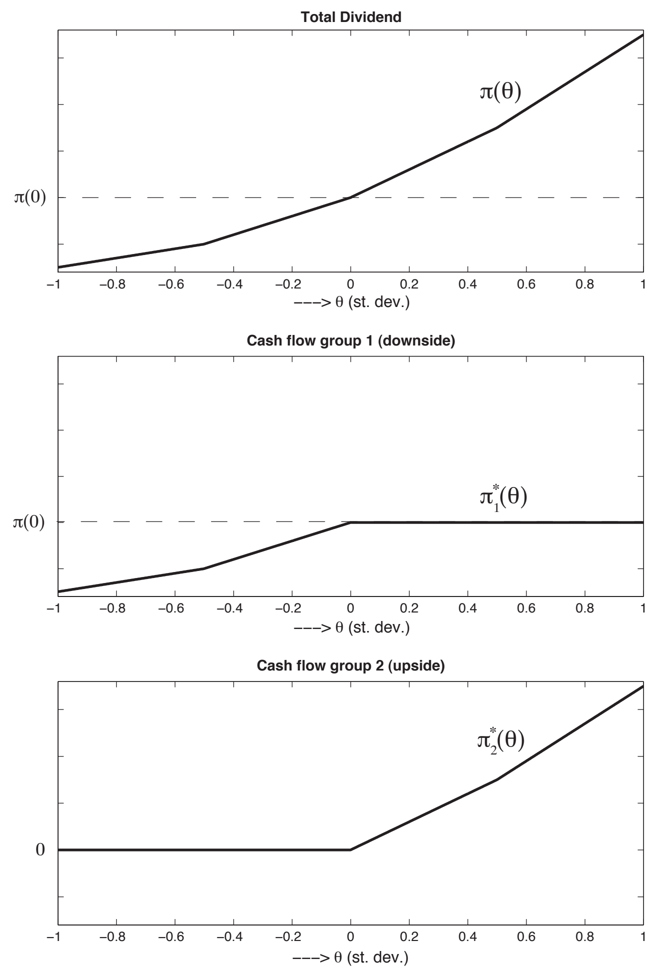
For given values of σ_P, the expected information aggregation wedge is additive across cash flows: $W_{\pi_1}(\sigma_P) + W_{\pi_2}(\sigma_P) = W_{\pi_1+\pi_2}(\sigma_P)$, for any σ_P, π_1 and π_2. If the two markets have identical characteristics, i.e. $\sigma_{P,1} = \sigma_{P,2}$, only the combined cash flow matters for the total wedge - i.e. the Modigliani-Miller result applies. If on the other hand the two markets have different informational characteristics, then the increasing difference property of $W_{\pi_1}(\cdot)$ implies that the seller's revenue is influenced by how the two cash flows are matched to the two markets, and the revenue is higher when the upside risk is matched with the market that has more severe information frictions (a higher value of σ_P). Intuitively, the seller exploits the information aggregation wedge to manipulate revenues, matching the pool of investors with high informational frictions with the upside risk, while selling the downside risks to an investor pool with lower informational frictions. This maximizes the gains from the positive wedge resulting on the upside, while it minimizes the losses from the negative wedge on the downside. This logic is pushed further by the next proposition, which considers how the seller can exploit the heterogeneity in investor pools if she gets to design the split of π into π_1 and π_2.

Proposition 4 (Designing Cash flows) *The seller maximizes her expected revenues by splitting cash flows according to* $\pi_1^*(\theta) = \min\{\pi(\theta), \pi(0)\}$ *and* $\pi_2^*(\theta) = \max\{\pi(\theta) - \pi(0), 0\}$*, and then assigning* π_1^* *to the investor pool with the lower value of* σ_P*.*

Figure 4 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 $\pi(.) = \pi(0)$ to the investor group with the lowest information friction parameter; $\sigma_{P,1}$, and the complement to the investor group with the highest friction; $\sigma_{P,2}$. For any other arbitrary division of cash flows $\{\pi_1(\cdot), \pi_2(\cdot)\}$ we have both $\pi_1^{*\prime}(\theta) - \pi_1^{*\prime}(-\theta) \leq \pi_1'(\theta) - \pi_1'(-\theta)$, and $\pi_2^{*\prime}(\theta) - \pi_2^{*\prime}(-\theta) \geq \pi_2'(\theta) - \pi_2'(-\theta)$. That is, π_1 has less downside risk than π_1^*, and π_2 has less upside risk than π_2^*. Due to the increasing difference property (Theorem 2, part *iv*), any such alternative split of cash flows between investor groups results in strictly lower revenue for the seller. 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 risk to the group that overvalues the claim the most with respect to its expected dividend (due to the high information frictions; $\sigma_{P,2}$). When $\pi(\cdot) > 0$, this split has a straight-forward interpretation in terms of debt and equity, with a default point on debt that is set at the prior median $\pi(0)$.

We conclude this section by stating a second version of the Modigliani-Miller theorem for realized

Figure 3: Optimal Cash-flow Design



revenues, $P_{\pi_1}(z_1, z_2) + P_{\pi_2}(z_1, z_2)$. The original Modigliani-Miller theorem holds also at an interim stage conditional on new information, as long as the traders with the threshold signal in the two markets hold identical beliefs for each realization (z_1, z_2).

Proposition 5 (Modigliani-Miller II) (i) *With informational segregation: The cash-flow split does not affect the seller's realized revenue, if and only if the noise trading is perfectly correlated across markets* $(\rho = 1)$*, and the two markets have identical informational characteristics* $(\beta_1 = \beta_2$ *and* $\sigma^2_{u,1} = \sigma^2_{u,2})$*.*

(ii) *With informational linkages: The cash-flow split does not affect the seller's realized revenue, if and only if* $\rho = 1$ *and either* $\beta_1 = \beta_2$ *and* $\sigma^2_{u,1} = \sigma^2_{u,2}$*, or* $\beta_1\sigma^{-2}_{u,1} \neq \beta_2\sigma^{-2}_{u,2}$*.*

An interim version of the theorem therefore requires perfect correlation in the noise in different markets, on top of identical informational characteristics. Perfect correlation reduces the noise to a single common shock. That this is necessary for the theorem to hold under segregation is immediate. It is also necessary for the case with informational linkages, because the market prices weight the two signals z_1 and z_2 differently in the two markets. In addition the signal distributions need to be the same, requiring that $\beta_1 = \beta_2$ and $\sigma_{u,1}^2 = \sigma_{u,2}^2$. Together these conditions imply that the two markets have identical informational characteristics.[20]

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 Modigliani and Miller's (1958) irrelevance result, and to show that the information frictions give the owner of a cash flow distinct possibility to manipulate its market value through security design.[21]

4.3 Dynamic Trading and Bubbles

A general implication of arbitrage-free asset pricing is the impossibility of persistent mis-pricing or rational bubbles for a general class of dynamic asset market economies (Tirole, 1982; Santos and Woodford, 1997). This is one of the classic no-arbitrage results: while the anticipation of higher prices in the future leads traders to bid up the price in the current period, a positive

[20]In the case with informational linkages, we need to consider the additional possibility that when $\beta_1\sigma_{u,1}^{-2} \neq \beta_2\sigma_{u,2}^{-2}$ and $\rho = 1$, the observations of two signals with different precision but perfectly correlated noise enables every trader to perfectly infer θ and u from the two prices regardless of the informational parameters, implying $P_{\pi_1}(z_1, z_2) = \pi_1(\theta)$ and $P_{\pi_2}(z_1, z_2) = \pi_2(\theta)$.

[21]We conjecture that, as long as investors are heterogeneous in their ability to acquire or process information, there will be a natural force towards segmentation between investors who actively trade on their private information, and migrate towards securities that are "information-sensitive" (like the equity tranch in our model), and investors that do not actively trade on private information and migrate towards securities that are less information sensitive (like the debt tranch in our model). The resulting segmentation would presumably not result in a perfect separation of upside from downside risks to limit the information sensitivity of the downside tranch, but the general principle would be similar. See Yang (2012) for progress on a related security design problem derived from investor information processing.

bubble component in the price is consistent with arbitrage by buy-and-sell strategies only if its date zero present value follows a martingale process. But this is inconsistent with the implication of discounting and the transversality condition, by which aggregate wealth and the present discounted value of aggregate consumption has finite present value, or merely with backwards induction, in the case of finite-lived securities, unless the bubble component is exactly zero.

Here, we show in a simple dynamic example how noisy information aggregation breaks this result and persistent mis-pricing, or even permanent over-valuation of securities becomes possible as an equilibrium outcome. Extending the insights of Theorem 2 to a dynamic environment, we show that if a security is expected to sustain a positive wedge (on average) in the future, then this anticipation increases prices in the current period, and can be sufficiently strong so that the security is priced above the present discounted value of future dividends in all periods and states.

To establish this result, consider the following environment. Time is discrete and infinite. There is an infinitely-lived security, which pays a dividend $\pi(\theta_{t-1})$ at the beginning of period t, where $\theta_t \sim \mathcal{N}\left(0, \sigma_\theta^2\right)$ is i.i.d. over time. The supply of this security is normalized to 1. In each period, a round of trading takes place between noise traders and informed traders, after the current dividend has been paid (and therefore θ_{t-1} is publicly known at the start of the period t market). In period t, a fraction $\Phi(u_t) \in (0,1)$ is bought by noise traders and held to period $t+1$, where $u_t \sim \mathcal{N}\left(0, \sigma_u^2\right)$ is i.i.d. over time. In addition, there are long-lived informed traders who are risk neutral and discount the future at a rate $\delta \in (0,1)$. In each period, these informed traders receive a noisy private signal $x_t \sim \mathcal{N}\left(\theta_t, \beta^{-1}\right)$, which is iid over time and across traders, and decide whether or not to purchase up to 1 unit of the security, in order to resell it in the following period.[22]

A trading strategy for informed traders determines demand $d \in [0,1]$ as a function of their past and current private signals, market prices and past fundamental realizations. A trader finds it optimal to purchase the security ($d = 1$), if and only if

$$P_t \leq \delta \mathbb{E}\left(\pi(\theta_t) + P_{t+1} \middle| \{x_{t-\tau}, P_{t-\tau}, \theta_{t-1-\tau}\}_{\tau=0}^{\infty}\right).$$

The equilibrium price P_t may in principle be a function of the entire sequence of current and past fundamentals and noise-trading shocks $\{\theta_{t-\tau}, u_{t-\tau}\}_{\tau=0}^{\infty}$. An equilibrium consists of posterior beliefs and a trading strategy of informed traders, and a price function for each period, such that in each period, traders behave optimally, update beliefs according to Bayes' rule, and the market clears. An equilibrium is stationary, if the equilibrium trading strategies and price functions are constant over time.

[22] This formulation is equivalent to one in which there are overlapping generations of traders, and in each period the old owners of the asset liquidate their positions and sell the asset to the next generation of traders.

Exploiting the recursive and forward-looking nature of the market, we show that there exists an equilibrium, in which traders condition their strategies only on current period signal x_t and price P_t, and the price is a function only of the current realizations of (θ_t, u_t):

Lemma 3 (Dynamic Equilibrium Characterization) *Define $z_t \equiv \theta_t + 1/\sqrt{\beta} \cdot u_t$. Then there exists a stationary, forward-looking equilibrium, in which current prices are a function of z_t only and take the form*

$$\begin{aligned} P_\pi(z_t) &= \delta \mathbb{E}(\pi(\theta_t) | x = z_t, z_t) + \delta P_\pi, \text{ where} \\ P_\pi &= \frac{\delta}{1-\delta} \mathbb{E}(\mathbb{E}(\pi(\theta_t) | x = z_t, z_t)). \end{aligned}$$

Informed traders acquire the asset if and only if $x \geq z_t$.

The equilibrium characterization can therefore be extended in a straight-forward manner from the static model to a dynamic one.[23] In this setting, the expected fundamental value of the asset conditional on P_t is

$$V_\pi(z_t) = \delta \mathbb{E}(\pi(\theta_t) | z_t) + \delta V_\pi, \text{ where } V_\pi = \frac{\delta}{1-\delta} \mathbb{E}(\pi(\theta_t))$$

We then have the following characterization of the dynamic information aggregation wedge:

$$\begin{aligned} W_\pi(z_t) &= w_\pi(z_t) + \delta \mathbb{E}(W_\pi(z)) = w_\pi(z_t) + \frac{\delta}{1-\delta} \mathbb{E}(w_\pi(z)), \text{ where} \\ w_\pi(z_t) &= \delta [\mathbb{E}(\pi(\theta_t) | x = z_t, z_t) - \mathbb{E}(\pi(\theta_t) | z_t)] \end{aligned}$$

is the wedge resulting from the next period's dividend, and $\mathbb{E}(w_\pi(z))$ its corresponding unconditional expectation. Thus, the information aggregation wedge in the dynamic setting depends on both the wedge resulting from current payoffs, and the expected discounted future wedge. Even when the current wedge is negative (at low realizations of z), the overall wedge may still be positive because traders anticipate higher share prices in the future. The following proposition formalizes this observation.

Proposition 6 (Sustainability of Bubbles) *Suppose that $\pi(\theta)$ is bounded below, increasing, and convex. Then, for any $\sigma_P > \sigma_\theta$, there exists $\hat{\delta} < 1$ such that for all $\delta > \hat{\delta}$, $W_\pi(z_t) > 0$, for all z_t.*

[23] We do not explore here the possibility that there exist other equilibria - for our purpose of highlighting the possibility of persistent mis-pricing, it suffices to characterize one such equilibrium. A general analysis would have to explore (i) whether equilibria may be conditioned on other (backwards-looking information), and (ii) whether the forward-looking nature of the equilibrium price characterization allows for additional non-stationary solutions.

Proposition 6 shows how claims that have a lower bound on payoffs (for example, requiring them to be non-negative), and that generate a positive expected wedge, can be priced in the market at a value exceeding expected dividends at all times and in all states of the world. Symmetrically, a claim whose payoffs are bounded above may be undervalued in all future states. The positive (negative) exponential payoff function from the example we considered exactly satisfies the required conditions for a permanent bubble (or discount).

The example illustrates the key forces that are at play to overturn the no-arbitrage argument against bubbles. First, with mean reversion in fundamentals and noise trading (captured by the i.i.d. assumption in shocks), the traders anticipation of future wedges are driven by the unconditional wedge. With upside risks, this is positive. Second, with bounded payoffs, there is a limit to how much the market's expectation of current dividends can be undervalued relative to the objective outsider's expectation. Third, the anticipation of a positive future wedge will dominate a negative current wedge, if traders are sufficiently patient.

This example is of course highly stylized, as a complete and exhaustive discussion of dynamic extensions of our model leads to additional difficulties on its own, which exceed the scope of this paper, and are left to future work. Nevertheless it is suggestive of the types of markets in which information-driven bubbles are likely to emerge, and when they are likely to occur, namely those that represent significant future upside opportunities, and/or markets in which investors face implicit protection against downside risks.

5 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 theorems impose no restrictions on the distribution of asset payoffs, and only the second theorem 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 quantitatively consistent with the empirical facts on return to skewness. And last but not least, our theory is tractable and easily lends itself to applications, such as our discussion of the Modigliani-Miller theorem and the sustainability of bubbles.

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, 2011a) already explored the potential for information as a 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. A fourth direction is to extend the analysis of a multi-period, and multi-asset extensions of our market model, both of which have already been touched upon in this paper in the context of specific examples. A final direction lies in the integration of financial market frictions with real decisions that endogenize the dividend payoff function we considered here. In a companion paper (Albagli, Hellwig, and Tsyvinski, 2011b), we consider one such model in which there is interplay between information aggregation, firm decisions and managerial incentives in a simple model of informational feedback.

References

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

[2] Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski (2011b). "Information Aggregation, Investments and Managerial Incentives," NBER Working paper No. 17330.

[3] Allen, Franklin, Stephen Morris, and Andrew Postlewaite (1993). "Finite Bubbles with Short Sale Constraints and Asymmetric Information," Journal of Economic Theory, 61, 206-229.

[4] Allen, Franklin, Stephen Morris, and Hyun Song Shin (2006). "Beauty Contests and Iterated Expectations in Asset Markets," Review of Financial Studies, 19, 161-177.

[5] Bacchetta, Philippe and Eric van Wincoop (2006). "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?," American Economic Review, 96 (3), 552-576.

[6] Bai, Yang, Eric C. Chang, and Jiang Wang (2006). "Asset Prices Under Short-Sale Constraints," working paper, MIT.

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

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

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

[10] Breon-Drish, Bradyn (2011). "Asymmetric Information in Financial Markets: Anything Goes," working paper, Stanford University.

[11] Breon-Drish, Bradyn (2012). "On Existence and Uniqueness of Equilibrium in Noisy Rational Expectations Economies," working paper, Stanford University.

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

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

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

[15] Chang, Bo Young, Peter Christoffersen, and Kris Jacobs (forthcoming). "Market Skewness Risk and the Cross-Section of Stock Returns," Journal of Financial Economics.

[16] Chen, Joseph, Harrison Hong and Jeremy Stein (2002)."Breadth of Ownership and Stock Returns," Journal of Financial Economics, 66 (2-3), 171-205.

[17] Conrad, Jennifer, Robert Dittmar, and Eric Ghysels (forthcoming). "Ex Ante Skewness and Expected Stock Returns," Journal of Finance.

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

[19] ———————————————— (1987). "Constraints on short-selling and asset price adjustment to private information," Journal of Financial Economics 18 (2), 277-311.

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

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

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

[23] Harrison, Michael and David Kreps (1978). "Speculative Investors in a Stock Market with Heterogeneous Expectations," Quarterly Journal of Economics, 92 (2), 323-336.

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

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

[26] Hong, Harrison, and David Sraer, (2011). "Quiet Bubbles," working paper, Princeton University.

[27] Hong, Harrison and Jeremy Stein, (2007). "Disagreement and the Stock Market," Journal of Economic Perspectives, 21 (2), 109-128.

[28] Milgrom, Paul (1981a), "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics* 12, 380-391.

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

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

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

[32] Ruf, Thomas (2012). "Limits to Market Making, Liquidation Risk and the Skewness Risk Premium in Options Markets," Unpublished Manuscript.

[33] Santos, Manuel, and Michael Woodford (1997). "Rational Asset Pricing Bubbles," Econometrica 65 (1), 19-57.

[34] Scheinkman, Jose and Wei Xiong (2003). "Overconfidence and Speculative Bubbles" Journal of Political Economy, 111 (6), 1183-1220.

[35] Tirole, Jean (1982). "On the Possibility of Speculation Under Rational Expectations," Econometrica, 50 (5), 1163-1181.

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

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

[38] Yang, Ming (2012). "Optimality of Debt under Flexible Information Acquisition," working paper, Fuqua School of Business, Duke University.

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

6 Appendix: Proofs

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

Lemma 4 *Suppose that θ is distributed according to some continuous bounded density $h(\cdot)$ and that the signal x satisfies assumption A1. 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 $\hat{\varepsilon} = \arg\max_{\varepsilon} f(\varepsilon)$, and notice that when $x > \theta + \hat{\varepsilon}$,

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

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 $x < \theta + \hat{\varepsilon}$,

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

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

Proof of Theorem 1. From the previous lemma, it follows immediately that for any smooth density $h(\cdot|P)$, $\mathbb{E}(\pi(\theta)\,|x,P)$ is strictly increasing in x and if P is on the interior of the support of $\pi(\theta)$ (that is, if there exist θ_1 and θ_2, such that $\pi(\theta_1) > P > \pi(\theta_2)$), then $\lim_{x\to-\infty} \mathbb{E}(\pi(\theta)\,|x,P) <$

$P < \lim_{x\to\infty} \mathbb{E}(\pi(\theta)|x,P)$. Thus, there exists a unique $z(P)$ for which $\mathbb{E}(\pi(\theta)|z,P) = P$, and combining with the equilibrium price function, we define the random variable $z(\theta,u) = z(p(\theta,u))$.

Now, $z(P)$ and P are informationally equivalent if and only if $z(P)$ is strictly increasing. Since log-concavity implies that $\mathbb{E}(\pi(\theta)|x,P)$ is strictly increasing in x, $z(P)$ is strictly increasing in P if and only if $\mathbb{E}(\pi(\theta)|x,P) - P$ is decreasing in P at $x = z(P)$. From the traders' first-order condition, $\mathbb{E}((\pi(\theta)-P)\cdot U'(d(\pi(\theta)-P))|x,P) = 0$, or

$$-cov\left(\pi(\theta)-P; U'(d(\pi(\theta)-P))|x,P\right) = \mathbb{E}((\pi(\theta)-P)|x,P)\cdot \mathbb{E}\left(U'(d(\pi(\theta)-P))|x,P\right).$$

Therefore, $sign(d(x,P)) = sign(\mathbb{E}(\pi(\theta)|x,P) - P)$, and $d(x,P) = 0$ is optimal if and only if $\mathbb{E}(\pi(\theta)|x,P) = P$. This implies that $z(P)$ is invertible if and only if condition 1 holds.

Now, if $z(P)$ is invertible, we have $P = \mathbb{E}(\pi(\theta)|z(P),P) \iff P(z) = \mathbb{E}(\pi(\theta)|x=z,z)$, which validates condition 2. If instead $z(P)$ is not invertible, then there exists values z', P' and P'', such that $z(P') = z(P'') = z'$, and since by construction, $P = \mathbb{E}(\pi(\theta)|z(P),P)$, it must be that $\mathbb{E}(\pi(\theta)|z',P') \neq \mathbb{E}(\pi(\theta)|z',P'')$, and therefore, either $\mathbb{E}(\pi(\theta)|x=z',z') \neq \mathbb{E}(\pi(\theta)|x=z',P') = P'$, or $\mathbb{E}(\pi(\theta)|x=z',z') \neq \mathbb{E}(\pi(\theta)|x=z',P'') = P''$, or both. ■

Proof of Lemma 1. By market-clearing, $z = \hat{x}(P(z))$ and $z' = \hat{x}(P(z'))$, and therefore $z = z'$ if and only if $P(z) = P(z')$. Since $P(z)$ is invertible, observing P is thus equivalent to observing $z = \hat{x}(P(z))$ in equilibrium. The characterization of $H(\cdot|x,P)$ follows immediately from Bayes' Law. ■

Proof of Proposition 1. Follows directly from Theorem 1 and Lemma 1. ■

Proof of Lemma 2. 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

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

With this, we compute $\mathbb{E}(P(z))$:

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

Therefore, W_π is

$$\begin{aligned} W_\pi &= \int_{-\infty}^{\infty} \pi(\theta)\left(\frac{1}{\sigma_P}\phi\left(\frac{\theta}{\sigma_P}\right) - \frac{1}{\sigma_\theta}\phi\left(\frac{\theta}{\sigma_\theta}\right)\right) d\theta \\ &= \int_{-\infty}^{\infty} \pi'(\theta)\left(\Phi\left(\frac{\theta}{\sigma_\theta}\right) - \Phi\left(\frac{\theta}{\sigma_P}\right)\right) d\theta \\ &= \int_{0}^{\infty} \left(\pi'(\theta) - \pi'(-\theta)\right)\left(\Phi\left(\frac{\theta}{\sigma_\theta}\right) - \Phi\left(\frac{\theta}{\sigma_P}\right)\right) d\theta, \end{aligned}$$

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 2, the definition of upside and downside risk, and the fact that $\Phi(\theta/\sigma_\theta) > \Phi(\theta/\sigma_P)$ for all θ (since $\sigma_P > \sigma_\theta$). For part (iv) notice that

$$\begin{aligned} W_{\pi_1}(\sigma_P) - W_{\pi_2}(\sigma_P) &= \int_{0}^{\infty} \Delta(\theta)\left(\Phi\left(\frac{\theta}{\sigma_\theta}\right) - \Phi\left(\frac{\theta}{\sigma_P}\right)\right) d\theta, \\ \text{where } \Delta(\theta) &= \pi_1'(\theta) - \pi_1'(-\theta) - (\pi_2'(\theta) - \pi_2'(-\theta)). \end{aligned}$$

Since π_1 is has more upside risk than π_2, $\Delta(\theta) \geq 0$ for all θ, which implies that $W_{\pi_1}(\sigma_P) - W_{\pi_2}(\sigma_P)$ is increasing in σ_P. ■

Proof of Proposition 3. If $\sigma_{P,1} = \sigma_{P,2} = \sigma_P$, then $W_{\pi_1}(\sigma_{P,1}) + W_{\pi_2}(\sigma_{P,2}) = W_\pi(\sigma_P)$, 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, $W_{\pi_1}(\sigma_{P,1}) + W_{\pi_2}(\sigma_{P,2}) > W_{\pi_1}(\sigma_{P,2}) + W_{\pi_2}(\sigma_{P,1})$, whenever $\sigma_{P,2} > \sigma_{P,1}$ (since π_2 has more upside risk than π_1). ■

Proof of Proposition 4. For any alternative split (π_1, π_2), the monotonicity requirements imply that $0 \leq \pi_1'(\theta) = \pi'(\theta) - \pi_2'(\theta) \leq \pi'(\theta)$. This in turn implies that for all $\theta \geq 0$, $\pi_1^{*\prime}(\theta) - \pi_1^{*\prime}(-\theta) = -\pi'(-\theta) \leq \pi_1'(\theta) - \pi_1'(-\theta)$ and $\pi_2^{*\prime}(\theta) - \pi_2^{*\prime}(-\theta) = \pi'(\theta) \geq \pi_2'(\theta) - \pi_2'(-\theta)$, i.e. π_1 has less downside risk and more upside risk than π_1^*, and π_2 has more downside risk and less upside risk than π_2^*. Moreover,

$$\left(\pi_1'(\theta) - \pi_1'(-\theta)\right) + \left(\pi_2'(\theta) - \pi_2'(-\theta)\right) = \pi'(\theta) - \pi'(-\theta) = \left(\pi_1^{*\prime}(\theta) - \pi_1^{*\prime}(-\theta)\right) + \left(\pi_2^{*\prime}(\theta) - \pi_2^{*\prime}(-\theta)\right)$$

But then, the expected revenue of selling π_1 to the investor pool with $\sigma_{P,1}$ and π_2 to the investor pool with $\sigma_{P,2}$ is $W_{\pi_1}(\sigma_{P,1}) + W_{\pi_2}(\sigma_{P,2}) = W_\pi(\sigma_{P,1}) + W_{\pi_2}(\sigma_{P,2}) - W_{\pi_2}(\sigma_{P,1})$, while the expected revenue from selling π_1^* to the investor pool with $\sigma_{P,1}$ and π_2^* to the investor pool with $\sigma_{P,2}$ is $W_{\pi_1^*}(\sigma_{P,1}) + W_{\pi_2^*}(\sigma_{P,2}) = W_\pi(\sigma_{P,1}) + W_{\pi_2^*}(\sigma_{P,2}) - W_{\pi_2^*}(\sigma_{P,1})$. The difference in revenues is therefore

$W_{\pi_2^*}(\sigma_{P,2}) - W_{\pi_2^*}(\sigma_{P,1}) - (W_{\pi_2}(\sigma_{P,2}) - W_{\pi_2}(\sigma_{P,1}))$, which is positive, since π_2^* contains more upside and less downside risk than π_2, and $\sigma_{P,2} \geq \sigma_{P,1}$ (Theorem 2, part (iv)). ■

Proof of Proposition 5.

Clearly, $z_1 = z_2 = z$ almost surely if and only if $\rho = 1$. If $\beta_1 = \beta_2$ and $\sigma_{u,1}^2 = \sigma_{u,2}^2$, it then follows that $P_{\pi_1}(z_1, z_2) + P_{\pi_2}(z_1, z_2) = P_\pi(z)$, almost surely, if and only if $\rho = 1$. Moreover, it follows from the characterizations of P_{π_1} and P_{π_2} that the price function is no longer additive (even if $\rho = 1$), whenever $\beta_1 \neq \beta_2$ or $\sigma_{u,1}^2 \neq \sigma_{u,2}^2$, unless the markets are informationally linked, and $\beta_1 \sigma_{u,1}^{-2} \neq \beta_2 \sigma_{u,2}^2$. In this last case, we find that signals have different precision, but perfectly correlated errors, so θ and the correlated error can be perfectly inferred from the two signals, i.e. $V \to \infty$ and the wedge disappears. ■

Proof of Lemma 3. The payoff to a share bought in period t is $\pi(\theta_t) + \delta P_\pi(z_{t+1})$, where $P_\pi(z_{t+1})$ is the price in period $t+1$, contingent on the period $t+1$ state z_{t+1}. Since in a forward-looking equilibrium, $P_\pi(z_{t+1})$ is independent of information available at time t (due to the iid assumption, if follows that $\mathbb{E}(P_\pi(z_{t+1}) | x, z_t)$ is simply a constant P_π corresponding to the unconditional expectation of the future price. Since $\mathbb{E}(\pi(\theta_t) | x, P_t)$ is monotone in x, it follows that demand is characterized by a threshold signal above which informed traders purchase the asset in period t. Thus, within period t, we have exactly the same equilibrium characterization as in the static model, and hence the characterization of the equilibrium price defined above. ■

Proof of Proposition 6. If $\pi(\cdot)$ is convex, then by Theorem 1, for any finite $\underline{w}$, there exists $\hat{\delta} < 1$, s.t. $\delta > \hat{\delta}$, $\delta \mathbb{E}(w(z)) > -(1-\delta)\underline{w}$. We therefore need to establish a lower bound for $w(z)$. But if $\pi(\cdot)$ is bounded below, then $\lim_{z \to -\infty} w(z) = 0$, and $w(z)$ is positive for sufficiently high z, so it is necessarily bounded. ■