Journal of Banking & Finance 35 (2011) 1563–1580

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Does more information in stock price lead to greater or smaller idiosyncratic return volatility?

Dong Wook Lee [a], Mark H. Liu [b,*]

[a] Korea University Business School, Seoul 136-701, Republic of Korea
[b] Gatton College of Business and Economics, University of Kentucky, Lexington, KY 40506, USA

A R T I C L E I N F O

Article history:
Received 3 February 2010
Accepted 1 November 2010
Available online 11 November 2010

JEL classification:
G12
G14

Keywords:
Idiosyncratic volatility
Noisy rational expectations equilibrium
Price informativeness

A B S T R A C T

We investigate the relation between price informativeness and idiosyncratic return volatility in a multi-asset, multi-period noisy rational expectations equilibrium. We show that the relation between price informativeness and idiosyncratic return volatility is either U-shaped or negative. Using several price informativeness measures, we empirically document a U-shaped relation between price informativeness and idiosyncratic return volatility. Our study therefore reconciles the opposing views in the following two strands of literature: (1) the growing body of research showing that firms with more informative stock prices have greater idiosyncratic return volatility, and (2) the studies arguing that more information in price reduces idiosyncratic return volatility.

1. Introduction

There is a recent debate in the literature about the relation between price informativeness and idiosyncratic return volatility. A growing body of studies show that more informative stock prices are associated with greater idiosyncratic return volatility. For example, Morck et al. (2000) find that stocks in countries with stronger property rights have higher idiosyncratic volatility. They argue that strong property rights promote informed arbitrage, which capitalizes firm-specific information and increases idiosyncratic return volatility.[1] Many empirical researchers use greater idiosyncratic return volatility as a measure of more informative stock prices (Brockman and Yan, 2009). Another group of studies have contradicting views on this issue and argue that higher idiosyncratic return volatility is an indication of more noise and pricing errors in stock prices (i.e., less informative stock prices). For example, West (1988) shows that more information in price about future dividends leads to lower idiosyncratic volatility. Kelly (2005) finds that firms with better information environments are associated with higher market-model R-square, i.e., smaller idiosyncratic return volatility. Many empirical studies use greater idiosyncratic return volatility as a measure of less informative stock prices and greater asymmetric information between firm insiders and outsiders (Krishnaswami and Subramaniam, 1999).

Understanding the true relation between price informativeness and idiosyncratic return volatility is important, given that an increasing number of studies use idiosyncratic volatility as a price informativeness measure, and the contradicting assumptions they make about the information content of idiosyncratic return volatility. Further, understanding the true information content of idiosyncratic volatility is important for practitioners and policy makers as well. For example, some critics of Regulation FD argue that the regulation has caused less disclosure by firms and has led to increases in stock return volatility, implicitly assuming that less information in stock price leads to higher idiosyncratic return volatility.[2] If higher idiosyncratic volatility is not an indication of less information in stock prices, then Regulation FD may not have the negative impact on corporate disclosure as critics claim.[3]

* Corresponding author. Tel.: +1 859 257 9842; fax: +1 859 257 9688.
E-mail addresses: donglee@korea.ac.kr (D.W. Lee), mark.liu@uky.edu (M.H. Liu).

[1] Jin and Myers (2006), who find that more transparent firms have higher idiosyncratic return volatility, argue that poor investor protection alone cannot explain the difference in idiosyncratic volatility, and opaqueness of the firm also plays an important role.

[2] See, e.g., "Disclosure Rule Gets A Bad Rap" by Jeff Opdyke and Michael Schroeder (the *Wall Street Journal*, June 5, 2001, page C1) and "Deals & Deal Makers: Direct Effects of Disclosure Rule Doubted" by Phyllis Plitch (the *Wall Street Journal*, July 24, 2001, page C14).

[3] We, however, do not take any side in the debate regarding the effectiveness of Regulation FD. We argue only that greater idiosyncratic volatility may not be an indication of less price informativeness.

0378-4266/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2010.11.002

This paper makes an attempt to understand the relation between price informativeness and idiosyncratic return volatility. We examine idiosyncratic return volatility in a multi-asset, multi-period noisy rational expectations model. Stocks are traded among three groups of investors: (1) liquidity traders, whose demand for a stock is exogenous and does not depend on the fundamental value of the stock; (2) informed traders, who incur a cost and obtain a noisy signal about the value of the stock and trade based on this private signal; (3) uninformed discretionary traders (UDTs), who do not have private information about the value of the stock, but infer the information conveyed by the stock price, and whose demand for the stock depends on the price. The equilibrium stock price is affected by both noise and the fundamental value of the stock. We decompose idiosyncratic return volatility into two parts: the noise component, which is caused by the demand of liquidity traders, and the information component, which is driven by the information regarding the fundamental value of the stock. We further decompose the information component of idiosyncratic return volatility into two parts. The first part, which we refer to as the information updating part, represents the fluctuation in price as private information about the fundamental value of the stock is incorporated into price. The second part, which we refer to as the uncertainty resolving part, represents the fluctuation in price due to the resolution of the remaining uncertainty in the stock value (i.e., through the realization of the fundamental stock value that was not previously reflected in price).

We show that the information component of idiosyncratic return volatility first decreases and then increases with price informativeness for the following reason. As more investors choose to produce information and more information is incorporated into the stock price, price informativeness increases. This, in turn, increases the information updating part of idiosyncratic return volatility. As more information is reflected in price, less uncertainty remains about the value of the stock, and the uncertainty resolving part of idiosyncratic return volatility becomes smaller. Therefore, the information updating part increases with price informativeness whereas the uncertainty resolving part decreases with price informativeness. The information component of idiosyncratic return volatility, which is the sum of the information updating part and the uncertainty resolving part, has a U-shaped relation with price informativeness since the average variance over time is the lowest when the uncertainty is resolved gradually. To use a numerical example to illustrate this intuition, suppose that the value of a stock should double from \$1 to \$2 in two periods. Consider three different scenarios: (1) The stock price is extremely uninformative and all information will be revealed only in the second period. The stock return will be 0 in the first period and 100% in the second period. The variance of return in this case is $\frac{1}{2}[(0-50\%)^2+(100\%-50)^2]=0.25$; (2) Information is incorporated into price gradually so that the stock price rises to \$1.5 at the end of the first period, then stock return will be 50% in the first period and 33.3% in the second period. The variance of return in this case is $\frac{1}{2}[(50\%-41.65\%)^2+(33.3\%-41.65\%)^2]=0.007$; (3) The stock price is extremely informative and all information will be revealed in the first period, then stock return will be 100% in the first period and 0 in the second period. The variance of return in this case is $\frac{1}{2}[(100\%-50)^2+(0-50\%)^2]=0.25$. Hence a U-shaped relation between price informativeness and the information component of return volatility.

We also show that the noise component of idiosyncratic return volatility decreases monotonically with price informativeness. This is because as more investors choose to produce information, liquidity trading has a lower impact on stock price since informed investors can better absorb liquidity traders' order flows, which makes the stock price less noisy. This, in turn, reduces the noise component of idiosyncratic return volatility. Therefore, the noise component of idiosyncratic return volatility decreases with price informativeness.

Our main theoretical results on the relation between price informativeness and idiosyncratic volatility are as follows. First, there exist no parameter values such that idiosyncratic return volatility increases monotonically with price informativeness. Second, there exist parameter values such that the relation between price informativeness and idiosyncratic return volatility is U-shaped. This happens when the variance of the demand from liquidity traders is relatively small compared to the variance in the firm's fundamental value. Finally, there exist parameter values such that idiosyncratic return volatility decreases monotonically with price informativeness. This happens when the variance of the demand from liquidity traders is relatively large compared to the variance in the firm's fundamental value. Empirically, we find a U-shaped relation between price informativeness and idiosyncratic volatility. Our sample spans from 1983 to 2004, with slightly more than 3000 US stocks in a given year. The U-shaped relation is observed in virtually every year, and it is robust to using as many as six different measures of price informativeness (which will be detailed in Section 4.1.3).

Our results contribute to the literature in several ways. First, this is the first study to theoretically model and empirically document a U-shaped relation between price informativeness and idiosyncratic volatility. Previous studies focus mainly on a monotonic relation between the two. Second, this paper helps us better understand the source of idiosyncratic return volatility: both noise and information influence return volatility, but while the noise component decreases monotonically with price informativeness, the information component first decreases and then increases with price informativeness. The information component of idiosyncratic return volatility can be further decomposed into two parts: the part representing the fluctuation in price as private information about the fundamental value of the stock is incorporated into price, and the part representing the fluctuation in price due to the resolution of the remaining uncertainty about the value of the stock when the true firm value is revealed. Finally, our results resonate with some studies arguing that idiosyncratic return volatility is not a good measure of how much information is reflected in stock prices; see, e.g., Ashbaugh-Skaife et al. (2005). In other words, our results show that researchers must be cautious when using idiosyncratic return volatility as a measure of price informativeness, since the relation between price informativeness and idiosyncratic volatility may not be monotonic.

The rest of the paper is structured as follows. We relate our work to the existing literature in Section 2. Section 3 develops the theoretical model. Section 4 reports empirical results. We conclude in Section 5. All proofs are confined to Appendix A.

2. Related literature

This paper is related to the theoretical work on the relation between price informativeness and idiosyncratic return volatility. Jin and Myers (2006) develop a model in which firm insiders can capture part of operating cash flows, which cannot be perfectly observed by outsiders. They show a positive relation between price informativeness and idiosyncratic volatility. In their model, the true firm value is never revealed to the market.[4] As a result, the uncertainty resolving part of the idiosyncratic return volatility is

[4] While some information about firm value will be revealed to the public at the end of a period (e.g., quarterly earnings and cash flows), other information will never be observed by investors (e.g., the intrinsic value of the firm and management effort). Therefore, assumptions in both Jin and Myers (2006) and our model are consistent with the reality to some degree. Our assumption is shared by many other studies in the literature (Kyle, 1985).

not modeled in Jin and Myers (2006), and this explains the difference in results between their work and this paper. In a one-period noisy rational expectations model with multiple assets, Ozoguz (2005) shows a negative relation between price informativeness and idiosyncratic return volatility. Her definition of idiosyncratic return volatility captures only the uncertainty resolving part, but not the information updating part of idiosyncratic return volatility. In contrast, we use a dynamic model and capture both parts of the idiosyncratic return volatility that are caused by information, and we show that the information updating part increases with price informativeness, consistent with Jin and Myers (2006), and the uncertainty resolving part decreases with price informativeness, consistent with Ozoguz (2005).

A growing number of studies focus on the properties of the market-model R-square of common stocks and idiosyncratic volatility. Roll (1988) observes that only a small proportion of the actual price movements of individual common stocks can be explained by market and industry influences. Campbell et al. (2001) find that idiosyncratic volatility of common stocks in the United States has increased significantly during the past few decades. Morck et al. (2000) find that firms in developed countries have higher idiosyncratic volatility compared with firms in developing countries, and they argue that this is due to the poor protection of investors' property rights in developing countries.

This paper joins the current debate on whether higher idiosyncratic volatility means more or less information in stock prices. Our results reconcile the opposing views expressed in the two strands of literature we have mentioned in Section 1. Unlike the existing literature, we find a non-monotonic relation between stock price informativeness and idiosyncratic volatility. Existing studies test only a monotonic relation between the two, but their results do not necessarily contradict ours. For example, Morck et al. (2000) compare idiosyncratic volatility in different countries. Since information is more efficient in the industry and country level than in the firm-specific level, their results fall in the right half of the U-shape we have documented, and this may explain the positive relation between stock price informativeness and idiosyncratic volatility in their studies. Kelly (2005) uses firm-specific data in the US, but he uses raw price informativeness measures instead of residual measures as we do. This explains why Kelly (2005) finds a negative relation between price informativeness and idiosyncratic volatility, while we find a U-shaped relation.[5] As we will explain later in Section 4.1.4, since profitability volatility and firm size produce a wide spread of idiosyncratic volatility and price informativeness, we need to control for these two factors to isolate the effect of price informativeness on idiosyncratic return volatility.

Finally, there is an ongoing debate in finance on the relation between idiosyncratic volatility and expected stock returns. Ang et al. (2006) and Guo and Savickas (2010) find a negative cross-sectional relation between idiosyncratic volatility and subsequent stock returns. Guo and Savickas (2008) find that idiosyncratic volatility has negative predictive power for aggregate stock market returns over time in G7 countries. In contrast, Fu (2009) finds a positive relation between idiosyncratic volatility and *contemporaneous* stock returns using the exponential GARCH models. Goyal and Santa-Clara (2003) find that equal-weighted total volatility (mainly idiosyncratic) is positively related to future stock market returns. Even though our focus is on how idiosyncratic volatility is related to price informativeness instead of expected stock returns, we contribute to the literature by helping researchers better understand the sources of idiosyncratic volatility. Specifically, we show that idiosyncratic volatility can be caused by either noise or uncertainty in the fundamental value, and the latter can be further decomposed into information updating volatility and uncertainty resolving volatility.

[5] In unreported results, we also find a negative relation between *raw* price informativeness and idiosyncratic volatility, similar to Kelly (2005).

3. The model

Consider an economy with one riskfree asset and $N+1$ risky assets. Assets 1 to N are individual stocks and asset M is the market index. All assets live for T periods. The liquidation value of asset n at time T is

$$V_{n,T} = \overline{V}_n + \sum_{t=1}^{T} \delta_{n,t}, \text{ for } n = 1, \dots, N, M.$$

$\overline{V}_n$ is the expected liquidation value of asset n, which is announced at time 0. $\delta_{n,t}$ is the innovation on the value of asset n in period t, which becomes known to the public at time t. We assume that for the N individual stocks, $\delta_{n,t}$ is influenced by a systematic component, m_t, and an idiosyncratic component, $f_{n,t}$,

$$\delta_{n,t} = \beta_n m_t + f_{n,t}, \text{ for } n = 1, \dots, N, \tag{1}$$

where β_n is the sensitivity of $\delta_{n,t}$ with respect to m_t, and m_t and $f_{n,t}$ are independent of each other and over time with the following distributions:

$$m_t \sim N(0, \sigma_m^2), \quad \text{and } f_{n,t} \sim N(0, \sigma_{f,n}^2). \tag{2}$$

For the market index, we have

$$\delta_{M,t} = m_t. \tag{3}$$

The risky asset $n \in \{1,\dots,N,M\}$ has a physical supply of Y_n. For simplicity, we assume that the riskfree asset is in perfectly elastic supply, and the net return on it is normalized to zero.

3.1. An equivalent economy

Given the structure of the payoffs, we can consider an equivalent representation of the original economy, similar to Ozoguz (2005). In the equivalent economy, there are $N+1$ risky assets. The liquidation value of asset n at time T is

$$u_{n,T} = \bar{u}_n + \sum_{t=1}^{T} \eta_{n,t}, \text{ for } n = 1, \dots, N, N+1,$$

where $\bar{u}_n = \overline{V}_n - \beta_n \overline{V}_M$ and $\eta_{n,t} = f_{n,t}$ for $n = 1,\dots,N$, and $\bar{u}_{N+1} = \overline{V}_M$ and $\eta_{N+1,t} = m_t$. That is, asset $n \in \{1,\dots,N\}$ in the equivalent economy is equivalent to a portfolio of one share of stock n plus $-\beta_n$ shares of the market index in the original economy. Therefore, in the equivalent economy, the physical supplies of the stocks are $y_n = Y_n$ for $n = 1,\dots,N$, and $y_{N+1} = Y_M + \sum_{n=1}^{N} \beta_n Y_n$. For simplicity, we call the liquidation value of asset n at time T as the fundamental value of asset n, i.e., $u_n = u_{n,T}$. We will first focus our analysis on the equivalent economy and derive stock prices and return volatility. Later, we will go back to the original economy and derive stock prices and idiosyncratic return volatility in the original economy.

3.2. Investors' information production and utility maximization problem

Three types of traders (investors) are in the market: liquidity traders, informed traders, and uninformed discretionary traders (UDTs). The aggregate demand (in terms of number of shares) from liquidity traders for asset $n \in \{1,\dots,N,N+1\}$ at time $t \in \{0,1,\dots,T-1\}$ is $z_{n,t+1}$, which has the following distribution:

$$z_{n,t+1} \sim N(0, \sigma_{z,n}^2). \tag{4}$$

We further assume that $z_{n,t+1}$ is independent across the $N+1$ securities and across time.

There is a continuum of UDTs over the interval [0,1]. UDTs do not know the value of $\eta_{n,t+1}$ at time t. They can, however, choose to either remain uninformed or acquire a noisy signal about $\eta_{n,t+1}$ at a cost and become informed.[6] Specifically, the value of $\eta_{n,t+1}$ can be decomposed into two parts

$$\eta_{n,t+1} = \theta_{n,t+1} + \epsilon_{n,t+1}, \tag{5}$$

where $\theta_{n,t+1}$ and $\epsilon_{n,t+1}$ are independent of each other, across securities, and across time,[7] with the following distributions:

$$\theta_{n,t+1} \sim N(0, \sigma^2_{\theta,n}), \quad \text{and } \epsilon_{n,t+1} \sim N(0, \sigma^2_{\epsilon,n}). \tag{6}$$

By incurring a cost of C_n at time t, a UDT observes the value of $\theta_{n,t+1}$ and becomes an informed trader.

All investors have the same exponential utility function of consumption over the time T wealth, W_T:

$$u(W_T) = -e^{-aW_T} \tag{7}$$

where a is the absolute risk-aversion coefficient. At each time $t \in \{0,1,\ldots,T-1\}$, the following events occur sequentially: (1) investor j starts with $X^j_t = (X^j_{1,t}, \ldots, X^j_{N+1,t})'$ shares of stocks and B^j_t dollars of cash (the riskfree asset); (2) the investor decides whether to acquire information on each of the $N+1$ stocks; (3) trading takes place, and stock prices $P_t = (P_{1,t},\ldots,P_{N+1,t})'$ are determined in equilibrium; (4) the investor carries X^j_{t+1} shares of stocks and B^j_{t+1} dollars of cash to the next period.

We use $I^j_{n,t}$ to denote investor j's decision on whether to acquire information about stock n at time t, with value 1 if yes and 0 otherwise. The cash holding of investor j changes over time as follows:

$$B^j_{t+1} = B^j_t + P'_t(X^j_t - X^j_{t+1}) - (I^j_t)'C,$$

where $I^j_t = (I^j_{1,t}, \ldots, I^j_{N+1,t})'$ and $C = (C_1, C_2, \ldots, C_{N+1})'$. At time T, the investor's final wealth is

$$W^j_T = B^j_T + u'_T X^j_T.$$

where $u = (u_1, \ldots, u_{N+1})'$ is the vector of the fundamental values.

3.3. Equilibrium in the last period

At time T, all information becomes public, so stock prices are as follows:

$$P_{n,T} = u_{n,T} \text{ for } n = 1, \ldots, N, N+1. \tag{8}$$

At time $T-1$, a number of $\mu_{n,T-1} \in (0,1)$ UDTs acquire information about stock n and become informed. The following proposition summarizes the equilibrium stock prices at time $T-1$:

Proposition 1 (Stock prices in the last period). *The price for stock $n \in \{1,\ldots,N,N+1\}$ at time $T-1$ is*

$$P_{n,T-1} = \bar{u}_n + \sum_{t=1}^{T-1} \eta_{n,t} + \alpha^\theta_{n,T-1}\theta_{n,T} + \alpha^z_{n,T-1}(z_{n,T} - y_n), \tag{9}$$

where $\alpha^\theta_{n,T-1} > 0$ and $\alpha^z_{n,T-1} > 0$ are given in Eqs. (A.12) and (A.13).

The equilibrium price of stock n at time $T-1$ is a linear combination of the private information held by informed traders, $\theta_{n,T}$, and the demand from liquidity traders, $z_{n,T}$. The price partially incorporates the private information but does not fully reveal it, since the price is also influenced by $z_{n,T}$. The UDTs can only infer part of the private information through price.

3.4. Decomposition of return volatility

The following proposition summarizes the price function at time $t \in \{0,1,\ldots T-2\}$.

Proposition 2 (Equilibrium stock prices). *The price of stock $n \in \{1,\ldots,N,N+1\}$ at time $t \in \{0,1,\ldots T-2\}$ is*

$$P_{n,t} = \bar{u}_n + \sum_{s=1}^{t} \eta_{n,s} + \alpha^\theta_{n,t}\theta_{n,t+1} + \alpha^z_{n,t}(z_{n,t+1} - y_n), \tag{10}$$

where $\alpha^\theta_{n,t}$ and $\alpha^z_{n,t}$ are characterized in Eqs. (A.24) and (A.25).

At any time from 0 to $T-2$, the stock price is a linear combination of the private information held by informed traders, $\theta_{n,t+1}$, and the demand by liquidity traders, $z_{n,t+1}$. The price partially incorporates the private information but does not fully reveal it. UDTs decide whether or not to produce information on each of the $N+1$ stocks at time t. Since the marginal benefit from acquiring information on stock n at time t decreases with the number of investors who choose to acquire information, equilibrium is reached when the marginal benefit equals the marginal cost of acquiring information on stock n.

Now, we define the return of stock n in period t as

$$r_{n,t} = P_{n,t} - P_{n,t-1} = (1 - \alpha^\theta_{n,t-1})\theta_{n,t} + \alpha^\theta_{n,t}\theta_{n,t+1} + \epsilon_{n,t} + \alpha^z_{n,t}(z_{n,t+1} - y_n) + \alpha^z_{n,t-1}(z_{n,t} - y_n).$$

The variance of the stock return is therefore

$$Var(r_{n,t}) = (1 - \alpha^\theta_{n,t-1})^2\sigma^2_{\theta,n} + (\alpha^\theta_{n,t})^2\sigma^2_{\theta,n} + \sigma^2_{\epsilon,n} + (\alpha^z_{n,t})^2\sigma^2_{z,n} + (\alpha^z_{n,t-1})^2\sigma^2_{z,n}.$$

The first component, $\left(1 - \alpha^\theta_{n,t-1}\right)^2\sigma^2_{\theta,n}$, is caused by the realization of the information about $\theta_{n,t}$ that is not previously reflected in price $P_{n,t-1}$. The second component, $\left(\alpha^\theta_{n,t}\right)^2\sigma^2_{\theta,n}$, is caused by the incorporation of information about $\theta_{n,t+1}$ in price $P_{n,t}$. The third component, $\sigma^2_{\epsilon,n}$, is caused by the imperfection of information production about $\eta_{n,t+1}$, i.e., the noise in the information production process. The fourth and fifth components, $(\alpha^z_{n,t})^2\sigma^2_{z,n}$ and $\left(\alpha^z_{n,t-1}\right)^2\sigma^2_{z,n}$, are caused by liquidity trading at time t and time $t-1$, respectively.

Instead of focusing on the time period, we focus on the sources of the return volatilities, $\theta_{n,t+1}$, $\epsilon_{n,t+1}$, and $z_{n,t+1}$. The stock return from time t to time $t+1$, $r_{n,t+1}$, is

$$r_{n,t+1} = \left(1 - \alpha^\theta_{n,t}\right)\theta_{n,t+1} + \alpha^\theta_{n,t+1}\theta_{n,t+2} + \epsilon_{n,t+1} + \alpha^z_{n,t+1}(z_{n,t+2} - y_n) + \alpha^z_{n,t}(z_{n,t+1} - y_n).$$

In $r_{n,t}$, $\left[\alpha^\theta_{n,t}\theta_{n,t+1} + \alpha^z_{n,t}(z_{n,t+1} - y_n)\right]$ is caused by $(\theta_{n,t+1}, z_{n,t+1})$, and in $r_{n,t+1}$, $\left[(1 - \alpha^\theta_{n,t})\theta_{n,t+1} + \epsilon_{n,t+1} + \alpha^z_{n,t-1}(z_{n,t} - y_n)\right]$ is caused by $(\theta_{n,t+1}, \epsilon_{n,t+1}, z_{n,t+1})$. We define instead

$$r'_{n,t} = \left[\alpha^\theta_{n,t}\theta_{n,t+1} + \alpha^z_{n,t}(z_{n,t+1} - y_n)\right] + \left[(1 - \alpha^\theta_{n,t})\theta_{n,t+1} + \epsilon_{n,t+1} + \alpha^z_{n,t-1}(z_{n,t} - y_n)\right]$$

This newly defined return process reflects the return components caused by $(\theta_{n,t+1}, \epsilon_{n,t+1}, z_{n,t+1})$ only. Note that the sum of $r'_{n,t}$ over time equals the sum of $r_{n,t}$ over time. Therefore, we will focus on this return instead. The variance of the newly defined return is

[6] Even though we assume that the stock value is affected by both the market return and the idiosyncratic component, our model setup is different from CAPM in the sense that CAPM assumes no asymmetric information among investors. In contrast, our model falls into the literature examining the asset pricing implications under asymmetric information. Whether CAPM holds under asymmetric information is an interesting research topic, but it is out of the scope of the current paper. We thank an anonymous referee for pointing this out.

[7] Since $\eta_{n,t+1}$'s are the idiosyncratic components of the N securities and the market factor in the original economy, they should be independent of each other cross-sectionally by definition. It is not unduly restrictive to assume that $\eta_{n,t+1}$'s are independent over time since they are the *innovative* parts of the fundamental values at time $t+1$.

$$Var(r'_{n,t}) = \left[(1-\alpha^{\theta}_{n,t})^2 + (\alpha^{\theta}_{n,t})^2\right]\sigma^2_{\theta,n} + \sigma^2_{\epsilon} + 2(\alpha^{z}_{n,t})^2\sigma^2_{z,n}$$

The first part of the return volatility,

$$V^{Info}(r'_{n,t}) = \left[\left(1-\alpha^{\theta}_{n,t}\right)^2 + \left(\alpha^{\theta}_{n,t}\right)^2\right]\sigma^2_{\theta,n},$$

is the return volatility caused by information, corresponding to the information component of return volatility that we refer to in the introduction. The second part,

$$V^{Noise}(r'_{n,t}) = \sigma^2_{\epsilon} + 2(\alpha^{z}_{n,t})^2\sigma^2_{z,n},$$

is the return volatility caused by noise, corresponding to the noise component of idiosyncratic return volatility that we refer to in the introduction. Furthermore, the information component of return volatility has two parts: the first part, $\left(\alpha^{\theta}_{n,t}\right)^2\sigma^2_{\theta,n}$, is the information updating part, which is caused by the incorporation of private information in the stock price; the second part, $\left(1-\alpha^{\theta}_{n,t}\right)^2\sigma^2_{\theta,n}$, is the uncertainty resolving part, which is caused by the realization of the residual private information that was not previously reflected in the stock price.

The following proposition shows that the information component of return volatility first decreases with the number of information producers, $\mu_{n,t}$, and then increases with $\mu_{n,t}$. That is, $V^{Info}\left(r'_{n,t}\right)$ has a U-shaped relation with respect to $\mu_{n,t}$. The noise component of return volatility decreases monotonically with $\mu_{n,t}$.

Proposition 3 (Relation between the number of informed traders and the information and noise components of return volatility).

(i) *There is a U-shaped relation between the number of information producers and the information component of return volatility. That is, for every stock* $n \in \{1,2,\ldots,N+1\}$, $\frac{\partial V^{Info}(r'_{n,t})}{\partial \mu_{n,t}} < 0$ *for* $\mu_{n,t} \in [0,\hat{\mu}_{n,t})$ *and* $\frac{\partial V^{Info}(r'_{n,t})}{\partial \mu_{n,t}} \geqslant 0$ *for* $\mu_{n,t} \in [\hat{\mu}_{n,t},1]$, *where*

$$\hat{\mu}_{n,t} = \frac{\sqrt{\left(2a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right)^2 + 4\sigma^2_{\theta,n}a^2\sigma^2_{z,n}\Sigma^2_{n,t+1}} - \left(2a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right)}{2\sigma^2_{\theta,n}}. \tag{11}$$

(ii) *The noise component of return volatility decreases monotonically with the number of information producers. That is, for every stock* $n \in \{1,2,\ldots,N+1\}$, $\frac{\partial V^{Noise}(r'_{n,t})}{\partial \mu_{n,t}} < 0$ *for any* $\mu_{n,t} \in [0,1]$.

The information updating part of return volatility, $\left(\alpha^{\theta}_{n,t}\right)^2\sigma^2_{\theta}$, increases with the number of information producers, $\mu_{n,t}$. The uncertainty resolving part of return volatility, $\left(1-\alpha^{\theta}_{n,t}\right)^2\sigma^2_{\theta}$, decreases with $\mu_{n,t}$. The sum of the two parts, which is the information component of return volatility, is the lowest when the number of information producers, $\mu_{n,t}$, is moderate so that neither part is too large. In contrast, when the number of information producers is either too large or too small, the sum is large. Therefore, we have a U-shaped relation between the number of information producers and information component of return volatility. We want to emphasize that the U-shaped relation holds for all parameter values, since $\hat{\mu}_{n,t}$ is always between 0 and 1. The noise component of return volatility decreases monotonically with $\mu_{n,t}$ since more information producers reduce the impact of noise on price and reduce the noise component of return volatility.

We use a numerical example here to show our results in Proposition 3.

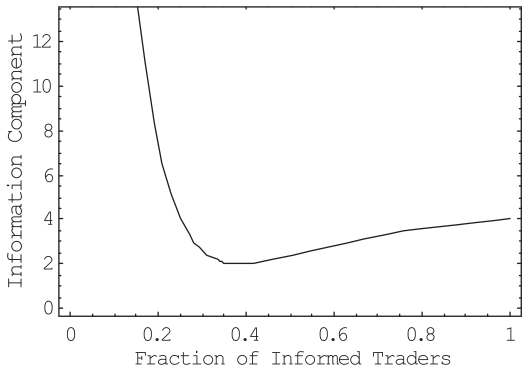

Fig. 1. Relation between the fraction of informed traders and information component return volatility.

Example 1. Assume the following parameter values: $\sigma_{\theta,n} = 2$, $\sigma_{z,n} = 1$, $\sigma_{\epsilon,n} = 0.5$, and $a = 1$. We first look at the relation between the fraction of informed traders and the information component of return volatility. When the fraction of informed traders increases from 0% to around 40%, the information component of return volatility decreases monotonically. When the fraction of informed traders increases from 40% to 100%, the information component of return volatility increases monotonically. Fig. 1 shows the U-shaped relation between the fraction of informed traders and the information component of return volatility. We then look at the relation between the fraction of informed traders and the noise component of return volatility. When the fraction of informed traders increases from 0% to 100%, the noise component of return volatility decreases monotonically. Fig. 2 shows the relation between the fraction of informed traders and the noise component of return volatility.[8]

3.5. Relation between price informativeness and return volatility

We define price informativeness as the uncertainty reduction in the value of the stock due to the knowledge of the price:

$$\Psi_{n,t} = 1 - \frac{Var_t(u_n|P_{n,t})}{Var_t(u_n)}. \tag{12}$$

where $Var_t(u_n|P_{n,t})$ is the variance of stock n's fundamental value conditional on the stock price, $P_{n,t}$, while $Var_t(u_n)$ is the unconditional variance of stock n's value at time t without observing the stock price.[9] This measure captures the fraction of information about the stock's fundamental value that is incorporated into the price. For example, when the price incorporates no information at all, then $Var_t(u_n|P_{n,t}) = Var_t(u_n)$ and $\Psi_{n,t} = 0$. In contrast, when the price is fully revealing, then $Var_t(u_n|P_{n,t}) = 0$ and $\Psi_{n,t} = 1$.

The following proposition characterizes the price informativeness of stock n, and its relation with the number of information producers in the economy and the information production cost.

Proposition 4 (Relation between the price informativeness and the number of informed traders). *The price informativeness of stock n at time t is*

[8] Because many of the parameters in our model (e.g., $\sigma^2_{\theta,n}$ and $\sigma^2_{\epsilon,n}$) are not directly observable, our model predictions are mostly qualitative instead of quantitative.

[9] Kyle (1985) interprets and defines price informativeness in a similar fashion.

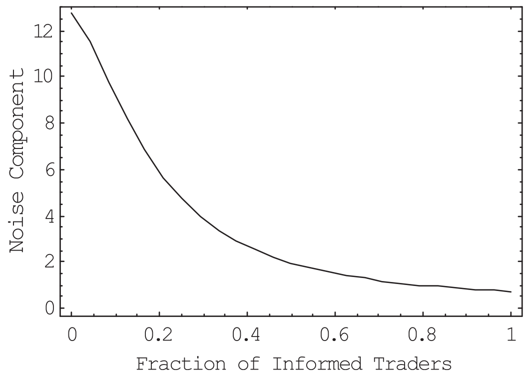

Fig. 2. Relation between the fraction of informed traders and noise component return volatility.

$$\Psi_{n,t} = \frac{\left(\alpha^{\theta}_{n,t}\right)^2 \sigma^4_{\theta}}{(T-t)\left(\sigma^2_{\epsilon}+\sigma^2_{\theta}\right)\left[\left(\alpha^{\theta}_{n,t}\right)^2\sigma^2_{\theta}+\left(\alpha^{z}_{n,t}\right)^2\sigma^2_{z}\right]}, \quad (13)$$

which increases with the number of informed traders about stock n, $\mu_{n,t}$, *and decreases with the information production cost for stock n,* C_n. *That is,* $\frac{\partial\Psi_{n,t}}{\partial\mu_{n,t}} > 0$ *and* $\frac{\partial\Psi_{n,t}}{\partial C_n} < 0$.

As more UDTs choose to produce information and become informed traders, the price becomes more informative. This is because the demand of the informed traders reflects the fundamental values of the stock and this information is partially incorporated into the price. Since in equilibrium the fraction of informed traders in the economy decreases with the information production cost, the price informativeness decreases with the information production cost.

The following proposition summarizes the relation between price informativeness and return volatility.

Proposition 5 (Relation between price informativeness and return volatility).

(i) *No parameter values exist such that the return volatility increases monotonically with the price informativeness* $\left(\frac{\partial Var(r'_{n,t})}{\partial\Psi_{n,t}} \geqslant 0 \text{ for all } \Psi_{n,t}\right)$; *(ii) There exist parameter values such that the relation between the return volatility and the price informativeness is U-shaped* $\left(\frac{\partial Var(r'_{n,t})}{\partial\Psi_{n,t}} < 0 \text{ for } \Psi_{n,t} < \widetilde{\Psi}_{n,t} \text{ and } \frac{\partial Var(r'_{n,t})}{\partial\Psi_{n,t}} \geqslant 0 \text{ for } \Psi_{n,t} \geqslant \widetilde{\Psi}_{n,t}\right)$; *(iii) There exist parameter values such that the return volatility decreases monotonically with the price informativeness* $\left(\frac{\partial Var(r'_{n,t})}{\partial\mu_{n,t}} \geqslant 0 \text{ for all } \Psi_{n,t}\right)$.

We use a numerical example here to show that parameter values exist so that there is a U-shaped relation between price informativeness and return volatility, as predicted by part (ii) of Proposition 5.

Example 2. Assume the following parameter values: $\sigma_{\theta,n} = 2$, $\sigma_{z,n} = 1$, $\sigma_{\epsilon,n} = 0.5$, $T - t = 1$, and $a = 1$. When the price informativeness increases from 0% to around 70%, the return volatility decreases monotonically. When the price informativeness increases from 70% to 100%, the return volatility increases monotonically. Fig. 3 shows the U-shaped relation between the price informativeness and the return volatility.

In the above example, the variance of the value of informed traders' private information is greater than the variance on the demand of liquidity trading ($\sigma^2_{\theta,n} = 4 > \sigma^2_{z,n} = 1$), and we observe

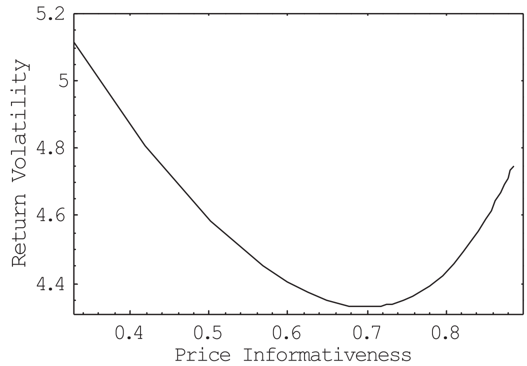

Fig. 3. Relation between price informativeness and return volatility when the variance of the noise is low.

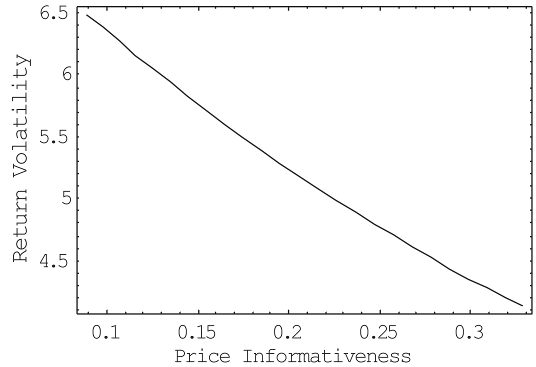

Fig. 4. Relation between price informativeness and return volatility when the variance of the noise is high.

a U-shaped relation between price informativeness and return volatility. In the next example, we show that when the variance of the value of informed traders' private information is smaller than the variance on the demand of liquidity trading, the return volatility decreases monotonically with the price informativeness. This next example also shows the existence of parameter values such that there is a negative relation between price informativeness and return volatility, as predicted by part (iii) of Proposition 5.

Example 3. Assume the following parameter values: $\sigma_{\theta,n} = 1$, $\sigma_{z,n} = 1.2$, $\sigma_{\epsilon,n} = 0.5$, $T - t = 1$, and $a = 1$. The return volatility decreases monotonically with the price informativeness. Fig. 4 shows the negative relation between the price informativeness and the return volatility.

3.6. Back to the original economy

A share of stock $n \in \{1,2,\ldots,N\}$ in the original economy is equivalent to a portfolio of one share of stock n and β_n shares of stock $N+1$ in the equivalent economy. We use the superscript O to denote prices in the original economy. The following proposition summarizes the stock prices over time in the original economy.

Proposition 6 (Stock prices in the original economy). *At time* $t \in \{0,1,\ldots,T\}$, *the price of stock* $n \in \{1,2,\ldots,N\}$ *in the original economy is*

$$P^{O}_{n,t} = P_{n,t} + \beta_n P_{N+1,t},$$

and the price of the market index (i.e., stock M) in the original economy is

$$P^O_{M,t} = P_{N+1,t},$$

where $P_{n,t}$ is given in Eqs. (8)–(10).

The stock prices of the N individual stocks are correlated, since they are all affected by the market index. The return volatility of stock $n \in \{1, 2, \ldots, N\}$ at time t is now

$$Var^O\left(r'_{n,t}\right) = \beta_n\left\{\left[\left(1-\alpha^{\theta}_{N+1,t}\right)^2 + (\alpha^{\theta}_{N+1,t})^2\right]\sigma^2_{N+1,\theta} + \sigma^2_{N+1,\epsilon} + 2(\alpha^z_{n,t})^2\sigma^2_{N+1,z}\right\} + \left\{\left[\left(1-\alpha^{\theta}_{n,t}\right)^2 + \left(\alpha^{\theta}_{n,t}\right)^2\right]\sigma^2_{n,\theta} + \sigma^2_{n,\epsilon} + 2\left(\alpha^z_{n,t}\right)^2\sigma^2_{n,z}\right\}, \quad (14)$$

That is, the return volatility can be decomposed into systematic return volatility and idiosyncratic return volatility. Following the empirical literature, we define the (relative) idiosyncratic return volatility as follows:

$$IV_{n,t} = \frac{\left[\left(1-\alpha^{\theta}_{n,t}\right)^2 + \left(\alpha^{\theta}_{n,t}\right)^2\right]\sigma^2_{n,\theta} + \sigma^2_{n,\epsilon} + 2\left(\alpha^z_{n,t}\right)^2\sigma^2_{n,z}}{Var^O(r'_{n,t})},$$

which is essentially one minus the market-model R-square.

In the original economy, the price informativeness of an individual stock depends on the informativeness of both the systematic component and the idiosyncratic component of the stock's fundamental value. However, since all stocks are influenced by the same systematic factor, the cross-sectional difference in price informativeness is driven only by the idiosyncratic component. For this reason, we keep our definition of price informativeness the same as that in the equivalent economy. That is, we define price informativeness of an individual stock in the original economy as the uncertainty reduction in the value of the stock's idiosyncratic value due to the knowledge of the price.

The following proposition summarizes the relation between the price informativeness and the (relative) idiosyncratic return volatility. The results follow naturally from the results in Proposition 5.

Corollary 1 (Relation between price informativeness and idiosyncratic return volatility).

(i) *No parameter values exist such that the idiosyncratic return volatility increases monotonically with the price informativeness* $\left(\frac{\partial Var^O(r'_{n,t})}{\partial \Psi_{n,t}} \geqslant 0 \text{ for all } \Psi_{n,t}\right)$; *(ii) There exist parameter values such that the relation between the idiosyncratic return volatility and the price informativeness is U-shaped* $\left(\frac{\partial Var^O(r'_{n,t})}{\partial \Psi_{n,t}} < 0 \text{ for } \Psi_{n,t} < \widetilde{\Psi}_{n,t} \text{ and } \frac{\partial Var^O(r'_{n,t})}{\partial \Psi_{n,t}} \geqslant 0 \text{ for } \Psi_{n,t} \geqslant \widetilde{\Psi}_{n,t}\right)$; *(iii) There exist parameter values such that the idiosyncratic return volatility decreases monotonically with the price informativeness* $\left(\frac{\partial Var^O(r_{n,t})}{\partial \mu_{n,t}} \geqslant 0 \text{ for all } \Psi_{n,t}\right)$.

4. Empirical tests

In this section, we empirically test the relation between price informativeness and idiosyncratic return volatility. We first describe our sample and how we measure price informativeness and idiosyncratic return volatility. Then, we report our results on the relation between the two. Finally, we perform an array of robustness checks to show that our empirical results are robust to different specifications.

4.1. Sample and data

4.1.1. Sample

We construct the sample with non-financial and non-utility stocks whose CUSIP identifier is either 10 or 11.[10] We also exclude stocks whose industry classification is not obvious (i.e., SIC codes are missing and thus have a value of zero). We examine their weekly stock returns and require stocks to have a full year of weekly return data. The sample period spans from 1983 to 2004.

4.1.2. Measuring idiosyncratic volatility

We measure idiosyncratic volatility by estimating the following equation:

$$r_{i,w,t} = \alpha_{i,t} + \beta_{i,t} \times r_{m,w,t} + \gamma_{i,t} \times r_{i2,w,t} + \varepsilon_{i,w,t}, \quad (15)$$

where $r_{i,w,t}$ is stock i's simple return in week w (Wednesday close to the next Wednesday close) in year t, $r_{m,w,t}$ and $r_{i2,w,t}$ are, respectively, the contemporaneous returns on the market portfolio and on the industry portfolio (based on the 2-digit SIC). To avoid any spurious results, both the market and the industry portfolios are constructed without stock i; remaining stocks are then value-weighted. The firm-specific return volatility (FSRV hereafter) for firm i in year t, $FSRV_{it}$, is defined as one minus the R-square from regression (15).[11,12]

4.1.3. Measuring price informativeness

We use six measures of price informativeness that are widely used in the literature. Each measure is transformed so that a higher value of the measure corresponds to greater price informativeness. Since we measure FSRV every year, the informativeness measures are also calculated on an annual basis.

4.1.3.1. Information-based trading (N_PIN). Our first measure is based on the PIN in Easley et al. (2010) and it is widely used in the literature as a measure of price informativeness (e.g., Kang, 2010). It is the market maker's estimate of the probability that a certain trade is based on private information about the stock. For firms with informative stock prices, since most firm-specific information is already incorporated into the price, the probability that the market maker is trading against an informed trader is small. On the contrary, for stocks with less informative stock prices, the probability that any given trade is information-based is higher; thus, the PIN will also be higher. We make the following transformation to create an informativeness measure based on PIN: $N\_PIN_t = -\log(\text{PIN for year } t)$. Our PIN measures cover all NYSE/Amex common stocks from 1983 to 2001 for which PINs can be estimated.

4.1.3.2. Price impact (N_PIM). Following Amihud (2002) and similar to Thapa and Poshakwale (2010), we measure price impact by the absolute daily return divided by the daily dollar volume of trade (in

[10] Both industry and CUSIP information is from CRSP. In particular, we use the historical SIC and CUSIP, not the ones in the header file.

[11] Our definition of FSRV can be viewed as relative FSRV, and this is the definition used in most studies in the literature (e.g., Morck et al., 2000). A few studies use absolute FSRV, which is the root mean square error of the regression of the firm's stock return on the market (and industry) return. As we will show later in robustness check, we obtain broadly consistent results when we use absolute FSRV.

[12] While our model explicitly predicts a U-shaped relation between absolute FSRV and price informativeness, it also implicitly predicts a U-shaped relation between relative FSRV and price informativeness. We assume that investors produce information about the firm-specific component but not about the systematic component. Therefore, while the absolute idiosyncratic volatility changes with how much information is produced by investors (i.e., price informativeness), the market volatility remains constant when price informativeness changes. That is, changes in relative FSRV are driven by changes in absolute FSRV in our model. We thank the anonymous referee for pointing this out.

millions), averaged over the year (namely, PIM). PIM measures how easily investors can liquidate a stock without severely affecting the price. A larger value of PIM means, upon a liquidity shock, one has to incur a greater dollar loss to sell the stock. Efficiently priced stocks tend to be more liquid.[13] Therefore, the price impact of any given trade is greater for stocks with less informative prices. We make the following transformation to create an information measure based on the price impact: $N\_PIM_t = -\log(0.0001 + PIM_t)$.

4.1.3.3. Analyst earnings forecast error (N_ERR). Following Krishnaswami and Subramaniam (1999), we assume that as analysts allocate more resources, their research will produce a more precise forecast, and more information about the fundamental value of the firm will be incorporated into the price. Following this logic, we use analyst earnings forecast error as another price informativeness measure. Each month, we calculate forecast error as the absolute value of the difference between the mean earnings forecast for the next fiscal year and the actual earnings per share (EPS), scaled by the stock price in that month; then, we average the monthly forecast errors over the year. We make the following transformation: $N\_ERR_t = -\log$ (0.0001 + average forecast error during year t).

4.1.3.4. Analyst earnings forecast dispersion (N_DSP). As analysts allocate more resources and intensify their research activities, it is also reasonable to assume that their opinions will converge to a correct parameter value. Following Krishnaswami and Subramaniam (1999), we use analyst earnings forecast dispersion as another price informativeness measure. Each month, we obtain forecast dispersion (i.e., standard deviation) for EPS of the next fiscal year, scaled by the mean forecast; we then average the monthly scaled forecast dispersions over the year. We make the following transformation: $N\_DSP_t = -\log$ (0.0001 + average forecast dispersion during year t).

4.1.3.5. The length of the firm's public trading history (AGE). Availability of public trading history will surely reduce the information production costs, and encourage the gathering and incorporation of information into stock prices. Specifically, we construct this measure by counting the number of years during which the stock is publicly traded (i.e., the number of days divided by 365). Following Pastor and Veronesi (2003), we make the following transformation: $AGE_t = -1/$(the number of publicly traded years as of the end of year t).

4.1.3.6. Institutional ownership (IO). It is widely accepted that institutional investors are more sophisticated than retail investors (Rubin and Smith, 2009). Therefore, institutional ownership for a stock can serve as a direct measure of the amount of information incorporated into the stock. Institutional ownership is measured as the fraction of shares that are held by institutions who file the 13F form with the Securities and Exchange Commission. If a stock is not held by any of those 13F-reporting institutions, we assume that the institutional ownership of that stock is zero. Since this is quarterly information, we use the average over four quarters within a year. We make the following transformation: $IO_t = \log(0.0001 +$ fraction of shares held by institutions for year t).

It is worth mentioning that in the theoretical part of the paper, we have defined price informativeness as the uncertainty reduction in the value of the stock due to the knowledge of the price, characterized by Eq. (12). Because we cannot directly observe and measure the unconditional variance of the stock value $Var_t(u_n)$, we can only use the above six measures to proxy for price informativeness. When the stock price becomes more informative, the value of $Var_t(u_n|P_{n,t})$ decreases, and the value of $\Psi_{n,t}$ (our theoretical definition of price informativeness) increases. Empirically, when the price becomes more informative, information-based trading should decrease, price impact of trading should be smaller, and analyst earnings forecasts should be more accurate and less dispersed. As the firm has a longer trading history and more ownership by sophisticated institutional investors, its stock price should also become more informative. That is why we choose these six widely-used measures to proxy for our theoretical definition of price informativeness $\Psi_{n,t}$.

4.1.4. Controlling for the volatility of profitability and firm size

Firms with more volatile cash flows and earnings, *ceteris paribus*, tend to have higher idiosyncratic stock return volatility (e.g., Pastor and Veronesi, 2003). It is also well documented in the literature that larger firms tend to have lower idiosyncratic volatility (e.g., Roll, 1988; Kelly, 2005). At the same time, both firm size and cash flow volatility greatly affect price informativeness. When estimating the relative importance of cash-flow and expected-return news for firm-level stock returns, he finds that the variance of expected-return news is approximately one-half of the variance of cash-flow news for excess returns. Given that firm size and cash flow volatility greatly affect idiosyncratic volatility and price informativeness, we first regress price informativeness on firm size and cash flow volatility and then examine the relation between *residual* price informativeness and idiosyncratic volatility.[14] This way, we can isolate the effects of price informativeness from the effects of firm size and cash flow volatility on idiosyncratic return volatility.

Firm size in year t is measured as the natural log of the firm's market capitalization as of the end of year $t-1$ (*MVE*). As a measure of the volatility of corporate profitability, we use the standard deviation of the firm's return on equity over the sample period, which we call *SROE*. We then estimate the following cross-sectional regression of one of our informativeness measures on *MVE* and *SROE*, and use the residuals as the informativeness measure in the analysis.

$$INFO_{i,t} = \alpha_t + \beta_t \times SROE_{i,t} + \gamma_t \times MVE_{i,t} + \varepsilon_{i,t} \tag{16}$$

where $INFO_{i,t}$ is one of our transformed informativeness measures for stock i in year t. We estimate this regression within the same NYSE size quintile to avoid imposing a simple linear relationship between firm size and a noise measure. We, however, also report results based on one single cross-sectional regression as a robustness check.[15]

4.2. Empirical results

We report empirical results in this section. We first provide the summary statistics for the variables used in this paper. Then, we document a U-shaped relation between price informativeness

[13] When a firm is informationally efficient, the market makers face less adverse selection from potential informed traders. The probability of any given trade is information-based is small, and the market maker tends to adjust the stock price less when there is an imbalance of trade order flows.

[14] Our theoretical model predicts a U-shaped relation between price informativeness and idiosyncratic volatility under certain parameter values (see Fig. 3 and Proposition 5). However, these results are based on comparative statics. Specifically, it shows that, *holding everything else constant* (including firm size and cash flow volatility), idiosyncratic volatility first decreases and then increases with price informativeness. Empirically, because different firms have different characteristics that affect idiosyncratic volatility and price informativeness, we have to control for these firm characteristics first before examining the relation between price informativeness and idiosyncratic volatility.

[15] If we do not control for firm size and profitability volatility in price informativeness, we find a monotonic negative relation between idiosyncratic volatility and *raw* price informativeness, similar to results in Kelly (2005). Alternatively, if we first regress idiosyncratic volatility on *MVE* and *SROE*, we find a monotonic negative relation between *residual* idiosyncratic volatility and *raw* price informativeness.

and FSRV. Finally, we provide some tests to show that our results are robust to different specifications.

4.2.1. Summary statistics

Table 1 provides the summary statistics for the main variables used in this paper. Panel A reports the summary statistics of $FSRV_{it}$. Specifically, we estimate summary statistics of FSRV each year and then average them over the sample period. The minimum FSRV in Panel A is thus the average minimum over the sample period. Consistent with earlier studies, our sample firms show significant FSRV, meaning that only a small portion of their stock returns, on average about 16%, are explained by the market or industry factors (as specified in Eq. (15)).

Panel B reports the correlation coefficients among firm size, two earnings volatility measures, and the six informativeness measures. As we have conjectured earlier, large firms tend to have smaller earnings volatility as is evidenced by the negative correlation between firm size and the volatility measures. Given that the six informativeness measures are available for samples with different size (see the average number of stocks for which the variables can be constructed), looking for a general pattern across them can further assure robustness of our results.

4.2.2. Relationship between FSRV and residual informativeness measures

We first plot FSRV against the residual informativeness measures and present six plots (one for each informativeness measure) in Fig. 5. Specifically, in each year, we first estimate Eq. (16) to obtain the residual informativeness measures. We then assign stocks into 10 groups by the value of their residual informativeness measures. Finally, we plot the mean and median FSRV for each of the 10 groups over the sample period.

All the six informativeness measures show a U-shaped relation to FSRV. In each plot, when the residual informativeness measure is small, FSRV decreases in the residual informativeness measure. This suggests that when the information environment of the firm is relatively poor, more information in the stock price leads to smaller FSRV, which is consistent with the argument of West (1988). The right half of the plot shows the opposite pattern as FSRV increases in the residual informativeness measure. This implies that when the information environment of the firm is relatively good, more information in stock price leads to greater FSRV, which is consistent with the findings of Morck et al. (2000).

To formally test our hypothesis and see if this change in the relation between stock price informativeness and FSRV is statistically significant, we estimate the following year-by-year cross-sectional regression:

$$FSRV_{i,t} = \alpha_t + \sum_{d=1}^{3}(\beta_{d,t} \times RG_{d,t}) + \sum_{d=1}^{3}(\gamma_{d,t} \times RG_{d,t} \times R\_INFO_{i,t}) + \varepsilon_{i,t}, \quad \text{s.t.}: \sum_{d=1}^{3}\beta_{d,t} = 0 \tag{17}$$

where $FSRV_{i,t}$ is firm-specific return variation for stock i in year t, $R\_INFO_{i,t}$ is the residual informativeness measure from regression (16), and $RG_{d,t}$ is a 0/1 dummy variable for one of three regions in the cross-section that is sorted by the residual informativeness

Table 1

Sample and summary statistics This table reports summary statistics of firm-specific return variation and other variables. The sample includes all non-financial and non-utility CRSP stocks whose CUSIP identifier is either 10 nor 11. Stocks whose SIC codes are missing or whose weekly return data are not available for the full year are excluded. Firm-specific return variation of a stock is estimated by deducting from one the R^2 of the year-by-year regression of the stock's weekly return on the weekly returns for the market and industry portfolios. Both portfolios are value-weighted and do not include the stock in question. Industry is determined by two-digit SIC codes. Firm size is the market capitalization as of the end of the previous year. Information measures include: probability of information based trading (PIN), price impact, analyst forecast error, analyst forecast dispersion, institutional ownership, and the number of publicly traded years. PIN is obtained from Soeren Hvidkjaer's web site. Price impact (PIM) is the absolute daily return over the daily dollar volume, averaged over the year and then multiplied by 10^6. Analyst forecast error (ERR) is measured as the absolute value of the difference between the mean estimate and the actual EPS (scaled by the stock price) for the next fiscal year. Analyst forecast dispersion (DSP) is measured as the standard deviation of analyst forecasts. At least two analysts are required for forecast error and dispersion. Institutional ownership is the fraction of shares held by 13f-filing institutional investors, averaged over the year. The number of publicly traded years is the number of days from the first public trading day to the last day of the year, divided by 365. All the information measures are transformed so that a higher value corresponds to more information: specifically, N_PIN = –log (PIN); N_ERR = –log (0.0001 + ERR); N_DSP = –log (0.0001 + DSP); N_PIM = –log (0.0001 + PIM); AGE = –1/the number of publicly traded days divided by 365; and IO = log (0.0001 + institutional ownership). MVE is the natural log of the market capitalization at the end of the previous year, SROE is the standard deviation of the return on equity (ROE) during the sample period. RROE is the mean square error from the regression of ROE on its 1-year lagged value, estimated over the sample period.

Panel A. Summary statistics of firm-specific return variation – in a given year

	n	Mean	Stdev	Min	q1	Median	q3	Max
	3073	0.837	0.145	0.199	0.764	0.879	0.949	1.000

Panel B. Other variables

Average number of sample stocks in a given year	SROE	RROE	MVE	N_PIN	N_PIM	N_ERR	N_DSP	AGE	IO
	3073	3073	3073	977	3073	1706	1648	3073	3073
Correlation coefficient with one another									
RROE	0.847								
	(0.000)								
MVE	−0.240	−0.124							
	(0.000)	(0.000)							
N_PIN	−0.073	−0.006	0.718						
	(0.000)	(0.407)	(0.000)						
N_PIM	−0.212	−0.110	0.912	0.715					
	(0.000)	(0.000)	(0.000)	(0.000)					
N_ERR	−0.205	−0.123	0.429	0.277	0.425				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
N_DSP	−0.147	−0.073	0.326	0.254	0.320	0.580			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
AGE	−0.189	−0.102	0.160	0.078	0.149	0.082	0.036		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)		
IO	−0.227	−0.153	0.613	0.439	0.606	0.281	0.259	0.193	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

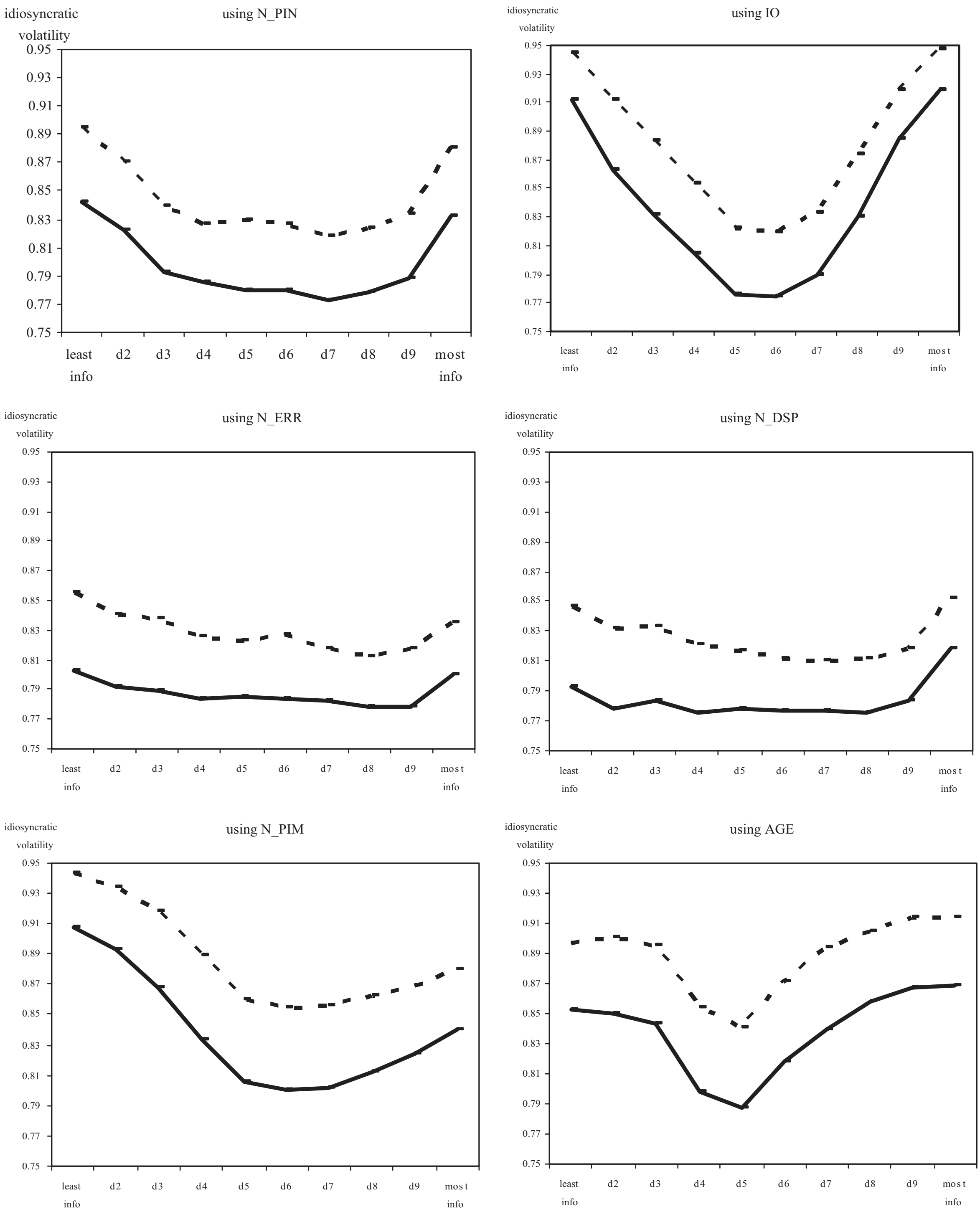

Fig. 5. Idiosyncratic volatility and price informativeness – after controlling for volatility of profitability and firm size. The above figures are a plot of idiosyncratic volatility (i.e., firm-specific return variation) against the residual price informativeness measure. Residual values are obtained from a year-by-year cross-sectional regression of the original price informativeness measure on cash flow volatility and firm size within the same size quintile. Firm size is the market capitalization as of the end of the previous year and volatility of profitability is the time-series standard deviation of the return on equity across all sample years.

measure ($RG_{1,t}$: first four deciles; $RG_{2,t}$: middle two deciles; $RG_{3,t}$: last four deciles). To avoid linear dependency, we impose the restriction that coefficients for the three dummies sum to zero.

Table 2 reports the time-series average of the regression coefficients and their Newey–West t-statistics, in the spirit of Fama and MacBeth (1973). The average coefficient on the interaction term between $RG_{1,t}$ and $R\_INFO_{i,t}$, $\gamma_{d,t}$, is negative and statistically significant for all six informativeness measures (the coefficient on *N_DSP* is only marginally significant with a t-statistic of 1.78). This means that for firms with poor information environments, more information in the stock price leads to smaller FSRV. The average coefficient on the interaction term between $RG_{2,t}$ and $R\_INFO_{i,t}$ is not statistically significant for four of the six informativeness measures. It is positively significant for the AGE measure, but negatively significant for the *N_PIM* measure. Overall, the results indicate that for firms with moderate information environments,

Table 2
Fama–MacBeth regression of firm-specific return variation on information measure. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) from the following year-by-year cross-sectional regressions: $FSRV_{i,t} = \alpha_t + \sum_{d=1}^{3}\beta_{d,t}RG_{d,t} + \sum_{d=1}^{3}\gamma_{d,t}(RG_{d,t} \times R\_INFO_{i,t}) + \varepsilon_{i,t}$, subject to $\sum_{d=1}^{3}\beta_{d,t} = 0$, where $FSRV_{i,t}$ is firm-specific return variation for stock *i* and year *t*, and $RG_{d,t}$ is a 0/1 dummy variable for one of three regions in the cross-section that is sorted by $R\_INFO_{i,t}$ (RG_1: first four deciles; RG_2: middle two deciles; RG_3: last four deciles). $R\_INFO_{i,t}$ is the residual price informativeness measure from following regression: $INFO_{i,t} = \alpha_t + \beta_t SROE_{i,t} + \gamma_t MVE_{i,t} + \varepsilon_{i,t}$, where $INFO_{i,t}$ is one of our transformed price informativeness measures (see Table 1 for details) for stock *i* and year *t*. We estimate this regression with all stocks within the same NYSE size quintile.

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.770	0.003	0.009	−0.012	−0.168	0.001	0.157	−0.325
(19 years)	(42.16)	(1.09)	(3.67)	−(4.96)	−(5.92)	(0.02)	(10.71)	−(8.12)
N_PIM	0.815	0.030	−0.012	−0.019	−0.027	−0.025	0.021	−0.048
(22 years)	(55.35)	(15.65)	−(5.57)	−(8.48)	−(7.54)	−(2.03)	(5.33)	−(6.57)
N_ERR	0.771	0.003	0.006	−0.009	−0.011	−0.032	0.014	−0.025
(22 years)	(44.37)	(1.10)	(2.79)	−(3.21)	−(3.00)	−(1.72)	(4.84)	−(6.52)
N_DSP	0.764	0.010	0.008	−0.018	−0.006	0.002	0.036	−0.042
(22 years)	(42.01)	(4.91)	(3.15)	−(5.65)	−(1.78)	(0.16)	(8.35)	−(9.27)
AGE	0.794	0.029	−0.055	0.026	−0.091	1.437	0.415	−0.506
(22 years)	(55.56)	(4.87)	−(4.67)	(2.19)	−(4.45)	(4.14)	(3.31)	−(4.07)
IO	0.797	0.025	−0.021	−0.004	−0.027	−0.014	0.062	−0.089
(22 years)	(52.37)	(8.69)	−(6.59)	−(1.10)	−(10.35)	−(0.87)	(14.37)	−(14.75)

Table 3
***F*-tests** from the year-by-year cross-sectional regressions. This table reports *p*-values of the *F*-tests for the hypothesis that $\gamma_{1,t}$ is equal to $\gamma_{3,t}$. in Eq. (17). As in Table 2, we report results based on residual information measures. These *F*-tests use the heteroscedasticity-consistent covariance matrix. Residual information measure is from Eq. (16) using all stocks within the same NYSE size quintile.

Year	Information measure					
	N_PIN	N_PIM	N_ERR	N_DSP	AGE	N_SPRD
1983	(0.001)	(0.001)	(0.791)	(0.147)	(0.000)	(0.000)
1984	(0.000)	(0.000)	(0.627)	(0.000)	(0.005)	(0.000)
1985	(0.000)	(0.000)	(0.404)	(0.172)	(0.000)	(0.000)
1986	(0.000)	(0.000)	(0.021)	(0.016)	(0.000)	(0.000)
1987	(0.000)	(0.000)	(0.000)	(0.006)	(0.000)	(0.000)
1988	(0.000)	(0.000)	(0.679)	(0.000)	(0.000)	(0.000)
1989	(0.000)	(0.000)	(0.006)	(0.000)	(0.000)	(0.000)
1990	(0.000)	(0.000)	(0.195)	(0.003)	(0.002)	(0.000)
1991	(0.000)	(0.000)	(0.010)	(0.451)	(0.536)	(0.000)
1992	(0.000)	(0.000)	(0.208)	(0.178)	(0.018)	(0.000)
1993	(0.000)	(0.001)	(0.463)	(0.787)	(0.749)	(0.000)
1994	(0.000)	(0.039)	(0.001)	(0.001)	(0.639)	(0.000)
1995	(0.000)	(0.158)	(0.000)	(0.000)	(0.003)	(0.000)
1996	(0.000)	(0.000)	(0.545)	(0.000)	(0.006)	(0.000)
1997	(0.186)	(0.005)	(0.002)	(0.000)	(0.000)	(0.000)
1998	(0.000)	(0.000)	(0.111)	(0.120)	(0.000)	(0.000)
1999	(0.017)	(0.000)	(0.019)	(0.064)	(0.000)	(0.000)
2000	(0.157)	(0.199)	(0.189)	(0.003)	(0.081)	(0.000)
2001	(0.000)	(0.000)	(0.000)	(0.003)	(0.077)	(0.000)
2002		(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
2003		(0.000)	(0.005)	(0.000)	(0.000)	(0.000)
2004		(0.000)	(0.510)	(0.001)	(0.000)	(0.000)

there is no obvious relation between the amount of information in the stock price and FSRV. Finally, the average coefficient on the interaction term between $RG_{3,t}$ and $R\_INFO_{i,t}$, $\gamma_{3,t}$, is positive and statistically significant for all six informativeness measures. This suggests that for firms with good information environments, more information in the stock price leads to greater FSRV.

The last column, DIFF, reports the average difference between $\gamma_{1,t}$ and $\gamma_{3,t}$, along with its Newey–West *t*-statistic. The difference is statistically significant at the 1% level for all six informativeness measures, which we interpret as evidence that information in the stock price affects FSRV differently depending on the information environments of the stock.

As another look at the difference between $\gamma_{1,t}$ and $\gamma_{3,t}$, Table 3 reports the *p*-value of the year-by-year *F*-test for the null hypothesis that $\gamma_{1,t}$ is equal to $\gamma_{3,t}$. We reject this null hypothesis in most of the sample years. For example, if we use *N_PIN* as an informativeness measure, out of the 19 years for which PIN is available, we reject the hypothesis at the 1% level in 16 years and at the 5% level in 17 years. This shows that the U-shaped relationship is present fairly consistently over time. The informativeness measure that offers the weakest rejection of the null hypothesis is the one from the analyst forecast errors (*N_ERR*). With this measure, we reject the null hypothesis at the 5% level in 12 out of the 22 sample years. This is probably due to that fact that the sample for this informativeness measure (*N_ERR*) is limited to firms in the IBES database, which have a relatively good information environment. Similar (but better relative to results based on *N_ERR*) results with

Table 4
Fama–MacBeth regression results without truncating extreme values. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the same specification as Table 2. Results in this table are based on a dataset where extreme values are not truncated.

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.770	0.004	0.008	−0.012	−0.165	−0.018	0.157	−0.322
(19 years)	(42.34)	(1.30)	(2.89)	−(4.94)	−(5.82)	−(0.22)	(10.51)	−(8.02)
N_PIM	0.816	0.030	−0.012	−0.018	−0.027	−0.023	0.020	−0.047
(22 years)	(55.75)	(15.11)	−(6.21)	−(8.64)	−(7.66)	−(1.67)	(5.38)	−(6.68)
N_ERR	0.771	0.000	0.008	−0.008	−0.013	−0.015	0.013	−0.026
(22 years)	(44.33)	(0.04)	(4.27)	−(2.72)	−(4.12)	−(0.76)	(4.05)	−(6.69)
N_DSP	0.769	0.007	0.004	−0.010	−0.007	0.009	0.022	−0.028
(22 years)	(42.40)	(2.01)	(1.57)	−(2.90)	−(2.06)	(0.87)	(14.50)	−(7.22)
AGE	0.797	0.027	−0.054	0.027	−0.089	1.361	0.371	−0.461
(22 years)	(55.54)	(4.35)	−(4.90)	(2.33)	−(4.42)	(4.43)	(3.16)	−(3.95)
IO	0.799	0.024	−0.021	−0.003	−0.027	−0.022	0.061	−0.088
(22 years)	(53.69)	(9.55)	−(7.28)	−(0.92)	−(10.38)	−(1.51)	(14.86)	−(14.72)

Table 5
Fama–MacBeth regression results with alternative measure of volatility of profitability. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the same specification as Table 2. Results in this table are based on an alternative measure of volatility of corporate profitability (i.e., squared root of errors from a regression of the annual return on equity on its 1-year lagged value, estimated over the sample period).

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.769	0.004	0.007	−0.011	−0.168	−0.029	0.155	−0.323
(19 years)	(42.24)	(1.27)	(2.53)	−(4.15)	−(5.88)	−(0.32)	(10.13)	−(7.92)
N_PIM	0.815	0.031	−0.013	−0.018	−0.027	−0.012	0.020	−0.047
(22 years)	(55.59)	(16.00)	−(6.92)	−(8.48)	−(7.79)	−(1.06)	(5.39)	−(6.81)
N_ERR	0.772	0.004	0.007	−0.011	−0.010	0.012	0.015	−0.025
(22 years)	(44.75)	(1.26)	(3.64)	−(4.11)	−(2.46)	(0.48)	(5.01)	−(5.74)
N_DSP	0.763	0.010	0.008	−0.019	−0.006	0.000	0.037	−0.043
(22 years)	(42.00)	(3.90)	(3.57)	−(5.68)	−(1.85)	(0.04)	(8.89)	−(9.45)
AGE	0.783	0.042	−0.059	0.017	−0.081	1.735	0.575	−0.656
(22 years)	(49.10)	(6.83)	−(5.57)	(1.56)	−(3.98)	(6.00)	(5.05)	−(5.57)
IO	0.798	0.023	−0.019	−0.004	−0.028	−0.027	0.061	−0.089
(22 years)	(53.78)	(8.96)	−(6.08)	−(1.00)	−(10.54)	−(1.81)	(15.21)	−(15.28)

Table 6
Fama–MacBeth regression results with residual information measures obtained from one single cross-sectional regression each year. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the same specification as Table 2. Results in this table are based on the residual information measures that are obtained from a single cross-sectional regression (instead of five regressions in each size quintile) each year.

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.772	0.007	0.009	−0.016	−0.145	−0.032	0.138	−0.283
(19 years)	(42.98)	(1.89)	(3.58)	−(5.05)	−(5.40)	−(0.42)	(11.95)	−(8.19)
N_PIM	0.824	0.028	−0.001	−0.028	−0.018	−0.052	0.013	−0.031
(22 years)	(57.36)	(8.37)	−(0.39)	−(6.95)	−(5.11)	−(4.73)	(2.86)	−(4.05)
N_ERR	0.772	0.006	0.005	−0.011	−0.009	−0.015	0.015	−0.024
(22 years)	(44.13)	(1.76)	(2.60)	−(3.81)	−(2.38)	−(1.13)	(4.55)	−(6.19)
N_DSP	0.765	0.007	0.008	−0.015	−0.007	0.002	0.032	−0.039
(22 years)	(41.53)	(3.81)	(2.98)	−(4.99)	−(2.00)	(0.14)	(7.90)	−(8.13)
AGE	0.795	0.043	0.025	−0.068	−0.044	−0.332	1.202	−1.245
(22 years)	(46.39)	(6.83)	(2.00)	−(4.38)	−(2.31)	−(1.13)	(5.59)	−(5.80)
IO	0.803	−0.018	0.001	0.017	−0.028	0.053	0.038	−0.066
(22 years)	(54.31)	−(7.13)	(0.27)	(4.09)	−(11.29)	(3.46)	(11.20)	−(12.73)

Table 7
Fama–MacBeth regression results with scaled information measures. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the same specification as Table 2. Results in this table are based on the scaled residual information measure that ranges from −0.5 to 0.5 (see, e.g., Mendenhall, 2004). Specifically, we first determine the percentile ranking from 0 to 99 in a given sample year, and deduct 49.5 from the ranking value and then divide by 99.

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.753	−0.004	0.025	−0.021	−0.199	0.008	0.197	−0.396
(19 years)	(38.27)	−(1.32)	(6.61)	−(7.56)	−(7.20)	(0.20)	(8.13)	−(8.35)
N_PIM	0.793	0.005	0.008	−0.014	−0.250	−0.062	0.128	−0.378
(22 years)	(45.83)	(1.56)	(2.21)	−(4.62)	−(9.01)	−(2.18)	(4.96)	−(7.39)
N_ERR	0.771	0.000	0.008	−0.007	−0.056	−0.064	0.049	−0.105
(22 years)	(43.81)	−(0.01)	(2.96)	−(2.00)	−(2.60)	−(2.16)	(2.86)	−(4.63)
N_DSP	0.760	0.003	0.012	−0.016	−0.055	0.003	0.128	−0.182
(22 years)	(40.46)	(0.95)	(5.14)	−(5.16)	−(2.11)	(0.11)	(6.57)	−(6.39)
AGE	0.802	−0.025	−0.002	0.027	−0.190	0.345	0.097	−0.286
(22 years)	(55.36)	−(2.49)	−(0.33)	(3.00)	−(4.91)	(4.36)	(4.72)	−(6.81)
IO	0.747	−0.002	0.027	−0.024	−0.354	−0.021	0.436	−0.790
(22 years)	(38.97)	−(0.60)	(8.17)	−(7.70)	−(10.28)	−(0.80)	(13.46)	−(12.14)

N_DSP further confirm this conjecture, since *N_DSP* is also calculated using the IBES database.

4.2.3. Robustness check

In this section, we conduct a variety of robustness checks to assure that our earlier results, the U-shaped relation between price informativeness and FSRV, are robust.

First, we report results based on an untruncated dataset (recall that thus far we have treated the observation as missing if *SROE*, *N_ERR*, or *N_DSP* is either below the 1st percentile or above the 99th percentile in a certain year). Table 4 reports results when we keep all those extreme values. The results are virtually identical: for all six informativeness measures, the average value of $\gamma_{1,t}$ is negative and significant, the average value of $\gamma_{3,t}$ is positive and significant, and the difference between $\gamma_{1,t}$ and $\gamma_{3,t}$ is statistically significant.

In Tables 2 and 3, we used *SROE*, the standard deviation of the firm's return on equity over the sample period, as a measure of

Table 8
Fama–MacBeth regression results with more variables controlled for. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the specification for Table 2 with slight modifications. Results in this table are based on an alternative specification for Eq. (17) where sector dummies, the number of segments within the firm, leverage ratio, and dividend payment dummy are included as control variables.

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.890	0.004	0.008	−0.012	−0.149	−0.034	0.126	−0.276
(19 years)	(37.76)	(1.63)	(3.96)	−(6.04)	−(6.13)	−(0.46)	(12.46)	−(8.57)
N_PIM	0.910	0.025	−0.007	−0.018	−0.025	−0.029	0.009	−0.034
(22 years)	(47.24)	(18.34)	−(4.67)	−(9.54)	−(8.05)	−(2.76)	(2.34)	−(5.25)
N_ERR	0.886	0.004	0.005	−0.009	−0.012	−0.031	0.011	−0.022
(22 years)	(35.64)	(1.33)	(3.07)	−(3.34)	−(3.64)	−(1.54)	(3.07)	−(4.58)
N_DSP	0.870	0.011	0.006	−0.017	−0.003	−0.006	0.025	−0.028
(22 years)	(34.28)	(6.20)	(2.17)	−(6.19)	−(1.22)	−(0.47)	(6.87)	−(6.79)
AGE	0.893	0.015	−0.041	0.026	−0.076	1.220	0.388	−0.464
(22 years)	(46.28)	(2.92)	−(4.70)	(2.36)	−(5.18)	(4.72)	(2.77)	−(3.28)
IO	0.885	0.022	−0.014	−0.008	−0.022	−0.021	0.052	−0.075
(22 years)	(44.67)	(8.85)	−(4.74)	−(2.40)	−(9.30)	−(1.31)	(13.12)	−(12.85)

Table 9
Fama–MacBeth regression results with FSRV estimated using the market model with four lagged market and industry returns. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the same specification as Table 2. Results in this table are based on firm-specific return variation that is obtained from a modified Eq. (15) where four lagged market returns and four lagged industry returns are included, as well as the contemporaneous market and industry returns.

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.636	0.001	0.008	−0.009	−0.134	0.032	0.121	−0.255
(19 years)	(39.12)	(0.49)	(3.50)	−(3.98)	−(5.86)	(0.41)	(8.55)	−(7.40)
N_PIM	0.672	0.025	−0.010	−0.015	−0.022	−0.023	0.018	−0.040
(22 years)	(50.28)	(15.29)	−(5.36)	−(9.04)	−(7.30)	−(2.02)	(5.19)	−(6.54)
N_ERR	0.636	0.002	0.006	−0.008	−0.009	−0.029	0.013	−0.022
(22 years)	(41.37)	(0.81)	(3.62)	−(3.13)	−(3.12)	−(1.67)	(4.62)	−(6.65)
N_DSP	0.629	0.007	0.009	−0.016	−0.006	−0.002	0.030	−0.036
(22 years)	(39.32)	(4.40)	(3.44)	−(5.74)	−(1.95)	−(0.18)	(7.03)	−(7.81)
AGE	0.653	0.025	−0.046	0.021	−0.071	1.207	0.360	−0.431
(22 years)	(51.67)	(5.36)	−(4.52)	(2.02)	−(4.15)	(4.12)	(3.36)	−(4.10)
IO	0.657	0.021	−0.017	−0.004	−0.023	−0.007	0.053	−0.075
(22 years)	(47.41)	(8.88)	−(5.74)	−(1.18)	−(10.28)	−(0.53)	(14.29)	−(14.85)

Table 10
Fama–MacBeth regression results with alternative residual information measures. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the same specification as Table 2. Results in this table are based on alternative residual information measures that are obtained by regressing each of them on earnings volatility alone (within the same size quintile).

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.772	0.000	0.007	−0.007	−0.173	−0.098	0.098	−0.270
(19 years)	(42.86)	(0.06)	(3.48)	−(1.73)	−(7.53)	−(1.52)	(7.80)	−(9.85)
N_PIM	0.807	0.025	−0.004	−0.022	−0.032	−0.028	0.016	−0.048
(22 years)	(52.32)	(12.89)	−(1.46)	−(9.28)	−(8.61)	−(4.42)	(7.72)	−(9.47)
N_ERR	0.771	0.002	0.007	−0.009	−0.013	0.007	0.012	−0.025
(22 years)	(44.16)	(0.81)	(4.30)	−(3.25)	−(3.91)	(0.56)	(4.01)	−(6.51)
N_DSP	0.764	0.009	0.008	−0.017	−0.008	0.007	0.031	−0.039
(22 years)	(41.67)	(4.05)	(2.59)	−(5.47)	−(2.43)	(0.43)	(6.40)	−(7.45)
AGE	0.792	0.031	−0.053	0.022	−0.084	1.329	0.482	−0.566
(22 years)	(57.11)	(4.16)	−(3.53)	(1.57)	−(3.77)	(3.12)	(3.37)	−(4.04)
IO	0.801	0.019	−0.039	0.021	−0.027	0.011	0.032	−0.059
(22 years)	(53.17)	(7.09)	−(5.91)	(3.85)	−(10.19)	(0.45)	(7.48)	−(9.76)

the volatility of corporate profitability. In Table 5, we obtain our results using *RROE*, namely, the mean squared error from a regression of the firm's return on equity on its 1-year lag over the sample period, as a measure of the volatility of profitability. The results are also qualitatively the same as those in Table 2.

So far, we have calculated the residual informativeness measures by estimating Eq. (16) within the same NYSE size quintile in each year. In Table 6, we report results when we calculate the residual informativeness measures by estimating Eq. (16) with all sample firms in each year. Again, the results are qualitatively the same as those in Table 2.

In Table 7, we report results when we scale the residual informativeness variables into a variable ranging from −0.5 to 0.5 (see, e.g., Mendenhall, 2004). It is basically a percentile ranking

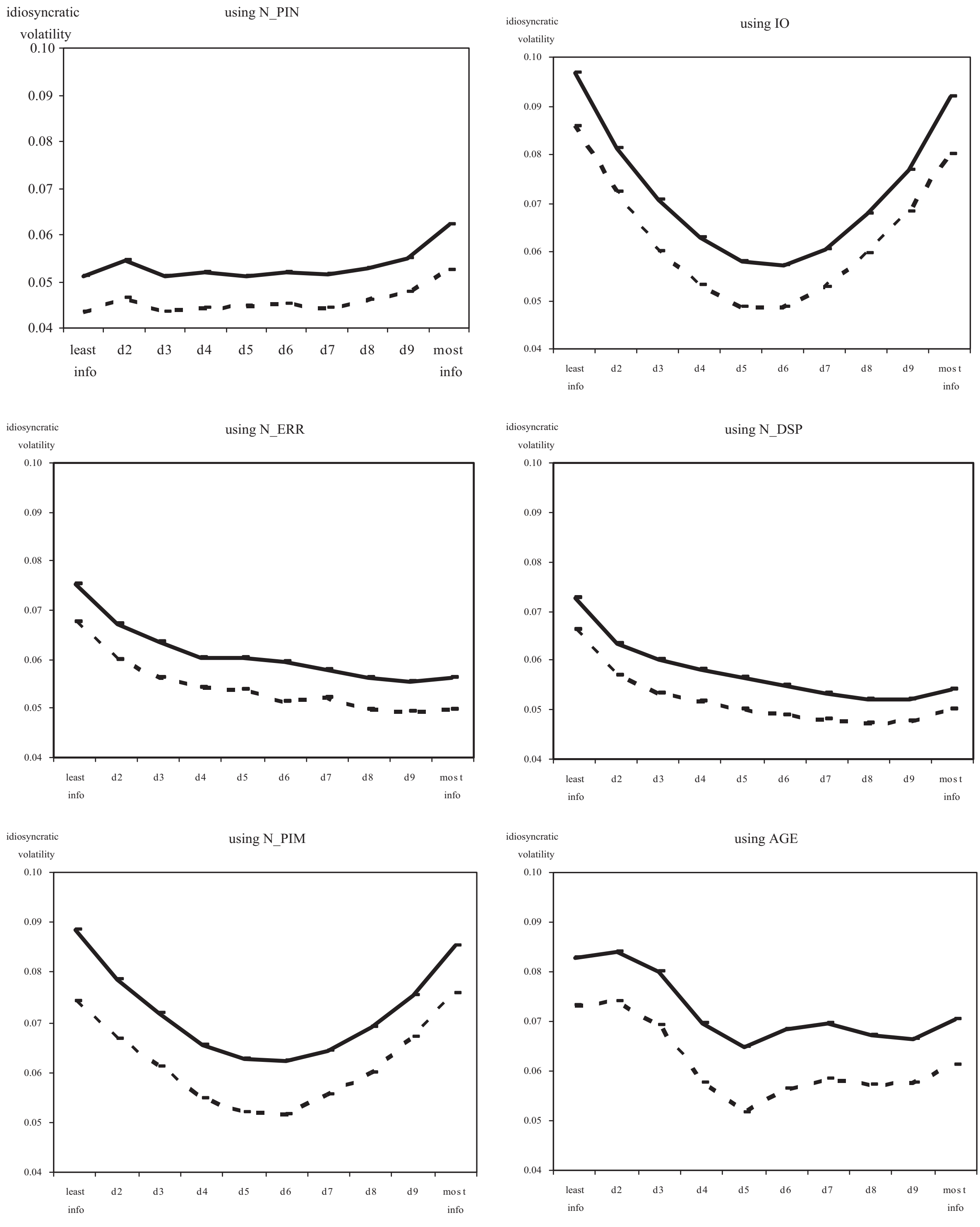

Fig. 6. Absolute measure of idiosyncratic volatility and price informativeness – after controlling for volatility of profitability and firm size. The above figures are a plot the absolute idiosyncratic volatility (i.e., absolute FSRV) against the residual price informativeness measure. Absolute idiosyncratic volatility is the mean root squared errors of the market model (Eq. (15)). Residual value is obtained from a year-by-year cross-sectional regression on earnings volatility and firm size within the same size quintile. Firm size is the market capitalization as of the end of the previous year and volatility of profitability is the time-series standard deviation of the return on equity over the sample years.

from 0 to 99 in a given sample year; we deduct 49.5 from the ranking value and then divide by 99, so that it ranges from −0.5 to 0.5. This addresses two potential problem of the original approach. First, it mitigates the outlier problem. Second, since the residual variables now ranges from −0.5 and 0.5, the coefficient can be interpreted as the change in FSRV when the residual measure changes from the lowest percentile to the highest percentile. Results are consistent with those in Table 2.

We further control for other factors which may affect FSRV, such as industry sectors, the number of segments in the firm, leverage, and dividends. Specifically, we include in Eq. (17) 14 industry sector dummies, the natural log of the number of segments, a

Table 11
Fama–MacBeth regression results with alternative residual information measures. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the same specification as Table 2. Results in this table are based on alternative residual information measures that are obtained by regressing each of the six price informativeness measures on firm size within the earnings volatility quintile.

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIM	0.772	0.006	0.011	−0.017	−0.145	0.007	0.145	−0.290
(19 years)	(42.68)	(1.61)	(4.97)	−(4.60)	−(5.45)	(0.10)	(11.05)	−(8.06)
N_PIM	0.824	0.027	0.001	−0.028	−0.019	−0.051	0.015	−0.033
(22 years)	(57.08)	(8.56)	(0.87)	−(8.20)	−(5.70)	−(5.44)	(3.33)	−(4.62)
N_ERR	0.771	0.003	0.008	−0.010	−0.012	0.002	0.015	−0.028
(22 years)	(44.03)	(0.84)	(4.76)	−(4.36)	−(3.30)	(0.18)	(4.97)	−(6.93)
N_DSP	0.764	0.010	0.009	−0.019	−0.007	0.012	0.035	−0.042
(22 years)	(41.23)	(3.78)	(3.42)	−(6.06)	−(1.90)	(0.85)	(8.00)	−(7.91)
AGE	0.800	0.032	−0.014	−0.018	−0.062	0.422	0.713	−0.775
(22 years)	(48.89)	(5.00)	−(1.19)	−(1.21)	−(3.37)	(1.63)	(3.40)	−(3.65)
IO	0.801	−0.012	0.001	0.011	−0.029	0.046	0.044	−0.073
(22 years)	(54.05)	−(5.87)	(0.29)	(3.06)	−(11.19)	(3.65)	(12.17)	−(12.90)

Table 12
Fama–MacBeth regression results with FSRV estimated using the Fama–French 3-factor model. This table reports Fama–MacBeth coefficients and Newey–West *t*-statistics (in parentheses) for Eq. (17) – the same specification as Table 2. Results in this table are based on firm-specific return variation that is obtained from the Fama–French 3-factor model.

Information measure	Intercept	Coefficient for						DIFF
		RG_1	RG_2	RG_3	$RG_1 \times R\_INFO$	$RG_2 \times R\_INFO$	$RG_3 \times R\_INFO$	
N_PIN	0.763	0.002	0.008	−0.011	−0.143	0.026	0.131	−0.273
(19 years)	(39.35)	(1.05)	(3.84)	−(4.65)	−(7.01)	(0.38)	(9.50)	−(8.90)
N_PIM	0.802	0.022	−0.009	−0.012	−0.028	−0.023	0.012	−0.040
(22 years)	(53.32)	(10.38)	−(4.76)	−(6.19)	−(8.28)	−(3.28)	(2.63)	−(5.31)
N_ERR	0.759	0.003	0.005	−0.009	−0.015	−0.026	0.009	−0.024
(22 years)	(44.31)	(1.36)	(4.08)	−(3.15)	−(6.07)	−(1.48)	(3.52)	−(9.77)
N_DSP	0.753	0.014	0.006	−0.019	−0.013	−0.016	0.027	−0.040
(22 years)	(43.13)	(7.78)	(2.44)	−(7.04)	−(3.51)	−(1.45)	(6.48)	−(11.80)
AGE	0.789	0.020	−0.043	0.023	−0.064	1.095	0.244	−0.308
(22 years)	(53.72)	(4.24)	−(4.49)	(2.80)	−(3.57)	(3.76)	(3.47)	−(3.99)
IO	0.784	0.024	−0.014	−0.009	−0.026	−0.018	0.056	−0.082
(22 years)	(51.14)	(7.63)	−(6.10)	−(2.60)	−(10.66)	−(1.32)	(13.29)	−(13.92)

dummy variable for firms who pay dividends during the year, and the leverage ratio.[16] Results are reported in Table 8. The results are qualitatively the same as those in Table 2.

Table 9 reports results based on an alternative specification for Eq. (15), the regression for FSRV estimation. Specifically, to estimate FSRV, we use four lagged market portfolio returns and four lagged industry portfolio returns, as well as their contemporaneous returns. Results are consistent with those in Table 2.

Our definition of the residual informativeness measure has been the residual from the regression of the raw informativeness measures on the volatility of corporate profitability and firm size. This robustness check, reported in Table 10, shows that even if we control for the volatility of profitability alone in calculating residual informativeness measures, the results still hold.

So far, we have been using the relative FSRV measure. Using relative FSRV is more appropriate since some businesses can be more susceptible to systematic shocks than others, and firm-specific events in these firms can be accordingly more intense. In other words, using relative FSRV helps control for such environmental volatility. As a robustness check, we see if our results hold if we use absolute FSRV instead. Fig. 6 shows that results based on the absolute measure are broadly consistent with those based on the relative measure.

In Table 2, we calculate the residual informativeness measures by estimating Eq. (16) within the same NYSE size quintile each year. In Table 11, we report results when we calculate the residual informativeness measures by regressing $INFO_{i,t}$ on $MVE_{i,t}$ and a constant in each *SORE* quintile each year. The results are qualitatively the same as those in Table 2.[17]

Most studies examining the relation between idiosyncratic volatility and expected stock returns define idiosyncratic volatility slightly differently. They use the Fama and French (1993) three-factor model instead of CAPM. In Table 12, we report results when we calculate FSRV by using Fama and French (1993) three factors instead of the market and industry returns in Eq. (15). We obtain results similar to those in Table 2.

5. Conclusion

Many recent studies use idiosyncratic return volatility as a measure of how much information is incorporated into the stock price. Alarmingly, while some studies use higher idiosyncratic volatility as a measure of more informative prices, others assume that higher idiosyncratic volatility means less price informativeness. Empirical evidence regarding the information content of idiosyncratic volatility is also mixed. A growing body of research shows that firms with more informative stock prices have higher idiosyncratic volatility (Morck et al., 2000). Another strand of studies find exactly the

[16] The 14 industry sector dummies are based on 2-digit SIC codes. Specifically, sector 1 between 1 and 9; sector 2 between 10 and 14; sector 3 between 15 and 19; sector 4 between 20 and 21; sector 5 between 22 and 23; sector 6 between 24 and 27; sector 7 between 28 and 32; sector 8 between 33 and 34; sector 9 between 35 and 39; sector 10 between 40 and 48; sector 11 between 50 and 52; sector 12 between 53 and 59; sector 13 between 70 and 79; and sector 14 between 80 and higher.

[17] If we instead calculate the residual informativeness measures by regressing $INFO_{i,t}$ on $MVE_{i,t}$, $SORE_{i,t}$, and a constant in each *SORE* quintile each year, the results are close to identical to those in Table 11. These results are available upon request to interested readers.

opposite (e.g., Kelly, 2005). Understanding the true relation between price informativeness and idiosyncratic volatility is important, given that an increasing number of studies use idiosyncratic volatility as a measure of price informativeness or information asymmetry. Further, understanding the true information content of idiosyncratic volatility is important for practitioners and policy makers as well.

This paper makes an attempt in this direction. We investigate the relation between price informativeness and idiosyncratic price volatility in a multi-asset, multi-period noisy rational expectations equilibrium. Idiosyncratic return volatility is decomposed into two parts: (1) the part caused by noise, and (2) the part caused by information regarding the firm's fundamental value. We show that the first component decreases with price informativeness, while the second component first decreases and then increases with price informativeness. Our main results are as follows. First, there exist no parameter values such that idiosyncratic return volatility increases monotonically with price informativeness. Second, there exist parameter values such that the relation between price informativeness and idiosyncratic return volatility is U-shaped. Finally, there exist parameter values such that idiosyncratic return volatility decreases monotonically with price informativeness. Using several price informativeness measures, we empirically document a U-shaped relation between price informativeness and idiosyncratic return volatility. Our study therefore reconciles the opposing views expressed in the following two strands of literature: (1) the growing body of research showing that firms with more informative stock prices have greater idiosyncratic return volatility (e.g., Morck et al., 2000; Jin and Myers, 2006), and (2) the studies arguing that more information in price reduces idiosyncratic return volatility (West, 1988; Kelly, 2005).

Acknowledgements

We thank Aziz Alimov, Brent Ambrose, Paul Childs, Art Durnev, Brad Jordan, Ike Mathur (the editor), Randall Morck, Don Mullineaux, Joe Peek, Wei Xiong, an anonymous referee, and seminar participants at the 2006 Financial Management Association Meetings, the First International Conference on Asia–Pacific Financial Markets, the 2007 International Finance Conference at Queen's University, Korea University, Seoul National University, and University of Kentucky for helpful comments. All errors and omissions are our own.

Appendix A. Proofs of propositions and corollaries

Proof of proposition 1. The information production and portfolio choice problem of investor j is

$$\max_{X_T} \quad E_{T-1}\left[-e^{-aW_T^j}|\Omega_{T-1}^j\right] \tag{A.1}$$

$$\text{s.t.}: W_T^j = B_{T-1}^j + P'_{T-1}\left(X_{T-1}^j - X_T^j\right) - (I_{T-1}^j)'C + u'X_T^j.$$

Since W_T^j is normally distributed and the utility function is negative exponential, it is well-known that the solution to the investor's portfolio choice problem is

$$X_T^j = \frac{1}{a}[Var(u|\Omega_{T-1}^j)]^{-1}\left[E_{T-1}(u|\Omega_{T-1}^j) - P_{T-1}\right]. \tag{A.2}$$

Given that the fundamental values of and the noisy demands for the $N+1$ stocks are independent of each other, we have

$$E_{T-1}(u_n|\Omega_{T-1}^j) = \begin{cases} \bar{u}_n + \sum\limits_{t=1}^{T-1}\eta_{n,t} + \theta_{n,T} & \text{if } I_{n,T-1}^j = 1 \\ \bar{u}_n + \sum\limits_{t=1}^{T-1}\eta_{n,t} + E_{T-1}(\eta_{n,T}|P_{n,T-1}) & \text{if } I_{n,T-1}^j = 0, \end{cases} \tag{A.3}$$

and $Var(u|\Omega_{T-1}^j)$ is an $(N+1)\times(N+1)$ diagonal variance–covariance matrix with the $(n,n)^{th}$ element as

$$Var(u_n|\Omega_{T-1}^j) = \begin{cases} \sigma_{\epsilon,n}^2 & \text{if } I_{n,T-1}^j = 1 \\ Var_{T-1}(\eta_{n,T}|P_{n,T-1}) & \text{if } I_{n,T-1}^j = 0, \end{cases} \tag{A.4}$$

where $P_{n,T-1}$ is independent of each other across the stocks, as we will show later. Plugging Eqs. (A.3) and (A.4) in Eq. (A.2) yields

$$X_{n,T}^j = \begin{cases} \frac{\bar{u}_n + \sum_{t=1}^{T-1}\eta_{n,t} + \theta_{n,T} - P_{T-1}}{a\sigma_{\epsilon,n}^2} & \text{if } I_{n,T-1}^j = 1 \\ \frac{\bar{u}_n + \sum_{t=1}^{T-1}\eta_{n,t} + E_{T-1}(\eta_{n,T}|P_{n,T-1}) - P_{T-1}}{aVar_{T-1}(\eta_{n,T}|P_{n,T-1})} & \text{if } I_{n,T-1}^j = 0. \end{cases} \tag{A.5}$$

The value of $X_{n,T}^j$ when $I_{n,T-1}^j = 1$ is the demand for stock $n$when investor j produces information about stock n, while the value of $X_{n,T}^j$ when $I_{n,T-1}^j = 0$ is the demand for stock n when the investor is uninformed about $\theta_{n,T}$. From Eq. (A.5), we know that for the fraction $\mu_{n,T-1}$ of investors who produce information and observe the value $\theta_{n,T}$, their demand for stock n is $\frac{\bar{u}_n + \sum_{t=1}^{T-1}\eta_{n,t} + \theta_{n,T} - P_{T-1}}{a\sigma_{\epsilon,n}^2}$, and for the fraction $1-\mu_{n,T-1}$ of investors who do not observe the value $\theta_{n,T}$, their demand is $\frac{E_{T-1}(u_n|P_{n,T-1}) - P_{n,T-1}}{aVar_{T-1}(u_n|P_{n,T-1})}$. The market clearing condition for stock n is therefore

$$\mu_{n,T-1}\frac{\bar{u}_n + \sum_{t=1}^{T-1}\eta_{n,t} + \theta_{n,T} - P_{T-1}}{a\sigma_{\epsilon,n}^2} + (1-\mu_{n,T-1})\frac{E_{T-1}(u_n|P_{n,T-1}) - P_{n,T-1}}{aVar_{T-1}(u_n|P_{n,T-1})} + z_{n,T} = y_n. \tag{A.6}$$

That is, the total demand for stock n from informed investors, UDTs, and liquidity traders equals the total physical supply of the stock. If we rearrange Eq. (A.6), we have

$$-\bar{u}_n - \sum_{t=1}^{T-1}\eta_{n,t} + P_{n,T-1} - \frac{1-\mu_{n,T-1}}{\mu_{n,T-1}}\sigma_{\epsilon,n}^2\frac{E_{T-1}(u_n|P_{n,T-1}) - P_{n,T-1}}{Var_{T-1}(u_n|P_{n,T-1})} + \frac{a\sigma_{\epsilon,n}^2}{\mu_{n,T-1}}y_n = \theta_{n,T} + \frac{a\sigma_{\epsilon,n}^2}{\mu_{n,T-1}}z_{n,T}. \tag{A.7}$$

UDTs observe everything at the left side of the above equation, which is a noisy signal of the private information held by informed traders, $\theta_{n,T}$. We define the left side of Eq. (A.7) as $S(P_{n,T-1})$, which is the set of information that is revealed to the UDTs by the price, $P_{n,T-1}$. If the beliefs of UDTs are consistent, we have

$$E_{T-1}(u_n|P_{n,T-1}) = E_{T-1}(u_n|S(P_{n,T-1})) \tag{A.8}$$

and

$$Var_{T-1}(u_n|P_{n,T-1}) = Var_{T-1}(u_n|S(P_{n,T-1})). \tag{A.9}$$

Using the properties of conditional normal distributions, we can show that

$$E_{T-1}(u_n|P_{n,T-1}) = \bar{u}_n + \sum_{t=1}^{T-1}\eta_{n,t} + \frac{\mu_{n,T-1}\sigma_{\theta,n}^2(\mu_{n,T-1}\theta_{n,T} + a\sigma_{\epsilon,n}^2 z_{n,T})}{\mu_{n,T-1}^2\sigma_{\theta,n}^2 + a^2\sigma_{\epsilon,n}^4\sigma_{z,n}^2} \tag{A.10}$$

and

$$Var_{T-1}(u_n|P_{n,T-1}) = \sigma_{\epsilon,n}^2 + \frac{\sigma_{\theta,n}^2 a^2\sigma_{\epsilon,n}^4\sigma_{z,n}^2}{\mu_{n,T-1}^2\sigma_{\theta,n}^2 + a^2\sigma_{\epsilon,n}^4\sigma_{z,n}^2}. \tag{A.11}$$

Plugging Eqs. (A.10) and (A.11) into Eq. (A.7), we obtain the equilibrium market price of stock n in Eq. (9), where

$$\alpha_{n,T-1}^{\theta} = 1 - \frac{(1-\mu_{n,T-1})a^2\sigma_{z,n}^2\sigma_{\epsilon,n}^4}{\mu_{n,T-1}^2\sigma_{\theta,n}^2 + a^2\sigma_{z,n}^2\sigma_{\epsilon,n}^4 + a^2\mu_{n,T-1}\sigma_{z,n}^2\sigma_{\theta,n}^2\sigma_{\epsilon,n}^2} \tag{A.12}$$

and

$$\alpha^z_{n,T-1} = \frac{a\sigma^2_{\epsilon,n}[\mu_{n,T-1}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\sigma^2_{\epsilon,n}(\sigma^2_{\theta,n}+\sigma^2_{\epsilon,n})]}{\mu^2_{n,T-1}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\sigma^4_{\epsilon,n} + a^2\mu_{n,T-1}\sigma^2_{z,n}\sigma^2_{\theta,n}\sigma^2_{\epsilon,n}}. \quad \square \tag{A.13}$$

Proof of proposition 2. We start with time $t = T-2$. Define $J_{T-1}(B_{T-1}, X_{n,T-1})$ as the expected payoff to an agent who has a portfolio of $(B_{T-1}, X_{n,T-1})$ and observed price P_{T-1} and a set of private signals S_{T-1}.[18] Since we have proved that UDTs will be indifferent between producing information and not-producing information on each stock, and the information production decisions are independent of each other, we can assume the agent chooses to produce information on all the $N+1$ stocks without loss of generality.

$$J_{T-1}(B_{T-1}, X_{n,T-1}) = \max_{X_T} \quad E_{T-1}[-e^{-aW_T}|\Omega_{T-1}]$$

$$= E_{T-1}[-e^{-a(B_{T-1}+\sum_{n=1}^{N+1}P_{n,T-1}(X_{n,T-1}-X_{n,T}) - \sum_{n=1}^{N+1}C_n + \sum_{n=1}^{N+1}u_{n,T}X_{n,T})}|\Omega_{T-1}] \tag{A.14}$$

$$= -e^{-a(B_{T-1}+\sum_{n=1}^{N+1}\{P_{n,T-1}(X_{n,T-1}-X_{n,T})-C_n+(\bar{u}_n+\sum_{t=1}^{T-1}\eta_{n,t}+\theta_{n,T})X_{n,T}-0.5aX^2_{n,T}\sigma^2_{\epsilon})} \tag{A.15}$$

The last equality follows Eqs. (A.3) and (A.4) when $I_{n,T-1} = 1$. Therefore, the objective of the investor at time $T-2$ is

$$\max_{X_{T-1}} \quad E_{T-2}[J_{T-1}(B_{T-1}, X_{n,T-1})|\Omega_{T-2}] \tag{A.16}$$

$$s.t.: B_{T-1} = B_{T-2} + \sum_{n=1}^{N+1} P_{n,T-2}(X_{n,T-2} - X_{n,T-1}) - \sum_{n=1}^{N+1} I_{n,T-2}C_n. \tag{A.17}$$

Or, equivalently,

$$\max_{X_{T-1}} -e^{-a(B_{T-2}+\sum_{n=1}^{N+1}\{P_{n,T-2}(X_{n,T-2}-X_{n,T-1})+E_{T-2}[P_{n,T-1}]X_{n,T-1}-0.5aVar_{T-2}[P_{n,T-1}]X^2_{n,T-1}\})}E_{T-2}[e^{-aK_{T-1}}] \tag{A.18}$$

where

$$K_{T-1} = \sum_{n=1}^{N+1}\left(-P_{n,T-1}X_{n,T} - I_{n,T-2}C_n - C_n + (\bar{u}_n + \sum_{t=1}^{T-1}\eta_{n,t} + \theta_{n,T})X_{n,T} - 0.5aX^2_{n,T}\sigma^2_{\epsilon}\right)).$$

FOC for $X_{n,T-1}$ leads to

$$-P_{n,T-2} + E_{T-2}[P_{n,T-1}] - aVar_{T-2}[P_{n,T-1}]X_{n,T-1} = 0,$$

or equivalently

$$X_{n,T-1} = \frac{E_{T-2}[P_{n,T-1}] - P_{n,T-2}}{aVar_{T-2}[P_{n,T-1}]}.$$

For the fraction $\mu_{n,T-2}$ of investors who produce information about stock n, their demand for stock n is $\frac{\bar{u}_n+\sum_{t=1}^{T-2}\eta_{n,t}+\theta_{n,T-1}-P_{T-2}}{a\sigma^2_{\epsilon,n}}$, and for the fraction $1-\mu_{n,T-2}$ of investors who do not observe the value $\theta_{n,T-1}$, their demand is $\frac{E_{T-2}(u_n|P_{n,T-2})-P_{n,T-2}}{aVar_{T-2}(u_n|P_{n,T-2})}$. The market clearing condition for stock n is therefore

$$\mu_{n,T-2}\frac{\bar{u}_n + \sum_{t=1}^{T-2}\eta_{n,t} + \theta_{n,T-1} - P_{T-2}}{a\sigma^2_{\epsilon,n}} + (1-\mu_{n,T-2})\frac{E_{T-2}(u_n|P_{n,T-2}) - P_{n,T-2}}{aVar_{T-2}(u_n|P_{n,T-2})} + z_{n,T-1} = y_n. \tag{A.19}$$

Following steps similar to Eqs. (A.8) through (A.11), we have

$$P_{n,T-2} = \bar{u}_n + \sum_{s=1}^{T-2}\eta_{n,s} + \alpha^{\theta}_{n,T-2}\theta_{n,T-1} + \alpha^z_{n,T-2}(z_{n,T-1} - y_n), \tag{A.20}$$

where

$$\alpha^{\theta}_{n,T-2} = 1 - \frac{(1-\mu_{n,T-2})a^2\sigma^2_{z,n}\Sigma^2_{n,T-1}}{\mu^2_{n,T-2}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma^2_{n,T-1} + a^2\mu_{n,T-2}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,T-1}},$$

$$\alpha^z_{n,T-2} = \frac{a\Sigma_{n,T-1}\left[\mu_{n,T-2}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma_{n,T-1}\left(\Sigma_{n,T-1} + \sigma^2_{\theta,n}\right)\right]}{\mu^2_{n,T-2}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma^2_{n,T-1} + a^2\mu_{n,T-2}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,T-1}},$$

and

$$\Sigma_{n,T-1} = \left(\alpha^{\theta}_{n,T-1}\right)^2\sigma^2_{\theta,n} + (\alpha^z_{n,T-1})^2\sigma^2_{z,n} + \sigma^2_{\epsilon,n}. \tag{A.21}$$

It is straightforward to use backward induction to show that

$$J_t(B_t, X_{n,t}) = \tag{A.22}$$

$$\max_{X_{t+1}} -e^{-a(B_t+\sum_{n=1}^{N+1}\{P_{n,t}(X_{n,t}-X_{n,t+1})+E_t[P_{n,t+1}]X_{n,t+1}-0.5aVar_t[P_{n,t+1}]X^2_{n,t+1}\})}E_{T-2}[e^{-aK_{t+1}}] \tag{A.23}$$

where

$$K_{t+1} = \sum_{n=1}^{N+1}(-P_{n,t+1}X_{n,t+2} - I_{n,t}C_n) + K_{t+2}$$

and the stock price at time t is given by Eqs. (10),

$$\alpha^{\theta}_{n,t} = 1 - \frac{(1-\mu_{n,t})a^2\sigma^2_{z,n}\Sigma^2_{n,t+1}}{\mu^2_{n,t}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\mu_{n,t}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}}, \tag{A.24}$$

$$\alpha^z_{n,t} = \frac{a\Sigma_{n,t+1}[\mu_{n,t}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma_{n,t+1}(\Sigma_{n,t+1} + \sigma^2_{\theta,n})]}{\mu^2_{n,t}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\mu_{n,t}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}}, \tag{A.25}$$

and

$$\Sigma_{n,t+1} = (\alpha^{\theta}_{n,t+1})^2\sigma^2_{\theta,n} + (\alpha^z_{n,t+1})^2\sigma^2_{z,n} + \sigma^2_{\epsilon,n}. \quad \square \tag{A.26}$$

Proof of proposition 3

(i) The derivative of $V^{Info}\left(r'_{n,t}\right)$ with respect to $\mu_{n,t}$ isSince

$$\frac{\partial V^{Info}(r'_{n,t})}{\partial \mu_{n,t}} = \frac{2a^2\Sigma^2_{n,t+1}\sigma^2_{z,n}\sigma^2_{\theta,n}\left[\mu_{n,t}(2-\mu_{n,t})\sigma^2_{\theta,n} + a^2\Sigma^2_{n,t+1}\sigma^2_{z,n} + a^2\sigma^2_{\theta,n}\sigma^2_{z,n}\Sigma_{n,t+1}\right]}{\left[\mu^2_{n,t}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\mu_{n,t}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right]^3} \times \left[\mu^2_{n,t}\sigma^2_{\theta,n} + (2\mu_{n,t}-1)a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\mu_{n,t}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right]$$

$\frac{2a^2\Sigma^2_{n,t+1}\sigma^2_{z,n}\sigma^2_{\theta,n}\left[\mu_{n,t}(2-\mu_{n,t})\sigma^2_{\theta,n}+a^2\Sigma^2_{n,t+1}\sigma^2_{z,n}+a^2\sigma^2_{\theta,n}\sigma^2_{z,n}\Sigma_{n,t+1}\right]}{\left[\mu^2_{n,t}\sigma^2_{\theta,n}+a^2\sigma^2_{z,n}\Sigma^2_{n,t+1}+a^2\mu_{n,t}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right]^3} > 0$ for all $\mu_{n,t} \in [0,1]$, while $\left[\mu^2_{n,t}\sigma^2_{\theta,n} + (2\mu_{n,t}-1)a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\mu_{n,t}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right]$ is negative for $\mu_{n,t} \in [0,\hat{\mu}_{n,t})$ and positive for $\mu_{n,t} \in [\hat{\mu}_{n,t}, 1]$, where we have $\frac{\partial V^{Info}(r'_{n,t})}{\partial \mu_{n,t}} < 0$ for $\mu_{n,t} \in [0,\hat{\mu}_{n,t})$ and $\frac{\partial V^{Info}(r_{n,t})}{\partial \mu_{n,t}} \geqslant 0$ for $\mu_{n,t} \in [\hat{\mu}_{n,t}, 1]$.

(ii) Since $\alpha^z_{n,t} > 0$, we have

[18] We drop the subscript j since all investors with the same portfolio $(B_{T-1}, X_{n,T-1})$ will have the same expected payoff.

$$\frac{\partial V^{Noise}(r'_{n,t})}{\partial \mu_{n,t}} = 2\alpha^z_{n,t}\sigma^2_{z,n}\frac{\partial \alpha^z_{n,t}}{\partial \mu_{n,t}} < 0. \qquad \square$$

Proof of proposition 4. We have

$$Var_t(u_n) = (T-t)(\sigma^2_\epsilon + \sigma^2_\theta),$$

and

$$Var_t(u_n|P_{n,t}) = \sigma^2_\epsilon + \sigma^2_\theta \frac{(\alpha^z_{n,t})^2\sigma^2_z}{(\alpha^\theta_{n,t})^2\sigma^2_\theta + (\alpha^z_{n,t})^2\sigma^2_z} + (T-t-1)(\sigma^2_\epsilon + \sigma^2_\theta).$$

Therefore,

$$\Psi_{n,t} = 1 - \frac{Var_t(u_n|P_{n,t})}{Var_t(u_n)} = \frac{\left(\alpha^\theta_{n,t}\right)^2\sigma^4_\theta}{(T-t)(\sigma^2_\epsilon+\sigma^2_\theta)\left[(\alpha^\theta_{n,t})^2\sigma^2_\theta + (\alpha^z_{n,t})^2\sigma^2_z\right]}$$
$$= \frac{\sigma^4_\theta}{(T-t)(\sigma^2_\epsilon+\sigma^2_\theta)\left[\sigma^2_\theta + \left(\frac{\alpha^z_{n,t}}{\alpha^\theta_{n,t}}\right)^2\sigma^2_z\right]}.$$

We have proved that $\frac{\partial \alpha^\theta_{n,t}}{\partial \mu_{n,t}} > 0$ and $\frac{\partial \alpha^z_{n,t}}{\partial \mu_{n,t}} < 0$. Therefore, $\frac{\alpha^z_{n,t}}{\alpha^\theta_{n,t}}$ decreases with $\mu_{n,t}$ since both $\alpha^\theta_{n,t}$ and $\alpha^z_{n,t}$are positive, which means that $\Psi_{n,t}$ increases with $\mu_{n,t}$; i.e., $\frac{\partial \Psi_{n,t}}{\partial \mu_{n,t}} > 0$. The result $\frac{\partial \Psi_{n,t}}{\partial C_n} < 0$ follows directly from $\frac{\partial \Psi_{n,t}}{\partial \mu_{n,t}} > 0$ and $\frac{\partial \mu_{n,t}}{\partial C_n} < 0$. $\square$

Proof of proposition 5. The derivative of return volatility with respect to $\mu_{n,t}$ is

$$\frac{\partial Var(r'_{n,t})}{\partial \mu_{n,t}} = \frac{2a^2\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma^2_{n,t+1}}{\left[\mu^2_{n,t}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\mu_{n,t}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right]^3}$$
$$\times \left\{\mu_{n,t}(1-\mu_{n,t})\sigma^2_{\theta,n}\left[\mu^2_{n,t}\sigma^2_{\theta,n} + (2\mu_{n,t}-1)a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\mu_{n,t}\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right]\right.$$
$$- \left[\mu_{n,t}\sigma^2_{\theta,n} + a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right]$$
$$\left.\times \left[\mu^2_{n,t}a^2\sigma^2_{\theta,n}\sigma^2_{z,n}\Sigma_{n,t+1} + a^4\sigma^2_{\theta,n}\sigma^4_{z,n}\Sigma^2_{n,t+1} + a^4\sigma^4_{z,n}\Sigma^3_{n,t+1}\right]\right\}$$

To prove that no parameter values exist such that $\frac{\partial Var(r'_{n,t})}{\partial \Psi_{n,t}} \geqslant 0$ for all $\Psi_{n,t}$, note that when $\mu_{n,t} = 0$,

$$\frac{\partial Var\left(r'_{n,t}\right)}{\partial \mu_{n,t}} = -\frac{2a^2\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma^2_{n,t+1}}{\left[a^2\sigma^2_{z,n}\Sigma^2_{n,t+1}\right]^3}\left[a^2\sigma^2_{z,n}\Sigma^2_{n,t+1} + a^2\sigma^2_{z,n}\sigma^2_{\theta,n}\Sigma_{n,t+1}\right]$$
$$\times \left[a^4\sigma^2_{\theta,n}\sigma^4_{z,n}\Sigma^2_{n,t+1} + a^4\sigma^4_{z,n}\Sigma^3_{n,t+1}\right]$$

which is negative. By the chain rule, we have

$$\frac{\partial Var(r'_{n,t})}{\partial \mu_{n,t}} = \frac{\partial Var(r'_{n,t})}{\partial \Psi_{n,t}}\frac{\partial \Psi_{n,t}}{\partial \mu_{n,t}},$$

since $\frac{\partial \Psi_{n,t}}{\partial \mu_{n,t}} > 0$ and when $\mu_{n,t} = 0, \frac{\partial Var(r'_{n,t})}{\partial \mu_{n,t}} < 0$, we have $\frac{\partial Var(r'_{n,t})}{\partial \Psi_{n,t}} < 0$ when $\mu_{n,t} = 0$. This proves part (i). To prove part (ii), note that when $\frac{\sigma^2_{\theta,n}}{\sigma^2_{z,n}}$ is large enough, $Var(r'_{n,t})$ is dominated by $Var^{Info}(r'_{n,t})$, which has a U-shaped relation with $\mu_{n,t}$ as we have proved in Proposition 5. Similarly, to prove part (iii), note that when $\frac{\sigma^2_{\theta,n}}{\sigma^2_{z,n}}$ is small enough, $Var(r'_{n,t})$ is dominated by $Var^{Noise}(r'_{n,t})$, which has a negative relation with $\mu_{n,t}$. $\square$

Proof of proposition 6. The proof follows directly from the law of one price and the fact that a share of stock $n \in \{1, 2, \ldots, N\}$ in the original economy is equivalent to a portfolio of one share of stock n and β_n shares of stock $N+1$ in the equivalent economy, and a share of stock M in the original economy is equivalent to a share of stock $N+1$ in the equivalent economy. $\square$

Proof of corollary 1. Following Eq. (14), we have

$$Var^O(r'_{n,t}) = \beta Var(r'_{N+1,t}) + Var\left(r'_{n,t}\right)$$

for $n = 1, 2, \ldots, N$. Since $Var\left(r'_{N+1,t}\right)$ is independent of $\Psi_{n,t}$, the relation between $Var^O\left(r'_{n,t}\right)$ and $\Psi_{n,t}$ depends only on the relation between $Var\left(r'_{n,t}\right)$ and $\Psi_{n,t}$, which is summarized in Proposition 5. $\square$

References

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. Journal of Finance 61, 259–299.

Ashbaugh-Skaife, H., Gassen, J., LaFond, R., 2005. Does stock price synchronicity represent firm-specific information? The international evidence. Working Paper, University of Wisconsin-Madison.

Brockman, P., Yan, X., 2009. Block ownership and firm-specific information. Journal of Banking and Finance 33, 308–316.

Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? Am empirical exploration of idiosyncratic risk. Journal of Finance 56, 1–43.

Easley, D., Hvidkjaer, S., O'Hara, M., 2010. Factoring information into returns. Journal of Financial and Quantitative Analysis 45, 293–309.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E., MacBeth, J., 1973. Risk, return, and equilibrium: empirical tests. Journal of Political Economy 71, 607–636.

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. Journal of Financial Economics 91, 24–37.

Goyal, A., Santa-Clara, P., 2003. Idiosyncratic risk matters! Journal of Finance 58, 975–1007.

Guo, H., Savickas, R., 2008. Average idiosyncratic volatility in G7 Countries. Review of Financial Studies 21, 1259–1296.

Guo, H., Savickas, R., 2010. Relation between time-seires and cross-sectional effects of idiosyncratic variance on stock returns. Journal of Banking and Finance 34, 1637–1649.

Jin, L., Myers, S., 2006. R^2 around the world: new theory and new tests. Journal of Financial Economics 79, 257–292.

Kang, M., 2010. Probability of information-based trading and the January effect. Journal of Banking and Finance 34, 2985–2994.

Kelly, P.J., 2005. Information Efficiency and Firm-specific Return Variation. Working Paper, Arizona State University.

Krishnaswami, S., Subramaniam, V., 1999. Information asymmetry, valuation, and the corporate spin-off decision. Journal of Financial Economics 53, 73–112.

Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Mendenhall, R., 2004. Arbitrage risk and post-earnings-announcement drift. Journal of Business 77, 875–894.

Morck, R., Yeung, B., Yu, W., 2000. The information content of stock markets: why do emerging markets have synchronous stock price movements? Journal of Financial Economics 58, 215–260.

Ozoguz, A., 2005. Informed Trading, Return Synchronicity and Equilibrium Asset Pricing. Working Paper, Queen's University.

Pastor, L., Veronesi, P., 2003. Stock valuation and learning about profitability. Journal of Finance 58, 1749–1789.

Roll, R., 1988. R^2. Journal of Finance 43, 541–566.

Rubin, A., Smith, D., 2009. Institutional ownership, volatility and dividends. Journal of Banking and Finance 33, 627–639.

Thapa, C., Poshakwale, S., 2010. International equity portfolio allocations and transaction costs. Journal of Banking and Finance 34, 2627–2638.

West, K., 1988. Dividend innovations and stock price volatility. Econometrica 56, 37–61.