Investor Overreaction, Cross-Sectional Dispersion of Firm Valuations, and Expected Stock Returns

Danling Jiang *

Fisher College of Business
The Ohio State University

First draft: April 29, 2005
This draft: May 25, 2006

---

[*] This paper is adapted from one chapter of my dissertation. I am grateful to the members of my dissertation committee: Bing Han and Kewei Hou for many helpful suggestions, and especially my chair David Hirshleifer for guidance, support, encouragement, and numerous insightful comments. I thank Malcolm Baker, Mike Cliff, Henrik Cronqvist, Phil Davies, Karl Diether, James Doran, Jim Hsieh, Andrew Karolyi, Bong Soo Lee, Sonya Lim, Ji-Chai Lin, Roger Loh, Angie Low, Dave Peterson, Christof Stahel, René Stulz, Mathijs van Dijk, and the participants in the Brown Bag seminar and Fin923 seminar at The Ohio State University, in a research seminar at Florida State University and University of San Diego for helpful comments. I thank Karl Diether and Siew Hong Teoh for providing me the short interest data and Kenneth French, Jeffrey Wurgler, and Yihong Xia for making their data available online. I appreciate the financial support from the Presidential Fellowship at The Ohio State University and the National Science Foundation (NSF). All errors remain mine. Correspondence to: Danling Jiang, 700 Fisher Hall, 2100 Neil Avenue, Columbus, OH 43210. E-mail: jiang.87@osu.edu; http://fisher.osu.edu/~jiang_87.

Abstract

I develop and test the theoretical predictions that when investor overreaction to market-wide news is larger, firm valuations in the cross section become more dispersed and stocks earn lower expected returns. Consistent with these predictions, measures of cross-sectional dispersion of firm valuations are negatively related to subsequent market and portfolio excess returns, especially for sets of firms with highly subjective valuations and significant limits to arbitrage. Further, these firms underperform those with the opposite characteristics in periods when beginning-of-period firm valuation dispersion is high. In contrast, they overperform when beginning-of-period firm valuation dispersion is low.

Introduction

In the past two decades, finance research has documented that high price-to-fundamental ratios tend to predict low subsequent returns for firms in the cross section and for the market as a whole in the time-series.[1] One possible explanation is that investor overreaction to information can cause stock price to overshoot; thus high current valuation relative to fundamentals tends to reverse and leads to low subsequent returns at the firm level and at the aggregate level (e.g., Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Hong and Stein (1999)). Behavioral theory also suggests that investor overreaction to market-wide information can simultaneously affect returns on the aggregate market and on individual stocks to different extents (Daniel, Hirshleifer, and Subrahmanyam (2001), hereafter DHS (2001)). However, so far this effect has received little empirical attention in the literature.

This paper provides new hypotheses and evidence for how investor overreaction affects the predictability of both the aggregate and the cross section of stock returns. Specifically, building upon the behavioral model of DHS (2001), I develop the prediction that when investor overreaction to aggregate information is larger, firm valuations in the cross section become more dispersed, and stocks on average earn lower expected returns, which I call the *overreaction hypothesis*. I therefore test whether cross-sectional dispersion of firm valuations can forecast stock returns. I form a composite measure of cross-sectional dispersion of firm-level valuations that incorporates the cross-sectional standard deviations of three logarithmic firm valuation ratios: book-to-market equity, dividend-to-price ratio, and earnings-to-price ratio, and call this composite measure *cross-firm valuation dispersion* (CVD).

Consistent with the model predictions, I find that (1) cross-firm valuation dispersion is a negative predictor of subsequent market and portfolio excess returns at least up to three years ahead and (2) the dispersion-return relationship is particularly strong among firms whose valuations are highly subjective and for which arbitrage is considerably limited. Result (1) is related to the value-spread literature (e.g., Brennan, Wang, and Xia (2001), Campbell and Vuolteenaho (2004)). This literature finds that the value spread, defined as the difference of the logarithmic book-to-market equity between value firms and growth firms (Cohen, Polk,

---

[1] There is an enormous literature on the relationship between firm price-to-fundamental ratios and the cross section of stock returns (e.g., Fama and French (1992)). There is also a large literature on the relationship between aggregate price-to-fundamental ratios and aggregate returns (e.g., Campbell and Shiller (1988), Fama and French (1988a), Kothari and Shanken (1997), Pontiff and Schall (1998), Lamont (1998), and Ang and Bekaert (2004)). The aggregate predictability is recently challenged by Goyal and Welch (2005). However, using a new method to adjust for the small sample bias, Lewellen (2004) finds strong predictive power of the fundamental-to-price ratios than previously documented.

and Vuolteenaho (2003)), forecasts one-month-ahead market excess returns. My measure CVD differs from the value spread in that CVD is based on cross-sectional standard deviation of firm valuation ratios, which is essential to distinguish the overreaction hypothesis from a hypothesis based on risk while the value spread is not (this point will be discussed in detail later).[2] Also, previous literature on the value spread does not address point (2), which provides evidence from the cross section of returns to support the overreaction hypothesis.[3]

Furthermore, I find that sets of difficult-to-value/difficult-to-arbitrage firms underperform those with opposite characteristics in periods with high beginning-of-period CVD—when overreaction is large. In contrast, when the beginning-of-period dispersion is low, these conceptually riskier firms overperform. This evidence suggests that the returns on these firms are more driven by mispricing during high dispersion periods but are more determined by risk premia when mispricing is relatively small during low dispersion periods. These results are robust to controls for the Fama-French 3-factor model (Fama and French (1993)), the 4-factor model that additionally includes a momentum factor (Carhart (1997)), and the ICAPM of Brennan, Wang, and Xia (2004) (hereafter BWX ICAPM).

In my model, investors overreact to private signals about cash flows of common economic factors. The source of overreaction is investor overconfidence about private information. I show that the variation in overconfidence can produce a negative correlation between firm valuation dispersion and expected aggregate returns. This negative correlation is attributed to the combination of two distinct effects of overconfidence about signal precision. First, overconfident investors incorrectly estimate expected aggregate cash flows and produce forecast error. I call this *the mean bias effect.* The forecast error generates an expected aggregate return that is distinctive from risk premia. This expected return is negative following good news and positive following bad news. The absolute expected market return is greater when overconfidence is stronger. Second, overconfident investors underestimate cash flow volatility; thus, they underestimate risk and reduce the market risk premium and the expected market returns. I call this the *risk premium reduction effect.* The expected market return is smaller when overconfidence is greater.

---

[2] Apart from the conceptual difference between CVD and the value spread, I find that CVD performs better in forecasting future aggregate returns. CVD captures not only dispersion in book-to-market equity, but also in dividend-to-price ratios and earnings-to-price ratios and it is defined to be free of a time trend in firm valuation dispersion during the sample period, both of which substantially enhance the ability of CVD over the value spread to forecast long-horizon aggregate returns.

[3] I show that CVD predicts the cross-sectional differences in returns after controlling for the BWX ICAPM or the other major asset pricing models, which provides an alternative view that the value spread, instead of a proxy for the state variable of the investment opportunities (Brennan, Wang, and Xia (2001)), is a proxy for investor overreaction.

At the firm level, firm valuation (defined as the difference between measures of firm fundamentals and stock price, denoted as $C-P$), reflects the expected market return according to firm beta. Thus, at each point in time, cross-sectional *dispersion* of firm $C-P$ is determined by the absolute expected market returns with cross-sectional *dispersion* of beta as a multiplier. Given the dispersion of beta, the more extreme the expected market return, the larger the cross-sectional dispersion of $C-P$. Hence, when the average investor becomes more overconfident, firm valuation dispersion is larger due to the mean bias effect and the expected aggregate returns are on average smaller due to the risk premium reduction effect.

In contrast, the CAPM produces opposite the predictions from the overreaction model. In my model, when investors are fully rational, the unconditional CAPM holds. The expected market return is solely determined by market risk premium, which is always positive. Thus, the greater the market risk premium, the greater the absolute expected market returns and the larger the firm valuation dispersion. Hence, when variation in expected market returns comes solely from variation in rational risk premia, larger firm valuation dispersion should be associated with greater expected market returns. I call this the *risk hypothesis*. The same predictions are offered in a general equilibrium setting by Gomes, Kogan, and Zhang (2003) in which a conditional CAPM holds.[4]

Taken together, CVD reflects the extremity of the expected market returns, which leads to opposite predictions under two competing hypotheses. In a market with investor overreaction, more extreme expected market returns indicate greater overconfidence and lower average expected market returns.[5] In contrast, in a market with fully rational investors, more extreme expected market returns suggest larger market risk premia and higher expected market returns. Therefore, I can test the overreaction hypothesis against the risk hypothesis by examining the correlation between CVD and future market excess returns.

In contrast, the value spread does not help to distinguish the two hypotheses. In my model, the value spread is determined by the expected market returns with the difference in beta between value stocks and growth stocks as a multiplier. Thus, the correlation between the value spread and the expected market returns is determined by the difference in beta regardless of whether investors are fully rational or subject to behavioral biases.

My empirical results support the overreaction hypothesis and reject the risk hypothesis. I

---

[4] Through simulation, Gomes, Kogan, and Zhang (2003) show that in their model cross-sectional dispersion of log book-to-market ratios is negatively correlated with aggregate log market-to-book equity. Since aggregate market-to-book equity is negatively correlated with expected aggregate returns, this result implies that cross-sectional dispersion of log book-to-market ratios should be a positive predictor of future aggregate returns.

[5] An alternative mechanism to generate the negative correlation between CVD and expected aggregate returns, shown by Jiang (2005a), is to combine investor overreaction and aggregate short-sale constraints.

find that, during the period 1963–2004, CVD is a negative predictor of market excess returns at least up to three years ahead. This result is consistent with the previously documented negative correlations between the value spread and future aggregate returns. In contrast with previous value spread findings which do not distinguish a behavioral explanation from a rational explanation, this finding for CVD helps distinguish the two. In addition, I uncover much stronger and more consistent negative correlation between CVD and future market returns across different return horizons.

CVD has a large economic effect on expected returns. Using the value-weighted market portfolio as an example, during my sample period a one standard deviation positive shock in dispersion leads to more than an 8% reduction in the following one-year returns and nearly a 19% reduction in the following three-year returns. For various forecast horizons, the predictive power of CVD is generally robust to controls for a number of standard aggregate return predictors.[6] Furthermore, CVD adds considerable incremental forecasting power to the set of standard predictors. It increases the fraction of explained variance of the value-weighted market returns from 2% to 6% when forecasting one-quarter-ahead returns, from 8% to 16% when forecasting one-year-ahead returns, and from 24% to 60% when forecasting three-year-ahead returns.

The overreaction hypothesis carries more unique predictions in the cross section. First, the risk premium reduction effect should be stronger among risky firms; a change in the market risk premium corresponds to a greater change in the firm risk premium when a stock has a large beta. For a similar reason, it should be most pronounced among firms whose cash flows are highly uncertain. By cash flow uncertainty, I mean the perceived ambiguity with respect to the cash flows inferred from information (Zhang (2006), Kumar (2005)). These firms are likely to load heavily on factors that govern these uncertain cash flows and, accordingly, that are more difficult to value and hence more subject to investor disagreement and overreaction (e.g., Hirshleifer (2001), Daniel, Hirshleifer, and Subrahmanyam (1998), DHS (2001)). Thus, risky firms, or firms with highly uncertain cash flows or large investor disagreement should be more affected by overreaction to information about economic factors that drive stocks' cash flows. As a result, the negative dispersion-return relation should be stronger among these firms.

Furthermore, when short sales are constrained, the mean bias effect can reinforce the

---

[6] They include aggregate dividend yield (Campbell and Shiller (1988), Fama and French (1988a), Lewellen (2004)), short-term interest rate (Fama and Schwert (1977)), past market returns (Fama and French (1988b)), default premium and term premium (Fama and French (1989)), and aggregate relative equity issuances (Baker and Wurgler (2000)).

relation between market-wide overreaction and returns on the above firms. As shown by Scheinkman and Xiong (2003a), Hong, Scheinkman, and Xiong (2006), and Jiang (2005a), when short-selling is costly, stock prices overreact to good news strongly but to bad news only slightly. On average, there is more upward bias in forecasting expected aggregate cash flows. Thus, firm valuation dispersion should further negatively forecast returns when the upward forecast error is corrected. In other words, I expect a larger correlation between dispersion and returns among firms with higher potential costs of short selling.[7]

Consistent with these predictions, I find that CVD better forecasts returns among four sets of firms: (1) firms that are risky, such as those with large beta or high return volatility; (2) firms that have more uncertain cash flows, such as those in the technology-related industries, with low age, few fixed assets, low profits, or low dividends; (3) firms that have large investor disagreement, such as those with high analyst forecast dispersion, or large trading volume; and (4) firms that have high potential cost of selling short, such as those with low institutional ownership or high short interest ratio. I call the four sets of firms the *difficult-to-value/difficult-to-short firms.* CVD exhibits stronger power to forecast returns on these sets of firms; the predictive slopes tend to be more negative and the R-squares tend to be larger for these portfolios than for their counterparts.

Prior research has documented that these difficult-to-value/difficult-to-short firms tend to underperform those with the opposite characteristics.[8] However, I find that such underperformance is almost entirely concentrated in periods with high beginning-of-period CVD—which represents larger overreaction, larger mean bias, and lower market risk premia. In particular, when the beginning-of-period dispersion is above the mean, these difficult-to-value/difficult-to-short firms underperform their counterparts by an average annual return of 3–16%. When the beginning-of-period dispersion is below the mean—which suggests smaller overreaction, smaller mean bias, and higher market risk premia—the return patterns are essentially reversed. During such periods, the presumably riskier stocks, such as those with large beta, high volatility, low age, low profits, or low institutional ownership, on average overperform their counterparts.

The overreaction hypothesis implies that CVD proxies for investor overreaction to market-

---

[7] There is a strand of literature on limits to arbitrage and mispricing (Shleifer and Vishny (1997)), particularly on short-sale constraints and overpricing when investors have divergent opinions about stock valuations (e.g., D'Avolio (2002), Diether, Malloy, and Scherbina (2002), Chen, Hong, and Stein (2002), and Jones and Lamont (2002)).

[8] These firms include those with high volatility (Ang, Hodrick, Xing, and Zhang (2006)), high short interest ratio (Asquith and Meulbroek (1995), Desai, Ramesh, Thiagarajan, and Balachandran (2002)), high analyst forecast dispersion (Diether, Malloy, and Scherbina (2002)), large trading volume (Lee and Swaminathan (2000)).

wide information. Alternatively, CVD could merely be a proxy for time-varying risk premia. I investigate this alternative interpretation in a few robustness tests. Several findings do not support this alternative hypothesis. First, CVD continues to predict the returns on the long-short characteristic portfolios after controlling for the return comovement with factors in major asset pricing models, including the Fama-French 3-factor model, the 4-factor model, and the BWX ICAPM. That is, the cross-sectional differences in the predictability of portfolio returns are not attributed to the cross-sectional differences in risk.

Second, the findings at the aggregate level are also inconsistent with the FF 3-factor model, the 4-factor model, or the BWX ICAPM. To test whether these models can explain the negative relation between firm valuation dispersion and the market excess returns, I develop a framework built upon the decomposition of book-to-market equity (BM) (Vuolteenaho (2000, 2002), Cohen, Polk, and Vuolteenaho (2003)), of dividend-to-price ratio (DP) (Campbell and Shiller (1988)), and of earnings-to-price ratio (EP). This framework relates firm BM, DP, and EP to stock returns that are described by each of these models. It suggests that the firm valuation dispersion should be positively related to future market excess returns in a nonlinear fashion. However, the tests clearly reject this nonlinear relationship for all of these models. Therefore, the relationship between CVD and the market premium is unlikely to be explained by these models.

A few papers have examined relationships between firm-level and aggregate valuation in different contexts. Baker and Wurgler (2000) and Eleswarapu and Reinganum (2004) test for effects on subsequent aggregate returns conditioning on a set of potentially mispriced firms.[9] My results on aggregate return predictability complement their findings. However, my aggregate mispricing measures are extracted from firm valuation ratios and from all firms in the three major exchanges.

Also Baker and Wurgler (2005) and Dong, Hirshleifer, Richardson, and Teoh (2003) test for effects on subsequent returns or behaviors of individual firms conditioning on possible aggregate mispricing.[10] My findings in the cross section are similar in spirit to the work of Baker and Wurgler (2005) that difficult-to-value and difficult-to-arbitrage firms are most sensitive to aggregate investor demand. However, I use a novel measure of aggregate overreaction and examine more direct proxies for shorting difficulty: institutional ownership (Nagel

---

[9] Baker and Wurgler (2000) focus on equity issuers and use the aggregate issuing activity to forecast market excess returns. Eleswarapu and Reinganum (2004) concentrate on growth firms and use the growth firm returns to predict future market excess returns.

[10] Baker and Wurgler (2005) develop an investor sentiment index, which by construction to some extent reflects aggregate mispricing. Dong, Hirshleifer, Richardson, and Teoh (2003) use cross-sectional standard deviation of price-to-book and price-to-value to proxy for aggregate mispricing but they focus on mergers and acquisitions.

(2005)) and short interest ratio (Boehme, Danielsen, and Sorescu (2005)).

The remainder of the paper is organized as follows. In Section II, I present the model and develop the test hypotheses. Section III describes the data and the key variables. Section IV shows the empirical results. Section V explores other alternative explanations. Section VI checks the robustness. Section VII summarizes and draws conclusions.

II. The Model

As in the model of DHS (2001), there is a finite mass of risk averse overconfident investors, denoted as C, who hold one riskless asset and N risky assets. There are three dates, $t = 0, 1, 2$. At date 0, investors start with their endowments and identical prior beliefs about the security payoffs. It is known to all investors that, at date 2, the riskless asset pays one unit per share and the risky asset pays θ_i, for all $i = 1, ..., N$. At date 1, investors receive an identical noisy private signal about the payoff of the common factor and exchange assets based on their beliefs. At date 2, the risky asset pays a liquidating dividend of θ and all consumption takes place.

Each risky asset has per capita supply of Q shares and its payoff at date 2 follows a single-factor structure:

$$\theta_i = \bar{\theta}_i + \beta_i F + \epsilon_i. \tag{1}$$

where $\bar{\theta}_i$ is the expected payoff of security i, ϵ_i is the firm-specific payoff, independently identically distributed as $N(0, 1/v^{\epsilon})$, and β_i is the loading of the ith security on the factor F. The common factor F is normally distributed as $N(0, 1/v)$. In addition, $E(F\epsilon_i) = 0$. The security loading, β, takes the values of β_1,..., β_N. I normalize the factor F to set the average β as one and denote the cross-sectional variance of β as $\sigma^2(\beta)$. The values of $\bar{\theta}_i$, β_i, and the distribution of ϵ_i and F are common knowledge, but the realizations of ϵ_i and F are not known until date 2.

The noisy private signal about the common factor payoff at date 1 takes the form

$$S = F + e, \tag{2}$$

where e is the noise of the signal. It is normally distributed as $N(0, 1/v^R)$, where v^R is the true/rational precision of the signal. Overconfident investors mistakenly believe that the variance of the noise is lower, $1/v^C < 1/v^R$ (i.e., $v^C > v^R$). Thus, the mean and variance of the common factor payoff conditional on the signals are given by:

$$\mu_C = \frac{v^C S}{v + v^C}, \quad \sigma_C^2 = \frac{1}{v + v^C}.$$

In contrast, if investors were rational, the mean and variance would be

$$\mu_R = \frac{v^R S}{v + v^R}, \quad \sigma_R^2 = \frac{1}{v + v^R}.$$

After receiving the private signal, each investor selects her portfolio to maximize a CARA utility function with a risk aversion coefficient of A for date 2 consumption. Following DHS (2001), before I solve for the equilibrium price of individual securities, I first solve for the equilibrium price of the *factor portfolio*, which is constructed to have an expected payoff of zero and a loading of one on the common factor F. The equilibrium price of the factor portfolio is given by

$$P = \mu_C - A\sigma_C^2 Q. \tag{3}$$

The first term is the investor expected factor cash flow. The second term is the price discount for risk, in which the conditional factor volatility is the investor perceived volatility. Due to the single factor payoff structure, the factor portfolio can also be defined as the market portfolio; therefore, the price of the market portfolio P_m is equal to P.[11]

Equation (3) suggests two distinct effects of overconfidence. First, overconfidence generates biased estimation of the expected cash flow. I define the factor cash flow mean bias, denoted as M, as the difference between the investor expected factor cash flow and the true expected factor cash flow, $M = \mu_C - \mu_R$. Upon a favorable signal $S > 0$, the expected factor cash flow is overestimated ($M > 0$). Conversely, upon an adverse signal $S < 0$, the expected factor cash flow is underestimated ($M < 0$). Further, it is easy to show that the forecast error, defined as the absolute mean bias ($|M|$), increases as overconfidence (measured by v^C) rises. I call this the mean bias effect.

Second, overconfidence produces lower perceived cash flow volatility. The conditional variance perceived by an overconfident investor is always smaller than the true variance ($\sigma_C^2 < \sigma_R^2$). This underestimation of cash flow volatility implies that, given the true cash flow volatility, the more overconfident the investors, the lower the risk premium they require. I call this the risk premium reduction effect.

Let π denote the risk premium $A\sigma_C^2 Q$. The aggregate price P_m can be decomposed into three components: the true expected aggregate cash flow (μ_R), the cash flow mean bias (M), and the risk premium (π). Then equation (3) can be rewritten as

$$P_m = \mu_R + M - \pi. \tag{4}$$

The above equation shows that the aggregate price is high when the expected aggregate cash

---

[11] When N is large enough, the idiosyncratic risk is diversified away in the market portfolio.

flow is high, the risk premium is low, or the expected cash flow is overestimated (M is more positive).

Equations (3) and (4) establishe a relationship between the expected aggregate return and overconfidence. The expected aggregate return, $\pi - M$, is equal to the difference between the true expected factor cash flow and the aggregate price, $\mu_R - P_m$. On average, the expected aggregate return is equal to π because cash flow mean overestimation ($M > 0$) and underestimation ($M < 0$) are symmetric.[12] Thus, upward bias and downward bias are cancelled out. As a result, the mean bias effect has on average no effect on expected aggregate return. In contrast, the risk premium π decreases with overconfidence. In sum, on average, the aggregate return is low when the overconfidence level is high.

To examine the relationship between firm valuation dispersion and the expected aggregate return, the next step is to derive the relationship between dispersion and overconfidence. In equilibrium, the price of asset i is determined by its unconditional expected cash flow and its cash flow sensitivity to the common factor.[13] That is,

$$P_i = \bar{\theta}_i + \beta_i P_m. \tag{5}$$

Let C_i be a noisy measure of the true expected cash flow of asset i conditional on the factor signal. Specifically,

$$C_i = \bar{\theta}_i + \beta_i \mu_R + \nu_i, \tag{6}$$

where ν_i is a firm-specific noise, independently identically distributed as $N(0, \sigma_\nu^2)$, and $E[\beta_i \nu_i] = 0$.[14] Empirically, C_i is proxied by fundamental measures such as book equity, dividends, or earnings.

We measure valuation of asset i by $C_i - P_i$, which can be empirically proxied by the difference between logarithmic fundamental measures and logarithmic price. For a sufficient large number of assets, the average $C_i - \bar{\theta}_i$, denoted as C_m, is equal to μ_R. Thus, the market valuation is measured by $C_m - P_m$. Combining equations (5) and (6) yields a simple relationship between asset i's valuation and the market's valuation,

$$C_i - P_i = \beta_i (C_m - P_m) + \nu_i. \tag{7}$$

---

[12] Since S is normally distributed with zero mean, cash flow misestimation is symmetric in both magnitude and probability. Thus, the unconditional mean of M is zero.

[13] This relationship can be derived from a non-arbitrage argument: since the payoff of asset i can be replicated by holding $\bar{\theta}_i$ units of the riskfree asset, β_i units of the market portfolio, and one unit of ϵ_i, the price of asset i should be equal to the sum of the prices of the three components in which ϵ_i is not priced since it is diversifiable risk.

[14] Alternatively, one can model C as a noisy measure of the true unconditional cash flow such that $C_i = \bar{\theta}_i + \nu_i$. This assumption leads to the same conclusion as the current model.

Replacing P_m in equation (7) with that in equation (5) gives

$$C_i - P_i = \beta_i(\pi - M) + \nu_i. \tag{8}$$

Taking the cross-sectional variance of both sides of equation (8) produces

$$\sigma^2(C - P) = \sigma^2(\beta)(\pi - M)^2 + \sigma^2_\nu. \tag{9}$$

Since σ^2_ν is a known constant, we can move it to the left-hand side of equation (11) and define the adjusted cross-sectional variance $\hat{\sigma}^2(C-P) = \sigma^2(C-P) - \sigma^2_\nu$. Thus, the adjusted firm valuation dispersion is[15]

$$\hat{\sigma}(C - P) = \sigma(\beta)\left|\pi - M\right|. \tag{10}$$

To obtain the intuition about the effects of overconfidence on firm valuation dispersion, I consider two extreme cases. In Case 1, there is no risk premium ($\pi = 0$), and hence $\hat{\sigma}(C-P) = \sigma(\beta)\left|M\right|$. When overconfidence rises, due to the mean bias effect, firm valuation dispersion becomes greater. In Case 2, there is no mean bias ($M = 0$) and hence $\hat{\sigma}(C-P) = \sigma(\beta)\pi$. When overconfidence is stronger, due to the risk premium reduction effect, firm valuation dispersion becomes smaller.

However, the two effects are not equally strong. In Appendix A, I show that the increase in the mean bias effect is faster than that in the risk premium reduction effect as overconfidence rises. This is because that overestimation of signal precision has a stronger effect on the conditional mean than on the conditional variance. Therefore, when overconfidence is strong enough, the mean bias effect will dominate the risk premium reduction effect to determine firm valuation dispersion, which leads to

---

[15] Equation (10) holds only when $\sigma(\beta)$ is non-zero, which is consistent with the empirical observation that there is a fairly large range of betas in the cross section. Alternatively, in my model, when there is no cross-sectional dispersion in betas and when investors also receive signals about the idiosyncratic payoff ϵ, $\hat{\sigma}(C-P)$ would be a measure of cross-sectional dispersion in idiosyncratic returns, which is similar in spirit as the measure of average idiosyncratic return volatility in Goyal and Santa-Clara (2003). However, $\hat{\sigma}(C-P)$ differs from the measure of Goyal and Santa-Clara by capturing the volatility of the future idiosyncratic returns implied in today's stock prices instead of the past return volatility. In that case, my result about the negative correlation between $\hat{\sigma}(C-P)$ and future market returns would be interpreted as a negative correlation between the average expected idiosyncratic volatility and the future market returns. This finding is then inconsistent with the finding of Goyal and Santa-Clara (2003) but consistent with the finding of Guo and Savickas (2006) for the aggregate returns and that of Ang, Hodrick, Xing, and Zhang (2006) about the cross section of returns. Further, my results about the cross section of returns suggest that interpreting CVD as a measure of average idiosyncratic volatility should be cautious. For example, I find returns on high beta stocks are more sensitive to the change in CVD than returns on low beta stocks, which suggests that CVD is unlikely to be solely associated with the idiosyncratic component of stock returns.

Proposition 1. *As overconfidence* v^C *increases,*
1. If overconfidence is sufficiently strong ($v^C > v^{C'}$*, where* $v^{C'}$ *is a constant), then on average the adjusted cross-sectional dispersion of firm valuations* $\hat{\sigma}(C-P)$ *increases; and*
2. However, the expected aggregate return on average decreases.

See Appendix A for proof of all propositions.

Thus, in a sufficiently overconfident market, firm valuation dispersion is a negative predictor of future market excess returns. When firm valuations in the cross section are more dispersed, the expected market return tends to be low.

In contrast, the variation in risk aversion or in factor cash flow volatility delivers an opposite prediction. When overconfidence is absent ($v^C = v^R$), the CAPM holds. Since no forecast error is produced ($M = 0$), the adjusted firm valuation dispersion collapses to

$$\hat{\sigma}(C-P) = \sigma(\beta)\pi. \tag{11}$$

The above equation leads to

Proposition 2. *As risk aversion* A *or factor cash flow volatility* $1/v$ *increases,*
1. The adjusted cross-sectional dispersion of firm valuation $\hat{\sigma}(C-P)$ *increases; and*
2. The expected aggregate return also increases.

Thus, if variation in expected returns is due to variation in rational risk premia, firm valuation dispersion should be a positive predictor of future market excess returns.[16] Intuitively, when risk aversion or factor cash flow volatility is high, the market risk premium is high, which increases both firm valuation dispersion and the expected aggregate return.

Taken together, Propositions 1 and 2 provide testable competing hypotheses; the overreaction hypothesis implies a negative correlation between firm valuation dispersion and aggregate returns while the risk hypothesis implies a positive correlation between the two.[17]

---

[16] In this model, I assume $\sigma(\beta)$ is a constant. However, $\sigma(\beta)$ can be time-varying in a conditional CAPM. For example, the model in Gomes, Kogan, and Zhang (2003) implies that cross-sectional dispersion of beta should be positively correlated expected market returns, which is consistent with Proposition 2. In contrast, Santos and Veronesi (2004) show that when beta is more determined by firm's fundamental risk than by market premium or firm's expected dividend growth, cross-sectional dispersion of beta is negatively correlated with market premium. Thus, the time-variation in β, under some condition, can also explain the negative correlation between firm valuation dispersion and future market returns. However, in Section IV C.1, I show that when firm valuation dispersion is high, β is negative correlated with the cross section of stock returns, which is inconsistent with the prediction of Santos and Veronesi (2004) that β should carry a positive premium. Thus, the model of Santos and Veronesi (2004) cannot fully explain my results.

[17] Note that the overreaction hypothesis holds only when the average investor is sufficiently overconfident. Thus, the test is a joint test of this hypothesis and a sufficiently high level of overconfidence.

In contrast, the correlation between the value spread and the expected aggregate return carries the same sign under the two hypotheses. Let the average beta of the value firms be β_v and that of the growth firms be β_g. When the number of firms in each group is sufficiently large, the value spread

$$(C - P)_v - (C - P)_g = (\beta_v - \beta_g)E(R_m). \tag{12}$$

Thus, the correlation between the value spread and the expected aggregate returns is determined by $\beta_v - \beta_g$ regardless whether the variation in the expected return is caused by variation in rational risk premium or in mispricing. A number of studies (e.g., Campbell and Vuolteenaho (2004), Ang and Chen (2005)) find that growth stocks on average have higher beta than value stocks after 1963. Therefore, equation (12) predicts that the value spread should be negatively correlated with the expected aggregate return, which is consistent with the findings in the value spread literature. However, this negative correlation between the value spread and expected returns does not help distinguish a behavioral theory from a rational theory.

Similarly, both models predict a positive correlation between $C_m - P_m$ and the expected aggregate returns since

$$C_m - P_m = E(R_m). \tag{13}$$

Thus, a positive correlation between aggregate fundamental-price ratios and subsequent aggregate returns does not differentiate between a fully rational market and an overconfident market, either.

Furthermore, the two hypotheses carry implications for the cross section of stock returns.

Proposition 3.

1. As overconfidence v^C increases, the reduction in the expected return of assets with higher betas is on average greater than that of assets with lower betas.

2. However, as risk aversion A or factor cash flow volatility $1/v$ increases, the increase in the expected return of assets with higher beta assets is greater than that of assets with lower betas.

Thus, based upon the overconfidence hypothesis, the negative dispersion-return relation should be stronger among sets of high beta firms; when overconfidence reduces market risk premium, it should reduce the risk premium of high beta stocks more than that of low beta stocks. Following a similar reasoning, the risk hypothesis predicts that high beta stocks should exhibit a stronger positive dispersion-return relationship.

In addition, the overreaction hypothesis predicts a stronger dispersion-return relation among firms whose valuations are highly subjective, which can hold even if these firms' total exposure to risk (beta) is no greater than that of the other firms. Psychology studies suggest that people tend to be more overconfident when a task is more difficult and feedback is more delayed or less conclusive (e.g., Einhorn (1980)). Firms with more subjective valuations usually have highly uncertain cash flows, which increases the difficulty in valuation task and the deviation of investor valuation from the rational valuation (Daniel, Hirshleifer, and Subrahmanyam (1998)). These firms are likely to load heavily on factors that generate highly uncertain cash flows and thus these firms are more likely to be subject to investor overreaction to factor payoffs.

To illustrate the intuition, suppose there are two common factors, F_1 and F_2, both of which are distributed as $N(0, 1/v)$. I also assume that investors are more overconfident about the precision of their signals about F_1's cash flow (i.e., $v_1^C > v_2^C$). For example, F_1 is a tech factor, F_2 is an oil factor, and investors are more overconfident when valuing the tech sector. Investors receive two independent noisy signals about payoffs of each common factor. Also there are two classes of risky assets whose payoffs follow

$$\theta_1 = \bar{\theta}_1 + \beta_{1,1} F_1 + \epsilon_1, \quad \text{and} \quad \theta_2 = \bar{\theta}_2 + \beta_{2,2} F_2 + \epsilon_2, \tag{14}$$

In addition, I assume that $\beta_{1,1} = \beta_{2,2}$ so that the two stock classes have the same total exposure to systematic risk. For example, asset class one are software firms, which load only on the tech factor, and asset class two are gas companies, which load only on the oil factor. If investor are more overconfident about the tech factor than about the oil factor, they will require a smaller risk premium for holding the tech portfolio than for holding the oil portfolio. Thus, the expected return on the software firms are on average lower than that on the gas companies in the presence of overconfidence.

In Appendix A, I show in a more general setting that the negative dispersion-return relation can be stronger among firms that load heavily on F_1 regardless of their loadings on F_2.

Proposition 4. *Suppose investors are more overconfident about the precision of F_1's signal than that of F_2's (i.e., for factor j, $v_j^C = v^R + \eta_j(v^C - v^R)$, and $\eta_1 > \eta_2$). There are two classes of risky assets whose payoffs follow*

$$\theta_1 = \bar{\theta}_1 + \beta_{1,1} F_1 + \beta_{1,2} F_2 + \epsilon_1, \quad \text{and} \quad \theta_2 = \bar{\theta}_2 + \beta_{2,1} F_1 + \beta_{2,2} F_2 + \epsilon_2. \tag{15}$$

When asset class one loads more heavily on F_1 *than asset class two (i.e.,* $\beta_{1,1} > \beta_{2,1}$*), as overconfidence* v^C *increases, the reduction in the expected return on asset class one is on average greater than that on asset class two if*
(1) $\beta_{1,2} \geq \beta_{2,2}$*; or*
(2) $\beta_{1,2} < \beta_{2,2}$ *and* $\eta_1 < \eta^*$*, where* $\eta^* < \dfrac{v + v^R}{v^C - v^R}$.

Thus, if an econometrician sorts stocks based on their loadings on F_1, she should observe a stronger dispersion-return relationship among stocks with larger F_1 loadings when the sorts create no cross-sectional differences in the F_2 loadings or when F_1 loadings are positively correlated with F_2 loadings. The same pattern can be observed even when F_1 loadings are negatively correlated with F_2 loadings and, meanwhile, the misperception of signal precision, measured by $\eta(v^C - v^R)$, is smaller to certain extent than the true precision of the signal, measured by $v + v^R$.[18] In contrast, the risk hypothesis predicts that only the total risk exposure matters and greater total risk exposure should be associated with a more positive dispersion-return relationship.

Finally, the overreaction hypothesis predicts that the negative dispersion-return relation should be stronger among firms that are more constrained from short selling. Models based on both investor overconfidence and short-sale constraints (Scheinkman and Xiong (2003a), Hong, Scheinkman, and Xiong (2006), and Jiang (2005a)) suggest that assets tend to be overpriced due to the average overestimation of expected cash flows. Jiang (2005a) shows that this overpricing is larger when overconfidence is stronger. As a result, the average cash flow overestimation should further reduce the expected returns in addition to the risk premium reduction effect, and this effect is stronger among firms with tighter short-sale constraints. In contrast, the risk hypothesis provides no prediction with respect to short-sale constraints.

Taken together, I test two sets of competing hypotheses.

Hypothesis I

a. (The overreaction hypothesis) Firm valuation dispersion should be negatively related to subsequent aggregate stock returns.

b. (The risk hypothesis) Firm valuation dispersion should be positively related to subsequent aggregate stock returns.

Hypothesis II

---

[18] This pattern is particularly strong when overconfidence about F_2 is fairly small (i.e., $\eta_2 \to 0$). In this case, the increase in v^C affects the expected return on π_2 only a little but that on π_1 a great deal. Therefore, the reduction in the expected return on asset class one will be greater than that on asset class two.

a. (The overreaction hypothesis) The negative dispersion-return relation should be most pronounced among firms that are risky, that have highly uncertain cash flows, large investor disagreement, or high costs to sell short.

b. (The risk hypothesis) The positive dispersion-return relation should be most pronounced among risky firms. However, there should be no cross-sectional differences in the positive dispersion-return relationship across firms with different cash flow uncertainty, investor disagreement, or short-sale costs when there is no cross-sectional differences in total risk exposure.

III. Data

The main sample includes all available firms listed in NYSE, AMEX, and NASDAQ from January 1963 to December 2004. Robustness checks also use the earlier period 1926–1962. Stock returns and other stock trading data are obtained from the Center of Research in Securities Prices (CRSP). Only common stocks (share code 10 and 11) are used. Stocks in the financial industry (SIC between 6000 and 6999) and the utility industry (SIC between 4900 and 4949) are excluded. Accounting information is obtained from COMPUSTAT. The analyst forecast data from 1976 to 2004 are from I/B/E/S. The institutional ownership data from 1980 to 2004 are from Thomson Financial 13F holdings. Short interest data are available for NYSE firms from 1991 to 2002 and for NASDAQ firms from 1992 to 2002.

A. Key variables

I use three firm valuation ratios: book-to-market equity (BM), dividend-to-price ratio (DP), and earnings-to-price ratio (EP). Firm book equity is calculated following Polk and Sapienza (2005), also consistent with Davis, Fama, and French (2000). Market equity is the product of stock price and shares outstanding. Firm book-to-market equity (BM) is book equity over market equity. Firm dividend yield (DP) is total dividend payout (data 21) over market equity. Firm earnings-to-price ratio (EP) is net income from continuing operation (data 178) over market equity. Zero or negative book equity, dividends, and earnings are excluded to calculate logarithmic ratios.[19]

I form a composite measure of dispersion in firm valuations based on the cross-sectional

---

[19] Excluding the negative values results in a truncated distribution of firm valuation ratios, which might raise a concern about whether the truncation drives the results. As a robustness check, I use sales-to-price ratios to calculate firm valuation dispersion and still observe the negative correlation between dispersion and aggregate returns. Sales are rarely positive and should include a more complete set of firms in my dispersion measure.

standard deviations of the logarithmic firm valuation ratios of all available firms. I denote the cross-sectional standard deviations of logarithmic BM, DP, and EP as σ(BM), σ(DP), and σ(EP) and calculate both value-weighted and equal-weighted standard deviations. The value-weighted measures more accurately reflect the statistical properties of the value-weighted market portfolio while the equal-weighted measures more precisely capture those of the equal-weighted market portfolio.[20]

——INSERT TABLE I, II and FIGURE 1 HERE——

Panels A to C of Figure 1 plot the time-series of the dispersion variables. The value-weighted measures exhibit similar patterns during the sample period. However, the equal-weighted measures have slightly different patterns for a few years. Further, each of the value-weighted measures exhibits a visible non-linear upward time-trend and each of the equal-weighted measures displays a linear upward time-trend. The source of the time-trend will be discussed later.

To avoid the influence of the time-trend on extracting the principal component of the three measures, I take out the trend by regressing each of the value-weighted measures on a time index, which takes the value of 1 to 42, and the squared time index, and use the residuals as the detrended series. I also regress each of the equal-weighted measures on a time index and take the residuals as the detrended series.

——INSERT TABLE III HERE——

Table III reports the regression results. Confirming my observation, the results show that the time trend accounts for a substantial fraction of the time variation in these measures. The adjusted R^2s are 77% to 88% for the value-weighted measures and 32% to 77% for the equal-weighted measures.

The residuals from the above regressions are denoted as *cd*(BM), *cd*(DP), and *cd*(EP). Their time series are plotted in Panels D to F of Figure 1. The summary statistics of these variables in both value-weighted and equal-weighted schemes are reported in Table I. Finally,

---

[20] To understand why, consider the mean book-to-market equity of a portfolio. In a value-weighted portfolio, the total market value of each stock is invested. Thus, the portfolio BM is the sum of book equity of all stocks over the sum of market equity of all stocks. It is also equal to the value-weighted mean of firm book-to-market equity. That is, $(\sum BE)/(\sum ME) = (\sum \frac{BE}{ME} ME)/(\sum ME)$. In an equal-weighted portfolio, one dollar is invested in each stock. Accordingly, each stock only contributes per dollar book equity and per dollar market equity to the portfolio. Thus, the portfolio BM is a simple average of BM of all stocks. That is, $(\sum \frac{BE}{ME})/(\sum \frac{ME}{ME}) = \sum \frac{BE}{ME}/N$, where N is the number of stocks in the portfolio. Similar arguments apply to the standard deviation.

cross-firm valuation dispersion (CVD) is defined as the first principal component of $cd(\text{BM})$, $cd(\text{DP})$, and $cd(\text{EP})$.

The same procedure is performed for the value-weighted measures and the equal-weighted measures separately. The procedures yield two parsimonious measures of the firm valuation dispersion:

$$\text{CVD}_{\text{vw}} = 0.357cd(\text{BM}) + 0.348cd(\text{DP}) + 0.347cd(\text{EP}), \qquad \text{and} \tag{16}$$

$$\text{CVD}_{\text{ew}} = 0.348cd(\text{BM}) + 0.396cd(\text{DP}) + 0.478cd(\text{EP}). \tag{17}$$

The value-weighted CVD explains 90% of the sample variance while the equal-weighted CVD explains 66%. For both measures, the first principal component is the only component with an eigenvalue greater than one and it is highly correlated with the underlying measures. As shown in Table II, the value-weighted CVD has a correlation 0.97 with $cd(\text{BM})$, 0.94 with $cd(\text{DP})$, and 0.94 with $cd(\text{EP})$ while the equal-weighted CVD has the three correlations of 0.69, 0.78, and 0.94. These numbers suggest that CVD captures major information contained in the underlying measures.

The two CVDs have a high correlation of 0.75 with each other. The time series of the value-weighted and the equal-weighted CVD are plotted in Panels G and H of Figure 1. The value-weighted CVD is positive for 1967–75, 1981, 1983, 1986–87, 1988–2001. Similarly, the equal-weighted CVD is positive for 1967–75. However, the equal-weighted CVD behaves somewhat differently after 1980; it is positive for 1980–87, 1990–91, 1996, and 1999–2001. Both time series have a big spike in the post-1997 bubble period and reach their peak in 2000. These high CVD periods mainly coincide with high investor sentiment periods in the anecdotal history collated by Baker and Wurgler (2005), suggesting investor sentiment can be at least partially driven by investor overreaction.

Comparing the two CVDs, the value-weighted CVD experiences little variation before 1997 relative to its big spike around 2000. In contrast, the equal-weighted CVD is more volatile before 1997 relative to its 2000 spike. Unreported plots show that behaviors of the value-weighted measures are dominated by only 5% of the firms.[21] Thus, when I conduct analysis in the cross section, I use the equal-weighted CVD to proxy for common overreaction in individual stocks. However, using the value-weighted CVD would not materially alter the results.

Going back to the time trend in $\sigma(\text{BM})$, $\sigma(\text{DP})$, and $\sigma(\text{EP})$, it is important to understand the source of the time trend to avoid potential data-mining. According to Campbell and

---

[21] These firms on average account for 65% of the total market cap during the sample period and 75% in June of 2000.

Shiller (1988) and Vuolteenaho (2000), BM and DP contain not only a component of future returns but also a component of future growth rate. In particular, BM reflects the growth rate of book equity, the return on equity (ROE), while DP captures the growth rate of dividend payout (ΔD). Accordingly, σ(BM) and σ(DP) should be linked to σ(ROE) and $\sigma(\Delta$D) which denote the cross-sectional standard deviations of ROE and ΔD. Similarly, σ(EP) should be linked to $\sigma(\Delta$E), where ΔE is the growth rate of earnings (see Appendix B for proof).

—–INSERT FIGURE 2 HERE—–

I conjecture that the time-trend in σ(BM), σ(DP), and σ(EP) is due to a time-trend in σ(ROE), $\sigma(\Delta$D), and $\sigma(\Delta$E). Indeed, as shown in Figure 2, the plots of σ(ROE), $\sigma(\Delta$D), and $\sigma(\Delta$E) display a clear upward time trend. Similar to the trend in σ(BM), σ(DP), and σ(EP), the trend in the value-weighted σ(ROE), $\sigma(\Delta$D), and $\sigma(\Delta$E) appears non-linear and that in the equal-weighted measures seems linear. The trend in the dispersion of firm return on equity, dividend growth, and earnings growth may be related to the change in the characteristics of public firms. There is evidence that firms going public have become younger (Fink, Fink, Grullon, and Weston (2005)), and the dispersion of firm growth rates and that of profitability have become larger (Fama and French (2004)).

Because the main focus of the paper is to examine how firm valuation dispersion is related to future expected returns rather than future growth rates, it helps to filter out the noise in my measure by taking out the time trend. Prior research usually filters out the time trend in a time series when examining the serial correlation between this measure and stock returns (e.g., Campbell, Grossman, and Wang (1993)). Further, obtaining CVD in absence of a time trend makes it more convenient to distinguish a state with strong overreaction (when CVD is high) from a state with weak overreaction (when CVD is low) for my subsequent tests.[22]

B. Firm characteristics

My test hypothesis concerns firms that are risky, have highly uncertain cash flows, have large investor disagreement, or have potentially high costs to sell short. I therefore gather variables that are proxies for riskiness, cash flow uncertainty, investor disagreement, and difficulty in selling short. There are ample proxies for each of the properties but probably none is pure and perfect. Hence, my choices depend largely on related prior research.

---

[22] However, detrending the time series using the full sample data might give rise to a concern of the "look-ahead bias" (Brennan and Xia (2005)). I will discuss this issue in detail in the robustness check in Section VI.

I use market beta (BETA) and total return volatility (VOL) as proxies for riskiness. BETA is estimated using a market model on 36 to 60 monthly available returns in the 5 years before July of year t.[23] VOL is the standard deviation of 10 to 12 monthly returns before July of year t. Both are proxies for the true market beta.[24] In addition, idiosyncratic volatility, one component of VOL, has also been linked to investor disagreement (e.g., Harris and Raviv (1993)) and investor overconfidence (e.g., Scheinkman and Xiong (2003b)), both of which suggest that firm valuations are highly subjective when VOL is high. Taken together, VOL is a composite proxy for risk and subjective valuations.

Four firm characteristics are used to measure cash flow uncertainty: firm age (AGE), earnings-to-book equity (E/BE), plant, property, and equipment (PPE), and dividends-to-book equity (D/BE). AGE is the number of years between the firm's first appearance on CRSP and t. E/BE is the ratio of net income from continuing operation over book equity, similar to return on equity. Firms with zero or negative earnings or book equity are excluded. PPE is plant, property, and equipment scaled by total assets. D/BE is total dividends paid scaled by book equity. These variables have been used in prior studies to identify stocks more subject to mispricing (e.g., Baker and Wurgler (2005), Zhang (2006)). I also consider stocks that are in technology-related industries as having high cash flow uncertainty. I use two tech industries based on the Fama-French 10 industry classfication: the computers, software, and electronic equipment (HiTec), and the telephone and television transmission (Telcm).[25]

Two widely used measures of investor disagreement are adopted: analyst forecast dispersion (ADISP) (Diether, Malloy, and Scherbina (2002)) and trading volume (TURN) (Lee and Swaminathan (2000)). Following Diether, Malloy, and Scherbina (2002), ADISP is defined as the standard deviation of analysts' current fiscal year earnings per share forecasts scaled by absolute mean forecasts. Unadjusted data are used as suggested by these authors. Zero mean forecasts are excluded. TURN is measured by turnover—trading volume scaled by shares outstanding. TURN is the average monthly turnover in the past 12 months before July of year t. Trading volume of NASDAQ firms is divided by two due to the double

---

[23] The reported results are based on the betas that are estimated from the time-series regression $R_{i,t} = \alpha + \beta_i R_{MKT,t} + \nu_i$. To account for the effects of nonsynchronous trading, I also estimate betas from the regression $R_{i,t} = \alpha + \beta_{i,1} R_{MKT,t-1} + \beta_{i,2} R_{MKT,t} + \nu_i$ and define $\beta_i = \beta_{i,1} + \beta_{i,2}$ (e.g., Fama and French (1992), Lewellen and Nagel (2005)). The main results still hold.

[24] The theory predicts that firms with high systematic risk are more affected by the risk premium reduction effect of investor overreaction. Thus, BETA seems a more appropriate measure of systematic risk than VOL. However, true market betas are not directly observable. It has also been noted that the regression estimated beta is subject to the error-in-variable problem (Fama and MacBeth (1973)). Thus, the regression estimated beta is only a noisy proxy for the true beta The total risk, VOL, serves as another proxy since it is larger when the true beta is larger.

[25] By excluding the utilities and the others, I used eight of the ten industry portfolios in my analysis.

counting problem in NASDAQ.

Finally, institutional ownership (IO) and short interest ratio (SIR) are used to proxy for shorting difficulty. IO is the total number of shares held by institutional investors scaled by shares outstanding. Only quarterly data is available in the database. SIR is measured by total number of shares shorted scaled by shares outstanding. Monthly data is available in the datafile. For both measures, stocks that have appeared at least once in the data file are included. Zero IO or SIR is assigned to stocks that do not appear in the file at time t. D'Avolio (2002) and Nagel (2005) suggest that stocks with high institutional ownership have more lendable shares since institutions are active stock leaders in the lending market. Therefore, it is relatively easier and cheaper for short-sellers to borrow these stocks when engaging in short-selling.[26] Boehme, Danielsen, and Sorescu (2005) show that short-selling cost, measured by the loan fee, increases with the short interest ratio. Therefore, it is costlier to short stocks with high short interest.

According to the overreaction hypothesis, I expect a stronger negative dispersion-return relation among sets of firms that have large beta, high volatility, low age, few profits, low PPE, low dividends, high analyst forecast dispersion, large trading volume, low institutional ownership, high short interest ratio, or in HiTec or Telcm industries. In contrast, based on the risk hypothesis, I expect a stronger positive dispersion-return relation among sets of firms with large beta or high volatility. However, the risk hypothesis predicts that, controlling for beta, there should be no cross-sectional differences in the dispersion-return relation across the other firm characteristics.

—–INSERT TABLE IV HERE—–

Based on each of the firm characteristics, I sort stocks into quintiles with equal number of firms in each quintile, except for BETA, AGE, IO, and D/BE. I also exclude stocks with prices less than \$5 (e.g., Jegadeesh and Titman (2001)). However, including them does not alter the main results. Prior research finds that high beta stocks tend to be small (Fama and French (1992)) and high IO stocks tend to be large (Nagel (2005)). In addition, in my sample, old firms tend to be large: the average correlation between firm size and AGE is 0.29. To avoid having these quintiles over-represent small firms, I use NYSE breakpoints for the three variables. For D/BE quintiles, firms with non-zero dividends are sorted into four portfolios based on NYSE breakpoints and those with zero dividends are grouped into

---

[26] Phalippou (2004) argues that IO also measures investor sophistication, which suggests that low IO firms tend to be mispriced by unsophisticated investors. This argument carries the same prediction as using IO as a proxy for short-sale constraints.

the quintile with the lowest D/BE. However, changing the breakpoint does not materially change the results.

Except for IO, SIR, and ADISP quintiles, I form portfolios in the end of each June. I then calculate the value-weighted monthly portfolio returns and the 12-month continuously-compounded returns from July of year t to June of year $t+1$. The IO quintiles are formed at the end of each quarter, and the 12-month returns from the beginning of the next quarter are used. The SIR and ADISP quintiles are sorted and updated at the end of each month based on the end-of-month SIR or ADISP, and the overlapping 12-month returns are used.

IV. Empirical Tests

A. CVD and aggregate stock returns

A.1 Sorts

I first test Hypothesis I by examining the correlation between CVD and future aggregate returns. I use the CRSP value-weighted and equal-weighted indices as proxies for the value-weighted and equal-weighted market portfolios. Based on whether the end-of-June CVD is above the mean (high) or below the mean (low), I first sort the following three consecutive 12-month continuously-compounded market excess returns. Shown in Figure 3, during the full sample period, when CVD is low, the average 12-month value-weighted market excess returns in the following three years are about 7–8%. However, when CVD is high, these returns are negative in the subsequent three years, ranging from −0.49% to −1.15%, which is inconsistent with the rational view that risk premium should never be negative. Equal-weighted market excess returns exhibit similar patterns. Following low CVD periods, the annual market excess returns are 10–15% while following high CVD periods, these returns are only 2–6%.

——INSERT FIGURE 3 HERE——

To make sure the poor performance following high CVD period is not solely driven by the post-1997 bubble period, I conduct sorts based on CVD from 1963 to 1996. The return patterns conditioning on CVD remain. Further, during this sample period, following high CVD, the next year's market performance is even worse; the average value-weighted annual return is merely −4.52%. Thus, these results suggest that CVD and future market excess returns are negatively correlated, supporting the overreaction hypothesis.

A.2 Univariate regressions

To formally evaluate the predictive power of CVD, I regress market excess returns on lagged CVD. The value-weighted (equal-weighted) CVD is used to predict value-weighted (equal-weighted) returns. I report results on three different return horizons: one quarter, one year, and three years. The results for other return horizons from 1 month to 3 years are qualitatively similar. For one-quarter returns, the predictors are updated at the end of each March, June, September, and December. For one-year or three-year returns, they are updated annually at the end of each June. Overlapping observations are used for the three-year return regressions.

Stambaugh (1986, 1999) points out that the OLS estimator in a predictive regression will be biased to favor the alternative in a small sample if (1) the regressor is highly persistent and (2) the innovation of the predicted variable and that of the forecaster are correlated. Shown in Table I, the first-order autocorrelation of CVD is 0.58 in the value-weighted scheme and 0.44 in the equal-weighted scheme, suggesting CVD is relatively persistent. Furthermore, preliminary analyses find that shocks to the expected market returns are correlated with shocks to CVD. Hence, the OLS estimator is likely to be biased, as are the OLS t-statistic and p-value.

To adjust for this bias, I simulate the empirical distribution of the predictive slope using the randomization method proposed by Nelson and Kim (1993), which has become one of the standard methods for such an adjustment. For each regression, under the null that there is no correlation between the predictor and the future returns, I obtain the distribution of the predictive slope with 2500 sets of simulated data. Then based on this distribution, I calculate the p-value of the estimated predictive slopes. This method provides a p-value that is free of both the bias in the OLS estimator derived from the regressor autocorrelation and the bias in the OLS standard error derived from overlapping observations. Throughout the remaining analyses, I report the simulated p-value for all predictive regressions.

—–INSERT TABLE V HERE—–

Table V reports the results of the predictive regression. Panel A reports the results with the full sample while Panel B reports the results after excluding the post-1997 period. As can be seen, the coefficients of CVD are all negative and significant at the 5% level.

The economic magnitude of the predictability using CVD is large. For the value-weighted market excess returns, in the full sample a one standard deviation positive shock to CVD results in a 2.38% reduction in future 1-quarter returns, a 8.34% reduction in future one-year returns and a 18.97% reduction in future three-year returns.[27] The three coefficients of CVD are 2.58%, 8.76% and 12.74% for the equal-weighted market excess returns.

Further, CVD alone explains 6% of one-quarter ahead value-weighted return variation, which is getting close to 9% using the consumption, wealth, and income ratio (*cay*) and much greater than 1% using dividend yield to forecast one-quarter ahead S&P Composite Index returns from 1952 to 1998 (Lettau and Ludvigson (2001)). In forecasting one-year ahead value-weighted returns, CVD produces a R^2 of 22%, larger than 18% using *cay* for the period 1952–1998 (Lettau and Ludvigson (2001)), 8% using dividend yields and 11% using aggregate relative equity issuances for the period 1963–1997 (Baker and Wurgler (2000)). Finally, CVD explains 51% of the value-weighted return variation, larger than 45% using dividend yield for the period 1957–1986 (Fama and French (1988a)).

Excluding the post-1997 period yields similar results. As reported in Panel B of Table V, all coefficients remain negative and highly significant, and furthermore, the magnitude of the coefficients becomes greater. For example, the coefficient of CVD is -16.60 in forecasting value-weighted one-year ahead returns as opposed to -8.34 with the inclusion of the post-1997 period. The change in coefficients is likely due to the relatively small variance of CVD after excluding the post-1997 period, during which CVD has a huge spike that increases its variance in the full sample.

In sum, the tests of Hypothesis I provide results that are consistent with the overreaction hypothesis that CVD is negatively related to future market excess returns and are inconsistent with the risk hypothesis which predicts a positive correlation between CVD and future aggregate returns.

A.3 Multivariate regression

Previous studies have identified a number of predictors of market returns. Some of the well-known ones include dividend yield (D/P) (Campbell and Shiller (1988), Fama and French (1988a), Lewellen (2004)), short-term interest rate (Fama and Schwert (1977)), past market returns (Fama and French (1988b)), default and term premium (Fama and French (1989)), and aggregate relative equity issuances (Baker and Wurgler (2000)). Although the main focus of this paper is not to propose a new aggregate return predictor, it is still interesting to examine whether CVD provides incremental power to forecast future aggregate returns.

Aggregate dividend yield (D/P) is computed following Fama and French (1988a). Annual dividends are used to avoid seasonal differences in dividend payments. Short term interest

---

[27] Since CVD has unit variance, the coefficients measure the effect of a one standard deviation shock to CVD on expected returns.

rate (TBILL) is measured by the one-month treasury bill rate obtained from Ibbotson Associates. Past market return (Rm) is the past return on the value-weighted CRSP index. Following Fama and French (1988b), past six-month, one-year, and three-year market returns are used, respectively, to forecast market returns six months, one year, and three years ahead. Aggregate relative equity issuance (S) is obtained from Wurgler's website over the period 1962–2002. Term premium (TERM) is defined as the difference between the 10-year treasury constant maturity rate and the one-year treasury constant maturity rate. Default premium (DEF) is defined as the difference between the Moody's seasoned Aaa corporate bond yield and the Moody's seasoned Baa corporate bond yield. The yield data are obtained from the Federal Reserve System website.[28]

The correlations of these predictors and CVD are reported in Panel C of Table II. The value-weighted CVD has a significant negative correlation with D/P, TERM, and DEF and the equal-weighted CVD has a significant positive correlation with TBILL. The correlations between CVD and business cycle proxies are not surprising. As shown in the model, CVD is affected by expected market returns, which are influenced by business cycle and economic conditions. However, if the variation in CVD is solely determined by business cycle and economic conditions, we would expect to observe a positive correlation between CVD and future market returns.

—–INSERT TABLE III and VI HERE—–

Table VI reports the results for forecasting market excess returns using all of the above predictors. I run two regressions for each of the forecast horizons. The first includes only the well-known predictors. The second adds CVD to the first regression.

In forecasting the value-weighted market excess returns, CVD remains significantly negative when being added to the set of standard predictors. The magnitude of the coefficients remains similar as used alone in the regression. CVD also appears to add incremental power in forecasting future value-weighted equity premium. It triples the R^2 (from 2% to 6%) in forecasting one-quarter-ahead returns, doubles the R^2 (from 8% to 16%) in forecasting one-year-ahead returns, and more than doubles the R^2 (from 24% to 60%) in forecasting three-year-ahead returns.

When it is added to forecast the equal-weighted market excess returns, CVD also remains negative and statistically significant in forecasting returns one-quarter or three-year ahead. Adding CVD mildly increases the forecasting power for one-quarter ahead returns and

---

[28] In a robustness check, I also use the term premium and the default premium obtained from Ibbotson Associates over the period 1963–1999. The main results remain unchanged.

three-year ahead returns. However, it does not seem to add incremental power in forecasting one-year ahead returns. In general, the results with both value-weighted and equal-weighted schemes suggest that CVD provides incremental information about expected market excess returns over and above these well-known predictors, particularly for the value-weighted returns.

B.2 Alternative measures of dispersion

As discussed in the introduction, a few studies have proposed several measures of cross-sectional dispersion of firm BM ratios, including the value spread (Cohen, Polk, and Vuolteenaho (2003)), the book-to-market (BM) spread, and the market-to-book (MB) spread (Liu and Zhang (2005)). The value spread is the difference between the logarithmic BM ratios of value firms and growth firms while the small value spread is defined only to include small firms. The BM (MB) spread is the difference of book-to-market (market-to-book) equity between the two classes of stocks. My model has shown that the key difference between CVD and the value spread is that the former helps to distinguish competing hypothesis while the latter does not. However, it is still interesting to compare the forecasting power of CVD with that of the value spread as well as of the other spreads.

—–INSERT TABLE VII HERE—–

I regress market excess returns on the BM spread, the MB spread, the value spread, and the small value spread with and without CVD and report the results in Table VII. In calculating the BM spread, the MB spread, and the value spread, I follow Liu and Zhang (2005) to define value firms and growth firms as the top and the bottom BM deciles. In calculating the small value spread, I follow Campbell and Vuolteenaho (2004) to define value firms, growth firms, and small firms based on the Fama-French 2 × 3 size-BM portfolios. Only one-year and three-year returns are used to run the predictive regressions.

The results in Table VII show that CVD is a stronger predictor than the other measures during my sample period. When used alone, both the BM spread and the MB spread have some forecasting power on equal-weighted 1 or 3 year returns but not on value-weighted 1 or 3 year returns over my sample period, which is consistent with the findings in Liu and Zhang (2005) during the period 1945–2001. The value spread alone exhibits no power to forecast future 1 or 3 year returns regardless of the weighting schemes of the market portfolios, also consistent with the findings in Liu and Zhang. When used alone, the small value spread forecasts the value-weighted returns but not the equal-weighted ones. When

CVD is added, only the BM spread remains significant with the expected signs in forecasting equal-weighted returns. The coefficients on other measures are either insignificant or switch to the unexpected sign. In contrast, CVD remains negative and highly significant when competing with each of the above valuation spreads. Therefore, during my sample period, CVD is a much more powerful predictor of aggregate returns than the other measures.

One possible reason for the empirical success of CVD over the value spread is that CVD is free of the noise evident in the time-trend. Another possible reason might be that CVD captures information from dispersion in BM as well as dispersion in DP and in EP. Consistently, I find that, after taking out the time trend in the value spread, the value spread alone becomes a significant predictor. However, the predictive power of the value spread is still subsumed by CVD when forecasting three-year returns. For brevity, these results are not reported.

C. CVD and the cross section of stock return

C.1 Two-way sorts

To test Hypothesis II about the relationship between CVD and the cross section of stock returns, I first sort the industry portfolios or firm characteristic quintiles based on the beginning-of-period CVD into two states: in one state, CVD is relatively high (CVD > 0) while in the other, CVD is relatively low (CVD < 0). I then examine the return patterns across quintiles in the cross section and across the two states. From now on, I focus on the overreaction hypothesis due to the rejection of the risk hypothesis at the aggregate level. For reasons discussed previously, I use the equal-weighted CVD to proxy for aggregate overreaction. CVD measured at the end of June of year t is used to conduct sorts from June of year t to May of year $t+1$.

——INSERT TABLE VIII, FIGURE 4 HERE——

Table VIII reports average annual returns on each portfolio in the high CVD state, in the low CVD state, and the return differentials between the two states. These returns are also plotted in Figure 4. The results show a few interesting patterns.

First, when CVD is high—presumably when the beginning-of-period overreaction is high—the difficult-to-value/difficult-to-short firms substantially underperform their counterparts. These stocks include those in the tech industries, that have large beta, high volatility, low age, low profits, few fixed assets, low dividends, low institutional ownership, high short interest ratio, high analyst forecast dispersion, or high volume. Specifically, when CVD is above the mean, the HiTec industry on average earns −2.42% per year and the Telcm industry earns −1.10% per year, the lowest two among all industries. The return differentials between the two extreme quintiles of risk or other characteristics are ranging from 3% to 19% per year. For example, the average annual return is −2.83% for the highest BETA quintile in contrast to 4.26% for the lowest BETA quintile. The average annual return is −2.37% for the lowest AGE quintile, 1.34% for the lowest IO quintile, and a mere −6.65% for the highest SIR quintile, in contrast to 2.64%, 5.25%, and 11.69% respectively for the other extreme quintiles.

In addition, following high dispersion periods, a number of portfolios on average earn negative excess returns. Interestingly, these portfolios tend to include stocks that are conceivably riskier than others, e.g., stocks with the largest beta, highest volatility, lowest age, or highest short interest ratio. These findings are at odds from the rational perspective that risk premium should be positive.

Second, when CVD is low—presumably when the beginning-of-period overreaction is low—the characteristic-return pattern is essentially reversed. The HiTec industry offers an annual return 12.63% and the Telcm industry provides 8.72% per year, the highest two among all industries. The highest BETA quintile earns 13.42% per year, 7.02% higher than the lowest BETA quintile. The lowest AGE quintile earns 15.90% per year, 6.41% higher than the highest AGE quintile. Similar patterns are observed for VOL, E/BE, PPE, D/BE, IO, SIR, and TURN. The highest ADISP quintile still underperforms its counterpart, but the return differential is reduced by more than two-thirds.

Third, the return differential between the high CVD state and the low CVD state is largest for these difficult-to-value/difficult-to-short firms, suggesting their returns are most influenced by CVD. For example, this differential is 15.05% for the HiTec industry and 9.82% for the Telcm industry, the largest two among the industry portfolios. It is 16.25% for the highest BETA quintile in contrast to 2.14% for the lowest BETA quintile, 15.43% for the lowest PPE quintile in contrast to 3.48% for the highest PPE quintile, and 20.55% for the highest SIR quintile in contrast to 7.61% for the lowest SIR quintile. Similar patterns apply for the remaining characteristics.

In sum, the two-way sorts support the overreaction hypothesis. When the beginning-of-period overreaction is high—when there is large forecast error and low risk premium—stocks with highly subjective valuation and significant limits to arbitrage substantially underperform. This evidence suggest these stocks are average overpriced during the high CVD periods.[29] Intriguingly, when conditioning on low beginning-of-period CVD, consistent with the traditional asset pricing model, these presumable riskier firms overperform. These patterns are reasonable under my model: when overreaction is large, stock returns are more determined by irrational beliefs and hence look anomalous; however, when overreaction is small, stock returns reflect more rational views and the risk-return tradeoff. Hence, my findings suggest that the anomalous stock return patterns are concentrated in periods when investor overreaction to market-wide information is greater.

C.2 CVD and characteristic portfolios

The sorts provide evidence that CVD has a stronger relationship with returns of stocks whose valuations are highly subjective and difficult to arbitrage. An alternative approach to examine the dispersion-return relation in the cross section is through regressions. Specifically, I regress returns of each of the characteristic portfolios on CVD and study the patterns of the CVD coefficients and R^2s across the portfolios.

———INSERT TABLE IX HERE———

Table IX reports the results. I report the p-value for each of the industry portfolios and for the hedging portfolios that are long on the lowest characteristic quintiles and short on the highest characteristic quintiles.

Consistent with the sorts, the coefficient of CVD is in general the most negative for the difficult-to-value/difficult-to-short stocks, suggesting a stronger dispersion-return relationship. The HiTec industry has a predictive slope of -14.14 and the Telcm industry has a slope of -6.20, both significant at the 5% level and also the two most negative slopes among the industry portfolios. The dispersion-return relation is uniformly negative for all characteristic quintiles. As we move from quintiles with the above characteristics to the other extreme, the coefficients become less negative. For example, the coefficient is -20.08 for the highest VOL quintile and it goes up to -2.47 for the lowest VOL quintile. Further, the R^2s are uniformly larger for the quintiles more subject to investor overreaction. For instance, the R^2 is 22% for the highest BETA quintile but only -1% for the lowest BETA quintile.

To test the statistical significance of the difference between the CVD coefficients of the two extreme quintiles, I report the p-value of the CVD coefficient of the long-short portfolio,

---

[29] In particular, this evidence suggests that the market and the unobserved factors on average earn negative returns during these periods. This can be explained by the findings that overconfidence tends to be greater in up markets (e.g., Gervais and Odean (2001), Statman, Thorley, and Vorkink (2006)). Therefore, aggregate expected cash flows tend to be overestimated more than underestimated, which results in a negative average aggregate return.

which is long on the top quintile and short on the bottom quintile of each firm variable. The predictive slopes of the long-short portfolios are all significant at the 5% level. A one standard deviation shock to CVD corresponds to a 2–18% change in annual expected returns of the long-short portfolios. Thus, the results suggest that the difficult-to-value/difficult-to-short stocks are most subject to aggregate investor overreaction.

Taken together, the empirical evidence is consistent with the overreaction hypotheses. Firm valuation dispersion is negatively related to subsequent aggregate returns and portfolio returns. Further, the negative dispersion-return relation is stronger among firms that have highly subjective valuations and significant limits to arbitrage.

V. Alternative explanations

A. The cross section of returns and multifactor models

A.1 Controlling for common comovement

Although CVD is shown to distinctively influence firms with opposite characteristics, it is possible that this difference is only due to a difference in firms' exposure to systematic risks. In other words, CVD might be a pure proxy for time-varying risk premia. Therefore, such tests need to control for return comovement with common risk factors. As a result of rejecting the CAPM (and the market model) in the previous test, I obtain common risk factors from available multifactor models.

I choose three multifactor models: the Fama-French 3-factor model (Fama and French (1993)), the 4-factor model (Carhart (1997)), and the ICAPM of Brennan, Wang, and Xia (2004). The first two models are empirically motivated while the last one stems from theory. With respect to the 3- or the 4-factor models, there has been a debate on whether the size factor (SMB), the book-to-market factor (HML), and the momentum factor (UMD) proxy for risk or mispricing (e.g., Fama and French (1995), Daniel and Titman (1997), and Daniel, Hirshleifer, and Subrahmanyam (2005)). Therefore, these factors, if related to mispricing, can possibly capture not only some of the risk reduction effect of overreaction but also some of the cash flow mean bias effect. Nonetheless, as long as CVD captures more effects of investor overreaction over these factors, the lagged CVD can possibly continue to forecast returns of the long-short portfolios even after controlling for these contemporaneous factors. In that case, these kinds of tests can at least distinguish whether CVD picks up novel effects beyond the well-known comovement in stock returns.

Specifically, I run each of the following time series regressions for the two tech industries and for each of the long-short risk or characteristic portfolios.

$$R_t = c + d\mathrm{CVD_{t-1}} + \beta\mathrm{MKT_t} + s\mathrm{SMB_t} + h\mathrm{HML_t} + u_t, \tag{18}$$

$$R_t = c + d\mathrm{CVD_{t-1}} + \beta\mathrm{MKT_t} + s\mathrm{SMB_t} + h\mathrm{HML_t} + m\mathrm{UMD_t} + u_t, \tag{19}$$

$$R_t = c + d\mathrm{CVD_{t-1}} + \beta\mathrm{MKT_t} + b\Delta\gamma_\mathrm{t} + e\Delta\eta_\mathrm{t} + u_t, \tag{20}$$

where MKT is the market excess returns, $\Delta\gamma$ is the estimated innovation in the instantaneous real interest rate, and $\Delta\eta$ is the estimated innovation in the instantaneous market Sharpe ratio. Regression (18) controls for the FF 3 factors, Regression (19) controls for the 4 factors, and Regression (20) controls for the innovations in the two state variables in addition to the market factor in the BWX ICAPM. The time series from 1963 to 2004 of the FF 3 factors and the momentum factor are obtained from French's website. The time series of both $\Delta\gamma$ and $\Delta\eta$ are obtained from Yihong Xia's website, which is available until the end of 2001.

—–INSERT TABLE X HERE—–

Table X reports the coefficients of CVD and the corresponding *p*-values in the three specifications. Annual excess returns are used as dependent variables. All of the factor returns are converted to annualized continuously-compounded returns. In Table X, the three specifications yield similar results. All coefficients remain in the expected direction except for the SIR long-short portfolio when controlling for the BWX model. Most coefficients remain significant at the 10% level with a majority significant at the 5% level, though it is insignificant for the ADISP, IO, and SIR long-short portfolios after controlling for BWX ICAPM. The magnitude of the coefficients is generally reduced from the univariate regressions but remains economically significant. Further, in unreported robustness check, I find that such predictability largely remains even after I allow for time-varying factor loadings that change with CVD.

Overall, CVD appears to be a stronge predictor of the returns on the long-short characteristic portfolios even after controlling for the well-known comovement in stock returns. Therefore, the cross-sectional differences in the predictability of portfolio returns using CVD are not fully explained by the cross-sectional differences in portfolios' exposure to a set of risk factors.

B. Aggregate returns and multifactor models

To explore the possibility that a rational multifactor pricing model can explain the negative correlation between CVD and aggregate returns, I develop a framework to link firm valuation dispersion to market excess returns through these asset pricing models. In addition, so far I have assumed that the fundamental-to-price ratios are only proxies for expected returns. However, it is known that they are also proxies for growth rates of firm fundamentals (e.g. Campbell and Shiller (1988), Vuolteenaho (2000)), whose impacts on my results have not been studied. Thus, another advantage of this framework is to take into consideration the effects of growth rates.

I start with the DP decomposition by Campbell and Shiller (1988), who provide the following log linear approximation for the log dividend yield:

$$d_{i,t-1} - p_{i,t-1} \approx \sum_{j=0}^{\infty} \bar{\rho}^j r_{i,t+j} - \sum_{j=0}^{\infty} \bar{\rho}^j \Delta d_{i,t+j} + c_{t-1}, \tag{21}$$

where d is the log dividend, p is the log price, r is the stock return, Δd is the log dividend growth rate, c is a constant, and ρ is the ratio of the ex-dividend to the cum-dividend stock price, which takes a typical value 0.97 (Polk, Thompson, and Vuolteenaho (2004)).

Equation (21) states that the logarithmic dividend-to-price ratio is the difference between discounted stock returns and discounted dividend growth rates plus a constant. This equation holds both *ex ante* and *ex post*. It is similar to the Gordon (1962) model: $DP = r - g$. However, it allows for time-varying returns and time-varying dividend growth rates.

I first derive this framework with a single factor and then generalize it to incorporate multiple factors. Polk, Thompson, and Vuolteenaho (2004) show that, if the Sharpe-Lintner CAPM holds,

$$d_{i,t-1} - p_{i,t-1} \approx \beta_i \sum_{j=0}^{\infty} \bar{\rho}^j (R_{m,t+j} - R_{f,t+j}) - \sum_{j=0}^{\infty} \bar{\rho}^j \Delta d_{i,t+j} + c'_{t-1}, \tag{22}$$

where R_m is the market returns, R_f is the riskfree rate, and c' is the sum of discounted riskfree rates plus c. Following these authors, I assume that the equity premium $R_m - R_f$ follows a first-order autoregressive process with the coefficient of γ_m. I additionally assume that the logarithmic dividend growth also follows an AR(1) with a coefficient of γ_d. Denoting the equity premium MKT$= R_{m,t} - R_{f,t}$, I can rewrite equation (22) as

$$d_{i,t-1} - p_{i,t-1} \approx \left(\frac{\beta_i}{1-\rho\gamma_m}\right) \mathrm{MKT} - \left(\frac{1}{1-\rho\gamma_d}\right) \Delta d_{i,t} + c'_{t-1}.$$

I then take cross-sectional variance of both sides and obtain

$$\sigma^2(d_{t-1} - p_{t-1}) = \frac{\sigma^2(\beta)}{(1-\rho\gamma_m)^2}\mathrm{MKT}^2 - \frac{cov(\beta, \Delta d_t)}{(1-\rho\gamma_m)(1-\rho\gamma_d)}\mathrm{MKT} + \frac{\sigma^2(\Delta d_t)}{(1-\rho\gamma_d)^2}.$$

The above equation shows that cross-sectional variance of the logarithmic dividend-to-price ratio is a function of MKT^2, MKT, and $\sigma^2(\Delta d_t)$. If the terms with β are constant over time, the above equation can be written as

$$\sigma^2(d_{t-1} - p_{t-1}) = b\mathrm{MKT}^2 + c\mathrm{MKT} + d\sigma^2(\Delta d_t), \tag{23}$$

$$\text{where} \quad b = \frac{\sigma^2(\beta)}{(1-\rho\gamma_m)^2}, \quad c = -\frac{cov(\beta, \Delta d_t)}{(1-\rho\gamma_m)(1-\rho\gamma_d)}, \quad \text{and} \quad d = \frac{1}{(1-\rho\gamma_d)^2}.$$

Equation (23) implies that there should be a relation between $\sigma(d{-}p)$ (the same as $\sigma(\mathrm{DP})$) and the equity premium MKT. However, this relationship is non-linear. In particular, b and d should be positive, and the sign of c is the opposite sign of $cov(\beta, \Delta d_t)$.

Similar to the DP decomposition, Vuolteenaho (2000, 2003) provides the following log linear approximation for BM:

$$b_{i,t-1} - p_{i,t-1} \approx \sum_{j=0}^{\infty} \bar{\rho}^j r_{i,t+j} - \sum_{j=0}^{\infty} \bar{\rho}^j roe_{i,t+j} + c_{t-1}, \tag{24}$$

where b is the log book equity and roe is the log return on equity. This equation states that the logarithmic book-to-price ratio is the difference between discounted returns and discounted return on equity plus a constant.

A similar procedure as for DP can lead to the following equation

$$\sigma^2(b_{t-1} - p_{t-1}) = b\mathrm{MKT}^2 + c\mathrm{MKT} + d\sigma^2(roe_t), \tag{25}$$

$$\text{where} \quad b = \frac{\sigma^2(\beta)}{(1-\rho\gamma_m)^2}, \quad c = -\frac{cov(\beta, roe_t)}{(1-\rho\gamma_m)(1-\rho\gamma_d)}, \quad \text{and} \quad d = \frac{1}{(1-\rho\gamma_d)^2}.$$

The above equation implies that $\sigma(d-p)$ (the same as $\sigma(\mathrm{BM})$) should be related to the equity premium MKT through a non-linear relationship.

If, instead of assuming CAPM holds, the FF 3 factor model holds for individual stock returns, I can derive the following relationship between $\sigma(\mathrm{DP})$ and the three factors.

$$\sigma^2(d_{t-1}-p_{t-1}) = b_m\mathrm{MKT}^2+b_s\mathrm{SMB}^2+b_h\mathrm{HML}^2+c_m\mathrm{MKT}+c_s\mathrm{SMB}+c_h\mathrm{HML}+d\sigma^2(\Delta d_t), \tag{26}$$

The definition of each term is in Appendix C. Similar to the CAPM case, b_m, b_s, b_h, and d should be positive but the signs of c_m, c_s, and c_h are unknown.

This framework can be easily extended to incorporate not only the FF 3 factors, but also the 4 factors and the BWX ICAPM factors (see Appendix C for details). For all of these models, the framework implies a nonlinear relationship between $\sigma(\mathrm{BM})$, $\sigma(\mathrm{DP})$, $\sigma(\mathrm{EP})$ and market excess returns. This nonlinear relationship may be approximated by a negative linear relation. Therefore, it remains a possible mechanism to generate the negative relation between firm valuation dispersion and subsequent market excess returns.

———INSERT TABLE XI HERE———

The specifications are listed in Table XI. I regress σ(BM), σ(DP), σ(EP) on a set of common factor returns, squared common factor returns, and the cross-sectional variance of their corresponding growth rates. These common factors always include the market factor. I only report the results for the value-weighted variances. Using equal-weighted variances offers similar results. To test this nonlinear relationship, I test whether the coefficient of the squared market factor (MKT^2) is positive, as predicted by this framework.

The results presented in Table XI unanimously reject the hypothesis. None of the coefficients of MKT^2 is positive. In fact, all of them are indistinguishable from zero. Interestingly, the coefficients of the squared HML and the squared UMD appear to be significantly positive for some specifications. Another interesting finding is that σ(BM), σ(DP), and σ(EP) are positively related to σ(ROE), $\sigma(\Delta \mathrm{D})$, and $\sigma(\Delta \mathrm{E})$, respectively. This result further confirms that dispersion in firm valuation ratios is linked to dispersion in growth rates of firm fundamentals. Finally, the R^2s of all of the specifications are very large, ranging from 75% to 95%, suggesting that this framework explains a major variation in the firm valuation dispersion. A substantial fraction of the R^2s is attributed to the time trend shared by the dispersion of valuation ratios and the dispersion of growth rates.

To summarize, the regression results reject the nonlinear relationship between cross-sectional dispersion of BM, DP and EP and future market excess returns. Therefore, the negative relation between the firm valuation dispersion and future market excess returns is less likely to be explained by the CAPM, the FF 3-factor model, the 4-factor model, or the BWX ICAPM.

VI. Robustness

In assessing the robustness of the empirical findings, I focus on three issues. First, I investigate the predictability of aggregate returns using CVD in the earlier period 1926-1962. Second, I examine the ability of the dispersion measures constructed based on only prior information to forecast aggregate returns. Third, I evaluate the sensitivity of the results by considering the change in portfolio composition.

First, throughout the above tests, I have focused on the post-1963 period. The primary reason is due to the unavailability of the accounting data to conduct cross-sectional analyses before 1963. Fortunately, data is available for the time series analysis during this earlier period. Thus, it is interesting to check whether firm valuation dispersion is negatively related to subsequent market returns during the period 1926–1962.

I calculate both the value-weighted and the equal-weighted σ(BM) and σ(DP) during this period. Firm book equity data is obtained from French's website. Firm dividend payout is constructed from CRSP stock returns with and without distribution (e.g., Fama and French (1988a)). Similar to the time series of the dispersion variables after 1963, I find a non-linear time trend in the value-weighted measures and a linear time trend in the equal-weighted measures. Thus, I detrend these variables accordingly through regressions. The detrended variables are *cd*(BM) and *cd*(DP). Finally, CVD is defined as the principal component of *cd*(BM) and *cd*(DP) during this period.

———INSERT TABLE XII HERE———

I report in Table XII the regression of future one-year and three-year market excess returns on CVD, *cd*(BM), or *cd*(DP) from 1926 to 1962. CVD exhibits strong negative correlations with the subsequent value-weighted market excess returns. The coefficients are respectively -13.30 and -29.78, both significant at the 1% level. The R^2 is 16% for the one-year return regression and 39% for the three-year return regression. Similar results apply to both *cd*(BM) and *cd*(DP). However, in forecasting the equal-weighted market excess returns, the coefficients on CVD are negative but insignificant. Panel B shows that the weak negative correlation is due to a lack of relationship between *cd*(BM) and future equal-weighted returns: the coefficients on *cd*(BM) are positive and insignificant. In contrast, the coefficients on *cd*(DP) are all negative and highly significant. Overall, the results show that firm valuation dispersion is a negative predictor of subsequent market excess returns over 1926–1962, particularly for the value-weighted returns.

Second, in the above tests, CVD is constructed from residuals by regressing the full-sample firm valuation dispersion on a time index. This might raise a concern about "look-ahead bias" (Brennan and Xia (2005)), which refers to strong forecasting power due to "ex post successfully fitting of the trend within the sample." To address this concern, I define CVD_{raw} as the first principal component of σ(BM), σ(DP), and σ(EP). Then I regress market excess returns on lagged CVD_{raw} together with a time index (t-index) or/and the squared time index (t^2-index). This method requires only prior information at each point of time to forecast returns in the sample.

———INSERT TABLE XIII HERE———

The results of forecasting one-year and three-year return are reported in Table XIII. The coefficients on CVD_{raw} are all negative and highly significant, except that it is insignificant when used to forecast one-year-head returns by only controlling for a linear time trend. Confirming the existence of a non-linear time trend in the value-weighted CVD, the coefficients on t-index and t^2-index are all statistically significant. Also consistent with a linear time trend in the equal-weighted CVD, t-index is significant when it is used as the only control; and the coefficients on both t-index and t^2-index are insignificant when they are used together as controls for the time trend. Finally, R^2s are also similar to the regressions that CVD is used as a sole independent variable. Thus, the results show that the forecasting power of CVD is not driven by a potential look-ahead bias.

Finally, I explore whether my results are driven by changes in firm composition instead of by changes in firm valuations. Firm valuation dispersion can become larger when the existing firms' valuations are more dispersed. Alternatively, it can be due to newly listed firms that tend to have extreme valuations. To examine whether the negative relation between CVD and subsequent aggregate returns holds for existing firms, I collect a group of firms that appear in my sample every year from 1965 to 2004. I then calculate CVD for this group of firms and find that CVD still negatively and significantly forecasts subsequent returns of this portfolio. Further, when I limit the sample to the largest size quintile, the largest age quintile, or the highest BM quintiles, I still observe the negative correlations between CVD and subsequent portfolio returns. Since firm valuation dispersion is calculated using firms only in each of those groups, it is less likely to be influenced by newly-listed firms. Therefore, the inclusion of newly-listed firms in my measure should not be responsible for my empirical results.

VII. Summary and Conclusion

In this paper, I present one model based on investor psychology and one based on investor full rationality, both of which predict a positive correlation between fundamental-price ratios and expected aggregate returns and a negative correlation between the value spread and expected aggregate returns. However, the two models deliver different predictions with respect to the correlation between cross-sectional dispersion of firm valuations and expected market returns.

The model based on investor overconfidence and overreaction predicts that higher firm valuation dispersion indicates greater overreaction and lower expected aggregate returns. In contrast, the model with fully rational investors predicts that higher firm valuation dispersion suggests greater market risk premium and higher expected aggregate returns. Thus, the two distinct predictions provide an interesting setting for testing the overreaction hypothesis against the risk hypothesis. In contrast, the aggregate fundamental-price ratios and the value spread do not help to distinguish the two theories.

I find that, supporting the overreaction hypothesis, during the period 1926–2004, my measure of cross-sectional dispersion of firm valuations is a negative predictor of subsequent market and portfolio excess returns up to three years ahead. Further, such predictability is particularly pronounced among firms whose valuations are highly subjective and for which arbitrage is considerably limited, including those that have large beta, high volatility, low age, low profits, few fixed assets, low dividends, low institutional ownership, high short interest ratio, high analyst forecast dispersion, large trading volume, or that are in the technology-related industries.

Most importantly, I show that the underperformance of these presumably riskier stocks, which has long plagued traditional asset pricing models, only becomes visible in the periods when the beginning-of-period firm valuation dispersion is high—when overreaction is large. When the beginning-of-period dispersion is relatively low, indicating small overreaction, the traditional norm seems to hold: investors get rewarded positively for holding risky positions. The most interesting finding is that following high dispersion periods, firms that are conceptually riskier, such as those with high beta, high volatility, and low age, on average earn negative excess returns, which challenges the explanation that firm valuation dispersion solely captures time-varying risk.

My results show that investor overreaction to market-wide information has systematic effects on both aggregate market returns and the cross section of stock returns. It suggests that investor overreaction is important for understanding asset pricing. In addition, through variation in aggregate overreaction, we can obtain a better understanding about waves in corporate behaviors and policies.

Appendix A: Proofs

Proof of Proposition 1

Since $\sigma(\beta)$ is a constant, for brevity, I assume that it is equal to one in all proofs. Thus,

$$\begin{aligned}E[\hat{\sigma}(C-P)] &= \int_{-\infty}^{S^*}(\pi - M)f(S)dS + \int_{S^*}^{+\infty}(M-\pi)f(S)dS \\ &= \int_{-\infty}^{S^*}\pi f(S)dS - \int_{S^*}^{+\infty}\pi f(S)dS + \int_{S^*}^{+\infty}Mf(S)dS - \int_{-\infty}^{S^*}Mf(S)dS,\end{aligned}$$

where $S^* = \dfrac{A(v+v^R)Q}{v(v^C-v^R)}$ and $M(S^*) = \pi$. When S is above S^* or below $-S^*$, $|M| > \pi$, the mean bias effect dominates the risk premium reduction effect in determining firm valuation dispersion. Conversely, when S is between $-S^*$ and S^*, $\pi > |M|$, the risk premium reduction effect dominates.

Further, let $\omega = 1/(v+v^C)$, and $M^* = M(S^*)$, taking derivative with respect to v^C yields

$$\begin{aligned}\frac{\partial E[\hat{\sigma}(C-P)]}{\partial v^C} &= -\omega\pi\left[2F(S^*)-1\right] - 2\frac{\partial S^*}{\partial v^C}\pi f(S^*) + 2\frac{\partial S^*}{\partial v^C}M^* f(S^*) \\ &\quad + \int_{S^*}^{+\infty} v\omega^2 Sf(S)dS - \int_{-\infty}^{S^*} v\omega^2 Sf(S)dS \\ &= v\omega^2\left\{\int_{S^*}^{+\infty} Sf(S)dS - \int_{-\infty}^{S^*} Sf(S)dS - \frac{AQ}{v}\left[2F(S^*)-1\right]\right\}\end{aligned}$$

If this derivative is greater than zero, then that in brace must be positive since $v\omega^2$ is positive. All else equal, this inequality becomes more likely to hold when overconfidence is strong.

To see that this derivative is positive when v^C is large enough, let us denote the component in brace as Ω, and taking derivative on Ω with respect to v^C yields

$$\frac{\partial\Omega}{\partial v^C} = -2f(S^*)\frac{\partial S^*}{\partial v^C}\left(S^* + \frac{AQ}{v}\right) > 0,$$

since $\dfrac{\partial S^*}{\partial v^C} < 0$. Thus, Ω becomes larger when v^C is larger. Consider two extreme cases. In Case 1, overconfidence is extremely low such that v^C approaches v^R. Then S^* approaches infinity and Ω is negative. In Case 2, overconfidence is extremely strong such that v^C approaches infinity. Then S^* approaches zero and Ω is positive. Thus, there must exist a threshold $v^{C'}$ $(0 < v^{C'} < +\infty)$, above which Ω is positive. □

Proof of Proposition 2

Let $v^C = v^R$ so there is no overconfidence, the CAPM holds in this model. The expected aggregate return is equal to the market risk premium $\pi = A\sigma_R^2 Q$. It is easy to show that

$$\frac{\partial \pi}{\partial A} = \sigma_R^2 Q > 0, \qquad \frac{\partial \pi}{\partial (1/v)} = \frac{v^2}{(v + v^R)^2} AQ > 0.$$

That is, when the risk aversion A or the factor cash flow volatility is greater (v is smaller), the risk premium π is greater. Since the expected aggregate return is equal to π and $\hat{\sigma}(C-P) = \pi$, greater risk premia lead to both higher expected aggregate return and larger $\hat{\sigma}(C-P)$. $\square$

Proof of Proposition 3

Suppose there are two assets that have market betas of β_1 and β_2, respectively, and $\beta_1 > \beta_2$. Then the average expected returns are $\beta_1 \pi$ and $\beta_2 \pi$. Hence, for a unit increase in overconfidence v^C, the average expected return on asset one is reduced by $\beta_1 \left| \frac{\partial \pi}{\partial v^C} \right|$ while that on asset two is reduced by $\beta_2 \left| \frac{\partial \pi}{\partial v^C} \right|$. Thus, the return reduction effect of overconfidence is stronger among asset one. $\square$

Proof of Proposition 4

When $v_j^C = v^R + \eta_j (v^C - v^R)$, the factor risk premium

$$\pi_j = \frac{AQ}{v + v^R + \eta_j (v^C - v^R)}. \tag{A-1}$$

In particular, $\eta_1 > \eta_2$ so that investors are more overconfident about the first factor than about the second factor. Since

$$\frac{\partial \pi_j}{\partial v^C} = -AQ \frac{\eta_j}{[v + v^R + \eta_j (v^C - v^R)]^2} < 0, \tag{A-2}$$

when overconfidence v^C increases, both factor portfolio expected returns are on average decreased. In addition,

$$\frac{\partial^2 \pi}{\partial \eta \partial v^C} = -AQ \frac{v + v^R - \eta (v^C - v^R)}{[v + v^R + \eta (v^C - v^R)]^3}. \tag{A-3}$$

It is easy to show that when $\eta < \frac{v + v^R}{v^C - v^R}$, $\frac{\partial^2 \pi_j}{\partial \eta \partial v^C} < 0$. In this case, the return reduction is greater for factors with a higher η (greater overconfidence).

The reduction in the average expected return of the two asset classes is given by, respectively,

$$\frac{\partial E_0[R_1]}{\partial v^C} = \beta_{11} \frac{\partial \pi_1}{\partial v^C} + \beta_{12} \frac{\partial \pi_2}{\partial v^C}$$
$$\frac{\partial E_0[R_2]}{\partial v^C} = \beta_{21} \frac{\partial \pi_1}{\partial v^C} + \beta_{22} \frac{\partial \pi_2}{\partial v^C}.$$

Thus,

$$\frac{\partial E_0[R_1]}{\partial v^C} - \frac{\partial E_0[R_2]}{\partial v^C} = \frac{\partial \pi_1}{\partial v^C}(\beta_{11} - \beta_{21}) + \frac{\partial \pi_2}{\partial v^C}(\beta_{12} - \beta_{22}) \tag{A-4}$$

It is easy to see that, given $\beta_{11} > \beta_{21}$, when $\beta_{12} \geq \beta_{22}$, the above difference is negative. That is,

$$\frac{\partial E_0[R_1]}{\partial v^C} < \frac{\partial E_0[R_2]}{\partial v^C}. \tag{A-5}$$

In this case, the return reduction is greater for assets with higher loadings on the first factor. This case is subsumed by Proposition 3. However, when $\beta_{12} < \beta_{22}$, there exists a threshold η^*, where $\eta^* < \frac{v + v^R}{v^C - v^R}$, and when $\eta_1 < \eta^*$ equation (A-5) can hold. When $\eta < \frac{v + v^R}{v^C - v^R}$, the risk premium reduction effect is greater for the first factor. Thus

$$\frac{\partial \pi_1}{\partial v^C} < \frac{\partial \pi_2}{\partial v^C}. \tag{A-6}$$

Hence, there must exist a range of η_1 and η_2 ($\eta_1 > \eta_2$), with which equation (A-5) holds. $\square$

Appendix B. Decomposition of earnings-to-price ratio

Let D be the dividend per share, E be the earnings per share, P be the stock price, d be the log dividend per share, e be the log earnings per share, p be the log price. Further, let δ denote the log earnings-to-price ratio, θ be the log dividend-to-earnings ratio, and Δe be the log earnings growth rate. Let r denote the log stock return, defined as

$$r_t = \log\left(\frac{P_t + D_t}{P_{t-1}}\right). \tag{B-1}$$

Substituting δ, θ, and Δe into equation (B-1) yields

$$r_t = \Delta e_t + \delta_{t-1} + \theta_t + \log(\exp(-(\delta_t + \theta_t))). \tag{B-2}$$

I approximate the stock returns by a first-order Taylor expansion and obtain

$$r_t = \Delta e_t + \delta_{t-1} + \theta_t - \rho(\delta_t + \theta_t) + \kappa_t.$$

where ρ is a parameter and κ is an approximation error plus a constant. If the firm pays any dividends then $\rho < 1$, and otherwise $\rho = 0$. Rearranging the terms yields

$$r_t - \Delta e_t - (1-\rho)\theta_t = \delta_{t-1} - \rho\delta_t + \kappa_t. \tag{B-3}$$

Using the linear form in equation (B-3), I iterate forward and express the EP ratio as an infinite discounted sum of future returns less future earnings growth rates and future dividend payout ratios:

$$\delta_{t-1} = \sum_{j=0}^{\infty} \rho^j r_{t+j} + \sum_{j=0}^{\infty} \rho^j(-\Delta e_{t+j}) + \sum_{j=0}^{\infty} \rho^j(\rho - 1)\theta_{t+j} + \sum_{j=0}^{\infty} \kappa_{t+j}. \tag{B-4}$$

If we further assume the dividend payout ratio is a constant, then

$$\delta_{t-1} = \sum_{j=0}^{\infty} \rho^j r_{t+j} + \sum_{j=0}^{\infty} \rho^j (-\Delta e_{t+j}) + \sum_{j=0}^{\infty} \kappa'_{t+j}. \tag{B-5}$$

The above decomposition shows that the log earnings-to-price ratio is approximately the sum of future returns less future earnings growth rates plus a constant.

Appendix C. Firm valuation dispersion and multifactor models

In this appendix, I extend the framework in Section V.B to incorporate multifactors. In equation (21) of Section V.B, instead of assuming that the CAPM holds for expected returns, I assume that Fama-French 3 factor model describes the expected firm returns. Hence,

$$\begin{aligned} d_{i,t-1} - p_{i,t-1} \approx & \sum_{j=0}^{\infty} \bar{\rho}^j \left\{ \beta_{im}(R_{m,t+j} - R_{rf,t+j}) + \beta_{is} R_{\text{SMB},t+j} + \beta_{ih} R_{\text{HML},t+j} \right\} && \text{(C-6)} \\ & - \sum_{j=0}^{\infty} \bar{\rho}^j \Delta d_{i,t+j} + c_{t-1}. && \text{(C-7)} \end{aligned}$$

I assume the SMB returns and HML returns also follow AR(1) processes with the first-order autocorrelations γ_s and γ_h. Further, I additionally assume that the three factor returns are independent. In addition, let MKT= $R_{m,t} - R_{f,t}$, SMB= $R_{\text{SMB},t}$, and HML= $R_{\text{HML},t}$. Equation (C-6) can be written as

$$d_{i,t-1} - p_{i,t-1} \approx \frac{\beta_{i,m}}{1-\rho\gamma_m}\text{MKT} + \frac{\beta_{i,s}}{1-\rho\gamma_s}\text{SMB} + \frac{\beta_{i,h}}{1-\rho\gamma_h}\text{HML} - \frac{1}{1-\rho\gamma_d}\Delta d_{i,t} + c'_{t-1}.$$

I then take the cross-sectional variance of both sides to yield

$$\sigma^2(d_{t-1} - p_{t-1}) = b_m\text{MKT}^2 + b_s\text{SMB}^2 + b_h\text{HML}^2 + c_m\text{MKT} + c_s\text{SMB} + c_h\text{HML} + d\sigma^2(\Delta d_t),$$

where $b_m = \dfrac{\sigma^2(\beta_m)}{(1-\rho\gamma_m)^2}$, $b_s = \dfrac{\sigma^2(\beta_s)}{(1-\rho\gamma_s)^2}$, $b_h = \dfrac{\sigma^2(\beta_h)}{(1-\rho\gamma_h)^2}$, $c_m = \dfrac{cov(\beta_m, \Delta d_t)}{(1-\rho\gamma_m)(1-\rho\gamma_d)}$, $c_s = \dfrac{cov(\beta_s, \Delta d_t)}{(1-\rho\gamma_s)(1-\rho\gamma_d)}$, $c_h = \dfrac{cov(\beta_h, \Delta d_t)}{(1-\rho\gamma_h)(1-\rho\gamma_d)}$, and $d = \dfrac{\sigma^2(\Delta d_t)}{(1-\rho\gamma_d)^2}$.

A similar procedure applies to the 4-factor model and the BWX ICAPM on σ(DP) yields the following relationship:

$$\begin{aligned} \sigma^2(\text{DP}) = {} & b_m\text{MKT}^2 + b_s\text{SMB}^2 + b_h\text{HML}^2 + b_u\text{UMD}^2 \\ & + c_m\text{MKT} + c_s\text{SMB} + c_h\text{HML} + c_u\text{UMD} + d\sigma^2(\Delta d_t), \\ \sigma^2(\text{DP}) = {} & b_m\text{MKT}^2 + b_b\Delta\gamma^2 + b_e\Delta\eta^2 + c_m\text{MKT} + c_b\Delta\gamma + c_e\Delta\eta + d\sigma^2(\Delta d_t). \end{aligned}$$

where b_m, b_s, b_h, b_u, b_b, and b_e are functions of first-order autocorrelation of the corresponding factors and should be positive, c_m, c_s, c_h, c_u, c_b, and c_e are functions of the first-order autocorrelation of the corresponding factor, the first-order autocorrelation of the corresponding growth rate, and the cross-sectional covariance between the corresponding factor loadings and the expected corresponding growth rates.

Similar equations can be derived for σ(BM) and σ(EP) except to replace Δd_t with roe_t and Δe_t, respectively. Please refer to Table XI for model specifications.

References

Ang, Andrew, and Geert Bekaert, 2004, Stock Return Predictability: Is it There?, Working Paper, Columbia University.

Ang, Andrew, and Joseph Chen, 2005, CAPM over the Long Run: 1926-2001, *Journal of Empirical Finance* forthcoming.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The Cross-Section of Volatility and Expected Returns, *Journal of Finance* 61, 259–299.

Asquith, Paul, and Lisa Meulbroek, 1995, An Emprical Investigation of Short Interest, M.I.T. working paper.

Baker, Malcolm, and Jeffrey Wurgler, 2000, The Equity Share in New Issues and Aggregate Stock Returns, *Journal of Finance* 55, 2219–2257.

Baker, Malcolm, and Jeffrey Wurgler, 2005, Investor Sentiment and The Cross-Section of Stock Returns, *Journal of Finance* forthcoming.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A Model of Investor Sentiment, *Journal of Financial Economics* 49, 307–343.

Boehme, Rodney D., Bartley R. Danielsen, and Sorin M. Sorescu, 2005, Short Sale Constraints, Differences of Opinion, and Overvaluation, *Journal of Financial and Quantitative Analysis* forthcoming.

Brennan, Michael J., Ashley W. Wang, and Yihong Xia, 2001, Intertemporal Capital Asset Pricing and the Fama-French Three-Factor Model, Working paper, UCLA and University of Pennsylvania.

Brennan, Michael J., Ashley W. Wang, and Yihong Xia, 2004, Estimation and Test of a Simple Model of Intertemporal Capital Asset Pricing, *Journal of Finance* 59, 1743–1776.

Brennan, Michael J., and Yihong Xia, 2005, *tay*'s as good as *cay*, *Finance Research Letters* 2, 1–14.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading Volume and Serial Correlation in Stock Returns, *Quarterly Journal of Economics* 108, 905–939.

Campbell, John Y., and Robert J. Shiller, 1988, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195–228.

Campbell, John Y., and Tuomo Vuolteenaho, 2004, Bad Beta, Good Beta, *American Economic Review* 94, 1249–1275.

Carhart, Mark M., 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 57–82.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein, 2002, Breadth of Ownership and Stock Returns, *Journal of Financial Economics* 66, 171–205.

Cohen, Randolph B., Christopher K. Polk, and Tuomo Vuolteenaho, 2003, The Value Spread, *Journal of Finance* 58, 609–642.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor Psychology and Security Market Under- and Overreactions, *Journal of Finance* 53, 1839–1886.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the Characteristics of Cross Sectional Variation in Stock Returns, *Journal of Finance* 52, 1–33.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 2001, Overconfidence, Arbitrage, and Equilibrium Asset Pricing, *Journal of Finance* LVI, 921–965.

Daniel, Kent D., David Hirshleifer, and Avanidhar Subrahmanyam, 2005, Investor Psychology and Tests of Factor Pricing Models, Preliminary Version, UCLA, Northwestern University, and Ohio State University.

Davis, James, Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, Covariances, and Average Returns: 1929-1997, *Journal of Finance* 55, 389–406.

D'Avolio, Gene, 2002, The Market for Borrowing Stock, *Journal of Financial Economics* 66, 271–306.

Desai, Hemang, K. Ramesh, S. Ramu Thiagarajan, and Bala V. Balachandran, 2002, An Investigation of the Informational Role of Short Interest in the Nasdaq Market, *Journal of Finance* 57, 2263–2287.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of Opinion and the Cross Section of Stock Returns, *Journal of Finance* 57, 2113–2141 5.

Dong, Ming, David Hirshleifer, Scott Richardson, and Siew Hong Teoh, 2003, Stock Market Misvaluation and Merger Booms, Manuscript, Ohio State University, York University and Wharton School.

Einhorn, Hillel J., 1980, Overconfidence in Judgment, *New Directions for Methodology of Social and Behavioral Science* 4, 1–16.

Eleswarapu, Venkat, and Marc R. Reinganum, 2004, The Predictability of Aggregate Stock Market Returns: Evidence Based on Glamour Stocks, *Journal of Business* 77, 275–294.

Fama, Eugene F., and Kenneth R. French, 1988a, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics* 22.

Fama, Eugene F., and Kenneth R. French, 1988b, Permanent and Temporary Components of Stock Prices, *Journal of Political Economy* 96.

Fama, Eugene F., and Kenneth R. French, 1989, Business Conditions and Expected Returns on Stocks and Bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene F., and Kenneth R. French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, *Journal of Finance* 50, 131–155.

Fama, Eugene F., and Kenneth R. French, 2004, New Lists: Fundamentals and Survival Rates, *Journal of Financial Economics* 73, 229–269.

Fama, Eugene F., and James D. MacBeth, 1973, Risk Return and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607–636.

Fama, Eugene F., and G. William Schwert, 1977, Asset Returns and Inflation, *Journal of Financial Economics* 5, 115–146.

Fink, Jason, Kristin E. Fink, Gustavo Grullon, and James P. Weston, 2005, IPO Vintage and the Rise of Idiosyncratic Risk, Working paper, James Madison University and Rice University.

Gervais, Simon, and Terrance Odean, 2001, Learning to be Overconfident, *Review of Financial Studies* 14, 1–27.

Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium Cross Section of Returns, *Journal of Political Economy* 111, 693–732.

Gordon, Myron, 1962, *The Investment, Financing, and Valuation of The Corporation.* (Irwin Homewood, IL).

Goyal, Amit, and Pedro Santa-Clara, 2003, Idiosyncratic Risk Matters!, *Journal of Finance* 58, 975–1007.

Goyal, Amit, and Ivo Welch, 2005, A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, Working paper, Emory University and Brown University.

Guo, Hui, and Robert Savickas, 2006, Idiosyncratic Volatility, Stock Market Volatility, and Expected Stock Returns, *Journal of Business & Economic Statistics* 24, 43–56.

Harris, Milton, and Artur Raviv, 1993, Difference of Opinion Make a Horse Race, *Review of Financial Studies* 6, 473–506.

Hirshleifer, David, 2001, Investor Psychology and Asset Pricing, *Journal of Finance* 56, 1533–1597.

Hong, Harrison, José Scheinkman, and Wei Xiong, 2006, Asset Float and Speculative Bubbles, *Journal of Finance* 61, 1073–1117.

Hong, Harrison, and Jeremy C. Stein, 1999, A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, *Journal of Finance* 54, 2143–2184.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Narasimhan Jegadeesh and Sheridan Titman, *Journal of Finance* 56, 699–720.

Jiang, Danling, 2005a, Overconfidence, Short-Sale Constraints, and Stock Valuation, Working paper, The Ohio State University.

Jones, Charles M., and Owen A. Lamont, 2002, Short Sale Constraints and Stock Returns, *Journal of Financial Economics* 66, 207–223.

Kothari, S.P., and Jay Shanken, 1997, Book-to-market, Dividend Yield, and Expected Market Returns: A Time-series Analysis, *Journal of Financial Economics* 44, 169–203.

Kumar, Alok, 2005, When Do Investors Exhibit Stronger Behavioral Baises?, University of Notre Dame, working paper.

Lamont, Owen, 1998, Earnings and Expected Returns, *Journal of Finance* 53, 1563–1587.

Lee, Charles M.C., and Bhaskaran Swaminathan, 2000, Price Momentum and Trading Volume, *Journal of Finance* 55, 2017–2069.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, Aggregate Wealth, and Expected Stock Returns, *Journal of Finance* 56, 815–849.

Lewellen, Jonathan, 2004, Predicting Returns with Financial Ratios, *Journal of Financial Economics* 74, 209–235.

Lewellen, Jonahan, and Stefan Nagel, 2005, The Conditional CAMP Does Not Explain Asset-Pricing Anomalies, *Journal of Financial Economics* forthcoming.

Liu, Naiping, and Lu Zhang, 2005, The Value Spread as Predictors of Returns, Working paper, University of Pennsylvania and University of Rochester.

Nagel, Stefan, 2005, Short Sales, Institutional Investors, and The Cross-Section of Stock Returns, *Journal of Financial Economics* forthcoming.

Nelson, Charles R., and Myung J. Kim, 1993, Predictable Stock Returns: The Role of Small Sample Bias, *Journal of Finance* 48, 641–661.

Phalippou, Ludovic, 2004, What Drives The Value Premium?, Working paper, INSEAD.

Polk, Christopher, and Paola Sapienza, 2005, The Real Effects of Investor Sentiment, working paper, Kellogg School of Management, Northwestern University.

Polk, Christopher, Samuel Thompson, and Tuomo Vuolteenaho, 2004, Cross-sectional forecasts of the equity premium, *Journal of Financial Economics* forthcoming.

Pontiff, Jeffrey, and Lawrence D. Schall, 1998, Book-to-Market Ratios as Predictors of Market Returns, *Journal of Financial Economics* 49, 141–160.

Santos, Tano, and Pietro Veronesi, 2004, Conditional Betas, Working paper, Columbia University and University of Chicago.

Scheinkman, Jose A., and Wei Xiong, 2003a, Overconfidence and Speculative Bubbles, *Journal of Political Economy* 111, 1183–1219.

Scheinkman, Jose A., and Wei Xiong, 2003b, Overconfidence, Short-Sale Constraints, and Bubbles, *Journal of Political Economy* 111, 1183–1219.

Shleifer, Andrei, and Robert Vishny, 1997, The Limits to Arbitrage, *Journal of Finance* 52, 35–55.

Stambaugh, Robert F., 1986, Biases in Regressions with Lagged Stochastic Regressors, CRSP Working Paper No. 156, University of Chicago, Graduate School of Business.

Stambaugh, Robert F., 1999, Predictive Regressions, *Journal of Financial Economics* 54, 375–421.

Statman, Meir, Steven Thorley, and Keith Vorkink, 2006, Investor Overconfidence and Trading Volume, *Review of Financial Studies.*

Vuolteenaho, Tuomo, 2000, Understanding The Aggregate Book-to-Market Ratio and Its Implications to Current Equity-Premium Expectations, Working paper, Harvard University.

Vuolteenaho, Tuomo, 2003, What Drives Firm-Level Stock Returns?, *Journal of Finance* 57, 233–264.

Zhang, X. Frank, 2006, Information Uncertainty and Stock Returns, *Journal of Finance* 61, 105–137.

Table I: Summary statistics of the dispersion variables

This table reports the descriptive statistics of the annual dispersion variables over the period 1963–2004, including the number of observations (Obs), the time series mean (Mean), the time series standard deviation (Std), the maximum (Max), the minimum (Min), and the autocorrelations in three lags. Each variable is formed based on all available firms in NYSE, AMEX, and NASDAQ. Variables in Panel A are in a value-weighted scheme and those in Panel B are in an equal-weighted scheme. The variables σ(BM), σ(DP), and σ(EP) are, respectively, the cross-sectional standard deviation of logarithmic firm book-to-market equity, dividend-to-price ratios, and earnings-to-price ratios. The variables cd(BM), cd(DP), and cd(EP) are the detrended σ(BM), σ(DP), and σ(EP). The variable CVD is the first principal component of cd(BM), cd(DP), and cd(EP).

Panel A: Value-weighted variables

						Autocorrelation		
	Obs	Mean	Std	Min	Max	1	2	3
σ(BM)	42	19.39	12.18	7.95	61.04	0.89	0.73	0.59
σ(DP)	42	31.65	23.91	10.69	99.87	0.92	0.82	0.71
σ(EP)	42	19.04	14.87	7.43	75.89	0.85	0.71	0.55
cd(BM)	42	0	5.84	−8.63	23.19	0.63	0.09	−0.25
cd(DP)	42	0	8.07	−28.66	31.24	0.49	0.17	−0.17
cd(EP)	42	0	7.06	−9.37	34.30	0.47	0.17	−0.16
CVD	42	0	1.00	−2.22	4.45	0.58	0.13	−0.21

Panel B: Equal-weighted variables

						Autocorrelation		
	Obs	Mean	Std	Min	Max	1	2	3
σ(BM)	42	0.79	0.12	0.59	1.27	0.69	0.47	0.41
σ(DP)	42	1.00	0.17	0.63	1.29	0.89	0.76	0.63
σ(EP)	42	0.92	0.15	0.63	1.37	0.80	0.63	0.49
cd(BM)	42	0	0.10	−0.16	0.39	0.52	0.20	0.10
cd(DP)	42	0	0.08	−0.18	0.15	0.66	0.34	0.08
cd(EP)	42	0	0.08	−0.13	0.28	0.38	−0.04	−0.26
CVD	42	0	1.00	−2.07	3.59	0.44	0.00	−0.23

Table II: Correlation matrix

Panel A reports the Pearson correlations of cross-firm valuation dispersion (CVD) and its underlying measures. Panel B reports the correlations of the well-known aggregate predictors and the value- and equal-weighted CVDs. The value-weighted (equal-weighted) CVD is denoted as CVD$_{\text{vw}}$ (CVD$_{\text{ew}}$) in Panel B. Aggregate dividend yield (D/P) is computed following Fama and French (1988a). Annual dividends are used. Short term interest rate (TBILL) is measured by the one-month treasury bill rate obtained from Ibbotson Associates. Past market return (Rm) is the past one-year CRSP value-weight index returns. Aggregate relative equity issuance (S) is obtained from Wurgler's website over the period 1962–2002. Term premium (TERM) is defined as the difference between the 10-Year treasury constant maturity rate and the one-year treasury constant maturity rate. Default premium (DEF) is defined as the difference between the Moody's seasoned Aaa corporate bond yield and the Moody's seasoned Baa corporate bond yield. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% level based on the two-tailed p-value.

Panel A: Correlations of dispersion measures

	VW			EW		
	CVD	cd(BM)	cd(DP)	CVD	cd(BM)	cd(DP)
cd(BM)	0.97***			0.69***		
cd(DP)	0.94***	0.88***		0.78***	0.18	
cd(EP)	0.94***	0.88***	0.81***	0.94***	0.56***	0.68***

Panel B: Correlations of aggregate return predictors

	CVD$_{\text{vw}}$	CVD$_{\text{ew}}$	Rm	D/P	TERM	DEF	TBILL
CVD$_{\text{ew}}$	0.75**						
Rm	−0.02	0.11					
D/P	−0.32**	−0.10	−0.23				
TERM	−0.27*	−0.18	0.07	−0.23			
DEF	−0.03***	0.25	−0.18	0.54***	0.19		
TBILL	0.11	0.34**	−0.11	0.67***	−0.48***	0.57***	
S	0.08	0.23	−0.28*	0.59***	−0.31**	0.44***	0.55***

Table III: Regression of the annual dispersion variables on a time index

This table reports the results by regressing the annual dispersion variables on a time index (and the squared time index) over the period 1963–2004. Each dependent variable is formed based on all available firms in NYSE, AMEX, and NASDAQ. The independent variable t-index is a time index taking values of 1, 2, 3... The independent variable t^2-index is the squared t-index. Dependent variables in Panel A are in a value-weighted scheme and those in Panel B are in an equal-weighted scheme. The variables σ(BM), σ(DP), and σ(EP) are, respectively, the cross-sectional standard deviation of logarithmic firm book-to-market equity, dividend-to-price ratios, and earnings-to-price ratio. OLS t-statistics are reported below the coefficient and the two-tailed p-values are reported in parenthesis. R-squares are adjusted for degree of freedom.

Panel A: VW			
	σ(BM)	σ(DP)	σ(EP)
Intercept	19.27	18.21	15.33
	6.62	*4.53*	*4.36*
	(0.00)	*(0.00)*	*(0.00)*
t-index	−1.33	−1.32	−1.24
	−4.25	*−3.07*	*−3.29*
	(0.00)	*(0.00)*	*(0.00)*
t^2-index	0.05	0.07	0.05
	6.79	*7.17*	*5.97*
	(0.00)	*(0.00)*	*(0.00)*
R^2	77%	88%	77%

Panel A: EW			
	σ(BM)	σ(DP)	σ(EP)
Intercept	0.67	0.75	0.71
	21.76	*30.23*	*28.32*
	(0.00)	*(0.00)*	*(0.00)*
t-index	0.01	0.01	0.01
	4.49	*11.82*	*9.96*
	(0.00)	*(0.00)*	*(0.00)*
R^2	32%	77%	71%

Table IV: Summary statistics of firm characteristic portfolios

Panels A to E report the summary statistics of industry portfolios, risk or characteristic quintiles. The industry portfolios are the Fama-French 10 industries, including consumer nondurables (NoDur), consumer durables (Durbl), manufacturing (Manuf), energy (Enrgy), computers, software, and electronic equipment (HiTec), telephone and television transmission (Telcm), wholesale and retail (Shops), utilities (Utils), and others (Other). Utils and Others are excluded. The two technology-related industries are HiTec and Telcm. The HiTec industry provides computers, software, and electronic equipment and services. The Telcm industry provides the telecommunication services. Market beta (BETA) is estimated using a market model on 36 to 60 monthly available returns in the 5 years before July of year t. Return volatility (VOL) is the volatility of 10 to 12 monthly returns before July of year t. Firm age (AGE) is the number of years between the firm's first appearance on CRSP and t. Earnings-to-book equity (E/BE) is the ratio of net income from continuing operation (data 178) over book equity. Plant, property, and equipment (data 7) (PPE) is scaled by total assets (data 6). Dividends-to-book equity (D/BE) is total dividends (data 21) over book equity. Institutional ownership (IO) is the total number of shares held by institutional investors scaled by shares outstanding. Short interest ratio (SIR) is measured by total number of shares shorted scaled by shares outstanding. Analyst forecast dispersion (ADISP) is defined as the standard deviation of analysts' current fiscal year earnings per share forecasts scaled by absolute mean forecasts. Trading volume (TURN) is the average turnover-trading volume scaled by shares outstanding-of no more than 12 months before July of year t. Trading volume of NASDAQ firms is divided by two due to the double counting problem in NASDAQ. All portfolios but for IO, ADISP, and SIR quintiles are formed at the end of June of year t and the annual returns are continuously-compounded value-weight returns measured from July of year t to June of year $t+1$ over the period 1963–2004. The IO quintiles are formed at the end of each quarter, and the 12-month continuously-compounded value-weight returns from the beginning of the next quarter are calculated over the period 1980–2004. The SIR and ADISP quintiles are formed at the end of each month, and the 12-month continuously-compounded value-weight returns from the next month are used. The SIR quintiles are available from 1991 to 2004. The ADISP quintiles are available from 1976 to 2004. For BETA, AGE, and IO, NYSE breakpoints are used. Time series averages of the value-weighted 12-month returns (Ret), of the average market equity (ME), and of the average book-to-market equity (BM) for each portfolio are reported.

Panel A: Industries											
		NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth		
Ret		6.36	3.19	3.99	5.47	4.18	3.21	4.96	6.55		
ME		801	683	611	1335	596	2393	513	853		
BM		0.52	0.94	0.70	0.77	0.52	0.36	0.81	1.02		
Panel B: Riskiness											
	L	2	3	4	H		L	2	3	4	H
			BETA						VOL		
Ret	5.20	5.20	5.13	4.54	4.30		5.65	4.66	4.75	4.73	5.58
ME	1121	1342	1315	1260	783		2477	1196	691	463	253
BM	0.99	0.93	0.90	0.88	0.79		0.87	0.90	0.90	0.90	0.86
Panel C: Cash flow uncertainty											
			AGE						E/BE		
Ret	5.65	5.70	4.62	5.50	4.39		5.26	4.37	3.81	4.72	5.58
ME	304	418	883	1229	4820		522	779	826	1405	2197
BM	0.74	0.88	0.94	0.99	0.94		1.28	1.01	0.82	0.66	0.47
			PPE						D/BE		
Ret	4.25	6.03	5.38	4.70	4.54		6.59	5.15	5.09	5.10	4.91
ME	635	714	1063	1122	1233		370	852	918	1356	4280
BM	0.71	0.76	0.82	0.91	1.03		0.80	1.20	1.03	0.83	0.57
Panel D: Investor disagreement											
			ADISP						TURN		
Ret	7.31	3.83	5.97	3.54	3.51		6.51	5.01	3.97	4.65	3.23
ME	2711	2117	1542	1177	771		549	1178	1305	1052	711
BM	0.55	0.57	0.66	0.77	0.94		0.99	0.91	0.88	0.85	0.74
Panel E: Difficulty in selling short											
			IO						SIR		
Ret	5.00	6.40	7.49	6.84	7.99		15.71	6.06	7.66	8.07	4.63
ME	1037	444	1370	2056	1904		198	1514	3745	2782	1399
BM	0.73	0.76	0.78	0.72	0.66		0.86	0.68	0.54	0.48	0.43

Table V: Predictability of market excess returns

This table reports results by regressing market excess returns on the lagged cross-firm valuation dispersion (CVD). The CRSP value-weighted and equal-weighted indices are used as the value-weighted and equal-weighted market portfolios. Subsequent one-quarter, one-year, and three-year returns in excess of the risk-free rate are used as dependent variables. Panel A shows the results for the full sample in which CVD from 1963 to 2003 is used. Panel B shows the results for the subperiod in which CVD from 1963 to 1996 is used. The value-weighted (equal-weighted) CVD is used to predict the value-weighted (equal-weighted) returns. In forecasting one-quarter returns, the predictors are updated at the end of each March, June, September, and December. In forecasting one-year or three-year returns, they are updated annually at the end of each June. Overlapping observations are used for the three-year return regressions. The coefficients of OLS regressions are reported, below which in parenthesis are the one-tail p-values of the coefficients in the simulated distribution using the method by Nelson and Kim (1993). R-squares are adjusted for degree of freedom.

Panel A: Full sample 1963-2003

	VW			EW		
	3-mon	1-yr	3-yr	3-mon	1-yr	3-yr
Intercept	1.22	4.74	13.68	2.17	7.96	22.97
	1.84	*2.10*	*4.94*	*2.28*	*2.40*	*4.79*
	(0.00)	*(0.00)*	*(0.00)*	*(0.01)*	*(0.01)*	*(0.00)*
CVD	−2.38	−8.34	−18.97	−2.58	−8.76	−12.74
	−3.39	*−3.47*	*−6.39*	*−2.61*	*−2.50*	*−2.49*
	(0.00)	*(0.00)*	*(0.00)*	*(0.01)*	*(0.01)*	*(0.00)*
R^2	6%	22%	51%	3%	12%	12%
Obs	163	41	40	163	41	40

Panel B: Subperiod 1963−1996

	VW			EW		
	3-mon	1-yr	3-yr	3-mon	1-yr	3-yr
Intercept	0.66	2.82	11.16	1.84	7.24	21.38
	0.95	*1.12*	*3.78*	*1.79*	*1.86*	*4.00*
	(0.01)	*(0.02)*	*(0.00)*	*(0.03)*	*(0.03)*	*(0.01)*
CVD	−4.61	−16.60	−31.88	−2.40	−11.10	−17.45
	−3.36	*−3.24*	*−5.31*	*−1.94*	*−2.21*	*−2.53*
	(0.01)	*(0.02)*	*(0.00)*	*(0.03)*	*(0.03)*	*(0.01)*
R^2	7%	22%	45%	2%	11%	14%
Obs	135	34	34	135	34	34

Table VI: Predictability of market excess returns with control variables

This table reports results by regressing market excess returns on cross-firm valuation dispersion (CVD) and a set of well-known predictors, including aggregate dividend yield (D/P), short term interest rate (TBILL), past market return (Rm), term premium (TERM), default premium (DEF), and aggregate relative equity issuance (S). The CRSP value-weighted and equal-weighted indices are used as the value-weighted and equal-weighted market portfolios. Subsequent one-quarter, one-year, and three-year returns in excess of the risk-free rate are used as dependent variables. Panel A shows the results for the value-weighted market excess returns and Panel B shows the results for the equal-weighted market excess returns. The value-weighted (equal-weighted) CVD and Rm are used to predict the value-weighted (equal-weighted) returns. In forecasting one-quarter returns, the predictors are updated at the end of each March, June, September, and December. In forecasting one-year or three-year returns, they are updated annually at the end of each June. Overlapping observations are used for the three-year return regressions. The coefficients of OLS regressions are reported, below which in parenthesis are the one-tail p-values of the coefficients in the simulated distribution using the method by Nelson and Kim (1993). R-squares are adjusted for degree of freedom.

Panel A: Value-weighted market excess returns

3 months

Intercept	CVD	Rm	D/P	TERM	DEF	TBILL	S	R2	Obs
−2.64		0.00	1.92	0.45	1.43	−8.32		2%	163
(0.41)		*(0.49)*	*(0.02)*	*(0.31)*	*(0.24)*	*(0.06)*			
−0.33	−2.13	−0.02	0.69	−0.19	3.26	−7.68		6%	163
(0.39)	*(0.01)*	*(0.39)*	*(0.26)*	*(0.41)*	*(0.05)*	*(0.06)*			
1 year									
−10.74		−0.10	6.63	5.76	−4.56	−4.60	−16.86	8%	41
(0.22)		*(0.27)*	*(0.18)*	*(0.06)*	*(0.33)*	*(0.37)*	*(0.35)*		
−3.33	−6.44	−0.11	2.54	3.10	1.42	−7.00	−0.37	16%	41
(0.25)	*(0.07)*	*(0.24)*	*(0.40)*	*(0.23)*	*(0.40)*	*(0.34)*	*(0.48)*		
3 years									
−19.00		−0.16	4.76	19.22	−40.04	70.89	45.12	24%	40
(0.39)		*(0.12)*	*(0.37)*	*(0.00)*	*(0.00)*	*(0.00)*	*(0.12)*		
4.89	−19.04	−0.16	−7.81	11.35	−18.27	56.88	89.88	60%	40
(0.36)	*(0.00)*	*(0.16)*	*(0.00)*	*(0.00)*	*(0.02)*	*(0.00)*	*(0.01)*		

Panel B: Equal−weighted market excess returns

3 months

Intercept	CVD	Rm	D/P	TERM	DEF	TBILL	S	R2	Obs
−1.14		−0.09	2.80	−0.45	6.12	−23.27		7%	163
(0.40)		*(0.14)*	*(0.01)*	*(0.37)*	*(0.01)*	*(0.00)*			
−0.78	−2.11	−0.09	1.49	−0.43	7.60	−18.38		8%	163
(0.42)	*(0.02)*	*(0.12)*	*(0.11)*	*(0.37)*	*(0.00)*	*(0.00)*			
1 year									
−5.39		−0.26	12.49	4.87	0.76	−32.59	−70.10	20%	41
(0.27)		*(0.03)*	*(0.04)*	*(0.16)*	*(0.38)*	*(0.06)*	*(0.07)*		
−5.67	−3.87	−0.23	10.08	4.25	4.22	−27.72	−56.57	20%	41
(0.28)	*(0.19)*	*(0.05)*	*(0.18)*	*(0.22)*	*(0.29)*	*(0.14)*	*(0.17)*		
3 years									
11.27		−0.14	12.52	8.47	−23.64	−1.70	−43.34	−1%	40
(0.24)		*(0.21)*	*(0.03)*	*(0.08)*	*(0.07)*	*(0.49)*	*(0.25)*		
13.08	−9.59	−0.14	6.31	7.09	−13.59	8.10	−17.73	2%	40
(0.23)	*(0.06)*	*(0.21)*	*(0.19)*	*(0.12)*	*(0.21)*	*(0.40)*	*(0.38)*		

Table VII: Predictability of market excess returns with valuation spreads

This table reports results by regressing market excess returns on the lagged cross−firm valuation dispersion (CVD) and several valuation spreads during 1963-2004. The BM (MB) spread is the difference of book-to-market (market-to-book) equity between value firms and growth firms. The value spread is defined as the difference in logarithmic book-to-market equity between value firms and growth firms. The small value spread is the value spread in small firms. Value firms and growth firms in the first three spreads are the top and bottom book-to-market deciles. Value firms, growth firms, and small firms in the small value spread are obtained from the Fama-French 2x3 sorts. Panel A shows the results for the value-weighted market excess returns and Panel B shows the results for the equal-weighted market excess returns. The value-weighted (equal-weighted) CVD is used to predict the value-weighted (equal-weighted) returns. Overlapping observations are used for the three-year return regressions. The coefficients of OLS regressions are reported, below which in parenthesis are the one-tail p-values of the coefficients in the simulated distribution using the method by Nelson and Kim (1993). R-squares are adjusted for degree of freedom.

Panel A: Value-weighted market excess returns

	1 year								3 years							
Intercept	4.30	4.26	4.33	4.36	4.71	4.78	4.76	4.72	12.14	12.08	12.13	12.14	13.64	14.13	14.10	13.88
	(0.31)	*(0.15)*	*(0.17)*	*(0.04)*	*(0.49)*	*(0.48)*	*(0.50)*	*(0.46)*	*(0.37)*	*(0.19)*	*(0.23)*	*(0.07)*	*(0.09)*	*(0.11)*	*(0.29)*	*(0.17)*
BM spread	−0.51				1.03				−0.91				2.59			
	(0.31)				*(0.44)*				*(0.37)*				*(0.19)*			
MB spread		−3.65				1.12				−4.79				7.78		
		(0.15)				*(0.31)*				*(0.19)*				*(0.00)*		
Value spread			−3.28				1.35				−3.22				9.58	
			(0.17)				*(0.30)*				*(0.23)*				*(0.00)*	
Small value spread				−5.08				−1.66				−4.96				5.25
				(0.04)				*(0.31)*				*(0.07)*				*(0.05)*
CVD					−8.54	−9.01	−9.11	−7.48					−19.50	−23.83	−24.79	−21.85
					(0.00)	*(0.01)*	*(0.01)*	*(0.01)*					*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*
R^2	−2%	3%	2%	8%	20%	20%	20%	20%	−2%	−1%	2%	2%	50%	56%	60%	53%
Obs	41	41	41	41	41	41	41	41	40	40	40	40	40	40	40	40

Panel B: Equal−weighted market excess returns

	1 year								3 years							
Intercept	7.30	7.48	7.56	7.58	7.75	7.88	8.02	7.97	21.36	21.88	21.94	21.92	22.45	22.76	23.46	23.44
	(0.08)	*(0.07)*	*(0.24)*	*(0.16)*	*(0.49)*	*(0.45)*	*(0.49)*	*(0.49)*	*(0.00)*	*(0.06)*	*(0.36)*	*(0.46)*	*(0.18)*	*(0.06)*	*(0.12)*	*(0.16)*
BM spread	6.53				6.89				16.68				17.21			
	(0.08)				*(0.05)*				*(0.00)*				*(0.00)*			
MB spread		6.28				1.96				9.49				3.52		
		(0.07)				*(0.43)*				*(0.06)*				*(0.28)*		
Value spread			3.21				2.95				1.49				10.01	
			(0.24)				*(0.28)*				*(0.36)*				*(0.05)*	
Small value spread				3.80				0.82				1.23				11.19
				(0.16)				*(0.41)*				*(0.46)*				*(0.02)*
CVD					9.05	7.50	10.66	9.22					13.49	10.37	19.57	19.40
					(0.00)	*(0.04)*	*(0.02)*	*(0.02)*					*(0.00)*	*(0.02)*	*(0.00)*	*(0.00)*
R^2	6%	6%	0%	0%	19%	10%	10%	9%	25%	7%	−2%	−2%	40%	10%	16%	19%
Obs	41	41	41	41	41	41	41	41	40	40	40	40	40	40	40	40

Table VIII: Two-way sorts on cross-firm valuation dispersion and firm characteristics

Panels A to E report the average annual returns of the two-way sorts based on firm characteristics and the equal-weighted cross-firm valuation dispersion (CVD). The industry portfolios are the Fama-French 10 industries excluding Utils and Other. The two technology-related industries are HiTec and Telcm. Other firm characteristics include market beta (BETA), return volatility (VOL), firm age (AGE), earnings-to-book equity (E/BE), plant, property, and equipment (PPE), dividends-to-book equity (D/BE), institutional ownership (IO), short interest ratio (SIR), analyst forecast dispersion (ADISP), and trading volume (TURN). CVD is positive for 1967–75, 1980–87, 1990–91, 1996, and 1999–2001. Difference is the return differential between the low CVD states (when CVD < 0) and the high CVD states (when CVD > 0).

Panel A: Industries													
			NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth			
CVD < 0			6.71	4.89	5.90	8.86	12.63	8.72	6.64	9.53			
CVD > 0			6.09	1.86	2.50	2.82	−2.43	−1.10	3.64	4.21			
Difference			0.62	3.03	3.40	6.04	15.05	9.82	3.00	5.32			
Panel B: Riskiness													
	L	2	3	4	H	L−H		L	2	3	4	H	L−H
			BETA							VOL			
CVD < 0	6.40	6.57	7.34	10.97	13.42	−7.02		7.21	8.22	11.32	14.57	19.53	−12.33
CVD > 0	4.26	4.13	3.40	−0.49	−2.83	7.09		4.43	1.87	−0.39	−2.98	−5.33	9.77
Difference	2.14	2.44	3.94	11.47	16.25	−14.11		2.77	6.34	11.71	17.55	24.87	−22.09
Panel C: Cash flow uncertainty													
			AGE							E/BE			
CVD < 0	15.90	13.05	10.27	8.17	6.63	9.27		12.61	7.77	7.87	7.96	9.44	3.16
CVD > 0	−2.37	−0.06	0.20	3.41	2.64	−5.01		−0.50	1.70	0.63	2.18	2.56	−3.06
Difference	18.26	13.11	10.07	4.76	3.99	14.28		13.10	6.07	7.24	5.78	6.88	6.22
			PPE							D/BE			
CVD < 0	12.90	12.75	9.37	8.58	6.49	6.41		17.89	12.20	9.81	8.47	6.77	11.12
CVD > 0	−2.53	0.76	2.25	1.67	3.02	−5.55		−2.26	−0.37	1.39	2.46	3.45	−5.71
Difference	15.43	11.99	7.12	6.92	3.48	11.96		20.14	12.57	8.42	6.00	3.32	16.83
Panel D: Investor disagreement													
			ADISP							TURN			
CVD < 0	8.06	5.95	8.40	6.52	6.27	1.80		8.94	7.15	9.85	12.78	15.06	−6.13
CVD > 0	6.49	1.53	3.33	0.32	0.51	5.98		4.60	3.35	−0.63	−1.70	−6.04	10.64
Difference	1.58	4.43	5.06	6.20	5.76	−4.18		4.33	3.80	10.48	14.48	21.10	−16.77
Panel E: Difficulty in selling short													
			IO							SIR			
CVD < 0	10.69	14.55	13.28	11.73	11.82	−0.16		19.30	10.14	14.25	15.79	13.90	6.88
CVD > 0	1.34	0.59	3.36	3.34	5.25	−3.91		11.69	1.02	−0.49	−1.46	−6.65	18.34
Difference	9.36	13.96	9.92	8.39	6.57	3.75		7.61	9.13	14.74	17.25	20.55	−11.46

Table IX: Predictability of portfolio excess returns

This table reports results by regressing excess returns of firm characteristic portfolios on the lagged equal-weighted cross-firm valuation dispersion (CVD). The industry portfolios are the Fama-French 10 industries excluding Utils and Other. The two technology-related industries are HiTec and Telcm. Other firm characteristics include market beta (BETA), return volatility (VOL), firm age (AGE), earnings-to-book equity (E/BE), plant, property, and equipment (PPE), dividends-to-book equity (D/BE),, institutional ownership (IO), short interest ratio (SIR), analyst forecast dispersion (ADISP), and trading volume (TURN). L−H is the long-short portfolio that is long on the lowest characteristic quintile and short on the highest characteristic quintile. The coefficients of OLS regressions are reported, below which in parenthesis are the one-tail p-values of the coefficients in the simulated distribution using the method by Nelson and Kim (1993). R-squares are adjusted for degree of freedom.

Panel A: Industries													
			NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth			
Coef			−0.98	−1.61	−3.46	−3.36	−14.14	−6.20	−2.71	−5.74			
p			*(0.37)*	*(0.30)*	*(0.10)*	*(0.14)*	*(0.00)*	*(0.03)*	*(0.21)*	*(0.03)*			
R2			−3%	−2%	0%	0%	21%	5%	−2%	5%			
Panel B: Riskiness													
			Coef							R2			
	L	2	3	4	H	L−H	p(L−H)	L	2	3	4	H	L−H
BETA	−1.60	−2.70	−3.71	−8.77	−15.17	13.57	*(0.00)*	−1%	1%	2%	14%	22%	26%
VOL	−2.47	−6.75	−9.29	−14.30	−20.08	17.62	*(0.00)*	0%	12%	18%	24%	34%	39%
Panel C: Cash flow uncertainty													
AGE	−16.54	−11.24	−8.95	−6.12	−3.40	−13.14	*(0.00)*	33%	21%	18%	10%	3%	45%
E/BE	−9.96	−5.97	−6.56	−6.91	−6.07	−3.89	*(0.03)*	20%	11%	12%	13%	9%	7%
PPE	−14.30	−10.21	−7.14	−6.70	−2.21	−12.09	*(0.00)*	25%	19%	12%	12%	0%	31%
D/BE	−17.05	−11.10	−7.91	−4.52	−2.81	−14.24	*(0.00)*	27%	19%	15%	6%	1%	33%
Panel D: Investor disagreement													
ADISP	−3.43	−6.25	−7.17	−8.01	−8.28	4.85	*(0.00)*	5%	15%	15%	18%	15%	9%
TURN	−3.32	−3.73	−7.97	−11.15	−16.69	13.37	*(0.00)*	3%	3%	13%	17%	24%	24%
Panel E: Difficulty in selling short													
IO	−10.75	−15.62	−12.42	−10.39	−8.28	−2.54	*(0.00)*	31%	35%	27%	28%	22%	9%
SIR	−6.68	−7.56	−10.98	−13.62	−17.60	11.00	*(0.00)*	19%	32%	50%	50%	49%	43%

Table X: Predictability of portfolio excess returns controlling for return comovement

This table reports the coefficient of the equal-weighted cross-firm valuation dispersion (CVD) by regressing annual excess returns of the two tech industry portfolios and the long-short portfolios on the lagged CVD while controlling for the return comovement with multifactors.

$$\text{3 factor:} R_t = c + d\mathrm{CVD_{t-1}} + \beta\mathrm{MKT_t} + s\mathrm{SMB_t} + h\mathrm{HML_t} + u_t, \tag{C-8}$$

$$\text{4 factor:} R_t = c + d\mathrm{CVD_{t-1}} + \beta\mathrm{MKT_t} + s\mathrm{SMB_t} + h\mathrm{HML_t} + m\mathrm{UMD_t} + u_t, \tag{C-9}$$

$$\text{ICAPM factor:} R_t = c + d\mathrm{CVD_{t-1}} + \beta\mathrm{MKT_t} + b\Delta\gamma_\mathrm{t} + e\Delta\eta_\mathrm{t} + u_t, \tag{C-10}$$

Where MKT is the market excess returns, SMB is the size factor returns, HML is the book-to-market factor returns, UMD is the momentum factor returns, $\Delta\gamma$ is the estimated innovation in the instantaneous real interest rate, and $\Delta\eta$ is the estimated innovation in the instantaneous market Sharpe ratio. The last two are factors in the ICAPM of Brennan, Wang, and Xia (2004). The two technology-related industries are HiTec and Telcm. The long-short portfolio is long on the lowest quintile and short on the highest quintile for the following firm characteristics: market beta (BETA), return volatility (VOL), firm age (AGE), earnings-to-book equity (E/BE), plant, property, and equipment (PPE), dividends-to-book equity (D/BE), institutional ownership (IO), short interest ratio (SIR), analyst forecast dispersion (ADISP), and trading volume (TURN). The coefficients of OLS regressions are reported. The one-tail p-values based on simulated distribution using the method by Nelson and Kim (1993) are reported in parenthesis. R-squares are adjusted for degree of freedom.

Panel A: Industries						
	3 factors		4 factors		ICAPM	
HiTec	−2.76	*(0.00)*	−2.54	*(0.00)*	−1.68	*(0.00)*
Telcm	−1.92	*(0.00)*	−1.99	*(0.00)*	−3.55	*(0.00)*
Panel B: Riskiness						
BETA	4.03	*(0.05)*	4.14	*(0.05)*	11.58	*(0.00)*
VOL	9.44	*(0.00)*	9.76	*(0.00)*	15.38	*(0.00)*
Panel C: Cash flow uncertainty						
AGE	−6.11	*(0.00)*	−6.10	*(0.00)*	−11.35	*(0.00)*
E/BE	−3.63	*(0.04)*	−3.91	*(0.03)*	−3.34	*(0.06)*
PPE	−3.24	*(0.04)*	−3.05	*(0.05)*	−10.73	*(0.00)*
D/BE	−6.39	*(0.00)*	−6.36	*(0.00)*	−11.89	*(0.00)*
Panel D: Investor disagreement						
ADISP	2.23	*(0.00)*	2.16	*(0.00)*	0.55	*(0.31)*
TURN	3.26	*(0.10)*	3.28	*(0.10)*	11.89	*(0.00)*
Panel E: Difficulty in selling short						
IO	−2.96	*(0.00)*	−2.51	*(0.05)*	−2.31	*(0.13)*
SIR	2.58	*(0.00)*	3.60	*(0.00)*	−0.09	*(0.43)*

Table XI: Firm valuation dispersion and common factor returns

This table reports the results by regressing the value-weighted σ^2(BM), σ^2(DP) , and σ^2(EP) on a set of common factors, squared common factors, and their corresponding value-weighted cross-sectional variance of firm growth rates. MKT is the market excess returns, SMB is the size factor, HML is the book-to-market factor, UMD is the momentum factor, $\Delta\gamma$ is the estimated innovation in the instantaneous real interest rate, and $\Delta\eta$ is the innovation in the instantaneous market Sharpe ratio. The last two are innovation in the state variables in the ICAPM of Brennan, Wang, and Xia (2004). The variables σ^2(ROE), σ^2(ΔD) , and σ^2(ΔE) are, respectively, the cross-sectional variance of log return-on-equity, dividend growth, and earnings growth. The OLS coefficients and t-statistics (in parenthesis) are reported. R-squares are adjusted for degree of freedom. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% level based on the two-tailed p-value.

Panel A: Dependent variable: σ^2(BM)

Intercept	MKT^2	SMB^2	HML^2	UMD^2	$\Delta\gamma^2$	$\Delta\eta^2$	MKT	SMB	HML	UMD	$\Delta\gamma$	$\Delta\eta$	σ^2(ROE)	Obs	R^2
67.41	0.05						−10.27***						17.51***	41	75%
(0.83)	*(0.31)*						*(−2.72)*						*(10.66)*		
21.25	0.06	0.12	10.54**				−12.74***						18.17***	41	77%
(0.24)	*(0.43)*	*(0.61)*	*(2.29)*				*(−3.36)*						*(11.37)*		
−6.69	−0.04	0.04	6.90**	1.17***			−2.61	7.53**					11.11***	41	92%
(−0.12)	*(−0.48)*	*(0.37)*	*(2.50)*	*(7.48)*			*(−0.99)*	*(−2.07)*					*(8.32)*		
−7.26	−0.03	0.02	6.25**	1.08***			−2.10	−5.66	0.27	−5.51			11.04***	41	92%
(−0.12)	*(−0.37)*	*(0.17)*	*(2.23)*	*(6.40)*			*(−0.79)*	*(−1.43)*	*(1.40)*	*(−1.08)*			*(8.21)*		
33.16	0.06				22.08*	−0.01	−12.68***				−87.23**	−1.06	17.59***	38	77%
(0.34)	*(0.37)*				*(1.75)*	*(−0.43)*	*(−3.29)*				*(−2.21)*	*(−0.97)*	*(10.67)*		

Panel B: Dependent variable: σ^2(DP)

Intercept	MKT^2	SMB^2	HML^2	UMD^2	$\Delta\gamma^2$	$\Delta\eta^2$	MKT	SMB	HML	UMD	$\Delta\gamma$	$\Delta\eta$	σ^2(ΔD)	Obs	R^2
−62.60	0.15						8.49						10.02***	41	88%
(−0.38)	*(0.52)*						*(−1.16)*						*(16.39)*		
−153.71	0.17	0.30	13.96				11.28						10.23***	41	88%
(−0.84)	*(0.58)*	*(0.74)*	*(1.50)*				*(−1.49)*						*(16.41)*		
−69.92	−0.02	0.03	6.86	2.01***			0.95	−18.59*					7.57***	41	92%
(−0.43)	*(−0.08)*	*(0.08)*	*(0.85)*	*(3.88)*			*(0.12)*	*(−1.76)*					*(8.84)*		
−64.36	−0.02	0.03	6.70	1.98***			1.06	−18.21	0.08	−2.45			7.58***	41	91%
(−0.35)	*(−0.07)*	*(0.09)*	*(0.79)*	*(3.45)*			*(0.13)*	*(−1.51)*	*(0.14)*	*(−0.16)*			*(8.54)*		
−99.09	0.09				32.68	−0.01	−10.17				−102.20	0.68	10.03***	38	87%
(−0.48)	*(0.27)*				*(1.26)*	*(−0.29)*	*(−1.28)*				*(−1.26)*	*(0.30)*	*(15.36)*		

Panel C: Dependent variable: σ^2(EP)

Intercept	MKT^2	SMB^2	HML^2	UMD^2	$\Delta\gamma^2$	$\Delta\eta^2$	MKT	SMB	HML	UMD	$\Delta\gamma$	$\Delta\eta$	σ^2(ΔE)	Obs	R^2
−147.48*	0.10						−9.90***						3.45***	41	89%
(−1.97)	*(0.71)*						*(−2.91)*						*(16.85)*		
−186.16**	0.10	0.15	7.15*				−11.32***						3.49***	41	89%
(−2.26)	*(0.77)*	*(0.79)*	*(1.68)*				*(−3.23)*						*(17.09)*		
−214.25***	0.01	0.04	6.97**	0.96***			−0.68	4.63					2.44***	41	95%
(−3.55)	*(0.09)*	*(0.30)*	*(2.31)*	*(4.08)*			*(−0.23)*	*(1.14)*					*(8.84)*		
−195.24***	−0.01	0.10	7.86***	1.07***			−1.48	1.44	−0.36*	5.15			2.47***	41	95%
(−2.96)	*(−0.09)*	*(0.73)*	*(2.63)*	*(4.50)*			*(−0.50)*	*(0.33)*	*(−1.69)*	*(0.92)*			*(9.13)*		
−185.20**	0.09				20.45*	0.00	−12.21***				−95.76***	−1.06	3.39***	38	90%
(−2.15)	*(0.65)*				*(1.85)*	*(0.08)*	*(−3.62)*				*(−2.76)*	*(−1.11)*	*(17.01)*		

Table XII: Predictability of market excess returns (1926-1962)

This table reports results by regressing market excess returns on the lagged cross-firm valuation dispersion (CVD), the detrended cross-sectional dispersion of book-to-market equity (*cd*(BM)), or the detrended cross-sectional dispersion of dividend-to-price ratios ((*cd*(DP)). The value-weighted *cd*(BM) (*cd*(DP)) are the residuals by regressing the value-weighted cross-sectional standard deviation of logarithmic BM (DP) from 1926–1962 on a time index (t-index) and the squared time index (t^2-index). The equal-weighted *cd*(BM) (*cd*(DP)) are the residuals by regressing the equal-weighted cross-sectional standard deviation of logarithmic BM (DP) from 1926–1962 on a time index. The dependent variables are the one-year ahead or three-year ahead market excess returns. The value-weighted (equal-weighted) predictor is used to predict the value-weighted (equal-weighted) returns. Overlapping observations are used for the three-year return regressions. The coefficients of OLS regressions are reported, below which in parenthesis are the one-tail p-values of the coefficients in the simulated distribution using the method by Nelson and Kim (1993). R-squares are adjusted for degree of freedom.

Panel A: CVD as a predictor				
	VW		EW	
	1-yr	3-yr	1-yr	3-yr
Intercept	7.53	22.35	10.54	31.18
	(0.02)	*(0.00)*	*(0.33)*	*(0.34)*
CVD	−13.30	−29.78	−7.12	−6.02
	(0.02)	*(0.00)*	*(0.33)*	*(0.34)*
R^2	16%	39%	0%	−2%
Panel B: *cd*(BM) as a predictor				
Intercept	7.53	22.35	10.54	31.18
	(0.06)	*(0.00)*	*(0.43)*	*(0.23)*
cd(BM)	−10.34	−31.04	5.52	11.90
	(0.06)	*(0.00)*	*(0.44)*	*(0.23)*
R^2	9%	42%	−1%	2%
Panel B: *cd*(DP) as a predictor				
Intercept	7.53	22.35	10.54	31.18
	(0.01)	*(0.00)*	*(0.01)*	*(0.02)*
cd(DP)	−13.35	−22.00	−17.48	−22.01
	(0.01)	*(0.00)*	*(0.01)*	*(0.02)*
R^2	16%	20%	16%	12%

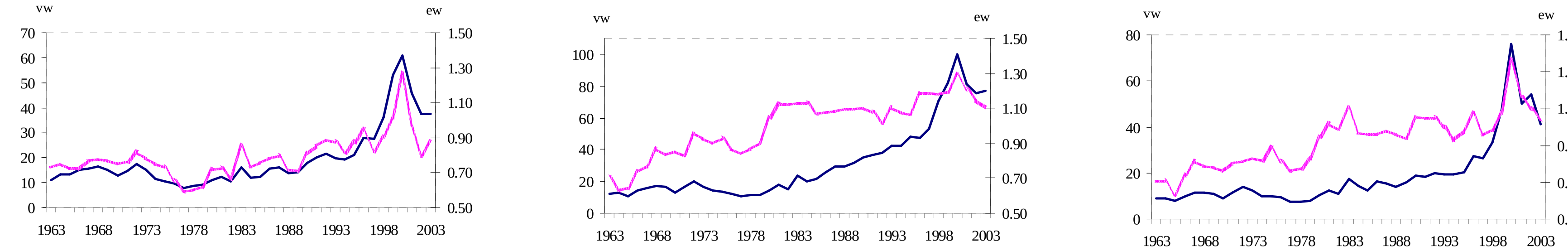
Table XIII: Predictability of market excess returns with undetrended CVD

This table reports results by regressing market excess returns on the undetrended lagged cross-firm valuation dispersion (CVD) while controlling for a linear or a nonlinear time trend. The value-weighted (equal-weighted) CVDraw is the first principal component of the value-weighted (equal-weighted) σ(BM), σ(DP), and σ(EP). The variable t-index is a time index that takes the value from 1, 2, 3 and the variable t^2-index is the squared t-index. The dependent variables are the one-year ahead or three-year ahead market excess returns. The value-weighted (equal-weighted) predictor is used to predict the value-weighted (equal-weighted) returns. Overlapping observations are used for the three-year return regressions. The coefficients of OLS regressions are reported, below which are the OLS t-statistics and in parenthesis are the one-tail p-values. The p-values of CVD$_{\text{raw}}$ are calculated based on the simulated distribution using the method by Nelson and Kim (1993). R-squares are adjusted for degree of freedom.

Panel A: Forecasting one-year ahead market excess returns									
		VW					EW		
Intercept	CVD$_{\text{raw}}$	t-index	t^2-index	R^2	Intercept	CVD$_{\text{raw}}$	t-index	t^2-index	R^2
−11.90	−9.73	0.75		9%	−24.73	−19.34	1.52		12%
−1.55	*−2.32*	*2.21*			*−1.87*	*−2.69*	*2.52*		
(0.00)	(0.16)	(0.03)			(0.31)	(0.00)	(0.02)		
−7.03	−24.01	−1.80	0.08	22%	−21.27	−19.11	1.07	0.01	10%
−0.96	*−3.72*	*−1.84*	*2.76*		*−1.32*	*−2.62*	*0.81*	*0.38*	
(0.02)	(0.00)	(0.07)	(0.01)		(0.69)	(0.01)	(0.42)	(0.70)	
Panel B: Forecasting three−year ahead market excess returns									
Intercept	CVD$_{\text{raw}}$	t-index	t^2-index	R^2	Intercept	CVD$_{\text{raw}}$	t-index	t^2-index	R^2
−30.68	−24.97	1.97		35%	−20.52	−23.97	2.03		8%
−3.10	*−4.54*	*4.52*			*−1.03*	*−2.24*	*2.19*		
(0.00)	(0.00)	(0.00)			(0.31)	(0.00)	(0.03)		
−21.38	−49.54	−2.61	0.15	54%	−9.50	−23.82	0.47	0.04	7%
−2.48	*−6.47*	*−2.19*	*4.03*		*−0.41*	*−2.22*	*0.25*	*0.91*	
(0.02)	(0.00)	(0.04)	(0.00)		(0.69)	(0.00)	(0.81)	(0.37)	

Figure 1: Time-series of firm valuation dispersion variables

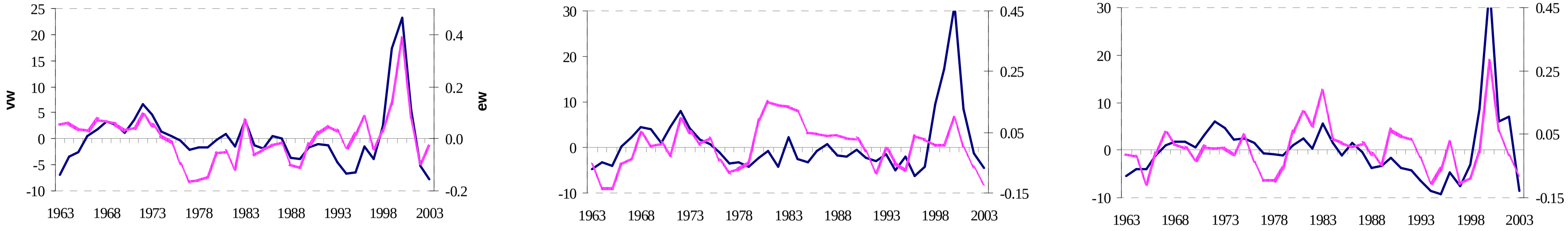
This figure plots the time series of the annual dispersion variables over the period 1963–2003. The variables σ(BM), σ(DP), and σ(EP) are, respectively, the cross-sectional standard deviation of logarithmic firm book-to-market equity, dividend-to-price ratios, and earnings-to-price ratio. Each variable is formed based on all available firms in NYSE, AMEX, and NASDAQ at the end of each June. The value-weighted *cd*(BM), *cd*(DP), and *cd*(EP) are, respectively, the residuals by regressing σ(BM), σ(DP), and σ(EP) on a time index and the squared time index. The equal-weighted *cd*(BM), *cd*(DP), and *cd*(EP) are, respectively, the residuals by regressing σ(BM), σ(DP), and σ(EP) on a time index. The variable CVD is defined as the first principal component of *cd*(BM), *cd*(DP), and *cd*(EP). CVD is standardized to have unit variance. Value-weighted measures are plotted with dark blue lines (darker) whil (lighter)



Panel A: Aggregate σ(BM)

Panel B: Aggregate σ(DP)

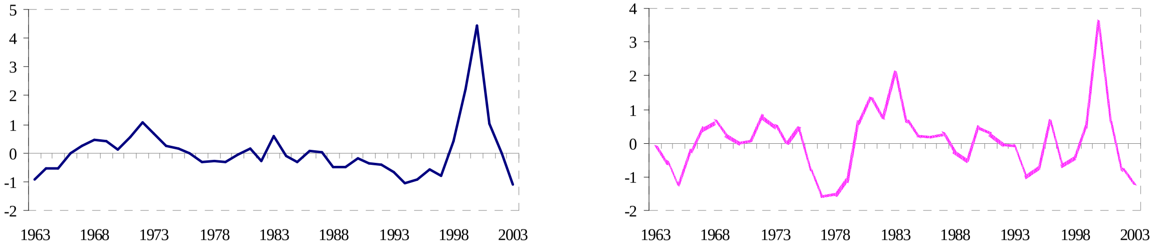
Panel C: Aggregate σ(EP)



Panel D: Aggregate *cd* (BM)

Panel E: Aggregate *cd* (DP)

Panel F: Aggregate *cd* (EP)



Panel G: Value-weighted CVD

Panel H: Equal-weighted CVD

Figure 2: Time-series of dispersion in growth rates

This figure plots the time series of the annual dispersion variables over the period 1963–2003. The variables $\sigma(\text{ROE})$, $\sigma(\Delta\text{D})$, and $\sigma(\Delta\text{E})$ are, respectively, the cross-sectional standard deviation of firm log return-on-equity, dividend growth, and earnings growth. Each variable is formed based on all available firms in NYSE, AMEX, and NASDAQ at the end of each June. Value-weighted measures are plotted with dark blue lines (darker) while equal-weighted measures are plotted with purple red lines (lighter).

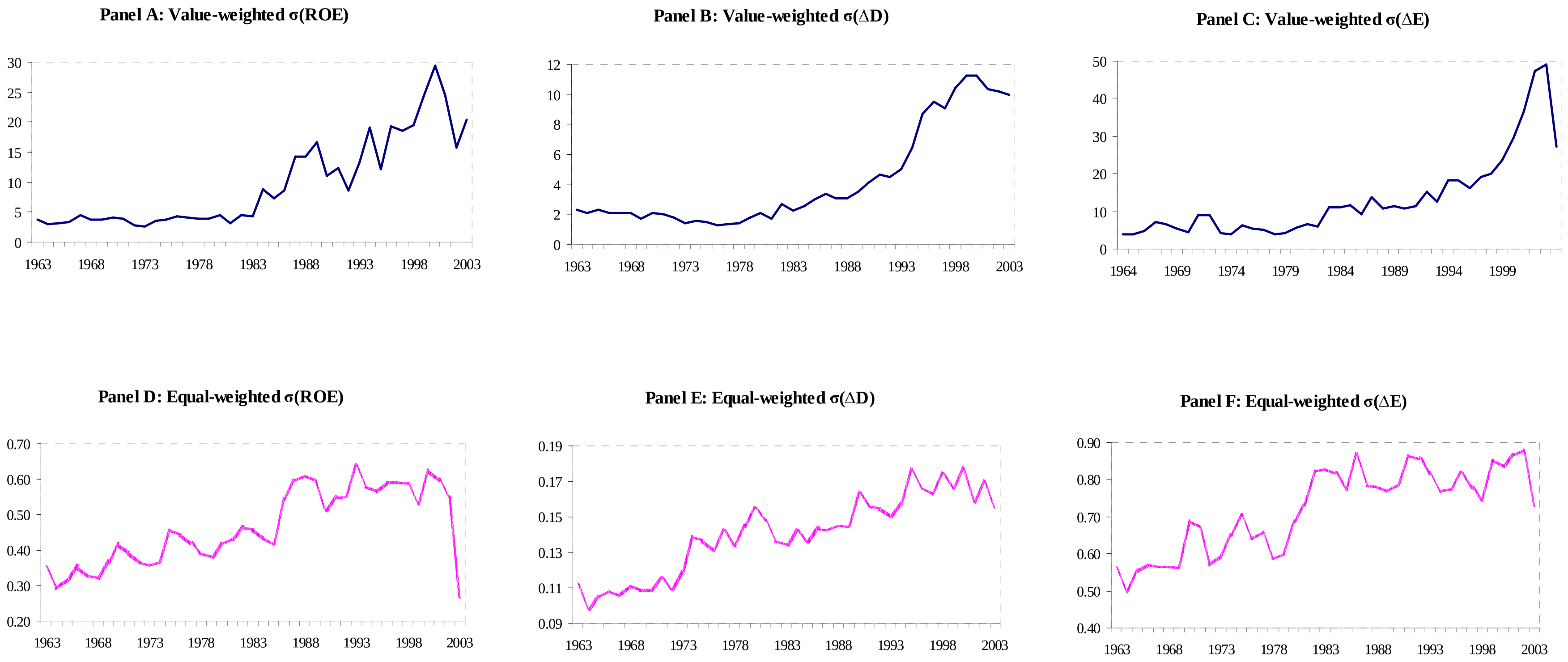


Figure 3: Market excess returns sorted based on CVD

This figure plots subsequent three 12-month market excess returns sorted based on whether the end-of-June CVD is below the mean (L) or above the mean (H). The first 12-month returns are plotted in blue bar. The second 12-month returns are plotted in red bar. The third 12-month returns are plotted in ivory bar. Panels A and B are sorted based on the CVD from 1963–2003. Panels C and D are sorted based on the CVD from 1963–1996.

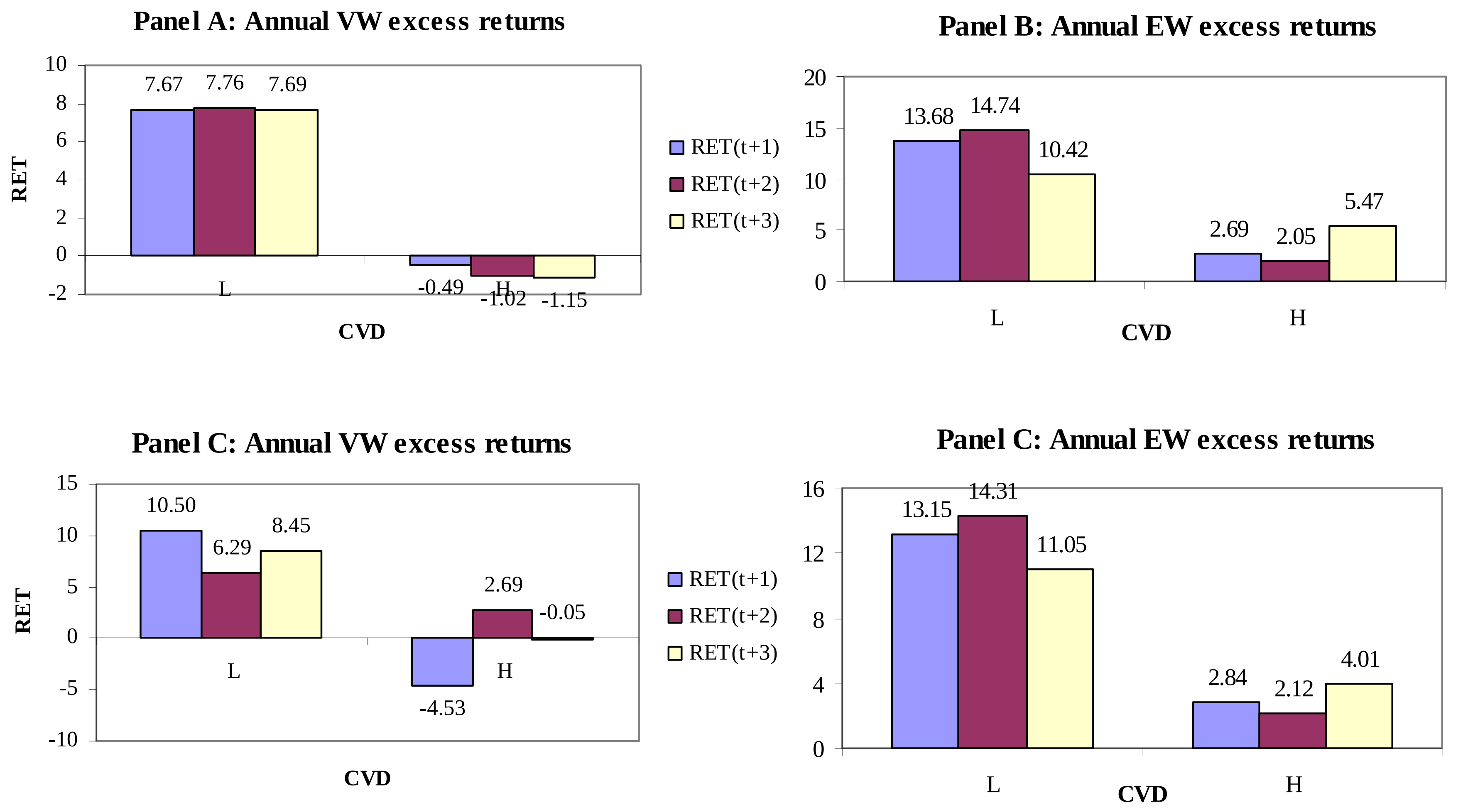


Figure 4: Portfolio excess returns of two-way sorts

This figure plots average annual value-weighted excess returns of industry portfolios or firm characteristic quintiles across high dispersion and low dispersion states over the period 1963–2003. The average excess returns with the beginning-of-period CVD greater than zero are plotted with blue bars (darker) while those with the beginning-of-period CVD less than zero are plotted with purple red bars (lighter). CVD is positive during 1967–75, 1980–87, 1990, 1991, 1996, and 1999–2001.