# Cross-Sectional Dispersion of Firm Valuations and Expected Stock Returns

Jiang, Danling

18 April 2008

# Cross-Sectional Dispersion of Firm Valuations and Expected Stock Returns

Danling Jiang *

The College of Business
Florida State University

April 18, 2008

*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 the following people for their helpful comments on early drafts of the paper: James Ang, Malcolm Baker, Jonathan Berk, Niki Boyson, Jeff Clark, Mike Cliff, Henrik Cronqvist, 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 seminars at The Ohio State University, Florida State University, University of San Diego, Sun Yat-Sen University in China, the 10th BDRM conference hosted by UCLA Anderson School of Management, and the FMA conferences at Chicago, IL. I appreciate Jonathan Berk, Kenneth French, Amit Goyal, Jeffrey Wurgler, and Yihong Xia for making their data and programs available online. All errors remain mine. E-mail: djiang@cob.fsu.edu; http://mailer.fsu.edu/∼djiang/.

Cross-Sectional Dispersion of Firm Valuations and Expected Stock Returns

**Abstract**

This paper develops two competing hypotheses for the relation between the cross-sectional standard deviation of logarithmic firm fundamental-to-price ratios ("dispersion") and expected aggregate returns. In models with fully rational beliefs, greater dispersion indicates greater risk and higher expected aggregate returns. In models with investor overconfidence, greater dispersion indicates greater mispricing and lower expected aggregate returns. Consistent with the behavioral models, the results show that (1) measures of dispersion are negatively related to subsequent market excess returns, (2) this negative relation is more pronounced among riskier firms, and (3) dispersion is positively related to aggregate trading volume, idiosyncratic volatility, and investor sentiment, and increases after good past market performance.

[Keywords] Return predictability, Dispersion, Overconfidence, Investor sentiment

*JEL Classification: G12, G14*

# I. Introduction

There is much evidence that fundamental-to-price ratios, such as the dividend-to-price and book-to-market equity ratios, forecast stock returns both at the aggregate and in the cross section.[1] This forecast ability, however, can arise either from time-varying risk premiums or from stock market mispricing.[2] Therefore, tests that use fundamental-to-price ratios often have both rational and behavioral explanations.

This paper takes a simple and fresh approach to test behavioral theories against rational theories at the aggregate level. Instead of focusing on the *first moment*—the aggregate fundamental-to-price ratios, this paper examines the *second moment*—the cross-sectional standard deviation of logarithmic firm valuation ratios, termed *dispersion*. Existing rational and behavioral models have opposite predictions regarding the relationship between dispersion and expected aggregate returns. Rational models of Sharpe (1964), Berk, Green, and Naik (1999), and Gomes, Kogan, and Zhang (2003) suggest a positive relation while behavioral theories of Daniel, Hirshleifer, and Subrahmanyam (2001) and Scheinkman and Xiong (2003) imply a negative one. Therefore, the dispersion-return relationship can help distinguish behavioral from rational explanations for the aggregate return predictability. This paper develops and tests these competing hypotheses. The empirical evidence supports the behavioral models.

Consider a single-period version of the Campbell and Shiller (1988) decomposition of the logarithmic dividend-to-price ratio with zero risk-free rates and no dividend growth, $(d-p)_i \simeq \mathrm{E}(r_i)$. Suppose expected log firm return, $\mathrm{E}(r_i)$, is approximately linearly related to expected log market returns, $\mathrm{E}(r_m)$, through beta, so $\mathrm{E}(r_i) \simeq \beta_i \mathrm{E}(r_m)$.[3] Taking cross-

[1] See Campbell and Shiller (1988), Fama and French (1988a), and Boudoukh, Michaely, Richardson, and Roberts (2007) for the dividend yield and the payout yield; Kothari and Shanken (1997), Pontiff and Schall (1998), and Fama and French (1992) for the book-to-market equity; Lamont (1998) for the earnings-to-price ratio, and Lee, Myers, and Swaminathan (1999) for the value-to-price ratio. Recent evidence includes Goyal and Welch (2007), Campbell and Thompson (2007), Cochrane (2007), among others.

[2] Examples of explanations based on risk premiums include Fama and French (1988a) and Berk (1995) and those based on mispricing include Lee, Myers, and Swaminathan (1999) and Daniel, Hirshleifer, and Subrahmanyam (2001).

[3] A linear relationship between beta and discrete returns under the (I)CAPMs approximately holds for log

sectional standard deviation of both sides in the decomposition equation yields $\sigma(d-p) = |\mathrm{E}(r_m)|\,\sigma(\beta) = \mathrm{E}(r_m)\sigma(\beta)$, with the second step based on a positive equity premium under a rational framework. Thus, when cross-sectional dispersion in beta is constant, the greater the expected equity premium, the larger the current cross-sectional dispersion in log firm valuation ratios—implying a positive dispersion-return relation.

When the dispersion of beta, however, changes over time, as considered in the conditional CAPMs of Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003), it tends to be positively correlated with the expected market return. This variation in the beta dispersion will reinforce the positive relation between firm valuation dispersion and expected market returns. Therefore, these rational pricing models predict that greater dispersion indicates greater risk and should positively forecast future aggregate returns.

In contrast, in a market with both rational and overconfident investors, firm valuation dispersion in the cross section is positively related to individual stock and market mispricing and negatively related to expected equity premium (Daniel, Hirshleifer, and Subrahmanyam (2001), Scheinkman and Xiong (2003)).[4] On the one hand, overconfident investors cause aggregate market price overreacting to market-wide information, and therefore increases aggregate mispricing, $|M|$, where $M$ can be positive (overpricing) or negative (underpricing). When expected aggregate return contains aggregate mispricing, then $\sigma(d-p)$ is partly determined by $|M|\,\sigma(\beta)$. So greater overconfidence leads to greater firm valuation dispersion. On the other hand, overconfident investors also tend to underestimate risk and lower market risk premiums and expected market returns. However, over a long period of time, overreaction to good and to bad news cancel out, leaving no net effect on average aggregate returns. Hence, greater overconfidence leads to both larger dispersion and lower average subsequent equity

returns with small risk-free rates or short intervals (e.g., Cochrane (2001) (pp.32, pp.103), Brennan, Wang, and Xia (2004)).

[4]Overconfidence refers to the tendency of investors to overestimate their own signal precision. It leads to overweighting private information or one's own judgement. As a behavioral trait, overconfidence has been widely identified by experimental studies in psychology (e.g., Alpert and Raiffa (1982), Lichtenstein, Fischhoff, and Phillips (1982), Camerer (1995)) and a growing finance literature has used overconfidence to explain a large set of stock market phenomena (e.g., Kyle and Wang (1997), Odean (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Scheinkman and Xiong (2003), Hong, Scheinkman, and Xiong (2006)).

premium.

In addition, when overreaction and mispricing also occur to firm-specific news, the effect of overconfidence on dispersion will be greater. In that case, firm valuation dispersion also reflects the dispersion of firm-specific mispricings. When overconfidence is stronger, firm mispricings, so firm valuations, will be more dispersed in the cross section. Built upon existing behavioral models, I derive the above effects and show that variation in overconfidence can induce a negative relation between dispersion and expected aggregate returns.

In the empirical tests, I form a composite measure of dispersion that incorporates cross-sectional standard deviations of three *logarithmic* firm valuation ratios: the book-to-market equity, dividend-to-price, and earnings-to-price ratios. I call this composite measure *cross-firm valuation dispersion* (CVD). Another measure, CVDr, is formed similarly but accounts for the cross-sectional differences in profitability or growth rates reflected in firm valuation ratios. Both dispersion measures are adjusted for an upward time trend, likely caused by the divergence of firm characteristics over time (Fama and French (2004)), and incorporated into the composite measures using only prior information that is available to the market at each point in time. Then I test whether the dispersion measures positively or negatively forecast subsequent aggregate returns.

My results are consistent with the behavioral models. I find that during 1963–2006 both dispersion measures are negative predictors of subsequent market excess returns at horizons as short as three months and as long as three years ahead. This forecast ability is, however, stronger for value-weighted excess returns than for equal-weighted ones. In addition, the dispersion measures are negatively correlated with aggregate log fundamental-to-price ratios, as measured relative to their long-run trend. When expected cash flows are given, a lower aggregate fundamental-to-price ratio generally indicates lower expected aggregate returns. Therefore, this evidence again suggests that expected equity premium tends to decrease when dispersion rises.

To further examine the behavioral explanation for the dispersion-return relationship, I provide three sets of additional evidence. First, I find that, supporting the behavioral models, the negative dispersion-return relationship is more pronounced among riskier firms, such as firms with higher beta or larger return volatility. This holds both before and after the controls for a set of common comovement in returns. Conditional on a high CVD state—indicating greater mispricing—a positive risk-return tradeoff is gone or even reverses out in the cross section.

Second, I find that the dispersion measures are positively correlated with other aggregate indicators of investor overconfidence, including aggregate trading volume, aggregate idiosyncratic volatility, and past market performance. These results are consistent with behavioral models (e.g., Odean (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Gervais and Odean (2001), and Scheinkman and Xiong (2003)) predicting that investor overconfidence generates excess trading and excess volatility, and that overconfidence tends to grow after past trading success, caused by biased self-attribution.[5]

Finally, I show an interesting link between the dispersion measures and the sentiment index of Baker and Wurgler (2006). My dispersion measures are highly positively correlated with the sentiment index. Additionally, the dispersion measures can forecast the sentiment index at one-year ahead. In particular, the peak of the dispersion measures occurs in 2000, which is also the peak of the sentiment index and that of the tech bubble in the late 90's.

My results have practical values for predicting aggregate returns. Goyal and Welch (2007) show that most, if not all, well-known aggregate predictors do not beat the historical mean equity premium in forecasting future market returns out of the sample.[6] Using a similar method, I find that the dispersion measures forecast the equity premium better than the historical mean premium in real time. In addition, trading strategies that are long on either the most or the least risky stocks, conditional on dispersion, beat a buy-and-hold market strategy by 118% to 260% during the period 1965–2006.

[5]Biased self-attribution refers to the tendency to attribute one's success to own ability but failure to bad luck.

[6]For dissenting opinions, see Campbell and Thompson (2007), Cochrane (2007), and Lettau and Van Nieuwerburgh (2007).

Related to the literature on the small value-spread (Brennan, Wang, and Xia (2004), Campbell and Vuolteenaho (2004)),[7] my results suggest that the ability of the value spread to forecast market returns may neither be unique to small firms nor be confined in short horizons. More importantly, I show in the model section that the value spread does not help disentangle risk from mispricing while the dispersion measures do. My findings also differ from those by Polk, Thompson, and Vuolteenaho (2006), who forecast aggregate returns using the slope coefficients from regressions of firm fundamental-to-price ratios on firm betas. Their measures, like the aggregate fundamental-to-price ratios, are directly linked to the expected market return, which do not help differentiate the behavioral from rational explanations. Also, Polk et al. show that the forecast power of their measures is strong from 1927–1965 but weak from 1965–2002, while mine are significant return predictors after 1965.

The remainder of the paper is organized as follows. In Section II, I present the models and develop the test hypotheses. Section III describes the data and the key variables. Section IV provides the empirical results. Section V summarizes and draws conclusions.

# II. The Model

In this section, I will first present the model with both rational and overconfident investors. The model with fully rational investors resembles the CAPM, and is developed as a special case of the model with overconfidence by setting the overconfidence level as zero.

The model setting is built upon that of Daniel, Hirshleifer, and Subrahmanyam (2001) (hereafter DHS (2001)), which ties a factor model with both investor overconfidence and multiple securities.[8] In the model, There are two equally populated groups of risk averse

[7] The small value-spread is defined as the difference of the logarithmic book-to-market equity between the value and the growth firms among small stocks. Cohen, Polk, and Vuolteenaho (2003) introduce the value spread across all firms, including small ones.

[8] The current model assumes risk aversion of rational investors as a form of limits to arbitrage. As an alternative approach, one can assume risk-neutral investors and add short-sale constraints to produce the negative dispersion-return relationship (Scheinkman and Xiong (2003)). In this approach, investors tend to bid up stock prices because they anticipate to sell the stock in the future to more optimistic investors at the presence of short-sale constraints. Thus, securities tend to be overpriced, so is the aggregate market, which leads to greater dispersion and lower expected returns.

investors, one rational, denoted as subscript $R$, and the other overconfident, denoted as subscript $C$.[9] Investors 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 $\theta_i$, 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 $\theta$ 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 $i$, $\epsilon_i$ is the firm-specific payoff, independently identically distributed as $N(0, 1/v^{\epsilon})$, and $\beta_i$ is the loading of the $i$th security on the factor $F$. The common factor $F$ is normally distributed as $N(0, 1/v)$. In addition, $\mathrm{E}(F\epsilon_i) = 0$. The security loading, $\beta$, takes the values of $\beta_1$,..., $\beta_N$. I normalize the factor $F$ to set the average $\beta$ as one and denote cross-sectional variance of $\beta$ as $\sigma^2(\beta)$. The values of $\bar{\theta}_i$, $\beta_i$, and the distribution of $\epsilon_i$ and $F$ are common knowledge, but the realizations of $\epsilon_i$ 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, for overconfident investors,

[9] The main conclusions of the model hold for any non-negligible fraction of overconfident population. As shown by DHS (2001), the equilibrium asset prices reflect the average overconfidence of investors. Thus, the assumption about the fraction of the overconfident population is not essential for deriving the main results.

the mean and variance of the common factor payoff conditioning 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, for rational investors, the conditional mean and variance are given by

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

Note that when good news arrives ($S > 0$), $\mu_C > \mu_R$, while when bad news arrives ($S < 0$), $\mu_C < \mu_R$. That is, the expected payoff is systematically mis-estimated by overconfident investors to the direction of the news. Also note that $\sigma_C^2 < \sigma_R^2$, suggesting that overconfident investors underestimate cash flow risks.

After receiving the private signals, each investor selects her portfolio to maximize a CARA utility function with a risk aversion coefficient of $A$ for date 2 consumption. Appendix A shows the solution to the equilibrium price of the market, $P_m = \mu_R + \gamma l - A\sigma_{RC}^2 Q$ (see Appendix A for definition of terms). For brevity, let $M$ denote the misestimation of expected aggregate cash flows, $\gamma l$, termed "cash flow mean bias," and $\pi$ denote the market risk premium, $A\sigma_{RC}^2 Q$. Then the aggregate price $P_m$ can be written as

$$P_m = \mu_R + M - \pi. \tag{3}$$

Accordingly, the rationally expected aggregate return, $\mathrm{E}(R_m)$, defined as $\mu_R - P_m$, is determined by both risk premium and cash flow mean bias:

$$\mathrm{E}(R_m) = \pi - M. \tag{4}$$

Similar to DHS (2001), valuation of asset $i$ is measured by $C_i - P_i$, where $C_i$ is a noisy measure of the true expected cash flow of asset $i$ conditional on the factor signal, and $P_i$ is equilibrium price of asset $i$. In the DHS (2001) model framework, return is measured in units of consumptions, so asset valuations are measured with the differences between cash flows and prices. Empirically, $C_i - P_i$ can be proxied by log fundamental-to-price ratios, which are approximately linearly related to expected log returns (Campbell and Shiller (1988)). It is easy to show that $P_i$ is determined by its unconditional expected cash flow and its cash flow sensitivity to the common factor.[10] Thus,

$$P_i = \bar{\theta}_i + \beta_i P_m, \qquad \text{and} \tag{5}$$

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

where $\nu_i$ is a firm-specific noise, independently identically distributed as $N(0, \sigma_\nu^2)$, and $E[\beta_i \nu_i] = 0$.

Combining equations (3), (5), and (6) yields a simple relationship between asset $i$'s valuation and the market's valuation,

$$C_i - P_i = \beta_i \mathrm{E}(R_m) = \beta_i(\pi - M) + \nu_i. \tag{7}$$

Taking the cross-sectional variance of both sides of equation (7) and define the adjusted cross-sectional variance $\hat{\sigma}^2(C-P) = \sigma^2(C-P) - \sigma_\nu^2$ produces

$$\hat{\sigma}(C-P) = \sigma(\beta) \left|\mathrm{E}(R_M)\right| = \sigma(\beta) \left|\pi - M\right|. \tag{8}$$

This equation shows that dispersion in firm valuations is determined by the absolute expected market returns multiplied by the dispersion of beta.

In equilibrium, both firm valuation dispersion and expected aggregate returns are endogenous. Thus, their relationship is usually determined by the exogenous variables that are shifted. In the model based on overconfidence, we consider the shift in the overconfidence parameter, $v^C$. In the model with full rationality, we examine the shift in either investor risk aversion, $A$, or cash flow volatility, $1/v$.

We first consider the shift in overconfidence. Variation in overconfidence is consistent with findings from experimental studies of psychology.[11] Prior overconfidence-based models

[10] This relationship can be easily 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, $\beta_i$ units of the market portfolio, and one unit of $\epsilon_i$, the price of asset $i$ should be equal to the sum of the prices of the three components in which $\epsilon_i$ is not priced since it is diversifiable.

[11] For example, it has been found that people tend to be more overconfident when individuals perform challenging judgment tasks (e.g., Lichtenstein, Fischhoff, and Phillips (1982)), and when feedback is vague and deferred (e.g., Einhorn (1980)).

(Odean (1998), Scheinkman and Xiong (2003) ) have studied the impacts of variation in overconfidence on price bubbles, excess volatility, and trading frequency.[12]

Through comparative statics we derive the following proposition:

**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. The expected aggregate return on average decreases.*

See Appendix A for proof of all propositions.

Thus, in a sufficiently overconfident market, larger firm valuation dispersion *on average* should precede lower aggregate returns.

The intuition is two-fold. First, consider the effects of overconfidence on firm valuation dispersion through two extreme cases. In Case 1 there is no risk premium ($\pi = 0$). Hence, $\hat{\sigma}(C-P) = \sigma(\beta)\,|M|$. When overconfidence rises, $|M|$ increases, and so is firm valuation dispersion. In Case 2 there is no mean bias ($M = 0$). Hence, $\hat{\sigma}(C-P) = \sigma(\beta)\pi$. When overconfidence is stronger, $\pi$ decreases, and so is dispersion. That is, both an increase in mean bias and a decrease in risk premium drive firm valuation dispersion. But the two effects are not equally strong. Appendix A shows that, as overconfidence rises, the increase in the mean bias dominates the decrease in risk premiums, owing to the stronger effect of overestimation of signal precision on the conditional mean than on the conditional variance. Therefore, when overconfidence becomes sufficiently strong, dispersion is determined by the aggregate mean bias.

Second, the decrease in risk premiums causes a decline in the average aggregate returns. Recall from equation (4) that the expected aggregate return conditional on private signals is $\pi - M$. On average the expected aggregate return is equal to $\pi$ because cash flow mean

[12]Empirical evidence in finance also suggests that overconfidence is not constant (Kumar (2008), Zhang (2006), Statman, Thorley, and Vorkink (2006)).

overestimation ($M > 0$) and underestimation ($M < 0$) are symmetric.[13] Thus, the upward bias and the downward bias cancel out. As a result, the mean bias effect does not influence the average aggregate return. The risk premium $\pi$, however, decreases with overconfidence due to the risk premium reduction effect. Therefore, the expected aggregate return is on average low when the overconfidence level is high. Taken together, variation in overconfidence can cause a negative dispersion-return relation.

So far the model considers only the impact of investor overreaction to market-wide news. But what if investors also receive and overreact to firm-specific news? A simple analysis shows that overreaction to firm-specific news can further enhance the negative dispersion-return relation. Denote firm-specific mispricing as $m_i$. Given the distribution of firm-specific payoffs, overconfidence and overreaction increase cross-sectional dispersion of firm-specific mispricing, $\sigma(m)$. Since in this case $\hat{\sigma}(C - P)$ is also determined by $\sigma(m)$, firm valuation dispersion is larger when overconfidence is stronger. Again, this implies a negative relation between dispersion and expected equity premium.

What about a rational framework? In this case, we consider how the shift in risk aversion or in factor cash flow volatility affects the dispersion-return relation. Now that overconfidence is absent ($v^C = v^R$) and the CAPM holds.[14] Since expectations are rational ($M = 0$), the adjusted firm valuation dispersion collapses to

$$\hat{\sigma}(C - P) = \sigma(\beta)\pi. \tag{9}$$

The above equation then 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 increases.*

Thus, if variation in expected returns is due to variation in rational risk premiums, greater

[13]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.

[14]DHS (2001) show that the unconditional CAPM holds when overconfidence is absent in their model.

firm valuation dispersion should on average precede higher aggregate returns. 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.

In the current model $\sigma(\beta)$ is assumed constant. In reality, however, $\sigma(\beta)$ can be time-varying. How does the change in $\sigma(\beta)$ affect the return predictability based on firm valuation dispersion? Existing rational models suggest that this change is likely to strengthen a positive dispersion-return relationship. For instance, in the model of Berk, Green, and Naik (1999), in each period firms choose investment projects based on the riskiness of the growth options. Other things equal, low beta projects are more attractive. Therefore, a larger cross-sectional dispersion of beta must indicate that firms are taking riskier projects at the margin. As a result, the aggregate market becomes riskier and is expected to deliver higher returns.

In Appendix C, using the same set of parameters as Berk, Green, and Naik (1999), I simulate a large panel of firm-month data. The three figures show that, under the Berk et al. (1999) model, firm valuation dispersion is positively related to expected aggregate returns and one-quarter/one-year ahead realized aggregate returns. Similar intuitions apply to the model of Gomes, Kogan, and Zhang (2003). Indeed, the model by Gomes et al. shows that (pp.723, Figure 5, and pp. 725, Figure 7) the cross-sectional dispersion of logarithmic book-to-market equity and the dispersion of firm betas have negative relationships with the log price-to-dividend ratio of the market. This implies a positive relationship between firm valuation dispersion, beta dispersion, and expected aggregate returns.

Taken together, Propositions 1 and 2 provide testable competing hypotheses; the over-confidence model predicts a negative dispersion-return relationship while the rational model predicts a positive relationship between the two. In contrast, the relationship between the value spread and the expected aggregate return carries the same sign under the two models. Let the average beta of the value firms be $\beta_v$ and that of the growth firms be $\beta_g$. When the number of firms in each group is sufficiently large, the value spread is

$$(C-P)_v - (C-P)_g = (\beta_v - \beta_g)\mathrm{E}(R_m). \tag{10}$$

Thus, the relationship between the value spread and the expected aggregate returns is determined by $\beta_v - \beta_g$ regardless of whether the variation in the expected return is caused by variation in rational risk premiums or in mispricing. A number of studies (e.g., Campbell and Vuolteenaho (2004)) find that growth stocks on average have higher beta than value stocks after 1963. Therefore, equation (10) predicts that the value spread should be negatively related to the expected aggregate return. This negative relationship between the value spread and expected aggregate returns, however, does not help differentiate a behavioral theory from a rational theory.

Furthermore, since $\beta$ is a multiplier on the expected aggregate return, the two hypotheses have implications for the cross section of stock returns.

**Proposition 3.**

*1. As overconfidence $v^C$ increases, ceteris paribus, the reduction in the expected returns of assets with higher betas are on average greater than those of assets with lower betas.*

*2. However, if investors are rational ($v^C = v^R$), as risk aversion $A$ or factor cash flow volatility $1/v$ increases, the increase in the expected return of assets with higher beta is greater than that of assets with lower betas.*

Thus, based upon the overconfidence model, the negative dispersion-return relation should be stronger among sets of high beta firms; when overconfidence reduces the 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 rational model predicts that high beta stocks should exhibit a stronger positive dispersion-return relationship.

Taken together, I develop two sets of competing hypotheses.

**Hypothesis I**

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

b. (The mispricing hypothesis) Firm valuation dispersion should be *negatively* related to subsequent aggregate stock returns.

**Hypothesis II**

a. (The risk hypothesis) The *positive* dispersion-return relation should be most pronounced among risky firms.

b. (The mispricing hypothesis) The *negative* dispersion-return relation should be most pronounced among risky firms.

It is worth remarking that the behavioral hypothesis does not exclude the influence of risk in the equity premium. Under the behavioral framework, a negative relation between dispersion and future aggregate returns only suggests the presence of mispricing and its stronger impact on the dispersion-return relation than risk.

# III. Data

The main sample includes common stocks (share code 10 and 11) listed in NYSE, AMEX, and NASDAQ from July 1963 to December 2006, excluding financials (SIC between 6000 and 6999) and utilities (SIC between 4900 and 4949). Stock trading data are obtained from the Center of Research in Securities Prices (CRSP). Accounting information is obtained from COMPUSTAT.

I calculate cross-sectional standard deviations of three logarithmic firm valuation ratios: log book-to-market equity $(b - p)$, log dividend-to-price $(d - p)$, and log earnings-to-price $(e - p)$; and denote the respective standard deviations as $\sigma(b - p)$, $\sigma(d - p)$, and $\sigma(e - p)$. These measures are later incorporated for a parsimonious composite measure. Following convention in the literature, book equity, dividends, and earnings in fiscal year end as of December of year $t - 1$ are used to form the dispersion measures from the end of June of year $t$ to the end of May of year $t + 1$. The annual dispersion measures are computed at the end of each June using the end-of-June market equity. The quarterly dispersion measures are computed at the end of each March, June, September, and December using the end-of-quarter market equity. Firm book equity is calculated following Polk, Thompson, and Vuolteenaho (2006). Market equity is the product of stock price and shares outstanding. Firm book-to-market equity is book equity over market equity. Firm dividend-to-price ratio is total dividends paid (data 21) over market equity. Firm earnings-to-price ratio 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.[15]

—–INSERT FIGURE 1 HERE—–

Panels A to C of Figure 1 plot the time series of the annual dispersion variables $\sigma(b-p)$, $\sigma(d-p)$, and $\sigma(e-p)$, measured at the end of each June. These series show an upward time trend. Preliminary analyses show that a time index accounts for 30-70% of the variation of these series. Common practice in the literature is to detrend a time series before making forecasts on returns (e.g., Baker and Wurgler (2006)). Before detrending, however, it is important to understand the source of the time trend to avoid potential data-mining.

According to Campbell and Shiller (1988) and Vuolteenaho (2000), log book-to-market equity $(b-p)$ and log dividend-to-price ratio $(d-p)$ contain not only a component of expected returns but also a component of expected growth rate. In particular, $b-p$ reflects the return on equity ($roe$) and $d-p$ captures the log growth rate of dividend payout ($\Delta d$) (see Appendix B for equations). Following a similar derivation, I show in Appendix B that the log earnings-to-price ratio $e-p$ is related to the growth rate of earnings ($\Delta e$). Accordingly, the cross-sectional dispersion in $b-p$, $d-p$, and $e-p$ should also reflect the dispersion in $roe$, $\Delta d$, and $\Delta e$ across firms. Hence, it is likely that the divergence of the firm expected growth rates and profitability cause a divergence of log firm valuation ratios over time. Indeed, Fama and French (2004)) find that the dispersion of firm growth rates and that of profitability have become larger.

To further verify my conjecture, for firms that are included in computing the firm valuation dispersion, I calculate cross-sectional standard deviation of the three log growth rates,

[15] 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 usually positive and should include a more complete set of firms in my dispersion measure.

denoted respectively as, $\sigma(roe)$, $\sigma(\Delta d)$, and $\sigma(\Delta e)$, in which $roe$ is defined as earnings over the average of current book equity and book equity lagged by one year. As shown in Panels D, E, and F of Figure 1, dispersion of firm growth rates exhibits a similar time trend to those in the valuation dispersion measures. That is, the time-trends in $\sigma(b-p)$, $\sigma(d-p)$, and $\sigma(e-p)$ are likely to be caused by a divergence in firm profitability over time that is not captured in my models. Therefore, filtering the time trend should not change the insights of the results to distinguish the two hypotheses.

—–INSERT FIGURE 2 HERE—–

To avoid a potential look-ahead bias (Brennan and Xia (2005)), my detrending method involves only prior information. Specifically, from June of year $t$ to May of $t+1$, I subtract each monthly dispersion measures by its 36-month moving average from year $t-3$ to $t-1$. That is, the long-run trend does not include dispersion within the most recent six to eighteen months. So this method preserves the short-term variation in dispersion, which intends to capture either the changes in general economic conditions or in misvaluation waves that usually last for a period of time.[16]

These detrended series, denoted as cd$(b-p)$, cd$(d-p)$, and cd$(e-p)$, are plotted on Panels A to C of Figure 2. Cross-firm valuation dispersion (CVD) is defined as the first principal component of cd$(b-p)$, cd$(d-p)$, and cd$(e-p)$, where the weight on each of the three components for CVD at month $t$ is generated using all data available in that month.[17] Again, the purpose is to avoid a possible look-ahead bias. The weights, in fact, have little variation over time and are usually around 0.34. I require at least 12-month observations to extract the principal component. Thus, the first observation of CVD starts from June 1965.

As discussed previously, log firm valuations also proxy for expected growth rates. To obtain a cleaner proxy for expected returns, I run a cross-sectional regression of each log

[16] Alternative definitions of long-term trends, such as the moving average over year $t$ to $t-5$, do not change the main findings. When the long-term trend is identified using the full sample, the main results are stronger.

[17] For example, CVD in June of 1986 is formed based on the weights of the three detrended series on the first principal component that are estimated using data from the beginning of the sample to June, 1986. These weighted are then updated in June of 1987 to form the annual CVD of that month.

firm valuation ratios on its corresponding growth rates, and defined the residuals as the residual log firm valuation ratios. I then define the cross-sectional standard deviations of the residuals as $\sigma(b-p)_{\mathrm{r}}$, $\sigma(d-p)_{\mathrm{r}}$, and $\sigma(e-p)_{\mathrm{r}}$. These series are detrended, as discussed previously, denoted as $\mathrm{cd}(b-p)_{\mathrm{r}}$, $\mathrm{cd}(d-p)_{\mathrm{r}}$, and $\mathrm{cd}(e-p)_{\mathrm{r}}$. Their first principal component, again using only prior information, is defined as the residual cross-firm valuation dispersion (CVDr).

—–INSERT TABLE 1 HERE—–

The summary statistics of all the dispersion measures are reported in Table 1. At each point in time CVD and CVDr are constructed to have zero mean and unit variance. Since the window is updated month by month, the full sample CVD and CVDr have means of −0.55 and −0.77, and standard deviations of 1.23 and 1.16. Comparing to the non-detrended dispersion and residual dispersion measures, the detrended dispersion measures, CVD, and CVDr have much smaller autocorrelations (around 0.50), suggesting less persistency after detrending, which can help make more reliable forecasts of future aggregate returns (Stambaugh (1986)).

The two CVDs have a high correlation of 0.95 at the monthly level (Panel C) and 0.94 at annual level (Panel B). They are also highly correlated with the underlying measures, shown in Panel B, suggesting that they capture major variation (which is around 85%) contained in the underlying measures. Their time series, plotted in Panels D of Figure 2, share remarkably similar trends and waves. Interestingly, three notable peaks in both CVD and CVDr, around the periods 1967 to 1968, 1980 to 1981, and 2000, all coincide with those in the sentiment index (Baker and Wurgler (2006)), as characterized by industry fads, hot IPOs, and technology bubbles, respectively.

# IV. Empirical Tests

I first present the main tests on the relationship between dispersion, aggregate valuation ratios, and aggregate returns. Then, I present additional evidence that dispersion is likely to measure aggregate investor overconfidence and sentiment.

## A. Forecasting aggregate returns

The rational hypothesis states a positive relation between dispersion and future aggregate returns, which implies a positive relation between dispersion and aggregate fundamental-to-price ratios. In contrast, the mispricing hypothesis predicts that dispersion is negatively related to future aggregate returns and to current aggregate fundamental-to-price ratios.

### 1. Correlation with aggregate valuation ratios

The aggregate book-to-market equity, dividends-to-price, and earnings-to-price ratios are calculated as the value-weighted average of firm valuation ratios including only firms that are used to compute the dispersion measure CVD. The results are similar if I use the aggregate valuation ratios based on S&P 500 index prices and fundamentals, reported by Goyal and Welch (2007). To linearize the relation between aggregate valuation ratios and expected equity premium, I take logarithm on the aggregate valuation ratios. In preliminary analyses, I find that all three aggregate log valuation ratios exhibit similar patterns: they rise during the 1960s until the early 1980s, then decline substantially through 2000, and then increase again. Similar to Goyal and Welch (2007), I also find that these aggregate valuation ratios have insignificant power to forecast equity premium after 1965.

The aggregate valuation ratios do not appear to have significant relationship with my dispersion measures, either. This lack of significant correlations can be caused by detrending the dispersion series. Recall that CVD and CVDr, by construction, reflect more short-term variation in dispersion relative to its long-run trend. The aggregate valuation ratios are dominated by their long-run trend. A diagnostic test, therefore, should use the change in the aggregate valuation ratios relative to its long-run trend to examine its correlation with CVD or CVDr that is independent from the common time trend. Thus, I subtract the monthly moving average from year $t-3$ to $t-1$ to aggregate valuation measures from June of $t$ to May of $t+1$, similar to that used for the dispersion measures, and denote the residual variables as $b-p$, $d-p$, and $e-p$. Then I test whether CVD and CVDr are positively or negatively correlated with $b-p$, $d-p$, and $e-p$.

Shown in Panel C of Table 1, CVD and CVDr are negatively correlated with aggregate $b-p$, $d-p$, and $e-p$, with all correlations significant at the 1% level.[18] In other words, firm valuations tend to become more dispersed when aggregate market prices relative to fundamentals become higher, following which the equity premium is likely to be lower. The evidence is consistent with the mispricing hypothesis that dispersion increases with aggregate market prices but inconsistent with the rational hypothesis that states the opposite.

## 2. Sorts

Next, I move to the correlation between CVD/CVDr and future aggregate returns. Before running predictive regressions, I first examine the aggregate return patterns conditioning on beginning-of-period CVD. The CRSP value-weighted and equal-weighted index compounded returns from 1963 to 2006 are used as aggregate stock returns. Based on whether the end-of-June CVD is above or below its historical mean up to each June, I sort the subsequent three consecutive 12-month market excess returns following June. In other words, the sorting results represent tradable portfolios at the point of the forecast. Shown in Figure 3, during the full sample period, in Panels A and B, when CVD is relatively low, the average next three 12-month value-weighted market excess returns are 6.82%, 8.80%, and 6.63%. In contrast, when CVD is relatively high, these average excess returns are 0.19%, 1.68%, and 2.95%, abysmal and even negative. The negative risk premium is at odds with the rational model where the risk premium must be positive.

——INSERT FIGURE 3 HERE——

Equal-weighted market excess returns exhibit similar patterns. Following low CVD periods the annual market excess returns are 9–10%, while following high CVD periods these

[18] I also find that non-detrended dispersion measures, $\sigma(b-p)$, $\sigma(d-p)$, and $\sigma(e-p)$, all have significant negative correlations with the non-detrended log aggregate valuation ratios. However, these correlations can be attributed an upward time trend in the raw dispersion measures and a downward time trend in the aggregate valuation measures from the early 1980s to 2000. Therefore, I focus on the tests with detrended measures.

excess returns are 3–5%. 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 and report the results in Panels C and D of Figure 3. The return patterns conditioning on CVD remain qualitatively similar. The unreported results conditioning on CVDr paint a similar picture. Thus, these results suggest a negative relationship between dispersion and future equity premium, which supports the mispricing hypothesis.

### 3. Univariate regressions

To formally evaluate the predictive power of CVD, I regress market excess returns on lagged CVD or CVDr. Both value- and equal-weighted market returns are used. I report results on three different return horizons: one quarter, one year, and three years. The results for other return horizons from one month to three years are qualitatively similar. For one-quarter-ahead returns, the predictors are updated at the end of each March, June, September, and December. For one-year-ahead or three-year-ahead returns, they are updated annually at the end of each June. Overlapping observations are used for the three-year return regressions. The OLS coefficients and $R^2$ are reported in Table 2. Since the OLS $t$-statistics may be subject to the small sample bias (Stambaugh (1986, 1999)), I report the one-tail $p$-value based on the simulated distribution of the predictive slope following Nelson and Kim (1993).[19]

——INSERT TABLE 2 HERE——

Panel A of Table 2 reports the predictive regression results using CVD and Panel B reports those with CVDr. As can be seen, the coefficients of CVD and CVDr are all negative. For value-weighted returns, the coefficients are −1.16%, −4.17% and −11.67%, all significant at the 5% level. Given a one standard deviation increase in CVD (which is 1.23), these coefficients imply a reduction of returns by 1.43%, 5.13%, and 14.35% at three horizons.

[19]Stambaugh 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 1, the first-order autocorrelations of CVD and CVDr are both 0.51, suggesting CVD is relatively persistent. Furthermore, preliminary analyses find that shocks to the expected market returns are correlated with shocks to CVD. Thus, the bias is likely to be present.

The $R^2$ range from 2-3% for one-quarter ahead returns to 14-21% for three-year ahead returns, suggesting non-negligible forecast power of CVD on future aggregate returns. The results using CVDr are similar, although with lower $R^2$ in regressions. For equal-weighted returns, the coefficients of CVD and CVDr are marginally significant for three-year returns but insignificant for one-quarter and one-year returns. However, they remain negative and economically significant.

### 4. Out-of-sample forecasts

A recent strand of literature debates about whether aggregate predictors have practical values for investors to forecast equity premium. In particular, Goyal and Welch (2007) argue that historical mean equity premium better forecasts aggregate excess returns than most, if not all, well-known aggregate return predictors. Given the debate, it is helpful to assess whether CVD and CVDr add power to forecast aggregate returns, relative to the historical mean equity premium.

Following Goyal and Welch, I perform the out-of-sample (OSS) forecasts based on CVD or CVDr. Specifically, in each quarter or year, I make forecasts of equity premium in the next quarter or year based on regression coefficients using information up to that point in time. The initial estimates are obtained using a 20-year period data, from June of 1965 to June of 1984. I compute the squared forecast error (SFE) from the model based on CVD or CVDr (the alternative), denoted as $\text{SFE}_A$, and from the model based on historical mean (the null), denoted as $\text{SFE}_N$. These squared forecast errors are cumulated period-by-period. Finally, the cumulative SFE difference is defined as the cumulative $\text{SFE}_N$ minus the cumulative $\text{SFE}_A$. That is, the better the alternative model performs, the lower the $\text{SFE}_A$ relative to $\text{SFE}_N$, and the greater the cumulative SFE difference.

—–INSERT FIGURE 4 HERE—–

The cumulative SFE differences based on CVD and CVDr are plotted in Figure 4, with quarterly forecasts in Panel A and annual forecasts in Panel B. An increase in the line means that the alternative model performs better, and vice versa when the line decreases. When the line stands above zero, the cumulative performance of the proposed model is superior to the null model that uses the historical mean equity premium. The picture is clear; the OSS performance of CVD-based forecasts beats that based on the historical mean equity premium. The cumulative SFE difference is always above zero, for both quarterly and annual forecasts, and tends to increase over time. The CVDr-based models perform slightly less strongly than CVD. In the 90's, the CVDr-based quarterly OSS performance is slightly beaten by the naive model. The forecast power of the null, however, is beaten by a significant margin by the end of 2006 for both the quarterly and annual CVDr-based forecasts.

In short, the results are consistent with the mispricing hypothesis that CVD is negatively related to future market excess returns. The results are inconsistent with the risk hypothesis which predicts a positive relationship between CVD and subsequent aggregate returns.

## B. Forecasting beta/volatility portfolio returns

Tests of Hypothesis I uncover a strong negative relationship between CVD and subsequent aggregate returns and suggest that mispricing is likely the dominate force in driving the dispersion-return relationship. I proceed to test Hypothesis II, which posits a stronger negative CVD-return relation among riskier stocks under the mispricing hypothesis but a stronger positive dispersion-return relation under the risk hypothesis.

For measures of risk, I use market beta (BETA) and total return volatility (VOL). BETA is estimated from a market model on 36 to 60 monthly available returns in the 5 years before July of year $t$. VOL is the standard deviation of 10 to 12 monthly returns before July of year $t$. After excluding stocks with prices less than \$5, I sort the stocks in the end of each June into equal-number BETA quintiles and VOL quintiles. Changing the breakpoints or including low price stocks does not materially change the results. I use CVD or CVDr, measured at the end of June of year $t$, to forecast the value-weighted monthly beta/volatility quintile returns from July of year $t$ to June of year $t+1$. Based on the mispricing hypothesis, I expect a stronger negative dispersion-return relation among firms with larger beta or greater return volatility. In contrast, based on the risk hypothesis, I expect a stronger positive dispersion-return relation among these sets of firms.

### 1. Two-way sorts

Before running regression, I conduct two-way sorts following Baker and Wurgler (2006) to examine which quintile of stocks are more affected by changes in CVD. Stocks are first sorted into BETA or VOL quintiles as described previously. Then the BETA and VOL quintiles monthly returns are sorted into two states based on the beginning-of-period CVD. In one state, CVD is above the sample mean while in the other CVD is below the sample mean.[20] The mispricing hypothesis predicts a greater return reduction among riskier stocks, when CVD switches from the low to the high state, while the risk hypothesis predicts the opposite.

—–INSERT TABLE 3 HERE—–

The two-way sort results, reported in Table 3, support the mispricing hypothesis. When moving from the low to the high CVD state, the return generally monotonically decreases across the quintiles; the decrease is the greatest among the highest BETA or VOL quintile and the smallest among the lowest quintiles. More interestingly, we see opposite return patterns across the two CVD states. During the low CVD states, the highest beta quintile *overperforms* the lowest beta quintile by 0.36% per month, or 4.32% per annum. In contrast, during the high CVD states, the highest beta quintile *underperforms* the lowest beta quintile by 0.43%, or 5.16% per annum. Similarly, the return differential between the highest VOL quintile and the lowest quintile is 0.69% per month during low CVD periods, but −0.16% per month during high CVD periods.

—–INSERT FIGURE 4 HERE—–

[20] According to Baker and Wurgler (2006), the two-way sorts help uncover the cross-sectional return patterns in the sample. However, since it involves using the full sample information to conduct the sorts, the long-short strategies presented in Table 3 are not tradable. To evaluate the validity of the long-short strategy, I also conduct sorts based on the change in CVD relative to its historical mean and find qualitatively similar results. This method, however, classify disproportionately more years into the low CVD states and less years into the high CVD states. Thus, I choose to report the sorts based on the full sample mean CVD.

The results imply a trading strategy conditional on the state of CVD: when the beginning-of-period CVD is relatively low, the portfolio is long on the highest BETA/VOL quintile; when CVD is high, it is long on the lowest BETA/VOL quintile, which I call "CVD-contingent" BETA/VOL strategies. These two strategies can potentially beat a naive buy-and-hold strategy on the market portfolio. To compare the three, I calculate the cumulative payoffs in the two strategies, assuming a $1,000 investment is made in June 1965. In the naive buy-and-hold strategy, the investment grows with returns on the value-weighted market portfolio. In the two CVD-based on strategies, when the end-of-June CVD is below its historical mean, the money is invested with the highest value-weighted BETA/VOL quintile; and, otherwise, with the lowest BETA/VOL quintile. I plot the cumulative dollar payoffs over time for the three strategies in Panel C of Figure 4. The results show that the CVD-contingent strategies significantly outperform the naive buy-and-hold strategy. Over a 42-year period, the $1000 grows to $65,417 with the market portfolio, to $142,630 with the CVD-contingent BETA portfolio, and to $233,502 with the CVD-contingent VOL portfolio. So, in real time, an investor could have been much better off through the CVD-contingent strategies.

The results confirm the predictions of the mispricing hypothesis: an increase in CVD causes a greater reduction in returns on riskier stocks. In particular, when beginning-of-period CVD is relatively low—suggesting less investor overconfidence and smaller aggregate mispricing—the cross section of return patterns reflect a positive risk-return trade-off. In contrast, when the beginning-of-period CVD is relatively high—indicating greater investor overconfidence and larger aggregate mispricing—the cross section of returns show no or even an anomalous negative risk-return trade-off.

### 2. Predicting portfolio returns using CVD

The results from the two-way sorts are further confirmed in regression analyses. To compare the difference in return sensitivity to change in dispersion, I regress monthly returns of each BETA or VOL quintile on CVD and then look for the cross-sectional patterns of the CVD coefficients and test whether the coefficient is significantly different between the two extreme quintiles.

———INSERT TABLE 4 HERE———

Table 4 confirms the mispricing hypothesis. The coefficients of CVD are all negative; they are the most negative for the highest BETA or VOL quintile and the least negative for the lowest quintile. The CVD coefficient on the long-short BETA portfolio is $-0.59$, suggesting that a one-standard-deviation increase in CVD on average reduces the highest beta quintile returns by 8.71% (i.e., $-0.59\% \times 1.23 \times 12$) per annum more than that of the lowest beta quintile returns. And the CVD coefficient on the long-short VOL portfolio is $-0.86$, representing even a greater impact of CVD on the cross section of VOL quintile returns.

Although CVD is shown to distinctively influence firms with different beta or volatility, it is possible that the differences in sensitivity are only due to differences in firms' exposure to other systematic risk factors. Therefore, I run the predictive regressions by adding multiple contemporaneous risk factors. I consider three multifactor models: the 3-factor model (Fama and French (1993)), the 4-factor model (Carhart (1997)), and the ICAPM of Brennan, Wang, and Xia (2004). There is disagreement in the literature about whether the Fama-French factors and the momentum factor are proxies for risk or mispricing (e.g., Fama and French (1995), Daniel and Titman (1997), and Daniel, Hirshleifer, and Subrahmanyam (2005)). These factors can possibly drive out the power of CVD to predict returns. However, if CVD captures more misvaluation than these factors, the lagged CVD may continue forecasting portfolio returns 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. Nevertheless, including these common factors makes the tests more conservative.

Specifically, I run the following time series regressions for each of quintiles and the long-short portfolios, a method used by prior studies (e.g., Lewellen (1999), Baker and Wurgler

(2006)).

$$R_t = c + d\mathrm{CVD}_{\mathrm{t-1}} + \beta\mathrm{MKT}_{\mathrm{t}} + s\mathrm{SMB}_{\mathrm{t}} + h\mathrm{HML}_{\mathrm{t}} + u_t, \tag{11}$$

$$R_t = c + d\mathrm{CVD}_{\mathrm{t-1}} + \beta\mathrm{MKT}_{\mathrm{t}} + s\mathrm{SMB}_{\mathrm{t}} + h\mathrm{HML}_{\mathrm{t}} + m\mathrm{UMD}_{\mathrm{t}} + u_t, \tag{12}$$

$$R_t = c + d\mathrm{CVD}_{\mathrm{t-1}} + \beta\mathrm{MKT}_{\mathrm{t}} + b\Delta\gamma_{\mathrm{t}} + e\Delta\eta_{\mathrm{t}} + u_t, \tag{13}$$

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 (11) controls for the FF 3 factors, Regression (12) controls for the 4 factors, and Regression (13) 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 2006 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 up to the end of 2001.

Panels B, C, and D in Table 4 report the results. It can be seen that the three specifications yield similar results. The CVD coefficients turn positive among quintiles with lower BETA or VOL and remain negative among quintiles with higher BETA or VOL. Nevertheless, the difference in CVD coefficients between the highest and the lowest quintiles all remain negative and statistically significant, suggesting a greater impact of CVD on riskier stocks even in the presence of contemporaneous common factors. Thus, the cross-sectional differences in the return sensitivities with respect to CVD are not fully explained by the cross-sectional differences in portfolios' exposure to a set of known common factors.

## 3. Correlation with indicators of aggregate overconfidence and investor sentiment

The results have shown that there exists a negative relation between firm valuation dispersion and expected aggregate returns, and it is stronger among riskier firms. These findings support the mispricing hypothesis that dispersion is an indicator of investor overconfidence and aggregate mispricing. Now I explore more direct links between dispersion and aggregate overconfidence and, more generally, investor sentiment.

As discussed previously, both theoretical and empirical research suggest that trading volume and idiosyncratic return volatility are associated with investor overconfidence. Thus, we expect to see greater aggregate trading volume and higher aggregate idiosyncratic volatility when CVD is larger. The monthly aggregate trading volume ($\mathrm{TURN}_a$) is calculated as logarithmic total monthly shares traded over total shares outstanding of all available firms in my sample. Due to the double counting problem in NASDAQ (e.g., Atkins and Dyl (1997)), the trading volume from NASDAQ firms is divided by 2. The monthly aggregate idiosyncratic volatility ($\mathrm{IVOL}_a$) is the equal-weighted average of firm idiosyncratic volatilities estimated through regressions of more than 17 daily returns on the FF 3 factors within the month (Ang, Hodrick, Xing, and Zhang (2006)). Similar to previous studies (e.g. Campbell, Lettau, Malkiel, and Xu (2001), Statman, Thorley, and Vorkink (2006)), I find an upward time trend in $\mathrm{TURN}_a$ and $\mathrm{VOL}_a$. To examine their correlation with dispersion that is independent from the time trend, I detrend the series with a moving average of past three years, a method similar to that used for the dispersion measures. I then examine the contemporaneous correlation between the dispersion and $\mathrm{TURN}_a$ or $\mathrm{VOL}_a$ through regressions.[21]

—–INSERT TABLE 5 HERE—–

Panel A of Table 5 reports the results of regressing monthly $\mathrm{TURN}_a$ and $\mathrm{VOL}_a$ on CVD or CVDr. The coefficients of CVD and CVDr are all positive and highly significant, suggesting that aggregate trading volume and idiosyncratic volatility tend to increase when firm valuations become more dispersion in the cross section, a phenomenon well predicted by the overconfidence-based models.

In addition, prior literature suggests that overconfidence tends to grow after past trading success due to biased self-attribution (DHS (1998), Gervais and Odean (2001)). For example, Statman, Thorley, and Vorkink (2006) find that trading volume tends to increase after a period of good market performance. Thus, we expect that dispersion, which captures aggregate overconfidence, should increase following a period of good past market returns. In

[21] I obtain similar results using Pearson correlations.

a similar spirit, I regress CVD and CVDr at the end of each June on the prior three or five-year CRSP value-weighted index returns. As reported in Panel B of Table 5, the coefficients on the past market returns are all positive and statistically significant, suggesting that CVD tends to rise after investors experience three to five year good stock market performance.

Finally, I examine the association between CVD/CVDr and investor sentiment index proposed by Baker and Wurgler (2006). The sentiment index incorporates six indicators of investor sentiment, ranging from the close-end fund discount, IPO markets, aggregate share turnover, to the dividend premium. Baker and Wurgler show that this sentiment measure is negatively related to subsequent aggregate returns. More importantly, it affects returns disproportionately more on firms that are difficult to value or to arbitrage. Similar to my two-way sorts, in particular, Baker and Wurgler show that there is a positive volatility-return relationship when the beginning-of-period sentiment is low, but a negative one when the initial sentiment is high. This finding is quite similar to my findings from the two-way sorts conditioning on CVD. Therefore, it would be interesting to find out to what extent my dispersion measures are correlated with the market-wide sentiment index.

Panel C of Table 5 reports the results of regressing the annual sentiment index (SENT) on the contemporaneous annual CVD or CVDr. The sentiment index is orthogonalized to economic indicators by Baker and Wurgler (2006). The results show a strong positive relation between dispersion and sentiment; the coefficients are all positive and significant. More interestingly, CVD and CVDr appear to forecast the sentiment index at one-year ahead. This predictive power remains even after I add the lagged SENT in the predictive regression. That is, high firm valuation dispersion tends to coincide with and precede periods of high investor sentiment. Overall, the evidence is consistent with the notion that firm valuation dispersion captures aggregate investor overconfidence and market-wide sentiment.

# V. Summary and Conclusion

Over the past two decades finance academia has seen the discovery of many return predictors. As evidence of stock return predictability mounts, the desire to more clearly understand predictability has increased as well. This paper attempts to contribute to the growing body of literature by documenting and understanding predictability, through developing and testing hypotheses from existing behavioral models against those from existing rational models.

I find that, over the period 1963 to 2006, cross-sectional dispersion of log firm valuations—including log book-to-market equity, log dividend-to-price, and log earnings-to-price ratios—is a negative predictor of subsequent equity premium of three months to three years ahead, and makes better forecasts than a naive model in real time. This predictability is more pronounced among riskier firms: those with larger beta and higher return volatility. More interestingly, in periods following high beginning-of-period dispersion, a positive risk-return trade-off in the cross section diminishes and even reverses out.

These results are less likely to be explained by a few risk-based models, and are better understood with existing overconfidence-based models that dispersion can increase with market mispricing, negatively forecast equity premium, and sometimes alter the positive risk-return relation in the cross section. While this paper sets up a simple framework to distinguish a behavioral explanation from a risk-based explanation for the aggregate return predictability, one should not interpret the evidence as incompatible with any risk-based models. Instead, my evidence only highlights the potential importance of understanding higher moments of firm characteristics from the perspective of rational risk premiums.

For behavioral research, this work uncovers interesting links between cross-sectional properties of firm valuations, aggregate trading, idiosyncratic volatility, past trading outcomes, and investor sentiment. Thus, higher moments of cross-sectional firm valuation ratios, as indicators of investor sentiment, may deserve a closer look in the future.

# References


Alpert, Murray, and Howard Raiffa, 1982, A Progress Report on the Training of Probability Assessors, in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds.: *Judgement under Uncertainty: Heuristics and Biases* (Cambridge University Press, Cambridge ).

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.

Atkins, Allen B., and Edward A. Dyl, 1997, Market structure and reported trading volume: NASDAQ versus the NYSE, *Journal of Financial Research* 20, 291–304.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor Sentiment and The Cross-Section of Stock Returns, *Journal of Finance* 61, 1645–1680.

Berk, Jonathan, 1995, A Critique of Size Related Anomalies, *Review of Financial Studies* 8, 275–286.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal Investment, Growth Options, and Security Returns, *Journal of Finance* 54, 1553–1607.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R. Roberts, 2007, On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing, *Journal of Finance* 62, 877–915.

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.

Camerer, Colin, 1995, Individual Decision Making, in John H. Kagel, and Alvin E. Roth, eds.: *The Handbook of Experimental Economics* (Princeton University Press, Princeton, NJ ).

Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu, 2001, Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk, *The Journal of Finance* 56, 1–43.

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 Samuel B. Thompson, 2007, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies* Forthcoming.

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.

Cochrane, John H., 2001, *Asset Pricing.* (Princeton University Press, Princeton and Oxford 41 William Stree, Princeton, New Jersey).

Cochrane, John H., 2007, The Dog That Did Not Bark: A Defense of Return Predictability, *Review of Financial Studies* forthcoming.

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* 54, 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.

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

Fama, Eugene F., and Kenneth R. French, 1988a, Dividend Yields and Expected Stock Returns, *Journal of Financial Economics* 22, 3–25.

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.

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.

Goyal, Amit, and Ivo Welch, 2007, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* forthcoming.

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

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, 2008, Hard-To-Value Stocks, Behavioral Biases, and Informed Trading, *Journal of Financial and Quantitative Analysis* forthcoming.

Kyle, Albert, and F. Albert Wang, 1997, Speculation Duopoly With Agreement to Disagree: Can Overconfidence Survive the Market Test?, *Journal of Finance* 52, 2073–2090.

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

Lee, Charles M.C., James Myers, and Bhaskaran Swaminathan, 1999, What is the Intrinsic Value of the Dow?, *Journal of Finance* 54, 1693–1741.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2007, Reconciling the Return Predictability Evidence, *Review of Financial Studies* forthcoming.

Lewellen, Jonathan, 1999, The Time-Series Relations among Expected Return, Risk, and Book-to-Market, *Journal of Financial Economics* 54, 5–43.

Lichtenstein, S., B. Fischhoff, and L.D. Phillips, 1982, Calibration of probabilities: The state of the art to 1980, in D. Kahneman, P. Slovic, and A. Tversky, eds.: *Judgment under uncertainty: heuristics and biases* (Cambridge University Press, Great Britain ).

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

Odean, Terrance, 1998, Volume, Volatility, Price and Profit When All Traders are Above Average, *Journal of Finance* 53, 1887–1934.

Polk, Christopher, Samuel Thompson, and Tuomo Vuolteenaho, 2006, Cross-Sectional Forecasts of the Equity Premium, *Journal of Financial Economics* 81, 101–141.

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

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

Sharpe, William F., 1964, Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk, *Journal of Finance* 19, 425–442.

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* 19, 1531–1565.

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

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

# Appendix A: Proofs

### Solution to equilibrium price

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

$$P = \mu_{RC} - A\sigma_{RC}^2 Q, \tag{A-1}$$

where $\mu_{RC} = \gamma\mu_C + (1-\gamma)\mu_R = \mu_R + \gamma l$, $\gamma = \dfrac{\sigma_R^2}{\sigma_R^2 + \sigma_C^2}$, and $\sigma_{RC}^2 = \dfrac{2\sigma_R^2\sigma_C^2}{\sigma_R^2 + \sigma_C^2}$.

In equation (A-1), the first term is an weighted average expected factor cash flows of two investor groups. The second term is the price discount for risk, in which the conditional factor volatility is determined by the average perceived volatility of two investor groups. Due to the single factor payoff structure, the factor portfolio can also be defined as the market portfolio, and then the price of the market portfolio $P_m$ is equal to $P$.[22]

Equation (A-1) 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 investor expected factor cash flow and the true expected factor cash flow (i.e., $M = \mu_{RC} - \mu_R = \gamma l$). The magnitude of mean bias, $|M|$, increases as overconfidence (measured by $v^C$) rises. I call this the "mean bias effect." Second, overconfidence leads to lower perceived cash flow volatility. Since overconfident investors overestimate the accuracy of their signals, they tend to underestimate risk and require a smaller risk premium than fully rational investors. I call this the "risk premium reduction effect." The combination of the mean bias effect and the risk premium reduction effect gives rise to a negative relation between firm valuation dispersion and expected aggregate returns.

### Proof of Proposition 1

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

Since $\sigma(\beta)$ is a constant, for brevity, I assume that it is equal to one in all proofs without loss of generality. Thus,

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

where $S^* = 2AQ\dfrac{v + v^R}{v(v^C - v^R)}$ and $M(S^*) = \pi$. When $S$ is greater than $S^*$ or less than $-S^*$, $|M| > \pi$, 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 $\omega = 1/(2v + v^C + v^R)$, and $M^* = M(S^*)$, taking derivative with respect to $v^C$ yields

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

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 within brace as $\Omega$, and taking derivative on $\Omega$ with respect to $v^C$ yields

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

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

### 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, i.e., $\pi = A\sigma_{RC}^2 Q$. It is easy to show that

$$\frac{\partial \pi}{\partial A} = \sigma_{RC}^2 Q > 0, \qquad \frac{\partial \pi}{\partial (1/v)} = 2v^2\pi\omega > 0.$$

That is, when the risk aversion $A$ or the factor cash flow volatility is greater ($v$ is smaller), the risk premium $\pi$ is greater. Since the expected aggregate return is equal to $\pi$ and $\hat{\sigma}(C-P) = \sigma(\beta)\pi$, 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 $\beta_1$ and $\beta_2$, 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. $\square$

## Appendix B. Decomposition of earnings-to-price ratio

According to Campbell and Shiller (1988) and Vuolteenaho (2000), log book-to-market equity $(b-p)$ and log dividend-to-price ratio $(d-p)$ can be approximated by

$$b - p_{t-1} = \sum_{j=0}^{\infty} \rho^j r_{t+j} + \sum_{j=0}^{\infty} \rho^j(-roe_{t+j}) + c; \tag{B-1}$$

$$d - p_{t-1} = \sum_{j=0}^{\infty} \rho^j r_{t+j} + \sum_{j=0}^{\infty} \rho^j(-\Delta d_{t+j}) + c; \tag{B-2}$$

$$\tag{B-3}$$

where $\rho$ is a time discount factor and usually set as close to one (Polk, Thompson, and Vuolteenaho (2006)).

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, and $p$ be the log price. Further, let $\delta$ denote the log earnings-to-price ratio, $\theta$ be the log dividend-to-earnings ratio, and $\Delta e$ be the log earnings growth rate. Let $r$ denote the log stock return, defined as

$$r_t = \log\left(\frac{P_t + D_t}{P_{t-1}}\right). \tag{B-4}$$

Substituting $\delta$, $\theta$, and $\Delta e$ into equation (B-4) yields

$$r_t = \Delta e_t + \delta_{t-1} + \theta_t + \log(\exp(-(\delta_t + \theta_t))). \tag{B-5}$$

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

$$r_t - \Delta e_t - (1-\rho)\theta_t = \delta_{t-1} - \rho\delta_t + \kappa_t. \tag{B-6}$$

Using the linear form in equation (B-6), 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:

$$\delta_{t-1} = \sum_{j=0}^{\infty} \rho^j r_{t+j} + \sum_{j=0}^{\infty} \rho^j(-\Delta e_{t+j}) + \sum_{j=0}^{\infty} \rho^j(\rho - 1)\theta_{t+j} + \sum_{j=0}^{\infty} \kappa_{t+j}. \tag{B-7}$$

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

$$\delta_{t-1} = \sum_{j=0}^{\infty} \rho^j r_{t+j} + \sum_{j=0}^{\infty} \rho^j(-\Delta e_{t+j}) + \sum_{j=0}^{\infty} \kappa'_{t+j}. \tag{B-8}$$

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 Aggregate Returns under Berk, Green, and Naik (1999)

These figures plot either expected or future realized aggregate returns against firm valuation dispersion under the model of Berk et al. (1999). The MatLab codes are obtained from Jonathan Berk's website. The programs generate a panel of 2000 firms over a 800-month period, using the same parameters as Berk et al. The first 200-month data are trimmed, following Berk el., to allow firms to mature. Then for each firm-month, I compute log book-to-market equity (b−p) and, for each month, define firm valuation dispersion as the cross-sectional standard deviation of (b−p), denoted as Stdev(b−p). The expected market return, denoted as E(Rm), is the cross-sectional average expected firm returns. The one-quarter (or one-year) ahead realized market return is the cumulative three-month (or 12-month) cross-sectional average of firm realized returns. The correlations between firm valuation dispersion and aggregate returns in all figures are all positive and significant at the 1% level. These figures show that, under the Berk et al. model, firm valuation dispersion should be positively related to future aggregate returns.

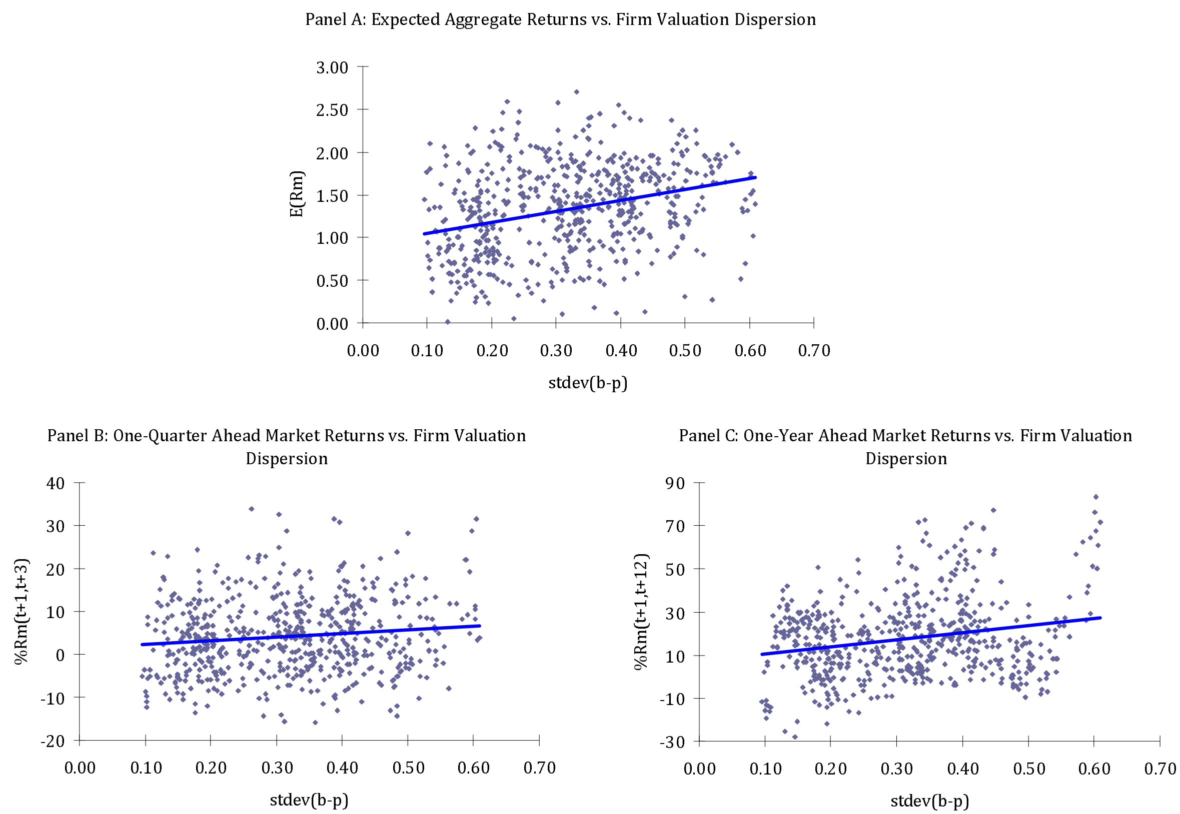


Table 1: Summary statistics

This table reports the descriptive statistics of the annual dispersion variables over the period 1963–2006. The variables $\sigma(b-p)$, $\sigma(d-p)$, and $\sigma(e-p)$ are the cross-sectional standard deviation of log firm book-to-market equity, log dividend-to-price ratios, and log earnings-to-price ratio, respectively. The variables $\sigma(b-p)_{\mathrm{r}}$, $\sigma(d-p)_{\mathrm{r}}$, and $\sigma(e-p)_{\mathrm{r}}$ are cross-sectional standard deviations of the residuals from cross-sectional regressions of each of the three log valuation ratios on their corresponding growth rates: log return on equity ($roe$), log dividend growth rate ($\Delta d$), and log earnings growth rate ($\Delta e$). The variables $\mathrm{cd}(b-p)$, $\mathrm{cd}(d-p)$, $\mathrm{cd}(e-p)$, $\mathrm{cd}(b-p)_{\mathrm{r}}$, $\mathrm{cd}(d-p)_{\mathrm{r}}$, and $\mathrm{cd}(e-p)_{\mathrm{r}}$ are the difference between $\sigma(b-p)$, $\sigma(d-p)$, $\sigma(e-p)$, $\sigma(b-p)_{\mathrm{r}}$, $\sigma(d-p)_{\mathrm{r}}$, and $\sigma(e-p)_{\mathrm{r}}$ of June of year $t$ and its average over a 36-month period from year $t-3$ to $t-1$, respectively. The variable CVD is defined as the first principal component of $\mathrm{cd}(b-p)$, $\mathrm{cd}(d-p)$, and $\mathrm{cd}(e-p)$, using information up to each month when CVD is formed. Similarly, CVDr is defined as the first principal component of $\mathrm{cd}(b-p)_{\mathrm{r}}$, $\mathrm{cd}(d-p)_{\mathrm{r}}$, and $\mathrm{cd}(e-p)_{\mathrm{r}}$, using only prior information. The variables $b-p$, $d-p$, and $e-p$ are defined as, respectively, log aggregate book-to-market equity, log dividend-to-price ratios, and log earnings-to-price ratios relative to its long-run trend. For example, aggregate book-to-market equity is the value-weighted average of firm book-to-market equity. For each month from June of year $t$ to May of $t+1$, the variable $b-p$ is defined as the log aggregate book-to-market equity ratio minus its monthly average from year $t-3$ to year $t-1$.

Panel A: Summary statistics of annual dispersion measures

| Variable | Obs | Mean | Std | AR(1) | Variable | Obs | Mean | Std | AR(1) |
|---|---|---|---|---|---|---|---|---|---|
| $\sigma(b-p)$ | 44 | 0.93 | 0.14 | 0.81 | $\mathrm{cd}(b-p)$ | 43 | −0.03 | 0.11 | 0.56 |
| $\sigma(d-p)$ | 44 | 1.00 | 0.16 | 0.90 | $\mathrm{cd}(d-p)$ | 43 | 0.00 | 0.07 | 0.60 |
| $\sigma(e-p)$ | 44 | 0.93 | 0.14 | 0.80 | $\mathrm{cd}(e-p)$ | 43 | −0.02 | 0.10 | 0.45 |
| CVD | 43 | −0.55 | 1.23 | 0.51 | | | | | |
| $\sigma(b-p)_{\mathrm{r}}$ | 43 | 0.84 | 0.12 | 0.79 | $\mathrm{cd}(b-p)_{\mathrm{r}}$ | 42 | −0.03 | 0.10 | 0.59 |
| $\sigma(d-p)_{\mathrm{r}}$ | 43 | 0.89 | 0.13 | 0.83 | $\mathrm{cd}(d-p)_{\mathrm{r}}$ | 42 | 0.00 | 0.07 | 0.51 |
| $\sigma(e-p)_{\mathrm{r}}$ | 43 | 0.81 | 0.12 | 0.74 | $\mathrm{cd}(e-p)_{\mathrm{r}}$ | 42 | −0.03 | 0.10 | 0.50 |
| CVDr | 42 | −0.77 | 1.16 | 0.51 | | | | | |

Panel B: Correlations of annual dispersion measures

| | $\mathrm{cd}(b-p)$ | $\mathrm{cd}(d-p)$ | $\mathrm{cd}(e-p)$ | CVD | $\mathrm{cd}(b-p)_{\mathrm{r}}$ | $\mathrm{cd}(d-p)_{\mathrm{r}}$ | $\mathrm{cd}(e-p)_{\mathrm{r}}$ | CVDr |
|---|---|---|---|---|---|---|---|---|
| $\mathrm{cd}(b-p)$ | 1.00 | | | | | | | |
| $\mathrm{cd}(d-p)$ | 0.66 | 1.00 | | | | | | |
| $\mathrm{cd}(e-p)$ | 0.92 | 0.74 | 1.00 | | | | | |
| CVD | 0.89 | 0.84 | 0.93 | 1.00 | | | | |
| $\mathrm{cd}(b-p)_{\mathrm{r}}$ | 0.96 | 0.62 | 0.87 | 0.85 | 1.00 | | | |
| $\mathrm{cd}(d-p)_{\mathrm{r}}$ | 0.70 | 0.81 | 0.78 | 0.81 | 0.68 | 1.00 | | |
| $\mathrm{cd}(e-p)_{\mathrm{r}}$ | 0.93 | 0.71 | 0.95 | 0.91 | 0.92 | 0.72 | 1.00 | |
| CVDr | 0.89 | 0.71 | 0.89 | 0.94 | 0.89 | 0.78 | 0.91 | 1.00 |

Panel C: Correlations with monthly log aggregate valuation ratios

| | CVD | CVDr | $b-p$ | $d-p$ | $e-p$ |
|---|---|---|---|---|---|
| CVD | 1.00 | | | | |
| CVDr | 0.95 | 1.00 | | | |
| $b-p$ | −0.26 | −0.30 | 1.00 | | |
| $d-p$ | −0.20 | −0.24 | 0.91 | 1.00 | |
| $e-p$ | −0.15 | −0.17 | 0.88 | 0.88 | 1.00 |

Table 2: Predictability of market excess returns

This table reports results by regressing market excess returns on the lagged cross-firm valuation dispersion (CVD and CVDr). 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. 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 two-tailed OLS $p$-values for the intercepts and the one-tailed $p$-values for CVD/CVDr based on the randomization method by Nelson and Kim (1993). R-squares are adjusted for degree of freedom. This table shows that CVD and CVDr are negative predictors of future aggregate excess returns.

Panel A: CVD as a predictor

| | Value-weighted | | | Equal-weighted | | |
|---|---|---|---|---|---|---|
| | 3-mon | 1-yr | 3-yr | 3-mon | 1-yr | 3-yr |
| Intercept | 1.43 | 2.00 | 7.91 | 2.57 | 5.16 | 16.38 |
| | (0.02) | (0.44) | (0.10) | (0.00) | (0.17) | (0.02) |
| CVD | −1.16 | −4.17 | −11.67 | −1.00 | −4.39 | −13.19 |
| | (0.02) | (0.03) | (0.00) | (0.11) | (0.09) | (0.01) |
| $R^2$ | 3% | 8% | 21% | 1% | 4% | 13% |
| Obs | 166 | 42 | 40 | 166 | 42 | 40 |

Panel B: CVDr as a predictor

| | Value-weighted | | | Equal-weighted | | |
|---|---|---|---|---|---|---|
| | 3-mon | 1-yr | 3-yr | 3-mon | 1-yr | 3-yr |
| Intercept | 1.29 | 1.13 | 5.82 | 2.39 | 4.44 | 15.71 |
| | (0.04) | (0.70) | (0.29) | (0.01) | (0.29) | (0.04) |
| CVDr | −0.96 | −4.17 | −10.53 | −0.66 | −3.57 | −6.94 |
| | (0.05) | (0.05) | (0.00) | (0.23) | (0.17) | (0.11) |
| $R^2$ | 2% | 7% | 14% | 0% | 1% | 2% |
| Obs | 162 | 41 | 39 | 165 | 41 | 39 |

Table 3: Portfolio Excess Returns from Two-way sorts

This table reports the average value-weighted monthly excess returns of the two-way sorts based on BETA/VOL and cross-firm valuation dispersion (CVD). At the end of each June, stocks are first sorted into BETA or VOL quintiles and the value-weighted monthly quintile returns from July of year $t$ to June of year $t+1$ are calculated. Depending on whether CVD at the end of June of year $t$ is above or below the sample mean, each BETA/VOL quintile is further sorted into two groups. CVD is above the mean (CVD–H) for 1966–1968, 1972–1973, 1979–1983, 1987, 1990–1991, 1996, 1999–2000. Return differentials refer to the quintile return differences between the low CVD states (CVD–L) and the high CVD states (CVD–H). The long-short portfolios (H−L) are long on the highest BETA or VOL quintile and short on the lowest BETA or VOL quintile. All returns are in percent. This table shows that, when beginning-of-period CVD is relatively low, risk is positively related to the cross section of stock returns. However, when CVD is relative high, the risk-return tradeoff becomes negative.

| | L | 2 | 3 | 4 | H | H−L |
|---|---|---|---|---|---|---|
| | | | | BETA | | |
| Average | 0.58 | 0.56 | 0.49 | 0.54 | 0.64 | 0.06 |
| CVD–L | 0.70 | 0.73 | 0.77 | 0.85 | 1.06 | 0.36 |
| CVD–H | 0.39 | 0.27 | 0.04 | 0.04 | −0.04 | −0.43 |
| Differential | 0.31 | 0.46 | 0.72 | 0.81 | 1.09 | 0.78 |
| | | | | VOL | | |
| Average | 0.54 | 0.49 | 0.56 | 0.69 | 0.90 | 0.36 |
| CVD–L | 0.75 | 0.73 | 0.81 | 1.02 | 1.44 | 0.69 |
| CVD–H | 0.22 | 0.10 | 0.15 | 0.17 | 0.05 | −0.16 |
| Differential | 0.53 | 0.62 | 0.65 | 0.85 | 1.38 | 0.85 |

Table 4: Predictability of portfolio excess returns

This table reports the regression results of value-weighted monthly excess returns of beta (BETA) and volatility (VOL) quintiles on the lagged CVD, with and without controls for the return comovement with a set of well-known common factors, including market excess returns (MKT), the size factor (SMB), the book-to-market factor (HML), the momentum factor (UMD), the innovation in the instantaneous real interest rate ($\Delta\gamma$), and the instantaneous market Sharpe ratio ($\Delta\eta$), obtained from French's and Xia's websites. The long-short portfolios (H−L) are long on the highest BETA or VOL quintile and short on the lowest BETA or VOL quintile. The coefficients on CVD are reported. The one-tail $p$-values based on the randomization method by Nelson and Kim (1993) are reported in parenthesis. R-squares are adjusted for degree of freedom. This table shows that the negative relation between CVD and future stock returns is more pronounced among riskier firms.

| | L | 2 | 3 | 4 | H | H−L | L | 2 | 3 | 4 | H | H−L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BETA | | | | | | VOL | | | | | |
| Panel A: $R_{i,t+1} = \alpha + \eta \mathrm{CVD}_t + \epsilon_i$ | | | | | | | | | | | | |
| $\eta$ | −0.08 | −0.20 | −0.33 | −0.41 | −0.67 | −0.59 | −0.20 | −0.35 | −0.39 | −0.55 | −0.86 | −0.66 |
| $p$-value | (0.31) | (0.12) | (0.05) | (0.05) | (0.02) | (0.01) | (0.11) | (0.04) | (0.06) | (0.03) | (0.01) | (0.01) |
| $R^2$ | 0% | 0% | 0% | 1% | 1% | 1% | 0% | 1% | 0% | 1% | 1% | 1% |
| Panel B: Controlling for the 3 factors: $R_t = \alpha + \eta \mathrm{CVD}_{\mathrm{t}-1} + \beta \mathrm{MKT}_{\mathrm{t}} + s \mathrm{SMB}_{\mathrm{t}} + h \mathrm{HML}_{\mathrm{t}} + u_t$ | | | | | | | | | | | | |
| $\eta$ | 0.16 | 0.08 | 0.02 | 0.02 | −0.10 | −0.26 | 0.09 | 0.01 | 0.04 | −0.04 | −0.26 | −0.35 |
| $p$-value | (0.99) | (0.90) | (0.59) | (0.62) | (0.16) | (0.04) | (0.98) | (0.56) | (0.73) | (0.34) | (0.02) | (0.01) |
| $R^2$ | 79% | 86% | 91% | 92% | 89% | 65% | 92% | 93% | 92% | 91% | 86% | 67% |
| Panel C: Controlling for the 4 factors: $R_t = \alpha + \eta \mathrm{CVD}_{\mathrm{t}-1} + \beta \mathrm{MKT}_{\mathrm{t}} + s \mathrm{SMB}_{\mathrm{t}} + h \mathrm{HML}_{\mathrm{t}} + m \mathrm{UMD}_{\mathrm{t}} + u_t$ | | | | | | | | | | | | |
| $\eta$ | 0.15 | 0.08 | 0.02 | 0.02 | −0.09 | −0.24 | 0.09 | 0.00 | 0.04 | −0.04 | −0.26 | −0.35 |
| $p$-value | (0.99) | (0.90) | (0.61) | (0.66) | (0.19) | (0.04) | (0.97) | (0.55) | (0.73) | (0.34) | (0.02) | (0.02) |
| $R^2$ | 79% | 86% | 91% | 92% | 89% | 67% | 92% | 93% | 92% | 91% | 86% | 66% |
| Panel D: Controlling for the ICAPM factors: $R_t = \alpha + \eta \mathrm{CVD}_{\mathrm{t}-1} + \beta \mathrm{MKT}_{\mathrm{t}} + b \Delta\gamma_{\mathrm{t}} + e \Delta\eta_{\mathrm{t}} + u_t$ | | | | | | | | | | | | |
| $\eta$ | 0.22 | 0.17 | 0.02 | −0.14 | −0.26 | −0.48 | 0.06 | 0.01 | −0.09 | −0.19 | −0.28 | −0.34 |
| $p$-value | (0.01) | (0.02) | (0.40) | (0.07) | (0.05) | (0.02) | (0.15) | (0.43) | (0.14) | (0.09) | (0.08) | (0.09) |
| $R^2$ | 78% | 86% | 92% | 90% | 80% | 35% | 90% | 93% | 90% | 81% | 73% | 28% |

Table 5: Dispersion, volume, idiosyncratic volatility, and investor sentiment

Panel A reports the regression results of the monthly aggregate trading volume and aggregate idiosyncratic volatility on the end-of-month cross-firm valuation dispersion (CVD). The monthly aggregate trading volume ($\text{TURN}_a$) is calculated as the total trading volume over the end-of-month total shares outstanding of all available firms in the sample. Trading volume is divided by 2 for NASDAQ firms. The aggregate idiosyncratic return volatility ($\text{VOL}_a$) is the average firm daily idiosyncratic return volatilities within the month, in which idiosyncratic returns are calculated as the residuals from the Fama-French 3 factor model. Panel B reports the annual regression results of the end-of-June CVD or CVDr on the past value-weighted market returns, Rm, from year $t-3$ to $t-1$, or from year $t-5$ to $t-1$. Panel C reports the annual regression of investor sentiment index (SENT) (Baker and Wurgler 2006) on CVD/CVDr or lagged CVD/CVDr. Two of the predictive regressions also control for the lagged sentiment index. For contemporaneous variables or lagged returns, the two-tailed OLS $p$-values are reported in parenthesis below the coefficients. For lagged CVD/CVDr, the one-tailed $p$-values are reported based on the randomization method by Nelson and Kim (1993). R-squares are adjusted for degree of freedom. This table shows that CVD and CVDr are positively related to aggregate trading volume, aggregate idiosyncratic volatility, investor sentiment index, and increase as stock markets perform well in the past.

Panel A: Explaining aggregate idiosyncratic volatility and trading volume

| | Dependent variable | | | |
|---|---|---|---|---|
| | $\text{IVOL}_t$ | | $\text{TURN}_t$ | |
| $\text{CVD}_t$ | 0.28 | | 0.05 | |
| | (0.00) | | (0.00) | |
| $\text{CVDr}_t$ | | 0.31 | | 0.05 |
| | | (0.00) | | (0.00) |
| Obs | 494 | 482 | 494 | 482 |
| $R^2$ | 23% | 25% | 7% | 7% |

Panel B: Predicting $\text{CVD}_t$ or $\text{CVDr}_t$

| | $\text{CVD}_t$ | | $\text{CVDr}_t$ | |
|---|---|---|---|---|
| $\text{Rm}(t-3, t-1)$ | 0.016 | | 0.017 | |
| | (0.04) | | (0.02) | |
| $\text{Rm}(t-5, t-1)$ | | 0.017 | | 0.019 |
| | | (0.01) | | (0.00) |
| Obs | 42 | 42 | 41 | 41 |
| $R^2$ | 7% | 15% | 11% | 22% |

Panel C: Explaining and Predicting Investor Sentiment

| | $\text{SENT}_t$ | | $\text{SENT}_{t+1}$ | | $\text{SENT}_{t+1}$ | |
|---|---|---|---|---|---|---|
| CVDt | 0.32 | | 0.36 | | 0.20 | |
| | (0.01) | | (0.00) | | (0.04) | |
| CVDrt | | 0.27 | | 0.33 | | 0.18 |
| | | (0.04) | | (0.01) | | (0.09) |
| SENTt | | | | | 0.50 | 0.53 |
| | | | | | (0.00) | (0.00) |
| Obs | 41 | 40 | 40 | 39 | 40 | 39 |
| $R^2$ | 14% | 8% | 19% | 13% | 40% | 38% |

Figure 1: Dispersion of Firm valuations and firm profitabilities

This figure plots the time series of the annual dispersion variables of firm valuations, profitabilities, and growth rates over the period 1963–2006. The variables $\sigma(roe)$, $\sigma(\Delta d)$, and $\sigma(\Delta e)$ are the cross-sectional standard deviation of firm log return-on-equity ratios, log dividend growth rates, and log earnings growth rates, respectively. The variables $\sigma(b-p)$, $\sigma(d-p)$, $\sigma(e-p)$, $\sigma(\text{b-p})_{\text{r}}$, $\sigma(d-p)_{\text{r}}$, and $\sigma(e-p)_{\text{r}}$ are measures of firm valuation dispersion, defined in Table 1. This figure shows that cross-sectional standard deviations of log firm valuations share a similar time trend to cross-sectional dispersion of firm profitability or growth rates.

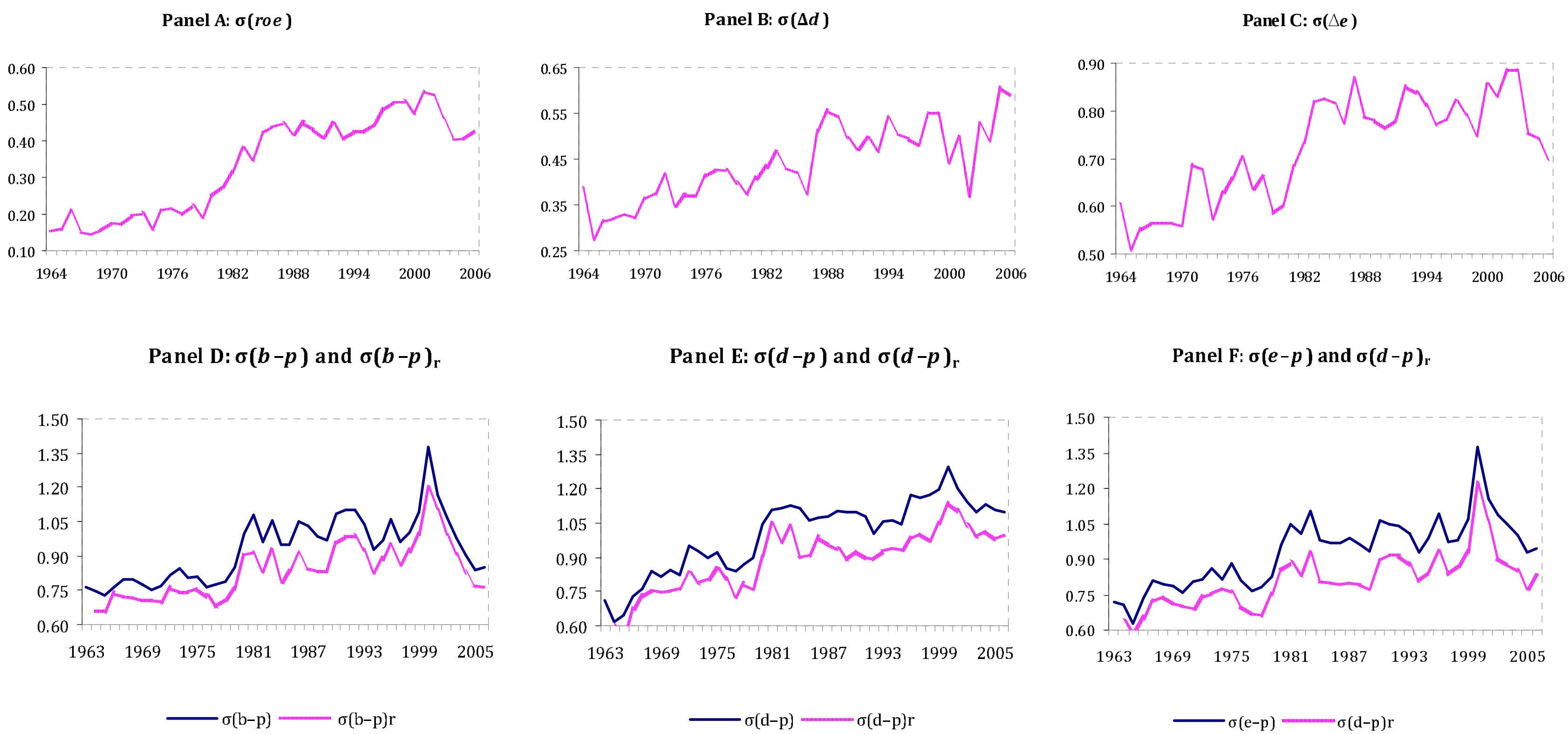
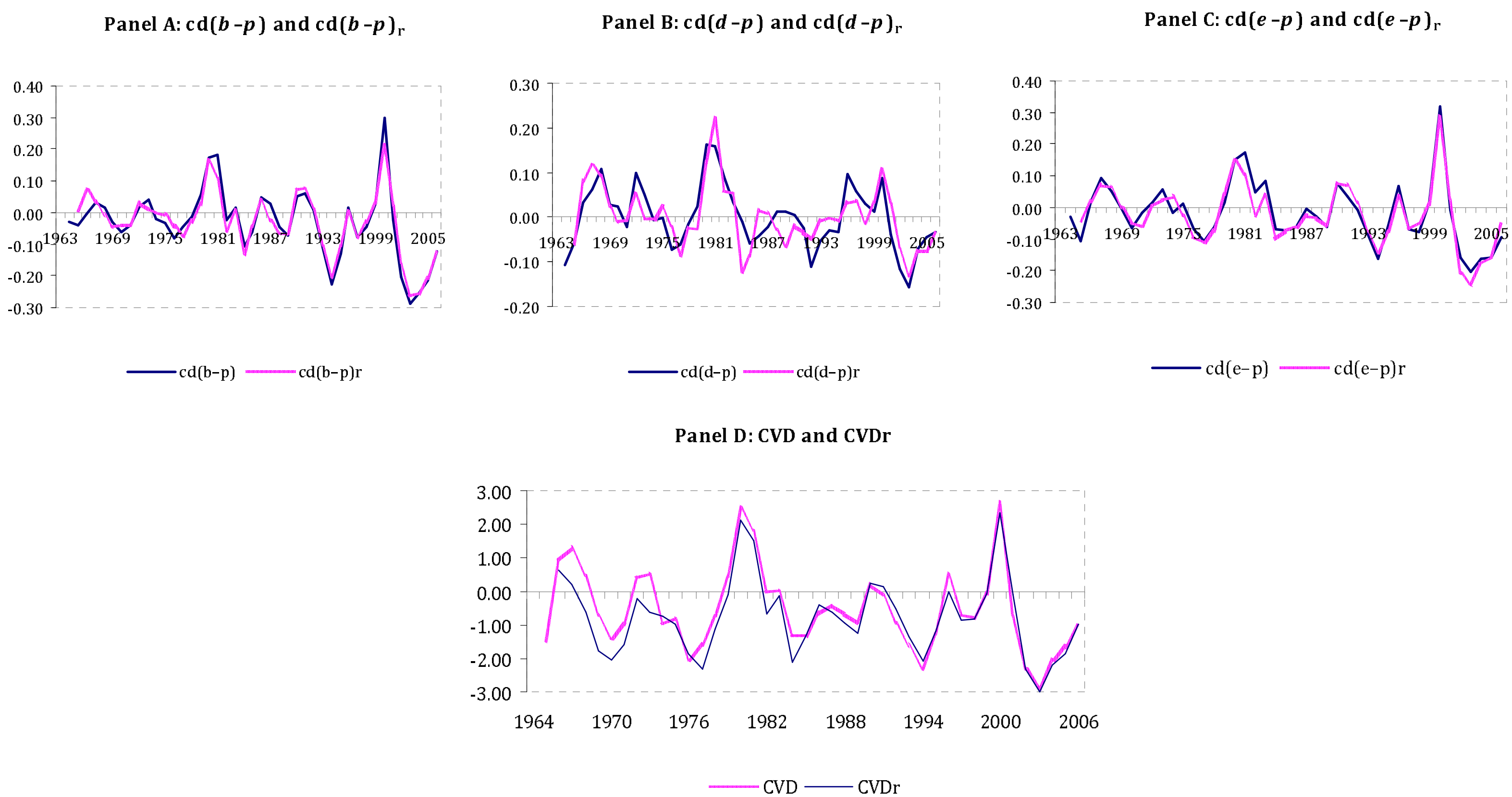


Figure 2: Detrended Dispersion Variables and Cross-firm valuation dispersion

This figure plots the time series of the annual dispersion variables over the period 1964–2006. The variables $\mathrm{cd}(b-p)$, $\mathrm{cd}(d-p)$, $\mathrm{cd}(e-p)$, $\mathrm{cd}(b-p)_{\mathrm{r}}$, $\mathrm{cd}(d-p)_{\mathrm{r}}$, and $\mathrm{cd}(e-p)_{\mathrm{r}}$ are detrended firm valuation dispersion measures, defined in Table 1. The variable CVD is defined as the first principal component of $\mathrm{cd}(b-p)$, $\mathrm{cd}(d-p)$, and $\mathrm{cd}(e-p)$, using information up to each month when CVD is formed. Similarly, CVDr is defined as the first principal component of $\mathrm{cd}(b-p)_{\mathrm{r}}$, $\mathrm{cd}(d-p)_{\mathrm{r}}$, and $\mathrm{cd}(e-p)_{\mathrm{r}}$, using only prior information. This figure shows that dispersion measures based on different valuation ratios exhibit similar variation over time.



## Figure 3: Future 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 (L) or above (H) its historical mean up to each year. Panels A and B are sorted based on the CVD from 06/1965–12/2006. Panels C and D are sorted based on the CVD from 06/1965–06/1996. This figure shows that high (low) beginning-of-period CVD is related to low (high) subsequent aggregate excess returns.

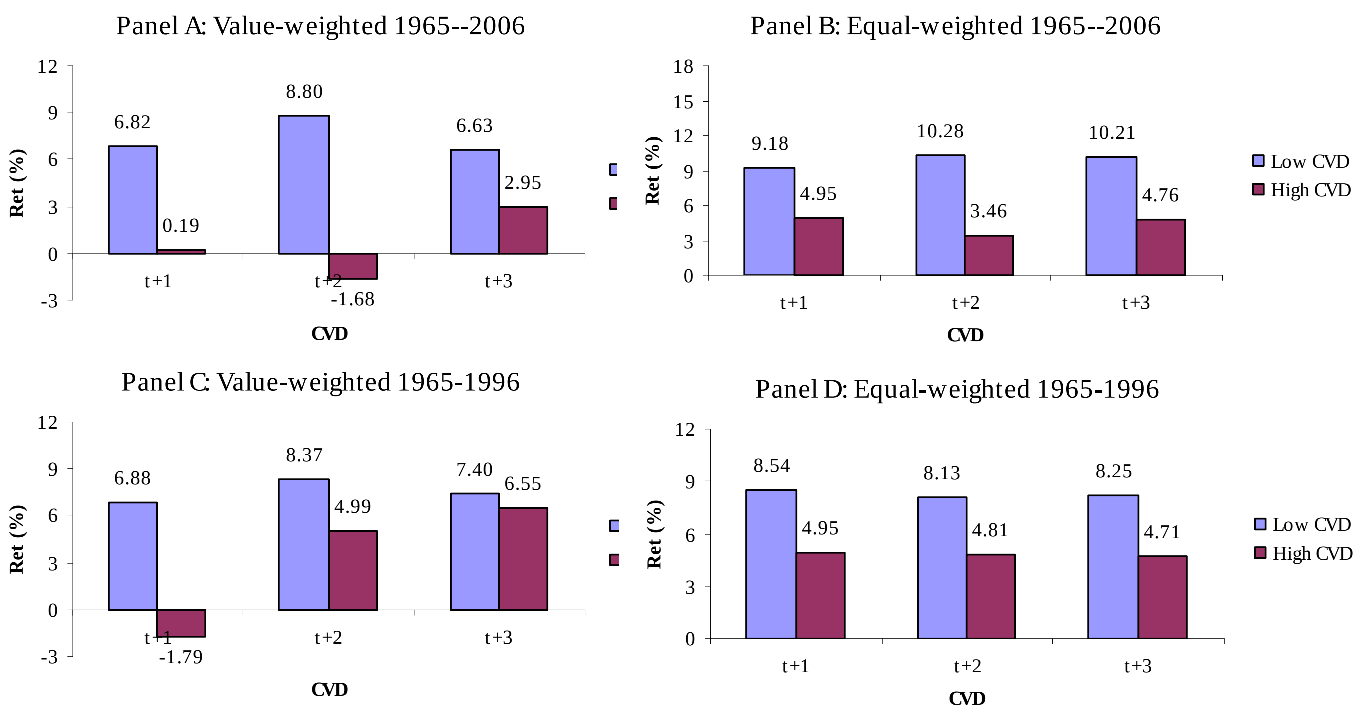


Figure 4: Out-of-Sample Performance and Trading Profits

Panels A and B plot the cumulative differences in squared out-of-sample forecast errors (SFE) between the null model, which is based on historical equity premium, and the alternative model, which is based on CVD or CVDr. Panel A forecasts quarterly aggregate excess returns using the regression coefficients based on data from the beginning of the sample to the quarter when the forecasts are made, with the initial forecasts starting from September 1984. Panel B forecasts annual aggregate excess returns, also using prior information, with the initial forecasts starting from June 1985. An increase of the line indicates better forecasts using the alternative model relative to the null. Panel C plots the cumulative investment payoffs of $1000 invested in June 1965 through December 2006, based on three trading strategies. The buy-and-hold market strategy invests in the value-weighted market index. The conditional BETA strategy is long on the lowest value-weighted BETA quintile from July of each year to the next June when the end-of-June CVD is below its historical mean (26 out of 42 years), and on the highest value-weighted BETA quintile, otherwise (16 out of 42 years). The conditional VOL strategy is similar to the conditional BETA strategy except that the quintiles are based on VOL. This figure shows that CVD and CVDr better forecast equity premium in real time than a naive model and produce profitable trading strategies that beat a passive buy-and-hold market strategy.