Market Crash Risk and Slow Moving Capital

Alexander Chernyakov Johnny Kang

First Draft: June 20, 2013

This Draft: November 14, 2013

Chernyakov is at Harvard University Department of Economics and Harvard Business School. Email: cherny@fas.harvard.edu. Kang is at AQR Capital Management. Email: Johnny.Kang@aqr.com. We are deeply grateful for suggestions from John Campbell, Lauren Cohen, Josh Coval, Robin Greenwood, and Adi Sunderam.

Abstract

Value stocks and junk bonds do poorly when the risk neutral probability of market crashes increases. However, investors are slow to fully incorporate this information into prices leading to significant predictability in value vs. growth stocks as well as junk vs. investment grade bonds. Using data on mutual fund flows, we find that investors rotate out of value stocks slowly following increases in the risk neutral probability of a market crash confirming the underreaction observed in returns. These findings set a higher bar for rational models that attempt to explain the risk premia in these two asset classes: because the returns to the value vs. growth trade and junk vs. investment grade trade are predictable, investors can achieve a significantly better return by actively rotating into/out of these securities. Our research highlights the role of slow moving capital in equity and bond markets.

1 Introduction

The distribution of returns on the wealth portfolio changes over time: some times are riskier than others. Finance has traditionally viewed market volatility as the source of risk and a tremendous literature has developed on understanding how variation in volatility of the market portfolio affects returns in the cross-section of assets (for example: Campbell et al. (2013); Ang et al. (2006); Chen (2002)). Recent work has begun to recognize that other moments of the distribution matter as well. In particular, the asymmetry in the distribution of returns is important to investors. Martin (2013) shows theoretically that investors with CRRA (or Epstein-Zin) utility require higher compensation for holding the market portfolio when its distribution of returns has a negative third cumulant. Our work attempts to understand if the cross-section of securities efficiently incorporates information regarding changes in the third cumulant of the market portfolio. We use the cross-section of equity securities as our primary test assets and verify our conclusions in the corporate bond market as well.

Our findings are two-fold: the value-minus-growth (HML)[1] trade performs poorly when the third cumulant of the market return distribution becomes more negative. HML incorporates information into its price slowly and thus continues to do poorly the following month as well. These conclusions also hold in the corporate bond market: the junk-minus-investment grade trade performs poorly when the third cumulant becomes more negative but also continues to perform poorly the following month. We measure changes in the distribution of returns by using options on the S&P 500 index. The benefit of using options is it allows us to extract investor expectations about the distribution of future returns without having to use a long time-series (which would be difficult anyway if the distribution changes frequently). Options, however, reveal the moments of the risk neutral distribution (instead of the physical) and thus we are unable to distinguish between changes in risk aversion and the distribution of physical returns.

HML has historically enjoyed high returns; the reason for this phenomenon is controversial as Chan and Lakonishok (2004) point out. The behavioral literature - as explained by Lakonishok et al. (1994) - has claimed that investors make mistakes regarding future earnings growth rates on value and growth stocks by incorrectly extrapolating recent earnings growth. Therefore, according

[1] We refer to a trade that goes long value and short growth as well as the formal value factor constructed by Fama and French (1992) as HML.

to this line of reasoning, value stocks are not riskier than growth stocks but are simply mispriced. Rational explanations claim that value stocks achieve high returns because they are in some way riskier (Fama and French (1992, 1993, 1996, 1998)). Campbell et al. (2013), for example, shows that HML performs poorly when volatility on the market portfolio increases. Our results are supportive of this interpretation: we find that HML also performs poorly when the third cumulant of the risk neutral distribution becomes more negative. However, we set a new bar for strictly rational explanations of the value premium because returns to HML are forecastable. This timing ability allows an investor to take HML risk opportunistically to achieve a significantly higher return stream. Accounting for this better performance in a strictly rational framework will be difficult. Furthermore, we show (using mutual fund flows) that investors underreact to changes in the third cumulant suggesting that either behavioral or institutional frictions will be important in explaining an active strategy that rotates into and out of HML.

This work contributes to two strands of literature: understanding the value premium and slow moving capital. In a series of papers Fama and French (1992, 1993, 1996, 1998) show that high book-to-market securities earn high returns relative to low book-to-market stocks. The rational asset pricing literature has advocated for an ICAPM style model where growth and value stocks have covariance with state variables. Campbell and Vuolteenaho (2004) examine the covariance of securities with discount rates and cash flows. Using an ICAPM they show that the price of risk associated with discount rate covariance is lower; value stock betas are mostly composed of cash flow betas while growth stock betas are discount rate betas. Therefore value stocks should command a higher return per unit of market beta. In a followup article Campbell et al. (2010) attempt to understand if this is related to sentiment: the authors show that the difference in value and growth stocks corresponds to the covariance of their fundamentals with market discount rate or cash flow news suggesting that there is a fundamental risk story. Petkova and Zhang (2005) show that market β of value stocks is higher in bad states of the world. In recent work, Campbell et al. (2013) report that value stocks have a negative β to volatility news. The authors use a long time series of realized volatility with a FIGARCH model as their proxy for volatility. Our work extends this literature by showing that returns to HML are poor when the third cumulant of the risk neutral distribution becomes more negative.

In a large separate literature, researchers have documented underreaction by investors to publicly available information. Hong and Stein (1999) present a model theoretically motivating how this could occur. Ang et al. (2012) show that individual stock options contain information regarding future returns of individual stocks. Duffie (2010) shows that arbitrage opportunities can remain unexploited during periods when investors experience impediments to trade. We contribute to this literature by showing that investors underreact to changes in the third cumulant of the risk neutral distribution.

2 Crash Risk

The third cumulant corresponds to the third central moment of the distribution. It is negative when the distribution is negatively skewed and thus has a long left tail, something investors (generally speaking) dislike. This aversion to a long left tail is captured in CRRA and Epstein-Zin preferences. Martin (2013) shows that[2]:

$$\mathbb{E}_t(R^e_{m,t}) \approx \gamma\overline{\kappa_2} - \frac{1}{2}\gamma^2\overline{\kappa_3} + \text{higher order stuff} \tag{2.1}$$

where $\overline{\kappa}_j$ is the j^{th} cumulant of the physical distribution of $R^e_{m,t}$ and γ is investor risk aversion[3]. A more negative $\overline{\kappa}_3$ causes a higher expected return on the market portfolio. We refer to this as crash risk because of the definitional relationship between skewness and $\overline{\kappa}_3$,

$$skewness \equiv \frac{\overline{\kappa}_3}{\overline{\sigma}^3} \tag{2.2}$$

where $\overline{\sigma}$ is the volatility of the physical distribution.

Namely, $\overline{\kappa}_3$ measures the interaction between skewness and volatility raised to the third power. The risk neutral version of $\overline{\kappa}_3$, which we denote κ_3, can be extracted from the options surface using the model free methods of Bakshi et al. (2003). Details are provided in the Appendix. Chang et al. (2013) examine the price of risk neutral skewness and do not find a relationship with HML. We differ in results because in our work, $\overline{\kappa}_3$ is the relevant theoretical quantity which is an interaction of volatility and skewness. Since κ_3 is a risk neutral measure, it can change because investor risk

[2]In his notation, with $\lambda = 1$, approximating $-\gamma(\gamma-1) \approx -\gamma^2$.

[3]We work directly with the market in this article which can be justified if the consumption to wealth ratio is approximately constant

aversion (γ) changes or because the characteristics of the physical distribution have changed. This relation is explained by Bakshi et al. (2003) under power utility[4]:

$$\kappa_3 \approx \overline{\kappa}_3 - \gamma \overline{\kappa}_4 \tag{2.3}$$

Therefore, risk due to the third cumulant can increase because the cumulant associated with the physical distribution has become more negative or because investors have become more risk averse.[5] As mentioned earlier, we are unable to disentangle these two effects. Because our work is concerned with understanding the degree of efficiency with which prices incorporate information, rather than estimating particular parameters of a model or explaining risk premia, this is less of a concern.

While we provide robustness results using the methodologically exact measurement of κ_3, we focus instead on an approximation analogous to the risk reversal used by Brunnermeier et al. (2008) to measure currency crash risk. Moreover, investors are used to thinking of risk as a positive number so we focus on approximating $-\kappa_3$ in our work. Specifically define

$$SKEW_t \equiv \frac{1}{|P|} \sum_{j \in P} IV_{j,t} - \frac{1}{|C|} \sum_{j \in C} IV_{j,t} \tag{2.4}$$

where P (C) is the set of 1 year S&P 500 out of the money (OTM) put (call) options and $|P|$ ($|C|$) its cardinality, $IV_{j,t}$ is implied volatility on option j at time t. Thus this is simply the average volatility on OTM S&P 500 put options minus the average volatility on OTM S&P 500 call options. We use the interpolated volatility surface from OptionMetrics to compute $SKEW$ (while the actual set of option prices is necessary to compute κ_3 which we discuss in the Appendix) monthly from 1996-2012 based on data availability. The surface provides a set of volatilities for put and call options in increments of five delta units and thus is symmetric around the at the money (ATM) point. Since our work will focus on monthly data and option markets close 15 minutes after equity markets, we use the second to last day of the month to compute our implied volatility related metrics to prevent any look ahead bias.

To relate $SKEW$ theoretically to risk neutral moments that we are familiar with, we look to

[4]Note that Bakshi et al. (2003) use κ to denote moments while we use it to denote cumulants.

[5]For completeness we note that this could also happen if $\overline{\kappa}_4$ has increased: the distribution has become more fat tailed

Proposition 2 of Backus et al. (2004). Using Gram-Charlier approximations the authors show that

$$IV_t(d) \approx \sigma_t \left[1 - \frac{RNSKEW_t}{3!} d - \frac{RNKURT_t}{4!}(1 - d^2)\right] \tag{2.5}$$

where $IV_t(d)$ is the implied volatility of a call option with moneyness d ($d \approx 0$ for at the money options, $d > 0$ in the money, and $d < 0$ out of the money)[6], σ_t is the risk neutral volatility of the underlying, $RNSKEW_t$ is the risk neutral skewness, $RNKURT_t$ is the risk neutral kurtosis. By put-call parity, in the money (ITM) call options will have the same volatility as OTM put options with the same strike. Therefore, $SKEW_t$ is equivalent to measuring the difference between the average volatility on ITM call options and OTM call options. Since we are using a standardized surface with perfectly symmetric d, by (2.5), the difference in implied volatilities between an ITM call option ($d > 0$) and the symmetric OTM call option ($d < 0$) is[7]:

$$IV_t(d) - IV_t(-d) \approx -\sigma_t \cdot RNSKEW_t \tag{2.6}$$

Therefore, $SKEW_t$:

$$SKEW_t \approx -\sigma_t \cdot RNSKEW_t \tag{2.7}$$

That is, just like κ_3 it measures the interaction of volatility and risk neutral skewness. When κ_3 is highly negative (high risk), $SKEW$ will be highly positive; it captures times of high risk neutral volatility *and* negative skewness.

There are several reasons we focus on using $SKEW$ as opposed to κ_3 in our work. First, extracting $SKEW$ from the set of data provided by OptionMetrics is simple and this database is used by academics and practitioners alike. Second, while κ_3 is meaningful to someone whose goal is to estimate the coefficient of risk aversion (γ), $SKEW$ is much more interpretable from a practitioner perspective. Risk reversals are frequently traded by options investors and are quoted in terms of volatility points - just like $SKEW$. One can have an intuitive understanding of what a one volatility point increase in $SKEW$ represents while an intuitive understanding of a one point decrease in κ_3 is more elusive. Thus another way to look at $SKEW$ is the following: when investors

[6]Formally, $d \equiv \frac{log(S_t/K)+r+\sigma^2/2}{\sigma}$

[7]$IV_t(d) - IV_t(-d) \approx -2\sigma_t \frac{RNSKEW_t}{3!} d$. Since we will be examining innovations, we can drop time invariant which yields the desired result.

want to protect their portfolio they can buy risk reversals. This transaction would increase the spread in implied volatility between OTM puts and OTM calls and push $SKEW$ higher. Therefore when investors are particularly fearful and want to protect their portfolios with such a transaction, $SKEW$ will be particularly high.

While we have related $SKEW$ to risk neutral moments from a theoretical perspective, it is helpful to empirically verify that our metric is capturing the desired quantities. We proxy for market volatility using the VIX index and extract $RNSKEW$ from the cross-section of S&P 500 options using the methodology described in the Appendix. As noted earlier, a high positive $SKEW$ represents risky states of the world from the perspective of the investor. Analogously a highly positive VIX represents high (risk neutral) forecasted levels of volatility; a highly *negative* $RNSKEW$ represents a *negatively* skewed distribution. The point to keep in mind is that a *negative* innovation in $RNSKEW$ is an *increase* in risk from the perspective of an investor. The opposite is true for VIX and $SKEW$: a *positive* innovation to these variables corresponds to a riskier distribution. We extract innovations at the monthly frequency using univariate ARMA models chosen by BIC for each quantity ($SKEW$, VIX and $RNSKEW$). All of the quantities are can be described by a low order ARMA model.

Our goal is to understand how $SKEW$ innovations, ε_{skew}, relate to innovations in risk neutral moments. We attack this problem parametrically using linear regression and non-parametrically using local polynomial regressions. Namely, we run two forms of regressions:

$$\varepsilon_{skew,t} = f(\varepsilon_{rnskew,t}, \varepsilon_{vix,t}) \tag{2.8}$$

$$\varepsilon_{skew,t} = a + \beta_v \varepsilon_{vix,t} + \beta_{rns} \varepsilon_{rnskew,t} + \beta_i (\varepsilon_{vix,t} \varepsilon_{rnskew,t}) + \eta_t \tag{2.9}$$

where $f(\cdot)$ is a second order local polynomial. Figure 1 presents a surface representing the results of fitting (2.8). Variables are standardized prior to running these regressions so that magnitudes can be more easily interpreted. The arrows point in the positive direction for each variable.

The figure plots the fitted values of $\varepsilon_{skew,t}$ from regression (2.8) against $\varepsilon_{vix,t}$ and $\varepsilon_{rnskew,t}$. It shows ε_{skew} is high when there is an increase in volatility *and* the distribution of returns becomes more negatively skewed. Thus $SKEW$ is a measure of the joint behavior of volatility and skewness, as desired. This can be examined in a linear regression context as well: the table below the figure presents results of regression (2.9). The results are the same as those explained by the plot: $SKEW$ increases when the distribution becomes more negatively skewed ($\varepsilon_{rnskew} < 0$) and volatility increases ($\varepsilon_{vix} > 0$). Additionally, there is a significant interaction effect that was highlighted by the plots: $SKEW$ increases particularly strongly when there is an increase in volatility and the distribution of returns becomes more negatively skewed.

It is helpful to understand how $SKEW$ varies through time in relation to the VIX and the business cycle since investors are generally familiar with the time-series pattern of these quantities. Figure 2 presents a plot of $SKEW$ for our sample along with the VIX and $SKEW$ orthogonalized to the VIX (using linear regression) labeled $OrthSKEW$; recessions are highlighted using gray bars. The variables have been standardized. One obvious pattern is the correlation that $SKEW$ exhibits with the VIX: roughly 60%. However, there are subtle differences in the pattern: $SKEW$ was higher relative to its normal levels than the VIX (relative to its normal levels) prior to the dot-com bubble bursting. It also spiked up prior to the 2008 financial crisis and has remained elevated after the crisis. We will see further in the article that these features are important in predicting HML returns.

3 Risk and Return

3.1 Equities

Martin (2013) provides the relationship between $\overline{\kappa}_3$ and returns on the market portfolio: a more negative third cumulant should correspond to higher expected returns. Translated to our metric, this implies a higher level of $SKEW$ should mean higher expected returns. We first verify empirically that this conjecture is true and $SKEW$ does, in fact, forecast the return on the market. The finance literature has already identified several variables that are considered state variables and forecast future market returns. We are careful to control for these other variables to understand the multivariate implications of $SKEW_t$ on market returns. The variables that we consider are the dividend yield on the S&P 500, the smoothed earnings yield, the term premium and the default premium. To understand the relationship between $SKEW_t$ and expected market returns, we run the following regression:

$$R^e_{m,t} = a + \rho_m R^e_{m,t-1} + \beta_s SKEW_{t-1} + \beta_{dp} dp_{t-1} + \beta_{dy} dy_{t-1} + \beta_{sey} sey_{t-1} + \beta_{tp} tp_{t-1} + \varepsilon_{m,t} \quad (3.1)$$

using monthly data from 1996 - 2012. Our sample is constrained by the availability of options data. Additionally, much of the expected return literature - see Cochrane (2011) for a recent summary - has reported stronger effects at longer horizons. Therefore, we also run a regression of 6 month market excess returns on predictor variables:

$$R^e_{m,t\to t+5} = a + \rho_m R^e_{m,t-1} + \beta_s SKEW_{t-1} + \beta_{dp} dp_{t-1} + \beta_{dy} dy_{t-1} + \beta_{sey} sey_{t-1} + \beta_{tp} tp_{t-1} + \varepsilon_{m,t} \quad (3.2)$$

Table 1 reports the results. As noted earlier, since options markets close later than equity markets, we skip an extra day between information on $SKEW_t$ and any equity return to prevent look-ahead bias. Thus in monthly data, $SKEW_t$ represents observations on the second to last day of the month. Note that the bandwidth selection for the HAC standard errors is automatic using the procedure of Newey and West (1994) and thus accommodates overlapping observations in regression (3.2). $SKEW$ has significant forecasting power for market returns even in the presence of other state variables. A one volatility point higher $SKEW$ corresponds to roughly 30 - 60 basis points of expected market returns the following month. Similarly the second half of the table shows that a one volatility point increase in $SKEW$ is related to 1.8% higher expected return over the following 6 months. The empirical return relationship is supportive of (2.1).

We next turn to the cross-section of equities: to determine if innovations in $SKEW$ are differentially important in the cross-section (this is not a pre-determined conclusion) we use the Fama-French 25 portfolios as our basis assets and regress excess returns of each portfolio, $R^e_{i,t}$, on ε_{skew} controlling for market returns and VIX innovations:

$$R^e_{i,t} = a + \beta_m R^e_{m,t} + \beta_v \varepsilon_{vix,t} + \beta_{\varepsilon_{skew}} \varepsilon_{skew,t} + \nu_t \quad (3.3)$$

Table 2 presents the multiple β of each portfolio with respect to ε_{skew}. The results here are pretty clear: HML has a negative exposure to ε_{skew}. That is, when the risk neutral distribution becomes riskier, value stocks do poorly while growth stocks do well. The results are especially strong in small and medium stocks. This is an empirical extension to higher cumulants of results reported by Campbell et al. (2013) who show that HML has a negative exposure to increases in volatility.

Since HML is clearly the factor of interest, we use the HML portfolio directly rather than each individual Fama-French portfolio in subsequent results. As mentioned in the introduction, we are interested in understanding how efficient securities in the cross-section are at incorporating information regarding changes in risk into their prices; we find a significant lag in the price adjustment process. In addition to doing poorly at time t when there is a positive innovation to $SKEW_t$, $R_{hml,t+1}$ is also highly negative; value stocks continue to underperform the following month. Figure 3 plots the cumulative response to $\varepsilon_{skew,t}$ from the regression of

$$R_{hml,t} = a + \beta_m R^e_{m,t} + \beta_{\varepsilon,j}\varepsilon_{skew,t-j} + \epsilon_t \tag{3.4}$$

for $j = 0...12$. The Newey-West error bounds are presented in the plot in dashes; j is measured on the x-axis. For each j, the figure plots $\sum_{k=0}^{j}\beta_{\varepsilon,k}$ and assumes that estimates of $\beta_{\varepsilon,j}$ are independent so that the cumulative standard error at each j is $\sqrt{\sum_{k=0}^{j} se_k(\beta_{\varepsilon,k})^2}$ where $se_j(\beta_{\varepsilon,j})$ is the Newey-West standard error associated with $\beta_{\varepsilon,j}$ in equation (3.4). This regression can be interpreted as asking: if there is a one volatility point increase in $SKEW$ at time t and a one volatility point increase in $SKEW$ at $t-1$, ..., $t-j$ then what is the cumulative effect on HML at time t?

We see a striking pattern: while $R_{hml,t}$ is indeed highly negatively correlated with $\varepsilon_{skew,t}$, there is a significant delay in the price adjustment process. The plot shows that the $j = 0$ and $j = 1$ coefficients for $R_{hml,t}$ are highly negative and significant (with associated t-statistics of -2.94 and -3.84, respectively). Following an increase in $SKEW$, HML continues to underperform the following month. In fact, we see that the price adjustment process actually takes up to 4 months to fully realize. This is a startling finding: to confirm these results we extend the sample internationally to Europe and Japan[8]. Note that our analysis is done from the perspective of a US investor (thus, for example, market returns refer to the CRSP value weighted market return). We also attempt to control for the effect of time-varying HML market β. As noted earlier, Petkova and Zhang (2005) found that the market β of HML varies through time; it is possible that a high level

[8] While this gives us comfort against data snooping concerns, we note that there is a 40% correlation between US HML returns and JPY HML returns; a 60% correlation between US and Europe HML returns.

of $SKEW_{t-1}$ forecasts a higher market β at time t. To be precise, imagine that

$$R_{hml,t} = \beta_{m,t} R^e_{m,t} + \eta_t \tag{3.5}$$

$$\beta_{m,t} = \beta_0 + \beta_1 SKEW_{t-1} \tag{3.6}$$

then the appropriate attribution regression to run for HML is

$$R_{hml,t} = \beta_0 R^e_{m,t} + \beta_1 (SKEW_{t-1} \cdot R^e_{m,t}) + \eta_t \tag{3.7}$$

We would like to eliminate the possibility that a time varying market β is driving our results so we include the interaction of $SKEW_{t-1}$ and $R^e_{m,t}$ into the regression. Therefore, for each region we run a forecasting regression of the form:

$$R_{hml,t} = a + \beta_m R^e_{m,t} + \beta_s \left(\sum_{j=1}^{4} \varepsilon_{skew,t-j} \right) + \beta_i (R^e_{m,t} \cdot SKEW_{t-1}) + \nu_t \tag{3.8}$$

where we use the sum of the last four innovations in $SKEW$ based on the results of Figure 3. Table 3 presents the summary statistics of HML returns in each region (Panel A) and reports the results from this regression (Panel B).

Panel A shows that HML in all three regions has a high return and a high alpha during our sample period. The units are left in their natural monthly frequency (the Sharpe ratio is annualized). Thus HML, in the US, has a CAPM alpha of .347% per month. The sample statistics in all three regions are similar with Europe having the highest Sharpe ratio. Panel B reports that a one volatility point unexpected increase in $SKEW$ over the last four months leads to a .6% lower return the following month in the US, .25% lower HML return in Japan and .37% lower HML return in Europe. The adjusted R^2 statistics are also quite large: a simple univariate forecasting regression in the US can explain 10% of the variation in HML returns in monthly data. The t-statistics are also very large: in the US the t-statistic associated with $\sum_{j=1}^{4} \varepsilon_{skew,t-j}$ is -6; this is directly linked to the Sharpe ratio of a strategy that one can construct Sharpe (1994).

To construct a realistic trading strategy from these regressions, we would like to avoid estimating an ARMA model for innovation extraction and also avoid estimating the distributed lag model for forecasting R_{hml}. While this is certainly the optimal method, our sample is relatively small and we want to avoid all possibility of look-ahead bias in parameters. To get around this constraint, we will simply use the level of $SKEW$: since several lags of innovations seem important (as noted earlier) and $SKEW$ is not terribly persistent, we hope that older innovations that are irrelevant will have decayed sufficiently and thus not erode our forecasting performance. To do this we simply run regression (3.8) but replace the term $\sum_{j=1}^{4} \varepsilon_{skew,t-j}$ with $SKEW_{t-1}$:

$$R_{hml,t} = a + \beta_m R^e_{m,t} + \beta_s SKEW_{t-1} + \beta_i (R^e_{m,t} \cdot SKEW_{t-1}) + \nu_t \tag{3.9}$$

Table 4 reports these results.

Comparing the two tables we can immediately see that our forecasting performance is worse using the level of $SKEW$ as to be expected: the t-statistic is cut in half in the US and so is the R^2. Thus, backtest results that we report using $SKEW$ as a forecasting variable is a lower bound on the true performance that an investor can achieve. Robustness of these results is presented in the Appendix: the results are robust to choice of volatility maturity and sub-samples (see Tables 12 and 13). Finally, we want to be sure that our results are not related to something else that $SKEW$ (being an approximation for $-\kappa_3$) is picking up. To verify this we run a forecasting regression of HML returns in all three regions but replace the $SKEW$ level (as in Table 4) with the level of $-\kappa_3$. Table 5 reports these results which are qualitatively no different (in terms of forecasting ability) than those with $SKEW$.

These results imply a strategy, which we refer to as active HML, that would selectively rotate into and out of HML being long value (growth) and short growth (value) stocks at different points in time. We are careful to prevent any look-ahead bias in the parameters and ensure this strategy is realistic. We begin by computing a forecast for returns on HML using

$$R_{hml,t-1} = a + \beta_s SKEW_{t-2} + \epsilon_{t-1} \tag{3.10}$$

allowing 36 months burn in period for estimation of β_s. Based on this model we compute the forecast for the following month's HML return, $\hat{R}_{hml,t}$. Assuming an endowment of $\mathcal{W}_{t-1}$, we build a portfolio by putting $w_{hml,t-1}$ of $\mathcal{W}_{t-1}$ into HML. Assuming that margin accounts pay no interest rate and require 50% of the absolute value of the position (so that if one wants to go long HML by purchasing "H" and selling "L" then one has to put up margin equivalent to half of the position), the remainder of the endowment, $(1 - |w_{hml,t-1}|)\mathcal{W}_{t-1}$, is invested into the risk free rate[9]. The weight is defined as $w_{hml,t-1} = \tanh\left(\frac{\hat{R}_{hml,t}}{\sqrt{\frac{1}{t}\sum_{j=1}^{t}(\hat{R}_{hml,j} - \overline{\hat{R}}_{hml,t})^2}}\right)$. This is the hyperbolic tangent of the forecasted HML return scaled by the standard deviation of previous HML forecasts. The hyperbolic tangent is applied so that $w_{hml,t-1} \in [-1, 1]$. The gross return to this portfolio $\frac{\mathcal{W}_t}{\mathcal{W}_{t-1}} = 1 + w_{hml,t-1}R_{hml,t} + (1 - |w_{hml,t-1}|)R_{free,t}$ where $R_{free,t}$ is the risk free rate realized at time t. Each successive month, the window over which the model is estimated expands but always only includes historical data. Rebalancing is done monthly for both active and passive HML. The monthly rebalancing for passive HML assures that the investor puts half of his wealth in being long "H" and half into being short "L" (ie equal weights). The cumulative return to passive and active HML is presented in Figure 4. The shaded regions represent times when the strategy is short HML while the white areas represent times when the strategy is long HML.

As is evident from the plot, the position direction is quite persistent; since HML is such a large aggregate, transaction costs here are also minimal. However, the performance of active HML is significantly better than passive HML. An investment of \$1.00 in HML in 1999 becomes roughly \$1.50 by the end of 2012. This same dollar invested in active HML becomes roughly \$2.50 by the end of 2012. The annualized information ratio, $IR \equiv \frac{E(R^{active}_{hml,t} - R^{passive}_{hml,t})}{\sigma(R^{active}_{hml,t})}$, for this strategy is .36. The times when this strategy is short HML correspond to significantly anomalous market conditions. For example we see that this strategy is short HML during the run-up in tech stocks of the late 90's. It is also short HML during/post the 2008 financial crisis. These two periods account for a substantial portion of the profit generated by this strategy. This is to be expected: we are attempting to pick up states of the world when the price investors are willing to pay for portfolio insurance spikes and these states should not be a frequent occurrence.

Given the relationship between $SKEW$, VIX and $RNSKEW$ we can horse-race them to see which contains the most relevant information for future HML returns. If, as claimed, $SKEW$ is the more relevant variable then it should perform better than VIX or $RNSKEW$. To answer this

[9] We assume that $\mathcal{W}_{t-1} > 0$; if at any point this condition is violated the strategy stops.

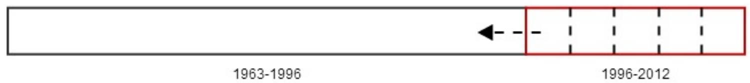
question Table 6 presents a regression of

$$R_{hml,t} = a + \beta_1 SKEW_{t-1} + \beta_2 VIX_{t-1} + \beta_3 RNSKEW_{t-1} + \eta_t \tag{3.11}$$

As the results clearly indicate, in a multivariate regression with $SKEW$ and VIX, the VIX becomes insignificant in predicting HML returns while the $SKEW$ is still highly significant. Similarly, in a regression using $SKEW$ and $RNSKEW$, $SKEW$ is the more relevant variable.

While we have been fortunate that our time sample includes diversity in business cycle conditions, we are still constrained by the availability of options data from OptionMetrics. To verify that our results work in other time samples - a truly out of sample test - we attempt to impute the value of $SKEW$ based on quantities that $SKEW$ should be picking up. To extend the sample in a principled way, we use the LASSO variable selection method of Tibshirani (1996). This methodology is an ℓ^1 penalized regression that selects variables among a candidate set that best capture the true relationship and kicks out all irrelevant ones. The set of possible variables that we include is $I_t \equiv \{\overline{\kappa}_3, R_{hml,t}, R^e_{m,t}, dy_t, tp_t, dp_t\}$ and I_{t-1}. As noted earlier, $\overline{\kappa}_3$ is the physical third cumulant of market returns computed over the past 3 months of daily data, $R_{hml,t}$ is the US HML return, $R^e_{m,t}$ is the market return, dy_t is the dividend yield, tp_t is the term premium, dp_t is the default premium. The logic for including these variables is simple: $\overline{\kappa}_3$ should capture the portion of $SKEW$ corresponding to the physical distribution, dy_t, tp_t, and dp_t could capture risk aversion that we know affect $SKEW$ as in equation (2.3), and we have already discussed how HML and market returns relate to $SKEW$.

To operationalize this technology, we use five-fold cross validation using the 1996-2012 sample, to fit the LASSO model to the $SKEW$ using I_t and I_{t-1}. The cross validation is needed to select the shrinkage parameter that LASSO uses to determine how aggressive it should be in shrinking regression coefficients. For a particular shrinkage parameter, we divide the sample (1996-2012) into 5 sections, pick a section to leave out, fit LASSO over the remaining four sections and compute the root-mean-square error of the model in forecasting the level of $SKEW$ in the left out section; we then leave a different section out and repeat the process. The average root-mean-square error for this particular value of the shrinkage parameter is recorded. A shrinkage parameter is selected that creates the lowest forecasting error. This shrinkage parameter corresponds to a particular set of variables out of the available set. A simple OLS model is then fit from 1996-2012 using the selected set of variables and is used to compute a fitted value of $SKEW$, termed $\widehat{SKEW}$, going back to 1963.



We validate that our results hold over this significantly longer non-overlapping sample period: we show that $\widehat{SKEW}$ has forecasting power for R_m^e and R_{hml} is slow to respond to innovations in $\widehat{SKEW}$, termed $\hat{\varepsilon}_{skew}$. Using the sample from 1963 - 1996, the first part of the Table 7 shows that R_{hml} responds slowly to innovations in $\hat{\varepsilon}_{skew}$: a one volatility point increase in $\widehat{SKEW}$ corresponds to 40 - 50 basis points poorer performance in HML the following month. The second part of the table shows that $\widehat{SKEW}$ is capable of forecasting the market return the following month as we saw in the 1996 - 2012 sample using $SKEW$.

3.2 Corporate Bonds

We have shown that HML responds to innovations in $SKEW$ slowly in three regions and over a lengthy time sample; these results led us directly to an implementable trading strategy. Do these results hold in other asset classes? We examine corporate bonds to try and answer this question. At the same time, it provides another out of sample test of our results regarding slow investor reaction to changes in crash risk. Corporate bonds are a natural asset class to examine because they contain a large cross-section (like equities) and have assets that are considered risky in absolute terms (junk bonds) by investors and those that are considered safe (investment grade bonds). Bank of America/ML provides total return indices by rating category (AAA through CCC) which we use as basis assets.

We first examine if these assets have differential exposure to innovations in $SKEW$ in Table 8. We see clearly that AAA bonds enjoy a positive return while CCC bonds have a negative return when $SKEW$ increases contemporaneously. Thus the long CCC, short AAA trade performs poorly when crash risk increases; this is the same type of result that we saw with value and growth stocks.

We next ask the question: do these securities incorporate changes in risk into their prices efficiently? We saw that investors active in equities do not and thus HML incorporates changes in $SKEW$ with a delay. To answer this question we run the analysis in equation (3.4) replacing $R_{hml,t}$ on the left hand side of the regression with $R_{CCC-AAA,t}$: the return of CCC bonds minus the return on AAA bonds. Figure 5 presents the results of this analysis. At lag zero we see the contemporaneous results presented in Table 8: the long junk short investment grade (CCC-AAA) trade performs poorly when there is a positive innovation in $SKEW$. However, we also see that it proceeds to perform poorly the following month as well (until reversing in month two). That is, these securities are also slow to fully incorporate all available information into their prices though they are more expedient than equities (which took 4 months). One can speculate regarding the reason for this: one story might be that corporate bonds have a more sophisticated investor base since fewer retail investors participate actively in the bond market; however, this is just speculation on our part.

These results can be presented in a regression framework using equation (3.8) replacing R_{hml} on the left hand side with returns on each bond rating category and using one lag of $SKEW$ innovations as opposed to four based on Figure 5. Table 9 presents the results of these regressions. We see that a one point increase in $SKEW$ predicts a -77 basis point return to $R_{CCC-AAA}$ the following month. These results confirm that underreaction to changes in crash risk is prevalent in the financial markets.

4 Investor Trading Behavior

As we pointed out early in the article, in a rational frictionless market, changes in risk/risk preferences should be reflected in stock returns contemporaneously. However, we have shown that prices are slow to fully incorporate information from $SKEW$ into HML. A large behavioral/frictional literature has documented significant delays in price reactions, discussed nicely by Hong and Stein (1999); Duffie (2010). Ang et al. (2012) show that options on individual securities have relevant information for future returns of those securities not captured by the standard risk factors. They propose that sophisticated investors express their views in the options market and this information is reflected in individual stock returns slowly. Duffie (2010) documents a number of such occurrences: in his evidence, investors are constrained due to leverage or lack of capital from exploiting arbitrage opportunities. Fresh capital arrives to the market slowly and thus anomalies persist.

Our work finds results that are similar: innovations in $SKEW$ contain information about future HML returns. While there is a clear contemporaneous reaction as shown in Figure 3, there is also a significant delay. We confirm that investors are in fact slow to react to this information by examining mutual fund flows. Just like HML returns, we find that flows into value and growth mutual funds are predictable using past innovations in $SKEW$. A positive innovation in $SKEW$ causes investors to withdraw money from value funds while not withdrawing money from growth funds.

We first classify funds by value style using their four factor exposure

$$R_{f,t} = a + \beta_{m,t}R_{m,t} + \beta_{smb,t}R_{smb,t} + \beta_{hml,t}R_{hml,t} + \beta_{umd,t}R_{umd,t} + \epsilon_{f,t} \tag{4.1}$$

using 36 months rolling regression as in Chan et al. (2002) among others. Funds are arranged into quintiles each month based on last month's HML exposure: $\beta_{hml,t-1}$. We further define

$$FLOW_{f,t} \equiv \frac{TNA_{f,t} - TNA_{f,t-1}(1 + R_{f,t})}{TNA_{f,t-1}} \tag{4.2}$$

for each fund f and month t where TNA is the total net assets and $R_{f,t}$ is the return of fund f. This is the percentage increase/decrease in the assets of the fund due to contributions/withdrawals by investors. Additionally as Chevalier and Ellison (1997) show, flows into mutual funds are highly dependent on the funds' past performance; we are sure to condition on this in our analysis so that any effect in investor behavior we find is due to information in $SKEW$ as opposed to past fund returns. Therefore, we also compute the 1 year rolling cumulative return of each fund: $R_{f,t-12\to t-1}$. Then, for each value style bucket (1 - 5) a TNA_{t-1} weighted average of $FLOW_{f,t}$ (termed $FLOW_{b,t}$ for $b \in \{1, 2, .., 5\}$) and $R_{f,t-12\to t-1}$ (termed $R_{b,t-12\to t-1}$) is taken. We regress:

$$FLOW_{b,t} = a + \beta_{b,skew}\varepsilon_{skew,t-1} + \beta_{b,ret}R_{b,t-12\to t-1} + \nu_t \tag{4.3}$$

Results in Table 10 show that $\beta_{b,skew}$ is highly significant and negative for value stocks but roughly zero or positive for growth stocks. Investors pull money out of value funds in response to an increase in market crash risk with a lag. The magnitudes are significant: one volatility point increase in $SKEW$ causes a .2% outflow from value funds relative to growth funds the following month. This underreaction by investors to the information expressed by option market participants drives the predictability of HML returns.

One may be concerned that mutual funds pre-position their portfolios in anticipation of flows and thus dampen the effect of lagged information on stocks. Consider a savvy value mutual fund manager who received an inflow of money this month: he may decide that there is a good chance that he will also receive an inflow of money the next month since flows are persistent. Knowing this fact, he uses his cash position (or takes a loan from a bank) to purchase value stocks this month anticipating to return his cash position (or pay the loan) to equilibrium the next month. This would serve to dampen the predictability of value stocks due to investor underreaction. To account for this fact we extract innovations from flows into each quintile, $\varepsilon_{b,t}^{flow}$, and treat these as unexpected flows to the fund manager. We then regress these unexpected flows on $\varepsilon_{skew,t-1}$ and past fund returns, $R_{b,t-12\to t-1}$:

$$\varepsilon_{b,t}^{flow} = a + \beta_{b,skew}\varepsilon_{skew,t-1} + \beta_{b,ret}R_{b,t-12\to t-1} + \nu_t \tag{4.4}$$

The results of this regression are presented in Table 11. This, however, does not alter our conclusions: innovations in $SKEW$ predict unexpected flows into growth and value stocks. Investors withdraw .1% from value funds relative to growth funds in response to one volatility point increase in $SKEW$ the previous month.

5 Discussion and Conclusion

We have shown that returns to HML (as well as corporate bonds) have a significant amount of predictability. Utilizing this predictability, we perform an "out-of-sample" test of performance: active HML significantly improves the returns to an investor relative to an investment in passive HML. A dollar invested in passive HML in 1999 grows to approximately $1.50 by the end of 2012. On the other hand, that same investment grows to roughly $2.50 in active HML. Furthermore, this predictability is easy to extract: it does not require complex computation just the implied volatility skew on the S&P 500. Using mutual fund flow data, we show that this predictability is due to delayed reaction by investors.

Our results have deep implications for theories attempting to explain high returns on HML:

theories must now consider explaining returns to active HML (a much more difficult thing to do given the favorable return profile). More generally, our results relate to a large literature on slow moving capital and segmented markets. We show how a large, heavily examined factor can have a significant amount of return predictability due to the slow rotation into and out of value/growth stocks by investors. One may regard slight economic frictions that prevent small stocks from re-pricing perfectly as unimportant. However, the return predictability that we identify here is on an aggregate, economically meaningful level. A significant portion of this predictability comes from periods when the market experiences stress: during the tech bubble and during the financial crisis of 2008. We highlight that these periods can generate significant mispricing for large aggregates.

A Appendix

A.1 Robustness

We perform several robustness checks. First we verify that our results are not solely driven by the 2008 financial crisis. We split the time period in half (by years): pre and post 2004 and test equation (4) in US, JPY, EUR. Results are robust in both sub-samples. It is difficult to split it up much more finely than this due to the low number of observations already in the sample (constraint is the options data). Table 12 reports the results.

The second set of robustness checks we run is to verify that the particular volatility maturity that we use (1 year) is not the only source of the results. To do this we run predictive regressions using other volatility maturities (from 1 month to 1 year): our results are highly robust to this. Table 13 reports the results.

A.2 Extracting Risk Neutral Moments

We follow a similar data cleaning methodology to Chang et al. (2013) to remove options with potentially erroneous quotes. In particular each day we remove options with prices of less than $.375, options that are in the money, days where there are less than 10 options quoted, options that have less than 10 days to maturity and options that violate arbitrage conditions. After this filtering, we extract implied volatilities from the remaining options. On each day, for each maturity we generate a volatility surface by interpolating the implied volatilities using a penalized cubic regression spline with a grid of 1000 evenly spaced strike to spot ratios (the strike to spot ratio is the ratio of the option strike to the index spot price) in $[.0001, 3]$. For values outside of the range available in the market we simply set the implied volatility to the nearest market available quote. This is to prevent the cubic spline from generating extreme (implausible) volatilities that would depend on the slope of the fit near potentially noisy tail options. Once we have a daily implied volatility surface, we compute the value of risk neutral skewness as follows. As in Bakshi et al. (2003), define:

$$V(t,\tau) \equiv \int_{S_t}^{\infty} \frac{2(1-\ln\frac{K}{S_t})}{K^2} C(t,\tau;K)dK + \int_0^{S_t} \frac{2(1+\ln\frac{S_t}{K})}{K^2} P(t,\tau;K)dK \tag{A.1}$$

$$W(t,\tau) \equiv \int_{S_t}^{\infty} \frac{6\ln\frac{K}{S_t} - 3(\ln\frac{K}{S_t})^2}{K^2} C(t,\tau;K)dK - \int_0^{S_t} \frac{6\ln\frac{S_t}{K} + 3(\ln\frac{S_t}{K})^2}{K^2} P(t,\tau;K)dK \tag{A.2}$$

$$\mu(t,\tau) \equiv e^{r\tau} - 1 - \frac{e^{r\tau}}{2}V(t,\tau) - \frac{e^{r\tau}}{6}W(t,\tau) - \frac{e^{r\tau}}{24}X(t,\tau) \tag{A.3}$$

$$RNSKEW(t,\tau) = \frac{e^{r\tau}W(t,\tau) - 3\mu(t,\tau)e^{r\tau}V(t,\tau) + 2\mu(t,\tau)^3}{(e^{r\tau}V(t,\tau) - \mu(t,\tau)^2)^{(3/2)}} \tag{A.4}$$

where $C(t,\tau;K)$ is the price of a call at time t with strike K and τ years until maturity, $P(t,\tau;K)$ is the price of a put at time t with strike K and τ years until maturity, and r is the risk free rate. $RNSKEW(t,\tau)$ is the value of risk neutral skewness as used by Chang et al. (2013). To be consistent in maturity with the VIX and previous studies of volatility such as Ang et al. (2006) we use the value of $\tau = 30/365$ and the VIX as a proxy for volatility. To build the constant maturity measures we interpolate using linear regression between measures computed using available maturities. The measures are then annualized (as the VIX is reported in annualized units).

References

Ang, A., Bali, T. and Cakici, N. (2012), 'The joint cross section of stocks and options', *Working* .

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X. (2006), 'The cross-section of volatility and expected returns', *The Journal of Finance* **61**(1), 259–299.

Backus, D., Foresi, S. and Wu, L. (2004), 'Accounting for biases in black-scholes', *Working Paper* .

Bakshi, G., Kapadia, N. and Madan, D. (2003), 'Stock return characteristics, skew laws, and the differential pricing of individual equity options', *The Review of Financial Studies* **16**(1), pp. 101–143.

Brunnermeier, M. K., Nagel, S. and Pedersen, L. H. (2008), 'Carry trades and currency crashes', *Working Paper* .

Campbell, J. and Vuolteenaho, T. (2004), 'Bad beta, good beta', *American Economic Review* **94**(5), 1249–1275.

Campbell, J. Y., Giglio, S., Polk, C. and Turley, R. (2013), An intertemporal capm with stochastic volatility, Technical report, Unpublished Manuscript.

Campbell, J. Y., Polk, C. and Vuolteenaho, T. (2010), 'Growth or glamour? fundamentals and systematic risk in stock returns', *Review of Financial Studies* **23**(1), 305–344.

Chan, L. K., Chen, H.-L. and Lakonishok, J. (2002), 'On mutual fund investment styles', *Review of financial studies* **15**(5), 1407–1437.

Chan, L. K. and Lakonishok, J. (2004), 'Value and growth investing: Review and update', *Financial Analysts Journal* pp. 71–86.

Chang, B. Y., Christoffersen, P. and Jacobs, K. (2013), 'Market skewness risk and the cross section of stock returns', *Journal of Financial Economics* **107**, 46 – 68.

Chen, J. (2002), 'Intertemporal capm and the cross-section of stock returns', *Working Paper* .

Chevalier, J. and Ellison, G. (1997), 'Risk taking by mutual funds as a response to incentives', *Journal of Political Economy* **105**(6), 1167–1200.

Cochrane, J. H. (2011), 'Presidential address: Discount rates', *The Journal of Finance* **66**(4), 1047–1108.

Duffie, D. (2010), 'Presidential address: Asset price dynamics with slow-moving capital', *The Journal of Finance* **65**(4), 1237–1267.

Fama, E. F. and French, K. R. (1992), 'The cross-section of expected stock returns', *the Journal of Finance* **47**(2), 427–465.

Fama, E. F. and French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of financial economics* **33**(1), 3–56.

Fama, E. F. and French, K. R. (1996), 'Multifactor explanations of asset pricing anomalies', *The Journal of Finance* **51**(1), 55–84.

Fama, E. F. and French, K. R. (1998), 'Value versus growth: The international evidence', *The Journal of Finance* **53**(6), 1975–1999.

Hong, H. and Stein, J. C. (1999), 'A unified theory of underreaction, momentum trading, and overreaction in asset markets', *The Journal of Finance* **54**(6), 2143–2184.

Lakonishok, J., Shleifer, A. and Vishny, R. W. (1994), 'Contrarian investment, extrapolation, and risk', *The journal of finance* **49**(5), 1541–1578.

Martin, I. W. (2013), 'Consumption-based asset pricing with higher cumulants', *The Review of Economic Studies* **80**(2), 745–773.

Newey, W. K. and West, K. D. (1994), 'Automatic lag selection in covariance matrix estimation', *The Review of Economic Studies* **61**(4), 631–653.

Petkova, R. and Zhang, L. (2005), 'Is value riskier than growth?', *Journal of Financial Economics* **78**(1), 187–202.

Sharpe, W. F. (1994), 'The sharpe ratio', *Journal of Portfolio Management* .

Tibshirani, R. (1996), 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society. Series B (Methodological)* **58**, pp. 267–288.

Table 1: $SKEW$ Forecasts of Future Market Excess Returns

We report the forecasting power of $SKEW$ for future market returns across the 1 month and 6 month horizon. We are careful to control for other variables that have been found to predict market returns in the literature.

Dependent Variable	$SKEW_{t-1}$	$R^e_{m,t-1}$	dp_{t-1}	dy_{t-1}	sey_{t-1}	tp_{t-1}	(Intercept)	R^2	N
$R^e_{m,t}$	0.260						-0.734	0.35%	201
	[1.221]						[-0.724]		
$R^e_{m,t}$	**0.377**	**0.162**					-1.357	2.31%	201
	[2.221]	**[1.978]**					[-1.429]		
$R^e_{m,t}$	**0.591**	0.171	**-3.659**	1.741	0.741	-0.315	**-4.078**	5.45%	201
	[2.176]	[1.934]	**[-2.051]**	[0.986]	[1.178]	[-1.031]	**[-2.291]**		
$R^e_{m,t\to t+5}$	**1.674**						-4.803	4.32%	196
	[2.025]						[-0.907]		
$R^e_{m,t\to t+5}$	**1.873**	0.272					-5.847	4.78%	196
	[2.346]	[1.233]					[-1.140]		
$R^e_{m,t\to t+5}$	2.142	0.300	-12.345	10.952	2.754	-1.208	-22.535	16.75%	196
	[1.516]	[1.305]	[-1.089]	[0.672]	[0.469]	[-0.466]	[-1.795]		

Table 2: Cross Sectional Sort by β to ε_{skew}

Multiple β of Fama-French 25 portfolio returns to innovations in $SKEW$; equation (3.3). Size of 1 indicates small stocks and size of 5 indicates large stocks. Value of 1 indicates low book-to-market stocks and value of 5 indicates high book-to-market stocks. Newey-West t-statistics are in brackets.

		Size						
		1	2	3	4	5	5-1	
Value	1	0.690	0.719	0.610	0.513	0.034	-0.627	[-0.942]
	2	0.300	0.246	0.132	-0.124	-0.130	-0.401	[-0.758]
	3	0.121	-0.158	-0.265	-0.275	-0.342	-0.434	[-0.959]
	4	-0.054	-0.225	-0.244	-0.166	-0.484	-0.401	[-0.917]
	5	-0.311	-0.140	-0.485	-0.146	-0.235	0.105	[0.211]
	5-1	**-0.972**	**-0.830**	**-1.066**	-0.630	-0.239		
		[-2.753]	**[-2.015]**	**[-2.455]**	[-1.594]	[-0.907]		

Table 3: Forecast of $R_{hml,t}$ Using $\sum_{j=1}^{4} \varepsilon_{skew,t-j}$ 1996-2012

Panel A shows summary statistics for HML in US, Japan (JPY) and Europe (EUR): all of these regions have a significant value effect. The results are from the perspective of a US investor (CAPM α is with respect to the US market). Panel B shows that lagged innovations in $SKEW$ have significant forecasting power for HML in all of these regions.

(a) HML Summary Statistics By Region

Region	$\overline{R_t}$	$\min R_t$	$\max R_t$	$\sigma(R_t)$	Sharpe	CAPM α	N
US	0.267	-12.600	13.840	3.495	0.265	0.347	201
JPY	0.457	-13.820	10.080	3.092	0.512	0.545	201
EUR	0.487	-9.570	10.960	2.653	0.636	0.481	201

(b) R_{hml} Forecasts

Region	$\sum_{j=1}^{j=4} \varepsilon_{skew,t-j}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-0.600**			**0.403**	10.87%	197
	[-6.010]			**[1.689]**		
US	**-0.576**	-0.154		**0.468**	15.07%	197
	[-5.159]	[-1.406]		**[1.773]**		
US	**-0.564**	-0.419	0.053	**0.428**	17.00%	197
	[-6.190]	[-1.692]	[1.403]	**[1.785]**		
JPY	**-0.255**			**0.505**	2.09%	197
	[-2.312]			**[1.939]**		
JPY	**-0.227**	**-0.183**		**0.582**	9.89%	197
	[-2.648]	**[-3.560]**		**[2.380]**		
JPY	**-0.223**	**-0.268**	0.017	**0.569**	9.74%	197
	[-2.412]	**[-3.108]**	[1.048]	**[2.399]**		
EUR	**-0.376**			**0.554**	7.15%	197
	[-4.621]			**[2.052]**		
EUR	**-0.379**	0.023		**0.545**	6.84%	197
	[-4.849]	[0.302]		**[1.837]**		
EUR	**-0.371**	-0.147	0.034	**0.520**	8.02%	197
	[-4.169]	[-0.822]	[1.261]	**[1.876]**		

Table 4: Forecast of $R_{hml,t}$ Using $SKEW_{t-1}$ 1996-2012

Reports the results of equation (3.9): forecasts of HML returns using the $SKEW$ level which is useful in turning our results into a trading strategy that avoids any look-ahead in parameters.

Region	$SKEW_{t-1}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-0.524**			**2.695**	6.26%	201
	[-3.778]			**[4.217]**		
US	**-0.484**	-0.154		**2.581**	10.41%	201
	[-3.506]	[-1.559]		**[4.298]**		
US	**-0.507**	-0.480	0.065	**2.650**	13.54%	201
	[-4.396]	[-1.885]	[1.568]	**[4.671]**		
JPY	**-0.285**			**1.778**	2.05%	201
	[-2.203]			**[3.067]**		
JPY	**-0.238**	**-0.180**		**1.646**	9.58%	201
	[-2.014]	**[-3.573]**		**[3.152]**		
JPY	**-0.246**	**-0.286**	0.021	**1.668**	9.61%	201
	[-2.002]	**[-3.131]**	[1.253]	**[3.084]**		
EUR	**-0.380**			**2.245**	5.65%	201
	[-3.651]			**[4.756]**		
EUR	**-0.386**	0.025		**2.263**	5.39%	201
	[-3.602]	[0.346]		**[4.221]**		
EUR	**-0.401**	-0.181	0.041	**2.307**	7.41%	201
	[-3.203]	[-0.956]	[1.397]	**[3.818]**		

Table 5: Forecast of $R_{hml,t}$ Using $-\kappa_{3,t-1}$ 1996-2012

Forecasts of HML using the level of $-\kappa_{3,t-1}$ to verify that our $SKEW$ metric's ability to forecast HML isn't due to some incidental characteristic but is directly related to crash risk.

Region	$-\kappa_{3,t-1}$	$R_{m,t}$	$-\kappa_{3,t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-2.424**			**1.130**	5.18%	201
	[-4.469]			**[3.530]**		
US	**-2.294**	-0.161		**1.159**	9.76%	201
	[-3.935]	[-1.328]		**[3.472]**		
US	**-2.773**	**-0.492**	**0.667**	**1.311**	23.37%	201
	[-5.728]	**[-4.102]**	**[4.397]**	**[4.024]**		
JPY	**-1.249**			**0.902**	1.42%	201
	[-1.999]			**[2.680]**		
JPY	**-1.102**	**-0.183**		**0.935**	9.30%	201
	[-2.282]	**[-3.494]**		**[3.064]**		
JPY	**-1.237**	**-0.277**	**0.189**	**0.978**	10.29%	201
	[-2.234]	**[-3.929]**	**[2.671]**	**[3.204]**		
EUR	**-1.592**			**1.054**	3.75%	201
	[-3.228]			**[3.082]**		
EUR	**-1.608**	0.019		**1.050**	3.39%	201
	[-3.283]	[0.246]		**[2.816]**		
EUR	**-1.901**	-0.183	**0.408**	**1.143**	12.04%	201
	[-2.826]	[-1.944]	**[3.601]**	**[3.128]**		

Table 6: Robustness to the VIX and $RNSKEW$

We show that $SKEW$ contains the relevant information for forecasting HML returns even after controlling for VIX or $RNSKEW$.

Dependent Variable	$SKEW_{t-1}$	VIX_{t-1}	$RNSKEW_{t-1}$	(Intercept)	R^2	N
$R_{hml,t}$	**-0.523**			**2.690**	6.23%	201
	[-3.771]			**[4.210]**		
$R_{hml,t}$		**-0.088**		**2.210**	3.66%	201
		[-2.980]		**[3.484]**		
$R_{hml,t}$			1.013	2.054	0.93%	201
			[1.683]	[1.841]		
$R_{hml,t}$	**-0.440**	-0.028		**2.918**	6.00%	201
	[-2.219]	[-0.602]		**[4.118]**		
$R_{hml,t}$	**-0.493**		0.509	**3.450**	6.10%	201
	[-3.533]		[0.878]	**[2.881]**		

Table 7: Extending the Sample: 1963 - 1996

We extend the sample to 1963 by replicating $SKEW$ using other variables, termed $\widehat{SKEW}$. This table reports the results of forecasting returns on HML using innovations in $\widehat{SKEW}$ and the market using $\widehat{SKEW}$ from 1963 - 1996.

Dependent Variable	$\hat{\varepsilon}_{skew,t-1}$	$R_{m,t}$	$\widehat{SKEW}_{t-1} \cdot R_{m,t}$	$\widehat{SKEW}_{t-1}$	$R_{m,t-1}$	(Intercept)	R^2	N
$R_{hml,t}$	**-0.519**					**0.461**	1.49%	396
	[-2.472]					**[3.078]**		
$R_{hml,t}$	**-0.413**	**-0.198**				**0.551**	12.70%	396
	[-2.274]	**[-5.239]**				**[3.936]**		
$R_{hml,t}$	**-0.423**	**-0.217**	0.006			**0.548**	12.50%	396
	[-2.069]	**[-2.400]**	[0.228]			**[3.977]**		
$R_{m,t}$				**0.443**		-0.781	0.85%	396
				[2.107]		[-1.277]		
$R_{m,t}$				**0.479**	0.071	-0.914	1.09%	396
				[2.360]	[1.364]	[-1.512]		

Table 8: Corporate Bond Returns and ε_{skew}

We document that AAA bonds perform well when $SKEW$ increases while CCC bonds perform poorly. Thus the trade that goes long CCC and short AAA bonds behaves like HML with respect to innovations in $SKEW$.

Dependent Variable	$\varepsilon_{skew,t}$	$\varepsilon_{vix,t}$	$R_{m,t}$	(Intercept)	R^2	N
$R_{aaa,t}$	0.228	-0.049	-0.006	**0.271**	1.08%	201
	[1.699]	[-0.532]	[-0.143]	**[2.741]**		
$R_{aa,t}$	0.154	-0.066	0.006	**0.277**	2.65%	201
	[1.527]	[-0.876]	[0.142]	**[2.734]**		
$R_{a,t}$	0.167	-0.113	0.021	**0.298**	8.28%	201
	[1.565]	[-1.053]	[0.419]	**[2.478]**		
$R_{bbb,t}$	-0.117	-0.055	0.072	**0.338**	11.48%	201
	[-0.798]	[-0.688]	[1.506]	**[2.547]**		
$R_{bb,t}$	-0.296	-0.117	**0.166**	**0.392**	39.26%	190
	[-1.454]	[-1.789]	**[2.948]**	**[2.384]**		
$R_{b,t}$	**-0.427**	**-0.115**	**0.259**	0.258	46.08%	190
	[-2.167]	**[-2.216]**	**[5.063]**	[1.549]		
$R_{ccc,t}$	**-0.720**	-0.121	**0.454**	0.337	46.06%	190
	[-2.259]	[-1.532]	**[3.696]**	[1.247]		
$R_{ccc,t} - R_{aaa,t}$	**-0.945**	-0.070	**0.465**	-0.181	44.34%	190
	[-2.413]	[-0.854]	**[3.810]**	[-0.650]		

Table 9: Bond Return Forecasts Using $\varepsilon_{skew,t-1}$

This table shows that lagged innovations in $SKEW$ are able to forecast corporate bond returns highlighting that this asset class is also slow to incorporate information from the options market into prices.

Dependent Variable	$\varepsilon_{skew,t-1}$	$R_{m,t}$	$\widehat{SKEW}_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
$R_{aaa,t}$	**0.397**	0.034	-0.006	**0.276**	7.45%	200
	[2.374]	[0.493]	[-0.494]	**[2.636]**		
$R_{aa,t}$	**0.318**	0.051	-0.004	**0.271**	6.49%	200
	[2.165]	[0.761]	[-0.401]	**[2.653]**		
$R_{a,t}$	**0.365**	0.146	-0.014	**0.276**	10.04%	200
	[2.176]	[1.365]	[-0.954]	**[2.127]**		
$R_{bbb,t}$	0.233	0.283	-0.032	**0.330**	16.27%	200
	[1.856]	[1.946]	[-1.438]	**[2.327]**		
$R_{bb,t}$	0.073	**0.445**	-0.033	**0.334**	36.43%	190
	[0.693]	**[2.537]**	[-1.271]	**[2.030]**		
$R_{b,t}$	-0.262	**0.500**	-0.024	0.198	43.02%	190
	[-1.670]	**[3.386]**	[-1.063]	[1.090]		
$R_{ccc,t}$	-0.376	**0.569**	0.009	0.226	43.44%	190
	[-1.504]	**[2.036]**	[0.201]	[0.793]		
$R_{ccc,t} - R_{aaa,t}$	**-0.775**	**0.545**	0.014	-0.277	43.70%	190
	[-2.897]	**[2.090]**	[0.337]	[-0.914]		

Table 10: Flows Into Value and Growth Funds

Mutual funds are sorted into style quintiles according to $\beta_{hml,t-1}$ in a regression of $R_{f,t} = a + \beta_{m,t}R^e_{m,t} + \beta_{smb,t}R_{smb,t} + \beta_{hml,t}R_{hml,t} + \beta_{umd,t}R_{umd,t}$ - where $R_{f,t}$ is the return of fund f at time t - as in Chan et al. (2002) among others. $FLOW_{f,t} \equiv \frac{TNA_{f,t}-TNA_{f,t-1}(1+R_{f,t})}{TNA_{f,t-1}}$ is computed for each fund f and month t. To control for the flow performance relationship cumulative returns over the past year are also computed and denoted by $R_{f,t-12\rightarrow t-1}$. Then for each quintile an asset (TNA_{t-1}) weighted average is taken across $FLOW_{b,t} \equiv \sum_{f\in b} w_{f,t-1}FLOW_{f,t}$ and $R_{b,t-12\rightarrow t-1} \equiv \sum_{f\in b} w_{f,t-1}R_{f,t-12\rightarrow t-1}$. Then for each quintile we regress $FLOW_t = a + \beta_{skew}\varepsilon_{skew,t-1} + \beta_{ret}R_{t-12\rightarrow t-1} + \nu_t$. The table shows that an increase in $SKEW_t$ causes investors to pull money away from value funds.

Value Style Quintile	$\varepsilon_{skew,t-1}$	$R_{t-12\rightarrow t-1}$	(Intercept)	R^2	N
1	0.007		-0.182	-0.50%	199
	[0.164]		[-1.036]		
	0.035	**0.016**	-0.362	16.88%	199
	[0.821]	**[2.678]**	[-3.111]		
2	-0.099		0.089	1.31%	199
	[-1.988]		[1.002]		
	-0.082	**0.010**	-0.019	5.22%	199
	[-1.592]	**[1.597]**	[-0.168]		
3	-0.027		0.007	-0.14%	199
	[-0.632]		[0.167]		
	-0.012	**0.008**	-0.066	4.30%	199
	[-0.258]	**[2.120]**	[-1.072]		
4	**-0.137**		-0.055	4.25%	199
	[-3.826]		[-0.518]		
	-0.117	**0.009**	-0.147	7.82%	199
	[-2.507]	**[2.192]**	[-1.535]		
5	**-0.208**		0.057	4.73%	199
	[-2.581]		[0.377]		
	-0.164	**0.017**	-0.139	12.76%	199
	[-2.039]	**[3.289]**	[-0.792]		
5-1	**-0.215**		0.239	1.66%	199
	[-2.076]		[0.675]		
	-0.163	**0.057**	0.215	53.75%	199
	[-1.980]	**[3.742]**	[1.387]		

Table 11: Unexpected Flows Into Value and Growth Funds

This table reports the results of equation (4.4) showing that innovations in $FLOW$ (relative to past values of $FLOW$) are predictable using lagged innovations in $SKEW$.

Value Style Quintile	$\varepsilon_{skew,t-1}$	$R_{t-12\to t-1}$	(Intercept)	R^2	N
1	**0.045**		-0.038	0.04%	199
	[2.108]		[-0.930]		
	0.042	-0.002	-0.020	-0.06%	199
	[1.208]	[-0.685]	[-0.412]		
2	-0.032		-0.019	-0.17%	199
	[-0.784]		[-0.421]		
	-0.035	-0.002	0.000	-0.44%	199
	[-0.868]	[-0.558]	[0.001]		
3	-0.017		-0.027	-0.36%	199
	[-0.364]		[-0.808]		
	-0.022	-0.003	-0.001	-0.17%	199
	[-0.479]	[-0.793]	[-0.011]		
4	-0.051		-0.025	0.62%	199
	[-1.600]		[-0.694]		
	-0.059	-0.003	0.009	1.02%	199
	[-1.490]	[-1.328]	[0.178]		
5	-0.050		0.007	0.32%	199
	[-0.961]		[0.175]		
	-0.047	0.001	-0.006	-0.09%	199
	[-0.895]	[0.826]	[-0.144]		
5-1	**-0.102**		0.013	1.40%	199
	[-2.328]		[0.238]		
	-0.104	-0.002	0.014	1.10%	199
	[-2.383]	[-0.372]	[0.259]		

Table 12: Robustness to Sub-Samples

We show that our results are robust to sub-samples by splitting the time-series in two. A finer split is difficult due to the lack of S&P 500 options data going back further in time.

(a) 1996-2004

Region	$SKEW_{t-1}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-0.793**			**3.799**	6.80%	95
	[-2.872]			**[3.162]**		
US	**-0.663**	**-0.473**		**3.487**	36.48%	95
	[-2.982]	**[-5.790]**		**[4.020]**		
US	**-0.685**	**-0.535**	0.014	**3.571**	35.82%	95
	[-2.996]	**[-2.010]**	[0.276]	**[3.788]**		
JPY	-0.399			**2.234**	1.50%	95
	[-1.187]			**[1.739]**		
JPY	-0.327	**-0.264**		**2.060**	12.72%	95
	[-1.097]	**[-3.280]**		**[1.867]**		
JPY	-0.383	**-0.423**	0.036	**2.280**	12.16%	95
	[-1.408]	**[-1.964]**	[0.651]	**[2.268]**		
EUR	**-0.650**			**3.777**	9.71%	95
	[-2.468]			**[3.688]**		
EUR	**-0.600**	**-0.182**		**3.657**	18.02%	95
	[-2.383]	**[-2.251]**		**[4.002]**		
EUR	**-0.643**	-0.305	0.028	**3.826**	17.50%	95
	[-2.330]	[-1.240]	[0.539]	**[3.705]**		

(b) 2004-2012

Region	$SKEW_{t-1}$	$R_{m,t}$	$SKEW_{t-1} \cdot R_{m,t}$	(Intercept)	R^2	N
US	**-0.360**			**1.918**	6.34%	106
	[-2.712]			**[2.783]**		
US	**-0.412**	**0.189**		**2.095**	18.60%	106
	[-2.948]	**[2.190]**		**[2.808]**		
US	**-0.394**	0.003	0.033	**1.989**	20.35%	106
	[-3.230]	[0.029]	[1.368]	**[3.159]**		
JPY	**-0.218**			**1.470**	2.34%	106
	[-2.351]			**[2.843]**		
JPY	**-0.193**	-0.090		**1.386**	5.06%	106
	[-2.213]	[-1.851]		**[2.837]**		
JPY	**-0.190**	**-0.124**	0.006	**1.367**	4.24%	106
	[-2.227]	**[-2.062]**	[0.725]	**[2.841]**		
EUR	-0.136			**0.700**	0.43%	106
	[-1.654]			**[1.960]**		
EUR	**-0.204**	**0.245**		**0.930**	28.91%	106
	[-2.367]	**[5.274]**		**[2.722]**		
EUR	-0.193	0.135	0.020	**0.867**	29.40%	106
	[-1.948]	[1.829]	[1.255]	**[2.012]**		

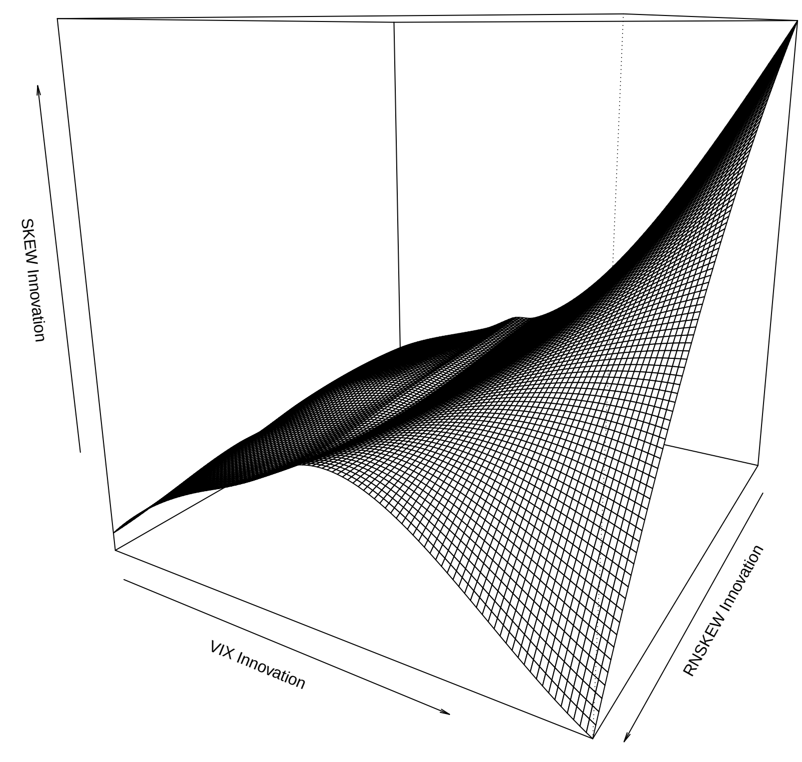
Table 13: Robustness: Other S&P 500 Implied Volatility Maturities

To demonstrate that our results are robust to using other maturity points on the S&P 500 volatility surface, we run two regressions: $R_{hml,t} = a + \beta_m R^e_{m,t} + \beta_s \varepsilon_{skew,t} + \epsilon_t$ where $\varepsilon_{skew,t}$ is the innovation in $SKEW_t$ from an AR(1) model (this validates the contemporaneous relationship). The other regression is the forecasting regression of $R_{hml,t} = a + \beta_s SKEW_{t-1} + \epsilon_t$ to show that one can forecast HML with other maturities as well.

Vol Tenure	$R_{m,t}$	$\varepsilon_{skew,t}$	$SKEW_{t-1}$	(Intercept)	R^2	N
1 month	**-0.246**	**-0.422**		0.398	10.56%	200
	[-1.978]	**[-3.496]**		[1.375]		
			-0.252	**1.306**	2.34%	200
			[-2.521]	**[3.036]**		
3 months	-0.274	**-0.623**		0.416	10.57%	200
	[-1.928]	**[-3.349]**		[1.430]		
			-0.442	**2.266**	5.65%	200
			[-3.924]	**[4.349]**		
6 months	-0.264	**-0.717**		0.413	9.85%	200
	[-1.807]	**[-3.026]**		[1.398]		
			-0.544	**2.834**	7.39%	200
			[-4.188]	**[4.507]**		
9 months	-0.276	**-0.876**		0.422	10.72%	200
	[-1.905]	**[-3.230]**		[1.423]		
			-0.539	**2.803**	6.69%	200
			[-3.934]	**[4.326]**		
12 months	-0.265	**-0.817**		0.417	9.43%	200
	[-1.849]	**[-3.001]**		[1.385]		
			-0.552	**2.849**	6.86%	200
			[-3.922]	**[4.347]**		

Figure 1: Nonparametric Relationship Between $SKEW$, VIX and $RNSKEW$ Innovations

We show a nonparametric surface that relates innovations in $SKEW$ to innovations in the VIX and $RNSKEW$. The surface is the result of a local polynomial regression using second order polynomials and 75% span. Below the plot we present the parametric version of the relationship from equation (2.9).



Dependent Variable	$\varepsilon_{rnskew,t}$	$\varepsilon_{rnskew,t} * \varepsilon_{vix,t}$	$\varepsilon_{vix,t}$	(Intercept)	R^2	N
$\varepsilon_{skew,t}$	**-0.167**	**-0.266**	**0.516**	0.062	41.02%	201
	[-3.421]	**[-2.074]**	**[9.369]**	[1.013]		

Figure 2: Time Series Plot of $SKEW$

Plot of $SKEW$, VIX and $OrthSKEW$ ($SKEW$ orthogonalized to the VIX through linear regression) from 1996-2012.

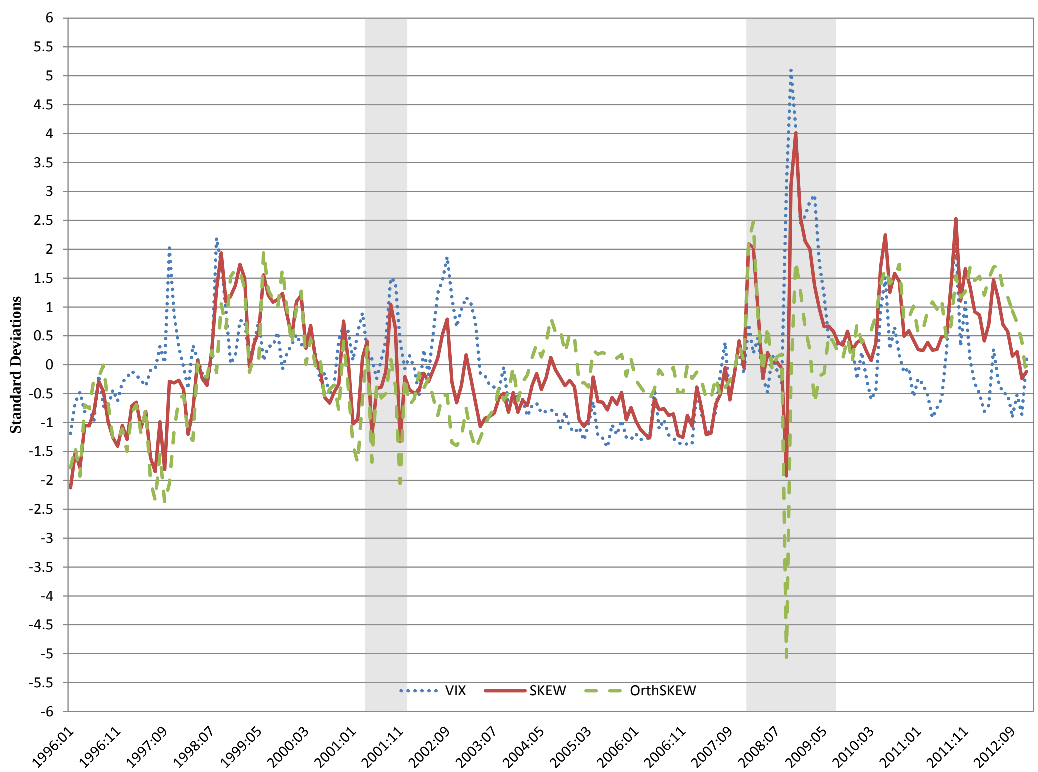


Figure 3: HML Underreaction to $SKEW$ Innovations

This plot demonstrates the underreaction of $R_{hml,t}$ to innovations in $SKEW_t$. We run regressions of $R_{hml,t} = a + \beta_m R_{m,t} + \beta_{\varepsilon,j}\varepsilon_{skew,t-j} + \epsilon_t$ for $j = 0...12$ and plot $\sum_{k=0}^{j} \beta_{\varepsilon,k}$. The error bars are plotted in dashes and assume that estimates of $\beta_{\varepsilon,j}$ are independent. As is clear from the plot, there is a significant predictability to $R_{hml,t}$ based on lagged innovations in $SKEW_t$.

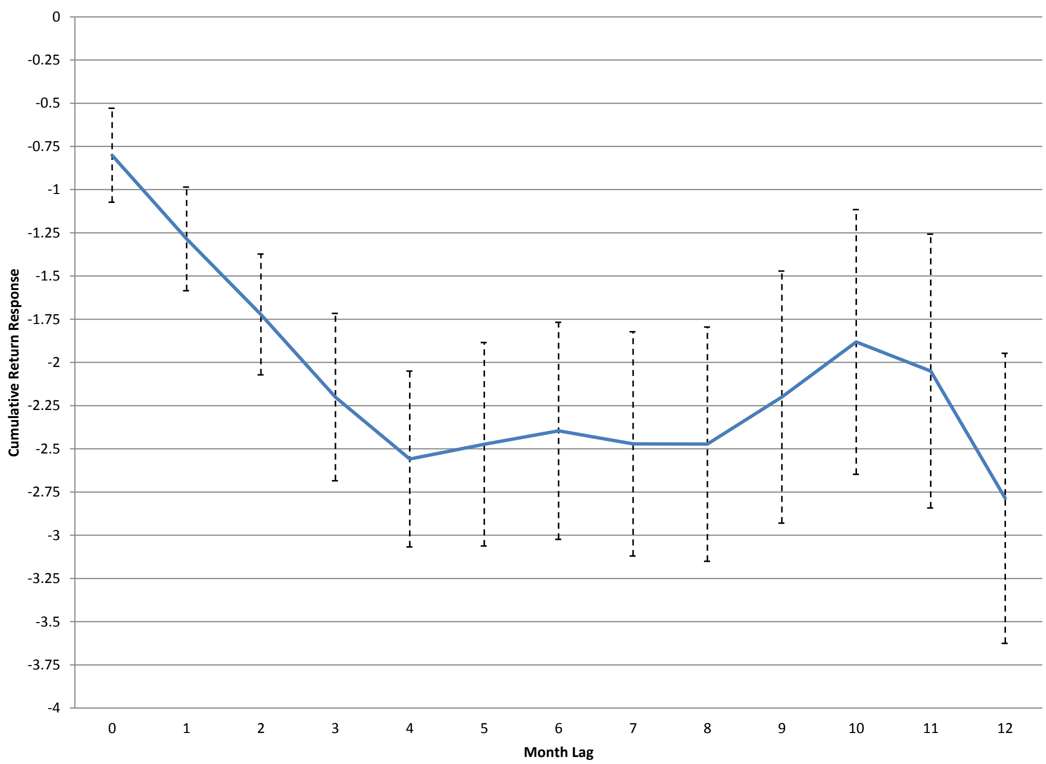


Figure 4: Cumulative Returns to Passive and Active HML Strategy

This plot demonstrates the out-of-sample forecasting performance of $SKEW_t$ in timing returns to $R_{hml,t}$. To construct the strategy we run a forecasting regression of $R_{hml,t-1} = a + \beta_s SKEW_{t-2} + \epsilon_{t-1}$ using 36 months of lagged data (this is the burn in period). Based on this model we compute the forecast for next month's HML return, $\hat{R}_{hml,t}$. To determine how much to invest each month we assume an endowment of $\mathcal{W}_{t-1}$ and Build a portfolio by putting $w_{hml,t-1}$ of it into HML and $(1 - |w_{hml,t-1}|)$ into the risk free rate where $w_{hml,t-1} = \tanh\left(\frac{\hat{R}_{hml,t}}{\sqrt{\frac{1}{J}\sum_{j=1}^{t}(\hat{R}_{hml,j} - \bar{\hat{R}}_{hml,t})^2}}\right)$. This is the hyperbolic tangent of the forecasted HML return scaled by the standard deviation of previous HML forecasts. The hyperbolic tangent is applied so that $w_{hml,t-1} \in [-1, 1]$. The return to this portfolio is $1 + R_{p,t} = 1 + w_{hml,t-1}R_{hml,t} + (1 - |w_{hml,t-1}|)R_{free,t}$. The mechanism described avoids look ahead bias in values and parameters. This cumulative return is plotted in the figure and labeled Active HML. Passive HML corresponds to monthly rebalancing strategy that invests equal weights into being long "H" and short "L".

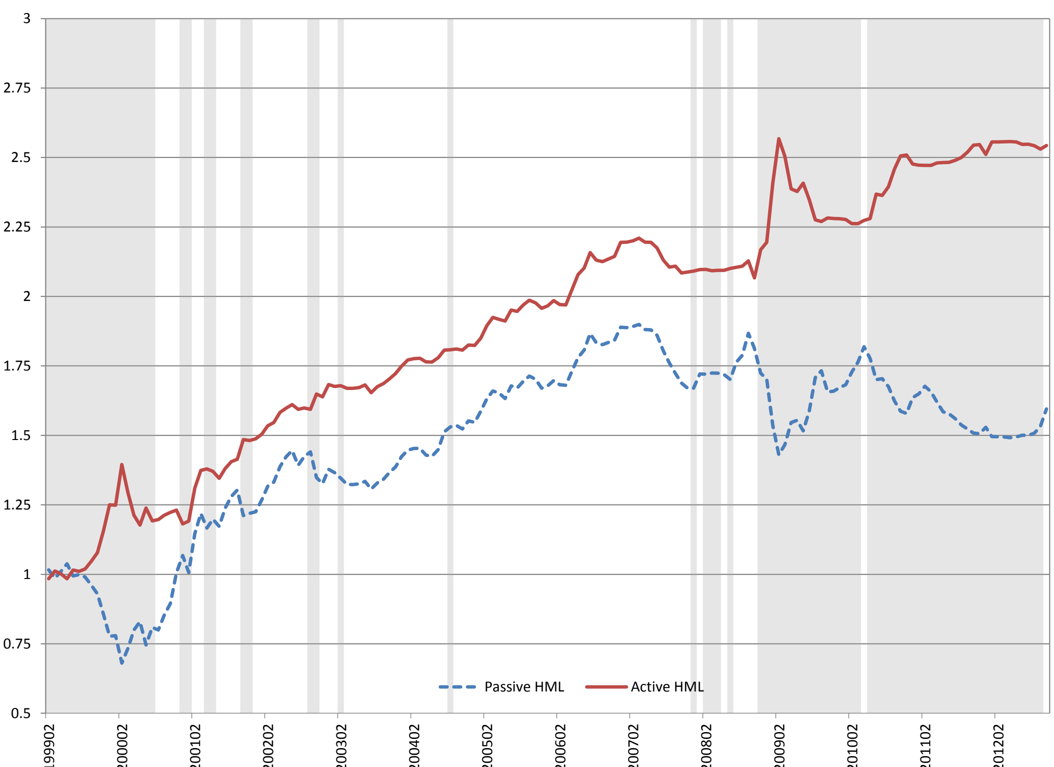


Figure 5: Cumulative CCC-AAA Response to ε_{skew}

Plot of the cumulative response to innovations in $SKEW$ of a trade that goes long CCC and short AAA bonds. Corporate bonds, like equities, are also slow to fully incorporate all information from $SKEW$.