A HIGH-LOW PRICE ANOMALY

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Mitchell Johnston

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Huseyin Gulen, Chair
Krannert School of Management
Dr. Sergey Chernenko
Krannert School of Management
Dr. Deniz Yavuz
Krannert School of Management
Dr. Adem Atmaz
Krannert School of Management

Approved by:
Dr. Yanjun Li
Head of the Graduate Program

ACKNOWLEDGMENTS

I would like to thank the members of my committee Huseyin Gulen (Chair), Deniz Yavuz, Sergey Chernenko and Adem Atmaz as well as Xiaoyan Zhang, Ravi Jagannathan, Kateryna Holland, Chengdong Yin, Ben McCartney, Logan Emery, Mike Woeppel and seminar participants at Purdue University for their helpful comments and suggestions. I would also like to thank Meghan Leiseberg for her unending support. I would also like to thank my family George, Jack, Susan, Charles, and Russ Johnston for all their support and guidance.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

ABSTRACT

Johnston, Mitchell Ph.D., Purdue University, August 2019. A High-Low Price Anomaly. Major Professor: Huseyin Gulen.

I examine movements in the closing price that are different than the movements of the high and low prices on a given day. Instances in which the closing price deviates from the movements in the midpoint between the high and low are a strong predictor of future abnormal returns. The predictive power of the *HLDiff* measure holds across size groups and sub-periods and holds in the presence of other common determinants of stock returns. The predictive power of *HLDiff* appears to be driven by the existence of market frictions. In particular, I find that an instrument for *HLDiff* based on attention proxies appears to account for the positive association with future returns. I also construct a factor based on *HLDiff* and find that the factor is consistent with market frictions and improves the pricing ability of the single-factor and five-factor models.

1. INTRODUCTION

Closing prices underlie many financial calculations. Academic research relies on the close to calculate returns and market capitalization. Mutual funds rely heavily on the close to determine their net asset value on a given day. Financial news often uses the closing price when reporting on a security. However, the close of the market has been shown to be unique. [1]. Ultimately, the close may move in an independent way relative to earlier prices. Given the prominence of the close price in financial calculations, it would be interesting to investigate movements in the closing price that are not reflective of movements in other prices, in particular I select the midpoint between the high and low price on the same day.

In particular, I calculate returns to the midpoint between the high and low price and compare the midpoint return to the closing price return. The difference between the high-low midpoint return and the closing return is accumulated over the preceding month. The use of the high-low midpoint is motivated by empirical literature that has shown the high and low to be informative (Beckers (1983), Yang and Zhang (2000), Corwin and Schultz (2012), and Abdi and Ranaldo (2017)). Beyond this literature, the high and low prices exhibit a variety of useful characteristics: (1) The high and low prices are readily observable to investors and are often quoted alongside closing prices, (2) the high and low bound all trading activity on a given day, and (3) the high and low are also unlikely to both be influenced by any unique trading mechanisms that may exist prior the exchange closure. To the extent that the high and low are genuinely informative, the close price would ideally reflect the movements in these prices such that the deviation of the close return from the high-low average return is not informative about future returns. Given the informativeness of the high and low

[1] These papers include but are not limited to: Admati and Pfleiderer (1988), Slezak (1994), Hong and Wang (2000), Bogousslavsky (2016), McInish and Wood (1992), Cushing and Madhavan (2000), Bogousslavsky (2017), Carhart, Kaniel, Musto, and Reed (2002), and Harris (1989).

and the theoretical and empirical research about the close, I hypothesize that market frictions may result in some of the information revealed in the high and low being omitted from the close. The extent to which prices fail to properly reflect information revealed in prior prices would be consistent with the closing price containing a degree of mispricing.

Mispricing that arises as a result of incomplete markets or information frictions is consistent with the investor attention hypothesis (Merton (1987)), in which a more visible firm will have information incorporated more quickly. Inadequate attention in this model is associated with the manifestation of various anomalies and variation in the timing of the correction of these anomalies. While attention does not account for all market frictions, it is a reasonable starting point to consider given the theoretical model of Merton (1987) and the empirical findings of Hou and Moskowitz (2005).Relying on the qualities of the high and low prices and the fact that frictions may influence the efficiency of information incorporation. I develop the measure *HLDiff* to capture the movements in the high and low prices not reflected in the closing price. Specifically, *HLDiff* is the monthly accumulated difference between the high-low midpoint return and the close return. Using all stocks traded on the NYSE, NASDAQ, and AMEX from 1965 through 2016, I document a strong positive correlation between the size of *HLDiff* and subsequent monthly stock returns. The highest decile portfolio formed based on *HLDiff* is associated with a value-weighted return of 1.15%. The lowest decile portfolio is associated with a value-weighted return of -0.36%. After adjusting for the Fama-French five-factor model, the spread portfolio alpha is economically and statistically significant at 1.52% (t-stat: 9.69). To verify the robustness of these results, I form three subgroups based on the market capitalization at the end of each month. The raw return and risk-adjusted alpha results are similar for these subgroups. Results also persist after dividing the sample period into sub-periods: 1965-1982, 1983-2000 and 2001-2016.

To further disentangle the effect of *HLDiff* from other return determinants, I run Fama-MacBeth regressions and find that the *HLDiff* effect remains positive and significant. Consistent with the portfolio results, *HLDiff* is also significant across sub-periods and size sub-groups after controlling for a variety of other common return determinants. The control variables included in the Fama-MacBeth regressions are: prior month return, prior six month return (excluding the most recent month), monthly turnover, monthly return volatility, log market capitalization (size), log book-to-market, and a proxy for bid-ask spread. The effect is strongest for small firms. While the small market capitalization group presents the largest coefficient and significance, the result is present for all subgroups. I also find that the significance of *HLDiff* is of similar magnitude or larger than all other control variables included in the Fama-MacBeth regression.

The construction of *HLDiff* makes it unlikely that the return predictability result is an artifact of market microstructure noise issues such as non-trading or bid-ask bounce. However, I adopt a variety of other safeguards to address market microstructure noise concerns such as excluding securities below $1 from my testing, using the bid-ask midpoint at the close to calculate returns, and adopting the weighted least squares (WLS) approach suggested by Asparouhova, Bessembinder, and Kalcheva (2010). Results reflecting both of these adjustments are reported in the robustness tests section. In the presence of these alternative procedures and adjustments, *HLDiff* continues to have a positive and significant relationship with subsequent returns.

Given the significance of the return predictability, I consider a variety of alternative drivers of the premium associated with *HLDiff*. First, I evaluate the timing of the return realization in detail by looking varying return horizons from one to twelve weeks. Hou and Moskowitz (2005) note that 'most stocks respond to information within a month's time'. Consistent with this, I find that the return predictability associated with *HLDiff* is almost entirely realized within the first four weeks. In addition, the relationship between *HLDiff* and returns is never negative. A buy and hold portfolio over the same 12 weeks presents similar positive and significant returns. All of these findings are consistent with a delayed response to information revealed by the high and low prices. *HLDiff* is also stronger when price delay is higher. That is to say when a stock has a delayed response to market information, the relationship between *HLDiff* and subsequent period returns is stronger.

These results are strongly indicative of the fact that the closing price does not always reflect the information contained in earlier prices. This would be consistent with a degree of mispricing for neglected securities when using the closing price to calculate returns. In addition, it is possible that other frictions may delay the closing price in responding to information. Miller (1977) argues that short selling constraints may restrict the ability of rational investors to rectify perceived overpricing and incorporate their information into prices. Stambaugh, Yu, and Yuan (2012) extend the logic of Miller (1977) by connecting short selling constraints to investor sentiment. They further demonstrate that following a period of high investor sentiment overpricing is more likely to exist. Stambaugh and Yuan (2016) utilize this motivation, in part, to develop a mispricing factor model as it is possible that investor sentiment may represent a systematic risk that rational investors require compensation for bearing. The closing price failing to reflect information contained in other prices is consistent with mispricing which is further solidified by the fact that the premium associated with *HLDiff* appears to be driven by a variety of frictions including attention and short selling constraints.

Given these facts, I construct a new factor and evaluate its pricing ability relative to 12 prominent anomalies. The 12 anomalies considered are the same as those examined by Stambaugh, Yu, and Yuan (2012) and Stambaugh and Yuan (2016) [2] to which I add idiosyncratic volatility[3]. Stambaugh, Yu, and Yuan (2012) note that these anomalies are consistent with mispricing. Accordingly, Stambaugh and Yuan (2016) use these anomalies in the construction of their mispricing factor model. To the extent that *HLDiff* captures a degree of mispricing, I anticipate that this factor may

[2]The anomalies are distress, O-score, net stock issues, composite equity issues, total accruals, net operating assets, gross profitability, asset growth, return on assets, investment to assets, and momentum.

[3]Stambaugh, Yu, and Yuan (2012) note a connection between mispricing and idiosyncratic volatility. Specifically, they note that higher idiosyncratic volatility translates into higher arbitrage risk which allows greater mispricing.

explain some of the alpha associated with these anomalies. After adding the *HLDiff* factor to the single-factor CAPM, the alpha associated with all but one anomaly is reduced in both magnitude and significance, with distress, gross profitability, and return on assets becoming insignificant under the single-factor model including the *HLDiff* factor.

As a further test of the model including the *HLDiff* factor, I conduct a Gibbons et al. (1989) (GRS) test. The null hypothesis of this test is that the selected factor model is sufficient to price the test assets. Under the single-factor and five-factor models, the GRS test rejects the null at the 1% level when *HLDiff* is excluded. However, after the addition of the *HLDiff* factor, the GRS test is unable to reject the null at the 1% level that the alphas for these 12 anomalies jointly equal zero for the given model. Stated differently, I am unable to reject the null at the 1% level that the two-factor and six-factor models including *HLDiff* factor are sufficient to jointly price the selected anomalies. All of these facts indicate that the inclusion of the *HLDiff* factor represents an improvement to these two factor models.

As a more direct examination of the relationship between *HLDiff* and mispricing, I conduct two additional batteries of tests. First, I regress the *HLDiff* factor on the Baker and Wurgler (2006) sentiment index and find that the sentiment index positively and significantly predicts the *HLDiff* factor loading. Further, I find that this relationship is stronger in the short leg portion of the spread portfolio consistent with Stambaugh, Yu, and Yuan (2012), who attribute a component of mispricing to short selling constraints. Second, I compare the *HLDiff* two-factor model to the mispricing M-4 model of Stambaugh and Yuan (2016) and the Fama French five-factor model. Consistent with *HLDiff* being associated with mispricing, the *HLDiff* two-factor model prices the performance mispricing factor while both the five-factor and M-4 models fall short of pricing the *HLDiff* factor. All of these findings support the fact that *HLDiff* is consistent with mispricing, and the relationship is separate and distinct from the factors in both of the competing models.

While traditional asset pricing models assume that stock markets are frictionless, numerous studies, both theoretical and empirical, document the existence of various frictions for which investors may demand compensation. It is possible, for instance, for investors to demand compensation for noisy prices which will exacerbate information asymmetry frictions. In this context, it would make sense to construct a factor that captures the aggregate level of mispricing. This is similar to approaches adopted by Hirshleifer and Jiang (2010) and Kozak, Nagel, and Santosh (2018). Hirshleifer and Jiang (2010) identify common misvaluation across firms and use this in the construction of a factor. Kozak, Nagel, and Santosh (2018) construct a model where fully rational investors trade with investors with distorted beliefs and deviations from the CAPM are caused by sentiment. The implication is that behavioral explanations for return premiums could be just as valid an explanation for the efficacy for factor models as rational risk-based explanations. As *HLDiff* appears to be related to sentiment and is predictive of future returns, the construction of a factor based on *HLDiff* is reasonable. Alternatively, it is possible that market-wide noise-trader sentiment may represent a systematic risk for which rational investors require compensation for bearing (Long, Shleifer, Summers, and Waldmann (1990)). Stambaugh, Yu, and Yuan (2012) present findings consistent with investor sentiment inducing mispricing when short selling constraints exist. Within this context, it is reasonable to develop and test a factor model that includes a factor to capture mispricing.

Ideally, the closing price would perfectly reflect all information revealed by prices earlier in the day. This is particularly important given the prominence of the closing price in financial research. The findings of this paper indicate that movements in prices observed earlier in the day seem to contain valuable information; the movement in these prices relative to the close is a strong predictor of stock returns; and this relationship appears to be driven by market frictions delaying the incorporation of information. The remainder of this paper is organized as follows. In Section **??**, I describe the data and methodology used to calculate *HLDiff*. In Section 3, I present both raw and risk-adjusted returns along with Fama-MacBeth regression results and an investigation of what drives *HLDiff*. In Section 4, I construct my factor and compare the *HLDiff* factor model with other factor models. Section 5 presents a variety of robustness tests. Section 6 concludes.

2. DATA

Using all NYSE, NASDAQ, and AMEX firms available from the CRSP daily stock file from 1965-2016, I obtain daily data including the high, low, and close price, as well as the bid and ask at the close. In addition to the price information, I also obtain daily returns as well as all relevant dividend information including stock splits and dividends. For the book-to-market calculations, I have obtained the book value of equity for the preceding year end from the Compustat database. That is to say, the book value of equity for the fiscal year ending 2015 is used for all book to market calculations in 2016. Market equity for these periods is based on the market equity for the firm from the prior month.

My main variable of interest is based on calculating the return between the midpoint between the high and low price from one day to the next and comparing the midpoint return to the closing price return. As an initial step, I recalculate all CRSP returns to within 0.01% using quoted closing prices. This ensures that my return calculations are accurate and fairly approximate the returns presented by the CRSP dataset. To obtain the return to the high-low midpoint, the close price in the above calculation is replaced with the midpoint between the high and low prices for a given day. In general, the high-low return is obtained through the following calculations.

$$HLPrc_{i,t} = \frac{High_{i,t} + Low_{i,t}}{2}, \tag{2.1}$$

$$HLRet_{i,t} = \ln\left(\frac{HLPrc_t + div_t}{HLPrc_{t-1}}\right), \tag{2.2}$$

In the above equations, $\mathrm{High}_{\mathrm{i,t}}$ represents the high price for firm i from day t and $\mathrm{Low}_{\mathrm{i,t}}$ represents the low price for firm i from day t. $\mathrm{HLPrc}_{\mathrm{i,t}}$ is the midpoint of the high and low prices for firm i on day t. $\mathrm{HLRet}_{\mathrm{i,t}}$ is the return calculated using the midpoint between the high and low prices for firm *i* on day *t* and day *t-1* including any dividends, stock splits, etc. Using this return, I compare the $\mathrm{HLRet_{i,t}}$ with the corresponding return to the CRSP calculated closing price return.

$$RDiff_{i,t} = HLRet_{i,t} - Ret_{i,t}, \tag{2.3}$$

In the above equation, $\mathrm{Ret_{i,t}}$ represents the natural log of either the closing price return for firm *i* on day *t*. **Figure A.1** provides a simple example of the return difference from a single day. As a robustness test, I have also calculated closing price returns using the the bid-ask midpoint at the close. This is to address the issue of any upward bias that may arise as a result of using daily closing prices to calculate returns as demonstrated by Blume and Stambaugh (1983). Using the bid-ask midpoint at the close helps to diminish this bias. As a further safe guard against market microstructure noise, I accumulate the daily return differences over all the trading days in a given month. Non-trading issues and bid-ask bounce will largely be mitigated by adopting this convention and *HLDiff* will be approximately zero. In the formation of my portfolios, firms with value of zero for *HLDiff* are excluded from my extreme portfolios. The exclusion of securities below $1 also assists in mitigating market microstructure noise issues.

$$HLDiff_{i,m} = \sum_{t=0}^{D} RDiff_{i,t}, \tag{2.4}$$

HLDiff represents the accumulated difference between the return calculated using the midpoint between the high and low prices on a given day and the closing price return. **Figure A.2** presents a simple example of the calculation of *HLDiff* over a 5-day window for the security, MYOK, from 12/23/2015 through 12/31/2015.

HLDiff is a reflection of a consistent movement in the midpoint between the high and the low that is not matched by the close. Since the high and low have previously been identified by researchers as informative, a natural question to ask is what can the movements tell me about future returns. Corwin and Schultz (2012) and Abdi and Ranaldo (2017) demonstrate that using the high and low prices on a given day can effectively estimate the bid-ask spread for a given security. The bid-ask spread is an important metric to understand when considering the cost of trading a particular security. A change in this spread may be reflective of an increase or decrease in the cost of trading that security. Accordingly, changes in the high and low prices on a given day would indicate a fundamental change in the cost of trading a security that in a perfectly efficient market should be accounted for in the closing price. Another body of literature uses the high and low prices on a given day in estimating volatility of a security. Beckers (1983) and Yang and Zhang (2000) are two such papers. I leave the investigation of the return relationship between transaction costs, bid-ask spread, and volatility proxies to the papers already mentioned. Ultimately, the empirical results using the high and low prices provide a logical starting point for selecting informative prices. An alternative approach would be to select a randomly realized price from earlier in the day, however, there is no guarantee that the selected price would be informative. Based on the existing literature, movements in the high-low midpoint should contain some information about either a change in bid-ask spread or volatility. The comparison to the close then is attempting to assess the extent to which the closing price either includes or excludes the information reflected in the movements of the high and low prices.

Table 1 contains the summary statistics for *HLDiff* along with all control variables included in the Fama-MacBeth regressions. I have included the mean, standard deviation, median, 25th, and 75th percentiles for each variable. Control variables include prior month returns and prior 6 months return as measured from t-7 through t-2. Prior month returns are commonly associated with short term reversal while the 6-month return is intended to capture momentum returns. I have also included controls for average daily turnover (lmto) as measured as total monthly volume divided by shares outstanding, daily return volatility, log size, log book-to-market, and spread as measured according to Abdi and Ranaldo (2017) with a two-day adjustment. Spread is included as a control to verify that the results are not simply a reflection of changes in transaction costs alone. A final control variable (rv) is constructed to capture the price range for the security. This variable is calculated as:

$$rv_{i,m} = \sum_{t=1}^{T} \frac{High_{i,t} - Low_{i,t}}{(High_{i,t} + Low_{i,t})/2} \tag{2.5}$$

where high and low represent the high and low prices on day t. These control variables have been included in the regressions that are discussed in detail in the following section.

3. RESULTS

In this section, I examine the relationship between *HLDiff* and future returns. Section 3.1 evaluates both the raw and risk-adjusted returns to decile portfolios formed based on *HLDiff*. In Section 3.2, I run Fama-MacBeth regressions at a firm level including a variety of different control variables commonly associated with return predictability. Finally, Section 3.3 presents results considering returns on a weekly basis.

3.1 Portfolio Tests

As a first step, I form decile portfolios based on the value of *HLDiff* at t-1. I present both equal and value-weighted raw returns and five-factor alphas. I limit my discussion here to the value-weighted results. All equal-weighted results are of the same sign and significance as those associated with the value-weighted returns and are larger in magnitude. I also present return results for three sub-periods as well as portfolios that control for market capitalization. Controlling for market capitalization is achieved by first sorting firms into three separate groups based on their market capitalization at the end of the preceding month. Once all securities are assigned to a size group (small, medium, or large), I form decile portfolios based on *HLDiff* at t-1. These portfolios are held for a single month and reformed at the beginning of the next month. The value-weighting and capitalization sub-groups ensure that the results are present across a broad cross-section of securities and not simply driven by small firms, while the sub-period results ensure that the findings are applicable across time and not simply isolated to a particular time frame.

Once each portfolio is formed, I obtain the time series of value-weighted returns to each portfolio for each month from January 1965 through December 2016. If movements in the high and low are informative and the close fails to incorporate this information, I would expect *HLDiff* to be positively and significantly associated with the subsequent monthly returns. The raw returns to each decile portfolio along with an arbitrage portfolio are presented in Table B.2. The Fama and French (2015) five-factor alphas are reported on Table B.3. The risk-adjusted alphas obtained from the five-factor model ensure that the results are not simply a reflection of the size, book-to-market, investment, or profitability factors. Controlling for common factor models helps ensure that the premium is unlikely to be associated with previously identified risks but also facilitates comparison with other anomalies. All t-statistics obtained from the factor regressions are heteroscedasticity-consistent following White (1980).

Table B.2 presents the value-weighted raw returns for each decile portfolio formed based on *HLDiff*. In addition to each portfolio return, a spread portfolio has been formed by subtracting the return to the decile 1 portfolio from the return to the decile 10 portfolio. A t-test has been performed as a way to assess if the return to this spread portfolio is significantly different from zero. Across sub-periods and size sub-groups, I find that the returns to these spread portfolios are positive and significant. This is consistent with the fact that movements in the high-low midpoint are informative and can be used to predict future returns when the closing price movements are not reflective of the movements in the high-low midpoint. The full sample spread portfolio has an average monthly value-weighted return of 1.52% with a t-statistic of 9.69. This return differential is fairly substantial and represents an annualized return of 19.84%. When the high and low drop lower relative to the close price (decile 1), the average value-weighted monthly returns are âĂŞ0.36%. However, when the high and low increase more relative to the close price (decile 10), the average value-weighted monthly returns are 1.15%. Table B.3 presents the value-weighted five-factor portfolio alphas. The decile 1 and decile 10 portfolios have average monthly alphas of âĂŞ0.74% and 0.72%, respectively. The spread portfolioâĂŹs alpha is positive and significant at 1.46% with a t-statistic of 7.78. After adjusting for the common risk factors, the returns remain virtually unchanged. The five-factor alpha is 1.46%, while the raw return is just 0.06% higher at 1.52%.

Both raw return and risk-adjusted alphas are consistent across size sub-groups as well. Table B.2 presents the raw returns for the size groups. Table B.3 presents the alphas for the size groups. The *HLDiff* effect is strongest for the small size group. The spread portfolio for this size group has an average monthly raw return of 2.77% with a t-statistic of 19.62 and an average monthly risk-adjusted alpha of 2.73% with a t-statistic of 17.77. The spread portfolio for the medium size group has an average monthly raw return of 1.78% with a t-statistic of 14.08, while the average monthly risk-adjusted alpha is 1.77% with a t-statistic of 12.73. For the large size group, the average monthly raw return is 1.14% with a five-factor alpha of 1.04% and t-statistics of 8.26 and 6.77, respectively. While the result is strongest for the small sub-group, the result remains economically and statistically significant even for the largest size group both in terms of raw and risk-adjusted returns. As before, returns for sub-groups are largely unchanged after adjusting for common risk factors. The largest decline in returns is within the large size group where the raw returns are reduced by 0.1%.

Table B.2 and Table B.3 both present the results across three sub-periods (1965-1982, 1983-2000, and 2001-2016). For the first sub-period, the average value-weighted monthly raw return is 0.77% with a risk-adjusted alpha of 0.81% and t-statistics of 4.38 and 4.16, respectively. Monthly value-weighted raw returns average 2.24% across the 1983-2000 sub-period with a t-statistic of 7.75 while the average monthly risk-adjusted alpha is 2.15% with a t-statistic of 5.07. The final sub-period from 2001-2016 presents an average raw return of 1.55% with a t-statistic of 4.69, while the risk-adjusted alpha is 1.44% with an associated t-statistic of 4.03. The effect is strongest in the 1983-2000 sub-period both in terms of raw and risk-adjusted returns. However, as with the size groups, the results remain both economically and statistically significant for all sub-periods. As all size groups and sub-periods present much the same way, it would seem that these results are robust and not driven by a particular subset of securities where transaction costs may make trading on this information prohibitive or isolated to a particular time period.

As further verification of my result, I also regress the spread portfolios on a variety of alternative factor models and obtain similar results. These factor models include the Sharpe-Lintner CAPM, the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model, and the Stambaugh and Yuan (2016) mispricing factor model (M-4). In the robustness section, I also completed a principal component analysis based on research by Giglio and Xiu (2018) to further verify that the risk-adjusted results are not simply the result of an omitted factor. I have presented results for the five-factor alphas for the sake of brevity and for comparability with other anomalies. I also find that these results persist after increasing the dollar cut off for stocks to $5 or $10.

All of these results reiterate that the movements in the high and low that are not reflected in the closing price are informative about future returns while common risk factors are not able to explain the excess returns realized by these securities. This result is consistent with prior research that has shown the high and low to be informative (Abdi and Ranaldo (2017), Corwin and Schultz (2012), Yang and Zhang (2000), and Beckers (1983)). If the close poorly reflects the information in other or earlier prices, it is reasonable to expect that the market will eventually adjust the price to reflect that information.

3.1.1 Comparing *HLDiff* to Other Return Determinants

This section presents results from Fama-MacBeth cross-sectional regressions of monthly stock returns on *HLDiff* and a variety of firm characteristics. These tests are run as a means to verify that the findings presented thus far are not simply a previously identified determinant of stock returns. The following regression equation is used for the purpose of this test.

$$ret_{i,t} = \beta_0 + \beta_1 HLDiff_{i,t-1} + \beta_2 Controls_{i,t-1} + \epsilon_{i,t}, \tag{3.1}$$

In this equation, $\mathrm{ret}_{\mathrm{i,t}}$ is the log stock return to stock i in month t. The control variables included in this regression are returns from month t-1, buy and hold returns from month t-7 to t-2, a proxy for the bid-ask spread proposed by Abdi and Ranaldo (2017), monthly turnover, daily return volatility, log market capitalization, and the log book-to-market ratio at the end of the preceding month. As before, I have conducted the analysis on a subset of firms based on their market capitalization and separated the full sample into three sub-periods as a further robustness check of my result. All standard error estimates have been Newey-West adjusted for 4 lags as suggested by Greene (2012).

Table B.4 presents the results for both CRSP closing price returns. In addition to the full sample, the results from running the Fama-MacBeth regressions on two sub-periods are reported to verify that these results are not isolated to a particular period. *HLDiff* remains a positive and significant determinant of returns in the presence of the other control variables. In terms of significance over the full sample period, *HLDiff* is the most statistically significant variable included in the regression. The coefficient for *HLDiff* for the full sample period is 0.2115 with a t-statistic of 23.95, which confirms the positive and statistically significant relationship between *HLDiff* and future stock returns. This result remains consistent across sub-periods. As with the raw and risk-adjusted returns presented, *HLDiff* is stronger in the middle sub-period. For the 1965-1982 sub-period, the coefficient on *HLDiff* is 0.1938 with a t-statistic of 14.72. For the middle time period, the coefficient increases to 0.2570 with a t-statistic of 25.03. In the final time period, the coefficient on *HLDiff* decreases to 0.1749 with a t-stat of 9.23. Irrespective of the time frame, the coefficient on *HLDiff* is positive and remains statistically significant. The coefficients of other control variables are consistent with expectations or insignificant.

Table B.5 presents the results from the Fama-MacBeth regressions for three size sub-groups. Again, these results present in a way consistent with those from the portfolio return analysis. *HLDiff* is largest for small firms with a coefficient of 0.2469 and t-statistic of 23.42. This coefficient decreases to 0.1661 with a t-statistic of 11.61 for the largest tercile of stocks. The results for the medium size tercile are positive and significant with a coefficient of 0.1741 and a t-statistic of 14.17. The magnitude and significance of the *HLDiff* coefficient of the medium size sub-group is similar to that of the large size sub-group. Irrespective of the size group, the *HLDiff* effect remains in the presence of each of the selected controls. The persistence of *HLDiff* is in contrast to some other common determinants of stock returns such as book-to-market or momentum which are both insignificant for the largest subset of firms while *HLDiff* presents the largest t-statistic across size groups. Irrespective of size group, it appears that the close may move in a way that does not reflect the movements in prices earlier in the day. It is possible that the closing price for small firms could be influenced by the increased trading volume at the close. However, these results confirm that the informativeness of *HLDiff* is not limited to small firms where trading around the close may have a more dramatic price impact. In unreported tests, I subdivide firms based on their stock price or monthly turnover. Across the stock price groups, the coefficient is positive and significant with the highest priced stock producing the smallest coefficient in terms of magnitude. The coefficient on *HLDiff* for the lowest priced stocks is 0.1593 with a t-statistic of 11.89. Similarly, all three turnover groups have a positive and significant coefficient on *HLDiff* with the highest turnover group producing the lowest coefficient of 0.1652 and a t-statistic of 12.73.[1].

One possibility is that *HLDiff* is consistent for a given security across time. If this is the case, I expect that *HLDiff* would be informative at horizons beyond one month. Additionally, this would indicate that either the market consistently fails to react to the information in the high and low prices, or that the returns realized at the close are partly driven by security characteristics, and *HLDiff* may in some way capture this non-transient security characteristic. To test this, I use returns at month

[1]To ensure that the result is not driven by firms with zero turnover and thus the least liquid firms, I omit firms in the low turnover group with a turnover value of 0. This also avoids multicollinearity issues that would arise from measuring *HLDiff* where no trading took place.

t+2, t+4, and t+6 as the dependent variables in the Fama-MacBeth regressions. For each of these alternative dependent variables, the coefficient on *HLDiff* is insignificant with the highest t-statistics of 1.59 for the returns 4 months ahead. This finding is consistent with the idea that the market incorporates the information contained in the high and low prices over the following month, and the difference between the return to the high-low midpoint and closing price return is not necessarily a reflection of a persistent security characteristic. This is further confirmed when regressing *HLDiff* on a lagged observation of *HLDiff*. The coefficient on the lagged *HLDiff* variable is negative and significant. That is to say that *HLDiff* reverses in the subsequent month. This is not entirely unexpected as *HLDiff* is higher when the closing price decreases, however, the results thus far indicate that the closing price will increase in the subsequent period, thus *HLDiff* is lower because the closing price relative to other prices is higher.

To this point, the results have shown that movements in the high and low midpoint are informative about future returns. However, it is possible that movements in the high or low alone are capable of explaining the predictive power of *HLDiff*. As a further test, I decompose *HLDiff* into movements in the high or low by recalculating *HLDiff* using only the high or low price. These results are presented in Table B.6. For *HLDiff* calculated using only the low price, I find similar positive and significant results, however, the magnitude of the coefficient is approximately half the size of the combined high and low price coefficient at 0.1062 with a t-statistic of 20.72.. *HLDiff* derived using only the high price is also positive and significant, but as with the low price decomposition, the magnitude of the coefficient is approximately half that of the coefficient on the combined high-low midpoint. The coefficient for *HLDiff* in this test is 0.1201 with a t-statistic of 18.28. Individually, movements in the high and low not reflected in the close are positive and significantly related to future returns. However, the midpoint, in terms of magnitude and significance, is larger than either price individually. To the extent that the high or low represent unique information, this result is not surprising and the combination of the two prices to obtain the midpoint provides a way of possibly capturing the information contained in each respective price. In addition, Corwin and Schultz (2012) find that the trades at the high price are almost always buyer-initiated, and the low price is almost always associated with a seller-initiated trade. Using both prices in some way then captures the movement of buyers and sellers while the high or low may only effectively capture the buyers or sellers, respectively.

I have also conducted a variety of robustness tests in addition to the results already presented. These results persist after controlling for different momentum strategies. In addition, the result is not sensitive to the use of log returns as the dependent variable. Log returns were originally used as the dependent variable as *HLDiff* is calculated using the log daily returns. However, the results are present for both sets of returns.

Ultimately, these cross-sectional regressions demonstrate that *HLDiff* persists in the presence of a variety of control variables, and the result presented thus far is not simply a manifestation of another previously identified return determinant. In all sub-periods and size groups, *HLDiff* remains a statistically significant and positive predictor of stock returns. The t-statistic associated with *HLDiff* is often the largest of all presented variables. This is particularly remarkable given that many documented predictors of future stock returns lose much of their predictive ability for the largest subgroup of securities. These results are consistent with others who find that prior price observations can be informative about future returns. Jegadeesh and Titman (1993) and George and Hwang (2004) both use previously realized prices or returns to predict future returns. This paper uses the closing price as a benchmark for comparison and finds that earlier trades are informative and may not be perfectly reflected in the close. Since the close does not reflect the information in these trades, it is interesting to consider if *HLDiff* can potentially explain other observed anomalies as it may capture information revealed to the market during the day but not fully reflected in the price at the close. In Section 5, I construct a factor based on *HLDiff* and evaluate the alphas associated with 12 anomalies after the inclusion of the *HLDiff* factor.

3.2 Weekly Testing

To the extent that movements in the high and low are informative, it seems reasonable for the market to react relatively quickly to this information. This is particularly true if the close fails to reflect the movements in the high and low prices. Hou and Moskowitz (2005) note that prices respond relatively quickly to information. In particular, information revealed to the market will be fully incorporated within a month's time. In addition, price changes that are meaningful are unlikely to reverse in subsequent periods as the shift reflects the efficient price for the security. To this point, the results are indicative of the fact that movements in the high and low that are not matched by the close do seem to provide meaningful information about future returns. The following tests attempt to identify how quickly the price reacts to *HLDiff* and if there is any subsequent reversal of the return. I form buy and hold portfolios and record returns up to 12 weeks in the future. I also run weekly Fama-MacBeth regressions on weekly returns up to 12 weeks in the future. These results are strongly indicative of the fact that movements in the high and low that are not captured by the close are largely incorporated into the price over the following week. However, *HLDiff* remains a positive predictor of returns up to 6 weeks in the future with declining magnitude for each week further in the future.

3.2.1 Weekly Portfolio Tests

As an initial consideration of the timing of the returns realization, I form weekly decile portfolios based on *HLDiff*. *HLDiff* is calculated based on the preceding 20 trading days, approximately one month, and weekly returns are compounded daily returns from Wednesday to the following Wednesday. The mid-week convention is adopted as it largely avoids issues with market holidays and any influence the week-end may have. However, results using returns calculated from other weekdays are qualitatively similar. The results of the weekly portfolio tests are reported in Table B.7. How the sake of brevity, I have presented the value-weighted raw return and five-factor alphas for the spread portfolio (portfolio 10 less portfolio 1) only. Portfolios are formed and held for up to 12 weeks. Weekly returns are positive and significant across all 12 weeks that the portfolios are held. The largest portion of the returns are realized in the first week after portfolio formation.

Within the first week of portfolio formation, the portfolio earns 0.65% with an associated t-stat of 13.53. The five-factor alpha is 0.66% with a t-stat of 13.60. This result is compelling in terms of statistical significance, and also economically with an annualized return of approximately 36%. While trading on a weekly basis may be prohibitive, the returns persist at horizons beyond one week. In the second week, the raw return increases slightly to 0.77%. The five-factor alpha increases to 0.78% in the second week. The raw return and alpha are positive and significant for every holding period up to 12 weeks. The results are largely consistent with the findings presented earlier with *HLDiff* positively predicting future returns. Beyond the confirmation of the earlier findings, these results also point to the fact that the market responds fairly quickly to the information content of the high and low prices. It is interesting to note that over the full holding period considered, the returns recognized in the first week are not realized. This would be consistent with the intuition that the movements in high and low prices reflect meaningful information that is impounded into the price. If the information was spurious or if *HLDiff* is not associated with any meaningful information, I would expect the returns to be corrected at longer horizons as the market adjusts its reaction to more appropriately reflect the true price and reduce the noise included in the price.

In unreported results, my findings are similar when I subdivide across size groups. The magnitude and significance is largest in the initial week following portfolio formation. The raw return and alpha remain positive and significant over the full sample period for all size groups. The magnitude and significance are largest for the small size sub-group and are smallest for the large size sub-group. These outcomes are also present across sub-periods. These results point to similar conclusions as those presented in the monthly portfolio testing. One issue that may arise is concern about market microstructure noise associated with compounding daily returns combined with weekly return measures. In the robustness section, I have recalculated returns using the bid-ask midpoint at the close to verify that the predictive power of *HLDiff* is not simply an artifact of market microstructure noise. The results associated with weekly portfolios are qualitatively similar after making both of these robustness adjustments.

3.2.2 Weekly Fama-MacBeth

As a further consideration of the timing of returns, I have run Fama-MacBeth regressions on weekly returns. Table B.8 reports the results for the Fama-MacBeth regressions along with control variables. *HLDiff* is positively related to the subsequent weeks returns. Consistent with previously presented results, the coefficient on *HLDiff* is the largest in terms of statistical significance and remains economically significant. In particular, a one interquartile increase in the value of *HLDiff* is associated with a 0.22% increase in weekly returns. When annualized, this represents a 14% return. If the movements in the high and low represent meaningful information that should be impounded in the price but is not reflected in the close, I would expect the market to respond quickly with the informativeness of *HLDiff* dropping in power as the market adjusts the price to fully incorporate the information. If *HLDiff* captures some sort of underlying risk, I would anticipate that the magnitude and significance will persist irrespective of when *HLDiff* is measured. To examine how quickly the price adjusts related to *HLDiff*, I run the following Fama-MacBeth regressions:

$$ret_{i,w+n} = \beta_0 + \beta_1 HLDiff_{i,t-1} + \beta_2 Controls_{i,t-1} + \epsilon_{i,w+n}, \tag{3.2}$$

where the dependent variable is the weekly returns *w+n* periods ahead. *HLDiff* is based on the preceding 20 trading days. Controls are the same as those in the monthly Fama-MacBeth regressions. The dependent variable is 1, 2, 4, 6, 8, 10, and 12 weekly returns in the future. On Table B.9, I report the coefficient associated with *HLDiff* for each Fama-MacBeth regression as well as the associated t-statistic. The findings are largely consistent with the weekly portfolio findings. *HLDiff* is strongest both in terms of magnitude and significance one week in the future with a coefficient of 0.1749 and t-statistic of 37.94. In the following week(*w+2*), the coefficient drops by 92.6% to 0.0136 with a similar decline in significance. However, *HLDiff* remains statistically significant, but is no longer economically significant. That is to say that the price quickly adjusts to incorporate the information that is reflected in the high and low prices. The decline continues up to 10 weeks with a small uptick at 12 weeks in the future.

While there is a general decline in the predictive power of *HLDiff*, the relationship between *HLDiff* and returns remains positive over the horizon considered. Similar to the portfolio findings, this indicates that the price adjusts relatively quickly. While the price responds relatively quickly, it is also worth noting that at longer horizons the relationship is never negative. Again, this indicates that the market does not reverse the effect of the original price movement associated with the *HLDiff* variable. The fact that the price responds relatively quickly and does not reverse is consistent with the movements in the high and low presenting meaningful information that is only incorporated after a delay. As with my portfolio results, I also consider both size sub-groups and sub-periods. All of these results are consistent with earlier reported results, with the coefficient remaining positive and significant across size groups and sub-periods. As before, I have also verified that the results hold for market microstructure adjustments. *HLDiff* remains a positive and significant predictor of future returns after shifting to bid-ask midpoint at the close returns or adopting the weighted-least squares regression suggested by Asparouhova, Bessembinder, and Kalcheva (2010).

Both sets of weekly results (portfolio and Fama-MacBeth) are indicative of the fact that *HLDiff* contains some information. The fact that prices respond relatively quickly to the information in *HLDiff* as well as the fact that the price movements are consistently positive and not reversed indicate that the high and low contain meaningful information, and that information is impounded into prices after a delay.

4. HLDIFF FACTOR

Thus far the results have shown a significant relationship between *HLDiff* and future stock returns. In this section, I explore the relationship between *HLDiff*, other anomalies, and mispricing. Stambaugh, Yu, and Yuan (2012) demonstrate that a large set of anomalies observed in empirical literature are, in part, driven by investor sentiment. They argue that high levels of market-wide sentiment will result in stronger overpricing due to short selling constraints consistent with Miller (1977). Consistent with this intuition, Stambaugh, Yu, and Yuan (2012) find that anomaly returns are stronger following periods of high sentiment. If *HLDiff* captures a degree of mispricing, investor sentiment should positively predict the premium earned by the *HLDiff* factor. Further, as many of the anomalies considered are, in part, driven by mispricing, I expect the *HLDiff* factor to explain the alpha for some of the anomalies if mispricing drives the results associated with *HLDiff*. In that spirit, I use *HLDiff* to create a 'factor' mimicking portfolio and perform a battery of tests. First, I establish the relationship between *HLDiff* and the Baker and Wurgler (2006) investor sentiment index. Next, I add the *HLDiff* factor to the Sharpe-Lintner single-factor and Fama-French five-factor models and document anomaly alphas before and after the inclusion. Lastly, I consider the relative ability of each of these models to explain the other factor models.

To construct the *HLDiff* factor, I adopt the approach of Carhart (1997). The factor portfolio return is obtained by subtracting the value-weighted return to the bottom 30 percent of stocks based on *HLDiff* from the value-weighted return to the top 30 percent of stocks based on *HLDiff*. The considered anomalies are generated from the simple closing price returns. Accordingly, I have used the difference between simple (not logged) returns calculated from the high-low midpoint for the factor portfolio formation.

Table B.10 presents the means, standard deviations, and correlations for a monthly series of the *HLDiff* factor with each of the factors in the Fama and French (2015) five-factor model. The *HLDiff* factor is negatively associated with the market and size factors and positively related to the value, profitability, and investment factors. The correlation magnitudes between *HLDiff* and the other five-factors are similar to the correlations amongst other factors.

4.1 ***HLDiff*** and Investor Sentiment

As discussed above, Stambaugh, Yu, and Yuan (2012) connect mispricing to investor sentiment. They extend this logic into the formation of their mispricing factor model in Stambaugh and Yuan (2016). I adopt a similar approach to the latter paper by connecting the *HLDiff* factor to sentiment by dividing the factor portfolio into long and short leg components and regressing the excess returns to these portfolios on lagged investor sentiment. Following a period of high sentiment, the most overpriced stocks will fall into the short leg portfolio, and thus the returns to this portfolio will be related the prior period's level of investor sentiment. Motivated by Miller (1977), short selling constraints may result in rational investors being unable to correct perceived overpricing. On the other hand, the long leg of anomaly arbitrage portfolios will be significantly less sensitive to investor sentiment as there is no constraint on the ability of investors to long a security. Underpricing will quickly be corrected as no market friction inhibits the ability of investors to respond to the observed underpricing. If the *HLDiff* factor captures some degree of mispricing, I would anticipate the short leg of the factor mimicking portfolio to be significantly correlated with investor sentiment. However, I anticipate that the long leg will not present any meaningful relationship between the excess return and investor sentiment. Finally, the spread portfolio should be positively and significantly related to the lagged investor sentiment measure.

To test the relationship between *HLDiff* and sentiment, I adopt the approach of Stambaugh and Yuan (2016) and run the following regression:

$$R_{i,t} = a_i + b \times S_{t-1} + u_{i,t}, \tag{4.1}$$

where $R_{i,t}$ represents the excess return for the short or long leg portfolio or the factor portfolio premium. S_{t-1} represents Baker and Wurgler (2006) investor sentiment from the prior month. Table B.11 reports the coefficient estimates for the above regression. Consistent with mispricing and the findings of Stambaugh, Yu, and Yuan (2012), I find that the coefficient on investor sentiment is negative and significant for the short leg portfolio and insignificant for the long leg portfolio. Further, the magnitude of the coefficient on the short leg portfolio is over two times as large as the coefficient on the long leg portfolio. The spread portfolio (long-short) is positive and significant. As discussed above, short selling constraints will limit the ability of investors to correct overpricing. The short leg is comprised of the firms that are most overpriced firms, or where the closing return has consistently exceeded the high-low midpoint return. The relationship between the short leg portfolio and investor sentiment is indicative of mispricing. These results are consistent with the findings of Stambaugh and Yuan (2016) who also find qualitatively similar loadings on their long and short factor portfolios and a positive relationship between their mispricing factors and prior sentiment. These results are reflective of the fact that the premium on the *HLDiff* factor is earned primarily by the most overpriced stocks. This asymmetry in the effect of sentiment on the excess returns for the long and short legs of the *HLDiff* factor is indicative of *HLDiff* capturing a degree of mispricing which, in this case, would be related to the inability of pessimistic investors to step in and correct the overpricing induced by the high sentiment of optimistic long investors.

4.1.1 Examining the HLDiff Factor's ability to explain anomalies

Table B.12 reports the alphas for the single-factor (Sharpe (1964) and Lintner (1965)) CAPM and five-factor (Fama and French (2015)) model before and after the inclusion of the *HLDiff* factor for various anomaly portfolios. The anomalies considered are: distress (Campbell, Hilscher, and Szilagyi (2008)), O-score (Ohlson (1980)), net stock issues (Ritter (1991) and Loughran and Ritter (1995)), composite equity issues (Daniel and Titman (2006)), total accruals (Sloan (1996)), net operating assets (Hirshleifer, Hou, Teoh, and Zhang (2004)), momentum (Jegadeesh and Titman (1993)), gross profitability (Novy-Marx (2013)), asset growth (Cooper, Gulen, and Schill (2008)), return on assets (Fama and French (2016)), Chen, Novy-Marx, and Zhang (2011)), investment to assets (Titman, Wei, and Xie (2004)), and idiosyncratic volatility (Ang, Hodrick, Xing, and Zhang (2006)). For each anomaly, spread portfolios have been formed based on the value-weighted return difference between the top and bottom deciles. All anomaly returns are from Robert Stambaugh's website[1]. These spread portfolio returns are then used as the dependent variable in the following regressions:

$$R_{i,t} = \alpha_i + \sum_{j=1}^{N} \beta_{i,j} \times F_{j,t} + u_{i,t}, \tag{4.2}$$

where $F_{j,t}$ represents the N factors in a given model. For instance, the five-factor model before the inclusion of the *HLDiff* factor would include the market, size, value, investment, and profitability premia while the five-factor model after inclusion would include market, size, value, investment, profitability, and the *HLDiff* factor. All t-statistics are heteroskedasticity-consistent following White (1980).

Of the anomalies considered, eleven have significant alphas in the single factor CAPM over the time period from 1965 to 2016. This is largely consistent with the fact that common risk factors are insufficient to explain the returns associated with these anomalies. In terms of absolute magnitude, the single-factor alphas range from

[1] `http://finance.wharton.upenn.edu/~stambaug/`

0.32% for the O-score anomaly to 1.56% for idiosyncratic volatility. The t-statistics range from 1.28 for the O-Score anomaly and 6.41 for the net stock issues anomaly. The O-Score anomaly is the only one that is unable to achieve significance over the time period considered in the single factor model. I also provide the average absolute alpha and t-statistic for all 12 anomalies as well as the the Gibbons, Ross, and Shanken (1989) (GRS Test) F-statistic and p-value. The single-factor average absolute alpha across all 12 anomalies is 0.73% while the t-statistic is 4.30. The F-statistic of the GRS Test is 6.61 with an associated p-value of 4.9772×10^{-11}. Under the five-factor model, the O-Score anomaly is now significant with a t-statistic of 3.73, while the asset growth anomaly is now insignificant. This item becoming insignificant is not surprising as the CMA factor is effectively asset growth. The average alpha for the five-factor model is 0.54% with an average t-statistic of 3.44. The GRS test yields an F-statistic of 5.55 with an associated p-value of 6.0331×10^{-9}. The GRS test for both the single and five-factor models rejects the null hypothesis that these models are adequate to explain the pricing errors of these anomalies.

After adding the *HLDiff* factor to the single-factor model, 10 anomalies are decreased in both magnitude and significance. The inclusion of my factor achieves a slightly better level of success than the five-factor model in explain the return of these anomalies. Under the five-factor model only one of these anomalies is no longer significant. However, with the inclusion of the *HLDiff* factor to the single-factor model, the alpha associated with three anomalies are now insignificant. These anomalies are distress, gross profitability and return on assets. In addition, four anomalies that were previously above the 3.00 t-statistic hurdle proposed by Harvey, Liu, and Zhu (2015) are less than 3.00 after the inclusion of the *HLDiff* factor. The average absolute alpha is now 0.44% with an average t-statistic of 2.77. The GRS test statistic is similarly improved, dropping to 7.69. Perhaps the most stark result is the increase in the p-value associated with the GRS Test. The null hypothesis of this test is that all alphas jointly equal zero for a given set of factors. The null is rejected at the 1% level both excluding and including the *HLDiff* factor. However, the p-value increases by 5 orders of magnitude. Contrasting this with the five-factor model p-value the additional of the *HLDiff* factor improves the single-factor model significantly more than the addition of the additional 4 factors of the five-factor model. Even though the null hypothesis is rejected, the addition of the *HLDiff* factor appears to represent a substantive improvement over the single-factor model alone. The reduction in the average alpha by over 38% further demonstrates the improvement in the explanatory ability of the model after adding *HLDiff*. In contrast, the addition of the size, value, investment, and profitability factors only improve the absolute average alpha by 27%.

Moving to the five-factor model with *HLDiff*, 9 of the 12 anomalies decrease in absolute magnitude with 10 anomalies decreasing in terms of statistical significance. However, only distress is insignificant after the inclusion of the *HLDiff* factor. The average alpha after the inclusion of *HLDiff* is 0.39% or a 27% decrease from the five-factor model excluding the *HLDiff* factor. The average t-statistic also decreases to 2.70. The GRS test statistic is 3.95. In this instance, four of the anomalies no longer exceed the Harvey, Liu, and Zhu (2015) t-statistic benchmark: gross profitability, return on assets, momentum, and idiosyncratic volatility. Again, the improvement in the GRS test statistic indicates that the inclusion of the *HLDiff* factor improves the five-factor model's ability to explain the observed pricing errors for the anomalies considered here. In this instance, the p-value improves by three orders of magnitude from 6.0331×10^{-9} to 7.9363×10^{-6}.

I have also considered the anomaly alphas for each sub-period(1965-1982, 1983-2000, and 2001-2016). In all sub-periods, *HLDiff* similarly improves the explanatory power of both the single and five-factor models. In the latest time period (2001-2016), five of the considered anomalies are insignificant after the addition of *HLDiff* to the single-factor model and the GRS test statistic unable to reject the null hypothesis at the 10% level. All of this again is indicative of the fact that the inclusion of *HLDiff* improves the ability of these models to explain the returns associated with this subset of anomalies.

The inclusion of an *HLDiff* improves the single-factor and five-factor model's ability to explain the returns generated by these 12 anomalies. I have also considered the Carhart (1997) four-factor model and Fama-French three-factor models and find similar results. As these anomalies are used by Stambaugh and Yuan (2016), I have not included the M-4 in the results here. However, the inclusion of *HLDiff* in this model improves the model's pricing ability. In fact, over the period considered, idiosyncratic volatility presents a statistically significant alpha under the M-4 model, but the alpha is insignificant after the inclusion of the *HLDiff* factor. I have provided a more direct comparison of the *HLDiff* model and the M-4 model relying on anomaly returns not used in the creation of either factor model in the next section.

4.2 Comparing *HLDiff* with Other Factor Models

One interesting question to consider is: To what extent can *HLDiff* price the factors in other models or vice versa? To this end, I have included the M-4 model of Stambaugh and Yuan (2016) as a benchmark for comparison along with the five-factor model as a way to distinguish *HLDiff* from existing factor models. I present the alphas for the following regression in Table B.13:

$$FP_{i,t} = \alpha_i + \sum_{j=1}^{N} \beta_{i,j} \times F_{j,t} + u_{i,t}, \tag{4.3}$$

where $\mathrm{FP_{i,t}}$ is the factor premium for a given factor and $\mathrm{F_{j,t}}$ are the factors for the competing model. By way of an example, if HML is the dependent variable and the competing model is the mispricing model (M-4), the independent variables would be the market risk premium, size[2], management, and performance mispricing factors. All models include the market risk premium as an independent variable. Panel B of Table B.13 reports the GRS test statistic and corresponding p-values. In this

[2]Stambaugh and Yuan (2016) calculates a unique size factor which is used as appropriate in these regressions.

instance, the GRS statistic is testing if the alphas for the alternative models' factors are jointly equal to zero.

Results across the various considered models are reported in Table B.13 for completeness. However, I limit my discussion here to those results related to *HLDiff*. For the five-factor model, SMB, HML, RMW, and CMA have mixed significance with HML, SMB and CMA alphas being positive and significant. On the other hand, RMW is insignificant with t-statistics of 1.35. Interestingly, this contrasts with the size factor of the M-4 model which was specifically constructed to avoid including mispriced stocks. The *HLDiff* both in terms of significance and magnitude the alpha for the M-4 size factor is larger than the Fama-French size factor. To the extent that *HLDiff* is consistent with mispricing, this could be interpreted as the improvement in pricing the Fama-French size factor by *HLDiff* stems from the fact that mispriced stocks are included in the factor construction. As for the M-4 model, none of the mispricing factors are priced by the *HLDiff* factor. *HLDiff* appears to do a better job of explaining the performance factor which includes the anomalies that the *HLDiff* model was able price in the Table B.12. The bottom of Panel A presents the alpha earned by the *HLDiff* factor in the five-factor and M-4 models. The alpha on *HLDiff* remains highly statistically significant under both models. The results in Panel B confirm these results. The *HLDiff* model is unable to price the factors for neither of the competing models. Similarly, both the five-factor and M-4 models fall short of pricing the *HLDiff* factor with GRS test statistics of 54.23 (p-value 5.8298×10^{-13}) and 30.73 (p-value 4.4074×10^{-8}), respectively. Each of these results further demonstrate the fact that the *HLDiff* factor is separate and distinct from the factors of the five-factor model and M-4 model. In addition, many of these results are also indicative of the relationship between *HLDiff* and mispricing.

It is difficult to directly compare the M-4 model with the two-factor model including *HLDiff* by looking at the anomalies used in the construction of the factors of either model. This is due to the fact that one model would have a relative advantage in explaining the returns associated with the anomaly or anomalies used in factor construction. As an additional test, I have compared the M-4 model and two-factor model using seven additional anomalies not used in the creation of either factor. These anomalies include idiosyncratic volatility(Ang, Hodrick, Xing, and Zhang (2006)), short term reversal (Jegadeesh (1990)), illiquidity (Amihud (2002)), book-to-market (Fama and French (1993)), long-term reversal (Bondt and Thaler (1985)), and price convexity (Gulen and Woeppel (2018)). Given these anomalies, the M-4 model is unable to reject the null of the GRS test at the 1% level (p-value: 0.0016). However, the two-factor model including *HLDiff* and the market risk premium is only able to reject the null at a 5% level (p-value: 0.019). This provides some evidence that *HLDiff* may be more effective at pricing errors for anomalies than the M-4 model when considering anomalies beyond those used in the construction of the respective factors.

The *HLDiff* factor is not priced by either the five-factor or the M-4 model. However, the *HLDiff* model appears to price the performance mispricing factor. All of these results indicate that *HLDiff* may in some way capture a degree of mispricing. This is further supported by the ability of the *HLDiff* model to explain a variety of anomalies which are, at least in part, driven by mispricing as shown by Stambaugh and Yuan (2016).

5. ROBUSTNESS

I conduct several checks on the robustness of the results presented in the preceding sections. The first tests help verify that the return and risk-adjusted returns associated with *HLDiff* are not simply momentum or idiosyncratic volatility calculated in another way. I have also implemented a test to verify that *HLDiff* factor is not an artifact of omitted factors or measurement error. This testing is consistent with the methodology suggested by Giglio and Xiu (2018). Finally, I adopt two separate approaches to address any market microstructure noise issues. One adjustment is to replace the closing price returns with returns calculated from the bid-ask midpoint at the close. Using the adjusted returns, I reform portfolios and obtain updated results for the Fama-MacBeth regressions. The second procedure relies on a weighted-least squares regression approach suggested by Asparouhova, Bessembinder, and Kalcheva (2010). I utilize the weighted-least squares method and re-run my Fama-MacBeth regressions. For all robustness tests, *HLDiff* remains a positive and significant predictor of subsequent period returns.

5.1 Double-Sorted Portfolios

To further distinguish the raw return and risk-adjusted return results associated with *HLDiff* from other prominent anomalies, I present double-sorted portfolio results for *Idiosyncratic Volatility* and *HLDiff* and *Momentum* and *HLDiff*.

5.1.1 Idiosyncratic Volatility

Table B.14 presents the raw returns and five-factor alphas for portfolios formed based on *Idiosyncratic Volatility* and *HLDiff*. Quintile portfolios are first formed based on *Idiosyncratic Volatility*. *Idiosyncratic Volatility* is calculated as the standard deviation of the daily return residual from a Fama-French Three-Factor model in a given month. This is consistent with Ang, Hodrick, Xing, and Zhang (2006). Each of these portfolios are then sorted into quintiles based on *HLDiff*. Panel A of Table B.14 presents the raw return for the resulting 25 portfolios along with long-short portfolios and their associated t-statistics. Panel B of Table B.14 presents the five-factor alphas for these portfolio along with the long-short portfolio alphas and their associated t-statistics. Irrespective of the *Idiosyncratic Volatility* quintile, the *HLDiff* long-short portfolios have positive and significant return differentials, both in terms of raw returns and risk-adjusted alphas. Stated differently, for a group of firms with similar levels of idiosyncratic volatility variation in *HLDiff* appears to still present both economically and statistically meaningful differences in returns. However, for *Idiosyncratic Volatility* only the first three quintiles present negative and significant raw and risk-adjusted returns. The remaining long-short portfolios are statistically insignificant both in terms of raw and risk-adjusted returns. If the return differentials presented in Section 3 were the result of *Idiosyncratic Volatility*, I would expect the long-short portfolios with similar levels of idiosyncratic volatility would have insignificant return differentials and alphas. However, the return differentials are positive and significant for *HLDiff* for firms with similar levels of idiosyncratic volatility. Therefore, it is unlikely that *HLDiff* is simply an alternative way of calculating idiosyncratic volatility.

5.1.2 Momentum

As with *Idiosyncratic Volatility* above, Table B.15 presents the raw returns and five-factor alphas for portfolios formed based on *Momentum* and *HLDiff*. Quintile portfolios are first formed based on *Momentum* using the 11-1-1 strategy consistent with Jegadeesh and Titman (1993) Each of these portfolios are then sorted into quintiles based on *HLDiff*. Panel A of Table B.15 presents the raw return for the resulting

25 portfolios along with long-short portfolios and their associate t-statistics. Panel B of Table B.15 presents the five-factor alphas for these portfolio along with the long-short portfolio alphas and their associated t-statistics. Irrespective of the *Momentum* quintile, the *HLDiff* long-short portfolios have positive and significant return differentials, both in terms of raw returns and alphas. State differently, for a group of firms with similar magnitudes of prior returns the variation in *HLDiff* appears to still present both economically and statistically meaningful differences in returns. For *Momentum* all portfolios present positive and significant raw and risk-adjusted returns. If the return differentials presented in Section 3 were the result of *Momentum*, I would expect the long-short portfolios with similar levels of momentum would have insignificant return differentials and alphas. However, the return differentials are positive and significant for *HLDiff* for firms with similar levels of momentum. Therefore, it is unlikely that the return differentials are simply an alternative measure of momentum.

5.2 Omitted Factor Testing

With the proliferation of various factors, it is difficult to assess if any particular factor is genuinely informative or if the ability to explain returns is spurious. One potential problem faced by researchers is the omitted variable bias. Giglio and Xiu (2018) note that "...omitted variable bias arises in standard risk premia estimates whenever the model used in the estimation does not fully account for *all* priced sources of risk." To address this problem, Giglio and Xiu (2018) propose a multi-step process. I leave a full explanation of the methodology to their paper. A brief description of the procedure follows: The first stage is to extract the principal components of a large panel of test assets to obtain the factor space. Next, cross-sectional regressions are run using the principal components excluding the factor of interest on the excess return for the test assets to obtain the factor premium. The third step regresses the time-series of the factor of interest on the principal components. The final step is to estimate the risk premium of the factor as the product of the loadings of the factor of interest (obtained in the third step) on the principal components and the associated risk premium estimated in the second step. The researchers note that this methodology guarantees that the estimate of the risk premium will be consistent. I implement this three step procedure based on the code available at Dacheng Xiu's website[1].

This test provides a solution to the problem of model selection by allowing the returns to a large group of test assets to identify the common components driving returns. This result also ensures that the premium associated with *HLDiff* is not simply an artifact of the benchmark model selected for testing. Table B.16 presents the results from the omitted factor testing. For this analysis, four, five, and six principal components have been extracted from the returns related to 202 separate portfolios. In addition to the factor premium estimate for *HLDiff*, the factor premiums for the five-factor model and the Pastor and Stambaugh (2003) aggregate liquidity factor are also reported in B.16. Next to each factor premium estimate, I have also reported a t-stat and a signal-to-noise ratio (R_g^2) also suggested by Giglio and Xiu (2018). The signal-to-noise ratio approximates how well the factor is estimated. A higher value indicates that factor is estimated with relatively little noise.

The factor premium and significance for the factors of the five-factor model and Pastor and Stambaugh (2003) aggregate liquidity are consistent with the findings of Giglio and Xiu (2018). The premium estimate for *HLDiff* factor with four principal components is positive and significant. The magnitude of the *HLDiff* factor premium is within 0.05% of the premiums associated with *SMB*, *HML*, *CMA*, and *RMW*. On the other hand, in terms of significance the *HLDiff* factor is the most significant of all considered factors. *HLDiff* is significant at the 1% level. These results are consistent across both five and six principal components as well. The final measure is the signal-to-noise ratio (R_g^2). *HLDiff's* signal-to-noise measure is 0.1919 which is only better than the Pastor and Stambaugh (2003) aggregate liquidity factor. This indicates that

[1] `http://dachxiu.chicagobooth.edu/`

the measure of *HLDiff* is measured with a degree of error, but not entirely without some amount of meaningful information.

All of the above are indicative of the fact that *HLDiff* appears to have a positive and significant premium, and this premium is not simply an artifact of the model selected or a narrow selection of assets for testing. The *HLDiff* premium also compares favorably in terms of magnitude and significance to all other factors considered in Table B.16.

5.3 Market Microstructure Noise Corrections

One potential concern is the presence of market microstructure noise in the measurement of *HLDiff* and in return calculations. A variety of researchers have documented the impact of microstructure noise such as Blume and Stambaugh (1983) and Asparouhova, Bessembinder, and Kalcheva (2010). In fact, the latter documents the fact that microstructure noise can be present even at a monthly level. Market microstructure noise is often attributed to issues associated with non-trading, bid-ask bounce or non-synchronous trading. It is worth noting, that in constructing *HLDiff* the issue of non-trading is largely mitigated. For a firm that has not traded, the quoted return will be based on the midpoint between the bid and ask for the day which will also be the associated low and high, respectively, for the day. In this instance, the value of the daily difference will be zero. Thus these firms will not be in extreme portfolios or have relatively low values of *HLDiff* relative to other firms. For the issue of bid-ask bounce, *HLDiff* will also be zero. A firm that is bouncing between the bid and ask prices will have daily differences that net to zero over any period where the prices are 'bouncing'. Thus the accumulation of *HLDiff* over the preceding month largely avoids the issue of bid-ask bounce. I also find that the results are robust to controlling for the number of trading days in the prior month as well as a variety of liquidity measures. As a final robustness, I find my results persist for a variety of dollar cut offs as well as eliminating the lowest 5% of firms by market capitalization. However, I have conducted additional tests to verify that the results are not simply a manifestation of market microstructure noise.

The first test reflects both the portfolio and Fama-MacBeth results based on adjusting all returns to be based on the bid-ask midpoint at the close. The use of the bid-ask midpoint is also an attempt to control for market microstructure noise as discussed by Blume and Stambaugh (1983). In particular, Blume and Stambaugh (1983) document that using the quoted closing prices could result in a an upward bias in the calculated returns and basing returns on the bid-ask midpoint at the close helps to mitigate the possibility that the effect is simply the result of any upwardly biased return calculation. To that end, the use of the bid-ask midpoint should effectively control for a degree of microstructure noise while not completely eliminating it as discussed by Fisher, Weaver, and Webb (2010). In addition to the bid-ask midpoint at the close, I also employ a weighted least squares regression as a way to address the microstructure noise problem. This approach is suggested by Asparouhova, Bessembinder, and Kalcheva (2010) and weights according to the prior period return plus 1.

5.3.1 Bid-Ask Midpoint at the Close

Table B.17 Panel A and Panel B report the raw returns and five-factor alphas, respectively, for portfolios formed on the basis of *HLDiff* where all returns are calculated from the bid-ask midpoint at the close. The time period for these tests are restricted to 1993-2016 due to the fact that bid-ask at the close is not available for all securities prior to 1993. All common stocks listed on the NYSE, AMEX, and NASDAQ with a price greater than $1 are including for the purposes of these tests. Three size sub-groups and two sub-periods are also reported. I have limited the analysis to value-weighted returns for the sake of brevity. Inferences are unchanged after transitioning to bid-ask midpoint at the close returns. The raw return and five-factor alpha is positive and significant with bid-ask midpoint at the close returns actually having a slightly higher raw return and five-factor alpha. The spread portfolio earns a 1.58% monthly return on average and a 1.55% monthly alpha with t-stats of 5.45 and 4.48 respectively. Both of these results are both economically and statistically significant. The annualized portfolio return is approximately 20.7%. While the t-stat associated with the five-factor alpha exceed the 3.00 t-stat threshold suggested by Harvey, Liu, and Zhu (2015).

Across size sub-groups these results are consistent with the results presented in Section 3. With the spread portfolio returns and five-factor alphas positive and significant across size sub-groups. The trend is also preserved with the largest returns in terms of magnitude and significance present for the small size sub-group. Economically, the results for all sub-groups remain meaningful with the lowest annualized return being 14.16% for the large sub-group. Unsurprisingly, the largest change in terms of the magnitude of both raw returns and alpha is in the small size sub-group when compared to the Section 3 results. This is expected since small firm returns are most likely to be impacted by market microstucture noise issues such as non-synchronous trading or non-trading. However, the raw return and alpha for this small sub-group remains positive and statistically significant even over the abbreviated sample period (1993-2016).

In addition to the size sub-groups, I subdivide the period from 1993-2016 into two separate periods from 1993-2004 and 2005-2016. *HLDiff* remains a positive and significant predictor of future returns across both sub-periods. *HLDiff* is stronger in the earlier period but remains positive and significant in the later periods as well. Raw returns and five-factor alphas persist adjusting returns are adjusted to be based on the bid-ask midpoint at the close.

I have also run Fama-MacBeth regressions based on the returns to the bid-ask midpoint at the close price. The results of these regressions are in Table B.18. As with the portfolio tests, the sample period is limited to 1993-2016 to ensure the bid-ask midpoint at the close is available for all securities throughout the testing period. The relationship between *HLDiff* and next months returns are positive and significant, in line with the results in Section 3. Here the returns are effectively equal-weighted. That is to say smaller firms are treated equally to larger firms in these regressions. Since market microstructure noise is more likely to have an impact on small firms, it is to be expected that the results attenuate more than the portfolio results presented above. The coefficient on *HLDiff* is 0.1341 with an t-stat of 9.40 for the full time period. This is in contrast to the coefficient for the closing price returns result of 0.2127 and a t-stat of 23.28. As in the Section 3 Fama-MacBeth regressions, *HLDiff* presents the largest t-stat of all coefficients included in the regression. To further address any concern about size, I have subdivided the sample into three size groups. Each size group has a positive and significant coefficient on the *HLDiff* variable. Consistent with earlier findings the result is strongest for the small group, but the result is not isolated to the small firms alone.

All of these results are strongly indicative of the fact that the predictive power of *HLDiff* does not simply manifest as a result of market microstructure noise. This approach is commonly adopted to address issues with market microstructure noise, but it is not necessarily a panacea. In the following section, I detail an alternative approach using a weighted least squares regression.

5.3.2 Asparouhova et al. (2010) Weighted Least Squares

To this point, the use of the bid-ask midpoint at the close has been predicated on the fact that the midpoint fairly reflects the efficient value of a firm. There are some theoretical reasons to believe that this may not always be true. In particular, Ho and Stoll (1980) develop a model in which liquidity providers may move quote midpoints in order to manage their security inventories. More recently, Fisher, Weaver, and Webb (2010) show that quote midpoints may be prevented from revealing efficient values. While the use of the midpoint may reduce microstructure bias, it cannot fully eliminate it. Asparouhova, Bessembinder, and Kalcheva (2010) note that 'the biases attributable to microstructure noise in prices can be effectively eliminated by a simple methodological correction where return premiums are estimated by a weighted least squares regressions that rely on individual security returns as the dependent variable, and that use the prior-period gross (one-plus) return as the weighting variable.' In addition, they note that microstructure noise (upward bias) in returns is largely attributable to illiquid stocks. Accordingly, I have adjusted my Fama-MacBeth cross-sectional regressions to weight according to the prior period return plus 1 (gross return) and dropped the least liquid firms from the sample.

As the returns in these regressions are effectively equal weighted, this allows for a consistent estimation of the coefficients. The results obtained from these weighted regressions are reflected on Table B.19 and are qualitatively similar to those presented in Section 3. I present four sets of results the first drops no firms the remaining three are after dropping the top 5%, 10%, and 25% most illiquid firms, respectively. Illiquidity is measured as:

$$ILLIQ_{i,m} = \frac{1}{D_{i,m}} \sum_{t=1}^{D_{i,m}} \frac{|R_{i,m,d}|}{VOLD_{i,m,d}}, \tag{5.1}$$

following Amihud (2002), where $D_{i,m}$ is the number of days where data is available for security i in month m, $R_{i,m,d}$ is the return to stock i on day d of month m, and $VOLD_{i,m,d}$ is the dollar trading volume on day d in month m for firm i. Before dropping any illiquid firms, the coefficient on *HLDiff* is 0.2067 with an associated t-stat of 22.78. The lowest coefficient is 0.1671 with a t-stat of 15.16 after dropping the 25% most illiquid stocks. The magnitude increases as more illiquid firms are included, but given these results it seems unlikely that the result is only associated with illiquid firms. To the extent that movements in the high and low prices are informative and not reflected in the closing price, a less liquid firm may take longer to incorporate this information. So this result and the diminishing magnitude and significance of the *HLDiff* coefficient is not entirely unexpected.

This weighted-least squares methodology is specifically suggested as a remedy to market microstructure noise problems. Even after adopting this methodology, *HLDiff* remains positive and significantly related to next months returns. Given these results combined with the bid-ask midpoint at the close results, it seems unlikely that the *HLDiff* relationship is simply an artifact of market microstructure noise. While this weighting methodology is specifically recommended for cross-sectional regressions of returns on firm variables, I also verify my portfolio results using this weighting convention and my inferences remain unchanged.

6. CONCLUSION

The preceding results demonstrate the substantial predictive power of price movements which are not reflected in the closing price. From 1965 to 2016, firms for which the high and low prices shifted up significantly more than the closing price earn subsequent value-weighted monthly returns of 1.15% on average. Firms for which the high and low prices shifted down significantly more than the closing price earn subsequent value-weighted monthly returns of -0.36% on average. The spread of 1.52% between these two portfolios is both economically and statistically significant. This result is strongest for small firms where the spread portfolio earns 2.77% per month on average. However, the result remains economically and statistically significant for large firms which earn 1.14% per month on average. The construction of *HLDiff* makes it unlikely that the relationship between *HLDiff* and subsequent returns is an artifact of market microstructure noise (i.e. bid-ask bounce or non-trading). However, the results are also robust to a variety of adjustments intended to address market microstructure biases that may exist in returns. This includes the utilization of bid-ask midpoints at the close as well as a weighted least squares regression. I also find that this effect remains strong in Fama-MacBeth regressions after including a variety of control variables. The coefficient on *HLDiff* exceeds, in terms of statistical significance, all other variables in the Fama-MacBeth regressions. These other variables include measures of size, book-to-market, reversal, return volatility, and momentum, among others.

The underlying mechanism driving the return predictability is most consistent with market frictions inhibit the correction of overpricing. Consistent with Miller (1977), Stambaugh, Yu, and Yuan (2012), and Stambaugh and Yuan (2016), the premium associated with the *HLDiff* factor is positively and significantly associated with the Baker and Wurgler (2006) sentiment index. In particular, the subset of firms that are most overpriced appear to present a significant relationship with lagged sentiment, while the long leg has no discernible association with lagged sentiment. These findings are consistent with mispricing and in particular, short-selling constraints limited the ability of investors to respond to observed overpricing.

Motivated by the above findings and the intuition of Hirshleifer and Jiang (2010), Stambaugh, Yu, and Yuan (2012), Stambaugh and Yuan (2016), and Kozak, Nagel, and Santosh (2018), I construct a mispricing factor based on *HLDiff* and evaluate its ability to explain the pricing errors of 12 well-known anomalies. Of the 12 anomalies considered, a two-factor model including the market risk premium and a factor created based on the *HLDiff* variable reduces the magnitude and significance of all but one of the anomalies. Further, three of the anomalies are statistically insignificant after adding the *HLDiff* factor to the single-factor CAPM. In addition, the mispricing factors suggested by Stambaugh and Yuan (2016) as well as the five-factor model of Fama and French (2016) are unable to price the *HLDiff* factor. I also employ the three-pass methodology of Giglio and Xiu (2018) and find that it is unlikely that my factor is the result of bias induced by omitted variables or measurement error.

The findings are also consistent with research connecting mispricing that arises from market frictions (Lamont and Thaler (2003) and Stambaugh, Yu, and Yuan (2012)) and seem to augment existing factor models by incorporating a factor to account for mispricing similar to Stambaugh and Yuan (2016). These findings indicate that a variety of financial metrics could be improved by incorporating additional prices. Examining possible ways to incorporate other prices into financial metrics and return calculations is an interesting area for further research.

REFERENCES

Abdi, F. and A. Ranaldo (2017). A simple estimation of bid-ask spreads from daily close, high, and low prices. *The Review of Financial Studies 30(12)*, 4437–4480.

Admati, A. R. and P. Pfleiderer (1988). A theory of intraday patterns: Volume and price variability. *Review of Financial Studies 1(1)*, 3–40.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets 5*(1), 31–56.

Amihud, Y. and H. Mendelson (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics 17(2)*, 223–249.

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

Asparouhova, E., H. Bessembinder, and I. Kalcheva (2010). Liquidity biases in asset pricing tests. *Journal of Financial Economics 96(2)*, 215–237.

Badrinath, S. G., J. R. Kale, and T. H. Noe (1995). Of shepherds, sheep, and the cross-autocorrelations in equity returns. *The Review of Financial Studies 8*(2), 401–430.

Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance 61(4)*, 1645–1680.

Barber, B. M. and T. Odean (2008). Do retail trades move markets? *The Review of Financial Studies 22*(1), 151–186.

Basak, S. and D. Cuoco (1998). An equilibrium model with restricted stock market participation. *The Review of Financial Studies 11(2)*, 309–341.

Beckers, S. (1983). Variances of security price returns based on high, low, and closing prices. *The Journal of Business 56(1)*, 97–112.

Ben-Rephael, A., Z. Da, and R. D. Israelsen (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies 30*(9), 3009–3047.

Blume, M. E. and R. F. Stambaugh (1983). Biases in computed returns: An application of the size effect. *Journal of Financial Economics 12(3)*, 387–404.

Bodnaruk, A. and P. Ostberg (2009). Does investor recognition predict returns? *Journal of Financial Economics 91(2)*, 208–226.

Bogousslavsky, V. (2016). Infrequent rebalancing, return autocorrelation and seasonality. *The Journal of Finance 71(6)*, 2967–3006.

Bogousslavsky, V. (2017). The cross-section of intraday and overnight returns. *Working Paper*.

Bondt, W. F. D. and R. Thaler (1985). Does the stock market overreact? *The Journal of Finance 40(3)*, 793–805.

Brennan, M. J., N. Jegadeesh, and B. Swaminathan (1993). Investment analysis and the adjustment of stock prices to common information. *The Review of Financial Studies 6*(4), 799–824.

Brennan, M. J. and A. Subrahmanyam (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics 41(3)*, 441–464.

Campbell, J. Y., J. Hilscher, and J. Szilagyi (2008). In search of distress risk. *The Journal of Finance 63(6)*, 2899–2939.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance 52(1)*, 57–82.

Carhart, M. M., R. Kaniel, D. K. Musto, and A. V. Reed (2002). Leaning for the tape: Evidence of gaming behavior in equity mutual funds. *The Journal of Finance 57(2)*, 661–693.

Chen, L., R. Novy-Marx, and L. Zhang (2011). An alternative three-factor model. Working Paper.

Cooper, M. J., H. Gulen, and M. J. Schill (2008). Asset growth and the cross-section of stock returns. *The Journal of Finance 63(4)*, 1609–1651.

Corwin, S. A. and P. Schultz (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance 67(2)*, 719–760.

Cushing, D. and A. Madhavan (2000). Stock returns and trading at the close. *Journal of Financial Markets 3(1)*, 45–67.

Daniel, K. and S. Titman (2006). Market reactions to tangible and intangible information. *The Journal of Finance 61(4)*, 1605–1643.

Drake, M. S., J. R. Thornock, and B. J. Twedt (2017). The internet as an information intermediary. *Review of Accounting Studies 22*(2), 543–576.

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

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116(1)*, 1–22.

Fama, E. F. and K. R. French (2016). Dissecting anomalies with a five-factor model. *Review of Financial Studies 29*, 69–103.

Fang, L. and J. Peress (2009). Media coverage and the cross-section of stock returns. *Journal of Financial Economics 64(5)*, 2023–2052.

Fisher, L., D. G. Weaver, and G. Webb (2010). Removing biases in computed returns. *Review of Quantitative Finance and Accounting 35(2)*, 137–161.

Frieder, L. and A. Subrahmanyam (2005). Brand perceptions and the market for common stock. *Journal of Financial and Quantitative Analysis 40*(1), 57–85.

George, T. J. and C.-Y. Hwang (2004). The 52-week high and momentum investing. *The Journal of Finance 59(5)*, 2145–2176.

Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica 57*, 1121–1152.

Giglio, S. and D. Xiu (2018). Asset pricing with omitted factors.

Greene, W. H. (2012). Econometric analysis.

Grullon, G., G. Kanatas, and J. P. Weston (2004). Advertising, breadth of ownership, and liquidity. *The Review of Financial Studies 17*(2), 439–461.

Gulen, H. and M. Woeppel (2018). Price-path convexity, extrapolation, and short-horizon return predictability. Working Paper, Purdue University.

Gurun, U. G. and A. W. Butler (2012). Don't believe the hype: Local media slant, local advertising, and firm value. *The Journal of Finance 67*(2), 561–598.

Harris, L. (1989). A day-end transaction price anomaly. *Journal of Financial and Quantitative Analysis 24(1)*, 29–45.

Harvey, C. R., Y. Liu, and H. Zhu (2015). ... and the cross-section of expected returns. *Review of Financial Studies 29(1)*, 5–68.

Hirshleifer, D. (1988). Residual risk, trading costs, and commodity futures risk premia. *The Review of Financial Studies 1(2)*, 173–193.

Hirshleifer, D., K. Hou, S. H. Teoh, and Y. Zhang (2004). Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics 38*, 297–331.

Hirshleifer, D. and D. Jiang (2010). A financing-based misvaluation factor and the cross-section of expected returns. *The Review of Financial Studies 23(9)*, 3401–3436.

Ho, T. and H. R. Stoll (1980). On dealer markets under competition. *The Journal of Finance 35(2)*, 259–267.

Hong, H. and J. Wang (2000). Trading and returns under periodic market closures. *The Journal of Finance 55(1)*, 297–354.

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *The Review of Financial Studies 20*(4), 1113–1138.

Hou, K. and T. J. Moskowitz (2005). Market frictions, price delay, and the cross-section of expected returns. *The Review of Financial Studies 18(3)*, 981–1020.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance 45(3)*, 881–898.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance 48(1)*, 65–91.

Kozak, S., S. Nagel, and S. Santosh (2018). Interpreting factor models. *The Journal of Finance 73(3)*, 1183–1223.

Lamont, O. A. and R. H. Thaler (2003). Can the market add and subtract? mispricing in tech stock carve-outs. *Journal of Political Economy 111*(2), 227–268.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance 20(4)*, 587–615.

Long, J. B. D., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990). Noise trader risk in financial markets. *Journal of Political Economy 98(4)*, 703–738.

Lou, D. (2014). Attracting investor attention through advertising. *The Review of Financial Studies 27*(6), 1797–1829.

Loughran, T. and J. R. Ritter (1995). The new issues puzzle. *The Journal of Finance 50(1)*, 23–51.

McInish, T. H. and R. A. Wood (1992). An analysis of intraday patterns in bid/ask spreads of nyse stocks. *The Journal of Finance 47(2)*, 753–764.

Mech, T. S. (1993). Portfolio return autocorrelation. *Journal of Financial Economics 34*(3), 307–344.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance 42(3)*, 483–510.

Miller, E. (1977). Risk uncertainty and divergence of opinion. *The Journal of Finance 32(4)*, 1151–1168.

Novy-Marx, R. (2013). The other side of value: Good growth and the gross profitability premium. *Journal of Financial Economics 56*, 335–381.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research 18*, 335–381.

Pastor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111(3)*, 642–685.

Ritter, J. R. (1991). The long-run performance of initial public offerings. *The Journal of Finance 46(1)*, 3–27.

Shapiro, A. (2002). The investor recognition hypothesis in a dynamic general equilibrium: Theory and evidence. *The Review of Financial Studies 15(1)*, 97–141.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19(3)*, 425–442.

Slezak, S. L. (1994). A theory of the dynamics of security returns around market closures. *The Journal of Finance 49(4)*, 1163–1211.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review 71*, 289–315.

Stambaugh, R. F., J. Yu, and Y. Yuan (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics 104(2)*, 288–302.

Stambaugh, R. F. and Y. Yuan (2016). Mispricing factors. *The Review of Financial Studies 30*, 1270–1315.

Titman, S., J. Wei, and F. Xie (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis 39(4)*, 677–700.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica 48*, 817–838.

Yang, D. and Q. Zhang (2000). Drift-independent volatility estimation based on high, low, open and close prices. *The Journal of Business 73(3)*, 477–492.

Yuan, Y. (2015). Market-wide attention, trading, and stock returns. *Journal of Financial Economics 116*(3), 548–564.

A. FIGURES

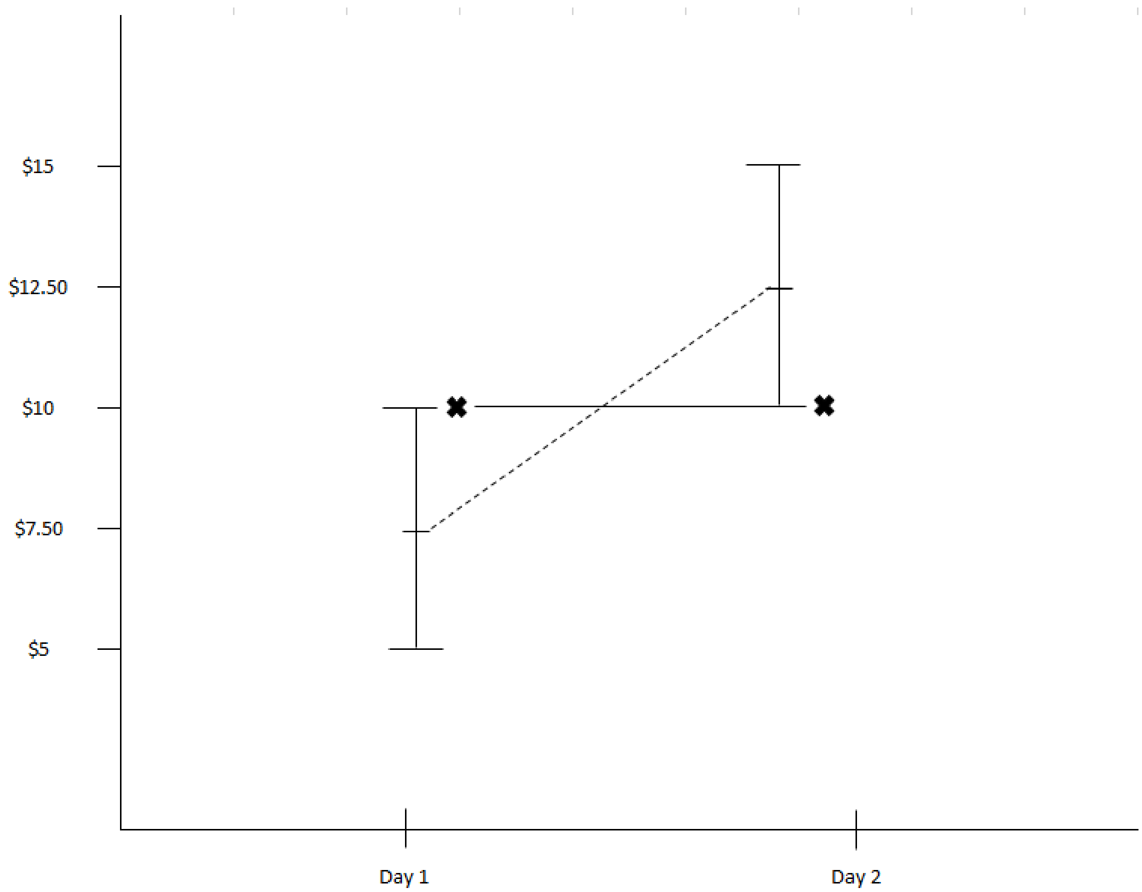
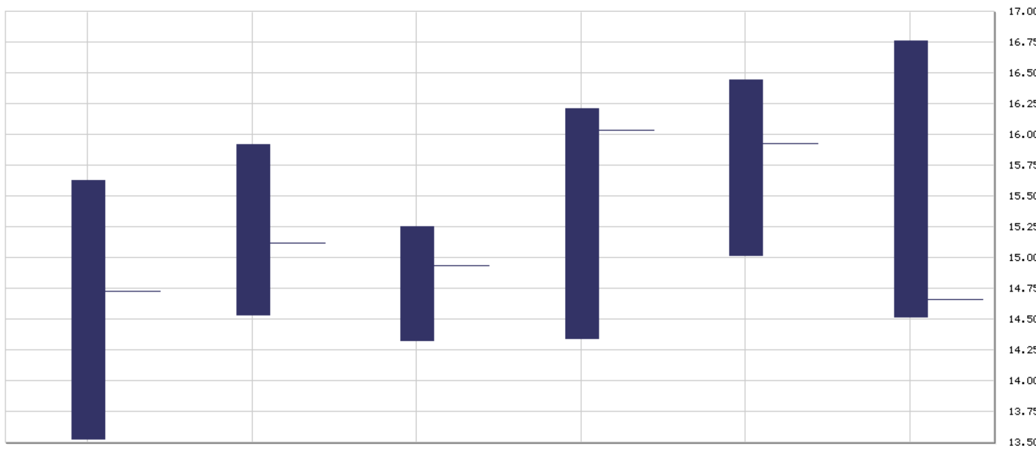


Figure A.1. The above figure is a simplified example of a shift in the high and low price that is not mirrored by a shift in the closing price. The closing price for each day is represented by the X on each respective day. The high and low prices are the horizontal lines for each day. The dotted line represents the return from the high and low midpoint moving from Day 1 to Day 2, while the solid line represents the close to close return. In the above example, the return to the high-low midpoint would be 66.67% ((12.5-7.5)/7.5). The close to close return, on the other hand, would be 0%. This would result in a difference of 66.67% for the purposes of calculating the variable of interest, *HLDiff*.



Date	12/23/2015	12/24/2015	12/28/2015	12/29/2015	12/30/2015	12/31/2015
High	15.63	15.92	15.25	16.21	16.45	16.76
Low	13.52	14.53	14.32	14.34	15.01	14.51
High-Low Midpoint	14.575	15.225	14.785	15.275	15.73	15.635
Close	14.73	15.12	14.94	16.04	15.93	14.66
HLRet		4.460%	-2.890%	3.314%	2.979%	-0.604%
Close Ret		2.648%	-1.190%	7.363%	-0.686%	-7.972%
Diff		1.812%	-1.700%	-4.049%	3.665%	7.368%
Accumulated Diff		1.812%	0.113%	-3.936%	-0.272%	7.097%

Figure A.2. The above chart presents the stock prices for MYOK from 12/23/2015 through 12/31/2015 to illustrate the calculation of *HLDiff*. This chart presents the high, low, and close prices for each of the corresponding days. Below the chart are the actual prices observed for the day, along with a simple return calculation for the high-low midpoint as well as the close to close return. The difference between these returns is also provided as well as the value of the accumulated difference over the time period presented.

B. TABLES

Table B.1.

Summary Statistics

The table below presents the summary statistics for the variable of interest, *HLDiff*, as calculated from the CRSP closing price returns. The sample period is from 1965 through 2016. The sample is limited to securities listed on the NYSE, NASDAQ, and AMEX stock exchanges and requires that stocks have a price above $1 at the end of the preceding month. *HLDiff* is the accumulated monthly difference between the return calculated using the midpoint between the high and low price on a given day and the associated closing price return. Daily return volatility is the standard deviation of CRSP closing price returns measured over the preceding month. The size is the log market capitalization of the firm. Book-to-Market is the log common equity (data item ceq) as measured at the end of the preceding fiscal year divided by the market equity at the end of the preceding month. Spread is the approximation of the daily spread as suggested by Abdi and Ranaldo (2017) using the two day adjustment. Turnover is the total volume observed in a given month divided by shares outstanding. Daily Range Volatility (rv) is the described in equation 2.5. All variable names for subsequent tables are included in this table parenthetically next to the associated variable. All control variables are measured at the end of the preceding month for the purposes of the Fama-MacBeth regressions.

Variable	**N**	**Mean**	**Std. Dev.**	**P5**	**P95**	**Q1**	**Q3**
HL Diff	2,792,747	-0.0001	0.0265	-0.0387	0.0385	-0.0084	0.0084
Ret(t-1)	2,769,831	0.0151	0.1695	-0.2033	0.2623	-0.0625	0.0732
Ret(t-2,t-7)	2,632,669	0.0885	0.4739	-0.4400	0.7500	-0.1426	0.2297
Daily Return Volatility (lvol)	2,790,868	0.0309	0.0247	0.0079	0.0737	0.0159	0.0386
Size (size)	2,772,128	11.5958	2.1386	7.46	15.40	10.00	13.00
Book-to-Market (lbm)	2,367,483	-0.5831	0.9482	-2.2371	0.7873	-1.0988	0.0252
Spread (lspread)	1,437,097	0.0162	0.01438	0.0030	0.0425	0.0072	0.0208
Turnover (lmto)	2,792,654	91.3878	233.6208	0.0000	337.810	12.5832	103.910
Range Volatility (rv)	2,792,747	1.0481	1.0853	0.2410	2.8952	0.4540	1.2255

Table B.2.: *HLDiff* Decile Returns: This table presents the average monthly value-weighted raw returns from January 1965 to December 2016 for portfolios formed based on *HLDiff*. *HLDiff*. Portfolios are formed every month from stocks listed on the NYSE, NASDAQ, and AMEX exchanges with prices in excess of $1. Panel A presents the equal-weighted returns for all firms as well as the three size subgroups and three sub-periods (1965-1982, 1983-2001, and 2002-2016). Panel B presents the value-weighted for all firms as well as the three size subgroups and sub-periods. Decile 1 represents the firms with the lowest value of *HLDiff*, and decile 10 represents the firms with the highest value of *HLDiff*. The spread portfolio returns are also presented, where the spread portfolio is formed by subtracting the return to decile 1 from the return to decile 10. The t-statistics reported are for the spread portfolio return. All values, excluding the t-statistic, are in decimal form where 0.01 is 1%.

Panel A

Equal Weighted Portfolios

	1 (Low)	2	3	4	5	6	7	8	9	10(High)	10-1	t-Stat
All Firms	-0.0048	0.0038	0.0042	0.0055	0.0067	0.0069	0.0082	0.0102	0.0116	0.0161	0.0209	21.41
Small	-0.0088	0.0013	0.0068	0.0071	0.0075	0.0064	0.0088	0.0113	0.0162	0.0212	0.0300	22.20
Medium	-0.0056	0.0041	0.0039	0.0057	0.0072	0.0075	0.0086	0.0091	0.0111	0.0135	0.0191	15.60
Large	-0.0008	0.0029	0.0045	0.0055	0.0061	0.0069	0.0072	0.0083	0.0103	0.0113	0.0123	10.56

Table B.2. (cont.)

Sub-periods

1965-1982	0.0010	0.0045	0.0031	0.0063	0.0074	0.0053	0.0084	0.0101	0.0101	0.0159	0.0149	11.64
1983-2000	-0.0100	0.0028	0.0046	0.0042	0.0050	0.0048	0.0060	0.0101	0.012597	0.0184	0.0285	18.49
2001-2016	-0.0026	0.0060	0.0061	0.0065	0.0084	0.0011	0.0010	0.0011	0.0122	0.0149	0.0175	8.54

Panel B

Value Weighted Portfolios

	1 (Low)	2	3	4	5	6	7	8	9	10(High)	10-1	t-Stat
All Firms	-0.0036	0.0003	0.0030	0.0038	0.0044	0.0056	0.0063	0.0077	0.0083	0.0115	0.0152	9.69
Small	-0.0098	0.0007	0.0053	0.0061	0.0070	0.0059	0.0083	0.0105	0.0140	0.0179	0.0277	19.62
Medium	-0.0047	0.0044	0.0040	0.0058	0.0068	0.0072	0.0088	0.0087	0.0111	0.0132	0.0178	14.08
Large	-0.0016	0.0011	0.0035	0.0039	0.0040	0.0051	0.0058	0.0079	0.0079	0.0097	0.0114	8.26

Sub-periods

1965-1982	0.0001	-0.0005	-0.0008	0.0006	0.0020	0.0040	0.0042	0.0026	0.0040	0.0078	0.0077	4.38
1983-2000	-0.0066	0.0012	0.0068	0.0063	0.0074	0.0081	0.0091	0.0139	0.0131	0.0158	0.0224	7.75
2001-2016	-0.0044	0.0003	0.0030	0.0045	0.0036	0.0045	0.0054	0.0063	0.0076	0.0106	0.0155	4.69

Table B.3.: *HLDiff* Five-Factor Alphas: This table presents the average monthly value-weighted five-factor alphas from January 1965 to December 2016 for portfolios formed based on the prior month's value of *HLDiff*. *HLDiff* is the accumulated monthly difference between the return to the high-low midpoint price and the closing price return on a given day. Portfolios are formed every month from stocks listed on the NYSE, NASDAQ, and AMEX exchanges with prices in excess of $1. Panel A presents the results for equal-weighted portfolios for all firms as well as the three size subgroups and sub-periods (1965-1982, 1983-2001, and 2001-2016). Panel B presents the results based value-weighted portfolios for all firms as well as the three size subgroups and sub-periods. Decile 1 represents the firms with the lowest value of *HLDiff*, and decile 10 represents the firms with the highest value of *HLDiff*. A spread portfolio alpha is also reported. The spread portfolio is formed by subtracting the decile 1 five-factor alpha from the decile 10 five-factor alpha. The t-statistics reported are for the spread portfolio. All values, excluding the t-statistic, are in decimal form where 0.01 is 1%. All t-statistics are heteroscedasticity-consistent following White (1980).

Panel A

Equal Weighted Portfolios

	1	2	3	4	5	6	7	8	9	10	10-1	t-Stat
All Firms	-0.0110	-0.0027	-0.0028	-0.0017	-0.0003	0.0000	0.0007	0.0029	0.0049	0.0095	0.0179	19.38
Small	-0.0157	-0.0057	-0.0002	0.0000	0.0016	-0.0005	0.0016	0.0041	0.0089	0.0140	0.0297	20.38
Medium	-0.0126	-0.0033	-0.0041	-0.0021	-0.0001	0.0003	0.0008	0.0014	0.0034	0.0065	0.0191	14.33
Large	-0.0056	-0.0031	-0.0019	-0.0010	-0.0006	0.0000	0.0005	0.0016	0.0042	0.0059	0.0115	9.07

Table B.3. (cont.)

Sub-periods

1965-1982	-0.0083	-0.0020	-0.0034	-0.0011	0.0008	-0.0006	0.0016	0.0025	0.0030	0.0060	0.0144	11.42
1983-2000	-0.0165	-0.0050	-0.0039	-0.0035	-0.0019	-0.0019	-0.0019	0.0022	0.0050	0.0133	0.0298	17.64
2001-2016	-0.0070	0.0002	-0.0001	0.0003	0.0012	0.0041	0.0035	0.0047	0.0061	0.0103	0.0173	8.16

Panel B

Value Weighted Portfolios

All Firms	-0.0074	-0.0039	-0.0026	-0.0019	-0.0014	-0.0005	0.0009	0.0029	0.0040	0.0072	0.0146	7.78
Small	-0.0167	-0.0068	-0.0021	-0.0015	0.0004	-0.0013	0.0006	0.0028	0.0066	0.0106	0.0273	17.77
Medium	-0.0114	-0.0028	-0.0039	-0.0019	-0.0002	0.0000	0.0011	0.0011	0.0036	0.0063	0.0177	12.73
Large	-0.0048	-0.0034	-0.0015	-0.0014	-0.0013	-0.0003	0.0001	0.0032	0.0037	0.0056	0.0104	6.77

Sub-periods

1965-1982	-0.0035	-0.0022	-0.0034	-0.0019	-0.0017	0.0007	0.0019	0.0005	0.0024	0.0046	0.0081	4.16
1983-2000	-0.0108	-0.0061	-0.0027	-0.0029	-0.0025	-0.0008	0.0000	0.0066	0.0059	0.0107	0.0215	5.70
2001-2016	-0.0085	-0.0036	-0.0013	-0.0002	-0.0004	-0.0010	0.0009	0.0021	0.0030	0.0059	0.0144	4.03

Table B.4.
Fama-MacBeth Regressions of Monthly Stock Returns on *HLDiff* and Other Variables

This table presents the results of the Fama-MacBeth regressions to determine if *HLDiff* is informative about future stock returns while controlling for other common return determinants. Monthly stock returns are regressed on lagged accounting and return variables. The sample extends from January 1965 through December 2016 and includes all stocks listed on the NYSE, NASDAQ, and AMEX exchanges with a stock price greater than $1. Control variables include returns from month t-1, buy and hold returns from month t-2 to t-7, a proxy for bid-ask spread as proposed by Abdi and Ranaldo (2017), monthly turnover calculated as monthly volume divided by shares outstanding (lmto), daily return volatility, which is the standard deviation of daily returns over the preceding month (lvol), size, which is log market capitalization at the end of the prior month(size), book-to-market which is the log of the book value of equity divided by the prior month's market capitalization (lbm), and a measure of daily range volatility (rv) as in equation 2.5. Along with the full sample, results for three sub-periods are also reported. Beta estimates are time-series averages of cross-sectional regression betas obtained from monthly cross-sectional regressions. All t-statistics have been Newey-West (1987) adjusted for 4 lags to address autocorrelation in coefficient estimates.

Sample Period	**1965-2016**		**1965-1982**		**1983-2000**		**2001-2016**	
Dependent Variable	$\mathrm{Ret}_{\mathrm{i,m}}$	t-stat	$\mathrm{Ret}_{\mathrm{i,m}}$	t-stat	$\mathrm{Ret}_{\mathrm{i,m}}$	t-stat	$\mathrm{Ret}_{\mathrm{i,m}}$	t-stat
Intercept	0.0186	5.38	0.0233	2.15	0.0082	1.11	0.0255	4.83
HL Diff	0.2115	23.95	0.1938	14.72	0.2570	25.03	0.1749	9.23
Ret(m-1)	-0.0250	-6.45	-0.0588	-8.01	-0.0011	-2.69	-0.0055	-1.19
Ret(m-7,m-2)	0.0090	5.68	0.0105	3.39	0.0198	9.22	0.0028	0.89
lspread	-0.0183	-0.42	0.2948	3.85	-0.0424	-1.04	-0.3172	-4.15
lmto	0.0000	0.80	0.0000	1.63	-0.0000	-0.58	-0.0000	-2.09
lvol	-0.2609	-9.11	-0.3720	-6.06	-0.3515	-12.31	-0.0469	-1.28
size	-0.0001	-0.40	-0.0009	-1.25	0.0008	1.77	-0.0004	-1.22
lbm	0.0384	5.98	0.0030	2.41	0.0055	5.40	0.0028	3.25
rv	-0.0082	-5.01	-0.0049	-1.52	-0.0026	-3.13	-0.0167	-5.28

Table B.5.
Size Sub-Group Fama-MacBeth Regressions of Monthly Stock Returns on *HLDiff* and Other Variables

This table represents Fama-MacBeth regressions to determine if *HLDiff* is informative about future stock returns across various size sub-groups while controlling for other common return determinants. Monthly stock returns are regressed on lagged accounting and return variables. The sample extends from January 1965 through December 2016 and includes all stocks listed on the NYSE, NASDAQ, and AMEX exchanges with a stock price greater than $1. Control variables include returns from month t-1, buy and hold returns from month t-2 to t-7, a proxy for bid-ask spread as proposed by Abdi and Ranaldo (2017), monthly turnover calculated as monthly volume divided by shares outstanding (lmto), daily return volatility which is the standard deviation of daily returns over the preceding month (lvol), size which is log market capitalization at the end of the prior month(size), and book-to-market which is the log of the book value of equity divided by the prior month's market capitalization (lbm), and a measure of daily range volatility (rv) as in equation 2.5. Beta estimates are time-series averages of cross-sectional regression betas obtained from monthly cross-sectional regressions. All t-statistics have been Newey-West (1987) adjusted for 4 lags to address autocorrelation in coefficient estimates.

Firm Size	**Small Sample**		**Medium Sample**		**Large Sample**	
Dependent Variable	$Ret_{i,m}$	t-stat	$Ret_{i,m}$	t-stat	$Ret_{i,m}$	t-stat
Intercept	0.0346	5.38	0.0227	3.09	0.0217	4.05
HL Diff	0.2469	23.42	0.1723	14.24	0.1661	11.61
Ret(m-1)	-0.0362	-7.29	-0.0160	-3.83	-0.0187	-4.12
Ret(m-7,m-2)	0.0099	5.78	0.0107	6.82	0.0065	3.21
lspread	-0.0642	-3.64	0.0261	0.42	0.0161	0.27
lmto	0.0000	0.07	0.0000	1.31	-0.0000	-0.03
lvol	-0.2426	-3.88	-0.3351	-7.07	-0.2161	-4.76
size	-0.0020	-0.04	-0.0003	0.48	-0.0004	-1.23
lbm	0.0053	4.18	0.0034	4.88	0.0015	2.06
rv	-0.0063	-4.12	-0.0118	-5.18	-0.0113	-4.26

Table B.6.

Decomposition of *HLDiff*

This table presents the Fama-MacBeth results after constructing *HLDiff* using only the high or low price from a given day. Monthly stock returns are regressed on *HLDiff* along with lagged accounting and return control variables. The sample extends from January 1965 through December 2016 and includes all stocks listed on the NYSE, NASDAQ, and AMEX exchanges with a stock price greater than $1. Control variables include returns from month t-1, buy and hold returns from month t-2 to t-7, a proxy for bid-ask spread as proposed by Abdi and Ranaldo (2017), monthly turnover calculated as monthly volume divided by shares outstanding (lmto), daily return volatility which is the standard deviation of daily returns over the preceding month (lvol), size which is log market capitalization at the end of the prior month(size), and book-to-market which is the log of the book value of equity divided by the prior month's market capitalization (lbm), and a measure of daily range volatility (rv) as in equation 2.5. Beta estimates are time-series averages of cross-sectional regression betas obtained from monthly cross-sectional regressions. All t-statistics have been Newey-West (1987) adjusted for 4 lags to address autocorrelation in coefficient estimates.

CRSP Close Returns	**Low Price**		**High Price**	
	Low Price	t-stat	High Price	t-stat
Intercept	0.0180	3.76	0.0109	2.72
HL Diff	0.1062	20.72	0.1201	18.28
Ret(m-1)	-0.0286	-7.36	-0.0266	-7.02
Ret(m-7,m-2)	0.0091	5.77	0.0093	5.88
lspread	-0.0185	-0.42	-0.0303	-0.68
lmto	0.0000	0.83	0.0000	0.92
lvol	-0.2498	-8.71	-0.2657	-9.17
size	-0.0001	-0.31	-0.0002	-0.50
lbm	0.0039	5.97	0.0384	5.97
rv	-0.0083	-5.06	-0.0083	-5.05

Table B.7.
Value-weighted Weekly Portfolio Results

This table presents the weekly buy-and-hold value-weighted raw returns and five-factor alphas for an arbitrage portfolio formed on the basis of *HLDiff*. The long leg of the portfolio is comprised of the decile of firms with the highest values of *HLDiff*, while the short leg is comprised of the decile firms with the lowest values of *HLDiff*. The sample extends from 1965 to 2016. All common stock traded on the NYSE, NASDAQ and AMEX with a share price greater than $1 are included for the sake of portfolio formation. Portfolios are held for up to 12 weeks with the results for holding periods of 1, 2, 4, 6, 8, 10, and 12 weeks reported. Weekly returns are calculated Wednesday to Wednesday. *HLDiff* is the daily return difference from the Tuesday immediately preceding portfolio formation. All t-statistics are heteroscedasticity-consistent following White (1980).

Holding Period	Raw Return	t-stat	5-Factor Alpha	t-stat
1 week	0.0065	13.53	0.0066	13.60
2 weeks	0.0077	12.43	0.0078	11.97
4 weeks	0.0079	10.10	0.0079	9.50
6 weeks	0.0084	9.97	0.0077	8.94
8 weeks	0.0072	7.34	0.0068	6.69
10 weeks	0.0076	6.90	0.0071	6.44
12 weeks	0.0085	7.19	0.0079	6.43

Table B.8.

Predicting Returns 1-week ahead

This table presents the results of a Fama-MacBeth regression of returns one week in the future regressed on *HLDiff* and a variety of control variables. The sample extends from January 1965 through December 2016 and includes all common stocks listed on the NYSE, NASDAQ, and AMEX exchanges with a stock price greater than $1. Weekly returns are calculated Wednesday to Wednesday. *HLDiff* is the daily difference between high-low midpoint returns and closing price returns accumulated over the preceding 20 trading days, which is regressed on the following weekly return. Control variables include returns from month t-1, buy and hold returns from month t-2 to t-7, a proxy for bid-ask spread as proposed by Abdi and Ranaldo (2017), monthly turnover calculated as monthly volume divided by shares outstanding (lmto), daily return volatility which is the standard deviation of daily returns over the preceding month (lvol), size which is log market capitalization at the end of the prior month(size), and book-to-market which is the log of the book value of equity divided by the prior month's market capitalization (lbm), and a measure of daily range volatility (rv) as in equation 2.5. Beta estimates are time-series averages of the cross-sectional betas obtained from monthly cross-sectional regressions. All t-statistics have been Newey-West (1987) adjusted for 4 lags to address autocorrelation in coefficient estimates.

Weekly Returns		
Dependent Variable	$\mathrm{Ret}_{i,w+1}$	t-stat
Intercept	0.0015	1.26
HL Diff	0.1756	38.27
Ret(m-1)	-0.0275	-20.78
Ret(m-7,m-2)	0.0038	8.62
lspread	-0.0405	-3.47
lmto	-0.0000	-0.07
lvol	-0.0735	-10.92
size	0.0002	1.90
lbm	0.0010	6.33
rv	-0.0108	-4.65

Table B.9.
Predicting Returns n-weeks ahead

This table presents the results of a Fama-MacBeth regression of returns n-weeks ahead regressed on *HLDiff* and a variety of control variables. The sample extends from January 1965 through December 2016 and includes all common stocks listed on the NYSE, NASDAQ, and AMEX exchanges with a stock price greater than $1. Weekly returns are calculated Wednesday to Wednesday. *HLDiff* is the daily difference between high-low midpoint returns and closing price returns accumulated over the preceding 20 trading days, which is regressed on the weekly return n-weeks ahead. Control variables include returns from month t-1, buy and hold returns from month t-2 to t-7, a proxy for bid-ask spread as proposed by Abdi and Ranaldo (2017), monthly turnover calculated as monthly volume divided by shares outstanding (lmto), daily return volatility which is the standard deviation of daily returns over the preceding month (lvol), size which is log market capitalization at the end of the prior month(size), and book-to-market which is the log of the book value of equity divided by the prior month's market capitalization (lbm), and a measure of daily range volatility (rv) as in equation 2.5. Beta estimates are time-series averages of the cross-sectional betas obtained from monthly cross-sectional regressions and are only reported for *HLDiff*. All t-statistics have been Newey-West (1987) adjusted for 4 lags to address autocorrelation in coefficient estimates.

	Weekly Returns	
# of weeks ahead	$\mathrm{Ret}_{\mathrm{i,w+1}}$	t-stat
1-week	0.1749	37.94
2-weeks	0.0136	6.00
4-weeks	0.0079	3.79
6-weeks	0.0051	2.57
8-weeks	0.0024	1.19
10-weeks	0.0003	0.14
12-weeks	0.0031	1.62

Table B.10.

Factor Summary Statistics and Correlations

This table reports the summary statistics for each of the factors from the five-factor model developed by Fama and French (2015) as well as the *HLDiff* factor. The table reflects the factor premiums, standard deviations, and associated t-statstics for each factor for the period from January 1965 to December 2016, as well as the correlation between each factor.

				Cross Correlations					
Factor Portfolio	**Excess Return**	**Std. Dev.**	**t-Stat**	**MKTRF**	**SMB**	**HML**	**RMW**	**CMA**	**HLDiff**
Mktrf	0.4817	4.484	2.68	1.00					
SMB	0.2671	3.080	2.16	0.28	1.00				
HML	0.3471	2.833	3.05	-0.27	-0.09	1.00			
RMW	0.2478	2.264	2.73	-0.23	-0.35	0.08	1.00		
CMA	0.3072	2.024	3.78	-0.39	-0.11	0.69	-0.03	1.00	
HLDiff	0.7683	2.445	7.83	-0.23	-0.17	0.02	0.28	0.13	1.00

Table B.11.
HLDiff and Investor Sentiment

The table presents the estimate of the coefficient from the regression of portfolio returns on lagged sentiment. The portfolio returns are the excess return from the long leg, short leg, and spread portfolio for the *HLDiff* factor. Investor sentiment is the prior month's Baker and Wurgler (2006) sentiment index. All t-statistics are heteroscedasticity-consistent following White (1980). The sample period is from January 1965 to December 2016.

	Long Leg		**Short Leg**		**Long-Short**	
	$\hat{\mathbf{b}}$	**t-Stat**	$\hat{\mathbf{b}}$	**t-Stat**	$\hat{\mathbf{b}}$	**t-Stat**
Intercept	0.8055	4.19	0.0445	0.20	0.7610	7.61
HLDiff	-0.2487	-1.17	-0.5013	-1.95	0.2525	2.13

Table B.12.

Anomaly Alphas with and without the *HLDiff* Factor

This table reports the alphas and t-statistics earned by the spread portfolios formed based on thirteen different anomalies. The alphas are reported for a single-factor model with and without the *HLDiff* factor and a five-factor model with and without the *HLDiff* factor. The sample periods for all anomalies is from January 1965 to December 2016. All t-statistics are heteroscedasticity-consistent following White (1980). All alphas reported are in decimal form where 0.01 is 1%. In addition to the alphas for each anomaly under the various models, I report the absolute average alpha and t-statistic for all anomalies under each model. At the bottom of the table, the Gibbons, Ross, and Shanken (1989) test statistics and p-values have been reported. The null hypothesis of the GRS test is that the alphas for all portfolios are jointly equal to zero. An effective factor model will produce a F-statistic from the GRS Test where the null hypothesis is not rejected.

	Single-Factor Model				**Five-Factor Model**			
	Without *HLDiff*		**With *HLDiff***		**Without *HLDiff***		**With *HLDiff***	
	Alpha	**t-Stat**	**Alpha**	**t-Stat**	**Alpha**	**t-Stat**	**Alpha**	**t-Stat**
Distress	0.0089	3.57	0.0023	0.90	0.0069	2.78	0.0036	1.49
O-Score	0.0032	1.28	0.0003	0.25	0.0041	3.73	0.0039	3.47
Net Stock Issues	0.0067	6.41	0.0052	4.77	0.0037	4.12	0.0034	3.53
Composite Equity Issues	0.0067	5.41	0.0057	4.44	0.0029	2.70	0.0035	2.33
Total Accruals	0.0050	3.67	0.0056	3.82	0.0054	4.00	0.0058	4.08
Net Operating Assets	0.0056	4.64	0.0052	4.11	0.0054	4.18	0.0050	3.90
Gross Profitability	0.0036	2.47	0.0011	0.74	0.0036	3.21	0.0028	2.39
Asset Growth	0.0054	4.14	0.0045	3.03	0.0003	0.32	-0.0004	-0.44
Return on Assets	0.0068	3.87	0.0021	1.25	0.0044	3.32	0.0026	1.97
Investment to Assets	0.0060	4.99	0.0060	4.80	0.0039	3.59	0.0036	3.31
Momentum	0.0136	5.33	0.0088	2.73	0.0139	4.55	0.0101	2.94
Idiosyncratic Volatility	-0.0156	-5.78	-0.0068	-2.42	-0.0098	-4.83	-0.0053	-2.52
Averages	0.0073	4.30	0.0044	2.77	0.0054	3.44	0.0039	2.70
	F-Statistic	**P-Value**	**F-Statistic**	**P-Value**	**F-Statistic**	**P-Value**	**F-Statistic**	**P-Value**
GRS Statisic	6.61	4.9772×10^{-11}	4.38	1.1513×10^{-6}	5.55	6.0331×10^{-9}	3.95	7.9363×10^{-6}

Table B.13.: Comparing the Ability of *HLDiff* to Explain the Factors of Other Models: On panel A of this table I present the alpha (in decimal form) and associated t-statistic for each factor with respect to the other models/factors being considered. All t-statistics are heteroscedasticity-consistent adopting the approach suggested by White (1980). The considered models are the five-factor model of Fama and French (2015) including the size(SMB), value(HML), profitability(RMW), and investment(CMA) factors and the mispricing factor model(M-4) of Stambaugh and Yuan (2016) including an alternative size (SMB) factor along with the management(MGMT) and performance(PERF) mispricing factors. The last item is the model including *HLDiff*, which is comprised of the market risk premium and the *HLDiff* factor. Panel B presents the GRS test results. This panel is testing if a given model produces zero alpha when used to explain the premium associated with another model's factors. P-values are reported along with the associated F-statistics. The sample period is from January 1965 to December 2016.

Panel A: Alphas and t-Statistics						
	Five-Factor Model		**M-4 Model**		***HLDiff* Factor**	
	Alpha	**t-Statistic**	**Alpha**	**t-Statistic**	**Alpha**	**t-Statistic**
Five-Factor Model Factors						
SMB	-	-	-0.0007	-1.58	0.0030	2.46
HML	-	-	-0.0002	-0.28	0.0047	4.09
RMW	-	-	0.0008	0.95	0.0012	1.35
CMA	-	-	-0.0007	-1.14	0.0036	4.23

Table B.13. (cont.)

M-4 Factors						
SMB	0.0017	4.43	-	-	0.0044	3.84
MGMT	0.0035	5.92	-	-	0.0066	6.27
PERF	0.0068	4.85	-	-	0.0039	2.32
HLDiff	0.0072	6.98	0.0054	4.89	-	-

Panel B: GRS F-Statistics and P-values

	F-Statistic	**P-Value**	**F-Statistic**	**P-Value**	**F-Statistic**	**P-Value**
SMB, HML, RMW, CMA	-	-	4.32	0.0019	9.15	3.5367×10^{-7}
SMB, MGMT, PERF	23.21	3.1197×10^{-14}	-	-	26.43	4.4409×10^{-16}
HLDiff	54.23	5.8298×10^{-13}	30.73	4.4074×10^{-8}	-	-

Table B.14.
Double-Sorted Portfolios on *HLDiff* and Idiosyncratic Volatility

This table presents the average monthly value-weighted raw returns and five-factor alphas from January 1965 to December 2016 for double-sorted portfolios. First, firms are separated into quintiles based on Idiosyncratic Volatility. The firms in each Idiosyncratic Volatility portfolio are then sorted into quintiles based on *HLDiff*. Idiosyncratic Volatility is consistent with Ang, Hodrick, Xing, and Zhang (2006) using the standard deviation of the daily return residuals from a Fama-French Three-Factor model. *HLDiff* is calculated as described in the body of the paper. Panel A presents the raw returns and associated long-short portfolios for both Idiosyncratic volatility and *HLDiff*. Panel B presents the five-factor alphas for both individual as well as long-short portfolios. t-statistics are presented for long-short portfolios. All regression t-statistics are heteroscedasticity-consistent following White (1980).

Panel A: Raw Returns

HLDiff Portfolios

Raw Returns

		1	2	3	4	5	(5-1)	t-stat
	1	0.0062	0.0080	0.0096	0.0106	0.0125	0.0063	6.56
Idio.	2	0.0059	0.0087	0.0090	0.0119	0.0145	0.0086	7.21
Volatility	3	0.0053	0.0081	0.0116	0.0116	0.0149	0.0096	6.55
Portfolios	4	0.0006	0.0051	0.0092	0.0114	0.0152	0.0146	8.12
	5	-0.0101	-0.0017	0.0008	0.0056	0.0095	0.0195	8.24
	(5-1)	-0.0163	-0.0097	-0.0088	-0.0051	-0.0031		
	t-stat	-4.94	-3.17	-3.13	-1.82	-0.97		

Panel B: Five-Factor Alphas

		1	2	3	4	5	(5-1)	t-stat
	1	-0.0032	-0.0017	0.0001	0.0007	0.0031	0.0063	6.13
Idio.	2	-0.0042	-0.0019	-0.0020	0.0017	0.0050	0.0091	7.07
Volatility	3	-0.0028	-0.0012	0.0016	0.0016	0.0057	0.0084	5.15
Portfolios	4	-0.0076	-0.0030	-0.0059	0.0027	0.0064	0.0140	6.89
	5	-0.0182	-0.0091	-0.0063	0.0009	0.0009	0.0191	7.40
	(5-1)	-0.0150	-0.0074	-0.0065	-0.0025	-0.0022		
	t-stat	-6.62	-3.60	-3.52	-1.33	-1.08		

Table B.15.
Double-Sorted Portfolios on *HLDiff* and Momentum

This table presents the average monthly value-weighted raw returns and five-factor alphas from January 1965 to December 2016 for double-sorted portfolios. First, firms are separated into quintiles based on Momentum. The firms in each Momentum portfolio are then sorted into quintiles based on *HLDiff*. Momentum is consistent with Jegadeesh and Titman (1993) using the 11-1-1 strategy. *HLDiff* is calculated as described in the body of the paper. Panel A presents the raw returns and associated long-short portfolios for both Momentum and *HLDiff*. Panel B presents the five-factor alphas for both individual as well as long-short portfolios. I report t-statistics for long-short portfolios only. All regression t-statistics are heteroscedasticity-consistent following White (1980).

HLDiff Portfolios

Panel A: Raw Returns

		1	2	3	4	5	(5-1)	t-stat
Momentum Portfolios	1	-0.0099	0.0003	0.0024	0.0064	0.0066	0.0166	7.27
	2	0.0001	0.0064	0.0087	0.0100	0.0117	0.0116	7.00
	3	0.0032	0.0064	0.0083	0.0108	0.0127	0.0095	7.47
	4	0.0049	0.0088	0.0110	0.0114	0.0135	0.0086	6.77
	5	0.0088	0.0095	0.0131	0.0165	0.0169	0.0082	4.97
	(5-1)	0.0188	0.0093	0.0107	0.0101	0.0103		
	t-stat	5.87	3.27	4.14	3.83	3.47		

Panel B: Five-Factor Alphas

		1	2	3	4	5	(5-1)	t-stat
Momentum Portfolios	1	-0.0189	-0.0093	-0.0086	-0.0037	-0.0037	0.0152	6.16
	2	-0.0102	-0.0031	-0.0020	-0.0002	0.0018	0.0120	6.56
	3	-0.0066	-0.0041	-0.0019	0.0005	0.0025	0.0092	6.46
	4	-0.0048	-0.0018	0.0004	0.0012	0.0029	0.0078	5.98
	5	0.0010	0.0007	0.0035	0.0068	0.0080	0.0070	3.89
	(5-1)	0.0199	0.0101	0.0121	0.0106	0.0117		
	t-stat	6.11	3.03	3.99	3.66	3.65		

Table B.16.

Omitted Factor Principal Components Testing

This table presents the estimated risk premium, t-statistic, and a signal-to-noise ratio (R_g^2) for the Fama-French five-factors, Pastor and Stambaugh (2003), and the *HLDiff* factor based on the methodology of Giglio and Xiu (2018). This analysis recovers the factor space by extracting principal components from 202 distinct portfolios. This is in contrast to the traditional methods of relatively arbitrary factor selection. I have presented the results extracting 4, 5, and 6 principal components. All t-statistics have been Newey-West (1987) adjusted for 4 lags.

	4 Principal Components			5 Principal Components			6 Principal Components		
Factor	Premium	t-Stat	R_g^2	Premium	t-Stat	R_g^2	Premium	t-Stat	R_g^2
Market	0.0037	1.85	0.9822	0.0037	1.76	0.9890	0.0035	1.59	0.9909
SMB	0.0023	1.77	0.9400	0.0023	1.77	0.9478	0.0023	1.77	0.9725
HML	0.0021	1.75	0.6760	0.0021	1.75	0.6889	0.0020	1.67	0.7917
RMW	0.0016	2.21	0.3843	0.0017	2.12	0.4141	0.0017	2.01	0.4687
CMA	0.0014	1.63	0.4486	0.0014	1.63	0.4638	0.0014	1.57	0.5570
Liquidity	0.0026	2.17	0.1199	0.0026	2.00	0.1202	0.0025	1.92	0.1202
HLDiff	0.0018	3.60	0.1919	0.0018	3.00	0.1967	0.0018	3.00	0.1971

Table B.17.: *HLDiff* Bid-Ask Midpoint Returns and Five-Factor Alphas: This table presents the average monthly value-weighted raw returns from January 1993 to December 2016 for portfolios formed based on the prior month's value of *HLDiff*. Portfolios are formed every month from common stocks listed on the NYSE, NASDAQ, and AMEX exchanges with prices in excess of $1. The sample period begins in 1993 as bid and ask prices at the close are not available for all securities in the CRSP database prior to this date. Panel A presents the raw returns for the complete sample of firms, three size sub-groups, and two sub-periods. Panel B presents the five-factor alphas for the complete sample of firms, three size sub-groups, and two sub-periods. Decile 1 represents the firms with the lowest value of *HLDiff*, and decile 10 represents the firms with the highest value of *HLDiff*. A spread portfolio is also reported. The spread portfolio is formed by subtracting the return to decile 1 from the return to decile 10. The t-statistics reported are for the spread portfolio return. All values, excluding the t-statistic, are in decile form where 0.01 is 1%. Alphas t-statistics are heteroscedasticity-consistent following White (1980).

Panel A

Raw Returns

	1 (Low)	2	3	4	5	6	7	8	9	10(High)	10-1	t-Stat
All Firms	-0.0042	0.0006	0.0044	0.0050	0.0051	0.0054	0.0074	0.0094	0.0079	0.0116	0.0158	5.45
Small	-0.0073	0.0032	0.0062	0.0078	0.0064	0.0084	0.0094	0.0090	0.0104	0.0104	0.0178	6.80
Medium	-0.0034	0.0051	0.0073	0.0074	0.0086	0.0081	0.0088	0.0099	0.0110	0.0105	0.0138	6.71
Large	-0.0021	0.0033	0.0043	0.0062	0.0049	0.0055	0.0066	0.0094	0.0093	0.0090	0.0111	4.51

Table B.17. (cont.)

Sub-periods

1993-2004	-0.0072	0.0000	0.0042	0.0048	0.0051	0.0068	0.0066	0.0110	0.0092	0.0127	0.0199	4.76
2005-2016	-0.0011	0.0011	0.0046	0.0052	0.0051	0.0039	0.0083	0.0078	0.0066	0.0105	0.0116	2.92

Panel B

Alphas

	1 (Low)	2	3	4	5	6	7	8	9	10(High)	10-1	t-Stat
All Firms	-0.0103	-0.0055	-0.0012	-0.0016	-0.0015	-0.0012	0.0012	0.0035	0.0018	0.0052	0.0155	4.48
Small	-0.0124	-0.0027	0.0003	0.0011	0.0010	0.0025	0.0032	0.0032	0.0046	0.0041	0.0165	7.33
Medium	-0.0093	-0.0022	-0.0005	-0.0009	0.0006	-0.0007	0.0008	0.0020	0.0040	0.0035	0.0128	5.95
Large	-0.0067	-0.0029	-0.0018	0.0000	-0.0019	-0.0009	0.0002	0.0037	0.0037	0.0029	0.0096	3.69

Sub-periods

1993-2004	-0.0118	-0.0074	-0.0026	-0.0029	-0.0031	-0.0004	0.0008	0.0057	0.0054	0.0053	0.0171	3.84
2005-2016	-0.0079	-0.0044	-0.0009	-0.0005	-0.0002	-0.0020	0.0021	0.0020	-0.0002	0.0047	0.0126	2.71

Table B.18.

Fama-MacBeth Regressions on Bid-Ask Midpoint Returns

This table presents the results of the Fama-MacBeth regressions to determine if *HLDiff* is informative about future stock returns while controlling for other common return determinants. The sample extends from January 1993 through December 2016 and includes all stocks listed on the NYSE, NASDAQ, and AMEX exchanges with a stock price greater than $1. The sample period begins in 1993 as bid and ask prices at the close are not available for all securities in the CRSP database prior to this date. Panel A reflects results obtained using the CRSP closing price returns. Panel B reports results obtained using the bid-ask midpoint at the close returns. Control variables include returns from month t-1, buy and hold returns from month t-2 to t-7, a proxy for bid-ask spread as proposed by Abdi and Ranaldo (2017), monthly turnover calculated as monthly volume divided by shares outstanding (lmto), daily return volatility, which is the standard deviation of daily returns over the preceding month (lvol), size, which is log market capitalization at the end of the prior month(size), and book-to-market which is the log of the book value of equity divided by the prior month's market capitalization (lbm). Beta estimates are time-series averages of cross-sectional regression betas obtained from monthly cross-sectional regressions. All t-statistics have been Newey-West (1987) adjusted for 4 lags to address autocorrelation in coefficient estimates.

Bid-Ask Midpoint Returns	**Full Sample**		**Small Sample**		**Medium Sample**		**Large Sample**	
Dependent Variable	$Ret_{i,m}$	t-stat	$Ret_{i,m}$	t-stat	$Ret_{i,m}$	t-stat	$Ret_{i,m}$	t-stat
Intercept	0.0193	4.55	0.0351	4.76	0.0174	1.71	0.0242	3.41
HL Diff	0.1328	9.57	0.1209	9.73	0.1153	6.81	0.1085	4.43
Ret(m-1)	0.0003	0.29	0.0013	0.30	0.0049	1.11	0.0003	0.04
Ret(m-7,m-2)	0.0132	5.78	0.0171	8.51	0.0127	5.16	0.0077	2.70
lspread	-0.2234	-3.64	-0.1910	-3.11	-0.2919	-3.42	-0.1109	-1.04
lmto	0.0000	0.07	0.0000	-1.67	0.0000	-0.16	-0.0000	-0.46
lvol	-0.1201	-3.88	-0.1623	-5.30	-0.0504	-1.00	-0.0187	-0.33
size	-0.0000	-0.04	-0.0016	-2.08	0.0002	0.29	-0.0004	-0.90
lbm	0.0038	4.18	0.0051	5.70	0.0036	3.77	0.0012	1.45
rv	-0.0124	-5.02	-0.0096	-4.60	-0.0163	-4.24	-0.0183	-3.91

Table B.19.

Weighted Least Squares Fama-MacBeth Regressions

This table presents the results of the Fama-MacBeth regressions utilizing a weighted least squares methodology suggested by Asparouhova et al. (2010). Monthly stock returns are regressed on lagged accounting and return variables where the regressions are weighted according to the prior months return plus 1. Results are reported for samples including only the most liquid percentiles of firms. The sample extends from January 1965 through December 2016 and includes all stocks listed on the NYSE, NASDAQ, and AMEX exchanges with a stock price greater than $1. Control variables include returns from month t-1, buy and hold returns from month t-2 to t-7, a proxy for bid-ask spread as proposed by Abdi and Ranaldo (2017), monthly turnover calculated as monthly volume divided by shares outstanding (lmto), daily return volatility, which is the standard deviation of daily returns over the preceding month (lvol), size, which is log market capitalization at the end of the prior month(size), and book-to-market which is the log of the book value of equity divided by the prior month's market capitalization (lbm), and a measure of daily range volatility (rv) as in equation 2.5. Beta estimates are time-series averages of cross-sectional regression betas obtained from monthly cross-sectional regressions. All t-statistics have been Newey-West (1987) adjusted for 4 lags to address autocorrelation in coefficient estimates.

	Full Sample		**95% Most Liquid**		**90% Most Liquid**		**75% Most Liquid**	
Dependent Variable	$Ret_{i,m}$	t-stat	$Ret_{i,m}$	t-stat	$Ret_{i,m}$	t-stat	$Ret_{i,m}$	t-stat
Intercept	0.0183	3.80	0.0195	4.09	0.0217	4.64	0.0231	4.83
HL Diff	0.2087	23.56	0.2057	22.88	0.1903	19.93	0.1741	16.92
Ret(m-1)	-0.0251	-6.79	-0.0243	-6.61	-0.0178	-4.82	-0.0142	-3.80
Ret(m-7,m-2)	0.0099	6.40	0.0099	6.41	0.0098	6.47	0.0092	5.87
lspread	-0.0695	-1.59	-0.0525	-1.23	-0.0158	-0.37	-0.0171	-0.39
lmto	0.0000	1.91	0.0000	1.94	0.0000	1.65	0.0000	1.39
lvol	-0.2353	-8.23	-0.2384	-7.83	-0.2546	-7.41	-0.2660	-6.88
size	-0.0001	-0.37	-0.0002	-0.58	-0.0003	-1.05	-0.0004	-1.29
lbm	0.0038	5.81	0.0037	5.69	0.0035	5.31	0.0032	4.87
rv	-0.0083	-5.10	-0.0091	-5.27	-0.0103	-5.28	-0.0108	-5.14

VITA

Mitchell S. Johnston

Purdue University Krannert School of Management

403 West State Street, West Lafayette, IN 47907

Contact Information

Mobile: 312-550-1854

E-Mail: johnst32@purdue.edu

Education

Purdue University Krannert School of Management *2014-2019*

PhD in Finance *(expected)*

Purdue University Krannert School of Management *2012-2014*

MBA, Concentration in Finance

Ohio State University Fisher College of Business *2001-2006*

BSBA Concentration in Accounting (with Honors, Summa Cum Laude)

Academic Interests

Research: Empirical Asset Pricing, Liquidity, Market Efficiency, Corporate Finance

Teaching: Investments, Financial Management, Financial Statement Analysis, Managerial and Financial Accounting

Job Market Paper

A High-Low Price Anomaly

I examine movements in the closing price that are different than the movements of the high and low prices on a given day. Instances in which the closing price deviates from the movements in the high-low midpoint are a strong predictor of future abnormal returns. The predictive power of my variable of interest, *HLDiff*, holds across size groups and sub-periods as well as in the presence of other common determinants of stock returns. I also find that the *HLDiff* factor is consistent with mispricing and is capable of explaining the pricing errors of six well-known anomalies.

Working Papers

The Real Effects of Balance Sheet Illiquidity

This paper proposes a simple and complete measure of balance sheet illiquidity and studies the real effects of firm illiquidity. I find that more illiquid firms reduce investment and net debt issuance and experience higher cost of debt and cost of equity. These results are consistent with the fact that illiquid firms are less flexible and less capable of redeploying assets to other ends including using them as stores of capital or as collateral to finance debt agreements. I also include the exogenous shock of the Lehman Brothers collapse and again find that more illiquid firms are disproportionately impacted by the shock to liquidity. All of these results highlight the importance of asset liquidity for firm decisions.

Work In Progress

Balance Sheet Liquidity and Stock Returns

Common stock represents a proportional ownership share of the underlying assets of the firm's assets. Utilizing a variety of liquidity measures to categorizes firms as liquid or illiquid relative to their industry peers, I consider the associated premium for holding a firm that is more illiquid than other firms in a similar industry.

Production Incentives and Optimal Production in Oil Firms *with Huseyin Gulen, Kateryna Holland, and William O'Brien*

Utilizing a hand collect data set of executive compensation objectives for oil firms, we examine the influence of production goals on decisions to produce when the price of oil may not necessarily justify the additional production.

Teaching Experience

Financial Management, Instructor with full responsibility, 2017, 2018

Teaching Evaluations (5.0/5.0)

Awards

Krannert Certificate for Distinguished Teaching, 2017

Krannert Scholar (Top 5%), 2013

Skills

Programming: Stata, SAS, Matlab

Professional Credentials: CPA Certified

Work Experience

Equity Residential

KPMG LLP

References

Huseyin Gulen (Chair)
Professor of Finance
Krannert School of Management
Purdue University
Phone: 765-496-2689
E-mail: hgulen@purdue.edu

Sergey Chernenko
Assistant Professor of Finance
Krannert School of Management
Purdue University
Phone: 765-494-4133
E-mail: schernen@purdue.edu

Deniz Yavuz
Associate Professor of Finance
Krannert School of Management
Purdue University
Phone: 765-494-1337
E-mail: myavuz@purdue.edu

Adem Atmaz
Assistant Professor of Finance
Krannert School of Management
Purdue University
Phone: 765-494-6725
E-mail: aatmaz@purdue.edu