Proceedings of the 2nd International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/17/20231064

Arbitrage Strategy Based on DHS Pricing Model

Jiaxuan Li[1,a,*], Yitong Chen[2,b], Yumin Wu[3,c], Dizhao Zhang[4,d], Zhutian Gao[5,e]

[1]The Chinese University of Hong Kong (Shenzhen), School of Management and Economics, Shenzhen, 518172, China
[2]School of Business, Hong Kong Baptist Univeristy, Hong Kong, 999077, China
[3]Shanghai University of International Business and Economics, Shanghai, 201620, China
[4]School of Information and Safety Engineering, Zhongnan University of Economics and Law, Wuhan, 430073, China
[5]New Jersey, USA
a. 120090093@link.cuhk.edu.cn, b. 19250835@life.hkbu.edu.hk, c. 21078009@suibe.edu.cn, d. 202021130112@stu.zuel.edu.cn, e. garygao333@gmail.com
**correspondence author*

Abstract: The Daniel-Hirshleifer-Sun (DHS) is a three-factor model based on the investor's psychology. It supplements the market factors of the CAPM model with two behavioral factors that capture commonalities in mispricing resulting from psychological biases. The DHS method focuses on two psychological biases affecting asset prices: overconfidence and limited attention. According to Daniel, Hirshleifer, and Sun, overconfidence in the investor tends to induce commonality in long-horizon mispricing.In contrast, the inattention of the investor tends to induce commonality in short-horizon mispricing. In this strategy, assets are priced according to the DHS model, and the unexplained return generated from this model is traded. According to the back-test, the explanation power of the DHS model is limited in Chinses market. As a result, the arbitrage strategy based on this model cannot generate a decent return in the long run. However, this strategy generates a significant positive return in turbulent market conditions. During these periods, investors tend to panic, and their psychology is especially unstable, so the two behavioral factors can explain the return efficiently.

Keywords: Daniel-Hirshleifer-Sun three-factor model, arbitrage strategy, Shanghai and Shenzhen stock market

1. Introduction

1.1. Idea

This paper examines a multi-factor stock selection arbitrage strategy using the Daniel-Hirshleifer-Sun three-factor model (DHS model), which is based on CAPM and considers investors' psychology. Daniel, K. D., D. A. Hirshleifer, and L. Sun released the model in 2020 [1].

1.2. Highlight

1.2.1. Economic Intuition

Two cognitive biases of the investor, overconfidence and limited attention, can be a reason for mispricing. Irrational investors are often overconfident, resulting in an insufficient response to listed companies' issuance and repurchase behavior. Besides, a large body of empirical evidence shows that stocks of companies that generate positive earnings surprises can earn higher returns in the following six to nine months compared to others with negative earnings surprises [1]. The underlying reason is that investors underreact to the latest earnings updates due to their limited attention. We will explore the Daniel-Hirshleifer-Sun multiple-factor model in this paper. Two factors, FIN and PEAD, are used to measure the bias. The two factors and the MKT factor compose the DHS model. Daily returns of the individual asset are regressed on MKT, FIN, and PEAD factors, and expected returns are then calculated. This pricing model works well in the US market. If it works well in the Chinese market, profit could be obtained from the mispricing of each asset.

1.2.2. Signal Generation

DHS model trading strategy will make regression-based FIN, PEAD, and MKT. MKT is excess return on the market portfolio. FIN captures investors' overconfident bias, and PEAD is the post-earnings announcement drift factor that captures limited investor attention [1]. Suppose the DHS model can sufficiently explain Shanghai and Shenzhen markets, as in the US market. In that case, there is no significant epsilon, stock return minus the projected return from the DHS model, in the market. Therefore, the assets with negative or positive epsilon are supposed to be mispriced. If the asset's epsilon is negatively significant, showing the asset is underpriced, we will remain the asset long. If epsilon is positively significant, showing an overpriced asset, we will short the asset.

1.2.3. Portfolio Construction

Assets are all traded in Shanghai and Shenzhen stock markets. The portfolio is rebalanced daily, and the last 20% of stocks with the most negative signal are purchased. Considering that the Chinese market has stringent regulations on short-selling and insufficient securities lending capacity, this strategy does not consider profit or loss generated from short-selling. Signal weighting will be applied for portfolio construction.

1.3. Performance Estimate

Our strategy is relatively new and has not been tested in the Chinese market. Results from testing on other similar strategies and the DHS model result based on the US market serve as performance estimations. The mere difference between Fama-French three-factor model (FF model) and the DHS model is that the DHS model uses FIN and PEAD instead of SMB (small minus big) and HML (high minus low) in the FF model. The methods used in the FF regression model and trading on epsilon are approximately the same as the DHS model.

According to Blanco [2], there is empirical evidence in favor of the Fame-French three-factor model. For example, Hu, Chen, Shao, and Wang (2019), tested the Fama-French model in the Chinese market, using data from 1995 to 2016, and got an average return of 1.23% [3]. It is also estimated that the DHS composite model outperforms the profitability-based model of Novy-Marx (2013), the five-factor model of Fama and French (2015), the q-factor model of Hou, Xue, and Zhang (2015), and the mispricing model of Stambaugh and Yuan (2017) in explaining the 34 anomalies in the US

DOI: 10.54254/2754-1169/17/20231064

market [1]. A study conducted by Lian, Liu, and Shi (2021) on measuring the overreaction and underreaction quantitatively to asset pricing in China stock market shows a maximum annual sharpe ratio of 2.02 [4].

1.4. Final Result

The performance of the DHS model has been invalid in the Shanghai and Shenzhen stock markets in recent times. The DHS model can merely explain 40% of the overall market condition. The model has limited explanatory power for the market, so the arbitrage strategy cannot generate a decent return and an ideal Sharpe ratio in the long term. According to the back-test, the Sharpe ratio attained in the Chinese market from 2008-2018 is only around **0.01964**. However, this strategy works effectively during turbulent market conditions, when the psychology of investors during these periods is especially unstable, so timing is a crucial part of this strategy.

2. Specification

2.1. Analysis

2.1.1. Qualitative Analysis

The DHS model is a three-factor model based on the investor's psychology. It supplements the market factors of the CAPM model with two behavioral factors that capture commonalities in mispricing using psychological biases. The DHS method focuses on two psychological biases affecting asset prices: overconfidence and limited attention. Overconfidence in the investor tends to induce commonality in long-horizon mispricing, while the inattention of the investor tends to induce commonality in short-horizon mispricing. The economic intuition behind the Long Horizon Financing Factor is that the investor's overconfidence, which results in an insufficient response to the issuance and repurchase behaviors of listed companies, will induce commonality in longer horizon mispricing [1]. Investors can take advantage of this mispricing by predicting it and positioning our portfolio accordingly. The long-horizon financing factor exploits the information in the manager's decision to issue or repurchase equity in response to persistent mispricing. The Long-Horizon Financing Factor, FIN, reflects returns associated with mispricing that generally happen in a time horizon greater than one year [1]. The Financing Factor is based on actions that increase issuance measures, such as seasoned issues or equity-financed acquisitions.

The economic intuition behind the Short Horizon Financing Factor, PEAD, is that the investor's limited attention makes them underreact to the latest earning updates, resulting in mispricing. The short-horizon earning surprise factor is motivated by the investor's inattention and evidence of short-term underreaction, which captures short-horizon mispricing [1]. The Post-Earning Announcement Drift Factor, or PEAD, reflects returns associated with mispricing that generally happen in a time horizon that is less than one year.

2.2. Quantitative Analysis

The overall explanation power is measured by R square and adjusted R square. The daily, annual, cumulative, and Sharpe ratios are calculated to measure performance and profitability. The strategy's risks will be estimated by evaluating volatility, higher moments, tail, and drawdown. Skewness and kurtosis are measured for return distribution. The return of the Shanghai composite index serves as a benchmark. The mean and standard deviation of the regression will be used to discover whether the model is predictable. Finally, we will conduct a t-test on return, FIN, and PEAD, respectively, to test their significance.

DOI: 10.54254/2754-1169/17/20231064

3. Data

3.1. Universe

The universe of our strategy includes all the stocks listed in Shanghai and Shenzhen stock markets. The number is according to the last trading day of last year. Then, when we use Tushare to get the stock price, we can get the tradable variety for each year.

The reason for our universe choice is that Shanghai and Shenzhen stock markets are China's largest capital markets, with a wealth of trading varieties. There are nearly 4,893 stocks that can be traded since 1990. In addition, the Shanghai and Shenzhen stock markets have become more mature in recent years. Many overseas index companies are selecting stocks on the Shanghai and Shenzhen stock markets to build index portfolios, such as the MSCI A50 index and the FTSE China A50 index. Finally, stock index futures in Shanghai and Shenzhen markets are becoming more and more abundant to help investors hedge.

3.2. Data Sets

a. ROE: We get ROE to calculate one-year net share issuance (NSI) and five-year composite share issuance (CSI). Then we can calculate FIN, which is the average of one-year net-share-issuance and five-year composite-share-issuance in the log.

b. Financial Disclosure date: We get this data to calculate PEAD, which is the cumulative excess rate of return relative to the market, measured by taking a cumulative excess rate of return during two trading days before the disclosure date of the latest financial report of the listed company to one trading day after disclosure.

c. Shanghai Composite Index: We use the Shanghai Composite Index as a benchmark to calculate the information ratio, beta, and other statistics to judge the strategy's effectiveness. If hedging is needed, we need to collect the price of Shanghai 50 stock index futures (IH) or CSI 300 stock index futures (IF).

d. Adjusted daily return: We use the stock price to calculate the return, construct the portfolio and rebalance it.

e. Risk-free return: SHIBOR rate

f. Repurchase scale and issuance scale: We use the scale of repo and issuance to calculate NSI.

3.3. Data Sources

a. China Stock Market & Accounting Research Database (CSMAR)

b. CSMAR

c. Tushare in Python

d. Tushare in Python

e. CSMAR

3.4. Data Range

1/1/2008 – 6/30/2018

IS: 1/1/2008 – 12/31/2018

We chose this decade because we believe it contains bull and bear markets and presents a variety of market conditions. These include the GFC, the 2015 stock market crash in Shanghai and Shenzhen, and the 2009 and 2014 bull markets. This data will be used to optimize the parameters.

OOS: 1/1/2019 – 6/30/2022

DOI: 10.54254/2754-1169/17/20231064

The data is new enough that it is close to the current state of the market. This ensures the timeliness of our research, and this data will be used to evaluate the profitability of the optimized model.

3.5. Strategy Detail

3.5.1. Signal Generation

The general idea of our strategy is that motivated by the DHS model (2020), we try to long assets with significant negative epsilon and short the assets with significant positive epsilon. The reasons are as follows:

a. Daniel-Hirshleifer-Sun multiple-factor model is a sufficient asset pricing model where PEAD and FIN factors can generate excess returns that other models cannot explain.

b. The three-factor model proposed by DHS can explain most factors in other models.

c. Generally, Daniel-Hirshleifer-Sun multiple-factor model serves as an advanced asset pricing model.

We construct the financing factor (FIN) based on the 1-year net share issuance (NSI) and 5-year composite share issuance (CSI). We assign firms to one of the two size groups (small "S" and big "B") based on whether that firm's market equity is below or above the Shang Hai and Shen Zhen exchange market median size breakpoint. Independently, we sort firms into one of the three financing groups (low "L," middle "M," or high "H") based on 1-year NSI and 5-year CSI, respectively. The three financing groups are created based on an NSI and CSI rankings index. Specifically, we first sort firms into three CSI groups (low, middle, or high) using 20% and 80% breakpoints. Six portfolios (SL1, SM1, SH1, BL1, BM1, and BH1) are formed based on the intersections of size and financing groups; each year, the value-weighted portfolio is constructed. The FIN factor return is calculated as the average daily return of the low financing portfolios (SL1 and BL1) minus the average daily return of the high financing portfolios (SH1 and BH1), that is,

$$FIN = \frac{r_{SL1}+r_{BL1}}{2} - \frac{r_{SH1}+r_{BH1}}{2} \tag{1}$$

FIN factor return is calculated daily.

PEAD is the post-earnings announcement drift factor intended to capture limited investor attention. It is again constructed in the fashion of Fama and French (1993). Following Chan, Jegadeesh, and Lakonishok (1996), the factor is measured as the 4-day cumulative abnormal return (CAR) around the most recent quarterly earnings announcement date, which is the return of 2 days before and one day after[1]. We first assign firms to one of two size groups (small “S” or big “B”) based on whether that firm’s market equity at the end of the month is below or above the Shanghai and Shenzhen exchange market median size breakpoint. Each stock is also independently sorted into one of three earnings surprise groups (low “L,” middle “M,” or high “H”) based on Cari at the end of month t − 1, using 20% and 80% breakpoints. Six portfolios (SL2, SM2, SH2, BL2, BM2, and BH2) are formed based on the intersections of the two groups, and a value-weighted portfolio is constructed each month. The PEAD factor return is then the average daily return of the high earnings surprise portfolios (SH2 and BH2) minus the average daily return of the low earnings surprise portfolios (SL2 and BL2), that is,

$$PEAD = \frac{r_{SH2}+r_{BH2}}{2} - \frac{r_{SL2}+r_{BL2}}{2} \tag{2}$$

PEAD factor is calculated daily.

DOI: 10.54254/2754-1169/17/20231064

We use 30-day rolling regression to regress the return of each asset on market factors MKT, FIN, and PEAD. The betas constantly generated from the previous 30 days are considered as the parameters on the current day. Then expected return is calculated as

$$E\left[R_{t,i}\right] = \beta_{[t-30,t-1],i} MKT_i + s_{[t-30,t-1],i} FIN_i + h_{[t-30,t-1],i} PEAD_i + \alpha_{[t-30,t-1],i} \tag{3}$$

The Epsilon is calculated as return minus expected return. If the Epsilon is negatively significant, we long. If Epsilon is positively significant, we are short. After regression of each asset on the three factors, we will get the formulas for each asset on each factor, where β, s, and h capture security's sensitivity to these three factors:

$$R_{t.i} = E[R_{t.i}] + e_{t.i} \tag{4}$$

($R_{t,\,i}$ = real return of each asset in the market, $e_{t,\,i}$ = error terms)

The asset's abnormal return, which indicates the difference between the expected return from the DHS model and real return in the market, should equal 0 under the hypothesized DHS pricing model if the asset is correctly priced. Otherwise, there are opportunities for profits from mispricing in the market. We use the average epsilon five days before trading day t as our signal and divide assets into five groups equally. (1-smallest epsilon, 5-biggest epsilon). We will short the assets in group five with significant positive epsilon and long the assets in group one with significant negative epsilon.

$$Signal = (e_{t-5,i} + e_{t-4,i} + e_{t-3,i} + e_{t-2,i} + e_{t-1,i})/5 \tag{5}$$

3.5.2. Portfolio Construction

Timing: According to Daniel, Hirshleifer, and Sun (2020), long-term mispricing will take a few years to correct, while short-term will take a few quarters [6]. Nevertheless, the return of each stock changes daily. The MKT, FIN, and PEAD are generated based on the daily return. Our regression and signal will thus be influenced and change daily. Therefore, we will rebalance our portfolio daily to capture the mispricing accurately.

Sizing: To increase the diversification and maximize the returns from two factors, we will long the last 20% of stocks with the slightest signal. We will trade the equities with signals calculated by epsilon in Shanghai and Shenzhen markets and use signal weighting to construct our portfolios as the signal indicating the extent of mispricing, thus leading to potential profits. We need to subtract each signal with its median to generate the signal used for portfolio construction.

Hedging & Money management: As all our assets in the portfolio are in Shanghai and Shenzhen stock markets, we need to consider market risks. To protect against downside risk in the event of any external disruptions, it would be better to short CSI 300 index future to avoid market risks.

3.5.3. Trade Execution

The Transaction Costs of the DHS strategy will include the bid-ask spread, possible slippage, latency costs, and a variety of fee costs.

Cai (2004) estimated that the Bid-Ask Spread for Shanghai A averages approximately 0.031 yuan, around 0.269 percent, while the bid-ask spread for Shenzhen A averages approximately 0.026 yuan, which is about 0.263 percent [6]. The bid-ask spread for the Shanghai A stock market is taken from an average of 492 Shanghai A stocks, while the bid-ask spread for the Shenzhen A stocks is taken from an average of 327 stocks. The percentage spread is calculated as [100*(Askit-Bidit)/Midit], where Midit is the midpoint of the bid (Bidit) and ask (Askit) quotes at the time of the transaction.

The borrowing fee for shorting stocks ranges from 0.3 to 3 percent. Therefore, we use 0.031 yuan as the bid-ask spread and 0.3 percent as the borrowing fee.

Slippage is also a possible transaction cost when investing in a low-cap firm since the firms listed in the Chinese stock market are generally lower in capital than the American stock markets. Latency costs will also be involved for the assets that got their signals from the short-horizon PEAD, especially for firms with low liquidity. As our trading volume is relatively low, we assume the slippage and latency cost to be 0 in the following implementation.

The Fee Costs of the order will involve the Chinese mainland taxes, Shanghai or Shenzhen exchange venue fees, and broker fees. The tax on capital gains is generally taxed at 20 percent. However, capital gains derived from Shanghai or Shenzhen stock exchanges are entitled to a 50 percent or even 100 percent tax reduction depending on the holding period. The venue fee for the Shanghai Exchange for A shares is 0.0049 percent of the trading value.

3.6. Development

3.6.1. Implementation

PnL Time Series.

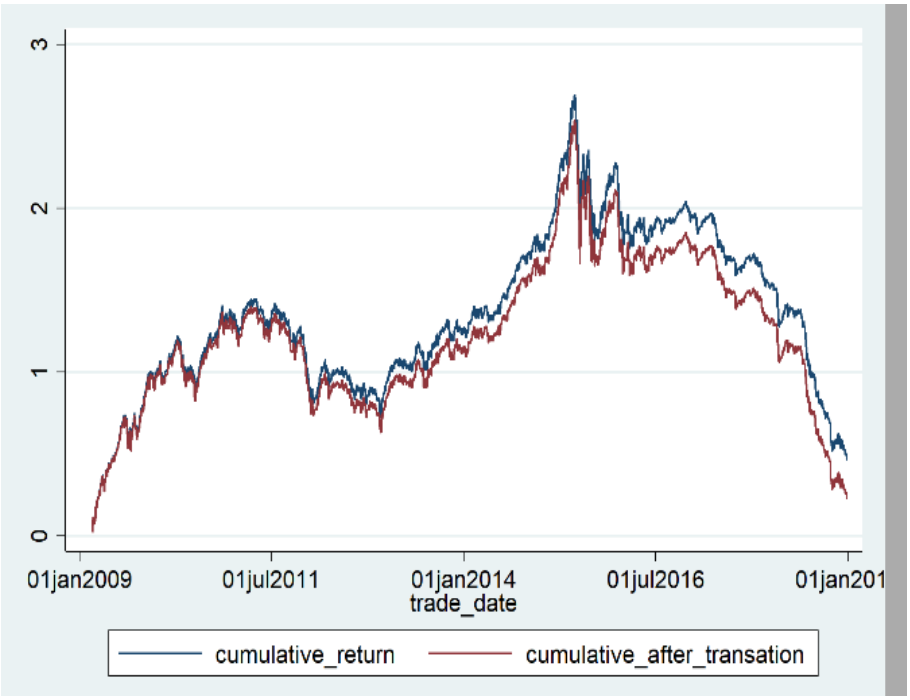

Figure 1: The cumulative return and cumulative return after transaction fee of the portfolio from 2009 to 2018.

DOI: 10.54254/2754-1169/17/20231064

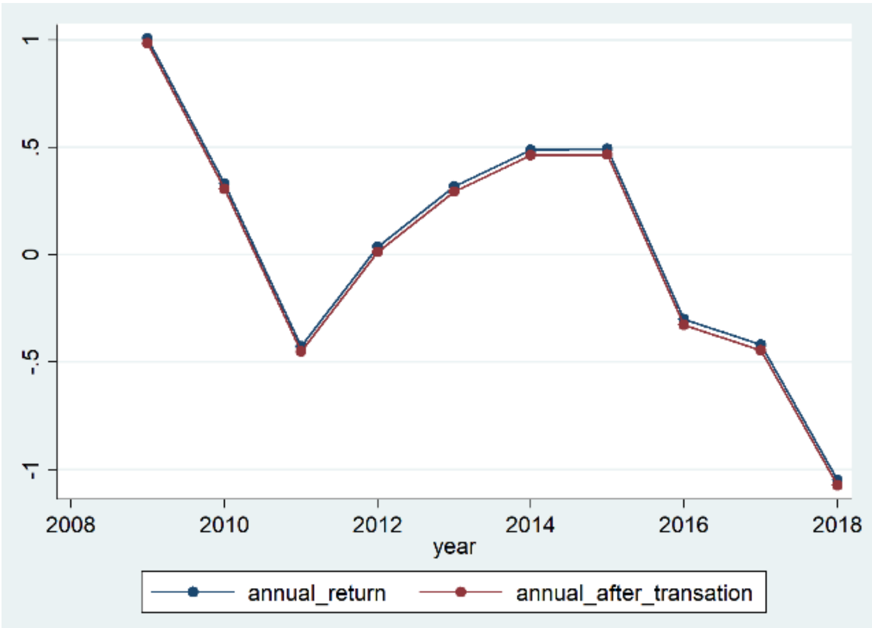

Figure 2: The annual return and annual return after transaction fee of the portfolio from 2009 to 2018.

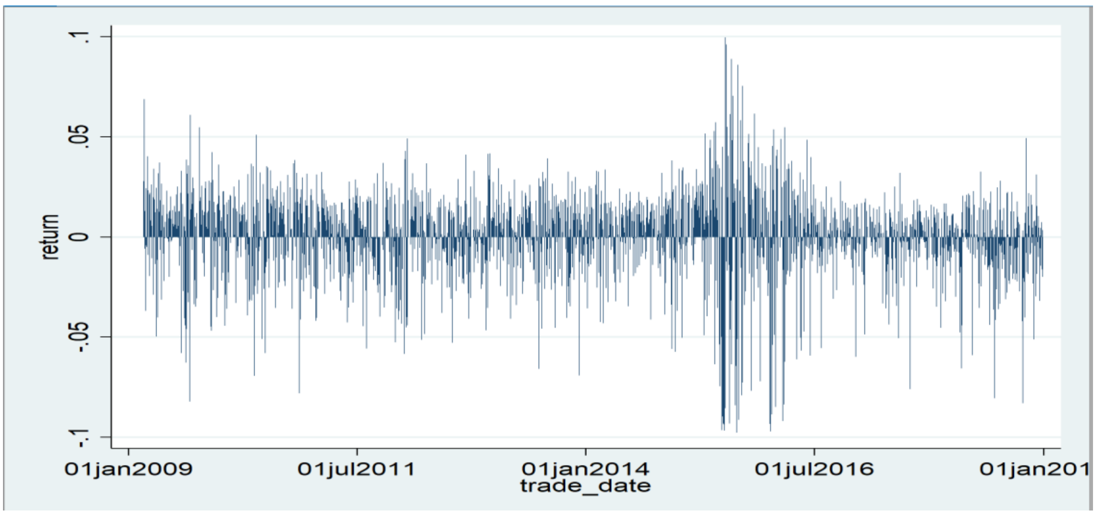

Figure 3: The portfolio returns of the daily bar chart from 2009 to 2018.

By the backtest, the cumulative return swells from 2009 to 2011 and slightly drops from 2011 to 2012. Then it began a steady climb toward its peak in 2015. A decline followed from 2016 to 2018. The bar chart of daily return clearly shows that the daily return fluctuates wildly and the extreme value achieved during 2015-2016. From the diagram of annual returns, it can be safely concluded that there were significant positive annual returns in 2009, 2014, and 2015.

The reason for large profits lies in the financial crisis in 2008 and the Chinese stock market turbulence that occurred from 2015 to 2016 with the stock market bubble popping. Enthusiastic individual investors created the stock market bubble by making massive stock purchases and borrowing funds faster than the economic development and returns of the firms they bought. As stated by Daniel, Hirshleifer, and Sun (2020), the model supplements the market factor with behavioral factors, acquiring mispricing with psychological biases, which is likely to be affected by overconfidence and limited attention [6]. The psychology of investors during these periods is especially unstable, leading to more

DOI: 10.54254/2754-1169/17/20231064

significant loadings on psychological factors and higher price fluctuations from overconfidence and limited attention. Consequently, our strategy is more profitable during these periods.

Table 1: The return summary of the portfolio.

RETURN SUMMARY (1)							
MEAN	Std. Dev.	Min	Max	Skewness	Kurtosis	Sharpe ratio	Annual-ized return
0.0001968	0.0214415	-0.0976847	0.099527	-0.8270534	6.64368	0.01964	0.0306

Table 2: The daily return summary of the portfolio.

Daily Return Summary (2)		
	Percentiles	Smallest
1%	-0.0767627	-0.0976847
5%	-0.0368519	-0.097151
10%	-0.0245652	-0.0967552
25%	-0.0087182	-0.0964489
50%	0.0024536	
75%	0.0125372	0.0888452
90%	0.0218617	0.0961361
95%	0.0285914	0.0989118
99%	0.0492646	0.099527

The distribution of return has a positive mean of around 0.02%. The skewness and kurtosis show that the return data are skewed left and have a "heavy-tailed" distribution. It could also be concluded from the return summary table (2), showing a median return of 0.24536%. The standard deviation in the table is the daily return standard deviation, and the annualized return standard deviation is 0.34037. We also annualized the cumulative return to calcite the Sharpe ratio.

```
. ttest return=0

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
  return |   3,280    .0000744    .0003549    .0203274   -.0006215    .0007704
------------------------------------------------------------------------------
    mean = mean(return)                                           t =   0.2098
Ho: mean = 0                                     degrees of freedom =     3279

    Ha: mean < 0                 Ha: mean != 0                 Ha: mean > 0
 Pr(T < t) = 0.5831         Pr(|T| > |t|) = 0.8339          Pr(T > t) = 0.4169
```

Figure 4: The t-test conducted on daily return.

According to the t-test conducted on daily return, there is not enough evidence to reject the hypothesis that the daily return differs from zero. Moreover, our portfolio obtains a Sharpe ratio of merely 0.0196. All the statistics show that the profitability of this strategy is limited.

DOI: 10.54254/2754-1169/17/20231064

Table 3: The regression summary of the portfolio.

Regression Summary (1)					
	Obs	**Mean**	**Std. Dev.**	**Min**	**Max**
Reg_r^2	9,036,809	0.4097615	0.2128085	7.77e-06	0.9816026
Adj_reg_ r^2	9,036,809	0.3416571	0.2373633	-0.1153759	0.9794798

Table 4: The regression summary of the portfolio in two factors, FIN and PEAD.

Variable	Mean	Std. Err.	Std. Dev.	95% confidence interval	t value
FIN	0.0002155	0.000055	0.0031399	[0.000108, 0.000323]	3.9305
PEAD	0.0014338	0.00011	0.006297	[0.001218, 0.001649]	13.0396

Based on the regression and t-test of two factors, the DHS model has a relatively reasonable explanation of stock return as the mean of r^2 is around 0.4 without more specific factors such as industry or company performance. Some stocks even got an r^2 of 0.98, which is a perfect return projection. Besides, the t-test summary indicates that both FIN and PEAD factors can generate a significant return in the Chinese market, which indicates a promising direction for the further study of the DHS model.

```
. ttest FIN=0

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
     FIN |   3,280    .0002155    .0000548    .0031399     .000108     .000323
------------------------------------------------------------------------------
    mean = mean(FIN)                                              t =   3.9305
Ho: mean = 0                                     degrees of freedom =     3279

    Ha: mean < 0                 Ha: mean != 0                 Ha: mean > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0001          Pr(T > t) = 0.0000

. ttest PEAD=0

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
    PEAD |   3,280    .0014338      .00011    .0062973    .0012182    .0016494
------------------------------------------------------------------------------
    mean = mean(PEAD)                                             t =  13.0396
Ho: mean = 0                                     degrees of freedom =     3279

    Ha: mean < 0                 Ha: mean != 0                 Ha: mean > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

Figure 5: The t-test of two factors, FIN and PEAD.

3.7. Difference from Estimations

Our strategy acquired a positive cumulative return from 2008 to 2018 and highly high returns during periods when investors are emotionally unstable. Thus, psychological factors influence more on stock returns more, and more opportunities are created. The previous strategy was tested in the US market and had a quiet, pleasant result in explaining the stock returns by Daniel, Hirshleifer, and Sun [7]. However, the market situation in the US and different from that in China. Our strategy has not been tested on the Chinese market, and no specific expected return data could be found. Therefore, we use another multi-factor stock-choosing model as a reference. As examined by Hu, Chen, Shao, and Wang (2019) tested the Fama-French model in the Chinese market, using data from 1995 to 2016 and a similar trading approach except for various factors [3]. Their portfolio got an average return of 1.23%. Our portfolio obtains an average annual return of 3.06%.

According to Guo, Kun, Sun, and Qian, Investor sentiments shown by traditional exchange indicator, investigation data and internet data may be essential in stock market investment [8]. However, the lack of a short-selling mechanism, the disposal effect, and the hype preference of individual stock investors for good news may be additional considerations when trading on the DHS strategy [9].

4. Refinement

We believe that the strategy can be restored to good performance in recent times by adjusting the parameters' size and using options to hedge.

First off, we will change the window of the composite share issuance (CSI) to calculate FIN. We are now constructing the FIN based on the 1-year net share issuance (NSI) and 5-year composite share issuance (CSI). However, we do not think 5-year composite share issuance is sensitive to the market. Five years are too long for FIN to respond well to the market, so we will narrow the window for CSI to calculate the FIN. We will use 2-year CSI instead of 5-year CSI because we think 2-year CSI can better reflect the market than 5-year CSI.

4.1. Implementation

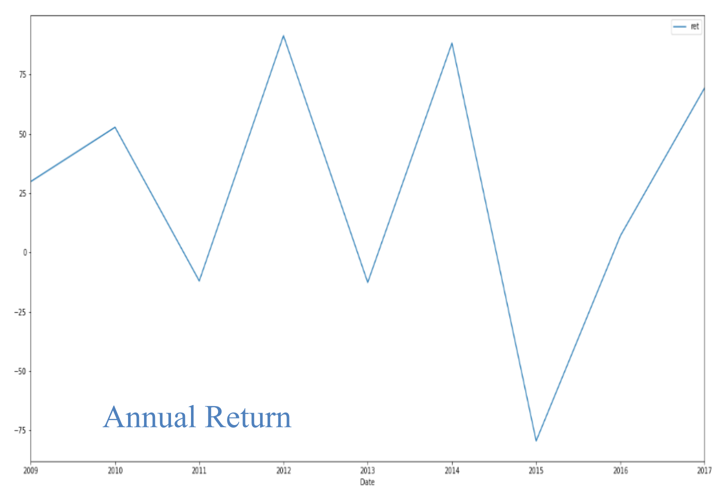

Figure 6: The annual return of the refined portfolio.

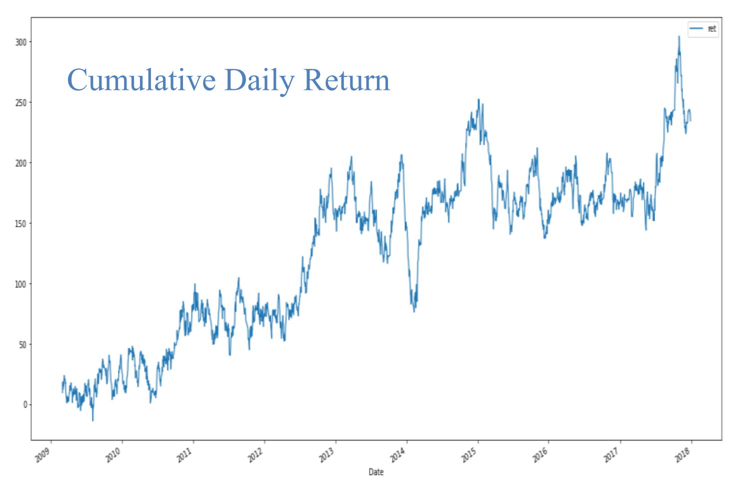

Figure 7: The cumulative daily return of the refined portfolio.

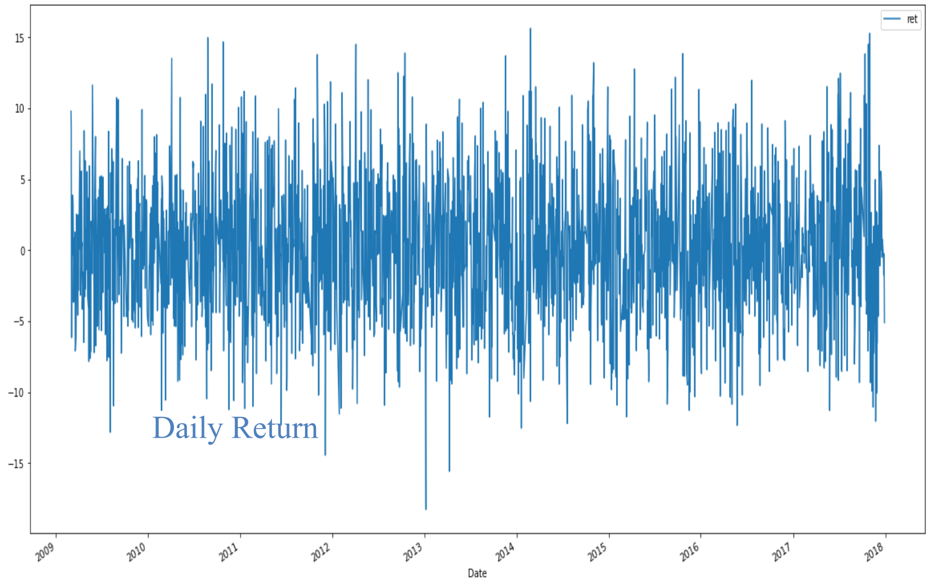

Figure 8: The return in the daily bar chart of the refined portfolio.

The above chart results from the back-test after using a 2-year CSI instead of a 5-year CSI. P.4 is the annual return. P.5 is the cumulative daily return. P.6 is a daily return. All numbers are presented in percentages.

By the backtest, the modified strategy has generally performed well, generating higher returns over ten years than the original strategy. However, there was still a massive pullback in the crash of 2015.

Table5: The return summary of the refined portfolio.

Return Summary						
Mean	**Std. Dev.**	**Min**	**Max**	**Skewness**	**Kurtosis**	**Sharpe ratio**
0.00108958	0.0494985	-0.182721	0.156378	0.109109	-0.04741	1.14969

The distribution of return has a positive mean of around 0.1089%. By the skewness and kurtosis, it can be found that the return data are skewed right and platykurtic. Sharpe's ratio also exceeds the original model. This shows that the strategy is more effective in the Shanghai and Shenzhen stock markets after modification.

4.2. Conclusion

4.2.1. Final Selection

Comparing the modified model with the original model, we find that the modified model performs better and can obtain more significant profits. 1.1 points optimized the Sharpe ratio. Therefore, we choose to use the modified model whose window of CSI is two years for an out-of-sample back-test.

4.2.2. Out-of-sample Back-test

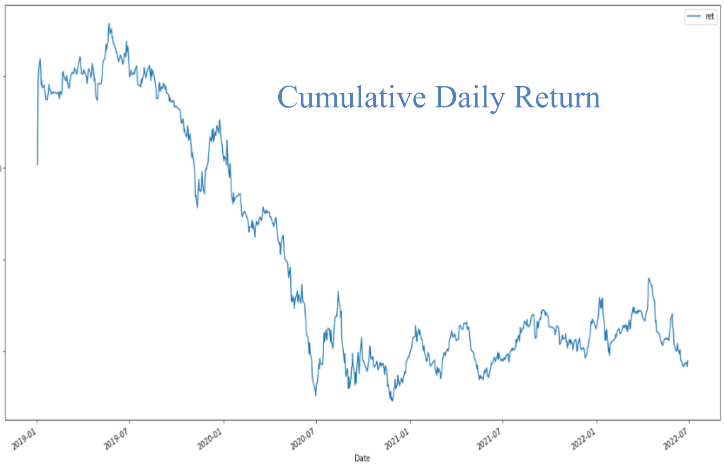

Figure 9: The annual return of the modified portfolio (out of sample).

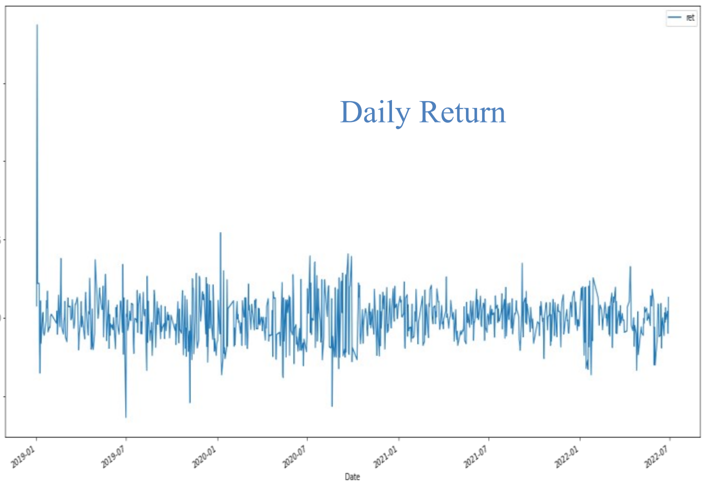

Figure 10: The return in the daily bar chart of the modified portfolio (out of sample).

The above chart results from the backtest after using 2-year CSI instead of 5-year CSI. P.9 is the annual return. P.10 is the cumulative daily return. P.11 is a daily return. All numbers are presented in percentages.

By the backtest, this strategy has not been very effective in the near term, and it has produced an enormous pullback. The global economy declined in 2019 and 2020 due to the COVID-19 pandemic. However, from April 2020, the Shanghai Composite Index showed a clear upward trend. Instead of making a profit, our revised strategy caused a drawdown of 50% in 2020. Our earnings were insignificant in 2021 and the first half of 2022. The portfolio ended up with a return of -42%. This shows that the modified model is invalid in the near term and cannot explain the market well and price each stock well.

Table 6: The return summary of the modified portfolio.

Return Summary						
Mean	**Std. Dev.**	**Min**	**Max**	**Skewness**	**Kurtosis**	**Sharpe ratio**
-0.0004977	0.0115383	-0.066377	0.186994	2.018948	25.9413	-0.6846

The distribution of return has a negative mean of around -0.04977%. By the skewness and kurtosis, it can be found that the return data are skewed right and leptokurtic. The Sharpe ratio is negative. Although no low-frequency high losses exist, continued losses make the strategy perform poorly.

4.3. Additional Considerations

We think changing the FIN window again will make the strategy more effective. In the back-test of our strategy, the in-sample data performed well, while the out-of-sample data performed exceptionally poorly, which may be due to overfitting. We think we can change FIN's windows to 3-year CSI or 4-year CSI so that we may avoid overfitting and get a more effective strategy.

In addition, adding a few common factors may improve the model pricing ability, making the strategy more profitable. This model's factors differ from other multi-factor models like the Fama-French three-factor model because the Fama-French three factors model can well explain the expected return rate of most Shanghai and Shenzhen stocks [10]. The factors of DHS model consider investor psychology for the first time. While this may differ from traditional pricing theories, we suggest considering size or value effects to make the model more accurate.

4.4. Trading Recommendation

We strongly recommend that a trader not use this strategy without other indicators, especially in the bear market. Although DHS is a new way of pricing and stock selection, and the main trading idea is the same as arbitrage trading using Fama French's three-factor model, according to the performance of back-test, this strategy has not performed well in recent years. We suggest that if this pricing model is used for stock selection, traders should use a small window to calculate FIN or add common factors such as size or value to the model to improve its pricing ability.

With some understanding of the multi-factor stock-picking type of strategy, we can then look at this strategy. Our strategy is complemented at its base by two behavioral factors that complement the CAPM model. These two factors describe the two psychological biases of overconfidence and limited attention span of traders, respectively. If you share the same opinion about the impact of these two psychological biases on the market, then you can try this strategy in the market. However, our long-term factors found that this strategy did not respond as well as expected to the market, and its returns were not as good as we expected. This strategy usually performs better when there is panic, excitement, and other excesses among traders in the market. The historical back-test results show that this strategy performed much better in 2008, 2009, 2014, and 2015. Overall, we recommend this strategy to capture the psychological bias of traders in good market conditions to gain excess returns.

The basis of our strategy is still a multi-factor stock strategy, so if you have different ideas, you can improve the strategy by making changes to the factors on top of that strategy. Likewise, you can also optimize the strategy by changing the calculation period of the strategy's trading cycle factor, etc. to upgrade the strategy.

5. Summary

Our strategy complements the market factors of the CAPM model with two behavioral factors proposed by the DHS model and uses psychological bias to capture mispricing as a trade-off for mispricing and to correct for the differential returns among mispricing. First, we characterize the financing factor and the drift factor after short-term earnings announcements by FIN and PEAD, respectively. After this, we regress the three factors, FIN, PEAD, and the market factor. The data period of the regression is chosen to be one month, and the resulting equation is then compared to the actual market price and predicted to have possible future spreads, gaining the benefit of correcting mispricing by rebalancing the portfolio daily. However, the strategy can also have significant limitations under different market mechanisms; for example, the lack of a shorting mechanism and other policy restrictions in the Chinese market may result in some gains not being achieved. Also, due to the strong correlation of the strategy to investor psychological factors, it may be affected by the speculation of individual stock investors.

References

[1] *Daniel, Kent, David Hirshleifer, and Lin Sun. 2019. "Short- And Long-Horizon Behavioral Factors". The Review Of Financial Studies 33 (4): 1673-1736. doi:10.1093/rfs/hhz069.*

[2] *Belen Blanco (2012). Using CAPM and Fama and French Three Factor Model: portfolios selection. Public and Municipal Finance, 1(2).*

[3] *Grace Xing Hu, Can Chen, Yuan Shao, and Jiang Wang, 2019. "Fama–French in China: Size and Value Factors in Chinese Stock Returns," International Review of Finance, International Review of Finance Ltd., vol. 19(1), pages 3-44, March.*

[4] *Lian, Xiangbin and Liu, Yangyi and Shi, Chuan, A Composite Four-Factor Model in China (September 21, 2021). Available at SSRN: https://ssrn.com/abstract=3928587 or http://dx.doi.org/10.2139/ssrn.3928587.*

[5] *Daniel, Kent D. and Hirshleifer, David A. and Sun, Lin, The FIN and PEAD Factors: Motivation, Construction, and Availability (September 11, 2020). Available at SSRN: https://ssrn.com/abstract=3691023 or http://dx.doi.org/10.2139/ssrn.3691023.*

[6] *Cai Jun, 2004. "Bid-Ask Spreads for Trading Chinese Stocks Listed on Domestic and International Exchanges". City University of Hong Kong. Available at SSRN: Https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.459.5398&rep=rep1&type=pdf.*

[7] *Daniel, K. D., Hirshleifer, D. A., & Sun, L. (2021). Teaching Slides on Short and Long Horizon Behavioral Factors. Columbia Business School Research Paper. Available at SSRN: https://ssrn.com/abstract=3849094 or http://dx.doi.org/10.2139/ssrn.3849094*

[8] *Guo, Kun, Yi Sun, and Xin Qian. 2017. "Can Investor Sentiment Be Used To Predict The Stock Price? Dynamic Analysis Based On China Stock Market". Physica A: Statistical Mechanics And Its Applications 469: 390-396. doi:10.1016/j.physa.2016.11.114.*

[9] *Wu, Yangru. 2010. "Momentum Trading, Mean Reversal And Overreaction In Chinese Stock Market". Review Of Quantitative Finance And Accounting 37 (3): 301-323. doi:10.1007/s11156-010-0206-z.*

[10] *Zhang, Lishuang. (2022) An Empirical Study on the Components of Shanghai Stock Exchange 50 Index in China Based on CAPM Model and Fama-French Three Factor Mode.Economic Research Guide , 26:78-80.*