The Profitability Premium:

The Effect of Arbitrage Activity on Abnormal Trading Profits

Trang Quynh Vu and Sanda Moldovan

Supervisor Francisco Santos

Master Thesis in Finance

NORWEGIAN SCHOOL OF ECONOMICS

PREFACE

This thesis is written as part of a Major in Finance, and marks the end of our two-year study at Norwegian School of Economics (NHH) and by that our Master of Science and Business Administration.

The process of finding an interesting subject was early determined as we found asset management and factor investing to be captivating, which motivated us to explore the topic further. This topic also enables us to advance the knowledge we have obtained through the different courses taken as a part of our major. The work with this thesis has been challenging but also a rewarding experience.

First and foremost, we would like to thank our advisor, Associate Professor Francisco Santos, for constructive feedback and professional guidance throughout the process. We also thank him for guiding us to find the inspiring topic, showing interest in our work and motivating us along the process. In addition, I would like to thank NHH IT Support for helping us with the software (SAS), which played an important role in this work.

Lastly, we are grateful to our families and friends for supporting us in this anchor leg in our studies at NHH.

Bergen, June 2019

Trang Quynh Vu — Moldovan S.

Trang Quynh Vu — Sanda Moldovan

ABSTRACT

Inspired by the Novy-Marx (2013) paper, the purpose of this thesis is to investigate the profitability premium and the effect of arbitrage activity on its abnormal trading profits over the period from June 1964 to December 2010. The primary contribution is to create a measure of arbitrage activity for the profitability strategy, which we dub coprofitability or CoPROF, mainly based on previous methodology of Lou and Polk (2013) and Huang, Lou and Polk (2016). This new measure is used to determine periods of relatively low and high arbitrage activity and evaluate whether trading in the strategy gets crowded. We show that during periods of low arbitrage activity, the majority of abnormal returns are statistically insignificant. In contrast, times of relatively high arbitrage activity are associated with positive and statistically significant abnormal returns as shown by Fama French 3 Factor model. Moreover, we do not find no indication of long run reversal and crash risk when the arbitrage activity is relatively high as opposed to the momentum and beta strategies.

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

Quality investing, which is designed to capture the documented excess returns of high-quality stocks over low-quality stocks, has become popular in recent years. Notwithstanding the lack of a precise definition, many quality-oriented investment strategies have been shown to deliver exceptional returns. As a result, a number of asset managers and investors aim quality investing as a complementary investment style and consider this strategy as a good hedge for value investing. This paper depicts an analysis of quality-based investment strategy, more specifically focused on the "profitability" anomaly: stocks with high profitability ratios tend to outperform on a risk-adjusted basis.

As is the case with other factors, profitability had been used for decades by practitioners such as Benjamin Graham and David Dodd. It has been recently documented by an expanding body of literature that profitability anomalies have significant power in predicting cross-section of average returns (Novy-Marx, 2013; Hou, Xue, and Zhang, 2015; Ball et al., 2015, 2016). These studies demonstrate that hedge strategies that are long in high profitability stocks and short in low profitability stocks deliver outstanding risk-adjusted abnormal returns and underline high Sharpe ratios and no crash risk (Lemperiere et al., 2015). Another evidence of the increasing interest in profitability as a substantial factor in equity investing was its extension in the widely known three-factor Fama French model, invented by Nobel laureate Eugene Fama and Kenneth French. Nowadays, the Fama French 5 Factor Model incorporates the profitability factor RMW (Robust Minus Weak) as part of its model regression equation.

Thus, in ***the first part of the thesis***, we implement an evaluation on whether gross profitability predicts the cross section of returns as documented in the previous research. We employ a Fama- MacBeth (1973) cross sectional regression on all individual stocks in the U.S stock market in the period 1964-2010 and follow the similar procedure from Fama French for portfolio sorting, based on the return- and accounting- data. In line with the Novy-Marx (2013) paper, the Fama and MacBeth regressions of returns on measures of profitability depict that gross profitability scaled by assets (GPA) appears to be the measure of basic profitability with the most power predicting the cross section of expected returns. It has roughly the same power as book-to-market regarding to t-statistics, and the results are even more significant after the industries are demeaned. In addition, we construct a univariate sort on GPA, which results in the abnormal returns of the profitable minus unprofitable return spread of *0.49%* per month, with a test-statistic of *4.2*, ignoring the effect of transactional costs.

Subsequently, we combine two strategies of value and quality investing, and test for the performance. Since the Spearman rank correlation between gross profits-to-assets and book-to-market ratios is -0.19 and highly significant, one would expect gross profitability to provide an excellent hedge for value. Thereby, we control for both factors (gross profits-to-assets and book-to-market) by constructing 25 portfolios using double sorts, all of which are rebalanced at the end of June. We find that profitability strategies could achieve significantly higher average returns once controlling for book-to-market (B-M) value and the performance of value strategies is enhanced while controlling for profitability. The average value spread across gross profits-to-assets is 0.65% per month and in every B-M portfolio is greater than the 0.38% per month spread on the unconditional value of strategy. Similarly, the average gross profit spread across B-M quintiles is equal to 0.51% per month and in every profitability portfolio exceeds the 0.28% per month spread on the unconditional profitability of strategy.

In fact, the predictability affiliated with profitability anomalies has developed a debate over whether these anomalies correspond to compensation for risk or systematic mispricing. If the profitability anomalies indicate systematic mispricing, arbitrageurs would exploit this opportunity and quickly eliminate the mispricing. In support of market efficiency under the Efficient Markets Hypothesis (EMH), profitable stock might generate higher return as it is fundamentally riskier, and thus investors demand a higher risk premium to holding these stocks. Nevertheless, while recognized risk premia strategies are veritably rewarding investors for carrying a substantial negative skewness risk (Harvey & Siddique, 2000), Landier (2015) documented that the returns of the arbitrage strategy on long position of the most profitable companies and short position of the less profitable ones yields a positive skewness, and very small tendency to crash. Thus, it is hard to account for this anomaly using a risk premium interpretation.

As a result, we detect that there is a gross profitability premium that arbitrageurs could capture, especially if they would be able to time the market correctly. However, it is complicated to measure the impact of arbitrageurs on stock prices due to the unknown composition of arbitrageurs in the market, unavailable information on the amount of arbitrage capital and other unobservable data such as leverage and short-selling activities (Lou and Polk, 2013). In an attempt to understand how investors could time the market while pursuing a profitability strategy, ***the second part of the paper*** introduces a new measure of arbitrage activity in profitability strategies, dubbed as CoPROF. In particular, we expand on the studies of Lou and Polk (2013) and Huang et al. (2016), who introduce similar measures for beta and momentum strategies. Our main contribution is the construction of the new measure while adjusting the methodology for CoPROF computation due to the limited availability of accounting data, which is published only once a year.

The methodology for constructing the measure of arbitrage activity in profitability strategies is based on the same premise as in the two above mentioned studies. We assume that when pursuing a profitability strategy, arbitrageurs tend to go long and short on a diversified portfolio of stocks at the same time, which results in *simultaneous* price impact on those stocks and hence leads to excess return comovement. Hence, we expect that the return correlation between profitable and unprofitable stocks should be high as these stocks are associated with periods of highest arbitrage activity.

We describe the construction of CoPROF variable and the relevant adjustments to the original methodology. To ensure that we are able to construct the measure of arbitrage activity in the profitability strategy correctly, we first replicate the CoBAR methodology and present our results in the Appendix relative to the Huang et al. (2016) paper. Afterwards, we modify the corresponding methodology by taking into account the unique characteristics of our dataset. In comparison to the CoBAR procedure from Huang et al. (2016) paper, we sort all stocks into deciles based on their GPA at the end of each June. These deciles will be kept the same for one year until the end of June next year. For every month, CoPROF is then computed as the average pairwise partial correlations using 2 months' worth of past daily returns rather than 52 prior weekly returns for all stocks in the *highest decile* whilst controlling for the Fama-French (1992) three factors. Due to limited availability of accounting information that does not allow for more frequent rebalancing of the portfolios than once a year, the composition of the portfolios stays the same throughout the year. Using 52 weekly returns to construct the CoPROF measure would lead to little variation in the underlying data and thus any significant movements in the CoPROF would not be revealed until the next rebalancing period when the composition of the portfolio would change. By using daily returns instead of weekly we induce enough variation in the underlying data to estimate the CoPROF measure more accurately as we move from one month to another.

In the next stage, we examine how our zero-cost portfolio, which goes long the value- weight portfolio of stocks in the highest gross-profits-to-assets (GPA) decile and short the value-weight portfolio of stocks in the lowest GPA decile, performs by analyzing the profitability and any potential long run reversal in the strategy returns relative to our CoPROF proxy. We do not find abnormal returns when arbitrage activity in profitability strategies is low as revealed by the Fama French 3 Factor and 5 Factor models. When the coprofitability is relatively high, then the abnormal returns are positive and statistically significant as shown by the Fama French 3 Factor model. When the 5 factors are used, the abnormal returns are positive, but statistically significant only in the first 6 months of the arbitrage strategy. Despite the difference in the magnitude of results, both models show that the profitability strategy does not become destabilizing, with no evidence of long run reversal and crash risk as opposed to the momentum and beta strategies. Overall, the results we obtained are consistent with the theory of price stabilization of arbitrageurs' activity.

Finally, in order to ensure the robustness of our main results, we design two series of tests. The first test runs the portfolio formation stage based on the CoPROF values conditional on stocks from decile 1 (i.e. unprofitable firms) and decile 5 (i.e. profitability neutral portfolio) and not from decile 10 as used in the main methodology. For decile 1, according to the Fama French 3 Factor model, the test shows that we obtain similar results when arbitrage activity is low, but large differences in the magnitudes of returns when the coprofitability is high. When the 5 factors are used, the profitability spread is similar to the one found in the main results. We do not find statistically significant abnormal returns, except for one holding period. For decile 5, surprisingly, the difference in the abnormal returns when coprofitability is high relative to when it is low is large and significant across all holding periods, except the second year. This result is unexpected because the trades in the profitability-neutral portfolio should not yield statistically significant returns when compared to the activity in the extreme deciles, in which the majority of the arbitrage activity takes place.

The second robustness test alters the original methodology of CoPROF construction by restricting the use of future information when forming the CoPROF quintiles. For example, for determining the CoPROF quintile breakpoints for the month of May of 2006, we would use only the sample up to this date, excluding all the months following May 2006. After getting the relevant breakpoints, we would assign the May 2006 date to the relevant quintile based on its CoPROF value at that point in time. This procedure is repeated for all the months in our dataset.

The test reveals that we do not obtain significantly different results (i.e. slightly smaller magnitude but nearly similar trend) as the ones presented in the main part.

The rest of the paper is organized as follows. Section 2 reviews the literature in this area to highlight our contributions. In section 3, we implement an evaluation of whether gross profitability explains the cross-section of expected returns and form portfolios based on univariate sorts on gross profitability, and bivariate sorts on gross profitability and book-to-market ratios. Section 4 outlines how we construct CoPROF, a measure of arbitrage activity in profitability strategies, as well as the procedure for generating profitability sorted portfolios. In section 5, we conduct a number of additional robustness tests to verify the sensitivity of our results. Finally, section 6 marks the conclusion of this study.

2 LITERATURE REVIEW

Our research is mainly motivated by two strands of literature. First of all, we present an increasing number of papers studying profitability as a priced factor in the cross-section of expected stock returns. In the second part, we identify literature that provides various ways of measuring the arbitrage activity, including return comovements, supporting our contribution on the CoPROF measure.

2.1 Profitability anomalies

Perhaps gross profitability is the version of a “quality” factor that is most conventional in academic circles. Many researchers have confirmed the existence of long-term profitability premium which is especially substantial. Taking the data for the US company´s stocks from July 1963 to July 2010, Novy-Marx (2013) illustrates that by univariate stock-sorting on gross profitability, the abnormal returns or alpha cannot be interpreted by other known violations of the CAPM such as the size, value and momentum anomalies. Furthermore, Novy-Marx (2013) finds that gross profitability performs relatively better than other quality strategies such as Graham’s quality, ROIC and earnings quality, especially among large-cap US stocks.

An increasing body of literature documents that profitability anomalies have significant power in cross section of stock returns (Novy-Marx, 2013; Fama and French, 2015; Hou, Xue, and Zhang, 2015; Ball et al., 2015, 2016). These studies show that hedge strategies that are long in high profitability stocks and short in low profitability stocks (hereafter long-short or hedge strategy) yield significant risk-adjusted abnormal returns. Furthermore, profitability anomalies subsume most of earnings-related anomalies (e.g., strategies based on price-to-earnings, or asset turn- over), a large number of seemingly unrelated anomalies (e.g., strategies based on default risk, or net stock issuance, failure probability, the distress risk, and free cash flow) and the accrual anomaly (Chen, Novy-Marx and Zhang, 2011; Novy-Marx, 2013; Ball et al., 2016; Wahal, 2019 and Linnainmaa and Roberts, 2018).

Novy-Marx (2013) argues that gross profitability, as measured in terms of gross profit scaled by total assets, is a cleaner measure of economic profit than alternative measures of profitability because it is unaffected by non-operating items, such as leverage and taxes. Consequently, he shows that a gross profitability measure predicts cross section of stock returns. Ball et al. (2015) argue that selling, general, and administrative (SG&A) expenses represent a significant proportion of business operations costs. To better match current expenses and revenues, they adjust gross profit by deducting SG&A expenses (excluding research and development expenditures), called as operating profit, and find that operating profitability (operating profit scaled by total assets) have a better explanatory power in cross section of stock returns than gross profitability. Finally, Ball, et al. (2016) find that an increase in operating profitability due to a non-cash earnings component (accruals) by Sloan (1996) has no relation with the cross section of stock returns. Thus, they exclude accruals from operating profit and find that the obtained cash-based operating profitability is a significant predictor of future stock performance that effectively subsumes the accrual anomaly (Dechow, 1994; Sloan, 1996).

Our thesis is inspired by Novy-Marx´s 2013 paper, which contributed with a new insight into the cross-section of stock returns. It investigated that a simpler quality measure, **gross profits-to-assets**[1], has as much power predicting stock returns as traditional value metrics (book-to-market value). Besides, Novy-Marx use Fama and MacBeth regressions of firms' returns on gross profits-to-assets, earnings- to-book equity, and free cash flow-to-book equity to prove that the gross profits-to-assets metric outperforms the two other accounting metrics. *In our paper, in line with the mentioned findings, we show that the strongest predictor of cross-section of returns is the gross profitability measure as compared to earnings to book equity and free cash flows to book equity measures.*

In addition to a strong profitability argument, Novy-Marx (2013) perceives a negative correlation between gross profitability and book-to-market, suggesting that a combination of these two factors will considerably enhance the performance of investment strategies relying on either factor. Indeed, strategies based on gross profitability are inherently growth strategies and strongly negatively correlated with strategies based on price signals, leading them exceptionally attractive to traditional value investors. These results have substantial influence on the design of both DFA's growth portfolios and AQR Capital Management's core equity funds.

It is important to note that Novy-Marx (2013) exploits incomparable profitability measure in comparison to the well-known Fama French five-factor model published by Fama and French (2015). The difference comes from the fact that Fama and French used **operating profitability** (OP) as a measure for a firm's profitability. More specifically, the profitability factor is

[1] Revenues minus cost of goods sold, scaled by book value of assets

calculated as the return of a well-diversified portfolio of stocks with high operating profitability minus the return of a well-diversified portfolio of stocks with low operating profitability, which is dubbed RMW (robust minus weak). In our thesis, we decide to focus on Novy-Marx (2013) paper using gross profits-to-assets as the profitability measure.

2.2 Measurement of Arbitrage activity

Regarding to Lou and Polk (2013) paper, arbitrage activity is extremely complicated to measure at any given point in time for various reasons such as unknown composition of arbitrageurs in financial markets; unavailable high-frequency data of institutional investors on capital under management; unobservable information regarding arbitrageurs' activities in leverage, short-selling, and derivatives; and changing liquidity of assets traded. In this section, we provide different ways to measure the arbitrage activity, thereby discussing its implications and effects arbitrageurs have on prices.

The first interpretation for stock returns is that **investor sentiment** affects asset prices (e.g. Baker & Wurgler, 2007; Berger & Turtle, 2012; Fong & Toh, 2014; Greenwood & Shleifer, 2014; Kim, Ryu, & Seo, 2014; Qian, 2014). Baker and Wurgler (2007) form six proxies for *market-wide investor sentiment*: the closed-end fund discount, NYSE share turnover, the number and average first-day returns on IPOs, the equity shares in new issues, and the dividend premium, and further empirically explain the cross-section of future stock returns. A later study indicates that a series of anomalies in financial market can be interpreted by market-wide sentiment. In addition to market-wide sentiment, Kumar and Lee (2006) use buy and sell trading records to form *retail investor sentiment*, and further conclude that systematic trading by retail investors could lead to stock return comovements.

Suggested by recent behavioral asset pricing papers, an alternative explanation for stock prices is that **crowded trades** affect asset prices (e.g. Stein, 2009; Pojarliev & Richard, 2011; Hanson & Sunderam, 2014; Menkveld, 2014; Barroso et al., 2018). Stein (2009) seeks to shed new light on the importance of crowded trades by conducting a formal study on the crowded-trade problem, and argues that investors push prices further away from fundamentals when an unexpectedly significant number of arbitrageurs pursue similar unanchored strategies. Then, the role of crowded trades on asset prices empirically become a prominent topic in a range of literatures. Pojarliev and Richard (2011) follow the approach used in PL (2008a) and measure carry-, trend- and value crowdedness to prove that "crowded trades harbour potential risk once a change in fundamentals or sentiment induces liquidation of positions". Also, built upon the Stein (2009) paper, Barroso et al. (2018) develop a model focused more specifically on momentum and investigate the role of non-linear concave beliefs. Using quarterly holdings of 13F institutions in the period from 1980 to 2015, the study constructs several proxies for momentum capital based on aggregate momentum trading.

The last explanation for stock returns is relied on the ideas of **return comovements**. It had been stated in the traditional theory based on frictionless markets that stock return comovement should be reflected by correlations in news about the fundamental value of securities. Nevertheless, empirical literature argues that stock return comovement is only partially explained by the comovement of firms' fundamentals and that systematic noise trading is a reasonable alternative explanation (Morck, Yeung, and Yu 2000). As a critical driving force of systematic noise trading, Barberis and Shleifer (2003) and Barberis, Shleifer, and Wurgler (2005) suggest that institutional aspects may play a significant role in the movement of stocks' discount rates, leading returns to comove above and beyond that indicated by their fundamentals. *We adopt a methodology to construct a measure of arbitrage activity that was originally published by Lou and Polk (2013) for momentum strategy, and later applied by Huang et al. (2016) for use in beta-strategy*.

In the specific context of momentum, Lou and Polk (2013) argue that crowding by momentum investors potentially accounts for negative skewness in momentum returns. Inspired by the comovement of stock prices presenting specific characteristics showed by Barberis & Shleifer (2003), the paper sheds new light on these issues by proposing a novel approach to measuring intensity of arbitrage activity based on high-frequency excess return comovement, which occurs when arbitrageurs tend to buy or sell a diversified portfolio of stocks at the same time. They dubbed their measure Comomentum (in this paper referred as Comom), which is defined as the average pairwise correlation of daily/weekly Fama-French (1992) three-factor residuals for winner/loser decile stocks in the ranking period. More specifically, the price correlation among momentum stocks is high during the periods of high arbitrage activity in the momentum strategy and low when there is little activity in corresponding strategy.

Huang et al. (2016) extended their previous research (Lou & Polk (2013)) to develop the measure CoBAR for the excess comovement of stocks involved in beta arbitrage, which exploits the low-beta premium suggested by Frazzini and Pedersen (2014). The proxy is a measure of arbitrage activity in the beta-strategy which are long the value-weighted lowest beta stock decile and short the value-weighted highest one. Their results suggest that beta arbitrage activity can have impact on the returns of the beta strategy. In fact, it is illustrated that the abnormal returns become negative for very high levels of arbitrage in the market.

In this paper, we are going to construct a measure of arbitrage activity by exploiting the comovement in stock returns of profitability anomaly. Overall, we find that the arbitrage activity pursuing a profitability strategy does not become destabilizing as there is no evidence of long run reversal and negative skewness for the abnormal returns. This means that the arbitrageurs' activity on stock prices has a stabilizing effect.

3 EMPIRICAL ANALYSIS ON GROSS PROFITABILITY

3.1 Data & Methodology

Following the approach of Novy-Marx (2013), the data used in this paper consist of monthly stock returns obtained from the Center for Research in Security Prices (CRSP) from July 1964 and December 2010. From CRSP, we also obtain the SIC codes of each firm for industry categorization. The accounting variables are retrieved from COMPUSTAT database. Finally, the monthly risk-free rate, market risk premium, and the Fama French factors are taken from Kenneth French's website.

Our dataset starts with all stocks traded on the three major U.S. exchanges (i.e. NYSE, Amex and Nasdaq) and it does not contain securities other than ordinary shares. To reduce survivorship bias, we impose for accounting data for a specific firm to be available on COMPUSTAT for at least two years (Fama & French, 1993). Further, to guarantee that we do not use any future accounting variables to explain stock returns, we match accounting data for all fiscal year ends in calendar year *t-1* with returns for July of year *t* to June of year *t+1* (Fama & French, 1992). The final dataset includes firms with non-missing market value of equity, book-to-market ratio, gross profit, book value of total assets, current month returns, and returns for the previous one-year period. We also do not include financial firms (i.e. those with a one-digit Standard Industrial Classification (SIC) code of six) in our sample.

The first measure of profitability used in the analysis is the gross profit to the book value of assets. We calculate gross profit as total revenue (Compustat item REVT) minus the cost of goods sold (Compustat item COGS). The second measure of profitability is earnings before extraordinary items (Compustat item IB) to the book value of equity. The final proxy for profitability is the free cash flow to the book value of equity. Free cash flow is calculated as net income (Compustat item NI), plus depreciation and amortization (Compustat item DP), minus changes in working capital (Compustat item WCAPCH), minus capital expenditures (Compustat item CAPX). We deflate all profitability measures by a book-based measure rather than by a market-based measure to separate the productivity proxy from book-to-market (Novy-Marx, 2013).

We compute firm size and book-to-market ratio following the same procedure as in Fama & French (1992,1993). In particular, we use a firm's market equity in June of year *t* to measure its size from July of year *t* to June of year *t+1*. For the computation of the book-to-market ratio from July of year *t* to June of year *t+1*, we use a firm's market equity at the end of December of year *t-1* and the book value for fiscal year end in calendar year *t-1*. The book value of equity is the shareholder's equity (Compustat item SEQ if available, or CEQ+PSTX if available, or AT-LT), plus balance sheet deferred taxes and investments tax credits (Compustat item TXDITC if available, or TXDB+ITCB), minus preferred stock. In cases when the values of deferred taxes and investment tax credits are missing, we set them to zero. To compute the value of preferred stock, we set it equal to the redemption value (Compustat item PSTKR) if available, or else to the liquidating value (Compustat item PSTKRL) if available, or else to the carrying value (Compustat item PSTK). We do not include firms with negative book value of equity. For Fama-MacBeth regressions, we transform size and book-to-market variables by taking the natural logarithm.

3.2 Gross profitability and the cross section of expected returns

3.2.1 Fama and MacBeth Regressions: Estimation Details

With regards to a cross-sectional asset pricing study, in order to examine if some factors can explain asset returns, one could use univariate and double-sorting portfolio techniques. Nonetheless, Grinblatt and Han (2005) argues that these techniques cannot markedly control for other variables that might influence returns, and sorting on three or more variables is impractical. Hence, to allow us to smoothly control for additional variables, Fama and MacBeth (1973) cross-sectional regression appears as a practical way of testing how these factors describe portfolio or asset returns.

In the first step, for each portfolio P, we run regression against one or more factor time series to obtain beta estimates, and to determine how exposed it is to each one (the "factor exposures"):

$$R_{P,t} - R_{f,t} = \widehat{\alpha_P} + \widehat{\beta_{1,P}}(F_{1,t} - R_{f,t}) + \widehat{\beta_{2,P}}(F_{2,t} - R_{f,t}) + \cdots + \widehat{\beta_{m,P}}(F_{m,t} - R_{f,t})e_{P,t}$$

In the second step, for t = 1, ..., T, we run T cross-section regressions against the factor exposures (m estimates of the $\hat{\beta}s$ calculated from the first step), to give a time series of risk premia coefficients for each factor. The insight of Fama-MacBeth is to then average these coefficients, once for each factor, to give the premium expected for a unit exposure to each risk factor over time:

$$R_{i,t} - R_{f,t} = \gamma_t + \lambda_{1,t}\widehat{\beta_{1,P}} + \lambda_{2,t}\widehat{\beta_{2,P}} + \cdots + \lambda_{m,t}\widehat{\beta_{m,P}} + \mu_i;$$

where the estimate of γ and λ are given as the average of the cross-sectional regression estimates:

$$\hat{\gamma} = \frac{1}{T}\sum_{t=1}^{T}\widehat{\gamma_t};$$

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T}\widehat{\lambda_t};$$

Also, t-stats for mth risk premium are:

$$\frac{\lambda_m}{\sigma_{\gamma_m}/\sqrt{T}};$$

In the framework of Fama and MacBeth (1973), one simply adds predetermined explanatory variables to the month-by-month cross-section regressions of returns on beta. If all differences in expected return are explained by beta, the average slopes on the additional variables should not be reliably different from zero.

Thus, to decide which profitability proxy has the most predictive power on the cross-section of expected returns, we use the regression approach of Fama and MacBeth (1973) described above. For every month, we regress the cross-section of stocks' returns on gross profits to assets, earnings to book equity, and free cash flows to book equity. Following earlier research (Novy-Marx, 2013), we include several control variables: book-to-market ratio (*log(B/M))*, market value of equity (*log(ME))*, the return for the previous month ($r_{1,0}$) and the return for the previous year, excluding month *t-1* ($r_{12,2}$). Our tests cover the period from July 1964 through December 2010. All independent variables are trimmed at 1% and 99% levels. In order for the coefficient estimates to be comparable across all the specifications, we trim the independent variables on a table-by-table basis. This implies that all regressions are run against the same set of observations.

For the comparison of the predictive power of the profitability measures, we will analyze the t-statistics. Ball, Gerakos, Linnainmaa & Nikolaev (2015) state that the average coefficient estimates in a Fama and MacBeth (1973) regression can be viewed as monthly returns on long-short trading strategies. Thus, the t-statistics of these estimates can be defined as the annualized Sharpe ratios multiplied by $\sqrt{T}$, where T represents the number of years in the dataset.

3.2.2 Fama and MacBeth Regressions: Estimation Results

Table 1 reports the time-series averages of the coefficients on the independent variables. From the 1st specification of the model, we find that the coefficient on gross profitability is positive and statistically significant (0.88 with a t-statistic of 6.54). The estimate we obtained is close to the one reported by Novy-Marx (2013) in Panel A of *Table 1* (0.75 with a t-statistic of 5.49). The small discrepancy in the estimates is due to differences arising from sample construction. The 2nd and 3rd specifications of the model use two alternative measures of profitability: earnings and free cash flow, both scaled by the book value of equity. We find that these measures have much less explanatory power on stocks' returns than gross profitability, which is in line with the findings of Novy-Marx (2013).

Table 1: Fama-Macbeth regressions of returns on measure of profitability.

This table reports results of Fama-MacBeth regressions of firms' returns on gross profits-to-assets, earnings- to-book equity, and free cash flow-to-book equity. Regressions include several control variables: book-to-market [log(B/M)], market value of equity [log(ME)], and past performance measured at horizons of one month ($r_{1,0}$) and 12 to two months ($r_{12,2}$). All independent variables are winsorized at the 1% and 99% levels. The sample excludes financial firms (those with one-digit standard industrial classification codes of six) and covers July 1964 to December 2010.

Independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gross profitability	0.88			0.85	0.78		0.78
	[6.54]			[6.40]	[5.95]		[5.96]
Earnings		0.31		0.08		-0.02	-0.15
		[1.80]		[0.45]		[-0.13]	[0.85]
Free cash flow			0.36		0.22	0.44	0.33
			[3.51]		[2.16]	[4.34]	[3.48]
log(B/M)	0.41	0.35	0.33	0.41	0.38	0.33	0.39
	[7.33]	[6.46]	[6.01]	[7.38]	[6.97]	[6.18]	[7.10]
log(ME)	-0.06	-0.08	-0.09	-0.06	-0.07	-0.08	-0.07
	[-1.51]	[-2.25]	[-2.29]	[-1.76]	[-1.89]	[-2.31]	[-1.84]
$r_{1,0}$	-5.56	-5.51	-5.52	-5.62	-5.62	-5.55	-5.64
	[-14.44]	[-14.44]	[-14.42]	[-14.75]	[-14.73]	[14.58]	[-14.86]
$r_{12,2}$	0.72	0.72	0.71	0.69	0.68	0.70	0.67
	[4.28]	[4.32]	[4.30]	[4.15]	[4.14]	[4.24]	[4.09]

In the 4th specification, we include both gross profitability and earnings in the Fama and MacBeth regression. The coefficient on earnings has decreased by more than half and is not statistically significant, while the coefficient on gross profitability is almost unchanged and stays highly significant (0.85 with a t-value of 6.40). In the 5th specification, we observe a similar situation. After controlling for free cash flow, the gross profitability measure remains a strong predictor of cross section of stock returns (0.78 with a t-value of 5.95). The results of the 6th specification demonstrate that the free cash flow measure absorbs the predictive power of the earnings measure. Finally, the 7th specification shows that gross profitability has the most explanatory power on the cross-section of expected returns. These results are again consistent with the ones reported by Novy-Marx (2013).

In summary, we find that profitable firms experience higher average returns than unprofitable firms. In addition, the strongest predictor of cross-section of returns is the gross profitability measure, which will be used for further analysis in this paper.

3.3 Portfolio sorts on gross profitability

3.3.1 Univariate Sort on Gross Profitability: Estimation Details

From the previous section of the paper, we saw that gross profits-to-assets can explain, on average, the stock returns. However, given the fact that our dataset consists mostly of small-cap stocks and that the previous regressions weighed all the observations equally, the results reported in *Table 1* are tilted towards these small-cap stocks. Portfolio sorts on gross profitability should provide a potentially robust method to assess the explanatory power by value-weighting the positions and by not imposing the parametric assumptions from the Fama and MacBeth regressions.

In this part, we will show the results of univariate sorts on profitability and, for comparison purposes, on the book-to-market ratio. We group the stocks in ascending order into one of the quintile portfolios, which are constructed based on New York Stock Exchange (NYSE) break points on the chosen measure (i.e. gross profit-to asset or B-M ratio). Monthly value-weighted returns are computed for each portfolio. We rebalance the portfolios annually at the end of June, and the dataset covers a time frame from July 1964 to December 2010. In addition, we form a zero-investment portfolio for each measure by going long in the highest quintile portfolio and short in the lowest quintile portfolio.

Further, the Fama French 3 Factor model abnormal returns for each quintile portfolio and High-Low portfolio are estimated by running the following regression:

$$r_{it} - r_{ft} = \alpha_i + b_i(r_{Mt} - r_{ft}) + s_i SMB_t + h_i HML_t + \varepsilon_{it}$$

In this equation, $R_{it}$is the return on security or portfolio i for the period t, $R_{Mt}$is the return on the value-weighted (VW) market portfolio, SMB_t (small minus big) is the difference between the returns on diversified portfolios of small and big stocks, and HML_t (high minus low) is the difference between the returns on diversified portfolios of high and low B/M stocks. If the slopes b_i, s_i and h_i capture all variation in the stock expected returns, the intercept α_i must be zero for all securities and portfolio i. Using thousands of random stock portfolios, Fama and French find that their model could explain as much as 95% of the return in a diversified stock portfolio.

3.3.2 Univariate Sort on Gross Profitability: Estimation Results

In *Table 2*, we report monthly value-weighted average excess returns to portfolios formed on gross profit-to-asset (*Panel A)* and on book-to-market (*Panel B).* We also show the regression results of each quintile portfolio and High-Low portfolio: the three-factor model alphas and loadings on the market (MKT), size (SMB) and value (HML) factor. r^e represents monthly average excess returns of the portfolio.

The results from *Panel A* suggest that the excess returns of the portfolios increase with gross profits-to-assets ratio. The highest quintile portfolio earns a monthly excess return of 0.59%, while the lowest quintile earns 0.31% per month. The difference between the two portfolios' excess returns (i.e. High – Low) generates a 0.28% return spread with a t-value of 2.19. The zero-investment portfolio strategy creates a significant excess return in spite of the significant negative loading on the HML factor. As a result, the three-factor model alpha obtained by pursuing such a trading strategy would achieve 0.49% per month, with a t-value of 4.20. In some way, the three-factor alpha shows the degree to which an investor could boost the portfolio Sharpe ratio by exposing his/her portfolio toward a profitability strategy while controlling for market, size and value factors to protect against any risks brought by those factors.

Table 2: Excess returns to portfolio sorted on profitability.

Panel A reports monthly value-weighted average excess returns to portfolios formed on gross profits-to-assets [(REVT - COGS)/AT], employing NYSE breakpoints, and results of time series regressions of five different quintile portfolios and High-Low portfolio on the Fama and French three-factor model alphas and loadings on the market factor (MKT), the size factor small-minus-large (SMB), and the value factor high-minus-low (HML), with test-statistics (in square brackets). Panel B illustrates similar results for portfolios formed on book-to-market. The sample excludes financial firms (those with one-digit standard industrial classification codes of six) and covers July 1964 to December 2010.). r^e represents monthly average excess returns of the portfolio.

Portfolio	$\boldsymbol{r^e}$	α	**MKT**	**SMB**	**HML**
Panel A: Portfolios sorted on gross profits-to-assets					
Low	0.31	-0.18	0.94	0.04	0.18
	[1.64]	[-2.48]	[57.31]	[1.61]	[7.19]
2	0.41	-0.07	1.01	-0.05	0.16
	[2.05]	[-1.01]	[58.47]	[-2.11]	[6.16]
3	0.50	0.03	1.01	0.04	0.05
	[2.47]	[0.51]	[71.26]	[1.95]	[2.38]
4	0.43	0.07	1.02	0.03	-0.24
	[1.99]	[1.10]	[70.16]	[1.41]	[-10.85]
High	0.59	0.31	0.92	-0.03	-0.29
	[2.96]	[4.61]	[58.91]	[-1.72]	[-12.27]
High - Low	0.28	0.49	-0.02	-0.07	-0.47
	[2.19]	[4.20]	[-0.93]	[-1.99]	[-11.56]
Panel B: Portfolios sorted on book-to-market					
Low	0.38	0.13	0.99	-0.08	-0.41
	[1.81]	[2.73]	[91.10]	[-5.16]	[-24.48]
2	0.46	0.00	0.98	0.06	0.03
	[2.31]	[0.05]	[78.73]	[3.30]	[1.36]
3	0.51	0.02	0.95	0.02	0.2
	[2.71]	[0.29]	[59.26]	[1.10]	[8.02]
4	0.64	0.03	0.93	0.11	0.48
	[3.45]	[0.41]	[61.86]	[5.22]	[20.80]
High	0.76	0.05	0.99	0.23	0.58
	[3.76]	[0.78]	[61.93]	[10.32]	[23.67]
High - Low	0.38	-0.07	-0.00	0.31	0.99
	[2.64]	[-0.94]	[-0.25]	[11.63]	[33.94]

In line with Spearman correlation between gross profitability and book-to-market ratio of -0.19 (*Table 3)*, we see that HML loadings decrease as we move toward more profitable firms. The lowest quintile portfolio has an HML loading of 0.18, with a t-statistic of 7.19, and the highest quintile portfolio registers an HML loading of -0.29, with a t-statistic of -12.27. Thus, profitable firms are more likely to be growth firms as shown by low book-to-market ratios, and unprofitable firms are more likely to be value firms as shown by high book-to-market ratios. However, while one may think that profitable firms and low book-to-market firms are similar given that both of them are growth firms, their expected returns are different. Firms with high gross profits-to-assets tend to earn higher returns than the market despite their low book-to-market ratios. Thus, a trading strategy formed on the basis of profitability would provide a great hedge for value strategies. We will further check this hypothesis by forming portfolios on two dimensions: gross profitability and book-to-market ratio.

Table 3: Spearman rank correlation between independent variables.

This table summarizes the time series averages of the cross section Spearman rank correlations between the independent variables used in the Fama-MacBeth regressions of Table 1: gross profitability [(REVT - COGS)/A], earnings (IB/A), free cash flow ((NI+DP-WCAPCH-CAPX)/A), book-to- market, market equity (ME), and past performance measured at horizons of one month ($r_{1,0}$) and 12 to two months ($r_{12,2}$), with test-statistics (in square brackets). The sample excludes financial firms (those with one-digit standard industrial classification codes of six) and covers July 1963 to December 2010.

Variables	**IB/E**	**FCF/E**	**B/M**	**ME**	$\boldsymbol{r_{12,2}}$	$\boldsymbol{r_{1,0}}$
Gross profitability (GP/A)	0.38	0.27	-0.19	-0.09	0.05	0.03
Earnings (IB/E)		0.54	-0.22	0.22	0.12	0.05
Free cash flow (FCF/E)			-0.09	0.19	0.10	0.04
Book-to-market (B/M)				-0.29	0.05	0.03
Market equity (ME)					0.01	0.04
Prior year's performance ($r_{12,2}$)						0.04

3.4 Profitability and Value

In the previous section, we hypothesized that profitability strategies could achieve higher average returns once we would control for book-to-market. In the same manner, the performance of value strategies could be enhanced by controlling for profitability. We could control for both factors by constructing portfolios using double sorts. As in the univariate sorts, we sort stocks independently into quintiles on gross profits-to-assets and B-M ratio, using NYSE breakpoints. Thus, we end up with 25 portfolios, all of which are rebalanced at the end of every June. We will report portfolios' average returns, as well as time series regression results of profitability and value strategies of High – Low portfolios (*Table 4)*.

Table 4: Double sorts on gross profits-to-assets and book-to-market.

This table shows the double sorted average excess returns for 25 value-weighted (VW) portfolios, all of which are rebalanced at the end of every June employing NYSE breakpoints, on gross profits-to-assets and book- to-market (5x5 GPA-B/M sorting), and results of time series regressions of these high minus low portfolios' returns on the Fama and French factors [the market (MKT), size SMB (small-minus-large) and value factor HML (high-minus-low)]. Test statistics are reported in square brackets. Our sample period starts in July 1964 and ends in December 2010.

	Gross profits-to-asset quintiles					**Profitability strategies**				
	L	2	3	4	H	r^e	α	β_{mkt}	β_{smb}	β_{hml}
B/M quintiles										
L	-0.06	0.19	0.22	0.32	0.54	0.60	0.78	-0.27	-0.26	0.03
						[3.2]	[4.40]	[-6.63]	[-4.58]	[0.49]
2	0.32	0.33	0.42	0.61	0.78	0.45	0.42	-0.09	0.17	0.08
						[2.60]	[2.35]	[-2.21]	[3.00]	[1.29]
3	0.33	0.35	0.67	0.71	1.00	0.67	0.48	0.04	0.56	0.05
						[3.82]	[2.96]	[0.96]	[10.69]	[0.94]
4	0.48	0.57	0.86	0.95	0.86	0.38	0.22	0.07	0.73	-0.21
						[1.92]	[1.34]	[1.91]	[13.49]	[-3.53]
H	0.64	0.69	1.00	1.05	1.10	0.46	0.38	-0.03	0.57	-0.18
						[2.45]	[2.23]	[-0.84]	[10.23]	[-2.98]
Value strategy										
r^e	0.70	0.50	0.78	0.73	0.56					
	[3.26]	[2.80]	[4.03]	[4.14]	[2.89]					
α	0.40	0.20	0.44	0.26	0.00					
	[2.55]	[1.28]	[2.55]	[2.04]	[0.03]					
β_{mkt}	-0.23	-0.06	-0.09	-0.08	0.01					
	[-6.36]	[-1.68]	[-2.35]	[-2.67]	[0.19]					
β_{smb}	-0.00	0.14	0.38	0.75	0.83					
	[-0.01]	[2.84]	[6.80]	[17.8]	[18.17]					
β_{hml}	1.06	0.77	0.74	0.79	0.84					
	[18.85]	[14.07]	[12.26]	[17.07]	[16.91]					

The results from *Table 4* confirm the hypothesis we presented in the paragraph from above. We observe that the returns for value strategies once we control for profitability are significantly higher than the ones presented in *Panel B* of *Table 2.* The average value spread across gross profits-to-assets quintiles is equal to 0.65% per month and in every B-M portfolio is greater than the 0.38% per month spread on the unconditional value of strategy. In the same way, the returns for profitability strategies are greater than the ones reported in *Panel A* of *Table 2.* The average gross profit spread across B-M quintiles is equal to 0.51% per month and in every profitability, portfolio exceeds the 0.28% per month spread on the unconditional profitability of strategy.

4 CoPROF – ACTIVITY, PORTFOLIO FORMATION AND PERFORMANCE

In the previous part of the paper, we showed that gross profitability scaled by assets (GPA) has the most predictive power on the cross-section of expected returns relative to the earnings to book equity and free cash flows to book equity measures. Also, we found that univariate sorts on GPA result in abnormal returns of profitable firms over unprofitable firms. Thus, we detect that there is a gross profitability premium that arbitrageurs could capture, especially if they would be able to time the market correctly. In an attempt to understand how investors could time the market while pursuing a profitability strategy, this chapter introduces a new measure of arbitrage activity in profitability strategies, dubbed as CoPROF. At a first glance, we describe the underlying data and the methodology of constructing the CoPROF, and examine the obtained results. Here, we utilize methodologies from the previous studies of Lou and Polk (2013) and Huang et al. (2016), and at the same time make relevant adjustments. Following that, we form a zero-cost portfolio that goes long the value- weight portfolio of stocks in the highest gross profitability to assets (GPA) decile and short the value-weight portfolio of stocks in the lowest GPA decile. We then track the cumulative abnormal returns of this zero-cost long-short portfolio in months 1 through 36 after portfolio formation, to detect any (conditional) long-run reversal to the profitability-arbitrage strategy.

4.1 Data & Methodology and Construction of CoPROF

4.1.1 Construction of CoPROF: Estimation Details

For the construction of the CoPROF measure, we apply the methodology outlined in the studies of Lou and Polk (2013) and Huang et al. (2016), who create the comomentum and COBAR proxies to measure arbitrage activity in momentum and beta strategies, respectively. More exactly, the CoPROF measure is an indicator of the impact of the arbitrage activity on stock prices, because it captures the past extent of abnormal return correlations among the stocks that the investor trades. Thus, we assume that when investors decide on which stocks to trade, their trading activity will have simultaneous effects on stock prices and hence generate return comovements.

The main dataset used in this section is based on the cleaned data that was applied for the calculation of the gross profits to assets measure and for the univariate sorts on GPA, and double sorts on GPA and book-to-market values in the first part of the paper. The clean dataset provides us with a list of stocks with their monthly returns adjusted for any delisting events, and the corresponding gross profits to asset ratio. This dataset contains all stocks that are traded on the NYSE, Amex and Nasdaq exchanges and it includes only common share stocks (i.e. share code 10 or 11). Any stocks that have accounting data in COMPUSTAT for less than two years, missing market value of equity, missing book-to-market ratio, missing gross profit, missing book value of total assets, missing current month returns and missing returns for the previous one-year period are excluded from our sample.

Further, with the remaining, now clean, dataset, we group the stocks in ascending order into one of the decile portfolios, which are constructed based on NYSE break points on gross profit to asset ratio at the end of June of every year. The portfolios are rebalanced every year at the end of June. After these computations, we end up with a dataset containing the list of stocks assigned to one of 10 portfolios that were formed based on the GPA ratio. For the construction of the CoPROF measure, we leave only the stocks assigned to decile 1,5 or 10. We hypothesize that the CoPROF measure of the profitable firms and unprofitable ones should be highly correlated through time. At the same time, variation in CoPROF in deciles 1 and 10 should not be correlated with variation in the profitability neutral decile (i.e. decile 5). This is explained by the fact that when pursuing a profitability strategy, arbitrageurs go long in profitable firms (i.e. decile 10) and short unprofitable firms (i.e. decile 1) at the same time, which results in *simultaneous* price impact on those stocks and hence leads to excess return comovement.

In the next stage, we have to calculate pairwise partial correlations using daily returns for the previous 2 months from the portfolio formation period for all stocks in each profitability decile. Given that we rebalance our portfolios at the end of June of every year, our sample only contains 47 formation periods. Thus, in order to be able to estimate a CoPROF value for every month, we also use 2 months' worth of past daily returns for each month in the post ranking period, which represents the following 11 months since for formation date of the portfolios. Because of the limited availability of accounting data that forces us to rebalance our portfolios only once in a year, we modify the methodologies of Lou and Polk (2013) and Huang et al. (2016) by using 2 months' worth of daily returns rather than 52 weekly returns for the estimation of pairwise partial correlations.

In the case of the studies of Lou and Polk (2013) and Huang et al. (2016), the authors are able to form portfolios every month, which implies that the composition of the portfolios changes on a monthly basis and thus there is enough variation in the underlying data to estimate COBAR and comomentum measures even after using 52 weekly returns in the portfolio ranking period. In our case, we form portfolios only once a year and as a result the composition of the portfolios stays unchanged for the remaining 11 months. If we were to use 52 weekly returns to estimate the CoPROF measure given that the composition of the portfolios stays unchanged for one year, then there is very little variation in the data, which makes the estimation of CoPROF problematic as there will be overlapping data for 11 months as we move from one month to another. Essentially, we will not see any significant movements in the CoPROF measure until the next rebalancing period. To address this issue, we use daily returns for the previous 2 months in order to estimate the CoPROF measure, allowing for enough variation in the underlying data. The choice of 2 months is justified by an appropriate number of observations (i.e. 42 trading days on average) while minimizing the number of overlapping observations.

On the implementation side, in order to estimate the pairwise partial correlations, we first select all distinct combinations of stocks (i.e. PERMNOs) and the end of month dates for deciles 1, 5, and 10. In addition, we create a complementary date variable, which is equal to the end of month date variable less than 2 months. Then, we join the PERMNOs from our new dataset with the corresponding PERMNO's daily returns downloaded from CRSP where the dates of the daily returns should be between the end of month date and the end of month date less than 2 months. We also import the variables of Fama French 3 Factor model measured on a daily basis from Ken French's data library and include it to the dataset of daily returns. We include the 3 factors to remove any comovements in stocks caused by common risk factors. After preparing the final dataset, we estimate the daily excess return for all stocks and compute the equal weighted daily returns for each portfolio, not taking into account stock *i*. This procedure is done for all deciles (i.e. 1, 5, and 10) separately. The following formula is applied:

$$retf_{-i} = \frac{(\sum_{i=1}^{N} Excess\ Return_i) - Excess\ Return_i}{N-1};$$

where $Excess\ Return_i$ represents the daily excess return of stock i and N is the number of stocks in each decile for a specific month. Next, we calculate the abnormal comovement of stocks involved in profitability arbitrage strategy (CoPROF) as the average pairwise correlation of the 3 factor residual of every stock in each of the GPA deciles with the remaining stocks corresponding to the same decile. This implies that we apply the following formula for a given month in our sample:

$$CoPROF = \frac{1}{N}\left(\sum_{i=1}^{N} partialCorr(retf_i, retf_{-i} \mid mktrf, smb, hml)\right);$$

where $retf_i$ represents the daily return of stock in its specific decile; $retf_{-i}$ is, as defined previously, the equal weighted daily return of the portfolio excluding stock i in its specific decile; and N represents the number of stocks in each of the deciles. Overall, we obtain 564 monthly figures of CoPROF for deciles 1, 5, and 10, ranging from June 1964 to the end of 2010.

4.1.2 Construction of CoPROF: Estimation Results

In this section, we examine simple characteristics of our arbitrage activity measure CoPROF. The study from Huang et al. (2016) documents a significant excess correlation among low-beta stocks on average and this pairwise correlation varies tremendously over time from the lowest value of 0,04 to the highest one of 0.20. *Table 5 Panel A* reports that the average value of CoPROF in the lowest GPA decile is 0.037 which is substantially lower than for CoBAR while the volatility of the two time series are roughly the same. An explanation could be limited attention from arbitrageurs to the profitability strategy, which became popular in academic literature after Novy-Marx's (2013) paper. On the other hand, the beta anomaly had been well-known for the long period of time since Black, Jensen, and Scholes (1972) provided a criticism of the CAPM that the security market line is too flat on average, or the risk-adjusted returns of high beta stocks are too low relative to those of low-beta stocks.

Table 5 Panel B depicts the correlation between different deciles for CoPROF, and examines CoPROF's correlation with existing measures linked to time variation in the expected abnormal returns to profitability-arbitrage strategies. The results show that the highest correlation, equal to 0.476, occurs between CoPROF decile 1 and decile 10, which is reasonable as the two deciles are the ones where the majority of trading activity takes place. Surprisingly, the correlation between CoPROF in decile 1 and 5 is around 0.470 and highly significant, and it is almost no different than the correlation obtained between decile 1 and 10. A potential explanation for this finding might be the fact that the profitability anomaly has been brought to our attention only in the recent years. Thus, a long time series will not capture significant differences in the correlation between decile 1, 5 and 10 for CoPROF variable. An alternative explanation could be that the methodology might not be the most appropriate for the profitability strategy. Opposite to the results we obtained, we were expecting that the profitability neutral portfolio (i.e. decile 5), indicating the average level of trading activity, would have a lower correlation with the two most crowded portfolios (i.e. decile 1 and 10) in terms of arbitrage activity.

Table 5: Summary statistics of the arbitrage activity measures

This table illustrates characteristics of CoPROF in three different deciles, the excess comovement in profitability strategies among stocks over the period 1964-2010, in comparison to the original CoBAR from Huang et al. (2016) paper. Panel A presents the number of observations, the average value of the whole period, volatility, minimum- and maximum values of the CoPROF measure. Panel B reports the correlation between variables. CoPROF dec.1 is made up by stocks with the lowest GPA values, dec.5 consists of stocks with median GPA values, and dec.10 contain stocks with the highest GPA values.

Panel A: Summary Statistics					
Variable	N	Mean	Std. Dev.	Min	Max
CoPROF dec.1	564	0.037	0.027	-0.055	0.134
CoPROF dec.5	564	0.019	0.027	-0.074	0.156
CoPROF dec.10	564	0.019	0.027	-0.049	0.134
COBAR	528	0.105	0.026	0.037	0.203

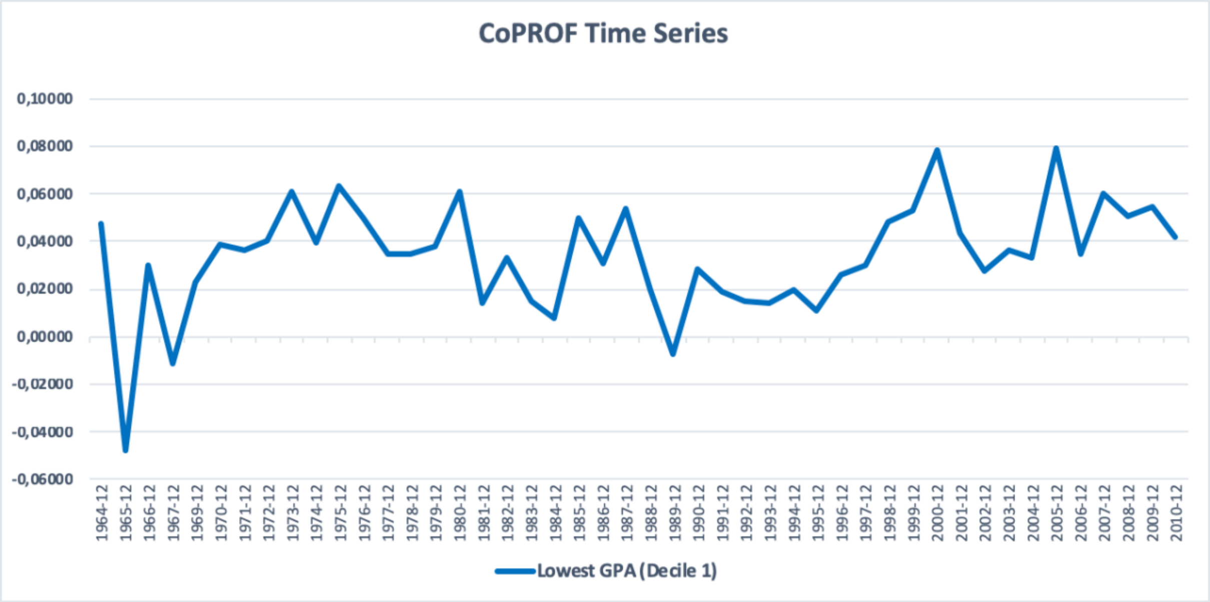
Panel B: Correlation			
	CoPROF dec.1	*CoPROF* dec.5	*CoPROF* dec.10
CoPROF dec.1	1		
CoPROF dec.5	0.470	1	
CoPROF dec.10	0.476	0.407	1

Figure 1 plots CoPROF in the two extreme deciles as the end of each December. Similar to CoBAR, we do not necessarily expect a trend in our measure. An increase in CoPROF basically implies that more capital flows to profitability-arbitrage strategy, which provides more liquidity to the market. After a substantial drop in the lowest decile CoPROF in the period of 1986-1989, the level of arbitrage activity started recovering and followed an upward trend for the rest of the sample. Furthermore, the business cycle seems to have a close relationship with arbitrage activity in the profitability strategy. More specifically, the CoPROF series tend to shift direction in advance of the business cycle. They turn down before an economic contraction and turn up before an economic expansion. For instance, the CoPROF trends correspond to the market crash in 1973-1974 ('Oil Crisis'), 1980-1982 ('Volcker crash'), 2000-2002 ('Dot-com bubble') and 2007-2009 ('Housing bubble').

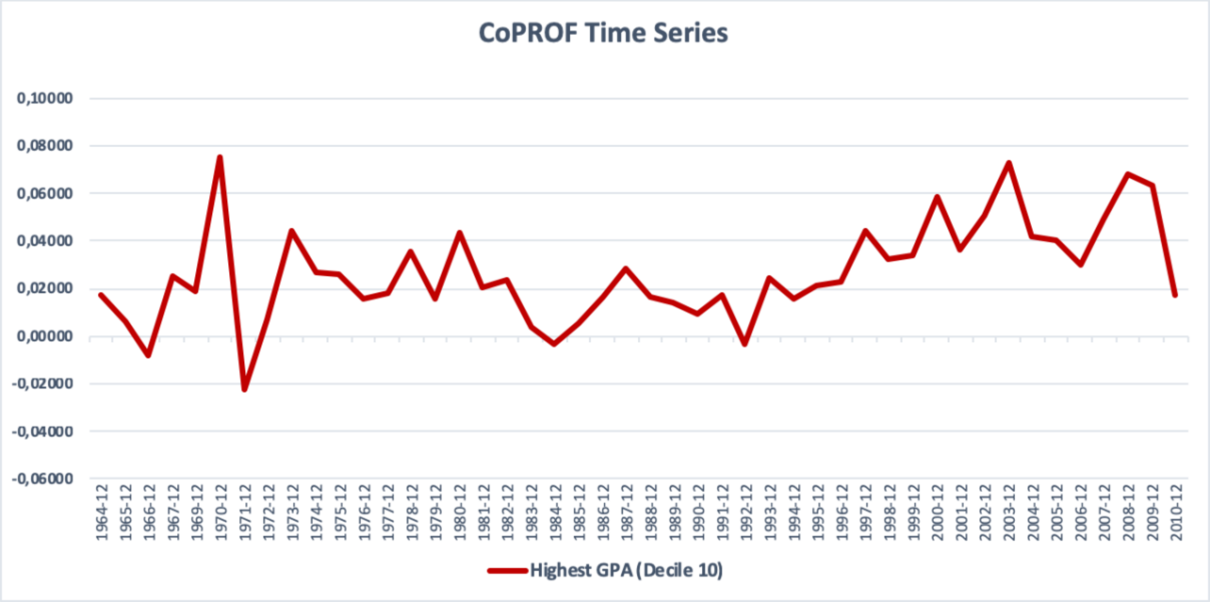
Figure 1: The time-series of CoPROF

The figure portrays the CoPROF time series over the period from 1964 to 2010. The blue line in Panel A depicts CoPROF formed on the lowest decile of gross profits to assets, while the red line in Panel B describes CoPROF formed on the highest decile of gross profits to assets. At the end of each June, all stocks are sorted into deciles based on their GPA calculated in the same year. CoPROF is computed as the average pairwise partial daily return correlation in the highest profitability decile over the past 2 months.

Panel A: CoPROF formed on the lowest decile of gross profits to assets



Panel B: CoPROF formed on the highest decile of gross profits to assets



4.2 Portfolio Formation in the Profitability-strategy

4.2.1 Portfolio Formation: Estimation Details

In this part of the paper, we turn our attention to portfolio formation in profitability strategies. Similar to the methodologies of Lou and Polk (2013) and Huang et. al (2016), we construct a value-weighted portfolio that takes a long position in profitable firms (i.e. decile 10) and a short position in unprofitable firms (i.e. decile 1). We choose to form the portfolios based on the CoPROF values that were computed using the stocks in decile 10. This is explained by the fact that the stocks in the decile 10 are usually larger, more liquid, less prone to distress and have longer cash flow durations in comparison to unprofitable stocks (i.e. decile 1) (Novy Marx, 2013). As a result, CoPROF variable of profitable firms is less likely to have issues caused by asynchronous trading and observational noise. However, for robustness purposes we repeat the portfolio formation stage taking into account CoPROF values of decile 1 and 5, and present the results of the robustness tests in chapter 5 of the paper.

We first start by importing the CoPROF monthly values of decile 10 we have calculated in the previous part of the paper. We split the whole dataset into CoPROF quintiles, such that every month in this dataset is assigned to a corresponding CoPROF quintile, denoting the level of arbitrage activity in profitability strategies in a given month. At this point, because we are using the whole sample to form the CoPROF quintiles, we are actually using future information when we define the level of arbitrage activity in profitability strategies of each month in our sample. To further account for this issue and ensure that the obtained results are not biased, we design an additional robustness test that will not allow for the use of future information when deciding to which CoPROF quintile a particular month is assigned to.

The second dataset that we import is the already cleaned dataset that was used for the calculation of CoPROF measure. It contains the monthly stock returns, the relevant PERMNOs, share price, number of shares outstanding, the gross profits to assets for each stock, and information regarding which profitability decile each stock is assigned to. For easier data tracking, we split this dataset into two smaller sets: one that only contains stocks from decile 1 and the second that only contains stocks from decile 10. In each of the smaller datasets, we create an additional date column that is equal to the monthly date variable (i.e. start date) plus 36 months (i.e. end date). In this way, we are able to join the PERMNOs from the 2 samples with the corresponding PERMNO's monthly returns downloaded from CRSP where the dates are between the start and end dates. We end up with 2 datasets that contain the original monthly data for each stock and similar information for the following 36 months for each stock. This data will further be used to evaluate the performance of the portfolios over different holding periods. During the next stage, we calculate the monthly value weighted returns for each portfolio in our sample.

The next step is to build our long-short portfolios and track their performance. We form the portfolios by subtracting the value-weighted returns of the portfolios in decile 1 from the value-weighted returns of the portfolio in decile 10. To evaluate the performance of the long-short profitability strategy, we will use the Fama French 3 Factor model and Fama French 5 Factor model. Once we get the value-weighted returns of the long-short portfolio, we append the 3 factors and the 5 factors to our sample. Next, we assign the long-short portfolios' monthly value-weighted returns and the corresponding 36 monthly returns to a corresponding CoPROF quintile from the first dataset we imported. Next, we evaluate the performance of our portfolio by analyzing the alphas that we obtain from the Fama French 3 Factor and Fama French 5 Factor regressions. Thus, for each of the quintiles and each of 36 holding periods, we first run the following regression:

$$r_{it} - r_{ft} = \alpha_i + b_i(r_{Mt} - r_{ft}) + s_i SMB_t + h_i HML_t + \varepsilon_{it}$$

where $r_{it}$is the return on portfolio i for the period t, $r_{Mt}$is the return on the value-weighted (VW) market portfolio, SMB_t (small minus big) is the difference between the returns on diversified portfolios of small and big stocks, and HML_t (high minus low) is the difference between the returns on diversified portfolios of high and low B/M stocks.

For Fama French 5 Factor model, the following regression should be run for each quintile and each of the 36 holding periods:

$$r_{it} - r_{ft} = \alpha_i + b_i(r_{Mt} - r_{ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + \varepsilon_{it}$$

where RMW_t is the difference between the returns on diversified portfolios of stocks with robust and weak profitability, and CMA_t is the difference between the returns on diversified portfolios of low and high investment stocks, which are regarded as conservative and aggressive investment strategies. Finally, as in the studies of Lou and Polk (2013) and Huang et al (2016), we control for autocorrelation and heteroskedasticity by applying Newey-West standard errors when calculating the t-statistics.

4.2.2 Portfolio Formation: Estimation Results

This section describes the abnormal returns (alpha) on the profitability-arbitrage strategy, conditional on CoPROF. Our purpose is to examine how crowded trades affect asset prices and detect any long-run reversal to the corresponding arbitrage strategy. We implement this by tracking the risk-adjusted returns of the zero-cost long-short portfolio (i.e., to go long the value-weight high-GPA decile and short the value-weighted low-GPA decile) for the first 6 months following the profitability-arbitrage trade, through month 7 to 12, and those occurring in year 2 and 3. The sample is sorted into five CoPROF quintiles, and the returns are estimated as a function of CoPROF value.

Table 6 summarizes our results. We apply both Fama and French (1993) three-factor (Panel A) and Fama and French (2014) five-factor (Panel B) to investigate abnormal returns. The results from panel A indicates that the arbitrage yields an abnormal three-factor return of 0.13%/month on average in the lowest quintile of CoPROF (rank 1) in the six months shortly subsequent to the profitability-arbitrage trade. This finding, however, is not statistically significant at 5% level with a t-statistic of 1.03. The alpha decreases in the next six months before an improvement in year 2 and 3, ending up at the level of 0.27%/month with a t-statistics of 3.48 in year 3. It is perceived that pursuing a profitability-arbitrage strategy when the arbitrage activity is at the lowest level takes patience as the investors can capture the abnormal returns starting only from year 2.

On the other hand, the abnormal returns arrive sooner and stronger following an increase in the levels of arbitrage activity (i.e. when we move from rank 1 to 5). More specifically, the difference in the monthly returns to the long-short strategy when CoPROF is high vs. when it is low in the first six months is large and statistically significant at 0.41%/month and with t-stat of 2.39. In addition, when the arbitrage activity is at its highest (rank 5), we notice that the average risk-adjusted returns in the FF-3 are strongly positive and statistically significant in all holding periods. Surprisingly, they pursue the same pattern as we examine in rank 1 of CoPROF. The alpha drops during the period of 7-12 months by 0.18%/month with a t-stat of 2.92, which is followed by a moderate improvement in year 2 before a peak in year 3, ending at the level of 0.57%/month with a t-statistics of 6.05. Furthermore, the difference between abnormal three-factor returns to profitability arbitrage in high and low CoPROF periods in year 3 is not only economically large at 0.30%/month but also statistically significant.

Table 6, Panel B exhibits the results using the Fama-French five-factor model (FF5F). The primary reason why we implement the five-factor model instead of Carhart four-factor model (FF3F + momentum factor) is that the former contains the profitability factor based on operating profitability (RMW). In this way we are able to test whether the model is able to capture all the alphas of the gross profitability. Our results show that the majority of the alphas disappear as compared to the ones obtained from FF3F model in *Table 6, Panel A*. This implies that the FF5F model performs better than FF3F to detect mispricing in the profitability-arbitrage strategy. However, we still find that when arbitrage activity is relatively high, abnormal returns on profitability-arbitrage strategies appear relatively quickly, within the first six months of the trade (t-stat = 2.79). It seems that the profitability factor, which is based on the operating profitability rather than gross profitability can only capture alphas when arbitrage activity is relatively low, but does not perform well when the level of arbitrage activity is relatively high.

In general, the **key finding** of our paper is that those quicker and stronger profitability-arbitrage returns following an increase in the level of arbitrage activity *cannot* be linked to consecutive reversal in the long run. During the periods of both high and low coprofitability, profitability strategies tend to be profitable and stabilizing; the returns to corresponding arbitrage activities reflect over-correction due to crowded arbitrage trading in the short period of time while the difference in year 3 abnormal five-factor returns is positive at 0.12%/month (t-statistic = 1.03), but not statistically significant. This conclusion is distinctive from what was found in Lou and Polk (2013) for momentum strategy and Huang, Lou and Polk (2014) for beta strategy. Specifically, according to the Huang et al. (2016) study, in year three and when $CoBAR$ is high, the abnormal four-factor return to beta arbitrage stays at a negative and statistically significant level of -0.93%/month. These abnormal returns are extremely different from their comparable values when $CoBAR$ is low; the difference in year 3 abnormal four-factor returns is -1.52%/month (t-statistic = -3.33).

Table 6: Forecasting Profitability-arbitrage Returns with CoPROF (decide 10)

The table outlines returns to the profitability strategy as a function of lagged CoPROF. At the end of each June, all stocks are sorted into 10 deciles based on their gross profitability scaled by assets (GPA) in the same year. We then sort CoPROF taken from decile 10, the average pairwise partial correlation for all stocks in the lowest-PROF decile over the past two months whilst controlling for the Fama-French (1992) three factors, into five different quintiles. Reported below are the risk-adjusted returns to the profitability strategy in each of the 36 months after portfolio formation between 1964 and 2010, following low to high CoPROF. "5-1" is the difference in monthly returns to the hedged strategy following high vs. low CoPROF. T-statistics, shown in parentheses, is measured by applying Bartlett kernel standard errors corrected for serial- dependence with 6 or 12 lags, depending on the holding periods. 5% statistical significance is underlined in bold.

Panel A: FF3F-Adjusted Profitability-arbitrage Returns

Decile 10	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	0.13 %	(1.02)	-0.05 %	(-0.39)	0.13 %	(1.68)	**0.27 %**	**(3.48)**
2	-0.03 %	(-0.28)	-0.01 %	(-0.06)	-0.06 %	(-0.73)	0.08 %	(1.00)
3	-0.17 %	(-1.36)	-0.16 %	(-1.20)	0.12 %	(1.16)	0.09 %	(1.05)
4	**0.38 %**	**(2.83)**	0.14 %	(1.12)	0.17 %	(1.60)	-0.09 %	(-0.98)
5	**0.53 %**	**(4.74)**	**0.36 %**	**(2.92)**	**0.38 %**	**(4.17)**	**0.57 %**	**(6.05)**
5-1	**0.40 %**	**(2.39)**	**0.41 %**	**(2.29)**	**0.25 %**	**(2.01)**	**0.30 %**	**(2.46)**

Panel B: FF5F-Adjusted Profitability-arbitrage Returns

Decile 10	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	-0.09 %	(-0.69)	-0.13 %	(-1.01)	0.02 %	(0.22)	0.02 %	(0.24)
2	-0.22 %	(-1.90)	-0.21 %	(-1.42)	**-0.24 %**	**(-2.57)**	-0.17 %	(-1.92)
3	**-0.28 %**	**(-2.10)**	-0.29 %	(-1.99)	-0.06 %	(-0.66)	0.01 %	(0.09)
4	**0.27 %**	**(2.12)**	0.05 %	(0.42)	0.06 %	(0.64)	**-0.27 %**	**(-2.93)**
5	**0.31 %**	**(2.79)**	0.14 %	(1.29)	0.12 %	(1.32)	0.15 %	(1.69)
5-1	**0.40 %**	**(2.32)**	0.27 %	(1.60)	0.10 %	(0.75)	0.12 %	(1.03)

5 ROBUSTNESS TESTS

In the main section of the results, we find that abnormal returns in profitability strategies do not crash in the long run, but just get diminished to 0 as the holding period increases. In order to make sure that these results still hold across various specifications within the methodology, we run several robustness tests.

The first series of robustness tests repeat the methodology for the portfolio formation stage, however now the breakpoints for the CoPROF quintiles are calculated based on the COPROF values of decile 1. As previously mentioned, we hypothesize that the abnormal returns obtained in a profitability strategy using CoPROF values of decile 1 should be similar to the ones obtained in a profitability strategy using CoPROF values of decile 10.

The second series of robustness tests put restrictions on the usage of future information when forming the CoPROF quintiles in the portfolio formation stage. More exactly, we do not allow for future information to be used when assigning a given month to a particular CoPROF quintile. For example, for determining the CoPROF quintile breakpoints for the month of May of 2006, we would use only the sample up to this date, excluding all the months following May 2006. After getting the relevant breakpoints, we would assign the May 2006 date to the relevant quintile based on its CoPROF value at that point in time. This procedure is repeated for all the months in our dataset. The main purpose of this robustness test is to ensure the validity and unbiasedness of the previously obtained results.

5.1 Portfolio Formation in the Profitability-strategy

This section highlights the adjusted profitability-arbitrage returns (alphas), applying the methodology described in *section 5.2*, conditional on CoPROFs from decile 1. In general, we rank CoPROFs in each decile into five quintiles so that the CoPROFs are assigned to one of the five groups in every month, where group 1 represents for the lowest CoPROF values and group 5 with the highest ones. We then compute long-short portfolio abnormal returns from the zero-cost hedged portfolio for the first 6 months following the profitability-arbitrage trade, through month 7 to 12, and those occurring in year 2 and 3.

As we mentioned in the previous section, we aim to identify how crowded trades affect the returns and detect any long-run reversal to the corresponding arbitrage strategy. In the following, we would like to emphasize the main outcomes, and examine the results of arbitrage returns when different deciles are taken into account. *Table 7* summaries our test on decile 1 using the Fama and French (1993) three-factor (*Panel A*) and Fama and French (2014) five-factor model (*Panel B*) to study abnormal returns.

Table 7: Forecasting Profitability-arbitrage Returns with CoPROF (decile 1)

The table shows returns to the profitability strategy as a function of lagged CoPROF. We sort CoPROF taken from decile 1 into five different quintiles. Reported below are the risk-adjusted returns to the profitability strategy in each of the 36 months after portfolio formation between 1964 and 2010, following low to high CoPROF. "5-1" is the difference in monthly returns to the hedged strategy following high vs. low CoPROF. 5% statistical significance is underlined in bold.

Panel A: FF3F-Adjusted Profitability-arbitrage Returns								
Decile 1	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	**0.37 %**	**(2.83)**	0.15 %	(1.41)	**0.19 %**	**(2.25)**	**0.30 %**	**(3.91)**
2	0.00 %	(0.03)	0.10 %	(0.80)	**0.22 %**	**(2.50)**	**0.19 %**	**(2.88)**
3	-0.10 %	(-0.73)	-0.27 %	(-1.96)	0.18 %	(1.84)	**0.21 %**	**(2.32)**
4	0.22 %	(1.45)	-0.06 %	(-0.52)	0.13 %	(1.25)	0.12 %	(1.11)
5	**0.43 %**	**(3.48)**	**0.34 %**	**(2.36)**	0.09 %	(0.86)	**0.29 %**	**(2.57)**
5-1	0.06 %	(0.32)	0.19 %	(1.02)	-0.10 %	(-0.76)	0.00 %	(-0.02)

Panel B: FF5F-Adjusted Profitability-arbitrage Returns								
Decile 1	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	0.08 %	(0.56)	**-0.26 %**	**(-2.07)**	-0.15 %	(-1.64)	0.06 %	(0.72)
2	-0.11 %	(-0.80)	-0.10 %	(-0.70)	0.02 %	(0.20)	-0.10 %	(-1.28)
3	-0.23 %	(-1.73)	**-0.36 %**	**(-2.49)**	0.05 %	(0.54)	0.11 %	(1.18)
4	0.07 %	(0.54)	-0.12 %	(-0.96)	-0.05 %	(-0.53)	-0.02 %	(-0.16)
5	0.17 %	(1.30)	0.14 %	(1.10)	-0.14 %	(-1.43)	-0.16 %	(-1.55)
5-1	0.09 %	(0.47)	**0.40 %**	**(2.25)**	0.01 %	(0.11)	-0.22 %	(-1.71)

Table 7, Panel A shows that most of three-factor adjusted returns in both rank 1 and 5 are statistically significant at 5% level. When the arbitrage activity is at its lowest (rank 1), the abnormal return is at its highest of 0.37%/month (t-stat = 2.83) in the six months shortly following the profitability-arbitrage trade. The returns tend to decrease in the next 6 months, prior to an improvement in the subsequent periods, ending at 0,30%/month in year 3 with t-stat of 3.91. In contrast, when the arbitrage activity is at its highest (rank 5), the abnormal return takes more time until year 3 so that it can be recovered at 0.29%/month with a t-stat of 2.57. Thus, during periods of high coprofitability, the profitability-arbitrage strategies do not tend to crash and revert in the long run. In comparison to the results in decile 10 **(***Table 6*, *Panel A***)**, the profitability spread (the difference in monthly returns to the hedged strategy following high vs. low CoPROF) obtained for decile 1 is substantially lower and no longer statistically significant. Additionally, the abnormal returns achieve the highest and statistically significant level in the first 6 months for decile 1 whereas the highest figures are recognized in year 3 for decile 10 result.

When we apply FF5F model **(***Table 7*, *Panel B***),** the significance of alphas in both the highest- and lowest rank disappears for all time-periods with the exception of month 7-12 in rank 1, where the alpha is significant at -0.26% /month. This implies that the profitability factor undoubtedly deflates alphas of profitability strategies when it is presented in the asset pricing model; however, RMW still cannot capture all alphas of gross profitability. Furthermore, the profitability spread for decile 1 is distinctive from the one obtained for decile 10. Across all holding periods, we do not observe statistically significant returns, except for the period of 7 to 12 months when the spread is equal to 0.40% with a t-statistic of 2.25. Compared to the decile 10 results, it seems to take more time to realize the abnormal returns in the profitability arbitrage strategy based on CoPROF, conditional on the decile of unprofitable firms.

Table 8: Forecasting Profitability-arbitrage Returns with CoPROF (decile 5)

The table shows returns to the profitability strategy as a function of lagged CoPROF based on decile 5. Reported below are the risk-adjusted returns to the profitability strategy in each of the 36 months after portfolio formation, following low to high CoPROF. "5-1" is the difference in monthly returns to the hedged strategy following high vs. low CoPROF. 5% statistical significance is underlined in bold.

Panel A: FF3F-Adjusted Profitability-arbitrage Returns								
Decile 5	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	-0.05 %	(-0.40)	-0.18 %	(-1.61)	0.07 %	(0.73)	0.06 %	(0.82)
2	0.19 %	(1.59)	0.08 %	(0.69)	0.00 %	(-0.01)	0.00 %	(0.01)
3	0.05 %	(0.38)	-0.05 %	(-0.40)	0.15 %	(1.81)	**0.22 %**	**(2.48)**
4	0.08 %	(0.64)	-0.02 %	(-0.14)	0.20 %	(1.90)	**0.21 %**	**(2.47)**
5	**0.54 %**	**(4.17)**	**0.43 %**	**(3.00)**	**0.25 %**	**(2.35)**	**0.53 %**	**(5.78)**
5-1	**0.59 %**	**(3.24)**	**0.61 %**	**(3.35)**	0.18 %	(0.33)	**0.47 %**	**(3.81)**

Panel B: FF5F-Adjusted Profitability-arbitrage Returns								
Decile 5	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	-0.20 %	(-1.48)	**-0.45 %**	**(-3.52)**	-0.17 %	(-1.74)	-0.06 %	(-0.70)
2	0.12 %	(0.88)	-0.07 %	(-0.52)	-0.17 %	(-1.74)	**-0.27 %**	**(-2.59)**
3	-0.08 %	(-0.61)	-0.15 %	(-1.21)	-0.01 %	(-0.16)	-0.02 %	(-0.25)
4	-0.09 %	(-0.74)	-0.14 %	(-1.16)	0.10 %	(1.01)	0.08 %	(0.94)
5	0.25 %	(1.80)	0.15 %	(1.11)	-0.05 %	(-0.52)	**0.23 %**	**(2.71)**
5-1	**0.45 %**	**(2.32)**	**0.60 %**	**(3.24)**	0.12 %	(0.88)	**0.28 %**	**(2.37)**

Table 8 illustrates results of adjusted profitability-arbitrage returns, conditional on CoPROF based on decile 5. When FF3F is employed to detect mispricing (Panel A), there are no statistically significant alphas for the lowest activity level (rank 1) whilst the opposite is true for the highest activity level (rank 5). When FF5F model is applied (Panel B), the abnormal returns appear in both highest and lowest activity levels. In contrast, we were not expecting to detect any significant abnormal returns in the profitability neutral decile, which is associated with the level of arbitrage activity of the typical stocks. This result may indicate that the methodology applied is not the most appropriate to measure the arbitrage activity for profitability strategies. Despite this fact, we still observe that when the activity is high, the abnormal returns are statistically significant across the three different specifications of the CoPROF variable.

5.2 Ignoring future information when forming the CoPROF quintiles

This section presents the results of abnormal returns from the profitability-arbitrage strategy given that restrictions on the usage of future information are taken into account when forming the CoPROF quintiles.

Table 9 describes our test results on decile 10 using the Fama and French (1993) three-factor (*Panel A*) and Fama and French (2014) five-factor model (*Panel B*) to study abnormal returns. When the level of arbitrage activity is at its lowest (quintile 1), the profitability-arbitrage strategy shows delayed correction, taking up to three years for abnormal returns to be fully realized (*Panel A*). The alpha increases to the highest level of 0.39%/month in year 3 with a t-stat of 4.19. On the other hand, when the level of arbitrage activity is at its highest (quintile 5), all three-factor adjusted returns are statistically significant at 5% level. The alphas decrease by time until year 2, followed by a strong expansion to 0.42%/month with a t-stat of 4.67 in year 3. In comparison to the outcome from Table 6, there are no significant differences for ranks 1 and 5; nonetheless, the profitability spread is only statistically significant in the first six months for the robust results. Panel B indicates the results when FF3F is replaced by FF5F model. It is no doubt that FF5F model performs better than FF3F in capturing significantly the profitability anomaly, but the alphas still exists in the highest quintile of CoPROF.

Table 9: Forecasting Profitability-arbitrage Returns with CoPROF with constraint (decile 10)

The table shows returns to the profitability strategy as a function of lagged CoPROF. We sort CoPROF taken from decile 10 into five different quintiles, while put restrictions on the usage of future information when forming the CoPROF quintiles in the portfolio formation stage. Reported below are the risk-adjusted returns to the profitability strategy in each of the 36 months after portfolio formation between 1964 and 2010, following low to high CoPROF. "5-1" is the difference in monthly returns to the hedged strategy following high vs. low CoPROF. 5% statistical significance is underlined in bold.

Panel A: FF3F-Adjusted Profitability-arbitrage Returns								
Decile 10	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	0.12 %	(0.74)	-0,03 %	(-0.20)	0.01 %	(0.13)	**0.39 %**	**(4.19)**
2	0.04 %	(0.35)	-0,15 %	(-1.23)	0.08 %	(1.06)	0.10 %	(1.47)
3	-0.17 %	(-1.02)	-0,05 %	(-0.33)	0.07 %	(0.69)	0.09 %	(0.95)
4	0.17 %	(1.22)	0,07 %	(0.57)	**0.23 %**	**(2.02)**	0.00 %	(-0.02)
5	**0.52 %**	**(4.61)**	**0,34 %**	**(2.78)**	**0.28 %**	**(3.19)**	**0.42 %**	**(4.67)**
5-1	**0.40 %**	**(2.06)**	0,37 %	(1.93)	**0.26 %**	**(2.01)**	0.03 %	(0.25)

Panel B: FF5F-Adjusted Profitability-arbitrage Returns								
Decile 10	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	-0.05 %	(-0.34)	-0.11 %	(-0.80)	-0.05 %	(-0.47)	0.10 %	(1.03)
2	**-0.26 %**	**(-2.17)**	-0.23 %	(-1.66)	-0.15 %	(-1.67)	-0.14 %	(-1.62)
3	-0.22 %	(-1.55)	-0.28 %	(-1.82)	-0.03 %	(-0.32)	-0.15 %	(-1.56)
4	0.10 %	(0.95)	0.11 %	(0.86)	0.14 %	(1.64)	**-0.25 %**	**(-2.70)**
5	**0.44 %**	**(3.92)**	0.13 %	(1.23)	0.00 %	(0.03)	**0.26 %**	**(2.84)**
5-1	**0.49 %**	**(2.73)**	0.24 %	(1.38)	0.05 %	(0.38)	0.16 %	(1.26)

In addition, *Table 10* outlines the abnormal returns to the profitability-arbitrage strategy by sorting CoPROF whose construction was based on stocks from the lowest profitability decile (decile 1). Compared to the previous results on *Table 8*, when the level of arbitrage activity is at its lowest, the profitability-arbitrage strategy displays statistically significant alphas only in the six months shortly subsequent to the profitability-arbitrage trade (*Panel A*). In addition, when the level of arbitrage activity is at its highest, the three-year adjusted abnormal returns tend to decrease over time until year 2, followed by an upward trend to the level of 0.20%/month with a t-stat of 2.27 in year 3. In *Panel B*, the FF5F exhibits its better performance to detect the abnormal returns, as the majority of alphas technically become zero when the profitability risk factor is incorporated in the pricing model. Similar to the result without constraint of future information (*Table 8, Panel B*), we still find a statistically significant alpha of -0.36% (t-stat = -2.19) in the lowest activity (quintile 1) in the period of months 7-12.

Table 10: Forecasting Profitability-arbitrage Returns with CoPROF with constraint (decile 1)

The table shows returns to the profitability strategy as a function of lagged CoPROF. We sort CoPROF taken from decile 1 into five different quintiles, while put restrictions on the usage of future information when forming the CoPROF quintiles in the portfolio formation stage. Reported below are the risk-adjusted returns to the profitability strategy in each of the 36 months after portfolio formation between 1964 and 2010, following low to high CoPROF. "5-1" is the difference in monthly returns to the hedged strategy following high vs. low CoPROF. 5% statistical significance is underlined in bold.

Panel A: FF3F-Adjusted Profitability-arbitrage Returns								
Decile 1	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	**0.35 %**	**(2.24)**	0.24 %	(1.42)	0.19 %	(1.61)	0.19 %	(1.86)
2	**0.30 %**	**(2.22)**	0.16 %	(1.21)	**0.19 %**	**(2.11)**	**0.33 %**	**(4.33)**
3	-0.10 %	(-0.74)	-0.18 %	(-1.25)	**0.35 %**	**(3.59)**	**0.25 %**	**(3.11)**
4	0.02 %	(0.15)	0.00 %	(0.03)	0.09 %	(0.94)	**0.20 %**	**(2.13)**
5	**0.31 %**	**(2.88)**	0.10 %	(0.91)	0.09 %	(1.06)	**0.20 %**	**(2.27)**
5-1	-0.04 %	(-0.21)	-0.14 %	(-0.71)	-0.10 %	(-0.70)	0.01 %	(0.09)

Panel B: FF5F-Adjusted Profitability-arbitrage Returns								
Decile 1	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	-0.01 %	(-0.08)	**-0.36 %**	**(-2.19)**	-0.25 %	(-1.93)	-0.07 %	-0.71)
2	0.15 %	(0.97)	-0.18 %	(-1.15)	-0.05 %	(-0.48)	0.05 %	(0.51)
3	-0.22 %	(-1.77)	**-0.29 %**	**(-2.04)**	0.12 %	(1.14)	0.03 %	(0.29)
4	-0.14 %	(-1.02)	-0.07 %	(-0.60)	-0.04 %	(-0.50)	0.02 %	(0.22)
5	0.16 %	(1.48)	0.01 %	(0.11)	-0.08 %	(-1.05)	-0.07 %	(-0.83)
5-1	0.17 %	(0.84)	0.37 %	(1.93)	0.16 %	(1.06)	0.00 %	(0.02)

Lastly, *Table 11* describes the results of alpha to the profitability-arbitrage strategy by sorting CoPROF whose construction was based on stocks from the neutral profitability decile (decile 5). We focus on comparing these results to the comparable ones in *Table 8* without constraint of future information. In general, there is no significant difference between the two. Here, we do not find abnormal returns in any of the holding periods in the lowest activity quintile when both asset pricing models are applied. In contrast, in the highest activity quintile, the alphas are still statistically significant but smaller both in magnitude and t-stat values. The FF5F cannot capture the alpha of 0.15% (t-stat = 2.29) in year 3, which is similar to the outcome in *Table 8, Panel B.*

Table 11: Forecasting Profitability-arbitrage Returns with CoPROF with constraint (decile 5)

The table shows returns to the profitability strategy as a function of lagged CoPROF. We sort CoPROF taken from decile 5 into five different quintiles, while put restrictions on the usage of future information when forming the CoPROF quintiles in the portfolio formation stage. Reported below are the risk-adjusted returns to the profitability strategy in each of the 36 months after portfolio formation, following low to high CoPROF. "5-1" is the difference in monthly returns to the hedged strategy following high vs. low CoPROF. 5% statistical significance is underlined in bold.

Panel A: FF3F-Adjusted Profitability-arbitrage Returns								
Decile 5	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	0.01 %	(0.04)	0.12 %	(0.67)	0.06 %	(0.48)	0.01 %	(0.10)
2	-0.07 %	(-0.52)	-0.20 %	(-1.61)	0.01 %	(0.10)	0.06 %	(0.64)
3	**0.40 %**	**(3.02)**	0.15 %	(1.22)	0.13 %	(1.38)	0.06 %	(0.58)
4	-0.04 %	(-0.36)	-0.07 %	(-0.56)	0.17 %	(1.92)	**0.28 %**	**(3.06)**
5	**0.30 %**	**(3.19)**	0.18 %	(1.65)	**0.22 %**	**(2.69)**	**0.36 %**	**(5.19)**
5-1	0.29 %	(1.39)	0.06 %	(0.29)	0.16 %	(1.08)	**0.35 %**	**(2.74)**

Panel B: FF5F-Adjusted Profitability-arbitrage Returns								
Decile 5	Months 1-6		Months 7-12		Year 2		Year 3	
Rank	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
1	-0.09 %	(-0.47)	-0.19 %	(-0.90)	-0.24 %	(-1.79)	-0.17 %	(-1.47)
2	-0.17 %	(-1.15)	**-0.40 %**	**(-3.12)**	-0.15 %	(-1.46)	-0.13 %	(-1.40)
3	**0.28 %**	**(2.05)**	0.03 %	(0.22)	-0.06 %	(-0.55)	**-0.20 %**	**(-2.01)**
4	-0.18 %	(-1.45)	-0.16 %	(-1.24)	0.05 %	(0.59)	0.07 %	(0.78)
5	0.07 %	(0.77)	-0.01 %	(-0.08)	0.02 %	(0.21)	**0.15 %**	**(2.29)**
5-1	0.16 %	(0.76)	0.18 %	(0.78)	0.26 %	(1.74)	**0.32 %**	**(2.30)**

6 CONCLUSION

In this paper, by testing on the U.S stock market in the period 1964-2010, we identify the potential explanations for the gross profitability anomaly, which is a strong predictor of the cross section of returns. The zero-investment portfolio strategy on long position of the most profitable companies and short position of the least profitable ones delivers a notable excess return regardless of the significant negative loading on the SMB and HML factors. In line with Novy-Marx´s prediction (2013), a value investor can achieve higher returns by controlling for gross profitability, while "*controlling for profitability dramatically increases the performance of value strategies*".

The second part of the paper investigates the effect of arbitrage activity in profitability strategy on abnormal trading profits. Applying the methodology to measure arbitrage activity invented by Lou and Polk (2013) for momentum strategy and then expanded by Huang et al. (2016) for beta strategy, we first show that we successfully replicate the CoBAR measure in the later paper. We then document our own measure of arbitrage activity to profitability-arbitrage strategy, CoPROF, which is computed as the average pairwise partial correlations using 2 months of past daily returns rather than 52 prior weekly returns for all stocks in the highest decile whilst controlling for the Fama-French (1992) three factors.

Specifically, our results suggest that when arbitrage activity is relatively low, abnormal returns on profitability-arbitrage strategies take much longer to materialize, occurring only three years after entering into such as strategy as revealed by the Fama French three-factor. Thus, an arbitrageur would be able to earn 0.27% (t-stat 3.48) by the end of the 3 year holding period. On the other hand, this return could not be realized under the Fama French five-factor model.

In contrast, in times when arbitrage activity is relatively high, abnormal returns appear relatively quickly, within the first six months of the trade as illustrated by both asset pricing models. An investor could earn 0.53% in the first 6 months of the profitability strategy according to the Fama French three-factor model and 0.31% in the first half of the year according to the Fama French five-factor model. Also, these positive abnormal returns appear both in the short and in the long run under FF3F.

Finally, we observe that the abnormal returns on the profitability-arbitrage strategies arrive sooner and stronger following an expansion of arbitrage activity level. We also find it interesting that those returns *cannot* be linked to long run reversal and crash risk as opposed to the momentum and beta strategies.

The first robustness test reveals that the results are sensitive to the various specifications of the deciles based on which the CoPROF variable is computed. This could be an indication of the fact that the methodology is not entirely appropriate for the measurement of the arbitrage activity for profitability strategies. However, despite the methodology's flaws, we still find that abnormal returns are dependent on the level of arbitrage activity. We find a positive relationship between the level of arbitrage activity and abnormal returns across all three specifications of deciles based on which the CoPROF is measured. In the second robustness test, we restrict the use of future information for the calculation of CoPROF quintile, only to find that our main results still hold.

7 APPENDIX

In this section, we present our replication of CoBAR over the period of 1985-2016, and compare our results relative to Huang et al. (2016) paper. In order to construct CoBAR, at the end of each month, all stocks are sorted into deciles based on their pre-ranking market betas calculated using daily returns in the past 12 months, while controlling for illiquidity and non-synchronous trading by including five lags of market excess returns on the right hand side of the OLS regression equation. CoBAR is then estimated as the partial pairwise correlations using the past 52 weekly returns for all stocks in the lowest decile, while controlling for the FF-3. In the Huang et al. paper, the authors focus on the low-beta deciles for the reason that these stocks tend to be "larger, more liquid, and have lower idiosyncratic volatility compared to the highest- beta decile", therefore CoBAR, their measurement of excess comovement will be less affected by asynchronous trading and measurement noise.

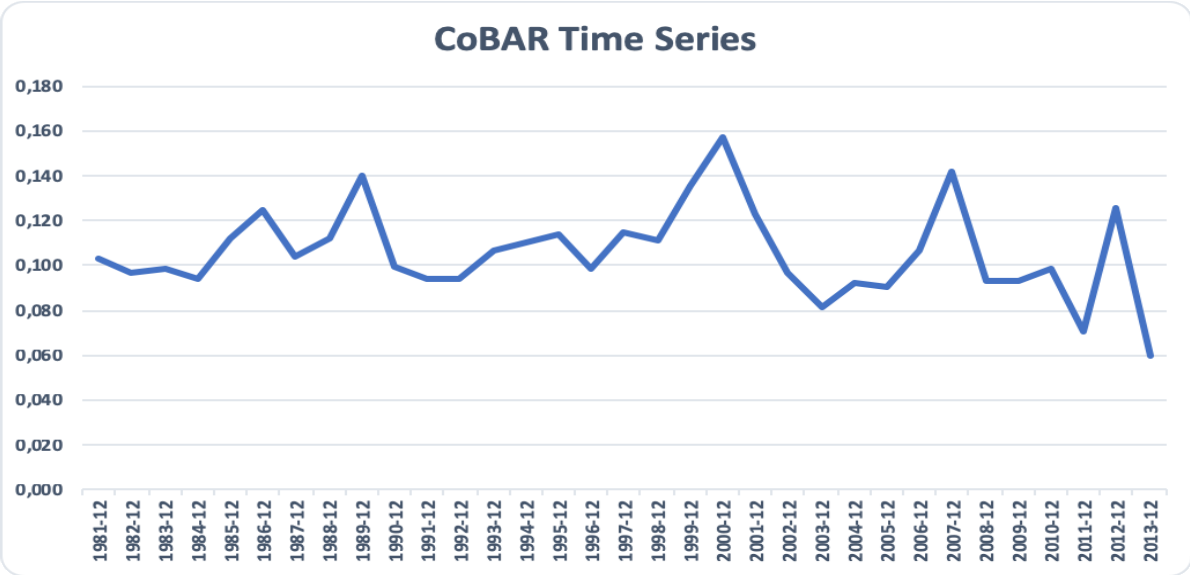
CoBAR shows the degree of beta arbitrage activity. A high value of CoBAR indicates more capital invested in beta-arbitrage strategies, whereas a low value of CoBAR implies a low level of arbitrage activity following these corresponding strategies. When arbitrageurs take long-short position on the beta strategies (ie. long the value-weighted lowest beta stock decile and short the value-weighted highest one), these trading activities can have temporary and simultaneous impacts on all beta stocks price, which leads to return comovement among these stocks. CoBAR is therefore a useful measure to examine the effect of beta-arbitrage activity on its abnormal trading profits.

Figure 2 plots our estimated $CoBAR$ as of the end of each December (Panel A) in comparison to the original CoBAR copied from Huang et al. (2016) study. We do not necessarily expect a clear trend of CoBAR. *Table 12* provides summary statistics of our estimated CoBAR and the original one. Overall, our estimated CoBAR is replicated nearly the same as the original CoBAR, in terms of trend and magnitude. Our CoBAR varies substantially over the period of 1985-2013 with an average value of 0.105 and a standard deviation of 0.021, which is slightly less than the corresponding value of the original CoBAR. It is due to the difference that comes from different time periods used in the original paper (1969-2013) and our paper (1981-2016).

Figure 2: The time-series of CoBAR from 1985 to 2013

The figure portrays the CoBAR time series at the end of December. The blue line illustrates our estimated CoBAR as the excess comovement among low beta stocks over the period from 1985 to 2013 (Panel A), whereas while Panel B shows the original time-series copied from Huang et al. (2016) paper. To measure CoBAR, at the end of each month, all stocks are sorted into deciles based on their lagged-12-month market beta computed using daily returns. CoBAR is then estimated as the average pairwise partial correlation using 52 (non-missing) weekly returns for all stocks in the lowest market beta decile in the portfolio ranking period while controlling for the Fama-French (1992) three factors.

Panel A: The estimated time-series of CoBAR (1985-2013)



Panel B: The original time-series of CoBAR (1970-2013)

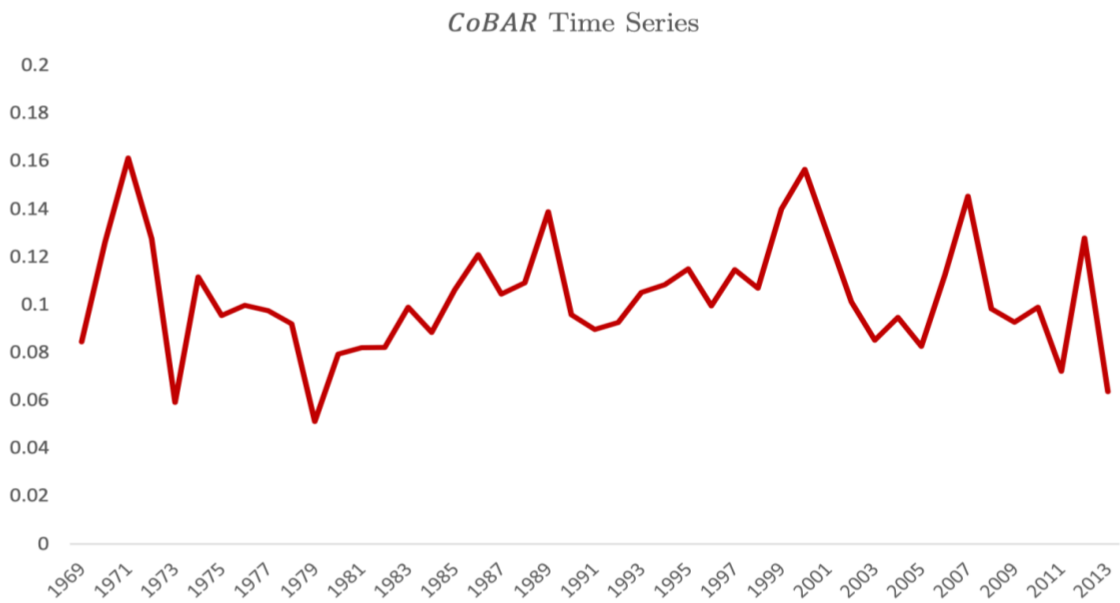


Table 12: Summary statistics of the original and the estimated CoBAR

The table displays the descriptive summary of our estimated CoBAR and the CoBAR from the original paper of Huang al et. (2014), in terms of the average value, standard deviation, maximum and minimum values.

	Estimated CoBAR	**Original CoBAR**
Average	0,105	0,106
Std. Dev.	0,021	0,027
Min	0,043	0,034
Max	0,168	0,202

8 REFERENCES

Baker, M., and Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives,* 21(2)

Ball, R., Gerakos, J., Linnainmaa, J. T. and Nikolaev, V. (2015). Deflating probability. *Journal of Financial Economics,* 117(2), pp. 225-248

Ball, R., Gerakos, J., Linnainmaa, J. T. and Nikolaev, V. (2016). Accruals, Cash Flows, and Operating Profitability in the Cross Section of Stock Returns. *Journal of Financial Economics,* 121(1), pp. 28-45

Barberis, N., and Shleifer, A. (2003). Style Investing. *Journal of Financial Economics,* 68(2), pp. 161-199

Barberis, N., Shleifer, A. and Wurgler, J. (2005). Comovement. *Journal of Financial Economics,* 75(2), pp. 283-317

Barroso, P., Edelen, R. M. and Karehnke, P. (2018). Institutional Crowding and the Moments of Momentum. *SSRN Electronic Journal*. doi:10.2139/ssrn.3045019

Berger, D., and Turtle, H. J. (2012). Cross-sectional Performance and Investor Sentiment in a Multiple Risk Factor Model. *Journal of Banking & Finance*, 36(4), pp. 1107–1121

Chen, L., Novy-Marx, R. and Zhang, L. (2011). An Alternative Three-Factor Model. *SSRN Electronic Journal.* doi:10.2139/ssrn.1418117

Dechow, P. M. (1994). Accounting Earnings and Cash Flows as Measures of Firm Performance: The Role of Accounting Accruals. *Journal of Accounting and Economics* 18(1), pp. 3-42

Fama, E. F. and French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance,* 47(2), pp. 427–465

Fama, E. F. and French, K.R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33, 3-56

Fama, E. F., and French, K. R. (2015). A Five-Factor Asset Pricing Model. *The Journal of Financial Economics,* 116(1), pp. 1-22

Fama, E. F. and MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, 81(3), pp. 607-636

Fong, W. M. and Toh, B. (2014). Investor Sentiment and the MAX Effect. *Journal of Banking & Finance,* 46, pp. 190–201

Frazzini, A. and Pedersen, L. H. (2014). Betting Against Beta. *Journal of Financial Economics,* 111(1), pp. 1-25

Grinblatt, M. and Han, B. (2005). Prospect Theory, Mental Accounting, and Momentum. *The Journal of Financial Economics,* 78(2), pp. 311–339

Greenwood, R. and Shleifer, A. (2014). Expectations of Returns and Expected Returns. *The Review of Financial Studies*, 27(3), pp. 714–746

Hanson, S. G. and Sunderam, A. (2014). The Growth and Limits of Arbitrage: Evidence from Short Interest. *Review of Financial Studies,* 27(4), pp. 1238–1286

Harvey, C. and Siddique, A. (2000). Conditional Skewness in Asset Pricing Tests. *The Journal of Finance,* 55, pp. 1263-1295

Hou, K., Xue C. and Zhang, L. (2015). Digesting Anomalies: An Investment Approach. *The Review of Financial Studies*, 28(3), pp. 650–705

Huang, S., Lou, D. and Polk, C. (2016). The Booms and Busts of Beta Arbitrage. *SSRN Electronic Journal.* doi:10.2139/ssrn.2666910

Kim, J. S., Ryu, D. and Seo, S. W. (2014). Investor Sentiment and Return Predictability of Disagreement. *Journal of Banking & Finance*, 42, pp. 166–178

Kumar, A. and Lee, C. (2006). Retail Investor Sentiment and Return Comovements. *The Journal of Finance*, 61(5), pp. 2451–2486

Landier, A., Simon, G. and Thesmar, D. (2015). The Capacity of Trading Strategies. *SSRN Electronic Journal.* doi:10.2139/ssrn.2585399

Lemperiere, Y., Deremble, C., Nguyen, T. T., Seager, P., Potters, M. and Bouchaud, J. P. (2015). Risk Premia: Asymmetric Tail Risks and Excess Returns. *Quantitative Finance*, 17(1), pp. 1-14

Linnainmaa, J. T. and Roberts, M. R. (2018). The History of the Cross-Section of Stock Returns. *Review of Financial Studies,* 31(7), pp. 2606-2649

Lou, D. and Polk, C. (2013). Comomentum: Inferring Arbitrage Activity from Return Correlations. *SSRN Electronic Journal.* doi:10.2139/ssrn.2023989

Menkveld, A. J. (2014). Crowded Trades: An Overlooked Systemic Risk for Central Clearing Counterparties. *SSRN Electronic Journal.* doi:10.2139/ssrn.2422250

Morck, R., Yeung, B. and Yu, W. (2000). The Information Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movements? *The Journal of Financial Economics*, 58(1-2), pp. 215-260

Novy-Marx, R. (2013). The Other Side of Value: The Gross Profitability Premium. *The Journal of Financial Economics,* 108(1), pp. 1-28

Pojarliev, M. and Richard, M. L. (2011). Detecting Crowded Trades in Currency Funds. *Financial Analysis Journal*, 01, pp. 26–39

Qian, X. (2014). Small Investor Sentiment, Differences of Opinion and Stock Overvaluation. *Journal of Financial Markets*, 19, pp. 219–246

Sloan, R. G. (1996). Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings?. *Accounting Review*, 71, pp. 289-315

Stein, J. C. (2009). Presidential Address: Sophisticated Investors and Market Efficiency. *The Journal of Finance,* 64(4), pp. 1517–1548

Wahal, S. (2019). The Profitability and Investment Premium: Pre-1963 Evidence. *The Journal of Financial Economics,* 131(2), pp. 362-3