11-2017

Dissecting arbitrage costs

F. Y. Eric LAM
Hong Kong Baptist University

Chishen WEI
Singapore Management University, cswei@smu.edu.sg

K. C John WEI
Hong Kong University of Science and Technology

Dissecting Arbitrage Costs[☆]

F.Y. Eric C. Lam
Department of Finance and Decision Sciences
Hong Kong Baptist University
Kowloon Tong, Kowloon, Hong Kong
Tel: 852-3411-5218
fyericcl@hkbu.edu.hk

Chishen Wei
Department of Banking and Finance
Nanyang Technological University
Singapore
Tel: 65-6592-1859
cswei@ntu.edu.sg

K.C. John Wei[*]
School of Accounting and Finance
The Hong Kong Polytechnic University
Hong Hum, Kowloon, Hong Kong
Tel: 852-2766-4953
Email: johnwei@ust.hk

This version: March 2017

[☆]We appreciate the helpful comments and suggestions from Bradley Paye, Raluca Stan (PFMC discussant), Avindhar Subrahmanyam, Yaqiong Yao (SFM discussant), and seminar participants at the Hong Kong Polytechnic University, the 2016 Paris Financial Management Conference (PFMC) at IPAG Business School, the 2016 Conference on the Theories and Practices of Securities and Financial Markets (SFM) at National Sun Yat-sen University, and the 2016 NCTU International Finance Conference at National Chiao Tung University, Taiwan. This paper won the Best Research Award at the 2016 SFM. We thank Alice Cheung for her editorial assistance. We acknowledge financial support from the Research Grants Council of the Hong Kong Special Administrative Region, China (project no. 16503514). All remaining errors are our own.

[*] Corresponding author: K.C. John Wei, School of Accounting and Finance, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. Tel: (852)-2766-4953; Fax: (852)-2330-9845. Email: johnwei@ust.hk.

Dissecting Arbitrage Costs

Abstract

This paper systematically examines the impact of nine popular arbitrage costs measures on cross-sectional mispricing based on ten well-known and robust anomalies. We show that binding arbitrage barriers slowly change over time. In early years with few publications documenting return anomalies, arbitrage costs have tiny impact even though mispricing is present. As anomalies become more widely known, arbitrage costs impact mispricing substantially. Arbitrage risk, ambiguity of fundamental value, round-trip broker's commission plus bid-ask spreads, and stock loan supply are binding on arbitrageurs. Only arbitrage risk is binding if larger cap stocks are emphasized. In recent years when market quality improves and some arbitrageurs become more creative, only round-trip broker's commission plus bid-ask spreads and stock loan supply remain binding on arbitrageurs. If larger cap stocks are emphasized, arbitrage costs do not matter at all because there is no longer mispricing. An empirical arbitrage costs model based on these simple dynamics subsumes annually-varying principal components of arbitrage costs in affecting mispricing. Incorporating our findings into future capital market efficiency research would mitigate type I and II errors in empirical tests applying the limits-to-arbitrage argument.

JEL Classification: G12; G14

Keywords: Arbitrage costs; arbitrage risk; information uncertainty; transaction costs; mispricing

1. Introduction

Proper functioning of and efficiency in the financial markets not only require sufficient capital flows to arbitrageurs (Akbas, Armstrong, Sorescu, and Subrahmanyam, 2016), but also depends on the effectiveness of carrying out arbitrage trades. Heightened arbitrage costs generally lessen incentives to eliminate mispricing in a timely manner desired by perfect capital markets (Shleifer and Vishny, 1997; Gromb and Vayanos, 2010). This limits-to-arbitrage theory, which predicts anomalies to be stronger when arbitrage costs are higher, formulates a main test in capital market efficiency research. An expanding body of empirical studies of anomalies in the cross section of average stock returns has reached a consensus for this argument.[1] However, papers use different measures of arbitrage costs and most of them involve individual anomalies or a set of similar anomalies. Besides, this literature does not address which of the arbitrage costs constitutes the most important barrier to arbitrage activity.

In this paper, we perform a systematic investigation on nine arbitrage costs measures that have strong theoretical motivations and large amounts of data, and are reliable and commonly used in the anomalies literature. These measures are idiosyncratic stock return volatility, cash flow volatility, analyst earnings forecast dispersion, analyst coverage, share price, dollar trading volume, Amihud (2002) illiquidity, effective bid-ask spread, and passive institutional ownership. To begin, we construct an overall arbitrage costs score from the nine measures. We establish the impact of overall arbitrage costs on extent of cross-sectional mispricing. Like the relative valuation approach in Stambaugh, Yu, and Yuan (2012, 2015), mispricing is identified by a misprice score comprised of ten well-known and robust anomalies in the cross section of stock returns. The anomaly

[1] Notable papers include Ali, Hwang, and Trombley (2003), Mashruwala, Rajgopal, and Shevlin (2006), Zhang (2006), Duan, Hu, and McLean (2010), Li and Zhang (2010), McLean (2010), Lam and Wei (2011), Lipson, Mortal, and Schill (2011), Cao and Han (2015), Stambaugh, Yu, and Yuan (2015), Li and Lou (2016), and Yan and Zheng (2017).

variables involved are total asset growth, investment-to-assets ratio, net operating assets, operating accruals, net share issuance, composite share issuance, Ohlson (1980) bankruptcy score, gross profitability, operating profitability, and prior-year stock return. Some of these variables are negatively correlated with future abnormal returns (e.g., total asset growth) while others are positively correlated with future abnormal returns (e.g., gross profitability). Since at least one arbitrage costs measure or anomaly variable is required in constructing a score, our sample size remains large even though we study multiple arbitrage costs measures and anomalies. We sort stocks by mispricing and examine how equal- and value-weighted three-factor alphas (Fama and French, 1993) on short-leg, long-leg, and long-short portfolios vary with overall arbitrage costs.[2] We also estimate interactive cross-sectional regressions to assess the impact of overall arbitrage costs on mispricing.

We split the nine arbitrage costs measures into three groups, namely arbitrage risk (idiosyncratic stock return volatility), information uncertainty (cash flow volatility, analyst earnings forecast dispersion, and analyst coverage), and transaction costs (share price, dollar trading volume, Amihud (2002) illiquidity, effective bid-ask spread, and passive institutional ownership), and we construct a score for each group. Arbitrage risk refers to firm-specific price movements, which are difficult for arbitrageurs to hedge (e.g., Pontiff, 1996, 2006; Cao and Han, 2015). Information uncertainty captures ambiguity of a stock's intrinsic or fundamental value, disagreement in future earnings among market participants, and information asymmetry in the stock's information environment (e.g., Hong, Lim, and Stein, 2000; Zhang, 2006). Transaction costs measure expenses and technical difficulties in entering and exiting arbitrage positions such

[2] Equal- and value-weighted monthly alphas on long-short portfolio based on quintiles of the misprice score are 0.91% (*t*-statistics = 8.61) and 0.66% (*t*-statistics = 5.58), respectively. These anomalous returns are comparable to those (0.69% (*t*-statistics = 6.44) and 0.50% (*t*-statistics = 3.35)) on long-short portfolio based on the fundamental-analysis quintiles Bartram and Grinblatt (2017) construct with market equity and 28 Compustat accounting items.

as round-trip broker's commission plus bid-ask spreads, liquidity, price impact, compensation for liquidity provision, and stock loan supply for short selling (e.g., Nagel, 2005; Novy-Marx and Velikov, 2015). We examine not only the impact of each group on mispricing individually, but also their impact jointly. To perform the joint investigation, we estimate interactive cross-sectional regressions of stock returns on the misprice score, arbitrage risk, the information uncertainty score, and the transaction costs score.

The overall arbitrage costs score, arbitrage risk, the information uncertainty score, and the transaction costs score always equally weight their components. To allow for time variation in relative informativeness of the nine arbitrage costs measures, we perform principal components analysis on the measures annually. The arbitrage costs space statistically stably reduces to three dimensions over time. On average, the first principal component is positively associated with idiosyncratic stock return volatility, cash flow volatility, effective bid-ask spread, and passive institutional ownership, and it is negatively associated with analyst coverage, share price, and dollar trading volume. The second principal component is positively associated with idiosyncratic stock return volatility, cash flow volatility, analyst coverage, dollar trading volume, and effective bid-ask spread, and it is negatively associated with passive institutional ownership. The third principal component is positively associated with analyst earnings forecast dispersion. We examine the impact of each of the three annually-varying principal components on mispricing individually as well as their impact jointly.

Moreover, we examine jointly the impact of the nine arbitrage costs measures on mispricing across three salient periods in the literature of return anomalies. We conveniently refer to these subperiods as the "pre-discovery period", the "golden age of mispricing", and the "attenuation period". The first subperiod is mainly motivated by the start of the anomalies literature. For example, only two of the 97 anomalies on the McLean and Pontiff (2016) list are published during this subperiod. Mispricing is well-known to be strong during the "golden age of mispricing". The third subperiod is motivated by Chordia, Subrahmanyam, and Tong (2014), who show that many anomalies attenuate towards zero around mid 2000s.

Portfolio and regression results from the U.S. sample covering monthly stock returns from 07/1963 to 12/2015 show that overall arbitrage costs explain a large portion of anomalous returns in the cross section. For example, value-weighted alpha on long-short portfolio based on the misprice score drops by 65% when we move from high to low overall arbitrage costs quintile. A much larger part of the alpha reduction comes from short-leg than from long-leg of the trading strategy.

Portfolio results indicate that arbitrage risk, information uncertainty, and transaction costs are individually important. The impact on short leg is larger than on long leg not only for arbitrage risk, but also for information uncertainty and transaction costs. The asymmetric short- versus long-position concern put forward by Stambaugh, Yu, and Yuan (2015) goes beyond the arbitrage costs represented by idiosyncratic stock return volatility.[3] Regression results show that all three arbitrage costs groups individually impact mispricing, except information uncertainty when larger cap stocks are emphasized. When the three arbitrage costs groups are investigated jointly, they all impact mispricing and yet the effect of transaction costs is the weakest. Furthermore, when we emphasize larger cap stocks, information uncertainty and transaction costs are redundant because only arbitrage risk matters.

[3] Almazan, Brown, Carlson, and Chapman (2004) document that mutual funds have low use of actual shorting due to investment policy restrictions. Furthermore, Hong and Sraer (2016) argue that short-sale impediments are prevalent because capital engaged in short selling, such as hedge funds, is tiny compared to capital not actively engaged in short selling, such as mutual funds.

Results from portfolios and regressions indicate that the first and second annually-varying principal components of arbitrage costs impact mispricing in a manner consistent with the limits-to-arbitrage argument. However, the third principal component does not. These findings are the same when the principal components are investigated individually and jointly.

Joint investigation of the nine arbitrage costs measures across subperiods show that the barriers slowing down correction of mispricing change over time. In early years when scientific knowledge on return anomalies is scarce (07/1963 to 06/1983), arbitrage costs have very little impact on mispricing despite its presence.[4] As anomalies become more widely known (07/1983 to 12/2004), mispricing remains strong but arbitrage costs have substantial impact on the extent of mispricing. Arbitrage risk (idiosyncratic stock return volatility), ambiguity of fundamental value (cash flow volatility), round-trip broker's commission plus bid-ask spreads (share price), and stock loan supply (passive institutional ownership) are binding on arbitrageurs. Only arbitrage risk matters if we emphasize larger cap stocks. In recent years when market quality improves and arbitrage strategy becomes more creative (01/2005 to 12/2015), mispricing is much weaker. Only round-trip broker's commission plus bid-ask spreads and stock loan supply remain binding on arbitrageurs. If we emphasize larger cap stocks, arbitrage costs do not matter at all because there is no longer mispricing.

Although the slow time variations in binding arbitrage costs we uncover might look simple, they are in fact practical: an empirical arbitrage costs model based on these simple dynamics subsumes the first and second annually-varying principal components of arbitrage costs in affecting mispricing. The subperiod findings also indicate that the joint impact of three arbitrage

[4] A change of variance analysis during the 1980s detects a breakpoint at 05/1983 on the time series of value-weighted alphas on long-short portfolio based on the misprice score. Limited by data availability, the principal components of arbitrage costs begin at 06/1983. We break at 06/1983 in order to compare an empirically motivated arbitrage costs model with the statistically oriented annually-varying principal components of arbitrage costs in affecting mispricing.

costs groups on mispricing we document from the entire sample period are specific to the "golden age of mispricing". The impact of arbitrage risk and the impact information uncertainty, mainly in the form of ambiguity of fundamental value, have largely disappeared while transaction costs, mainly in terms of round-trip broker's commission plus bid-ask spreads and stock loan supply, remain binding on arbitrageurs in recent years. Ambiguity of fundamental value might have become less important in recent years because young ventures from the dot-com period have aged. Our findings also echo Ljungqvist and Qian (2016), who report that increasing number of creative arbitrageurs recently practice an information disclosure strategy designed to circumvent noise trader risk.[5]

Our empirical results fit into the literature on limits to arbitrage and anomalies. Early work establishes positive correlations between arbitrage costs and individual anomalies or a set of similar anomalies. For example, Ali, Hwang, and Trombley (2003) study the book-to-market anomaly, Mashruwala, Rajgopal, and Shevlin (2006) study the accruals anomaly, Zhang (2006) study price momentum, Duan, Hu, and McLean (2010) study the short interest anomaly, Li and Zhang (2010) study capital investment anomalies, McLean (2010) study price momentum and reversal, Lam and Wei (2011) and Lipson, Mortal, and Schill (2011) study the asset growth anomaly, and Li and Lou (2016) study the cash holdings anomaly. Although emerging papers go further to involve multiple anomalies (e.g., Cao and Han, 2015; Stambaugh, Yu, and Yuan, 2015; Yan and Zheng, 2017), they either focus only on arbitrage risk represented by idiosyncratic stock return volatility or study some arbitrage costs measures that are quite arbitrarily chosen. Even though this literature reaches an agreement that limits to arbitrage are crucial for mispricing, an

[5] The arbitrage strategy begins by shorting an overpriced stock. Instead of hedging the position and quietly wait for price convergence, arbitrageurs actively release private information on overvaluation to persuade existing shareholders to sell their holdings in order to accelerate price discovery and prevent adverse price movements (e.g., Delong, Shleifer, Summers, and Waldmann, 1990).

open question remains: which of the various arbitrage costs constitute the most crucial arbitrage barrier? This question is important because empirical tests using arbitrarily chosen measures might produce misleading conclusions.

We are the first to systematically study an extensive list of popular arbitrage costs measures and anomalies.[6] Like other papers, our extensive results largely support the limits-to-arbitrage story that mispricing is stronger when arbitrage costs are higher. Unlike other papers, we are able to sort out which of the arbitrage costs constitute the most important arbitrage barriers by pitting arbitrage costs variables against one another and examine their impact on mispricing.

Our findings have fundamental implications for empirical capital market efficiency research that applies the limits-to-arbitrage argument in analyzing anomalies. First, using an arbitrary basket of arbitrage costs measures would overstate the amount of evidence for the mispricing explanation for an anomaly or against the null of rational pricing, and it also provides mistaken interpretations. For example, arbitrage risk is the only arbitrage costs that matter for larger cap stocks. Due to correlations between idiosyncratic stock return volatility, share price, dollar trading volume, illiquidity, and effective bid-ask spread, supportive evidence obtained from these transaction costs measures and value-weighted returns are redundant. These results also lead to the false conclusion that transaction costs themselves are important for an anomaly and large firms.

Second, using conventional sample of monthly returns starting from 07/1963 would overstate the scope of evidence for the mispricing explanation for an anomaly or against the null of rational pricing. For example, arbitrage risk affects anomalous returns on larger cap stocks from 07/1983 to 12/2004. A study finding supportive evidence from arbitrage risk and value-weighted returns

[6] Our objective is to dissect which arbitrage costs matter for mispricing. We do not aim to test whether limits to arbitrage completely explain cross-sectional anomalies.

for the period 07/1963 to 12/2015 provides the wrong perception that the evidence is based on more than 53 years of history even though the results are significant for just about 22 years.

Third, using ineffective arbitrage costs measures would ran a chance of falsely declining the mispricing explanation for an anomaly in favor of the null of rational pricing. For example, using equal-weighted returns and only arbitrage risk for a period starting around mid 2000s might not produce any evidence for the mispricing explanation but using round-trip broker's commission plus bid-ask spreads and loan stock supply might.

We proceed as follows. The next section reviews the nine arbitrage costs measures. Section 3 describes the misprice score. Section 4 describes our data and presents summary statistics. Section 5 reports empirical results. Section 6 summarizes the study and provides concluding remarks.

2. Arbitrage Costs

Abnormal profit in a mispriced stock attracts arbitrageurs to correct the price towards fundamental value. When arbitrage costs binding on arbitrageurs are higher, correction of mispricing takes longer hence a stock return anomaly should be more pronounced. The contribution of our analysis is to sort out which of the various arbitrage costs constitute the most important arbitrage barrier. The set of firm-level arbitrage costs measures we analyze is determined by five criteria. First, the measures must have strong theoretical foundations. Second, the literature has shown or it is obvious that the measures adequately cover the underlying arbitrage costs concepts. Third, the literature has documented positive correlations between the measures and some anomalies. Fourth, the measures are somewhat but not extremely correlated. Finally, the measures must have large amounts of data for effective statistical inferences. We discuss the shortlist of nine arbitrage costs measures below.[7]

2.1 Arbitrage risk

Arbitrage is deemed risky when risk-adverse arbitrageurs cannot entirely diversify or hedge away the risk of an arbitrage position they intend to hold. As arbitrageurs are typically under-diversified and it is difficult to locate the perfect substitutes required to hedge an arbitrage trade, idiosyncratic stock return volatility (IVOL) is likely to be a serious concern when they arbitrage mispriced stocks (e.g., Pontiff, 1996; Wurgler and Zhuravskaya, 2002). Pontiff (2006) shows that arbitrageurs prefer to hold less of the stocks that carry higher idiosyncratic stock return volatility.[8]

Ali, Hwang, and Trombley (2003), Mashruwala, Rajgopal, and Shevlin (2006), Duan, Hu, and McLean (2010) and McLean (2010) show that the negative relations between abnormal returns and prior book-to-market equity ratio, accounting accruals, short interests, and three- to five-year stock returns, respectively, are stronger when IVOL is higher. Li and Zhang (2010) find that the negative relation between capital investment and future returns is stronger when IVOL is higher. Similarly, Lam and Wei (2011), and Lipson, Mortal, and Schill (2011) find that the negative relation between total asset growth and future returns is stronger when IVOL is higher. Furthermore, Li and Lou (2016) find that the positive relation between cash holdings and future returns is stronger when IVOL is higher. We use 36 months of stock returns prior to the end of June of year *t* to estimate the market model and use the standard deviation of residuals to measure

[7] Some arbitrage costs concepts such as synchronization risk (e.g., Abreu and Brunnermeirer, 2003) is not included here because there is no reliable measure that fits our criteria.

[8] A strand of research documents that idiosyncratic stock return volatility and future stock returns are negatively associated (e.g., Ang, Hodrick, Xing, and Zhang, 2006, 2009). Therefore, exposure to higher idiosyncratic stock return volatility is not necessary compensated with a risk premium.

IVOL.[9] To standardize the range across arbitrage costs measures, we annually assign independently sorted decile rankings to firms based on the measures. The sorting order assigns higher arbitrage costs to higher ranking. Thus, arbitrage risk (ARBRISK) of a stock is the decile ranking of IVOL.

2.2 *Information uncertainty*

There are three typical layers of information uncertainty. First, when a firm's cash flow is more volatile, its fundamental value is more ambiguous. The ambiguity of fundamental value reduces the precision of arbitrageurs in identifying mispricing, thereby, they would be less willing to arbitrage. Zhang (2006) shows that the post-analyst forecast revision price drift and price momentum are stronger when cash flow volatility (CFVOL) is higher. We use the standard deviation of asset-scaled cash flow from operations over the previous five fiscal years to measure CFVOL.

Second, when market's opinions about a firm's future earnings are more diverse, arbitrageurs are less confident in their valuations. Again, arbitrageurs would be more reluctant to engage to arbitrage because they are more uncertain about whether the stock is mispriced or not. Analyst earnings forecast dispersion (DISP) is associated with disagreement about future earnings among market participants (e.g., Diether, Malloy, and Scherbina, 2002; Zhang, 2006; Cen, Wei, and Yang, 2016). Zhang (2006) shows that price momentum is stronger when DISP is higher. We use the standard deviation of one-year-ahead analyst earnings forecasts scaled by closing stock price at the end of June of year *t* as to measure DISP.

[9] Our findings remain similar when we use past weekly returns to estimate the market model or use residuals from the Fama and French (1993) three-factor model.

Third, arbitrageurs would be more reluctant to arbitrage when they possess less information about a firm. Hong, Lim, and Stein (2000) interpret higher analyst coverage (COV) as lower information asymmetry. Gleason and Lee (2003) show that the post-analyst forecast revision price drift is stronger when COV is lower. Besides, Zhang (2006) shows that the price momentum is more pronounced when COV is lower. We use the number of analysts following the stock at the end of June of year *t* to measure COV.

We construct an information uncertainty score (INFO-score) that captures the three layers of information uncertainty. Each year we independently sort stocks into deciles by CFVOL and DISP in ascending order and COV in descending order. INFO-score of a stock is the average of these three decile rankings, whichever is available.

2.3 Transaction costs

The costs and technical difficulties of establishing and exiting an arbitrage position would deter arbitrage activity. High transaction costs can turn trading against mispricing unprofitable. Arbitrage capital would stay on the sidelines when it is inflexible to move capital in and out of an arbitrage trade. There are five typical aspects of transaction costs.

First, Bhardwaj and Brooks (1992) show that round-trip broker's commission and bid-ask spreads are inversely related to share price (PRICE). We use the closing share price at the end of June of year *t* to measure PRICE. Second, Bhushan (1994) shows that the time required to fill an order or to trade a large block of shares is inversely related to dollar trading volume (DVOL). We use the average daily dollar trading volume, which is closing share price multiplied by the trading day's share trading volume, over the year prior to the end of June of year *t* to measure DVOL. Third, the impact of order flow on stock price restricts the amount of capital that can be invested in and liquidated from the stock at a specific price. We measure price impact with Amihud (2002) illiquidity (ILLIQ), which is the average of absolute value of daily returns scaled by the trading day's dollar trading volume over the year prior to the end of June of year *t*.

Fourth, effective bid-ask spread (BIDASK) refers to the trading expenses for compensating dealers for making markets and providing liquidity. We use the average daily closing bid-ask spread over the year prior to the end of June of year *t* to measure BIDASK. Fifth, it is harder to borrow shares of stocks that have lower passive institutional ownership (IO) because these stocks have low stock loan supply (D'Avolio, 2002; Nagel, 2005). These stocks are also more exposed to the risk of short squeeze (Dechow, Hutton, Meulbroek, and Sloan, 2001). Following Nagel (2005), we measure potential lending supply as the percentage of outstanding shares held by Vanguard 500 index fund and Dimensional Fund Advisors at the end of June of year *t*.[10] Lam and Wei show that the asset growth anomaly is stronger when these measures signal higher transaction costs.

We construct a transaction costs score (TCOST-score) that captures the five aspects of transaction costs. Each year we independently sort stocks into deciles by PRICE, DVOL, and IO in descending order and BIDASK and ILLIQ in ascending order. TCOST-score of a stock is the average of these five decile rankings, whichever is available.

2.4 Overall arbitrage costs

We construct an overall arbitrage costs score (ARBCOST-score) of a stock by averaging the annual decile rankings of nine arbitrage costs measures, whichever is available. We individually examine the impact of ARBCOST-score, ARBRISK, INFO-score, and TCOST-score on cross-

[10] Vanguard 500 index fund and Dimensional Fund Advisors are large index and passive mutual fund families that provide most of the stock loan supply. Our findings are similar when we include all institutional shareholders.

sectional mispricing. We then examine the impact of ARBRISK, INFO-score, and TCOST-score jointly.

2.5 Time-varying arbitrage costs

The handy arbitrage costs groups might be too restrictive because they constantly equal weight a fixed set of components over time. Alternatively, we perform principal component analysis on the nine arbitrage costs measures each year. The statistical combinations of arbitrage costs measures of a stock can change over time, depending on the time variation in the measures' empirical cross-sectional distributions. After we individually examine the impact of annually-varying principal components of arbitrage costs on cross-sectional mispricing, we examine their impact jointly.

Furthermore, we examine jointly the impact of the nine arbitrage costs measures on cross-sectional mispricing across three subperiods we identify in the return anomalies literature as follows. The capital asset pricing model of Sharpe (1964) and Lintner (1965), together with efficient capital markets (e.g., Fama, 1965), are the standards in the asset pricing literature and research publications documenting return anomalies are scarce until early 1980s. For example, only two of the 97 anomalies McLean and Pontiff (2016) examine are published before 1980.[11] Seven of the 97 anomalies are published during the 1980s.[12] Change of variance analysis using the binary segmentation of Scott and Knott (1974) on the time series (01/1980 to 12/1989) of monthly value-weighted alphas of long-short portfolio based on the misprice score detects a change point

[11] The publications are Blume and Husic (1973) on stock price and Langetieg (1978) on existence of prior acquisition.
[12] The publications are Banz (1981) on firm size, Barry and Brown (1984) on firm age, Foster, Olsen, and Shevlin (1984) on earnings surprise, Debondt and Thaler (1985) on long term reversal, Amihud and Mendelson (1986) on bid-ask spread, Bhandari (1988) on financial leverage, and Jegadeesh (1990) on short term reversal.

at 05/1983.[13] Due to data availability, the principal components of arbitrage costs begin at 06/1983. We therefore break at 06/1983 so later we can pit an arbitrage costs model motivated by the subperiod findings against the statistically oriented annually-varying principal components of arbitrage costs. We refer to this first subperiod from 07/1963 to 06/1983 as the "pre-discovery period". Chordia, Subrahmanyam, and Tong (2014) document that many anomalies attenuate towards zero around mid 2000's. Chow test on the time series (07/1983 to 12/2015) of monthly value-weighted alphas of long-short portfolio based on the misprice score indicates that the breakpoint 12/2004 is significant at the 5% level. For this reason, we end the second subperiod at the middle of 2000s, i.e., 12/2004. As mispricing is well-known to be strong during this subperiod, we refer to it as the "golden age of mispricing". Finally, we refer to the third subperiod from 01/2005 to 12/2015 as the "attenuation period" because mispricing turns weaker in general.

3. The Mispricing Score

At the end of June of year *t*, we independently sort stocks into 10 anomaly deciles in descending order of total asset growth (TAG), investment-to-assets ratio (I/A), net operating assets (NOA), operating accruals (OA), net share issuance (NSI), composite stock issuance (CSI), and Ohlson (1980) bankruptcy score (O-score) and in ascending order of gross profitability (GP), operating profitability (OP), and prior-year stock return (PRET). We then combine the trading strategies into a comprehensive one. The misprice score (MISPRICE-score) of a stock is the average of these anomaly rankings, whichever is available. Stocks with higher (lower) MISPRICE-score are expected to have higher (lower) abnormal stock returns from the sorting date to the end

[13] This technique applies a single change point method sequentially and, to avoid overfitting, minimizes a cost function that penalizes larger number of breakpoints.

of June of year t+1. Thus, high (low) MISPRICE-score proxies for relative underpricing (overpricing).

We briefly review the 10 anomalies below.

(1) TAG: Cooper, Gulen, and Schill (2008), Lipson, Mortal, and Schill (2011), Titman, Wei, and Xie (2013), and Watanabe, Xu, Yao, Yu (2013) show that firms with higher growth in total assets earn lower future abnormal returns.

(2) I/A: Titman, Wei, and Xie (2004), Lyandres, Sun, and Zhang (2008), Xing (2008), and Polk and Sapienza (2009) show that firms with higher capital investment earn lower future abnormal returns.

(3) NOA: Hirshleifer, Hou, Teoh, and Zhang (2004) show that firms with higher net operating assets earn lower future abnormal returns.

(4) OA: Sloan (1996) shows that firms with higher operating accruals earn lower future abnormal returns.

(5) NSI: Pontiff and Woodgate (2008) show that firms issuing more new shares on net earn lower future abnormal returns.

(6) CEI: Daniel and Titman (2006) and Fama and French (2006) show that firms with higher composite equity issuance earn lower future abnormal returns.

(7) O-score: Dichev (1998), Griffin and Lemmon (2002), Campbell, Hilscher, and Szilsagyi (2008) show that firms with higher bankruptcy likelihood earn lower future abnormal returns.

(8) GP: Novy-Marx (2013) and Sun, Wei, and Xie (2016) show that firms with higher gross profitability earn higher future abnormal returns.

(9) OP: Ball, Gerakos, Linnainmaa, Nikolaev (2015) and Fama and French (2015, 2016) show

that firms with higher operating profitability earn higher future abnormal returns.

(10) PRET: Jegadeesh and Titman (1993, 2000) and others show that firms with higher stock returns during the past three to 12 months earn higher future abnormal returns in the following three to 12 months.

4. Data and Summary Statistics

Our sample contains firms listed on the NYSE, AMEX, and NASDAQ. Annual financial statements are obtained from Compustat. Monthly and daily stock market data are taken from the Center for Research in Security Prices (CRSP). Financial analyst data are obtained from I/B/E/S. Institutional holdings records are taken from Thomson Reuters (13f) Institutional Holdings.[14] Following Fama and French (1992, 1993), we only include non-financial common stocks. We merge monthly stock returns from the end of June of year t to the end of June of year t+1 with financial statements for fiscal year t–1 and firm attributes observed at the end of June of year t. We further require firms to have valid CAPM beta, market equity, and book-to-market equity ratio as well as stock prices being higher than five dollars at the end of June. We use delisting returns to reduce survivorship bias.[15] The sample period covers fiscal year 1962 to fiscal year 2014. The monthly holding period returns extend from 07/1963 to 12/2015. Passive institutional ownership starts at 1981.[16] Analyst forecast dispersion and coverage begin from 1983. The annual principal component analysis begins from 1983 because all nine arbitrage costs measures are required.[17]

---

[14] While information from Data Explorers is used in some recent studies, we are unable to use it in ours because it only covers a five-year period that coincides with a time of high market turmoil.

[15] Shumway (1997) suggests that the returns of stocks delisted for poor performance (delisting codes 500 and 520 to 584) are usually unavailable. Following Shumway and Warther (1999), we use the delisting return when the return is missing for a CRSP month date. When the delisting return is not available, we use -30% for poor performance delisting and 0% for other cases. The results are almost the same when we use raw returns of -65% or -100% for poor performance delisting, which suggests that poor performance delisting plays a negligible role in our findings.

[16] The ownership records of Vanguard 500 index fund and Dimensional Fund Advisors begin from 1981.

[17] Although I/B/E/S provides data that begin from 1976, analyst related variables start at 1983 since I/B/E/S has limited

Panel A of Table 1 presents the mean, standard deviation, minimum, 25th/50th/75th percentiles, and maximum of arbitrage costs and anomaly variables. We only present the first, second, and third principal components (PC1, PC2, and PC3) as they are the only ones that carry eigenvalues greater than or equal to one. We also present the statistics of CAPM beta (BETA), market equity (ME), and book-to-market equity ratio (B/M), which are used as controls in our Fama and MacBeth (1973) regressions. The mean values of ARBCOST-score, ARBRISK, INFO-score, and TCOST-score are 5.498, 5.500, 5.491, and 5.503, respectively. The corresponding median values are 5.530, 5.708, 5.588, and 5.434. These arbitrage costs variables do not seem to be skewed. Their standard deviations are 1.420, 2.872, 1.331, and 1.689, suggesting that they have rather similar variations. ARBRISK has a slightly higher variation, which is consistent with the slightly wider range it has. The mean, median, and standard deviation of MISPRICE-score, respectively, are 5.501, 5.442, and 1.273. Similarly, the misprice score does not seem to be skewed and it possess meaningful variation. As the three principal components take rather different values from those of ARBCOST-score, ARBRISK, INFO-score, and TCOST-score, we also use independently sorted annual deciles of the principal components in Fama and MacBeth (1973) regressions to ensure fair comparisons.

[Table 1 here]

Panel B of Table 1 presents the sample correlations between misprice score and arbitrage costs variables. First, MISPRICE-score has low correlations with arbitrage costs variables. For example, the correlations of MISPRICE-score with ARBCOST-score, ARBRISK, INFO-score, TCOST-score, PC1, PC2, and PC3, respectively, are -6%, -8%, -4%, -6%, -15%, -5%, and 7%. The correlations of MISPRICE-score with the nine arbitrage costs measures range from -11% (BIDASK) to 11% (PRICE). These indicate that the arbitrage costs variables we develop are rather

inclusion of stocks prior to 1983 (e.g., Cen, Wei, and Yang, 2016).

orthogonal to the misprice score. Thus, our findings from double-sorted portfolios are not likely driven by the arbitrage costs sorts simply refining the misprice score sort.

ARBCOST-score exhibits high correlations with ARBRISK (73%), INFO-score (67%), and TCOST-score (93%). This is not surprising given ARBCOST-score is made to be an overall arbitrage costs measure. The correlations between ARBRISK, INFO-score, and TCOST-score are not as high and range from 42% to 56%, suggesting that there is some commonality among the arbitrage costs group scores.

By design, PC1, PC2, and PC3 are uncorrelated. PC1 is highly correlated with ARBCOST-score (88%). PC1 is also correlated with ARBRISK (76%), INFO-score (66%), and TCOST-score (81%). As such, PC1 highly resembles an overall arbitrage costs measure, but PC1 contains less information on transaction costs than ARBCOST-score. As PC1 already picks up a substantial amount of arbitrage costs information, PC2 is much less correlated with ARBCOST-score (11%). PC2 is correlated with ARBRISK (35%) and INFO-score (18%) but not much with TCOST-score (-5%). Essentially, PC2 picks up the residual information about arbitrage risk and information uncertainty. PC3 has very low correlations with ARBCOST-score (0%), ARBRISK (5%), INFO-score (-7%), and TCOST-score (3%). Comparatively, PC3 is not as informative.

The correlations between nine arbitrage costs measures range from 29% (IVOL and PRICE) to 56% (IVOL and BIDASK). Correlations that are larger than 30% in absolute values include the one between IVOL and CFVOL (38%), the one between COV and DVOL (44%), and the one between PRICE and DVOL (38%). The nine arbitrage costs measures are somewhat but not extremely correlated. It is necessary to pit the measures against each other when examining their impact on cross-sectional mispricing.

5. Empirical Results

Table 2 reports Fama and French (1993) three-factor alphas on value-weighted portfolios sorted at the end of June of year *t* by quintiles of MISPRICE-score or each component anomaly variable. The portfolios are held until the end of June of year *t*+1. The long-short strategy (Long-short) simultaneously buys the high MISPRICE-score quintile (Underpriced) and shorts the low MISPRICE-score quintile (Overpriced). For brevity, we do not report results from equal-weighted portfolios as all of them are similar to, and for various occasions being stronger than, those from value-weighted portfolios.

[Table 2 here]

The long-short strategy based on MISPRICE-score generates an alpha of 0.66% (*t*-statistic = 5.58) per month. The strategies based on component anomaly variables generate alphas from 0.12% (*t*-statistic = 1.25) to 0.56% (*t*-statistic = 4.86). The long-short alpha based on MISPRICE-score is higher but less volatility than those based on component anomaly variables. MISPRICE-score contains more precise information about cross-sectional mispricing as it combines information and diversifies noise from individual anomalies away.

5.1 Arbitrage costs scores

5.1.1 ARBCOST-score

Table 3 reports alphas on portfolios independently sorted at the end of June of year *t* by quintiles of MISPRICE-score and quintiles of ARBCOST-score. The portfolios are held until the end of June of year *t*+1. For each ARBCOST-score quintile, the long-short strategy (Long-short) simultaneously buys the high MISPRICE-score quintile (Underpriced) and shorts the low MISPRICE-score quintile (Overpriced). For each MISPRICE-score quintile, the high-low strategy

(High-low) simultaneously buys the high ARBCOST-score quintile (High costs) and shorts the low ARBCOST-score quintile (Low costs).

[Table 3 here]

In the high ARBCOST-score quintile, overpriced portfolio generates an alpha of -0.92% (*t*-statistic = -7.14) while underpriced portfolio generates an alpha of 0.35% (*t*-statistic = 2.47). The long-short strategy generates an alpha of 1.27% (*t*-statistic = 6.52). 72% (0.92%/1.27%) of the long-short alpha comes from the short leg, suggesting that high overall arbitrage costs impede shorting overpriced stocks more than longing underpriced stocks. The long-short alpha declines from the high ARBCOST-score quintile to the low ARBCOST-score quintile. In the low ARBCOST-score quintile, overpriced portfolio generates an alpha of -0.19% (*t*-statistic = -2.28) while underpriced portfolio generates an alpha of 0.25% (*t*-statistic = 3.73). The long-short strategy generates an alpha of 0.44% (*t*-statistic = 4.38). The difference in long-short alpha between high and low ARBCOST-score quintiles reported at the bottom right corner cell is 0.83% (*t*-statistic = 4.50). Low overall arbitrage costs significantly reduce mispricing in the cross section. In the Overpriced quintile, high costs portfolio generates an alpha of -0.92% (*t*-statistic = -7.14) while low costs portfolio generates an alpha of -0.19% (*t*-statistic = -2.28). The high-low strategy generates an alpha of -0.73% (*t*-statistic = -4.74). In the Underpriced quintile, high costs portfolio generates an alpha of 0.35% (*t*-statistic = 2.47) while low costs portfolio generates an alpha of 0.25% (*t*-statistic = 3.73). The high-low strategy only generates an alpha of 0.10% (*t*-statistic = 0.73). High overall arbitrage costs impede shorting overpriced stocks more significantly than longing underpriced stocks.

We analyze the impact of overall arbitrage costs in the time series. Figure 1 shows the 5-year moving average of annualized long-short alphas for low ARBCOST-score quintile

(ARBCOST-score1) and high ARBCOST-score quintile (ARBCOST-score5).[18] The long-short strategy in ARBCOST-score5 outperforms the strategy in ARBCOST-score1 except in 1967, 1968, and 1969. These are consistent with the results in Table 3.

[Figure 1 here]

The long-short strategy in ARBCOST-score5 generates positive average alphas every year. The larger average alphas occur between early 1980's and mid 2000's. After 2005, the average alphas become much smaller. The impact of overall arbitrage costs also weakens in recent years, which is consistent with the recent improvements in market quality and financial innovations (e.g., Chordia, Subrahmanyam, and Tong, 2014; Ljungqvist and Qian, 2016). The long-short strategy in ARBCOST-score1 generates positive but comparatively small average alphas until 2005 and the average alphas become smaller and more volatile afterwards.

Figure 2 shows the 5-year moving average of annualized high-low alphas for low MISPRICE-score quintile (OVERPRICE) and high MISPRICE-score quintile (UNDERPRICE). The high-low strategy in UNDERPRICE generates positive average alphas except for 1972, 2004, 2007, 2009, 2010, 2011, and 2012. In 86% (42/49) of the time, high overall arbitrage costs increase the alpha of underpriced portfolio. However, some of the average alphas are rather small (e.g., the average alphas for 1969, 1971, 1986, 1987, 2003, 2006, 2008, and 2014). For the 31 years between 1973 and 2003, the average alphas are continuously positive, but the streak breaks in recent years. From 2004 to 2015, half of the average alphas, namely those for 2004, 2007, 2009, 2010, 2011, and 2012, are negative. It seems that recent market improvements substantially reduce the impact of overall arbitrage costs on underpriced stocks.

[Figure 2 here]

---

[18] The moving average reduces noise and produces a smoother graph. The patterns are similar without the moving average.

The high-low strategy in OVERPRICE generates consistently negative average alphas since 1978. In 88% (43/49) of the time, high overall arbitrage costs reduce the alpha of overpriced portfolio. The exceptions cluster in the earlier years 1967, 1968, 1969, 1975, 1976, and 1977. The magnitudes of average high-low alphas in OVERPRICE are generally larger than in UNDERPRICE, which echoes the concept of arbitrage asymmetry. The average high-low alphas in OVERPRICE somewhat drop in recent years, but not as much as the average high-low alphas in UNDERPRICE. Recent market improvements seem to influence the impact of overall arbitrage costs on underpriced stocks more than on overpriced stocks.

5.1.2 ARBRISK, INFO-score, and TCOST-score

Table 4 repeats the portfolio analysis in Table 3 with ARBCOST-score replaced by ARBRISK (Panel A), INFO-score (Panel B), or TCOST-score (Panel C). The long-short alphas are 1.24% (*t*-statistic = 5.36), 1.20% (*t*-statistic = 5.02), and 1.16% (*t*-statistic = 7.41), respectively, in high ARBRISK, INFO-score, and TCOST-score quintiles. More than half of the long-short alpha comes from the short leg, suggesting that high overall arbitrage costs impede shorting overpriced stocks more than longing underpriced stocks, for all three arbitrage costs groups.

[Table 4 here]

The long-short alphas decline from the high arbitrage costs group quintile to the low arbitrage costs group quintile. The long-short alphas are 0.32% (*t*-statistic = 3.16), 0.65% (*t*-statistic = 5.07), and 0.56% (*t*-statistic = 4.54), respectively, in low ARBRISK, INFO-score, and TCOST-score quintiles. The differences in long-short alpha between high and low arbitrage costs group quintiles are 0.92% (*t*-statistic = 4.08), 0.55% (*t*-statistic = 2.26), and 0.61% (*t*-statistic = 3.84), respectively, for arbitrage risk, information uncertainty, and transaction costs. For all groupings, low arbitrage costs significantly reduce mispricing in the cross section. In the Overpriced quintile, the high-low alphas are -0.84% (*t*-statistic = -4.95), -0.60% (*t*-statistic = -3.41), and -0.48% (*t*-statistic = -4.13), respectively, for arbitrage risk, information uncertainty, and transaction cost. In the Underpriced quintile, the high-low alphas are 0.08% (*t*-statistic = 0.34), -0.05% (*t*-statistic = -0.33), and 0.13% (*t*-statistic = 0.94), respectively, for arbitrage risk, information uncertainty, and transaction costs. For all groupings, high arbitrage costs impede shorting overpriced stocks more significantly than longing underpriced stocks. The impact on mispricing seems to vary across the arbitrage costs groups. For example, the difference in long-short alpha between high and low arbitrage risk is 0.92%, which is higher than the difference based on information uncertainty (0.55%), transaction costs (0.61%), and even overall arbitrage costs (0.83%).

5.1.3 Multivariate Fama-MacBeth regressions

Table 5 reports regression results for investigating overall arbitrage costs, arbitrage risk, information uncertainty, and transaction costs individually and jointly. To ease the presentation, we multiply slopes by 10. Monthly cross-sectional regressions in Panel A are estimated by OLS. Column (1) reports slope estimates of the baseline regression that includes MISPRICE-score and controls. The MISPRICE-score slope is 2.183 (*t*-statistic = 9.28), which echoes the positive long-short alpha in portfolio analysis. Column (2) adds ARBCOST-score and the interaction between MISPRICE-score and ARBCOST-score. The MISPRICE-score slope is -0.556 (*t*-statistic = -1.09) and the interactive slope is 0.476 (*t*-statistic = 4.91). Consistent with double portfolio sorts, firms with higher MISPRICE-score generate higher adjusted stock returns that are even higher when ARBCOST-score is higher. When ARBCOST-score is replaced by ARBRISK in column (3), INFO-score in column (4), and TCOST-score in column (5), the interactive slopes, respectively, are 0.264 (*t*-statistic = 4.94), 0.383 (*t*-statistic = 4.50), and 0.332 (*t*-statistic = 4.40). Consistent with double portfolios sorts, each arbitrage costs group impacts mispricing according to the limits-to-arbitrage argument.

[Table 5 here]

The key regression reported in column (6) simultaneously includes MISPRICE-score, ARBRISK, INFO-score, TCOST-score, and the interactions between MISPRICE-score and arbitrage costs group variables. The MISPRICE-score slope is -0.559 (*t*-statistic = -1.06). The interactive slopes for ARBRISK, INFO-score, and TCOST-score, respectively, are 0.177 (*t*-statistic = 3.22), 0.173 (*t*-statistic = 2.22), and 0.125 (*t*-statistic = 1.68). The arbitrage costs groups jointly impact cross-sectional mispricing according to the limits-to-arbitrage argument. However, the impact of transaction costs is comparatively the weakest.

Monthly cross-sectional regressions in Panel B are estimated by WLS that weighs an observation by its prior-month market equity. These regressions emphasize on larger cap stocks. Column (1) reports slope estimates of the baseline regression. The MISPRICE-score slope is 1.434 (*t*-statistic = 5.38). Column (2) adds ARBCOST-score and the interaction between MISPRICE-score and ARBCOST-score. The MISPRICE-score slope is 0.127 (*t*-statistic = 0.20) and the interactive slope is 0.355 (*t*-statistic = 2.39). When ARBCOST-score is replaced by ARBRISK in column (3), INFO-score in column (4), and TCOST-score in column (5), the interactive slopes are, respectively, 0.245 (*t*-statistic = 2.80), 0.038 (*t*-statistic = 0.26), and 0.375 (*t*-statistic = 2.58). Arbitrage risk and transaction costs impact mispricing according to the limits-to-arbitrage argument. Information uncertainty no longer impacts mispricing in linear regression when larger cap stocks are emphasized.

The key regression reported in column (6) simultaneously includes MISPRICE-score, ARBRISK, INFO-score, TCOST-score, and the interactions between MISPRICE-score and arbitrage costs group variables. The MISPRICE-score slope is 1.305 (*t*-statistic = 1.70). The interactive slopes for ARBRISK, INFO-score, and TCOST-score, respectively, are 0.216 (*t*-statistic = 2.15), -0.183 (*t*-statistic = -1.30), and 0.084 (*t*-statistic = 0.52). Only arbitrage risk impacts cross-sectional mispricing. These findings are consistent with larger cap stocks having better information quality (e.g., Ohlson, 1995; Hong, Lim, and Stein, 2000) and being less expensive to trade (e.g., Novy-Marx and Velikov, 2016).

5.2 Annually-varying principal components of arbitrage costs

Panel A of Table 6 presents year-by-year percentages of total variance explained by each of the nine principal components (PC1 to PC9). The percentages are generally stable over the years. On average, the first, second, and third principal components together capture 59% of the total variance. Panel B presents year-by-year eigenvalues of each principal components. The eigenvalues are stable throughout the years. The first, second, and third principal components carry eigenvalues greater than or equal to one every year. As a result, they also have eigenvalues greater than or equal to one on average.

[Table 6 here]

Panel C presents time-series average eigenvectors of the PC1, PC2, and PC3. PC1 positively loads on IVOL, CFVOL, BIDASK, and IO and negatively loads on COV, PRICE, and DVOL. A higher value of PC1 is associated with higher arbitrage risk, higher information uncertainty in terms of higher cash flow volatility and lower analyst coverage, and higher transaction costs via higher effective bid-ask spread and lower share price and dollar trading volume. However, it is associated with lower transaction costs in terms of higher passive institutional ownership. PC2 positively loads on IVOL, CFVOL, COV, DVOL, and BIDASK and negatively loads on IO. A higher value of PC2 is associated with higher arbitrage risk, higher information uncertainty in terms of higher cash flow volatility, and higher transaction costs via higher effective bid-ask spread and lower passive institutional ownership. However, it is associated with lower information uncertainty in terms of higher analyst coverage and lower transaction costs in terms of higher dollar trading volume. PC1 and PC2 are similar in how they relate to IVOL, CFVOL, and BIDASK but they are different in how they relate to COV, PRICE, DVOL, and IO. PC3 loads on DISP, on which PC1 and PC2 do not.

Table 7 reports alphas of value-weighted portfolios sorted by quintiles of MISPRICE-score and quintiles of PC1 (Panel A), PC2 (Panel B), or PC3 (Panel C). The long-short alphas are 2.22% (*t*-statistic = 5.26) and 1.05% (*t*-statistic = 5.03) in the high PC1 and PC2 quintiles, respectively. These long-short alphas decline to 0.52% (*t*-statistic = 3.56) and 0.20% (*t*-statistic = 1.08) as we move to the low quintiles. The corresponding differences in long-short alpha are 1.70% (*t*-statistic = 4.28) and 0.85% (*t*-statistic = 3.66). PC1 and PC2 significantly impact mispricing in the cross section. In the Overpriced quintile, the high-low alphas are -1.16% (*t*-statistic = -4.02) and -0.88% (*t*-statistic = -3.84) for PC1 and PC2, respectively. In the Underpriced quintile, the high-low alphas are 0.43% (*t*-statistic = 2.30) and -0.03% (*t*-statistic = -0.18) for PC1 and PC2, respectively. Resembling overall arbitrage costs and arbitrage costs groups, high arbitrage costs identified by the first and second principal components impede shorting overpriced stocks more significantly than longing underpriced stocks. However, PC3 does not impact mispricing according to the limits-to-arbitrage argument. The long-short alpha is merely 0.39% (*t*-statistic = 1.94) in the high PC3 quintile but it is 1.00% (*t*-statistic = 4.11) in the low PC3 quintile. The difference in long-short alpha is -0.62% (*t*-statistic = -2.46). Furthermore, portfolio and regression results show that the remaining principal components do not impact mispricing.

[Table 7 here]

Table 8 reports regression results for investigating the first, second, and third annually-varying principal components of arbitrage costs individually and jointly. Monthly regressions reported in columns (1) to (5) are estimated with OLS. Column (1) reports slope estimates of the baseline regression. The MISPRICE-score slope is 2.300 (*t*-statistic = 6.81). Columns (2/3/4) add PC1/PC2/PC3 decile (PC1_dec/PC2_dec/PC3_dec) sorted independently at the end of June each year and the interaction between MISPRICE-score and principal component decile. The interactive slopes for PC1_dec, PC2_dec, and PC3_dec, respectively, are 0.344 (*t*-statistic = 4.55), 0.282 (*t*-statistic = 4.59), and -0.002 (*t*-statistic = -0.04). Consistent with double portfolio sorts, PC1 and PC2 does but PC3 does not impact mispricing according to the limits-to-arbitrage argument.

[Table 8 here]

The key regression reported in column (5) simultaneously includes MISPRICE-score, PC1_dec, PC2_dec, and the interactions between MISPRICE-score and principal component deciles. The interactive slopes for PC1_dec and PC2_dec are 0.264 (*t*-statistic = 3.34) and 0.249 (*t*-statistic = 4.45), respectively. The first and second principal components jointly impact cross-sectional mispricing in a similar manner.

Monthly regressions reported in columns (6) to (10) are estimated with WLS, which emphasizes on larger cap stocks. Column (6) reports slope estimates of the baseline regression. The MISPRICE-score slope is 1.280 (*t*-statistic = 3.39). Columns (7/8/9) add PC1/PC2/PC3 decile and the interaction between MISPRICE-score and principal component decile. The interactive slopes for PC1_dec, PC2_dec, and PC3_dec, respectively, are 0.155 (*t*-statistic = 1.83), 0.109 (*t*-statistic = 1.17), and -0.152 (*t*-statistic = -1.82).

The key regression reported in column (6) simultaneously includes MISPRICE-score, PC1_dec, PC2_dec, and the interactions between MISPRICE-score and principal component deciles. The interactive slopes for PC1_dec and PC2_dec are 0.200 (*t*-statistic = 2.21) and 0.198 (*t*-statistic = 2.03), respectively. The first and second principal components jointly impact cross-sectional mispricing in a similar manner even when larger cap stocks are emphasized.[19]

5.3 Subperiod analysis of nine arbitrage costs measures

Table 9 reports regressions for three subperiods. The baseline monthly cross-sectional regressions in Panel A are estimated by OLS. The MISPRICE-score slope is 2.019 (*t*-statistic = 7.40) in the pre-discovery period (07/1963-06/1983). The slope rises to 2.951 (*t*-statistic = 7.91) during the golden age of mispricing (07/1983-12/2004) and it falls to 0.835 (*t*-statistic = 2.21) in the attenuation period (01/2005-12/2015). Mispricing becomes economically weaker but continues to exist in recent years.

[Table 9 here]

The baseline monthly cross-sectional regressions in Panel B are estimated by WLS, which emphasizes on larger cap stocks. The MISPRICE-score slope is 1.753 (*t*-statistic = 5.37) in the pre-discovery period. The slope drops to 1.496 (*t*-statistic = 1.21) during the golden age of mispricing and it further drops to 0.796 (*t*-statistic = 2.21) in the attenuation period. When larger cap stocks are emphasized, mispricing becomes economically weaker over time. In the recent years, the MISPRICE-score slope is volatile and mispricing is absent.

---

[19] The findings remain the same if we add PC3_dec and MISPRICE-score × PC3_dec to the key regressions.

Panel C reports the key OLS regressions that simultaneously include MISPRICE-score, deciles of the nine arbitrage costs measures independently sorted at the end of June each year, and the interactions between MISPRICE-score and arbitrage costs deciles. In the pre-discovery period, eight out of the nine interactive slopes are insignificant. The interactive slope for CFVOL_dec, which is 0.090 (*t*-statistic = 1.68), is the exception. Yet, it is not only marginally significant at the 10% level but also economically insignificant. Arbitrage costs have tiny impact on mispricing.

During the golden age of mispricing, the interactive slopes for ARBRISK, CFVOL_dec, PRICE_dec, and IO_dec, respectively, are 0.348 (*t*-statistic = 4.89), 0.097 (*t*-statistic = 1.82), 0.278 (*t*-statistic = 2.97), and 0.140 (*t*-statistic = 2.85). The other interactive slopes are insignificant. On the one hand, arbitrage risk, ambiguity of fundamental value, round-trip broker's commission plus bid-ask spreads, and stock loan supply jointly impact cross-sectional mispricing according to the limits-to-arbitrage argument. On the other hand, divergence in opinion on future earnings, information asymmetry, liquidity, price impact, and compensation for liquidity provision are not binding on arbitrageurs.

In the attenuation period, the interactive slope for PRICE_dec and IO_dec are 0.218 (*t*-statistic = 1.66) and 0.339 (*t*-statistic = 4.94), respectively. The interactive slope is -0.799 (*t*-statistic = -2.02) for DVOL_dec and 0.828 (*t*-statistic = 1.87) for ILLIQ_dec. It is counterintuitive that the two transaction costs measures have opposite effects on mispricing. In view of the negative correlation between these measures, we estimate the regression without DVOL_dec or ILLIQ_dec. The interactive slope for DVOL_dec is 0.097 (*t*-statistic = 1.47) without ILLIQ_dec and the interactive slope for ILLIQ_dec is 0.057 (*t*-statistic = 0.70) without DVOL_dec. In both cases, the interactive slopes for PRICE_dec and IO_dec are positive and significant. In all three specifications, other interactive slopes are insignificant. That is, only round-trip broker's commission plus bid-ask spreads and stock loan supply impact cross-sectional mispricing and are binding on arbitrageurs.

Panel D reports the key WLS regressions that simultaneously include MISPRICE-score, nine arbitrage costs deciles, and the interactions between MISPRICE-score and arbitrage costs deciles. These regressions emphasize on larger cap stocks. In the pre-discovery period, none of the interactive slopes is significant. Arbitrage costs do not impact cross-sectional mispricing. During the golden age of mispricing, the interactive slope for ARBRISK is 0.354 (*t*-statistic = 2.01). The other interactive slopes are insignificant. Only arbitrage risk impacts mispricing and is binding on arbitrageurs. In the attenuation period, all interactive slopes are insignificant. Given the absence of mispricing in this subperiod, it is not surprising that arbitrage costs are not binding.

5.4 *Empirical arbitrage costs model vs. annually-varying principal components*

Findings in Section 5.3 suggest an empirical model that uses arbitrage risk, ambiguity of fundamental value, round-trip broker's commission plus bid-ask spreads, and stock loan supply in the golden age of mispricing, and round-trip broker's commission plus bid-ask spreads and stock loan supply in the attenuation period as the binding arbitrage costs. When larger cap stocks are emphasized, the model uses arbitrage risk in the golden age of mispricing. Columns (1) and (2) in Panel A of Table 10 presents this empirical arbitrage costs model.[20] The OLS regression for the golden age of mispricing simultaneously includes MISRPICE-score, deciles of ARBRISK, CFVOL, PRICE, and IO, and the interactions between MISRPICE-score and arbitrage costs deciles. The interactive slopes for ARBRISK, CFVOL_dec, PRICE_dec, and IO_dec, respectively, are 0.324 (*t*-statistic = 4.36), 0.103 (*t*-statistic = 1.87), 0.194 (*t*-statistic = 2.70), and 0.174 (*t*-

[20] For fair comparison, we require firms to have valid observations for all nine arbitrage costs measures.

statistic = 3.64). The OLS regression for the attenuation period simultaneously includes MISRPICE-score, deciles of PRICE and IO, and the interactions between MISRPICE-score and the deciles. The interactive slopes for PRICE_dec and IO_dec are 0.277 (*t*-statistic = 2.52) and 0.336 (*t*-statistic = 5.82), respectively.

Columns (3) and (4) present OLS regressions that simultaneously include MISRPICE-score, deciles of PC1 and PC2, and the interactions between MISRPICE-score and the deciles. During the golden age of mispricing, the interactive slopes for PC1_dec and PC2_dec are 0.365 (*t*-statistic = 3.68) and 0.257 (*t*-statistic = 3.72), respectively. The principal components jointly impact mispricing. In the attenuation period, the interactive slopes for PC1_dec and PC2_dec are 0.036 (*t*-statistic = 0.42) and 0.233 (*t*-statistic = 2.44), respectively. Mirroring the empirical model, fewer of the arbitrage costs are binding on arbitrageurs.

Columns (5) and (6) pit the empirical arbitrage costs model against the annually-varying principal components of arbitrage costs. The OLS regression for the golden age of mispricing simultaneously includes MISRPICE-score, deciles of ARBRISK, CFVOL, PRICE, IO, PC1, and PC2, and the interactions between MISRPICE-score and the deciles. The interactive slopes for ARBRISK, CFVOL_dec, PRICE_dec, IO_dec, PC1_dec, and PC2_dec, respectively, are 0.283 (*t*-statistic = 3.50), 0.078 (*t*-statistic = 1.39), 0.191 (*t*-statistic = 2.77), 0.171 (*t*-statistic = 3.51), 0.057 (*t*-statistic = 0.72), and 0.063 (*t*-statistic = 0.99). The OLS regression for the attenuation period simultaneously includes MISRPICE-score, deciles of PRICE, IO, PC1, and PC2, and the interactions between MISRPICE-score and the deciles. The interactive slopes for PRICE_dec, IO_dec, PC1_dec, and PC2_dec, respectively, are 0.268 (*t*-statistic = 2.05), 0.278 (*t*-statistic = 4.61), -0.013 (*t*-statistic = -0.13), and 0.123 (*t*-statistic = 1.27). In the golden age of mispricing and the attenuation period, the empirical model subsumes the annually-varying principal components.

Column (1) in Panel B presents the empirical arbitrage costs model when larger cap stocks are emphasized. The WLS regression for the golden age of mispricing simultaneously includes MISRPICE-score, ARBRISK, and the interaction between MISRPICE-score and ARBRISK. The interactive slope for ARBRISK is 0.517 (*t*-statistic = 3.43). Column (2) presents the WLS regression that simultaneously includes MISRPICE-score, deciles of PC1 and PC2, and the interactions between MISRPICE-score and the deciles. The interactive slopes for PC1_dec and PC2_dec are 0.289 (*t*-statistic = 2.49) and 0.269 (*t*-statistic = 2.26), respectively. Both principal components impact mispricing.

Column (3) pits the empirical arbitrage costs model against the annually-varying principal components of arbitrage costs. The WLS regression simultaneously includes MISRPICE-score, ARBRISK, deciles of PC1 and PC2, and the interactions between MISRPICE-score and the deciles. The interactive slopes ARBRISK, PC1_dec, and PC2_dec, respectively, are 0.500 (*t*-statistic = 2.56), -0.073 (*t*-statistic = -0.48), and 0.128 (*t*-statistic = 1.01). When larger cap stocks are emphasized, the empirical model still subsumes the annually-varying principal components.

6. Conclusion

For any cross-sectional stock return anomaly, there are two economic explanations: one is rational pricing and the other one is behavioral mispricing. The rational explanation argues that stocks earn higher future returns because they are discounted more to compensate for higher systematic risks. The behavioral explanation suggests that stocks are mispriced due to investors' psychological biases. To test the mispricing explanation for an anomaly, a current norm is to apply the limits-to-arbitrage argument, which predicts the anomaly to be stronger when arbitrage costs are higher. The literature documents positive correlations between various anomalies and arbitrage costs. However, there is a tendency to use different arbitrage costs measures across different papers because the literature has not studied out which of the various arbitrage costs constitutes the most crucial arbitrage barrier.

In this paper, we take nine arbitrage costs measures that have strong theoretical motivations and large amounts of data, and are reliable and commonly used in the anomalies literature as given, and we systematically perform an empirical investigation on their impact on cross-sectional mispricing. We utilize an extensive list of well-known and robust anomalies to ensure our findings are general. Results based on monthly U.S. stock returns from 07/1963 to 12/2015 indicate that overall arbitrage costs impact, mainly through the short leg, anomalous returns in the cross section. Three typical arbitrage costs groups, namely arbitrage risk, information uncertainty, and transaction costs, are jointly important. Only arbitrage risk matters when larger cap stocks are emphasized.

The main finding we uncover is that binding arbitrage barriers slowly change over time. In the early years when there are few publications documenting return anomalies, arbitrage costs have tiny impact even though mispricing is present. As anomalies become more extensively known, arbitrage costs impact mispricing substantially. Arbitrage risk (idiosyncratic stock return volatility), ambiguity of fundamental value (cash flow volatility), round-trip broker's commission plus bid-ask spreads (share price), and stock loan supply (passive institutional ownership) are binding on arbitrageurs. Only arbitrage risk is binding if larger cap stocks are emphasized. In recent years, as market quality improves and some arbitrageurs become more creative, only round-trip broker's commission plus bid-ask spreads and stock loan supply remain binding on arbitrageurs. When larger cap stocks are emphasized, arbitrage costs do not matter at all because there is no longer mispricing. This indicates that the significance of some arbitrage costs groups in the entire sample is periodic. Besides, the significance of information uncertainty stems from ambiguity of fundamental value while the significance of transaction costs stems from round-trip broker's commission plus bid-ask spreads and stock loan supply. An empirical arbitrage costs model based on these simple dynamics subsumes annually-varying principal components of arbitrage costs in affecting mispricing. Incorporating our findings into future capital market efficiency research would mitigate type I and II errors in empirical tests using the limits-to-arbitrage argument.

References

Abreu, Dilip, and Markus K. Brunnermeier, 2003. Bubbles and crashes, *Econometrica* 71, 173–204.

Almazan, Andres, Keith C. Brown, Murray Carlson, and David A. Chapman, 2004. Why constrain your mutual fund manager? *Journal of Financial Economics* 73, 289–321.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006. The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2009. High idiosyncratic volatility and low returns: international and further U.S. evidence, *Journal of Financial Economics* 91, 1–23.

Ali, Ashiq, Lee-Seok Hwang, and Mark A. Trombley, 2003. Arbitrage risk and the book-to-market anomaly, *Journal of Financial Economics* 69, 355–373.

Amihud, Yakov, and Haim Mendelson, 1986. Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

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

Ball, Ray, Joseph Gerakos, Juhani T. Linnainmaa, and Valeri V. Nikolaev, 2015. Deflating profitability, *Journal of Financial Economics* 117, 225–248.

Banz, Rolf W., 1981. The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3–18.

Barry, Christopher B., and Stephen J. Brown, 1984. Differential information and the small firm effect, *Journal of Financial Economics* 13, 283–294.

Bhandari, Laxmi Chand, 1988. Debt/Equity ratio and expected common stock returns: empirical evidence, *Journal of Finance* 43, 507–528.

Bhardwaj, Ravinder K., and Leroy D. Brooks, 1992. The January anomaly: effects of low share price, transaction costs, and bid-ask bias, *Journal of Finance* 47, 553–575.

Bhushan, Ravi, 1994. An informational efficiency perspective on the post-earnings announcement drift, *Journal of Accounting and Economics* 18, 45–66.

Blume, Marshall E., and Frank Husic, 1973. Price, beta, and exchange listing, *Journal of Finance* 28, 283–299.

Brav, Alon, J.B. Heaton, and Si Li, 2010. The limits of the limits of arbitrage, *Review of Finance* 14, 157–187.

Campbell, John, Y., Jens Hilscher, and Jan Szilsagyi, 2008. In search of distress risk, *Journal of Finance* 63, 2899–2939.

Carhart, Mark M., 1997. On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Cao, Jie and Bing Han, 2016. Idiosyncratic risk, costly arbitrage and the cross-section of stock returns, *Journal of Banking and Finance* 73, 1–15.

Cen, Ling, K.C. John Wei, and Liyan Yang, 2016. Disagreement, underreaction, and stock returns, *Management Science*, forthcoming

Chen, Te-Feng, Sun Lei, K.C. John Wei, and Feixue Xie, 2016. On the explanations for the gross profitability effect: insights from international equity markets, working paper presented at the 2014 European Finance Association Annual Meeting.

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, 1986. Economic forces and the stock market, *Journal of Business* 59, 383–403.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008. Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1651.

Daniel, Kent, and Sheridan Titman, 2006. Market reaction to tangible and intangible information, *Journal of Finance* 52, 1–33.

D'Avolio, Gene, 2002. The market for borrowing stock, *Journal of Financial Economics* 66, 271–306.

De Bondt, Werner F.M., and Richard Thaler, 1985. Does the stock market overreact? *Journal of Finance* 40, 793–805.

Dechow, Patricia M., Amy P. Hutton, Lisa K. Meulbroek, and Richard G. Sloan, 2001. Short-sellers, fundamental analysis, and stock returns, *Journal of Financial Economics* 61, 77–106.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, Robert J. Waldmann, 1990. Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Dichev, Ilia D., 1998. Is the risk of bankruptcy a systematic risk? *Journal of finance* 53, 1131–1147.

Diether, Karl, Christopher Malloy, and Anna Scherbina, 2002. Differences of opinion and the cross-section of stock returns, *Journal of Finance* 57, 2113–2141.

Duan, Ying, Gang Hu, and R. David McLean, 2010. Costly arbitrage and idiosyncratic risk: evidence from short sellers, *Journal of Financial Intermediation* 19, 564–579.

Fama, Eugene F., 1965. The behavior of stock market prices, Journal of business 38, 34–105.

Fama, Eugene F., and Kenneth R. French, 1992. The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993. Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2006. Profitability, investment and average returns, *Journal of Financial Economics* 82, 491–518.

Fama, Eugene F., and Kenneth R. French, 2015. A five-factor asset pricing model, *Journal of Financial Economics* 116, –22.

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

Fama, Eugene F. and James D. MacBeth, 1973. Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Foster, George, Chris Olsen, and Terry Shevlin, 1984. Earnings releases, anomalies, and the behavior of security returns, *The Accounting Review* 59, 574–603.

Gleason, Cristi A., and Charles M.C. Lee, 2003. Analyst forecast revisions and market price discovery, *The Accounting Review* 78, 193–225.

Griffin, John M., and Michael L. Lemmon, 2002. Book-to-market equity, distress risk, and stock returns, *Journal of Finance* 57, 2317–2336.

Gromb, Denis, and Dimitri Vayanos, 2010. Limits of arbitrage, *Annual Review of Financial Economics* 2, 251–275.

Hirshleifer, David, Kewei Hou, Siew H. Teoh, and Yinglei Zhang, 2004. Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics* 38, 297–331.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265–295.

Hong, Harrison, and David Sraer, 2016. Speculative betas, *Journal of Finance* 71, 2095–2144.

Jegadeesh, Narasimhan, 1990. Evidence of predictable behavioral of security returns, *Journal of Finance* 45, 881–898.

Lam, F.Y. Eric C., and K.C. John Wei, 2011. Limits-to-arbitrage, investment frictions, and the asset growth anomaly, *Journal of Financial Economics* 102, 127–149.

Langetieg, Terence C., 1978. An application of a three-factor performance index to measure stockholder gains from merger, *Journal of Financial Economics* 6, 365–383.

Li, Xiafei, and Di Luo, 2016. Investor sentiment, limited arbitrage, and the cash holding effect, *Review of Finance*, forthcoming.

Li, Dongmei, and Lu Zhang, 2010. Does *q*-theory with investment frictions explain anomalies in the cross-section of returns? *Journal of Financial Economics* 98, 297–314.

Lintner, John, 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statist*ics 47, 13–37.

Lipson, Marc L., Sandra Mortal, and Michael J. Schill, 2011. On the scope and drivers of the asset growth effect, *Journal of Financial and Quantitative Analysis* 46, 1651–1682.

Ljungqvist, Alexander, and Wenlan Qian, 2016. How constraining are limits to arbitrage? *Review of Financial Studies* 29, 1975–2028.

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2008. The new issues puzzle: testing the investment-based explanation, *Review of Financial Studies* 21, 2825–2855.

Mashruwala, Christina A., Shivaram Rajgopal, and Terry J. Shevlin, 2006. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs, *Journal of Accounting and Economics* 42, 3–33.

McLean, R. David, 2010. Idiosyncratic risk, long-term reversal, and momentum, *Journal of Financial and Quantitative Analysis* 45, 883–906.

McLean, R. David, and Jeffrey Pontiff, 2016. Does academic research destroy stock return predictability? *Journal of Finance* 71, 5–31.

Nagel, Stefan, 2005. Short sales, institutional investors, and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.

Newey, Whitney K., and Kenneth D. West, 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Novy-Marx, Robert, 2013. The other side of value: the gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Novy-Marx, Robert, and Mihail Velikov, 2016. A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Ohlson, James A., 1980. Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109–13.

Ohlson, James A., 1995. Earnings, book values, and dividends in equity valuation, *Contemporary Accounting Research* 11, 661–687.

Polk, Christopher, and Paolo Sapienza, 2009. The stock market and corporate investment: a test of catering theory, *Review of Financial Studies* 22, 187–217.

Pontiff, Jeffrey, 1996. Costly arbitrage: evidence from closed-end funds, *Quarterly Journal of Economics* 111, 1135–1151.

Pontiff, Jeffrey, 2006. Costly arbitrage and the myth of idiosyncratic risk, *Journal of Accounting and Economics* 42, 35–52.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008. Share issuance and cross-sectional returns, *Journal of Finance* 63, 921–945.

Scott, A. J., and M., Knott, 1974. A cluster analysis method for grouping means in the analysis of variance, *Biometrics* 30, 507–512.

Sharpe, William F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Shleifer, Andrei, and Robert W. Vishny, 1997. The limits of arbitrage, *Journal of Finance* 52, 35–55.

Shumway, Tyler, 1997. The delisting bias in CRSP data, *Journal of Finance* 52, 327–340.

Shumway, Tyler, and Vincent A. Warther, 1999. The delisting bias in CRSP's Nasdaq data and its implications for the size effect, *Journal of Finance* 54, 2361–2379.

Sloan, Richard G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings, *The Accounting Review* 3, 289–315.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2012. The short of it: investor sentiment and anomalies, *Journal of Financial Economics* 104, 288–302.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle, *Journal of Finance* 70, 1903–1948.

Titman, Sheridan, K.C. John Wei, and Feixue Xie, 2004. Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39, 677–700.

Titman, Sheridan, K.C. John Wei, and Feixue Xie, 2013. Market development and the asset growth effect: international evidence, *Journal of Financial and Quantitative Analysis* 48, 1405–1432.

Watanabe, Akiko, Yan Xu, Tong Yao, and Tong Yu, 2013. The asset growth effect: insights from international equity markets, *Journal of Financial Economics* 108, 529–563.

Wurgler, Jeffrey A., and Ekaterina Zhuravskaya, 2002. Does arbitrage flatten demand curves for stocks? *Journal of Business* 75, 583–607.

Xing, Yuhang, 2008. Interpreting the value effect through the *q*-theory: an empirical investigation, *Review of Financial Studies* 21, 1767–1795.

Yan, Xuemin (Sterling), and Lingling Zheng, 2017, Fundamental analysis and the cross-section of stock returns: A data-mining approach, *Review of Financial Studies,* forthcoming.

Zhang, X. Frank, 2006. Information uncertainty and stock returns, *Journal of Finance* 61, 105–136.

Appendix. Variable definitions

The individual arbitrage costs measures

IVOL: Idiosyncratic stock return volatility, which is the standard deviation of residuals from the market model of monthly stock returns as dependent variable and S&P 500 return as independent variable, estimated at the end of June of calendar year *t* with a full history of 36 months of observations. Data source: CRSP.

CFVOL: Cash flow volatility, measured as the standard deviation of cash flow from operations for fiscal years *t*–5, *t*–4, *t*–3, *t*–2, and *t*–1. At least three years of observations are required. Cash flow from operations is earnings before extraordinary items (item IB) minus accruals, scaled by the average of total assets (item AT) over fiscal year *t*. Accruals is change in current assets (item ACT) less change in cash and short-term investments (item CHE) less change in current liabilities (item LCT) less depreciations (item DP) plus change in debt included in current liabilities (item DLC) plus change in income taxes payable (item TXP) from fiscal year *t*–1 to fiscal year *t*. Data source: Compustat.

DISP: Analyst earnings forecast dispersion, calculated as the latest available standard deviation of one-year ahead earnings forecasts on the stock from the beginning of January of calendar year *t* to the end of June of calendar year *t*, scaled by closing share price at the end of June of calendar year *t*. Data source: I/B/E/S and CRSP.

COV: Analyst coverage, measured as the latest available number of analysts following the stock from the beginning of January of calendar year *t* to the end of June of calendar year *t*. Data source: I/B/E/S.

PRICE: Closing share price, or the average of bid and ask quotes if share price is unavailable, at the end of June of calendar year *t*. Data source: CRSP.

DVOL: Daily dollar trading volume, which is closing share price times the trading day's share trading volume, averaged over the year at the end of June of calendar year *t*. Data source: CRSP.

ILLIQ: Amihud (2002) illiquidity, measured as the time-series average of absolute value of daily returns scaled by the trading day's dollar trading volume over the year at the end of June of calendar year *t*. Data source: CRSP.

BIDASK: Bid-ask spread, which is calculated as 2×|(Price–(Ask+Bid)/2)|/Price at the end of a trading day, averaged over the year at the end of June of calendar year *t*. Price is closing share price and Ask (Bid) is the ask (bid) quote. Data source: CRSP.

IO: Institutional ownership, measured as the latest available percentage of outstanding shares held by Vanguard 500 index fund or Dimensional Fund Advisors from the beginning of January of calendar year *t* to the end of June of calendar year *t*. Data source: Thompson Reuters (13F) Institutional Holdings and CRSP.

The individual anomaly variables

TAG: Total asset growth, calculated as total assets (item AT) at the end of fiscal year *t*–1 minus total assets at the end of fiscal year *t*–2, scaled by total assets at the end of fiscal year *t*–2. Data source: Compustat.

I/A: Investment-to-asset ratio, calculated as change in the sum of inventories (item INVT) and gross property, plant, and equipment (item PPEGT) from fiscal year *t*–2 to fiscal year *t*–1, scaled by total assets (item AT) at the end of fiscal year *t*–2. Data source: Compustat.

NOA: Net operating assets, calculated as the difference between operating assets and operating liabilities at the end of fiscal year *t*–1 scaled by total assets (item AT) at the end of fiscal year *t*–2. Operating assets is total assets minus cash and short-term investments (item CHE). Operating liabilities is total assets less current liabilities (item DLC), long-term debt (item DLTT), minority interests (item MIB), preferred stocks (item PSTK), and common equity (item CEQ). Data source: Compustat.

OA: Operating accruals, calculated as change in current assets (Compustat item ACT) less change in cash and short-term investments (item CHE) less change in current liabilities (item LCT) less depreciations (item DP) plus change in debt included in current liabilities (item DLC) plus change in income taxes payable (item TXP) from fiscal year *t*–2 to fiscal year *t*–1, scaled by the average of total assets (item AT) over the year. Data source: Compustat.

NSI: Net share issuance, calculated as the natural logarithm of the ratio of split-adjusted shares outstanding (item CSHO multiplied by item ADJEX_C) at the end of fiscal year *t*–1 to that at end of fiscal year *t*–2. Data source: Compustat.

CEI: Composite equity issuance, calculated as the difference between continuously compounded growth in market equity from the end of June of calendar year *t*–5 to the end of June of calendar year *t* and continuously compounded growth in share price from the end of June of calendar year *t*–5 to the end of June of calendar year *t*. Data source: CRSP.

O-score: Ohlson (1980) bankruptcy score, calculated as

$$O\text{-}score = -4.07\times Ln(A) + 6.03\times(L/A) - 1.43\times(CA\text{-}CL)/A + 0.0757\times CL/CA$$
$$-2.37\times NI/A + 0.285\times Loss - 1.72\times NegBook - 0.521\times\Delta NI - 1.83\times Op/L,$$

where *Ln*(*A*) is natural logarithm of total assets (item AT), *L* is total liabilities (item LT), *A* is total assets (item AT), *CA* is current assets (item ACT), and *CL* is current liabilities (item LCT) at the end of fiscal year *t*–1. *NI* is net income (item NI) for fiscal year *t*–1. *Loss* is equal to one if net income (item NI) for fiscal year *t*–1 and net income (item NI) for fiscal year *t*–2 are negative and zero otherwise. *NegBook* is equal to one if *L* is greater than *A* and zero otherwise. Δ*NI* is change in net income (item NI) from fiscal year *t*–2 to fiscal year *t*–1, scaled by the sum of the absolute values of the net income (item NI) over the two years. *Op*, funds from operations, is income before extraordinary items (item IB) plus income statement deferred tax (item TXDI), if available, plus equity's share of depreciation expenses for fiscal year *t*–1, which is depreciation expenses (item DP) multiplied by market equity and divided by total assets (item AT) minus book value of equity plus market equity at the end of fiscal year *t*–1. Book equity is total assets (item AT) minus liabilities (item LT), plus balance sheet deferred taxes (item TXDB) and investment tax credits (item ITCI), minus preferred stock liquidation value (item PSTKL) if available, or redemption value (item PSTKRV) if available, or carrying value (item PSTK) if available. Market equity is closing share price times number of shares outstanding. Data source: Compustat and CRSP.

GP: Gross profitability, measured as gross profit (item GP) for fiscal year *t*–1 scaled by total assets (item AT) at the end of fiscal year *t*–1. Data source: Compustat.

OP: Operating profitability, measured as gross profit (item GP) less selling and general administrative expenditures (item XSGA) plus R&D expenditures (item XRD) for fiscal year *t*–1, scaled by total assets (item AT) at end of fiscal year *t*–1. Data source: Compustat.

PRET: Prior-year stock return, calculated as monthly stock return continuously compounded from the end of June of calendar year *t*–1 to the end of May of calendar year *t*. Data source: CRSP.

<u>Control variables in Fama-MacBeth regressions</u>

BETA: Capital asset pricing model (CAPM) beta, which is the slope of the time-series regression of monthly stock return in excess of the risk-free rate as dependent variable and the monthly market premium as independent variable, estimated at the end of June of calendar year *t* with a full history of 36 months of observations. Data source: CRSP and Kenneth French Data Library.

ME: Market equity, calculated as closing share price times number of shares outstanding at the end of June of calendar year *t*. Data source: CRSP.

B/M: Book-to-market equity ratio, calculated as book value of equity at the end of fiscal year *t*–1 divided by market equity at the end of calendar year *t*–1. Book equity is total assets (item AT) minus liabilities (item LT), plus balance sheet deferred taxes (item TXDB) and investment tax credits (item ITCI), minus preferred stock liquidation value (item PSTKL) if available, or redemption value (item PSTKRV) if available, or carrying value (item PSTK) if available. Market equity is closing share price times number of shares outstanding. Data source: Compustat and CRSP.

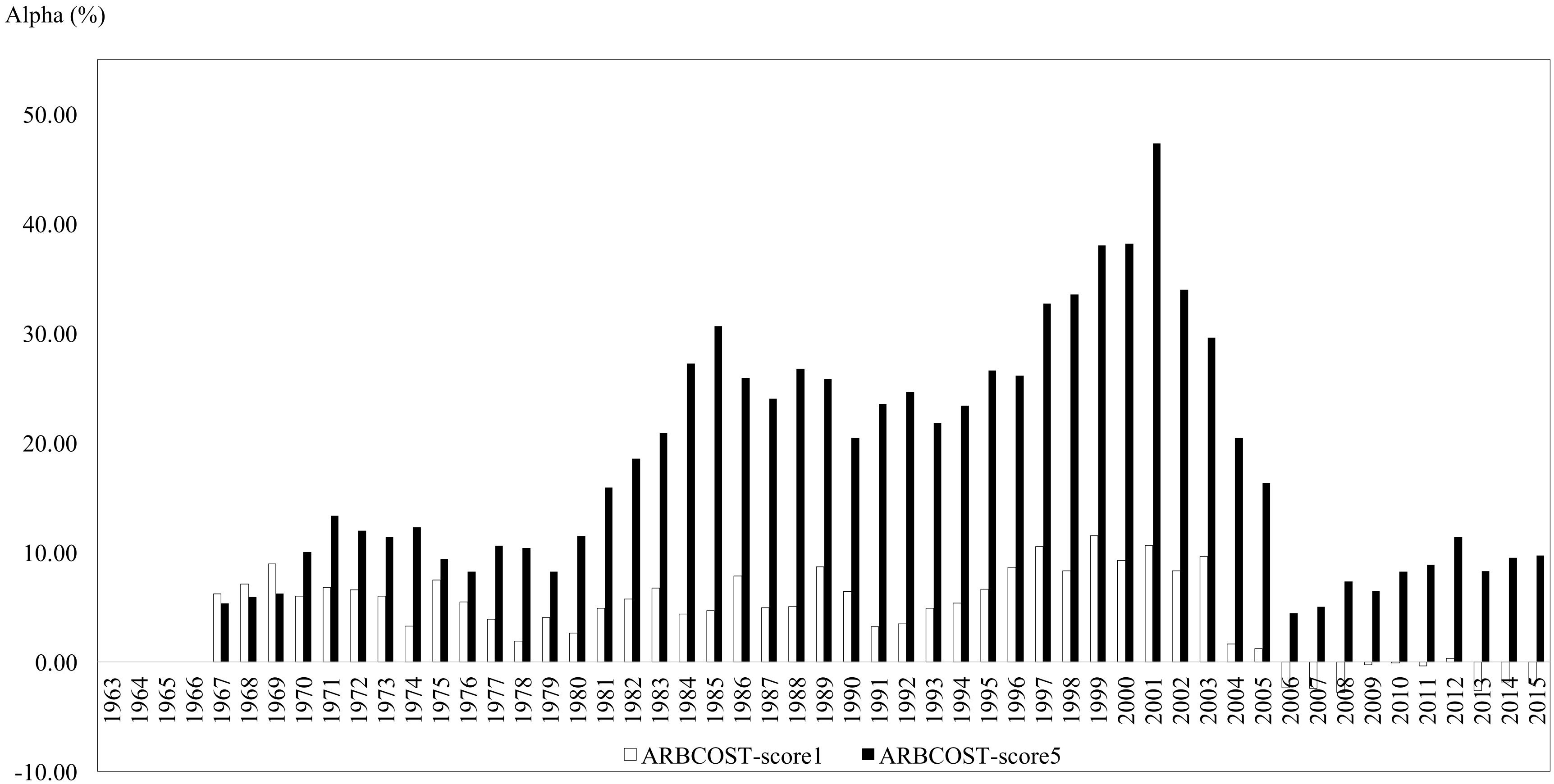

Figure 1. 5-year moving average of annualized alphas of value-weighted long-short MISPRICE-score portfolios in extreme ARBCOST-score quintiles

This figure presents 5-year moving average of annualized Fama-French three-factor alphas of long-short MISPRICE-score portfolios in the low (ARBCOST-score1) and high (ARBCOST-score5) ARBCOST-score quintiles. The sample period is from the end of June of 1963 to the end of December of 2015.

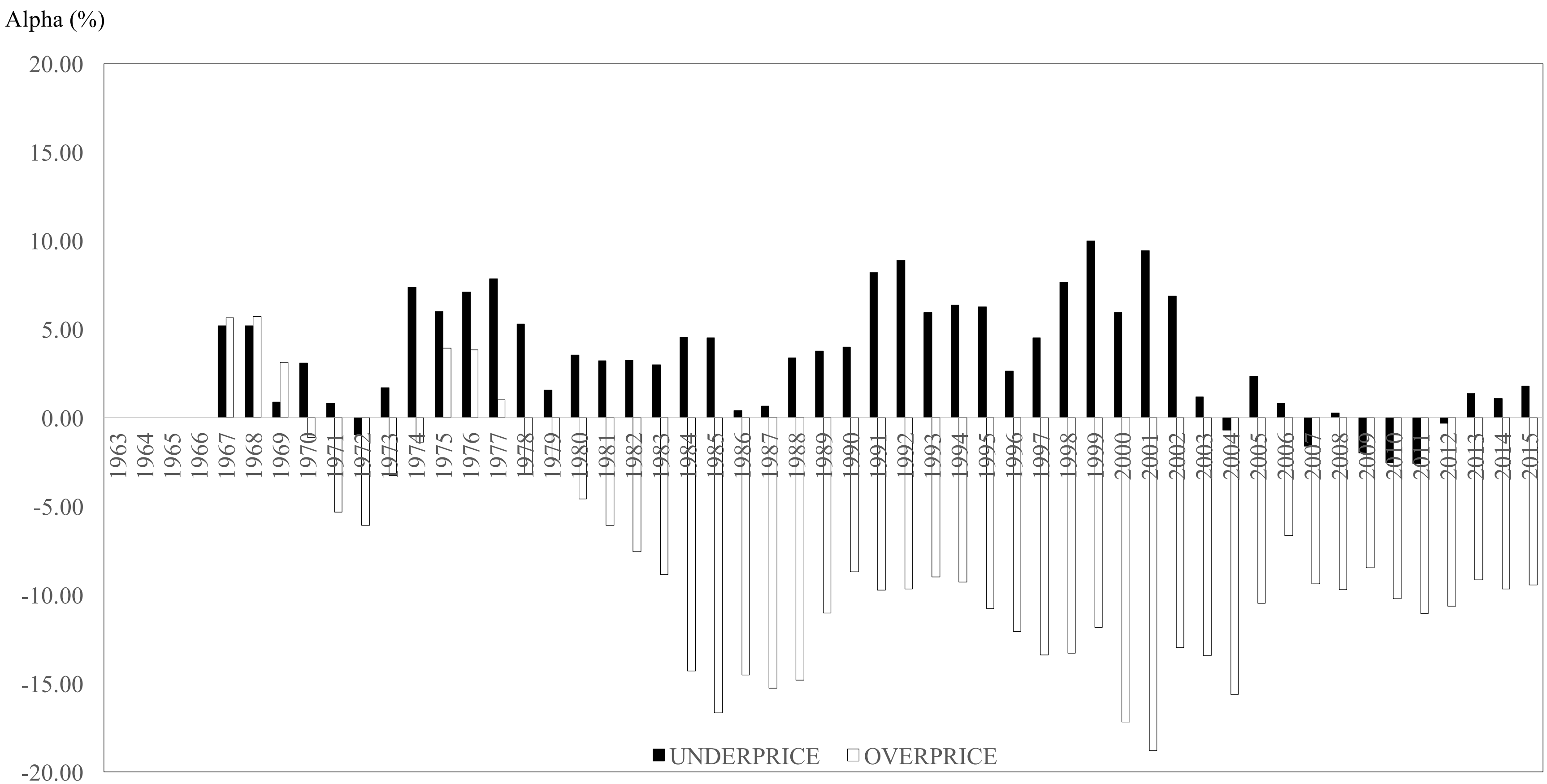

Figure 2. 5-year moving average of annualized alphas of long-short value-weighted ARBCOST-score portfolios in extreme MISPRICE-score quintiles

This figure presents 5-year moving average of annualized Fama-French three-factor alphas of high-minus-low ARBCOST-score portfolios in the first (OVERPRICE) and fifth (UNDERPRICE) MISPRICE-score quintiles. The sample period is from the end of June of 1963 to the end of December of 2015.

Table 1. Summary statistics and sample correlations

Panel A reports average cross-sectional mean (MEAN), standard deviation (SD), minimum (MIN), 25%/50%/75% percentiles (P25/P50/P75), and maximum (MAX) of the variables. N is the total number of firms. Panel B reports average cross-sectional correlations between arbitrage costs variables and the composite mispricing score. ARBCOST-score is the overall arbitrage costs score, which is the average of decile rankings of nine arbitrage costs measures, whichever is available. ARBRISK is arbitrage risk, which is the decile ranking of idiosyncratic stock return volatility (IVOL). INFO-score is the information uncertainty score, which is the average of decile rankings of cash flow volatility (CFVOL), analyst earnings forecast dispersion (DISP), and analyst coverage (COV), whichever is available. TCOST-score is the transaction costs score, which is the average of decile rankings of stock price (PRICE), dollar trading volume (DVOL), the Amihud (2002) illiquidity (ILLIQ), bid-ask spread (BIDASK), and institutional ownership (IO), whichever is available. PC1, PC2, and PC3 are the first, second, and third time-varying principal components of arbitrage costs. MISPRICE-score is the composite mispricing score, which is the average of decile rankings of total asset growth (TAG), investment-to-assets ratio (I/A), net operating assets (NOA), operating accruals (OA), net share issuance (NSI), composite equity issuance (CEI), Ohlson (1980) bankruptcy score (O-score), gross profitability (GP), operating profitability (OP), and prior-year stock return (PRET), whichever is available. CAPM beta (BETA), market equity (ME), and book-to-market equity ratio (B/M) are controls in Fama-MacBeth regressions. The Appendix defines the variables in detail. The sample period is from fiscal year 1962 to fiscal year 2014. The principal components are extracted from nine arbitrage costs measures at the end of June between 1983 and 2015.

Panel A. Summary statistics

	MEAN	SD	MIN	P25	P50	P75	MAX	N
ARBCOST-score	5.498	1.420	2.040	4.454	5.530	6.566	8.878	107,926
ARBRISK	5.500	2.872	1.000	3.000	5.708	8.000	10.000	107,926
INFO-score	5.491	1.331	2.208	4.544	5.588	6.487	7.981	107,926
TCOST-score	5.503	1.689	1.906	4.228	5.434	6.762	9.536	107,926
PC1	0.000	1.680	-11.141	-1.110	-0.034	1.082	7.851	44,013
PC2	0.000	1.214	-3.971	-0.724	-0.137	0.546	11.604	44,013
PC3	0.000	1.014	-8.193	-0.303	-0.975	0.238	13.765	44,013
IVOL	0.104	0.052	0.024	0.069	0.094	0.127	0.734	107,926
CFVOL	0.072	0.161	0.002	0.031	0.052	0.086	5.150	97,921
DISP (%)	0.796	3.382	0.000	0.051	0.180	0.522	45.564	47,331
COV	5.572	7.388	0.000	0.000	2.530	8.500	43.121	76,417
PRICE	26.944	30.332	5.095	12.342	21.096	33.981	781.980	107,926
DVOL (10^8)	14.460	52.633	0.001	0.471	2.388	10.155	1464.753	102,405
ILLIQ (10^{-9})	10.092	37.645	0.004	0.332	1.573	6.947	938.451	102,405
BIDASK (%)	1.724	0.850	0.083	1.055	1.615	2.230	6.641	107,926
IO (%)	1.593	1.975	0.000	0.202	0.686	2.144	10.093	81,134
MISPRICE-score	5.501	1.273	1.892	4.583	5.442	6.365	9.408	107,926
TAG	0.166	0.472	-0.644	0.015	0.085	0.192	12.083	107,505
I/A	0.099	0.230	-0.913	0.022	0.067	0.132	5.574	104,771
NOA	0.720	0.362	-0.922	0.579	0.715	0.832	8.044	95,845
OA	-0.025	0.082	-0.638	-0.065	-0.029	0.011	0.642	102,843
NSI	0.027	0.147	-1.465	-0.003	0.005	0.024	2.085	107,348
CEI	0.376	0.539	-2.436	0.009	0.235	0.664	3.241	94,344
O-score	0.801	5.259	0.000	0.026	0.084	0.243	78.170	104,539
GP	0.381	0.263	-0.859	0.202	0.343	0.507	2.277	107,837
OP	0.162	0.105	-0.666	0.106	0.153	0.213	0.908	107,822
PRET	0.166	0.811	-0.778	-0.124	0.061	0.293	20.520	107,926
BETA	1.160	0.682	-1.498	0.694	1.084	1.539	5.089	107,926
ME (10^9)	2.218	8.757	0.004	0.112	0.347	1.218	183.836	107,926
B/M	0.855	0.848	0.019	0.437	0.709	1.063	17.894	107,926

Table 1 (Continued)

Panel B. Sample correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) MISPRICE-score																
(2) ARBCOST-score	-0.06															
(3) ARBRISK	-0.08	0.73														
(4) INFO-score	-0.04	0.67	0.42													
(5) TCOST-score	-0.06	0.93	0.56	0.42												
(6) PC1	-0.15	0.88	0.76	0.66	0.81											
(7) PC2	-0.05	0.11	0.35	0.18	-0.05	0.00										
(8) PC3	0.07	0.00	0.05	-0.07	0.03	0.00	0.00									
(9) IVOL	-0.08	0.65	0.87	0.37	0.50	0.75	0.43	0.05								
(10) CFVOL	-0.06	0.44	0.36	0.61	0.27	0.50	0.41	0.11	0.38							
(11) DISP	-0.09	0.21	0.15	0.26	0.15	0.21	0.30	-0.51	0.15	0.09						
(12) COV	-0.02	-0.60	-0.27	-0.58	-0.54	-0.67	0.48	0.06	-0.23	-0.14	-0.03					
(13) PRICE	0.11	-0.54	-0.33	-0.25	-0.57	-0.64	0.18	0.07	-0.29	-0.15	-0.11	0.29				
(14) DVOL	0.04	-0.35	-0.16	-0.16	-0.40	-0.48	0.55	0.28	-0.13	-0.07	-0.03	0.44	0.38			
(15) ILLIQ	0.07	0.35	0.17	0.14	0.39	0.25	-0.18	0.47	0.19	0.10	-0.01	-0.14	-0.20	-0.11		
(16) BIDASK	-0.11	0.54	0.59	0.26	0.50	0.74	0.40	0.06	0.56	0.24	0.17	-0.16	-0.24	-0.06	0.01	
(17) IO	0.07	0.14	0.10	0.17	0.10	0.40	-0.40	0.20	0.03	0.00	-0.03	-0.28	-0.16	-0.12	0.00	0.12

Table 2: Monthly Fama-French three-factor alphas (%) of value-weighted portfolios sorted by MISPRICE-score or anomaly variable

Stocks are sorted into quintiles at the end of June each year. Portfolios are held until the end of June next year. Firms that are overpriced/underpriced are those in the first/fifth quintile. Long-short is the trading strategy which longs the underpriced portfolio and shorts the overpriced portfolio. The sample is from 07/1963 to 12/2015. Newey-West *t*-statistics are shown in parenthesis.

	MISPRICE-score	TAG	I/A	NOA	OA	NSI	CEI	O-score	GP	OP	PRET
Overpriced	-0.41	-0.03	-0.15	-0.24	-0.21	-0.21	0.01	-0.16	-0.17	-0.33	-0.11
	(-4.73)	(-0.45)	(-1.82)	(-3.40)	(-2.69)	(-2.62)	(0.24)	(-1.75)	(-2.53)	(-3.95)	(-1.22)
2	-0.15	0.14	0.06	-0.14	-0.01	0.11	0.08	-0.01	-0.11	-0.14	-0.05
	(-2.39)	(2.62)	(0.82)	(-2.18)	(-0.17)	(1.79)	(1.44)	(-0.17)	(-1.31)	(-2.18)	(-0.83)
3	0.09	0.05	0.08	0.16	0.14	0.09	-0.04	0.02	0.05	0.01	-0.01
	(1.97)	(1.01)	(1.92)	(3.12)	(2.87)	(1.41)	(-0.57)	(0.37)	(0.73)	(0.25)	(-0.23)
4	0.17	0.08	0.15	0.17	0.14	0.09	0.08	0.06	0.08	0.08	0.14
	(3.35)	(1.54)	(2.68)	(3.19)	(2.28)	(1.63)	(1.35)	(1.05)	(1.25)	(1.40)	(2.35)
Underpriced	0.25	0.10	0.11	0.24	0.10	0.17	0.14	0.14	0.35	0.23	0.27
	(4.39)	(1.21)	(1.68)	(3.75)	(1.32)	(2.76)	(1.74)	(2.45)	(4.44)	(4.40)	(2.94)
Long-short	0.66	0.13	0.26	0.49	0.31	0.37	0.12	0.30	0.51	0.56	0.38
	(5.58)	(1.21)	(2.59)	(4.56)	(2.71)	(3.32)	(1.25)	(2.54)	(4.19)	(4.86)	(2.81)

Table 3. The impact of overall arbitrage costs on mispricing

This table reports monthly Fama-French three-factor alphas (%) for 25 value-weighted portfolios independently sorted by ARBCOST-score and MISPRICE-score. Firms subject to low/high arbitrage costs are those in the first/fifth quintile. Long-short is the trading strategy which longs the underpriced portfolio and shorts the overpriced portfolio. High-low is the trading strategy which longs the high arbitrage costs portfolio and shorts the low arbitrage costs portfolio. The sample period is from 07/1963 to 12/2015. Newey-West *t*-statistics are shown in parenthesis.

MISPRICE-score	High costs	4	3	2	Low costs	High-low
Overpriced	-0.92	-0.97	-0.65	-0.45	-0.19	-0.73
	(-7.14)	(-5.79)	(-4.29)	(-3.41)	(-2.28)	(-4.74)
2	-0.47	-0.16	0.15	-0.23	-0.14	-0.33
	(-2.98)	(-1.44)	(1.37)	(-2.46)	(-1.76)	(-2.03)
3	-0.21	0.00	0.19	0.07	0.10	-0.31
	(-1.52)	(0.02)	(1.19)	(0.62)	(1.97)	(-2.08)
4	-0.04	-0.01	0.15	0.14	0.19	-0.23
	(-0.31)	(-0.08)	(1.61)	(1.12)	(3.09)	(-1.53)
Underpriced	0.35	0.18	0.47	0.17	0.25	0.10
	(2.47)	(1.61)	(3.06)	(1.60)	(3.73)	(0.73)
Long-short	1.27	1.15	1.12	0.62	0.44	0.83
	(6.52)	(5.68)	(4.58)	(3.32)	(4.38)	(4.50)

Table 4. The impact of arbitrage risk, information uncertainty, and transaction costs on mispricing

This table reports monthly Fama-French three-factor alphas (%) for 25 value-weighted portfolios independently sorted by ARBRISK, INFO-score, or TCOST-score and MISPRICE-score. Long-short is the trading strategy which longs the underpriced portfolio and shorts the overpriced portfolio. High-low is the trading strategy which longs the high arbitrage costs portfolio and shorts the low arbitrage costs portfolio. The sample period is from 07/1963 to 12/2015. Newey-West *t*-statistics are shown in parenthesis.

Panel A: Arbitrage risk (ARBRISK)

MISPRICE-score	High costs	4	3	2	Low costs	High-low
Overpriced	-0.95	-0.76	-0.51	-0.22	-0.11	-0.84
	(-6.69)	(-3.84)	(-3.09)	(-1.91)	(-1.21)	(-4.95)
2	0.04	0.01	-0.06	-0.15	-0.21	0.25
	(0.21)	(0.06)	(-0.52)	(-1.28)	(-2.26)	(1.15)
3	0.06	0.15	0.12	0.09	0.08	-0.02
	(0.30)	(0.84)	(0.94)	(1.05)	(1.13)	(-0.11)
4	-0.02	0.48	0.27	0.21	0.09	-0.10
	(-0.08)	(2.56)	(2.08)	(2.35)	(1.14)	(-0.44)
Underpriced	0.29	0.35	0.21	0.34	0.21	0.08
	(1.41)	(2.24)	(1.56)	(3.61)	(2.51)	(0.34)
Long-short	1.24	1.10	0.72	0.56	0.32	0.92
	(5.36)	(4.26)	(3.03)	(3.55)	(3.16)	(4.08)

Panel B: Information uncertainty (INFO-score)

MISPRICE-score	High costs	4	3	2	Low costs	High-low
Overpriced	-0.95	-0.56	-0.43	-0.22	-0.35	-0.60
	(-5.76)	(-3.61)	(-2.70)	(-1.77)	(-3.46)	(-3.41)
2	-0.13	0.00	-0.13	-0.15	-0.11	-0.02
	(-0.92)	(0.00)	(-1.13)	(-1.32)	(-1.13)	(-0.09)
3	0.16	-0.17	-0.03	0.04	0.17	-0.01
	(0.99)	(-1.21)	(-0.28)	(0.50)	(2.49)	(-0.05)
4	-0.19	0.30	-0.03	0.14	0.16	-0.35
	(-1.22)	(1.98)	(-0.29)	(1.54)	(2.13)	(-1.91)
Underpriced	0.25	0.33	0.14	0.24	0.30	-0.05
	(1.65)	(2.64)	(1.16)	(2.42)	(4.35)	(-0.33)
Long-short	1.20	0.88	0.57	0.46	0.65	0.55
	(5.02)	(3.88)	(2.67)	(2.81)	(5.07)	(2.26)

Panel C: Transaction costs (TCOST-score)

MISPRICE-score	High costs	4	3	2	Low costs	High-low
Overpriced	-0.80	-0.63	-0.50	-0.46	-0.32	-0.48
	(-6.82)	(-4.96)	(-3.61)	(-3.83)	(-3.39)	(-4.13)
2	-0.24	-0.11	0.00	-0.12	-0.15	-0.09
	(-2.12)	(-1.04)	(-0.04)	(-1.51)	(-1.98)	(-0.79)
3	-0.08	-0.03	0.06	0.11	0.11	-0.19
	(-0.77)	(-0.34)	(0.58)	(1.07)	(2.11)	(-1.61)
4	0.06	0.14	0.14	0.11	0.19	-0.12
	(0.55)	(1.46)	(1.70)	(1.18)	(3.09)	(-0.95)
Underpriced	0.36	0.20	0.27	0.27	0.23	0.13
	(2.96)	(1.98)	(2.77)	(3.24)	(3.89)	(0.94)
Long-short	1.16	0.83	0.77	0.73	0.56	0.61
	(7.41)	(5.37)	(4.78)	(4.77)	(4.54)	(3.84)

Table 5. Fama-MacBeth regressions: Mispricing, overall arbitrage costs, arbitrage risk, information uncertainty, and transaction costs

This table reports regressions of monthly returns on MISPRICE-score and MISPRICE-score interacted with ARBCOST-score, ARBRISK, INFO-score, and TCOST-score. Cross-sectional regressions are estimated with OLS (Panel A) or WLS that weighs an observation by its prior-month market equity (Panel B). Slopes are multiplied by 10. The sample period is from 07/1963 to 12/2015. Newey-West *t*-statistics are shown in parenthesis.

Panel A: OLS

	(1)	(2)	(3)	(4)	(5)	(6)
MISPRICE-score	2.183	-0.556	0.614	0.023	0.284	-0.559
	(9.28)	(-1.09)	(2.44)	(0.05)	(0.66)	(-1.06)
MISPRICE-score × ARBCOST-score		0.476				
		(4.91)				
MISPRICE-score × ARBRISK			0.264			0.177
			(4.94)			(3.22)
MISPRICE-score × INFO-score				0.383		0.173
				(4.50)		(2.22)
MISPRICE-score × TCOST-score					0.332	0.125
					(4.40)	(1.68)
ARBCOST-score		-3.165				
		(-4.15)				
ARBRISK			-1.472			-0.877
			(-3.98)			(-2.38)
INFO-score				-2.866		-1.722
				(-4.97)		(-3.45)
TCOST-score					-1.885	-0.668
					(-3.77)	(-1.43)
BETA	0.717	0.961	0.707	1.008	0.756	0.814
	(0.68)	(1.03)	(0.93)	(0.98)	(0.74)	(1.09)
ME	-0.528	-0.896	-0.510	-0.843	-0.507	-0.679
	(-1.55)	(-3.11)	(-1.82)	(-2.62)	(-1.53)	(-2.26)
B/M	2.082	1.807	1.844	1.880	1.951	1.754
	(3.41)	(2.99)	(3.29)	(3.12)	(3.21)	(3.16)
Adjusted R^2	4.05%	4.55%	4.85%	4.22%	4.35%	5.13%
Number of firm-month	1,255,049	1,255,049	1,255,049	1,255,049	1,255,049	1,255,049

Table 5 (Continued)

Panel B: WLS

	(1)	(2)	(3)	(4)	(5)	(6)
MISPRICE-score	1.434	0.127	0.607	1.401	0.165	1.305
	(5.38)	(0.20)	(1.76)	(1.85)	(0.27)	(1.70)
MISPRICE-score × ARBCOST-score		0.355				
		(2.39)				
MISPRICE-score × ARBRISK			0.245			0.216
			(2.80)			(2.15)
MISPRICE-score × INFO-score				0.038		-0.183
				(0.26)		(-1.30)
MISPRICE-score × TCOST-score					0.375	0.084
					(2.58)	(0.52)
ARBCOST-score		-2.333				
		(-2.51)				
ARBRISK			-1.162			-0.812
			(-2.38)			(-1.42)
INFO-score				-0.593		0.644
				(-0.65)		(0.78)
TCOST-score					-2.483	-1.186
					(-2.98)	(-1.21)
BETA	0.159	0.248	-0.147	0.303	0.261	-0.153
	(0.11)	(0.25)	(-0.12)	(0.21)	(0.19)	(-0.13)
ME	-0.737	-0.902	-0.602	-0.854	-0.892	-0.903
	(-2.55)	(-2.69)	(-1.84)	(-2.92)	(-2.53)	(-2.47)
B/M	0.848	0.758	0.712	0.755	0.761	0.686
	(1.17)	(1.04)	(0.97)	(1.0045)	(1.06)	(0.96)
Adjusted R^2	9.69%	11.00%	11.32%	10.94%	10.60%	13.11%
Number of firm-month	1,255,049	1,255,049	1,255,049	1,255,049	1,255,049	1,255,049

Table 6. Time-varying principal components of arbitrage costs

Panel A presents year-by-year, average, and cumulative (starting from PC1) percentages of total variance explained by each of the nine principal components (PC1 to PC9). Panel B presents year-by-year and average eigenvalues of each principal component. Panel C presents average eigenvectors of the first, second, and third principal components. The principal components are extracted from nine arbitrage costs measures at the end of June between 1983 and 2015.

Panel A: Proportions of total variance explained by a principal component

Year	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
1983	28%	15%	13%	11%	10%	8%	7%	6%	4%
1984	29%	16%	11%	11%	10%	8%	6%	5%	4%
1985	31%	15%	11%	11%	10%	9%	6%	4%	3%
1986	33%	15%	11%	10%	10%	9%	6%	4%	3%
1987	33%	16%	11%	10%	9%	8%	6%	4%	3%
1988	35%	14%	11%	10%	9%	8%	6%	3%	3%
1989	36%	14%	11%	10%	8%	8%	6%	4%	3%
1990	34%	15%	11%	10%	9%	8%	6%	4%	3%
1991	35%	15%	11%	10%	8%	8%	6%	3%	3%
1992	35%	15%	11%	10%	9%	7%	6%	3%	3%
1993	34%	15%	11%	10%	9%	8%	6%	4%	3%
1994	35%	15%	11%	11%	8%	7%	6%	4%	3%
1995	36%	15%	11%	11%	8%	7%	5%	4%	3%
1996	34%	16%	11%	11%	8%	7%	6%	4%	2%
1997	33%	16%	11%	11%	9%	7%	6%	5%	2%
1998	31%	18%	11%	11%	10%	7%	6%	4%	2%
1999	31%	18%	11%	11%	9%	7%	6%	4%	2%
2000	26%	21%	11%	11%	10%	8%	7%	5%	2%
2001	26%	21%	11%	11%	9%	8%	7%	4%	2%
2002	28%	19%	11%	11%	10%	8%	7%	4%	2%
2003	28%	18%	11%	11%	10%	8%	7%	4%	3%
2004	30%	16%	11%	11%	10%	8%	7%	4%	2%
2005	30%	16%	11%	11%	10%	8%	7%	4%	3%
2006	29%	16%	11%	11%	10%	8%	8%	4%	3%
2007	30%	15%	11%	11%	10%	8%	7%	5%	3%
2008	27%	14%	11%	11%	11%	10%	8%	5%	3%
2009	29%	14%	11%	11%	11%	9%	7%	5%	3%
2010	31%	16%	12%	10%	9%	7%	6%	5%	4%
2011	30%	15%	11%	11%	10%	7%	6%	5%	4%
2012	30%	16%	12%	11%	10%	7%	6%	5%	3%
2013	30%	17%	11%	11%	10%	7%	7%	5%	2%
2014	31%	18%	11%	10%	10%	8%	6%	5%	2%
2015	32%	18%	11%	10%	9%	8%	5%	5%	2%
Average	31%	16%	11%	11%	10%	8%	6%	4%	3%
Cumulative	31%	47%	59%	69%	79%	87%	93%	97%	100%

Table 6 (Continued)

Panel B: Eigenvalues of a principal component

Year	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
1983	2.53	1.35	1.13	1.01	0.86	0.70	0.60	0.50	0.32
1984	2.64	1.44	1.00	0.96	0.93	0.73	0.51	0.43	0.34
1985	2.77	1.33	1.00	0.97	0.89	0.82	0.52	0.40	0.31
1986	2.97	1.34	1.00	0.92	0.86	0.79	0.51	0.36	0.26
1987	2.99	1.41	1.00	0.92	0.84	0.74	0.50	0.35	0.26
1988	3.11	1.30	1.00	0.91	0.81	0.76	0.54	0.31	0.25
1989	3.22	1.24	1.00	0.94	0.74	0.70	0.53	0.33	0.29
1990	3.06	1.33	1.01	0.93	0.80	0.70	0.56	0.33	0.27
1991	3.16	1.32	1.01	0.94	0.75	0.69	0.54	0.31	0.27
1992	3.14	1.39	1.01	0.94	0.77	0.65	0.54	0.28	0.27
1993	3.08	1.32	1.01	0.94	0.83	0.68	0.54	0.32	0.28
1994	3.15	1.33	1.01	0.96	0.76	0.67	0.53	0.32	0.28
1995	3.21	1.35	1.01	0.95	0.74	0.66	0.48	0.33	0.28
1996	3.06	1.45	1.02	0.99	0.74	0.65	0.50	0.36	0.22
1997	2.95	1.46	1.00	0.98	0.83	0.65	0.50	0.41	0.21
1998	2.82	1.58	1.01	0.98	0.86	0.66	0.54	0.36	0.20
1999	2.75	1.64	1.02	0.96	0.81	0.64	0.57	0.40	0.19
2000	2.32	1.91	1.00	0.96	0.87	0.73	0.63	0.43	0.14
2001	2.34	1.89	1.02	1.00	0.85	0.68	0.64	0.39	0.20
2002	2.50	1.70	1.02	0.97	0.90	0.69	0.60	0.39	0.22
2003	2.53	1.59	1.01	0.97	0.92	0.73	0.62	0.38	0.25
2004	2.68	1.46	1.01	0.98	0.93	0.73	0.59	0.40	0.20
2005	2.67	1.46	1.01	0.99	0.94	0.70	0.59	0.40	0.24
2006	2.63	1.41	1.00	0.96	0.91	0.73	0.70	0.39	0.27
2007	2.71	1.39	1.01	1.00	0.88	0.69	0.65	0.42	0.26
2008	2.47	1.26	1.01	1.00	0.98	0.87	0.69	0.45	0.26
2009	2.60	1.22	1.02	1.00	0.98	0.83	0.64	0.41	0.28
2010	2.75	1.44	1.05	0.93	0.85	0.67	0.54	0.41	0.36
2011	2.69	1.38	1.01	0.98	0.92	0.64	0.56	0.43	0.40
2012	2.71	1.46	1.05	0.95	0.94	0.61	0.57	0.42	0.28
2013	2.74	1.51	1.01	0.95	0.93	0.64	0.59	0.47	0.17
2014	2.80	1.61	1.00	0.92	0.86	0.69	0.53	0.44	0.16
2015	2.88	1.61	1.02	0.91	0.80	0.69	0.49	0.42	0.18
Average	2.81	1.45	1.02	0.96	0.86	0.70	0.57	0.39	0.25

Panel C: Average eigenvectors of the first, second, and third principal components

	PC1	PC2	PC3
IVOL	0.45	0.38	-0.01
CFVOL	0.30	0.37	0.02
DISP	0.05	0.09	0.42
COV	-0.41	0.39	-0.01
PRICE	-0.39	0.15	0.00
DVOL	-0.29	0.47	0.02
ILLIQ	0.15	-0.14	0.16
BIDASK	0.44	0.35	0.01
IO	0.25	-0.31	-0.02

Table 7. The impact of time-varying principal components of arbitrage costs on mispricing

This table reports monthly Fama-French three-factor alphas (%) for 25 value-weighted portfolios independently sorted by the first, second, or third principal component of arbitrage costs (PC1, PC2, or PC3) and MISPRICE-score. The principal components are extracted from nine arbitrage costs measures at the end of June between 1983 and 2015. Long-short is the trading strategy which longs the underpriced portfolio and shorts the overpriced portfolio. High-low is the trading strategy which longs the high arbitrage costs portfolio and shorts the low arbitrage costs portfolio. The sample period is from 07/1983 to 12/2015. Newey-West *t*-statistics are shown in parenthesis.

Panel A: First principal component of arbitrage costs (PC1)

MISPRICE-score	High costs	4	3	2	Low costs	High-low
Overpriced	-1.39	-0.92	-0.60	-0.40	-0.23	-1.16
	(-4.82)	(-3.88)	(-2.76)	(-1.95)	(-1.95)	(-4.02)
2	-0.62	-0.23	0.14	0.02	-0.34	-0.28
	(-2.84)	(-1.40)	(0.90)	(0.20)	(-1.90)	(-1.05)
3	-0.04	-0.05	0.33	0.04	0.24	-0.28
	(-0.19)	(-0.20)	(2.01)	(0.29)	(3.16)	(-1.37)
4	-0.10	-0.04	0.05	0.08	0.09	-0.19
	(-0.45)	(-0.27)	(0.28)	(0.59)	(1.40)	(-0.84)
Underpriced	0.83	0.07	0.23	0.14	0.30	0.54
	(3.57)	(0.36)	(1.55)	(1.28)	(3.29)	(2.30)
Long-short	2.22	0.99	0.83	0.54	0.52	1.70
	(5.26)	(3.39)	(3.29)	(2.26)	(3.56)	(4.28)

Panel B: Second principal component of arbitrage costs (PC2)

MISPRICE-score	High costs	4	3	2	Low costs	High-low
Overpriced	-0.74	-0.44	-0.20	0.08	0.14	-0.88
	(-4.18)	(-2.49)	(-1.10)	(0.64)	(1.03)	(-3.84)
2	-0.46	-0.07	-0.03	0.15	0.29	-0.74
	(-2.29)	(-0.55)	(-0.28)	(1.13)	(2.69)	(-3.13)
3	0.20	0.08	0.15	0.27	0.25	-0.05
	(1.38)	(0.63)	(1.16)	(2.35)	(2.13)	(-0.23)
4	0.06	0.04	0.26	0.27	0.21	-0.16
	(0.62)	(0.37)	(2.44)	(2.15)	(1.88)	(-1.17)
Underpriced	0.31	0.16	0.26	0.36	0.34	-0.03
	(2.69)	(1.65)	(1.74)	(2.54)	(1.96)	(-0.18)
Long-short	1.05	0.60	0.46	0.28	0.20	0.85
	(5.03)	(3.15)	(2.76)	(2.10)	(1.08)	(3.66)

Panel C: Third principal component of arbitrage costs (PC3)

MISPRICE-score	High costs	4	3	2	Low costs	High-low
1: Overpriced	0.02	-0.27	-0.55	-0.37	-0.95	0.97
	(0.10)	(-1.38)	(-3.56)	(-2.14)	(-4.17)	(3.59)
2	0.27	0.10	-0.05	-0.11	-0.65	0.92
	(1.52)	(0.79)	(-0.38)	(-0.69)	(-3.13)	(3.09)
3	0.57	0.31	0.09	0.18	-0.34	0.91
	(3.83)	(2.20)	(0.63)	(1.64)	(-2.11)	(3.62)
4	0.20	0.13	0.10	0.39	-0.02	0.22
	(1.52)	(1.13)	(1.02)	(3.10)	(-0.13)	(1.02)
5: Underpriced	0.41	0.36	0.20	0.28	0.05	0.35
	(3.24)	(3.23)	(1.58)	(2.24)	(0.35)	(1.87)
Long-short	0.39	0.64	0.76	0.65	1.00	-0.62
	(1.94)	(2.66)	(4.38)	(3.77)	(4.11)	(-2.46)

Table 8. Fama-MacBeth regressions: Mispricing and time-varying principal components of arbitrage costs

This table reports regressions of monthly returns on MISPRICE-score and MISPRICE-score interacted with deciles of the first, second, and third principal components of arbitrage costs (PC1_dec, PC2_dec, and PC3_dec). The principal components are extracted from nine arbitrage costs measures at the end of June between 1983 and 2015. The deciles are independently sorted at the end of June each year. Cross-sectional regressions in columns (1) to (5) are estimated with OLS and regressions in columns (6) to (10) are estimated with WLS that weighs an observation by its prior-month market equity. Slopes are multiplied by 10. The sample period is from 07/1983 to 12/2015. Newey-West *t*-statistics are shown in parenthesis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
MISPRICE-score	2.300	0.329	0.684	2.314	-0.631	1.280	0.835	0.528	2.189	-0.745
	(6.81)	(0.85)	(2.01)	(5.03)	(-1.26)	(3.39)	(1.73)	(0.86)	(3.56)	(-1.00)
MISPRICE-score × PC1_dec		0.344			0.264		0.155			0.200
		(4.55)			(3.34)		(1.83)			(2.21)
MISPRICE-score × PC2_dec			0.282		0.249			0.109		0.198
			(4.59)		(4.45)			(1.17)		(2.03)
MISPRICE-score × PC3_dec				-0.002					-0.152	
				(-0.04)					(-1.82)	
PC1_dec		-2.011			-1.457		-0.792			-1.170
		(-4.48)			(-3.21)		(-1.45)			(-2.33)
PC2_dec			-2.034		-1.828			-1.097		-1.523
			(-4.97)		(-4.83)			(-1.68)		(-2.36)
PC3_dec				0.296					1.404	
				(0.82)					(2.59)	
BETA	0.171	0.277	0.596	0.252	0.570	0.379	0.280	0.673	0.571	0.692
	(0.13)	(0.22)	(0.49)	(0.19)	(0.49)	(0.20)	(0.15)	(0.37)	(0.31)	(0.39)
ME	0.162	0.112	0.247	0.144	0.268	-0.404	-0.346	-0.144	-0.171	-0.210
	(0.56)	(0.41)	(0.80)	(0.50)	(0.94)	(-1.04)	(-0.90)	(-0.35)	(-0.44)	(-0.53)
B/M	1.873	1.764	1.569	1.823	1.525	0.373	0.329	0.196	0.342	0.197
	(2.54)	(2.40)	(2.17)	(2.52)	(2.10)	(0.44)	(0.39)	(0.23)	(0.42)	(0.23)
Adjusted R^2	3.21%	3.57%	3.59%	3.40%	3.84%	8.91%	9.90%	10.21%	10.01%	10.91%
Number of firm-month	881,702	881,702	881,702	881,702	881,702	881,702	881,702	881,702	881,702	881,702

Table 9. Subperiod Fama-MacBeth regressions: Mispricing and nine arbitrage costs measures

This table reports regressions of monthly returns on MISPRICE-score and MISPRICE-score interacted with deciles of nine arbitrage costs measures (ARBRISK, CFVOL_dec, DISP_dec, COV_dec, PRICE_dec, DVOL_dec, ILLIQ_dec, BIDASK_dec, and IO_dec) over three subperiods. The deciles are independently sorted at the end of June each year. Cross-sectional regressions are estimated either with OLS (Panels A and C) or WLS that weighs an observation by its prior-month market equity (Panels B and D). Slopes are multiplied by 10. Newey-West *t*-statistics are shown in parenthesis.

Panel A: MISPRICE-score regressions (OLS)

	(1)	(2)	(3)
	07/1963-06/1983	07/1983-12/2004	01/2005-12/2015
MISPRICE-score	2.019	2.951	0.835
	(7.40)	(7.91)	(2.21)
BETA	1.290	-0.066	0.704
	(0.75)	(-0.04)	(0.39)
ME	-1.518	0.241	-0.017
	(-2.23)	(0.61)	(-0.05)
B/M	2.744	2.640	0.147
	(2.56)	(2.88)	(0.14)
Adjusted R^2	5.30%	3.67%	2.36%
Number of firm-month	388,174	640,321	241,381

Panel B: MISPRICE-score regressions (WLS)

	(1)	(2)	(3)
	07/1963-06/1983	07/1983-12/2004	01/2005-12/2015
MISPRICE-score	1.753	1.496	0.796
	(5.37)	(3.29)	(1.21)
BETA	-0.461	0.195	0.793
	(-0.20)	(0.08)	(0.30)
ME	-1.226	-0.354	-0.517
	(-3.18)	(-0.71)	(-0.88)
B/M	1.881	0.808	-0.607
	(1.46)	(0.74)	(-0.51)
Adjusted R^2	10.98%	9.22%	8.22%
Number of firm-month	388,174	640,321	241,381

Table 9 (Continued)

Panel C: Interactive MISPRICE-score regressions (OLS)

	(1)	(2)	(3)
	07/1963-06/1983	07/1983-12/2004	01/2005-12/2015
MISPRICE-score	1.216	-1.004	-3.186
	(2.26)	(-1.30)	(-2.70)
MISPRICE-score × ARBRISK	0.016	0.348	-0.172
	(0.19)	(4.89)	(-1.39)
MISPRICE-score × CFVOL_dec	0.090	0.097	-0.077
	(1.68)	(1.82)	(-0.96)
MISPRICE-score × DISP_dec		-0.021	0.126
		(-0.30)	(1.38)
MISPRICE-score × COV_dec		0.039	0.027
		(0.60)	(0.47)
MISPRICE-score × PRICE_dec	0.046	0.278	0.218
	(0.51)	(2.97)	(1.66)
MISPRICE-score × DVOL_dec	-0.215	-0.106	-0.799
	(-0.88)	(-0.45)	(-2.02)
MISPRICE-score × ILLIQ_dec	0.146	0.004	0.828
	(0.57)	(0.02)	(1.87)
MISPRICE-score × BIDASK_dec	0.037	-0.119	0.155
	(0.47)	(-1.48)	(1.28)
MISPRICE-score × IO_dec		0.140	0.339
		(2.85)	(4.94)
ARBRISK	0.358	-1.917	0.948
	(0.58)	(-4.69)	(1.19)
CFVOL_dec	-0.641	-0.736	0.234
	(-1.88)	(-2.20)	(0.55)
DISP_dec		-0.589	-1.348
		(-1.39)	(-2.65)
COV_dec		-0.372	-0.295
		(-0.97)	(-0.79)
PRICE_dec	-0.298	-1.653	-0.970
	(-0.47)	(-2.68)	(-0.95)
DVOL_dec	2.003	1.552	4.627
	(1.34)	(1.14)	(1.96)

Table 9 (Continued)

ILLIQ_dec	-1.388	-0.958	-4.488
	(-0.95)	(-0.80)	(-1.59)
BIDASK_dec	-0.019	0.969	-0.998
	(-0.04)	(1.92)	(-1.20)
IO_dec		-1.021	-2.204
		(-3.51)	(-6.14)
BETA	1.017	0.327	1.002
	(0.89)	(0.35)	(0.77)
ME	-1.061	0.173	0.417
	(-1.46)	(0.34)	(0.33)
B/M	2.354	2.274	-0.652
	(2.50)	(3.51)	(-0.57)
Adjusted R^2	7.41%	5.40%	4.43%
Number of firm-month	373,347	640,321	241,381

Table 9 (Continued)

Panel D: Interactive MISPRICE-score regressions (WLS)

	(1)	(2)	(3)
	07/1963-06/1983	07/1983-12/2004	01/2005-12/2015
MISPRICE-score	1.681	-0.580	-0.391
	(2.61)	(-0.50)	(-0.26)
MISPRICE-score × ARBRISK	-0.008	0.354	-0.135
	(-0.08)	(2.01)	(-0.50)
MISPRICE-score × CFVOL_dec	-0.129	0.063	0.027
	(-1.51)	(0.63)	(0.22)
MISPRICE-score × DISP_dec		-0.028	0.039
		(-0.21)	(0.26)
MISPRICE-score × COV_dec		-0.090	0.060
		(-0.92)	(0.68)
MISPRICE-score × PRICE_dec	0.041	0.118	0.088
	(0.26)	(0.80)	(0.41)
MISPRICE-score × DVOL_dec	-0.036	0.165	-0.309
	(-0.13)	(0.56)	(-0.59)
MISPRICE-score × ILLIQ_dec	0.066	-0.265	0.101
	(0.23)	(-0.90)	(0.16)
MISPRICE-score × BIDASK_dec	0.096	0.154	0.050
	(0.87)	(0.71)	(0.16)
MISPRICE-score × IO_dec		0.035	0.124
		(0.30)	(1.15)
ARBRISK	0.218	-1.290	1.132
	(0.31)	(-1.36)	(0.72)
CFVOL_dec	0.555	-0.517	-0.169
	(1.13)	(-0.91)	(-0.25)
DISP_dec		-0.671	-0.634
		(-0.99)	(-0.67)
COV_dec		0.900	-0.401
		(1.43)	(-0.76)
PRICE_dec	-0.420	-0.502	-0.645
	(-0.40)	(-0.53)	(-0.47)
DVOL_dec	-0.264	-0.955	2.117
	(-0.17)	(-0.48)	(0.70)

Table 9 (Continued)

ILLIQ_dec	0.487	0.339	-0.926
	(0.29)	(0.19)	(-0.26)
BIDASK_dec	-0.388	-1.095	-0.921
	(-0.62)	(-0.73)	(-0.46)
IO_dec		-0.533	-0.935
		(-0.86)	(-1.64)
BETA	-0.897	-0.564	2.002
	(-0.46)	(-0.40)	(0.87)
ME	-0.958	-0.583	-0.714
	(-1.99)	(-0.77)	(-0.79)
B/M	1.532	1.749	-0.738
	(1.18)	(1.86)	(-0.75)
Adjusted R^2	16.68%	16.10%	15.92%
Number of firm-month	373,347	640,321	241,381

Table 10. An empirical model of arbitrage costs versus the first and second time-varying principal components of arbitrage costs

This table reports regressions of monthly returns on MISPRICE-score and MISPRICE-score interacted with deciles of arbitrage risk (ARBRISK), cash flow volatility (CFVOL_dec), share price (PRICE_dec), passive institutional ownership (IO_dec), and the first and second principal components of arbitrage costs (PC1_dec and PC2_dec) over two subperiods. The deciles are independently sorted at the end of June each year. Cross-sectional regressions are estimated with OLS (Panel A) or WLS that weighs an observation by its prior-month market equity (Panel B). Slopes are multiplied by 10. Newey-West *t*-statistics are shown in parenthesis.

Panel A: OLS

	(1)	(2)	(3)	(4)	(5)	(6)
	07/1983-12/2004	01/2005-12/2015	07/1983-12/2004	01/2005-12/2015	07/1983-12/2004	01/2005-12/2015
MISPRICE-score	-1.687	-2.772	-0.556	-0.800	-1.990	-3.074
	(-2.39)	-(3.09)	(-0.86)	(-1.20)	(-2.45)	(-3.56)
MISPRICE-score × ARBRISK	0.324				0.283	
	(4.36)				(3.50)	
MISPRICE-score × CFVOL_dec	0.103				0.078	
	(1.87)				(1.39)	
MISPRICE-score × PRICE_dec	0.194	0.277			0.191	0.268
	(2.70)	(2.52)			(2.77)	(2.05)
MISPRICE-score × IO_dec	0.174	0.336			0.171	0.278
	(3.64)	(5.82)			(3.51)	(4.61)
MISPRICE-score × PC1_dec			0.365	0.036	0.057	-0.013
			(3.68)	(0.42)	(0.72)	(-0.13)
MISPRICE-score × PC2_dec			0.257	0.233	0.073	0.123
			(3.72)	(2.44)	(0.99)	(1.27)
ARBRISK	-1.783				-1.511	
	(-4.51)				(-3.51)	
CFVOL_dec	-0.798				-0.624	
	(-2.35)				(-1.85)	
PRICE_dec	-1.247	-1.570			-1.220	-1.407
	(-2.59)	(-1.69)			(-2.66)	(-1.37)
IO_dec	-1.231	-2.211			-1.186	-1.795
	(-4.27)	(-6.51)			(-4.00)	(-5.23)
PC1_dec			-1.959	-0.328	-0.311	-0.205
			(-3.45)	(-0.56)	(-0.66)	(-0.31)
PC2_dec			-1.765	-1.968	-0.593	-1.268
			(-3.91)	(-2.80)	(-1.31)	(-1.78)
BETA	0.277	0.537	0.273	1.240	0.336	1.072
	(0.26)	(0.32)	(0.18)	(0.74)	(0.32)	(0.65)

Table 10 (Continued)

ME	-0.022	0.088	0.385	0.007	0.085	0.052
	(-0.06)	(0.29)	(1.00)	(0.02)	(0.23)	(0.17)
B/M	1.975	-0.432	2.411	-0.466	1.968	-0.766
	(2.73)	(-0.40)	(2.87)	(-0.39)	(2.72)	(0.67)
Adjusted R^2	4.76%	2.84%	4.17%	3.07%	4.90%	3.40%
Number of firm-month	640,321	241,381	640321	241,381	640,321	241,381

Panel B: WLS

	(1)	(2)	(3)
	07/1983-12/2004	07/1983-12/2004	07/1983-12/2004
MISPRICE-score	-0.065	-1.314	-0.696
	(-0.11)	(-1.41)	(-0.76)
MISPRICE-score × ARBRISK	0.517		0.500
	(3.43)		(2.56)
MISPRICE-score × PC1_dec		0.289	-0.073
		(2.49)	(-0.48)
MISPRICE-score × PC2_dec		0.269	0.128
		(2.26)	(1.01)
ARBRISK	-2.580		-2.269
	(-3.50)		(-2.46)
PC1_dec		-1.475	0.363
		(-2.33)	(0.42)
PC2_dec		-1.860	-1.207
		(-2.43)	(-1.49)
BETA	-0.403	0.388	-0.594
	(-0.24)	(0.17)	(-0.34)
ME	-0.211	-0.115	0.179
	(-0.34)	(-0.23)	(0.30)
B/M	0.710	0.628	0.680
	(0.65)	(0.58)	(0.63)
Adjusted R^2	10.92%	11.10%	12.44%
Number of firm-month	640,321	640,321	640,321