Is Informational Inefficiency Priced in Stock Markets? A Comparison between the U.S. and Chinese Cases

Abstract

Using a sample of U.S. and Chinese stocks between July 1999 and June 2016, we investigate the pricing role of informational inefficiency in stock markets. We find that the relations between returns and the informational inefficiency factor statistically change from significantly positive, to insignificant, and further to significantly negative as informational efficiency increases. This finding provides new insights into the common belief that emerging markets are less efficient than developed markets. We propose new factor models for less efficient markets. Our conclusions are robust to altering the ways of sorting portfolios, to various subsample analyses, and to alternative factor models.

"Market efficiency remains central to the study of financial markets, but while several measures of stock price efficiency have been proposed, little is known about how these measures vary over time, whether they comove across stocks and with each other, and what economic forces drive time variation in systematic market efficiency."   (Rösch et al., 2017)

1. Introduction

Informational efficiency is closely related with the degree to which stock prices correctly and quickly reflect information and thus the true value of an underlying asset (e.g., Fama, 1970; Bai et al., 2016; Rösch et al., 2017). There is a long-standing debate on whether financial markets are informationally efficient between the efficient markets hypothesis (EMH) and behavioral finance (e.g., Fama, 1991; Jacobs, 2016; Stambaugh and Yuan, 2017). Against the backdrop of this debate, Lo (2004, 2005) argues that the informational efficiency of a market is time-varying and driven by the intrinsic rules of economic selection, known as the adaptive markets hypothesis (AMH). Following the AMH, it is increasingly acknowledged that the informational efficiency of a market is changing over time (e.g., Neely et al., 2009 and the references therein). While informational efficiency is well discussed at the market level, there are few attempts to study it at the portfolio level. We bridge this gap by investigating whether informational inefficiency, at the stock portfolio level, is priced in the U.S. and Chinese stock markets, what the mechanisms are, and the difference between these two markets in terms of informational efficiency.

This study is motivated by the time-varying feature of portfolio efficiency in the spirit of the AMH and mixed evidence on whether developed markets are more informationally efficient than emerging ones. In searching for evidence on the

hypotheses that informational inefficiency is priced in excess returns and that the U.S. stock market is more efficient than the Chinese market, we examine informational inefficiency in the U.S. and Chinese stock markets between July 1999 and June 2016 using standard and informational inefficiency augmented factor models. We find that the relations between informational inefficiency and excess returns in both markets change from positive to negative as informational efficiency improves, which is consistent with most of the AMH implications[i]. In view of this mechanism, we argue that the U.S. market is more efficient than the Chinese market. Moreover, the informational inefficiency augmented models perform better in the less efficient Chinese market. Finally, we provide the evidence that these findings are robust to alternative ways of sorting portfolios, to various subsample analyses, and to different factor models.

The present study complements and extends existing work in at least four directions. First, while the literature on informational efficiency is mostly at the market level (e.g., Lo, 2004, 2005), we take it one step further to the portfolio level by examining whether informational inefficiency is priced via sorting portfolios. The recent literature applies various linear and non-linear tests to one or several market indices and documents the existence of informational inefficiency, while empirical evidence at the portfolio level is scarce, especially on the relationship between informational inefficiency and returns.

Moreover, we provide fresh evidence supporting the theoretical argument that mispricing induces the return premium (e.g., Brennan and Wang, 2010) through an

inverse feedback mechanism in the U.S. and Chinese stock markets. Specifically, we find that when informational inefficiency is low, the negative relation between returns and informational inefficiency stops investors arbitraging for informational inefficiency, which makes informational inefficiency expand even further. As informational inefficiency reaches a relatively high level, the relation turns positive, which drives investors to trade against informational inefficiency to make the market more informationally efficient.

In addition, we contribute to research on multifactor asset pricing models by introducing the informational inefficiency factor (IIF), which characterizes the evolution of market mispricing. Different from the mispricing factors in Stambaugh and Yuan (2017), which are synthesized from existing anomalies, we construct the IIF by taking advantage of entropy to characterize collective market imperfections given that anomalies change over time (Jones and Pomorski, 2017).

Finally, our study adds to the existing literature on the comparison of the informational efficiency between developed markets and emerging ones, which provides new insights into market efficiency. In contrast to the findings of Jacobs (2016), we suggest that the U.S. stock market is more informationally efficient than the Chinese stock market since it has a stronger inverse feedback mechanism.

The rest of this paper is organized as follows. Section 2 briefly summarizes the debate on market (in)efficiency. Section 3 presents the data and methodologies used in this study. Section 4 reports the empirical results. Section 5 concludes.

2. Background literature

Informational efficiency, which is so-called market efficiency at the market level, has been discussed in the literature. Fama (1970, 1991) proposes the EMH, which evolves from the random walk theory of asset prices. The role that information plays in price formation is termed "market efficiency." The EMH assumes that security prices at any time "fully reflect" all available information under the assumption that investors are rational and homogeneous in a frictionless market. Three forms of market efficiency are proposed—weak, semi-strong, and strong—based on three information sets. The weak-form EMH implies that market prices reflect all information about past prices. The information set further includes publicly available information under the semi-strong-form EMH and privately held information under the strong-form EMH.

Evidence against the EMH has accumulated. For instance, Simon (1955) argues that investors have bounded rationality and make satisfactory instead of optimal choices because of costly optimization and limited computational abilities. Kahneman and Tversky (1979) argue that investors tend to be risk averse when faced with gains and risk seeking when faced with losses, which casts doubt on the expected utility theory dealing with decision making under risk. Following Kahneman and Tversky (1979), psychologists and economists have found that many types of behavioral biases lead investors to make poor decisions (Bailey et al., 2011). The types of biases that result in informational inefficiency include overconfidence

(Gervais and Odean, 2001; Gervais et al., 2011; Radzevick and Moore, 2011), over-reaction and under-reaction (Bondt and Thaler, 1985; Chan, 2003; Jiang and Zhu, 2017), loss aversion (Shefrin and Statman, 1985; Odean, 1998; Bodnaruk and Simonov, 2016), herding (Huberman and Regev, 2001; Choi and Sias, 2009), home bias (Lin and Viswanathan, 2015), psychological accounting (Grinblatt and Han, 2005), and regret (Bell, 1983).

To reconcile the EMH and aforementioned behavioral finance studies, Lo (2004, 2005) proposes the AMH from the evolutionary perspective. Lo (2004, 2005) argues that natural selection determines the survival of the fittest in a dynamic market and that investors learn from their mistakes to adapt to the ever-changing environment. Within the AMH framework, informational efficiency is a time-varying concept rather than an all-or-nothing one that evolves with dynamic market conditions and determines the evolving nature of return predictability.

Following the seminal works of Lo (2004, 2005), a small but growing strand of studies has supported the AMH. Various modified linear and non-linear tests have been applied to test the AMH and the statistics of these tests or the measurements based on them have been used to measure market efficiency or relative return predictability. For instance, Ito and Sugiyama (2009) find that the relative inefficiency of the U.S. stock market is time-varying from 1955 to 2006 by taking the time-varying autocorrelation of stock returns as a proxy of market inefficiency. To detect non-linear dependence, Kim et al. (2011) introduce the generalized spectral test, which can capture market inefficiency both linearly and non-linearly. They find that

6

the return predictability of the Dow Jones Industrial Average is time-varying and driven by changing market conditions. Similar studies include but are not limited to Charles et al. (2012), Urquhart and Hudson (2013), Levich and Potì (2015), and Urquhart and McGroarty (2016).

Overall, recent studies have focused on informational efficiency at the market level. A single market is taken as a whole, and its index is used as an item in a sample. From this macro perspective, markets are not fully informationally efficient and market efficiency changes periodically. To the best of our knowledge, no study in the existing literature investigates the effect of informational inefficiency on stock returns, especially at portfolio levels. This is an important task, given the background of the debate between the EMH and AMH, and our study bridges this gap.

3. Methodology and data

In this section, we first describe the methods used to measure informational efficiency and our empirical approaches and then present the data description and preliminary analysis.

3.1. Measurement of informational efficiency

One key aspect of our study is measuring informational efficiency. As stated by the EMH, stock returns follow a random walk process and are unforecastable in a fully informationally efficient market because of investors' constant arbitrage activities. This principle indicates that the extent to which a return series is random is the very

7

essence of informational efficiency. In this spirit, various measurements have been proposed by the recent literature, including entropy-based measurements (Alvarez-Ramirez et al., 2012), variance ratios (Lo and MacKinlay, 1989), return predictability (Boehmer and Wu, 2007), pricing error relative to the efficient price (Hasbrouck, 1993), and the extent to which markets obey the law of one price (Cremers and Weinbaum, 2010). Among these measurements, we choose an entropy measurement for our research for the following reasons.

Firstly, entropy measurements have a solid theoretical foundation for capturing the randomness of a time series. Generally speaking, approximate entropy (AE), which we use herein, is an applicable derivation for short series of Shannon entropy, whose continuous version is defined as

$$H(P) = - \int_{-\infty}^{+\infty} P(x) \log[P(x)] \, dx, \tag{1}$$

where $H(P)$ is the Shannon entropy of the density function, $P(x)$, of the random variable, x, and $\int_{-\infty}^{+\infty} P(x) dx = 1$. In information theory, Shannon entropy represents the degree of the uncertainty of a complex system composed of x outcomes. For a time series formed by x outcomes, the more random the series is, the higher entropy x has. Therefore, in a fully informationally efficient market, stock return series are random, and their entropies reach their highest level. By contrast, in real-world stock markets where a degree of market inefficiency can be observed, information fails to get incorporated into stock prices accurately and immediately because of market imperfections, which produces lower entropy values than the value associated with

the perfectly efficient market. In this spirit, entropy provides a method of measuring time-varying informational efficiency.

Second, entropy measurements perform well in picking up the linear and non-linear features of data series from complex systems. However, measures such as the variance ratio based on autocorrelation, cannot fully capture the non-linear characteristics of informational efficiency. Instead, AE's ability to measure information on the patterns and variations that a system can display makes it an appropriate proxy for informational efficiency.

Finally, entropy measurements can be applied to both the Chinese and the U.S. stock markets. Although measures based on the law of one price can capture non-linear characteristics, they are constrained by limited data sources in China, which has a nascent options market.

Although entropy is a powerful tool for characterizing the diversity of the patterns contained in a time series, its application has been hindered by its strict data requirements. After Gulko (1999) first introduced the concept of entropy to financial studies, Pincus (1991) proposed AE under the assumption that if the joint probability measures that describe two systems differ, their marginal distributions on a fixed partition are likely different. Further, Pincus and Kalman (2004) demonstrate the applicability of AE by considering empirical data and models in financial markets. They argue that as a model-independent measure of sequential irregularity, AE is excellent for dealing with financial time series, even if the sequences are short, as well as for signaling the level of market stability. Since then, entropy has been used to

9

measure market efficiency in various markets such as stock markets (Alvarez-Ramirez et al., 2012) and crude oil markets (Martina et al., 2011).

Our measure of informational efficiency by AE is more direct than the approach of Alvarez-Ramirez et al. (2012), who measure informational efficiency using AE divided by the minimum AE of 5000 random sequences for each timescale. As AE is divided by the same minimum entropy for each time scale, it does not affect the rank of stocks. Specifically, we calculate AE using the algorithm in the Appendix.

3.2. Empirical models

We check whether various asset pricing models, which include factors such as size, book-to-market, momentum, operating profitability, investment, and informational inefficiency, can explain the time-series variations in stock returns. We focus on examining whether the IIF plays an essential role in explaining time-series variations in stock returns.

We adopt the following factor models to conduct ordinary least squares regressions. These are presented in equations (2) to (9), which denote the CAPM, FF3F, IIF3F, WML4F, FF5F, IIF4F, IIF5F, and IIF6F, respectively. These models are used for main two aims. The first is to evaluate the ability of the IIF to price assets. As market and size factors outperform the other factors (see Section 4.1), we construct IIF3F to compare the pricing ability of the IIF and factors except for MP and SMB in FF3F, WML4F, and FF5F. (See the notes to Table 1 for the definitions of the studied factors.) In other words, we aim to compare models (3), (4), (5), and (6). The second

10

aim is to check the robustness of the relationship between the IIF and excess returns. We attain our second aim by adding the IIF into the FF3F, WML4F, and FF5F models, from which equations (7), (8), and (9) are obtained. The CAPM, given in equation (2), is a benchmark model for asset pricing.

$$R_{pt} - R_{ft} = a_p + b_p MP_t + \varepsilon_{pt} \tag{2}$$

$$R_{pt} - R_{ft} = a_p + b_p MP_t + s_p SMB_t + h_p HML_t + \varepsilon_{pt} \tag{3}$$

$$R_{pt} - R_{ft} = a_p + b_p MP_t + s_p SMB_t + \psi_p IIF_t + \varepsilon_{pt} \tag{4}$$

$$R_{pt} - R_{ft} = a_p + b_p MP_t + s_p SMB_t + h_p HML_t + w_p WML_t + \varepsilon_{pt} \tag{5}$$

$$R_{pt} - R_{ft} = a_p + b_p MP_t + s_p SMB_t + h_p HML_t + r_p RMW_t + c_p CMA_t + \varepsilon_{pt} \tag{6}$$

$$R_{pt} - R_{ft} = a_p + b_p MP_t + s_p SMB_t + h_p HML_t + \psi_p IIF_t + \varepsilon_{pt} \tag{7}$$

$$R_{pt} - R_{ft} = a_p + b_p MP_t + s_p SMB_t + h_p HML_t + w_p WML_t + \psi_p IIF_t + \varepsilon_{pt} \tag{8}$$

$$R_{pt} - R_{ft} = a_p + b_p MP_t + s_p SMB_t + h_p HML_t + r_p RMW_t + c_p CMA_t + \psi_p IIF_t + \varepsilon_{pt} \tag{9}$$

where $R_{pt} - R_{ft}$ is the excess return of the portfolios; MP_t is the market excess return[ii]; SMB_t is the size factor; HML_t is the book-to-market factor; WML_t is the momentum factor; RMW_t is the profitability factor; CMA_t is the investment factor; IIF_t stands for the orthogonalized IIF; ε_{pt} denotes the pricing error term; and the factor sensitivities or loadings, b_p, s_p, h_p, w_p, c_p, r_p, and φ_p, are the slope coefficients for MP, SMB, HML, WML, CMA, RMW, and IIF, respectively. We use Newey and West's (1987) standard error estimator for the asymptotically correct standard error estimation under possible heteroscedasticity or autocorrelation of unknown form.

Following Fama and French (2015), we construct 25 portfolios by bivariate or univariate sorts at the end of June each year. We form portfolios in three ways, which are based on (1) informational inefficiency and size (AE-size), (2) informational

inefficiency and the book-to-market ratio (AE-BM), and (3) informational inefficiency only (AE). When forming these 25 AE-size portfolios, we assign each stock to five informational inefficiency portfolios according to the AE rank. Independently, we sort the stocks into five equal-sized portfolios according to the market capitalization rank. We then form the 25 portfolios by taking the intersection between the AE and size groups; in a similar way, we obtain the 25 portfolios using AE and the book-to-market ratio. When forming the 25 informational inefficiency portfolios, we assign all stocks to one of the 25 equal-sized portfolios by AE rank. After constructing these portfolios, we compute the value-weighted monthly returns for each portfolio. We calculate the excess portfolio returns by taking the difference between the daily portfolio returns and risk-free rate. We rebalance the portfolios at the end of June every year from 1999 to 2016.

We construct the Fama–French three factors and momentum factor (MP, SMB, HML, and WML) following Fama and French (1993, 2012). The IIF is constructed as follows. At the end of each June, firms are ranked by size (market capitalization) and assigned to either a small-sized portfolio or a large-sized portfolio. Then, these stocks are independently ranked and assigned to three portfolios by their AE. The high-AE L1 portfolio contains the top 30% of stocks, while the low-AE L3 portfolio contains the bottom 30% of stocks. The middle 40% of stocks are assigned to the L2 portfolio. Then, six portfolios (S/L1, S/L2, S/L3, B/L1, B/L2, and B/L3) are formed at the intersections of the independent sorts on size and AE. The value-weighted daily returns of these six portfolios are calculated each day over the year

after portfolio formation. Repeating this procedure for each year produces 4116 value-weighted daily returns from July 1999 to June 2016 for every portfolio. The IIF is the simple average of the difference in returns between the L3 portfolios (with high portfolio returns) and L1 portfolios (with low portfolio returns):

$$IIF = [(S/L3 - S/L1) + (B/L3 - B/L1)]/2 \tag{10}$$

To show how informational inefficiency is priced, we orthogonalize the IIF using the FF3F, WML4F, and FF5F models. For instance, we perform the orthogonalization using the FF5F model as follows:

$$IIF_t = a_e + b_e MP_t + s_e SMB_t + h_e HML_t + r_e RMW_t + c_e CMA + \varepsilon_{et} \tag{11}$$

The orthogonalized IIF is constructed as the sum of the intercept a_e and residual ε_{et}. When orthogonalized using WML4F and FF5F, this procedure is repeated with the corresponding models. Unless otherwise specified, the IIFs hereafter (except those in Section 4.1) are the orthogonalized IIFs.

3.3. Data description and preliminary analysis

We obtain the U.S. stock trading data and annual accounting data from the Centre for Research in Security Prices and Compustat databases. All data for the Chinese stock market are from the China Stock Market and Accounting Research database. We calculate the daily risk-free rates using the one-month U.S. Treasury bill rate and People's Bank of China's one-year deposit rate, respectively.

In line with previous studies (Fama and Macbeth, 1973; Fama and French, 1992), we adopt the following four data selection criteria. First, we select the daily return data of non-financial companies with appropriate adjustments for capital changes such as splits and stock dividends. Second, we exclude firms with negative book-to-market ratios. Third, we delete stocks that have more than 50 trading days of missing returns in rolling windows of 500 consecutive trading days to rule out thin-trading stocks, which are likely to have different return characteristics to other types of stocks[iii]. Fourth, following Wang and Xu (2004), we exclude initial public offering (IPO) returns in the first month of individual stocks from the Chinese sample because first-month IPO stock returns are abnormally high in the Chinese A-shares market, with most being more than 50%. Hence, excluding IPO returns in the first month of individual stocks can help rule out extreme returns, which could severely bias our test results. Finally, our sample consists of firms with daily individual stock returns with dividend reinvestment, market capitalization, book-to-market ratios, operating profitability, growth in total assets, and monthly trading volumes. Our final sample includes an average of 816 listed firms per day from the Chinese stock market and 802 stocks per day from the U.S. stock market from July 1999 to June 2016.

Table 1 reports the summary statistics and pairwise correlation coefficients among the explanatory variables in the U.S. and Chinese stock markets. In Panel A, two points are worth mentioning. First, the corresponding factors in both countries have a different magnitude, which indicates dissimilar market conditions in the two economies. For instance, U.S. SMB has a mean of 13 base points, while the Chinese

14

SMB is much smaller. Second, although some factors have a mean of zero, all factors have notable deviations from their means. To give an illustration, the IIF has notable 5% and 95% percentiles like the other factors, which suggests time-varying premiums for informational inefficiency in both markets. In line with Brennan and Wang's (2010) argument, we show in Section 4.3 that informational inefficiency can induce a premium even with an average of zero.

[Insert Table 1 around here]

Panel B of Table 1 reports the Pearson pairwise correlations among our variables. Except for the slightly high correlation coefficients between HML and CMA in the U.S. market, most of the other coefficients are below 0.30, which suggests that multicollinearity is not severe in our study.

As mentioned before, the performance of the IIF varies over time, as the AMH implies. To provide a further illustration, we investigate its dynamic feature in moving subsamples of different lengths separately.

[Insert Figure 1 around here]

Figure 1 shows the P-values of one-sample t-tests on the U.S. and Chinese IIFs in the moving subsamples with a length of 1000, 300, 100, and 50 days, respectively. The horizontal axis denotes the dates when the subsamples begin. The vertical axis denotes the P-values, which represent the probability that the IIF has a mean of zero in a subsample under the zero-mean hypothesis. The horizontal line with a P-value of 0.1 shows the 10% significance level, below which the subsamples have significant IIFs. As we can see, when the length of the subsamples is set to 1000 days,

the IIF varies over time, but there is no sign of significance in either the U.S. or the Chinese stock markets. However, when the subsamples are shortened to 300, 100, and 50 days, more and more subsamples with significant IIFs appear. This finding implies that the process that information related to intrinsic values is absorbed into asset prices takes time. Consequently, when processing information, the shorter a rolling window is, the more likely a market is to see a significant IIF. Moreover, compared with the United States, the Chinese stock market witnesses subsamples with more significant IIFs, which suggests it is more informationally inefficient. These results are further corroborated by the orthogonalized IIF results in Section 4.

[Insert Figure 2 around here]

Likewise, the performance of the other factors varies as the subsamples become shorter, which is consistent with the AMH. Figure 2, similar to Figure 1, shows the performance of SMB in the U.S. and Chinese stock markets. Different from the IIF, the subsamples with a fixed length of 1000 days see impressive long-short returns on size in the U.S. market. Instead, SMB performs unstably as the IIF in the 50-day subsamples. Moreover, other factors including MP, HML, WML, RMW, and CMA behave in a similar way in the short run regardless of their behavior in the long run. To save space, the plots of these factors are not shown, but they are available on request. The dynamic performance of such factors may result from stock prices failing to incorporate information in the short term. We show in Section 4.3 that the process of integrating information is achieved through a time-consuming negative feedback mechanism rather than immediately.

16

4. Empirical results

We present the results from various standard tools that permit us to gauge from different angles the answer to the central question in our study. Section 4.1 discusses the necessity of the IIF. Section 4.2 considers the performance of the eight factor models and Section 4.3 examines the role that the IIF plays in asset pricing. Sections 4.4 and 4.5 present the results on model evaluation and further robustness tests, respectively.

4.1. IIF

When performing time-series regression, the tests on whether a new factor can be added into an existing factor model are identical to those on whether the alpha declines with the new factor introduced. When factors are correlated with each other, adding a new factor might influence the coefficients of existing factors without lowering the alpha. A practical way to overcome this problem is to orthogonalize the new factor on the existing factors and form an orthogonalized new factor, which we use in this study to test whether it can be added into the existing factor model.

According to the AMH, the additional pricing ability provided by the IIF relative to existing models may vary over time as market conditions change. On the one hand, during a certain period, the pricing information in the IIF that results from market imperfections could be covered by existing factor models, which indicates that the IIF helps little in improving pricing accuracy. In such a case, the newly orthogonalized factor is insignificantly different from zero, which indicates that the

17

market has relatively high informational efficiency, as informational inefficiency has little extra effect on returns. On the other hand, if the IIF cannot be explained by existing factor models, the newly orthogonalized factor is significantly different from zero and can be added into the existing factor model. Accordingly, the market is less informationally efficient. To avoid sampling bias, we use subsamples of fixed lengths varying from 40 to 1500 days.

[Insert Figure 3 around here]

Figure 3 reports the percentages of the subsamples of fixed lengths in which the orthogonalized IIF using FF5F is significantly different from zero. The horizontal axis shows the lengths of the moving subsamples. To save space, similar results when using FF3F and WML4F to the orthogonalized IIF are not shown but are available upon request. As the figure shows, the orthogonalized IIFs perform better in shorter moving windows, even better than SMB in terms of 40 days in both the U.S. and the Chinese stock markets. This finding indicates that the IIF carries extra information beyond the existing factor models, especially in the short run when new information on returns cannot be fully processed. Therefore, it is necessary to add the IIF into the factor models, as without it, the factor models do not perform well in shorter sample periods.

4.2. Factor model performance

We use the aforementioned eight factor models and report the numbers of significant factor coefficients of these factor models in Table 2. Panels A, B, and C report the

results for portfolios constructed in different ways. We report the numbers of significant coefficients at the 5% level for each model. The results for the 10% significance level are available on request.

[Insert Table 2 around here]

The significance of the coefficients of a factor signals whether an asset has risk exposure to this factor. As the table shows, several points are worthy of mention. First, interestingly, the magnitude of the significant coefficients for MP, SMB, HML, and IIF is large, which underlies their economic significance. Moreover, MP and SMB are the most significant and robust across all three panels in both markets, followed by HML and IIF. This result shows that most of the portfolios have risk exposure on the market, size, and BM factors as well as the IIF. Second, the number of significant IIF coefficients is higher than that of the WML, CMA, and RMW coefficients, which also suggests that the IIF is an important risk factor in asset pricing models. Furthermore, compared with the U.S. market, the Chinese market has more risk exposure to the IIF, which signals that it might be more informationally inefficient.

One measure of the ability of an asset pricing model is the number of assets it can price. Therefore, the number of significant intercepts is an important indicator. The fewer significant intercepts there are, the better is the pricing model. On the one hand, the U.S. and Chinese markets share two common features in terms of model performance. First, the CAPM has the most substantial number of significant intercepts at the 5% significance level when dealing with portfolios constructed by

19

different ways of sorting. Second, on the basis of IIF3F, there is little improvement as the number of factors increases. On the other hand, in the Chinese stock market, IIF3F has the fewest significant intercepts at the 5% significance level in Panels A and B of Table 2 and has similar performance to FF5F and IIF6F in Panel C, which means that IIF3F can price these portfolios better than the other models. On the contrary, the IIF models in the U.S. market improve only a little by introducing the IIF into the traditional model. The more important role the Chinese IIF plays in asset pricing indicates that the Chinese stock market is less informationally efficient. The common belief that the U.S. stock market is more informationally efficient is supported by not only the more pricing information the IIF carries in the Chinese market, but also the stronger mechanism in the U.S. market, as we show in Section 4.5.

4.3. Role of the IIF in asset pricing

Table 3 reports the coefficients of the IIF in the time-series regressions with factor models including IIF3F and IIF6F in the U.S. and Chinese stock markets. Panel A shows the coefficients of the IIF from the regressions using the different pricing models on the portfolios constructed by AE and size. The coefficients of the risk factors should be non-negative because investors require more compensation to bear more risk. Further, the coefficients of the IIF across the AE-size portfolios for all four models show three distinct patterns. The first pattern shows a number of negative coefficients that cannot be ignored; the positive coefficients are mainly distributed on the left-hand side of the table, whereas the negative coefficients are distributed on the

20

right-hand side. The second pattern is that the most statistically significant coefficients are mainly concentrated on the extremes of the AE groups and the coefficients in some moderate AE groups are both positive and negative. Moreover, the number of portfolios that have significantly positive exposure to the IIF in the U.S. stock market from the different models is less than that in the Chinese market (six portfolios in the U.S. market and 15 in the Chinese market on average).

Overall, as AE increases, the coefficients of the IIF statistically change from significantly positive, through insignificant, to significantly negative. As expected, not all the coefficient rows follow this pattern perfectly because of the limited range of informational inefficiency. For instance, in the second largest groups in the U.S. stock market, the coefficients change from insignificantly negative to significantly negative as AE increases.

[Insert Table 3 around here]

Panel B of Table 3 reports the coefficients of the IIF from the regressions with the different pricing models on the portfolios constructed by AE and BM. The number of negative coefficients changes slightly. In every row, the transition of coefficients from significantly positive to significantly negative is similar to that in Panel A. In addition, the Chinese market still has more significantly positive coefficients than the U.S. market.

Panel C shows the coefficients of the IIF from the regressions with the different pricing models on the portfolios constructed by AE only. Compared with the results for the portfolios formed by two variables, Panel C seems to show mixed

21

evidence on the patterns in Panels A and B. For the Chinese market, the patterns are rather straightforward. Except for some data surprises, they are similar, with significantly positive coefficients in the top part, insignificant ones in the middle part, and significantly negative ones in the bottom part. Instead, for the U.S. market, it seems far-fetched that the pattern exists. In view of the patterns in the former panels and complexity of informational inefficiency resulting from various market imperfections, we further study the correlation between AE, size, and BM to dissect the unexpected results in Panel C. We find that AE in the Chinese market correlates with size and BM with coefficients of 0.30 and 0.48, respectively. By contrast, AE in the U.S. market has high correlation coefficients of 0.85 and 0.60, respectively. Hence, the results in Panel C are consistent with those in Panels A and B.

Informational inefficiency may induce a negative premium, as it is closely related to behavioral bias, which might be negatively priced. By dividing the mispricing premium into four elements, Brennan and Wang (2010) show that the premiums related to under-reaction and overpricing are negative. Similarly, Hong and Sraer (2016) show that when aggregate investor disagreement is high, expected returns can be negatively related to the beta.

The fact that the coefficients of the IIF statistically change from significantly positive to significantly negative as informational inefficiency declines shows that trading on informational inefficiency is profitable (unprofitable) when it is high (low). This finding concurs with the AMH. Under the AMH framework, the relationship between risk and returns is unstable because of changing market conditions (e.g., the

policy environment and preferences of market participants). For participants that have bounded rationality, the turning point from a positive relationship to a negative one is unknown. Investors tend to trade based on their experience before they realize there has been a reversal in the IIF–return relationship. Moreover, this may account for the malfunctioning of their investment strategies that once functioned well.

Similar to Grossman (1976) and Grossman and Stiglitz (1980), this study considers investors to be driven by profit on informational advantages. When informational inefficiency is high, the returns of trading on informational advantages cover the cost of doing so, and the more informational advantages investors exploit, the more they will earn. In this case, investors keep arbitraging for informational inefficiency and thus the market becomes more informationally efficient. Hence, there is a positive relationship between returns and risk on informational inefficiency under high informational inefficiency. However, when the levels are low, the cost of trading on informational advantages is more than its return, and the more informational advantages investors exploit, the more they lose. Although investors that have bounded rationality do not know the exact turning point of the sign of the IIF–return relationship, investors stop (begin) arbitraging as soon as they realize the turnaround, which results in higher (lower) levels of informational inefficiency relative to the level at the turning point. In this negative–positive circle originating from market imperfections and driven by time-varying market conditions, loss and gain alternate when trading on informational inefficiency. We call this cycle an inverse feedback

mechanism, which raises informational efficiency when it is too low and lowers it when it is too high.

4.4. Factor model evaluation

As Section 4.2 shows, relative to traditional models, models including the IIF can price these portfolios better in the Chinese market, whereas they have average performance in the U.S. market. In this section, we further evaluate the performance of these models following the literature (Fama and French, 2015).

[Insert Table 4 around here]

If an asset pricing model covers all the risk factors that determine stock returns, the intercept should be zero in regressions of an asset's excess returns on the model's factor returns. Following Fama and French (2015), we use the average absolute value of the intercepts (AAVI) in the regressions and the GRS F-test (Gibbons et al., 1989) to assess the performance of the models. The GRS statistic has been formulated under the assumption of normal error terms that are homoscedastic and uncorrelated over time, and it is subject to a chi-square distribution under the null hypothesis that the intercepts from the regressions on a perfect model should be jointly indistinguishable from zero. We examine eight of the factor models (CAPM, FF3F, IIF3F, WML4F, IIF4F, FF5F, IIF5F, and IIF6F) and evaluate their overall explanatory power. Panels A, B, and C in Table 4 show the results for these eight models regressed on the portfolios constructed by AE and size, AE and BM, and AE only, respectively.

A better model would have a lower AAVI or a higher insignificant result (or lower GRS-Stat) in the GRS F-test. Overall, the models including the IIF show an equal or better performance than traditional models. For instance, in Panel A, AAVIs decrease significantly by introducing the IIF (3.83–3.76 from FF3F to IIF4F, 5.76–5.43 from WML4F to IIF5F in the U.S. market) and IIF3F outperforms the other models in terms of model concision and precision. In Panel C, adding the IIF also helps improve the pricing ability of traditional models. The U.S. stock market in Panel B appears to be an exception, probably because it is more informationally efficient. In a highly efficient market, the IIF plays a less vital role in asset pricing, which results in it adding more noise relative to the contributions it makes.

4.5. Additional robustness checks

Seeking to add additional robustness to our main findings, we conduct subsample analyses of the IIF models using a fixed moving window varying from 40 to 1500 days. In Section 4.3, we found that the inverse feedback mechanism is constituted by three parts of the IIF–return relationship, namely being significantly positive, insignificant, and significantly negative. We show that parts of or all the mechanisms in Section 4.3 disperse in the different subsamples and nothing goes beyond the pattern. To save space, we report the case regressed on IIF6F using a fixed moving window of 1000 days to illustrate the dynamic feature of the inverse feedback mechanism. The results using the other window lengths or from the regressions on the other models are not shown but are available upon request.

[Insert Figure 4 around here]

Figure 4 shows the patterns of the coefficients of the IIF from the time-series regressions on the portfolios formed by AE and size using IIF6F in the moving subsamples of a fixed length of 1000 days. As the legend ticks show, the seven colors represent the seven kinds of coefficients (i.e., significantly positive or negative at the three significance levels of 1%, 5%, and 10% and insignificant coefficients at the 10% significance level) in the negative feedback mechanism. The blue (red) bars denote positive (negative) coefficients. Therefore, the color bar from its blue end to red end can be regarded as the complete changing process of the coefficients of the IIF as informational efficiency improves. The vertical axis shows the portfolios. For example, S1E1 is the portfolio with the smallest size and lowest AE in a subsample. The horizontal axis denotes the dates when the subsamples start.

Figure 4 highlights several of the attractive features of the inverse feedback mechanism. First, for all five AE portfolios within the same size portfolios, the transitions of the IIF coefficients as informational efficiency increases are subject to the negative feedback mechanism in both the U.S. and the Chinese stock markets. In the Chinese market, most size groups across all the moving subsamples are dominated by the entire mechanism that moves from dark blue, through light yellow, and ends at deep red from the E1 groups to the E5 groups. By contrast, a minority of the U.S. size groups (e.g., the smallest size groups in the subsamples that start after June 10, 2009) hold the complete mechanism. A majority of the U.S. subsamples are overshadowed by the light yellow, with the red area the second largest. The more non-positive IIF–

26

return relations indicate that the U.S. stock market is more informationally efficient than the Chinese market, which supports the widely held belief that emerging markets are less efficient than developed markets.

Second, megacap portfolios tend to be more informationally efficient than microcaps. In support of this, the blue area in every size group shrinks from S1 to S5 in the U.S. and Chinese markets. The slower process of microcaps incorporating information into prices may result from their illiquidity (Amihud and Mendelson, 1986; Lin et al., 2018) or behavioral biases (Shleifer and Vishny, 1997; Carpentier et al., 2018).

Third, the positions of the shifts between the positive and negative IIF–return relations vary over time. The light yellow area in every five size groups consists of turning points that signal changes in the signs of the IIF–return relations. Over time, the insignificant territories broaden or narrow inside the space enclosed by the red and blue bars as in the Chinese market or occupy most of the 25 portfolios between June 25, 2003 and June 10, 2009 as in the U.S. market. This time-varying characteristic of the mechanism mirrors the dynamic market conditions in an evolving market, which is consistent with the literature (e.g., Lo, 2004, 2005; Urquhart and Hudson, 2013; Urquhart and McGroarty, 2014).

As the mechanism implies, the U.S. and Chinese stock markets are driven by the trade-off between profit that induces arbitraging and loss that inhibits such an activity. This finding suggests that significantly positive or negative IIF–return relations are unstable as investors keep changing their decisions. We argue that in a

relatively stable market, the IIF–return relation finally converges to be insignificant as investors learn from their trading. Further, the more efficient a market is, the less time convergence takes. Indeed, the mechanism's ability to converge is central to market efficiency in an imperfect market, which complements the traditional view that higher informational efficiency stands for a more efficient market. In this sense, we suggest that the U.S. stock market is more informationally efficient than the Chinese market.

When we change the length of the subsamples, regression model, or ways of sorting portfolios, the new results are consistent with our earlier findings. The robustness of our findings thus provides strong support for the AMH and shows how informational efficiency in the Chinese stock market evolves.

5. Concluding remarks

Inspired by the dynamic view of market efficiency incorporated into the AMH, we study how informational inefficiency is priced in portfolio returns and complement recent studies at the market level using U.S. and Chinese stock data from July 1999 to June 2016. We draw the following four main conclusions.

First, we find that as informational efficiency increases, the relations between returns and the IIF statistically change from significantly positive, through insignificant, to significantly negative, which indicates an inverse feedback mechanism in the U.S. and Chinese stock markets[iv]. This mechanism appears to be a result of investors' trade-off between profit and loss when arbitraging, which makes the market "adaptively efficient." That is, whenever informational inefficiency is too

high or low, the invisible hand pulls it back. Therefore, the EMH can be considered to be a special case of the AMH. This finding concurs with that of Grossman and Stiglitz (1980) and the implications of the AMH that market efficiency (arbitrage opportunity or the performance of investment strategies) varies over time.

Second, we observe that the U.S. stock market is more informationally efficient, suggesting that both the pricing information in the IIF and the power of the mechanism reflect the level of market efficiency in an imperfect market setting. In this spirit, this finding is consistent with the widely held belief that emerging markets are less efficient than developed markets.

Third, we construct factor models, which are more suitable for less efficient stock markets, by introducing the IIF to capture the shifts between the positive and negative IIF–return relations and better price portfolios in a less efficient market.

Finally, we check the robustness of our findings using various subsample analyses, altering the ways of sorting the portfolios, and regressing on different factor models. Our findings strongly support the AMH and suggest a potential opportunity to build a trading strategy by equally investing in the stock quantiles sorted on the IIF. We leave the profitability of such a strategy to future research.

References

Alvarez-Ramirez, J., E. Rodriguez, and J. Alvarez. 2012. "A Multiscale Entropy Approach for Market Efficiency." *International Review of Financial Analysis* 21: 64–69.

Amihud, Y., and H. Mendelson. 1986. "Asset Pricing and the Bid-ask Spread." *Journal of Financial Economics* 17(2): 223–249.

Bai, J., T. Philippon, and A. Savov. 2016. "Have Financial Markets Become More Informative?" *Journal of Financial Economics* 122(3): 625–654.

Bailey, W., A. Kumar, and D. Ng. 2011. "Behavioural Biases of Mutual Fund Investors." *Journal of Financial Economics* 102(1): 1–27.

Bell, D. E. 1983. "Risk Premiums for Decision Regret." *Management Science* 29(10): 1156–1166.

Bodnaruk, A., and A. Simonov. 2016. "Loss-averse Preferences, Performance, and Career Success of Institutional Investors." *Review of Financial Studies* 29(11): 3140–3176.

Boehmer, E., and J. Wu. 2007. "Order Flow and Prices." *Working Paper*, University of Georgia.

Bondt, W. F., and R. Thaler. 1985. "Does the Stock Market Overreact?" *Journal of Finance* 40(3): 793–805.

Brennan, M. J., and A. W. Wang. 2010. "The Mispricing Return Premium." *Review of Financial Studies* 23(9): 3437–3468.

Carpentier, C., Romon, F., and Suret, J. M. 2018. "Are Investors Rational When Valuing Loss Firms?" *Journal of Behavioural Finance* 19(2): 177–189.

Chan, W. S. 2003. "Stock Price Reaction to News and No-news: Drift and Reversal after Headlines." *Journal of Financial Economics* 70(2): 223–260.

Charles, A., O. Darné, and J. H. Kim. 2012. "Exchange-rate Return Predictability and the Adaptive Markets Hypothesis: Evidence from Major Foreign Exchange Rates." *Journal of International Money and Finance* 31(6): 1607–1626.

Choi, N., and R. W. Sias. 2009. "Institutional Industry Herding." *Journal of Financial Economics* 94(3): 469–491.

Cremers, M., and D. Weinbaum. 2010. "Deviations from Put-call Parity and Stock Return Predictability." *Journal of Financial and Quantitative Analysis* 45: 335–67.

Fama, E. F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25(2): 383–417.

Fama, E. F. 1991. "Efficient Capital Markets: II." *Journal of Finance* 46(5): 1575–1617.

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

Fama, E. F., and K. R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33(1): 3–56.

Fama, E. F., and K. R. French. 2012. "Size, Value, and Momentum in International Stock Returns." *Journal of Financial Economics* 105(3): 457–472.

Fama, E. F., and K. R. French. 2015. "A Five-factor Asset Pricing Model." *Journal of Financial Economics* 116(1): 1–22.

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

Gervais, S., J. B. Heaton, and T. Odean. 2011. "Overconfidence, Compensation Contracts, and Capital Budgeting." *Journal of Finance* 66(5): 1735–1777.

Gervais, S., and T. Odean. 2001. "Learning to be Overconfident." *Review of Financial Studies* 14(1): 1–27.

Gibbons, M. R., S. A. Ross, and J. Shanken. 1989. "A Test of the Efficiency of a Given Portfolio." *Econometrica* 57(5): 1121–1152.

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

Grossman, S. 1976. "On the Efficiency of Competitive Stock Markets where Trades have Diverse Information." *Journal of Finance* 31(2): 573–585.

Grossman, S. J., and J. E. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review* 70(3): 393–408.

Gulko, L. 1999. "The Entropic Market Hypothesis." *International Journal of Theoretical and Applied Finance* 2(03): 293–329.

Hasbrouck, J. 1993. "Assessing the Quality of a Security Market: A New Approach to Transaction-cost Measurement." *Review of Financial Studies* 6: 191–212.

Hong, H., and D. A. Sraer. 2016. "Speculative Betas." *Journal of Finance* 71(5): 2095–2144.

Huberman, G., and T. Regev. 2001. "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar." *Journal of Finance* 56(1): 387–396.

Ito, M., and S. Sugiyama. 2009. "Measuring the Degree of Time-varying Market Inefficiency." *Economics Letters* 103(1): 62–64.

Jacobs, H. 2016. "Market Maturity and Mispricing." *Journal of Financial Economics* 122(2): 270–287.

Jiang, G. J., and K. X. Zhu. 2017. "Information Shocks and Short-term Market Underreaction." *Journal of Financial Economics* 124(1): 43–64.

Jones, C. S., and L. Pomorski. 2017. "Investing in Disappearing Anomalies." *Review of Finance* 21(1): 237–267.

Kahneman, D., and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2): 263–291.

Kim, J. H., A. Shamsuddin, and K. P. Lim. 2011. "Stock Return Predictability and the Adaptive Markets Hypothesis: Evidence from Century-long U.S. Data." *Journal of Empirical Finance* 18(5): 868–879.

Levich, R. M., and V. Potì. 2015. "Predictability and "Good Deals" in Currency Markets." *International Journal of Forecasting* 31(2): 454–472.

Lin, Z. Y., C. C. Chang, and Y. H. Wang. 2018. "The Impacts of Asymmetric Information and Short Sales on the Illiquidity Risk Premium in the Stock Option Market." *Journal of Banking & Finance* 94: 152–165.

Lin, M., and S. Viswanathan. 2015. "Home Bias in Online Investments: An Empirical Study of an Online Crowdfunding Market." *Management Science* 62(5): 1393–1414.

Lo, A. W. 2004. "The Adaptive Markets Hypothesis." *Journal of Portfolio Management* 30(5): 15–29.

Lo, A. W. 2005. "Reconciling Efficient Markets with Behavioral Finance: The Adaptive Markets Hypothesis." *Journal of Investment Consulting* 7(2): 21–44.

Lo, A. W., and C. MacKinlay. 1989. "The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation". *Journal of Econometrics* 40: 203–38.

Martina, E., E. Rodriguez, R. Escarela-Perez, and J. Alvarez-Ramirez. 2011. "Multiscale Entropy Analysis of Crude Oil Price Dynamics." *Energy Economics* 33(5): 936–947.

Neely, C. J., P. A. Weller, and J. M. Ulrich. 2009. "The Adaptive Markets Hypothesis: Evidence from the Foreign Exchange Market." *Journal of Financial and Quantitative Analysis* 44(2): 467–488.

Newey, W. K., and K. D. West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55(3): 703–708.

Odean, T. 1998. "Are Investors Reluctant to Realize their Losses?" *Journal of Finance* 53(5): 1775–1798.

Pincus, S. 1991. "Approximate Entropy as a Measure of System Complexity." P*roceedings of the National Academy of Sciences (PNAS)* 88(6): 2297–2301.

Pincus, S., and R. E. Kalman. 2004. "Irregularity, Volatility, Risk, and Financial Market Time Series." *Proceedings of the National Academy of Sciences (PNAS)* 101(38): 13709–13714.

Radzevick, J. R., and D. A. Moore. 2011. "Competing to Be Certain (but Wrong): Market Dynamics and Excessive Confidence in Judgment." *Management Science* 57(1): 93–106.

Rösch, D. M., A. Subrahmanyam, and M. A. van Dijk. 2017. "The Dynamics of Market Efficiency." *Review of Financial Studies* 30(4): 1151–1187.

Shefrin, H., and M. Statman. 1985. "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence." *Journal of Finance* 40(3): 777–790.

Shleifer, A., and R. W. Vishny. 1997. "The Limits of Arbitrage." *Journal of Finance* 52(1): 35–55.

Simon, H. A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69(1): 99–118.

Stambaugh, R. F., and Y. Yuan. 2017. "Mispricing Factors." *Review of Financial Studies* 30(4): 1270–1315.

Urquhart, A., and R. Hudson. 2013. "Efficient or Adaptive Markets? Evidence from Major Stock Markets Using Very Long Run Historic Data." *International Review of Financial Analysis* 28: 130–142.

Urquhart, A., and F. McGroarty. 2014. "Calendar Effects, Market Conditions and the Adaptive Market Hypothesis: Evidence from Long-run U.S. Data." *International Review of Financial Analysis* 35: 154–166.

Urquhart, A., and F. McGroarty. 2016. "Are Stock Markets Really Efficient? Evidence of the Adaptive Market Hypothesis." *International Review of Financial Analysis* 47: 39–49.

Wang, F., and Y. Xu. 2004. "What Determines the Chinese Stock Returns?" *Financial Analysts Journal* 60(6): 65–77.

Figure 1: P-values of the one-sample t-tests on the U.S. and Chinese IIFs
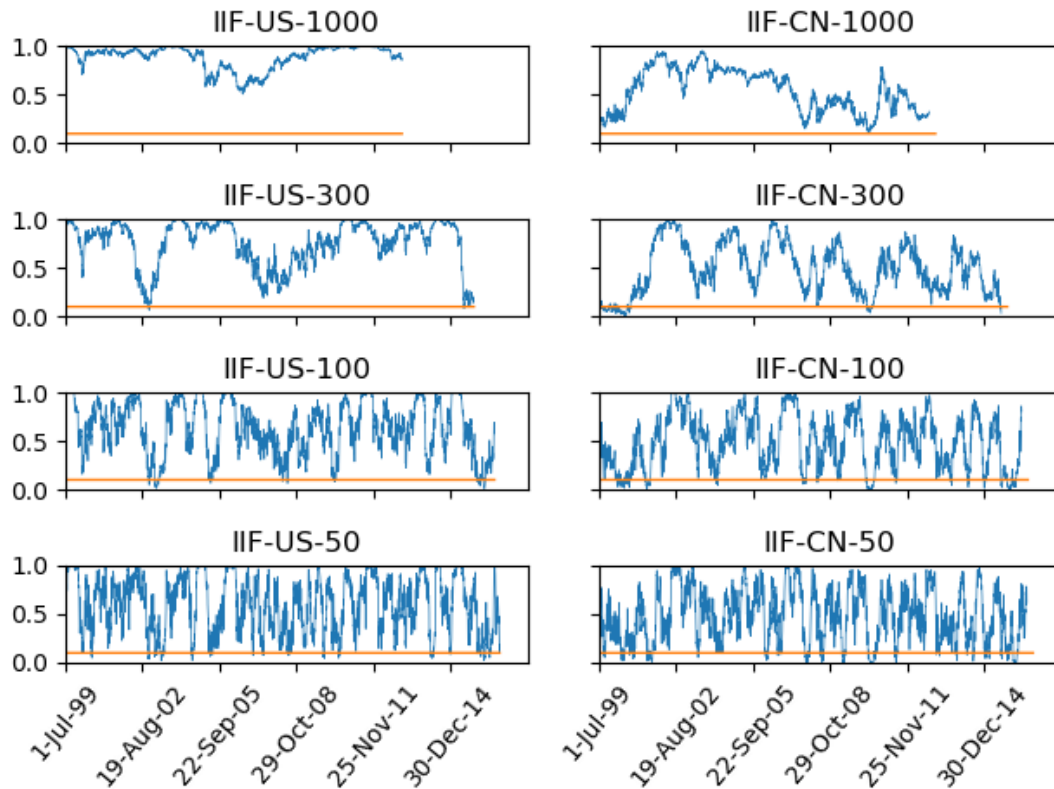
Note: This figure shows the P-values of the one-sample t-tests on the U.S. and Chinese IIFs in the moving subsamples with a length of 1000, 300, 100, and 50 days. The horizontal axis denotes the dates when the subsamples begin. The vertical axis denotes the P-values that are the probability that the IIF has a mean of zero in a subsample under the zero-mean hypothesis. The horizontal line with a P-value of 0.1 shows the 10% significance level, below which the subsamples have significant IIFs.

Figure 2: P-values of the one-sample t-tests on the U.S. and Chinese SMB
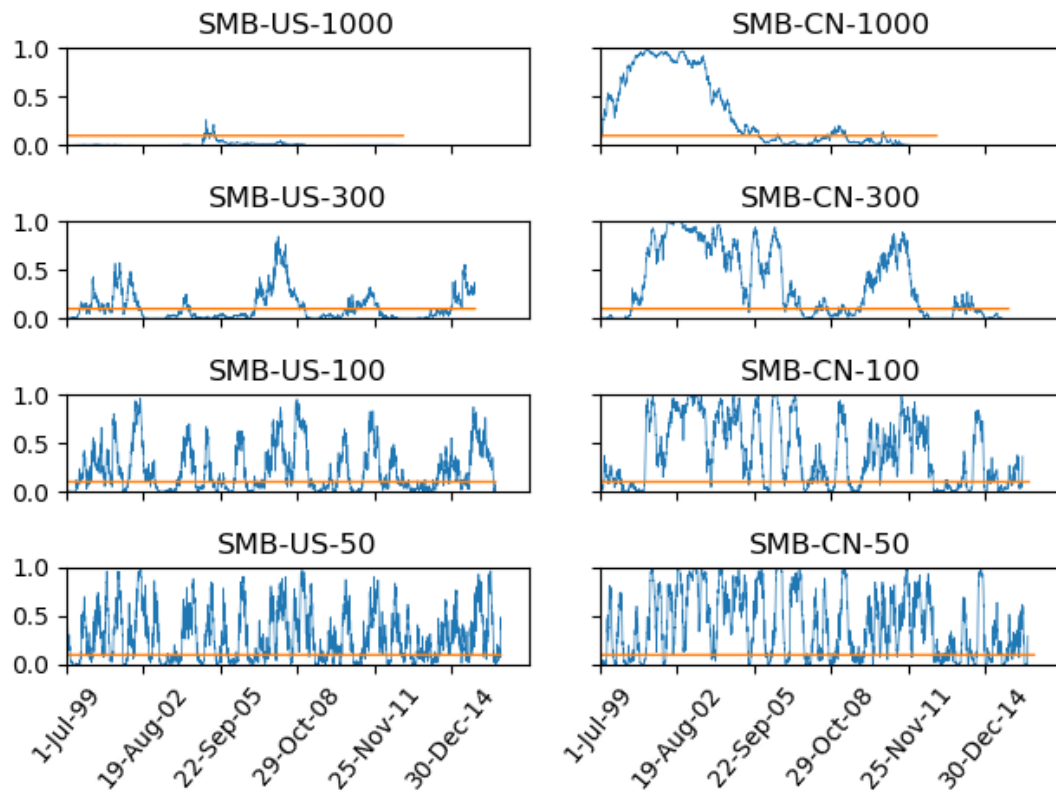
Note: This figure shows the P-values of the one-sample t-tests on the U.S. and Chinese SMB in the moving subsamples with a length of 1000, 300, 100, and 50 days. The horizontal axis denotes the dates when the subsamples begin. The vertical axis denotes the P-values that are the probability that SMB has a mean of zero in a subsample under the zero-mean hypothesis. The horizontal line with a P-value of 0.1 shows the 10% significance level, below which the subsamples have significant SMB.

Figure 3: Percentages of the subsamples with significant factors
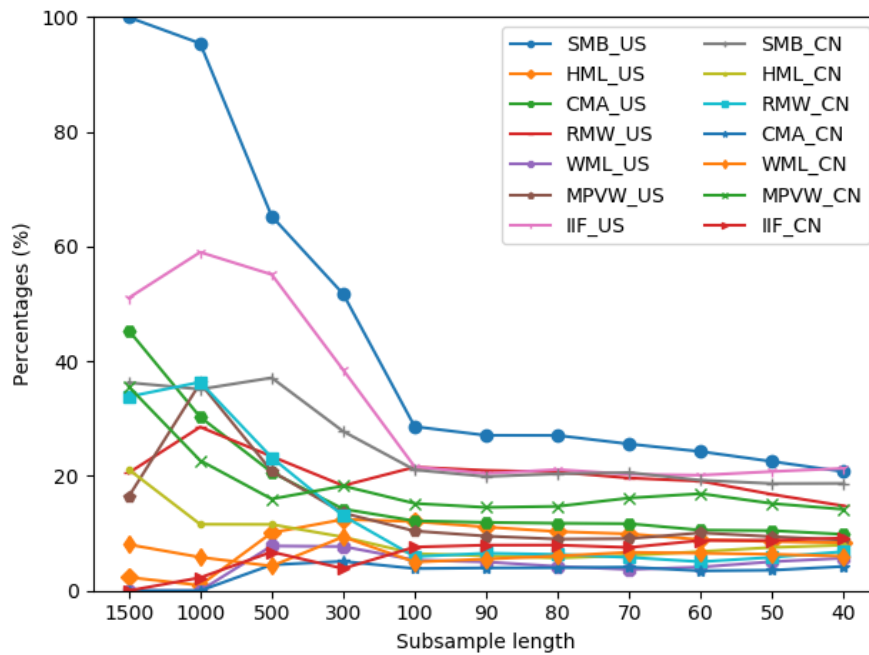
Note: This figure reports the percentages of the subsamples of a fixed length of 1500, 1000, 500, 300, 100, 90, 80, 70, 60, 50, and 40 days in which the orthogonalized IIF using FF5F is significantly different from zero. To save space, similar results when using FF3F and WML4F to orthogonalize the IIF are not shown but are available upon request.

Figure 4: Negative feedback mechanisms over time
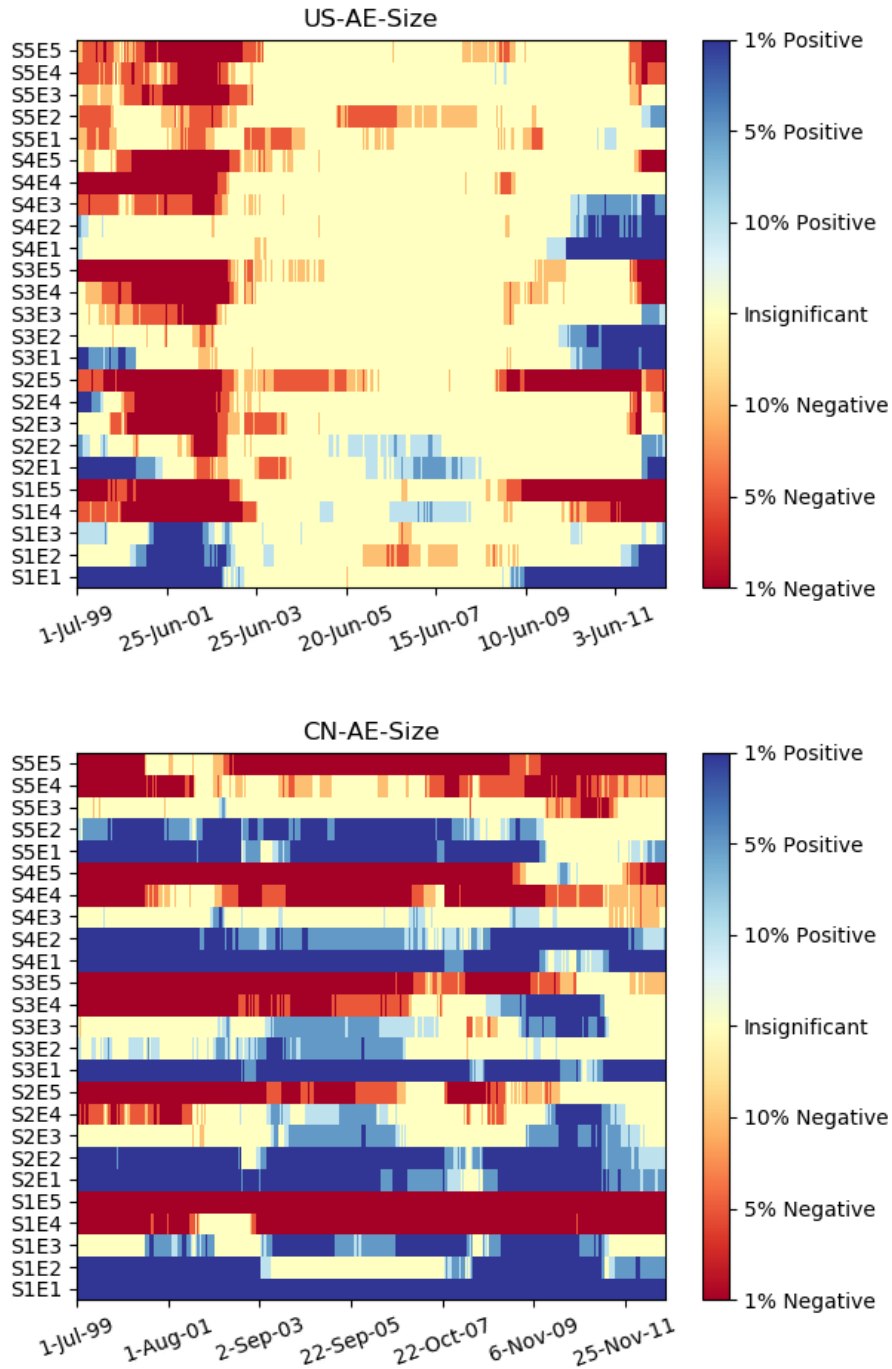
Note: This figure shows the patterns of the coefficients of the IIF from the time-series regressions on the portfolios formed by AE and size using factor model IIF6F in the moving subsamples of a fixed length of 1000 days. The seven colors in the legend represent the seven kinds of coefficients (significantly positive or negative at the three significance levels of 1%, 5%, and 10% and insignificant coefficients at the 10% significance level) in the negative feedback mechanism. The blue (red) bars denote positive (negative) coefficients. The vertical axis shows the portfolios. For example, S1E1 is the portfolio with the smallest size and lowest AE in a subsample. The horizontal axis denotes the dates when the subsamples start.

Table 1: Descriptive statistics of the explanatory variables

Panel A: Summary statistics

Variable	Mean	Standard deviation	Skewness	5% percentile	25% percentile	50% percentile	75% percentile	95% percentile
US Market								
MP	0.0002	0.0127	-0.0432	-0.0196	-0.0055	0.0006	0.0062	0.0183
SMB	0.0013	0.0155	0.0880	-0.0235	-0.0066	0.0017	0.0095	0.0243
HML	0.0000	0.0073	0.3206	-0.0112	-0.0036	0.0000	0.0036	0.0114
WML	0.0000	0.0100	-0.0177	-0.0138	-0.0046	-0.0003	0.0043	0.0148
IIF	-0.0003	0.0059	-0.2753	-0.0095	-0.0032	-0.0002	0.0027	0.0086
RMW	0.0005	0.0121	0.2708	-0.0180	-0.0059	0.0005	0.0069	0.0191
CMA	0.0002	0.0075	-0.5072	-0.0101	-0.0033	0.0000	0.0035	0.0117
CN Market								
MP	0.0005	0.0178	-0.3132	-0.0278	-0.0076	0.0010	0.0022	0.0267
SMB	0.0004	0.0067	-0.6232	-0.0105	-0.0028	0.0007	0.0013	0.0102
HML	0.0000	0.0062	0.2443	-0.0098	-0.0035	-0.0001	0.0005	0.0101
WML	0.0002	0.0050	3.9984	-0.0066	-0.0022	0.0000	0.0003	0.0068
IIF	0.0000	0.0045	-0.2761	-0.0073	-0.0021	0.0001	0.0005	0.0069
RMW	0.0001	0.0049	0.5165	-0.0074	-0.0027	-0.0001	0.0005	0.0083
CMA	-0.0001	0.0034	-0.0485	-0.0057	-0.0020	-0.0001	0.0003	0.0053

Panel B: Correlations between the variables

	MP	SMB	HML	WML	IIF	RMW	CMA
US Market							
MP	1.000						
SMB	0.177	1.000					
HML	0.058	0.093	1.000				
WML	-0.171	-0.195	-0.376	1.000			
IIF	0.259	0.232	-0.039	-0.099	1.000		
RMW	0.201	0.316	0.400	-0.300	-0.058	1.000	
CMA	-0.180	-0.223	-0.461	0.361	-0.126	-0.209	1.000
CN Market							
MP	1.000						
SMB	0.140	1.000					
HML	0.199	-0.213	1.000				
WML	-0.033	0.034	0.019	1.000			
IIF	0.064	0.241	-0.150	0.066	1.000		
RMW	-0.234	-0.396	-0.305	0.055	-0.119	1.000	
CMA	-0.008	0.203	0.166	0.045	0.030	-0.297	1.000

Note: This table reports the descriptive statistics of the model variables. Panel A reports the summary statistics, and Panel B displays the correlations between the variables. The daily time-series statistics are reported for the factors. MP is the value-weighted daily market excess return; SMB is the daily return, which is the difference in the returns of portfolios between small stocks and large stocks; HML is the daily return, which is the difference in the returns of portfolios between high BM stocks and low BM stocks; WML is the daily return, which is the difference in the returns of portfolios between past winners and past losers; RMW is the daily return, which is the difference in the returns of portfolios between stocks with robust profitability and stocks with weak profitability; and CMA is the daily return, which is the difference in the returns of portfolios between stocks of low investment firms and stocks of high investment firms. The construction of these factors follows those of Fama and French (1992, 2012, 2015). The IIF is the daily return, which is the difference in the returns of portfolios between stocks of low informational efficiency and stocks of high informational efficiency.

Table 2: Number of significant factor coefficients in the factor models

Panel A: 25 portfolios by the five size-quintile and five AE-quintile portfolios

PANEL A	CAPM		FF3F		IIF3F		IIF4F		WML4F		IIF5F		FF5F		IIF6F	
	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN
Intercept	25	17	15	5	15	5	15	6	16	5	16	6	15	6	17	6
MP(VW)	25	25	23	25	23	25	23	25	24	25	24	25	25	25	25	25
SMB			25	25	25	25	25	25	25	25	25	25	25	25	25	25
HML			19	23			19	23	20	23	20	23	18	23	19	24
WML									14	6	18	6				
RMW													17	13	18	13
CMA													22	22	22	22
IIF					17	21	17	21			17	21			17	21

Panel B: 25 portfolios by the five BM-quintile and five AE-quintile portfolios

PANEL B	CAPM		FF3F		IIF3F		IIF4F		WML4F		IIF5F		FF5F		IIF6F	
	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN
Intercept	25	6	10	2	14		11	1	9	2	9	1	7	1	8	1
MP(VW)	25	25	20	25	21	25	21	25	21	25	21	25	24	25	24	25
SMB			25	22	25	23	25	22	25	22	25	22	25	23	25	23
HML			23	25			23	25	23	25	23	25	22	23	22	24
WML									19	7	21	8				
RMW													15	7	15	8
CMA													18	17	18	18
IIF					20	22	21	22			22	22			21	22

Panel C: 25 AE-quintile portfolios

PANEL C	CAPM		FF3F		IIF3F		IIF4F		WML4F		IIF5F		FF5F		IIF6F	
	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN	US	CN
Intercept	25	7	4	5	7	5	6	5	4	4	6	5	3	5	2	5
MP(VW)	25	25	24	25	25	25	25	25	24	25	25	25	24	25	24	25
SMB			25	23	25	23	25	23	25	23	25	23	25	22	25	23
HML			21	19			21	19	20	19	20	19	18	20	18	20
WML									16	7	20	7				
RMW													16	10	16	10
CMA													18	12	18	14
IIF					21	21	21	21			20	21			21	21

Note: This table shows the numbers of significant factor coefficients in the U.S. and Chinese stock markets from the factor models at the 5% significance level. The US and CN columns show the results for the U.S. and Chinese stock markets, respectively. The eight models are the CAPM (Eq. (2)), FF3F (Eq. (3)), IIF3F (Eq. (4)), WML4F (Eq. (5)), FF5F (Eq. (6)), IIF4F (Eq. (7)), IIF5F (Eq. (8)), and IIF6F (Eq. (9)). At the end of July in each year, all stocks are sorted according to their market capitalization (size), the book-to-market ratio of equity (BM), and AE. Firms are then assigned to quintiles independently by size and AE, BM and AE, or AE only. Then, 25 portfolios are constructed at the intersection of the five size-quintile and five AE-quintile portfolios (Panel A), five BM-quintile and five AE-quintile portfolios (Panel B), or 25 AE-quintile portfolios only (Panel C).

Table 3: Coefficients of the IIF in the regressions with IIF3F, IIF4F, IIF5F, and IIF6F

Panel A: Coefficients of the IIF in the AE-size quintiles

			AE Quintiles				
			Low				High
IIF3F_US	**Size Quintiles**	Small	0.097*	0.060	-0.212***	-0.347***	-0.324***
			(1.57)	(1.10)	(-4.30)	(-5.58)	(-5.87)
			0.281***	0.091*	-0.116**	-0.229***	-0.555***
			(4.94)	(1.46)	(-1.78)	(-5.55)	(-11.39)
			0.245***	0.020	0.035	-0.164***	-0.367***
			(5.93)	(0.43)	(0.78)	(-2.97)	(-9.58)
			-0.020	0.020	-0.002	-0.271***	-0.231***
			(-0.46)	(0.63)	(-0.06)	(-6.20)	(-6.66)
			0.548***	0.085**	0.059	-0.042	-0.776***
		Large	(9.33)	(1.86)	(0.78)	(-0.84)	(-7.85)
IIF6F_US	**Size Quintiles**	Small	0.097*	0.060	-0.212***	-0.347***	-0.324***
			(1.57)	(1.09)	(-4.41)	(-6.18)	(-6.05)
			0.281***	0.091**	-0.116**	-0.229***	-0.555***
			(5.90)	(1.96)	(-2.05)	(-6.55)	(-12.66)
			0.245***	0.020	0.035	-0.164***	-0.367***
			(8.46)	(0.57)	(0.96)	(-3.95)	(-9.74)
			-0.020	0.020	-0.002	-0.271***	-0.231***
			(-0.55)	(0.73)	(-0.07)	(-6.41)	(-6.61)
			0.548***	0.085***	0.059	-0.042	-0.776***
		Large	(13.71)	(2.51)	(1.17)	(-0.83)	(-8.29)
IIF3F_CN	**Size Quintiles**	Small	0.312***	0.216***	0.152***	-0.049	-0.135***
			(7.42)	(6.04)	(3.27)	(-1.21)	(-3.36)
			0.386***	0.227***	0.052*	-0.153***	-0.316***
			(8.82)	(6.02)	(1.46)	(-3.34)	(-7.47)
			0.345***	0.208***	0.115***	-0.041	-0.184***
			(10.04)	(5.16)	(2.46)	(-1.14)	(-4.38)
			0.295***	0.240***	0.123***	0.017	-0.150***
			(6.23)	(4.66)	(2.78)	(0.40)	(-3.77)
			0.763***	0.240***	0.195***	-0.604***	-0.598***
		Large	(14.18)	(3.31)	(4.56)	(-6.05)	(-9.27)
IIF6F_CN	**Size Quintiles**	Small	0.299***	0.206***	0.141***	-0.058*	-0.144***
			(8.11)	(6.56)	(3.25)	(-1.57)	(-3.75)
			0.385***	0.228***	0.047*	-0.159***	-0.322***
			(9.25)	(6.65)	(1.45)	(-3.51)	(-8.12)
			0.341***	0.206***	0.111***	-0.041	-0.189***
			(11.08)	(5.65)	(2.54)	(-1.25)	(-4.57)
			0.297***	0.242***	0.121***	0.017	-0.148***
			(7.01)	(5.00)	(2.97)	(0.42)	(-3.89)
			0.770***	0.247***	0.194***	-0.613***	-0.585***
		Large	(14.14)	(3.60)	(4.88)	(-6.14)	(-9.10)

Panel B: Coefficients of the IIF in the AE-BM quintiles

			AE Quintiles				
			Low				High
IIF3F_US	**BM Quintiles**	Low	0.650***	0.185**	-0.032	-0.139***	-0.404***
			(8.76)	(2.07)	(-0.31)	(-2.51)	(-4.80)
			0.347***	-0.019	0.156***	-0.279***	-0.725***
			(6.87)	(-0.49)	(2.59)	(-5.61)	(-5.81)
			0.232***	0.083**	0.069**	-0.128***	-0.525***
			(5.21)	(1.85)	(1.79)	(-2.97)	(-9.43)
			0.287***	0.074	-0.016	-0.257***	-0.202***
			(4.08)	(1.24)	(-0.34)	(-4.68)	(-3.64)
			0.568***	0.138*	-0.040	-0.117**	-0.448***
		High	(5.08)	(1.64)	(-0.65)	(-1.77)	(-8.18)
IIF6F_US	**BM Quintiles**	Low	0.640***	0.167***	-0.081*	-0.160***	-0.434***
			(11.64)	(2.88)	(-1.28)	(-3.20)	(-6.93)
			0.332***	-0.010	0.155***	-0.288***	-0.733***
			(7.18)	(-0.26)	(2.51)	(-6.25)	(-6.13)
			0.232***	0.085**	0.074**	-0.115***	-0.517***
			(5.45)	(1.84)	(2.04)	(-3.46)	(-9.30)
			0.296***	0.090**	-0.005	-0.259***	-0.192***
			(4.76)	(1.95)	(-0.17)	(-4.67)	(-4.60)
			0.516***	0.162***	-0.027	-0.094**	-0.435***
		High	(5.54)	(2.74)	(-0.60)	(-1.97)	(-9.79)
IIF3F_CN	**BM Quintiles**	Low	0.449***	0.278***	0.152**	-0.013	-0.251***
			(8.04)	(4.14)	(2.17)	(-0.20)	(-3.93)
			0.478***	0.213***	0.119**	-0.034	-0.329***
			(8.05)	(4.05)	(2.11)	(-0.65)	(-6.61)
			0.626***	0.156**	0.094**	-0.571***	-0.389***
			10	(2.10)	(1.90)	(-4.30)	(-5.96)
			0.759***	0.376***	0.235***	-0.067	-0.243***
			(8.99)	(7.18)	(5.02)	(-1.17)	(-4.33)
			0.308***	0.246***	0.191***	-0.219***	-0.316***
		High	(7.23)	(4.78)	(3.69)	(-4.23)	(-5.64)
IIF6F_CN	**BM Quintiles**	Low	0.448***	0.276***	0.147***	-0.015	-0.253***
			(12.01)	(5.64)	(2.93)	(-0.29)	(-5.00)
			0.483***	0.215***	0.113***	-0.031	-0.332***
			(9.81)	(5.03)	(2.44)	(-0.71)	(-7.08)
			0.608***	0.164***	0.093**	-0.579***	-0.389***
			(11.01)	(2.41)	(1.90)	(-4.40)	(-6.17)
			0.758***	0.377***	0.224***	-0.065	-0.235***
			(8.61)	(7.20)	(4.96)	(-1.10)	(-4.13)
			0.308***	0.248***	0.183***	-0.233***	-0.307***
		High	(8.05)	(4.70)	(3.44)	(-4.64)	(-6.28)

	IIF3F_US	IIF6F_US	IIF3F_CN	IIF6F_CN
Low	0.433***	0.433***	0.511***	0.509***
	(8.60)	(8.63)	(9.57)	(11.62)
	0.081**	0.081**	0.404***	0.399***
	(2.24)	(2.22)	(9.24)	(9.43)
	0.244***	0.244***	0.515***	0.514***
	(3.20)	(3.75)	(9.54)	(10.33)
	-0.005	-0.005	0.543***	0.537***
	(-0.06)	(-0.07)	(10.78)	(11.17)
	-0.162***	-0.162***	0.756***	0.755***
	(-2.81)	(-3.09)	(8.83)	(8.82)
	-0.088**	-0.088**	0.479***	0.486***
	(-1.89)	(-1.89)	(10.06)	(10.08)
	0.041	0.041	0.306***	0.299***
	(0.70)	(0.77)	(5.57)	(6.18)
	-0.095***	-0.095***	0.391***	0.388***
	(-2.54)	(-2.56)	(6.62)	(7.30)
	-0.057	-0.057	0.093**	0.099**
	(-1.12)	(-1.13)	(2.11)	(2.25)
	-0.085*	-0.085*	0.052	0.064
	(-1.33)	(-1.63)	(0.64)	(0.85)
	-0.297***	-0.297***	0.199***	0.189***
	(-5.16)	(-6.13)	(4.29)	(4.44)
	0.893***	0.893***	0.156***	0.157***
	(10.19)	(14.42)	(2.87)	(3.09)
	-0.404***	-0.404***	0.210***	0.195***
	(-7.27)	(-7.70)	(3.70)	(3.61)
	-0.685***	-0.685***	0.116***	0.117***
	(-5.23)	(-5.27)	(2.42)	(2.57)
	-0.239***	-0.239***	0.052	0.055
	(-4.74)	(-4.58)	(0.94)	(1.04)
	-0.398***	-0.398***	0.123***	0.115***
	(-5.29)	(-5.22)	(2.39)	(2.38)
	-0.584***	-0.584***	0.076*	0.082*
	(-9.65)	(-9.90)	(1.36)	(1.52)
	-0.591***	-0.591***	-0.012	-0.012
	(-6.93)	(-8.18)	(-0.25)	(-0.24)
	0.335***	0.335***	-0.917***	-0.932***
	(4.98)	(6.60)	(-8.89)	(-9.01)
	0.315***	0.315***	-0.418***	-0.420***
	(8.35)	(8.36)	(-7.11)	(-6.92)
	0.228***	0.228***	-0.349***	-0.343***
	(3.49)	(4.32)	(-4.87)	(-4.79)
	0.063	0.063	-0.259***	-0.253***
	(1.01)	(1.04)	(-6.08)	(-6.04)
	0.343***	0.343***	-0.276***	-0.272***
	(5.47)	(7.09)	(-5.49)	(-5.46)
	-0.053*	-0.053*	-0.339***	-0.342***
	(-1.44)	(-1.37)	(-6.25)	(-6.22)
	-0.254***	-0.254***	-0.584***	-0.578***
High	(-2.79)	(-2.95)	(-8.40)	(-8.51)

(AE Quintiles — label along left axis)

Note: This table reports the coefficients of the IIFs in the time-series regressions with the factor models, including IIF3F, IIF4F, IIF5F, and IIF6F. At the end of July in each year, stocks are ranked into five size quintiles and five AE quintiles independently. The intersections of the two types produce 25 AE-size portfolios (Panel A). Similarly, 25 AE-BM portfolios (Panel B) are produced, while 25 AE portfolios (Panel C) are produced by ranking stocks. The IIF is the orthogonalized IIF. The t statistics are in parentheses and *, **, and *** denote the significance of the factor loading at the 10%, 5%, and 1% levels, respectively. The positive and significant coefficients are dark-shaded, and the negative and significant ones are light-shaded.

Table 4: GRS test results

Model	US stock market				Chinese stock market			
	GRS-stat	GRS-P	AAVI	Aver_r2	GRS-stat	GRS-P	AAVI	Aver_r2
Panel A: 25 AE-size portfolios								
CAPM	12.44	0.00	12.23	0.63	1.71	0.02	3.07	0.79
FF3F	5.82	0.00	3.83	0.72	1.34	0.12	1.14	0.87
IIF3F	5.52	0.00	3.76	0.73	1.29	0.15	1.11	0.87
WML4F	5.76	0.00	3.82	0.73	1.37	0.10	1.13	0.87
IIF4F	5.46	0.00	3.76	0.73	1.34	0.12	1.14	0.87
FF5F	6.22	0.00	4.09	0.73	1.39	0.09	1.13	0.87
IIF5F	5.43	0.00	3.75	0.73	1.38	0.10	1.13	0.87
IIF6F	5.86	0.00	4.10	0.74	1.41	0.09	1.13	0.88
Panel B: 25 AE-BM portfolios								
CAPM	5.86	0.00	6.65	0.64	0.77	0.78	1.48	0.78
FF3F	1.79	0.01	1.92	0.71	0.68	0.89	0.89	0.81
IIF3F	2.21	0.00	2.28	0.69	0.68	0.89	0.92	0.80
WML4F	1.72	0.01	1.80	0.71	0.70	0.86	0.88	0.81
IIF4F	2.22	0.00	2.08	0.72	0.68	0.88	0.89	0.82
FF5F	1.23	0.20	1.34	0.71	0.71	0.85	0.87	0.81
IIF5F	2.18	0.00	1.97	0.72	0.70	0.87	0.88	0.82
IIF6F	1.70	0.02	1.54	0.72	0.71	0.85	0.87	0.82
Panel C: 25 AE portfolios								
CAPM	4.74	0.00	6.47	0.64	1.35	0.12	1.64	0.80
FF3F	0.87	0.64	1.29	0.68	1.27	0.17	1.17	0.81
IIF3F	0.81	0.74	1.47	0.69	1.39	0.09	1.15	0.81
WML4F	0.80	0.74	1.17	0.69	1.36	0.11	1.19	0.81
IIF4F	0.74	0.82	1.27	0.69	1.40	0.09	1.17	0.81
FF5F	0.52	0.98	0.83	0.69	1.19	0.24	1.13	0.81
IIF5F	0.70	0.87	1.15	0.70	1.45	0.07	1.20	0.82
IIF6F	0.46	0.99	0.81	0.70	1.29	0.15	1.12	0.82

Note: This table reports the model evaluation results. At the end of July in each year, stocks are ranked into five size quintiles and five AE quintiles independently. The intersections of the two types produce 25 AE-size portfolios (Panel A). Similarly, 25 AE-BM portfolios (Panel B) are produced, while 25 AE portfolios (Panel C) are produced by ranking stocks. GRS-stat shows the value of the GRS statistics. GRS-p shows the corresponding P-value of the GRS statistics. AAVI denotes the average absolute value of the intercepts. Aver_r2 provides the average adjusted R square.

Appendix: Calculating AE

We denote an $N \times 1$ time series of the stock return data as $r(1), r(2), ..., r(N)$.

Step 1: Form a positive integer, m, and a positive real number, ε. m denotes the length of the compared run of data and r specifies the filtering level.

Step 2: Form a sequence of vectors, $x(1), x(2), ..., x(N-m+1)$. $x(i)$ is a vector of $m \times 1$, which is defined by $x(i) = [r(i), r(i+1), ..., r(i+m-1)]$.

Step 3: Select $x(i)$ and $x(j)$ from the sequence $x(1), x(2), ..., x(N-m+1)$. $x(i)$ and $x(j)$ are considered to be similar if their distance is smaller than a given tolerance, ε. The distance is defined as $d(x, x^*) = \max_a \left| r(a) - r^*(a) \right|$.

Step 4: For each $x(i)$, $1 \leq i \leq N-m+1$, define its frequency of similarity as

$C_i^m(\varepsilon) = \dfrac{n}{N-m+1}$, where n is the number of $x(j)$, which is similar to $x(i)$.

Step 5: Define $\Phi^m(\varepsilon) = (N-m+1)^{-1} \sum\limits_{i=1}^{N-m+1} \log(C_i^m(\varepsilon))$.

Step 6: Define AE as $AE = \Phi^m(\varepsilon) - \Phi^{m+1}(\varepsilon)$.

Typically, in applications, the parameters are specified by $m = 2$ and $\varepsilon = 0.15\sigma$, where σ is the standard deviation of the time series.

[i] According to Lo (2004, 2005), the AMH implies that in an ever-changing market, (1) the relationship is unstable over time, (2) arbitrage opportunities arise occasionally, (3) the performance of investment strategies varies over time, and (4) innovation is the key factor to survival.

[ii] We calculate the market excess returns by both value-weighted and equally weighted ways, which does not affect our conclusions. To save space, we report only the value-weighted results.

[iii] We also obtain robust results when deleting stocks that have more than 40, 60, 70, or 80 trading days of missing returns in the rolling windows of 500 consecutive trading days. The results are available on request.

[iv] We mainly focused on statistical significance in our paper. In addition, we also find that the IIF is of economical significance in terms of coefficient magnitude by benchmarking against HML, RMW, and CMA in Fama and French (2015). However, economical significance also relates to transaction cost, tax, risk, etc. In view of this, a comprehensive analysis of the economical significance of IIF will probably be explored in our future research.