Journal of Financial Economics 122 (2016) 270–287

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Market maturity and mispricing[☆]

Heiko Jacobs[*]

Finance Department, University of Mannheim, L5, 2, Mannheim 68131, Germany

ARTICLE INFO

Article history:
Received 18 August 2015
Revised 11 December 2015
Accepted 9 January 2016
Available online 10 August 2016

JEL Classification:
G12
G14
G15
F37

Keywords:
Anomalies
Asset pricing
Behavioral finance
International stock markets
Emerging markets

ABSTRACT

Relying on the Stambaugh, Yu, and Yuan (2015) mispricing score and on 45 countries between 1994 and 2013, I document economically meaningful and statistically significant cross-sectional stock return predictability around the globe. In contrast to the widely held belief, mispricing associated with the 11 long/short anomalies underlying the composite ranking measure appears to be at least as prevalent in developed markets as in emerging markets. Additional support for this conjecture is obtained, among others, from tests for biased expectations based on the behavior of anomaly spreads surrounding earnings announcements as well as from within-country variation in development.

1. Introduction

In their marketing materials, mutual fund companies often claim that emerging markets yield better opportunities for stock picking than developed markets.[1] However, the evidence is mixed. Dyck, Lins, and Pomorski (2013) and Huij and Post (2011) indeed find that active management outperforms passive management in emerging markets or is at least successful enough to cover its expenses. In contrast, Busse, Goyal, and Wahal (2014), Eling and Faust (2010), Ferreira, Keswani, Miguel, and Ramos (2013), or Kang, Nielsen, and Fachinotti (2011) report that mutual funds tend to underperform traditional benchmarks, and find little to no evidence for stock picking skill, superior performance, or performance persistence in emerging markets.

With respect to more specific measures of potential mispricing, particularly studies with early sample periods, such as Bekaert and Harvey (2002) or Bhattacharya, Daouk, Jorgenson, and Kehr (2000), tend to conclude that there could be larger inefficiencies in emerging markets. More recent results in Griffin, Kelly, and Nardari (2010) point to higher transaction costs and information costs in emerging markets, but also show that proxies for the violation of the weak form of market efficiency as well as the post-earnings-announcement drift are similar in developed and emerging markets. Other studies find that specific return

☆ I wish to express my thanks to Martin Weber, Hongjun Yan, and seminar participants at the University of Mannheim for valuable comments. I thank an anonymous referee for helpful comments and suggestions that significantly improved the paper. All remaining errors are my own.

* Fax: +49 621 181 1534.
E-mail address: jacobs@bank.bwl.uni-mannheim.de

[1] For instance, M&G Investments (2015, p. 18) states that "emerging markets are less efficient than developed markets, many market participants have short-term time horizons, and rapid swings in investor sentiment mean prices can often deviate from fundamentals." Fidelity (2014, p. 7) states that "emerging markets are widely accepted to be less efficient than developed stock markets. (...) These factors coupled with greater incidence of risks augur for an active approach."

http://dx.doi.org/10.1016/j.jfineco.2016.01.030
0304-405X/

phenomena even tend to be stronger in markets deemed to be more developed. Examples include Titman, Wei, and Xie (2013) and Watanabe, Xu, Yao, and Yu (2013) on the asset growth effect, Eisdorfer, Goyal, and Zhdanov (2014) on the financial distress anomaly, or Barber, George, Lehavy, and Trueman (2013) on the earnings announcement premium.

In essence, the contrasting views can be illustrated with two quotes from recent interviews[2]: "Emerging markets are less efficient than developed markets" (Richard Thaler). There is "nothing convincing we know of" to support such an assertion (Eugene Fama). In sum, the empirical evidence is far from conclusive. In this paper, I aim to revisit this controversial debate. My findings pose a challenge to the widespread perception of necessarily stronger cross-sectional mispricing in emerging markets.

Based on the Morgan Stanley Capital International (MSCI) market classification, I first construct a comprehensive international stock market data set, which covers 115 million firm days between January 1994 and December 2013. I then implement the cross-sectional composite mispricing metric proposed in Stambaugh, Yu, and Yuan (2015). Their methodological innovation is to condense the information contained in 11 well-established or recently proposed anomalies in an aggregate mispricing score for each stock month. Stambaugh, Yu, and Yuan (2015) show that both the alpha and the associated *t*-statistic are much higher in their U.S. sample when sorting on the mispricing score as opposed to averaging the estimates for the individual anomalies. In other words, the approach appears to capture inefficiencies particularly well.

Additional credibility for this conjecture comes from Akbas, Armstrong, Sorescu, and Subrahmanyam (2015). They show that "dumb money" (as proxied for by mutual fund flows) exacerbates mispricing as indicated by the metric, whereas "smart money" (as proxied for by hedge fund flows) attenuates mispricing. Further supporting evidence is provided in Stambaugh, Yu, and Yuan (2012, 2014) who show that investor sentiment drives the dynamics of each of the 11 individual anomalies underlying the mispricing score. In sum, the Stambaugh, Yu, and Yuan (2015) score arguably represents a state-of-the-art approach to identify cross-sectional mispricing based on publicly available information. For brevity, I will thus refer to this metric as "mispricing" in the remainder of the paper.

I find strong evidence for mispricing around the globe, with point estimates exceeding U.S. estimates for about a third of the 45 developed and emerging markets considered in the baseline analysis. For the average country and based on long/short mispricing quintiles, the equally weighted (value-weighted) alpha in local currency relative to a country-specific (Fama and French, 1993) three-factor model is about 107 (84) basis points (bp) per month over the 1994–2013 period.

Notably, mispricing associated with the 11 (Stambaugh, Yu, and Yuan, 2015) anomalies appears to be at least as prevalent in developed markets as in emerging markets. In fact, the alpha difference between developed and emerging markets tends to be positive, and it is often statistically significant and economically meaningful. This key finding is robust. It holds among different firm-level return weighing schemes (equally weighted or value-weighted), different country-level return weighing schemes (country average or country composite), different asset pricing models (raw returns, local factor models, global factor models), and different treatment of currency effects (local currency or USD).

All anomalies underlying the mispricing score as well as the return predictive power of the score itself were originally documented in the U.S. stock market. In this context, my key finding could be driven by two different aspects of data mining, broadly defined. First, statistical biases in the sense of Fama (1991), McLean and Pontiff (2016), or Schwert (2003) could have inflated the historical magnitude of seemingly anomalous returns in the U.S. stock market. However, many countries produce larger long/short spreads, and my results also hold after the exclusion of the U.S. as well as in post-publication years of anomalies. These findings suggest that data snooping is not a major issue.

Second, and more relevant for my purpose, the academic effort of identifying variables that reliably predict differences in cross-sectional average returns has been mainly concentrated on developed markets so far. For instance, Harvey, Liu, and Zhu (2016, p. 5) document that there are "hundreds of papers and factors" focusing solely on the U.S. market. In contrast, emerging markets appear to be "comparatively under-researched" (Fidelity, 2014, p. 7). This asymmetric attention likely has led to a better understanding of which factors truly have predictive power for returns in more mature stock markets, and the Stambaugh, Yu, and Yuan (2015) mispricing score could be partly based on such variables.[3] It is thus important to stress that my results are subject to the caveat that mispricing in emerging markets could be associated with other anomalies, perhaps yet undiscovered.

Furthermore, and as discussed in Griffin, Kelly, and Nardari (2010), comparing the relative degree of Stambaugh, Yu, and Yuan (2015) mispricing across markets is challenging as the level and the cost of information production are hard to measure. While by no means conclusive, my attempts to better understand and interpret the findings continue to support the insights from the baseline analysis.

Most notably, I explore the predictability of the market reaction around earnings announcements as well as of sell-side analysts' forecast errors. Engelberg, McLean, and Pontiff (2015) perform a similar analysis for a broad range of cross-sectional return phenomena in the U.S. market. They

2 http://media.pimco.com/Documents/15-0088-03-DCD-AprilThaler.pdf and https://www.dimensional.com/famafrench/questions-answers/qa-seeking-the-inefficient-asset-class.aspx.

3 Nevertheless, the findings in Green, Hand, and Zhang (2014), Hou, Xue, and Zhang (2015), and Jacobs (2015) collectively indicate that many of the individual anomalies underlying the Stambaugh, Yu, and Yuan (2015) mispricing score do not necessarily belong to the strongest return predictors in the U.S. stock market in terms of economic magnitude and statistical significance. In addition, I find that more than 80% of individual anomaly spreads produced in developed markets are as large or larger than those produced in emerging markets.

conclude that the return predictability is the result of mispricing caused by biased beliefs, which are partly corrected upon news arrival. My findings extend these insights to an international level. In line with the idea of biased cash-flow expectations, Stambaugh, Yu, and Yuan (2015) spreads around the globe are particularly large surrounding earnings announcements. Differences between developed and emerging markets are consistent with the hypothesis that the cross-country differences in average long/short spreads are driven by different degrees of mispricing associated with the underlying anomalies.

In addition, I exploit within-country variation in market development in two distinct settings. First, I analyze the consequences of MSCI market reclassification. If anything, the findings suggest that relative mispricing increases in countries that have been upgraded to developed markets. Second, I analyze sudden changes in the information environment caused by mandatory International Financial Reporting Standards (IFRS) adoption. Difference-in-differences estimates provide only weak evidence that this shock systematically affects mispricing.

Finally, panel regressions indicate that mispricing is positively related to firm-specific return variation, to trading activity, and, to a lesser extent, to analyst forecast dispersion. These results appear to be consistent with noise trader-based interpretations of these variables as brought forward in, for instance, Baker and Stein (2004), Baker and Wurgler (2006), Daniel, Hirshleifer, and Subrahmanyam (1998), or Hou, Peng, and Xiong (2013).

Under the assumption of institutional trading being comparatively more prevalent in developed markets, my results are consistent with the intriguing insights of Edelen, Ince, and Kadlec (2016). The authors study institutional trading in the U.S. stock market with respect to those 11 individual anomalies that enter the Stambaugh, Yu, and Yuan (2015) mispricing score. Edelen, Ince, and Kadlec (2016) show that “institutions have a strong tendency to buy stocks classified as overvalued” (p. 472), and conclude that their “evidence strongly rejects the sophisticated institutions hypothesis” (p. 473). The authors argue that friction-based limits to arbitrage cannot account for their findings. They instead propose partial explanations based on biased cash-flow expectations and, in particular, institutions’ tracking of common firm characteristics due to agency conflicts in the sense of Lakonishok, Shleifer, and Vishny (1994). Similar in spirit, the findings of DeVault, Sias, and Starks (2015) for the U.S. stock market suggest that “institutional investors, rather than individuals, are the traders who drive sentiment-induced mispricings” (p. 1).

My findings add to several streams of the literature. First, they provide novel insights into the price discovery process in emerging markets. Even though these countries are still widely neglected in the literature, they are economically highly important. For instance, based on the 2014 World Economic Outlook of the International Monetary Fund, MSCI emerging markets account for about 1/3 of the world’s gross domestic product measured in current USD. Firms in emerging markets also account for more than 14% of the 2014 Financial Times Global 500, which is a snapshot of the world’s largest companies by market capitalization. Finally, based on the most recent data available from the World Bank (as of 2012), emerging markets represent ten of the 25 countries with the largest stock market capitalization.

Second, I am the first to study the Stambaugh, Yu, and Yuan (2015) mispricing measure outside the U.S. stock market. I thereby extend the growing literature which tests for the existence and potential drivers of individual cross-sectional return phenomena (but not of a composite measure of mispricing) in an international setting. A non-exhaustive list includes Barber, George, Lehavy, and Trueman (2013), Chui, Titman, and Wei (2010), Fama and French (2012), McLean, Pontiff, and Watanabe (2009), Rouwenhorst (1998), Titman, Wei, and Xie (2013), and Watanabe, Xu, Yao, and Yu (2013). A comprehensive international analysis can enrich or challenge our understanding of price formation. For instance, out-of-sample tests help to assess to what extent seemingly abnormal returns are robust and whether they are driven by the same factors as in the original sample.

Third, my analysis may offer industry professionals insights into ways to optimize their investment process. For instance, my findings, coupled with the assumption of higher transaction costs in emerging markets, suggest that popular trading strategies based on public information tend to be more profitable in developed markets. These results also provide a partial explanation for why the mutual fund literature surprisingly often finds that active management in emerging markets is not more successful than in developed markets.

Fourth, my work contributes to the emerging literature that aims at understanding the “big picture” of anomalies. This work addresses the critique brought forward in Subrahmanyam (2010) and Richardson, Tuna, and Wysocki (2010) who highlight that we still know little about which, where, and why anomalies work (or do not work). Papers recently progressing on this front include Engelberg, McLean, and Pontiff (2015), Green, Hand, and Zhang (2013, 2014), Harvey, Liu, and Zhu (2016), Hou, Xue, and Zhang (2015), Jacobs (2015), McLean and Pontiff (2016), or Stambaugh, Yu, and Yuan (2012). The aforementioned studies explore, among others, the role of publication effects, statistical dependencies, investor sentiment, or limits to arbitrage for the magnitude of anomalies in the cross-section or the time-series of (exclusively) U.S. stocks. I add to this literature by investigating the role of financial market development by exploiting cross-country and within-country variation in a large international sample.

2. Empirical approach

2.1. Data

I gather daily equity market data at the individual firm level from the Center for Research in Security Prices (CRSP) in the case of the U.S. as well as from Datastream for all international markets. I obtain accounting data from Compustat and Worldscope, respectively. Analyst data for all markets are gathered from the Institutional Brokers’ Estimate System (I/B/E/S).

The sample period starts in January 1994 and ends in December 2013. The start date is somewhat arbitrarily set and meant to be a compromise between maximizing the length of the sample period and maximizing the number of stocks in the cross-section. Many markets, in particular emerging markets, have limited data availability in prior years. In addition, and as discussed in Griffin, Kelly, and Nardari (2010), emerging markets are widely considered to have integrated with world markets by 1994. To quantify market maturity, I use information provided by MSCI as a leading index provider.[4] I start by considering all countries which, at least at some point during the sample period, are classified as a developed market, an emerging market, or a frontier market, and for which there are stock market data available via Datastream. Nevertheless, for most tests in this paper, I concentrate on developed and emerging markets only.[5]

I consider both active and dead stocks, as reported by Datastream. I drop observations with missing identifier, return, or market capitalization. Moreover, I require that the home country of a firm equals the country in which its stock is traded. To assure that the findings are not driven by non-common equity, outliers, data errors, or the smallest and most illiquid firms, I perform a number of screens following previous work on international stock market data from Datastream (e.g., Chui, Titman, and Wei, 2010; Griffin, Kelly, and Nardari, 2010; Griffin, Hirschey, and Kelly, 2011; Hou, Karolyi, and Kho, 2011; Ince and Porter, 2006). More specifically, I apply the generic filter rules proposed in Griffin, Kelly, and Nardari (2010). The approach involves using industry code and name filters to identify and exclude preferred stock, American Depositary Receipts (ADRs), mutual funds, index funds, warrants, investment trusts, Real Estate Investment Trusts (REITs) and other forms of non-common equity. In addition, I follow Ince and Porter (2006) in deleting all firm-level return observations from the end of the sample period to the first nonzero return to make sure that already delisted firms do not distort my analysis. As an additional check, I only consider the period before the “inactive date” as reported by Worldscope.

Any daily (monthly) return over 100% (300%) which is reversed on the following day (in the next month) is treated as missing. For all markets, I drop firm months with a lagged market capitalization of less than ten million USD. Moreover, I exclude observations in which the stock market capitalization is larger than 90% of the country market capitalization. Return data and market capitalization data are winsorized at the 0.1% and the 99.9% levels. Finally, I manually screen the data for potential remaining errors. The resulting initial stock market data set (including frontier markets) is comprised of about 115 million firm days from 78 countries.

Panel A on the left side of Table 1 displays descriptive statistics for countries which, at least at some point in time during the sample period, are classified as an emerging or a developed market. The largest developed (emerging) markets based on eligible firm months are the U.S., Japan, the U.K., and Canada (China, Korea, India, and Taiwan).

2.2. Quantifying mispricing

The Stambaugh, Yu, and Yuan (2015) metric represents a bottom-up approach that synthesizes information from the following 11 cross-sectional individual anomalies:

- Financial distress (Campbell, Hilscher, and Szilagyi, 2008)
- O-Score bankruptcy probability (Ohlson, 1980)
- Net stock issues (Ritter, 1991; Loughran and Ritter, 1995)
- Composite equity issues (Daniel and Titman, 2006)
- Accruals (Sloan, 1996)
- Net operating assets (Hirshleifer, Hou, Teoh, and Zhang, 2004)
- Price momentum (Jegadeesh and Titman, 1993)
- Gross profitability (Novy-Marx, 2013)
- Asset growth (Cooper, Gulen, and Schill, 2008)
- Return on assets (Fama and French, 2006; Chen, Novy-Marx, and Zhang, 2011)
- Investment-to-assets (Titman, Wei, and Xie, 2004; Xing, 2008)

Detailed descriptions of these anomalies are provided in Stambaugh, Yu, and Yuan (2012, 2015). My implementation largely follows these papers. The international setting requires a few adjustments that likely result in conservative alpha estimates.[6]

To measure aggregate mispricing, and for each anomaly-month-country combination, I first rank stocks in a way that the presumably most underpriced (overpriced) stock receives the lowest (highest) rank. Ranks are standardized to be uniformly distributed over the interval (0,1] in each country-month. A stock's composite rank is then computed as the arithmetic average of its individual anomaly ranks. Following (Stambaugh, Yu, and Yuan, 2015), I require at least five individual anomalies to construct a composite rank for a given stock month. Finally, I again normalize this aggregate mispricing measure to be uniformly distributed over the interval (0,1]. For each country-month, I then construct a long/short strategy which goes long (short) stocks in the bottom (top) quintile of mispricing. In other words, the approach is country-neutral in that the number of underpriced stocks (roughly) equals the number of overpriced stocks for each country.

The construction of the mispricing score requires the availability of lagged stock market data and, in particular, accounting data. As the Online Appendix shows in more detail, valid accounting data appear to be available for

[4] See https://www.msci.com/market-classification for details.

[5] For frontier markets, the sample period is much shorter and data availability and reliability turn out to be lower. In addition, the economic importance is lower. Finally, focusing on developed and emerging markets facilitates presentation. I make use of the information contained in frontier markets in Section 5.1.

[6] First, I rely on yearly (as opposed to quarterly) accounting data due to limited data availability. Second, and to assure real-time data availability and comparability across countries, I impose a conservative lag of at least 6 months after the fiscal year end (e.g., Fama and French, 1993). Third, I focus on portfolio quintiles instead of deciles due to the low number of stocks in some countries. Information on the construction of the anomalies is provided in the Online Appendix.

Table 1
Descriptive statistics.

Panel A shows descriptive characteristics for all countries that have eligible stock market data (as described in detail in Section 2.1) and that are classified as a developed or emerging market at least at some point during the sample period ranging from January 1994 to December 2013. *DM/EM/FM* denotes developed/emerging/frontier markets. Firm size is measured in millions USD. Panel B shows descriptive statistics for stocks with a valid mispricing score. In Panel B, the *start* and *end* of the sample period are determined by the first and last date that satisfy the following two criteria. First, there are least 25 eligible stocks with non-missing mispricing scores at least 36 months in a row. Second, the country is classified as a developed or emerging market. Idiosyncratic volatility (*idio vola*) is computed as the standard deviation of the residual obtained from rolling regressions of daily excess returns on a local (Fama and French, 1993) three-factor model over the months $t-12$ to $t-1$. *Fraction zero return days* is based on returns in local currency.

Country	Market	Total no. firms	Min no. firms	Max no. firms	Firm months (in 1000)	Mean size	Median size	Start	End	Total no. firms	Mean size	Idio. vola	Fraction zero return days
	Panel A: Firms in the baseline stock market data set							*Panel B: Firms with valid mispricing score*					
Argentina	EM/FM	103	34	69	14	557	88	May-00	May-09	66	647	2.14%	44%
Australia	DM	2,278	286	1,262	172	771	54	Jan-94	Dec-13	1,917	955	3.69%	28%
Austria	DM	157	59	87	18	868	140	Jan-94	Dec-13	134	888	1.92%	38%
Belgium	DM	211	74	129	25	1,674	153	Jan-94	Dec-13	182	1,858	1.91%	28%
Brazil	EM	221	37	154	21	2,736	371	Mar-98	Dec-13	190	3,297	2.58%	39%
Canada	DM	2,586	220	1,688	176	991	57	Jan-94	Dec-13	1,762	1,404	4.17%	24%
Chile	EM	253	115	164	36	804	166	Jan-94	Dec-13	211	909	1.37%	62%
China	EM	2,465	135	2,366	281	1,049	343	Jul-95	Dec-13	2,434	1,012	2.00%	14%
Colombia	EM	93	27	52	11	1,368	194	Jul-07	Dec-13	55	4,172	1.24%	54%
Czech	EM	75	4	61	6	903	84						
Denmark	DM	286	120	185	36	753	83	Jan-94	Dec-13	260	773	2.08%	41%
Egypt	EM	165	4	113	18	451	100	Jul-04	Dec-13	119	687	2.18%	20%
Finland	DM	176	42	124	25	1,435	169	Jan-94	Dec-13	167	1,519	2.25%	28%
France	DM	1,419	46	720	144	2,021	96	Jan-94	Dec-13	1,198	2,321	2.44%	27%
Germany	DM	1,238	371	752	138	1,769	100	Jan-94	Dec-13	1,087	1,925	2.59%	29%
Greece	EM/DM	387	85	301	50	381	65	Jan-94	Dec-13	358	430	2.54%	21%
Hong Kong	DM	186	108	156	32	2,922	180	Jan-94	Dec-13	169	3,315	2.71%	30%
Hungary	EM	66	8	32	6	793	79						
India	EM	2,995	458	1,786	235	505	50	Jul-94	Dec-13	2,041	762	2.68%	10%
Indonesia	EM	484	90	351	50	606	72	Jan-94	Dec-13	419	671	3.18%	53%
Ireland	DM	89	35	57	10	1,205	184	Feb-96	Dec-13	78	1,201	3.01%	26%
Israel	EM/DM	666	169	385	67	303	37	Aug-98	Dec-13	381	570	2.22%	27%
Italy	DM	449	162	270	53	1,986	243	Jan-94	Dec-13	412	2,086	1.99%	13%
Japan	DM	4,610	2,394	3,613	755	1,022	133	Jan-94	Dec-13	4,401	1,082	2.34%	25%
Jordan	EM/FM	201	89	165	13	221	32	Jul-07	Nov-08	123	170	1.99%	37%
Korea	EM	2,469	464	1,702	281	384	50	Jan-94	Dec-13	2,042	513	3.07%	14%
Malaysia	EM	1,105	326	742	140	344	53	Jan-94	Dec-13	1,056	396	2.56%	55%
Mexico	EM	206	93	115	25	1,713	320	Jan-94	Dec-13	168	2,023	1.81%	29%
Morocco	EM/FM	86	19	71	12	577	110	Jul-06	Nov-13	69	981	1.87%	44%
Netherlands	DM	238	75	171	28	2,890	228	Jan-94	Dec-13	201	2,692	2.09%	21%
New Zealand	DM	211	56	107	18	327	77	Jul-97	Dec-13	145	304	2.14%	43%
Norway	DM	380	73	199	35	899	107	Jan-94	Dec-13	328	1,021	2.86%	36%
Pakistan	EM/FM	265	66	179	31	186	46	Jul-94	Dec-08	157	246	2.28%	35%
Peru	EM	16	2	7	1	63	36						
Philippines	EM	225	71	166	28	477	77	Jul-94	Dec-13	212	548	3.13%	53%
Poland	EM	624	12	369	40	436	54	Aug-98	Dec-13	427	494	2.70%	20%
Portugal	EM/DM	127	35	85	14	828	112	Jan-94	Dec-13	97	940	2.15%	34%
Russia	EM	473	9	318	29	2,669	111	Jul-05	Dec-13	356	3,718	2.89%	48%
Singapore	DM	744	160	527	87	511	62	Jan-94	Dec-13	725	577	3.42%	42%
South Africa	EM	724	187	365	60	970	143	Jan-94	Dec-13	498	1,265	2.64%	36%
Spain	DM	207	89	147	30	3,758	499	Jan-94	Dec-13	184	4,031	1.70%	22%
Sri Lanka	EM/FM	230	16	176	17	74	27						
Sweden	DM	650	103	339	57	1,222	93	Jan-94	Dec-13	538	1,337	2.70%	22%
Switzerland	DM	386	162	259	50	3,479	270	Jan-94	Dec-13	297	3,820	1.96%	28%
Taiwan	EM	1,958	262	1,589	226	477	96	Jul-94	Dec-13	1,814	504	2.12%	18%
Thailand	EM	668	165	469	76	377	55	Jan-94	Dec-13	603	404	2.48%	33%
Turkey	EM	392	90	310	52	548	75	May-94	Dec-13	333	689	2.50%	19%
UK	DM	3,040	870	1,418	261	1,872	100	Jan-94	Dec-13	2,484	2,094	2.27%	45%
USA	DM	13,366	3,503	6,782	1,191	2,481	209	Jan-94	Dec-13	11,458	2,861	3.27%	9%
Venezuela	EM	48	10	32	5	290	72						

a significantly larger fraction of companies in developed markets than in emerging markets in the first half of the sample period (1994–2003). In contrast, differences in accounting data availability are indistinguishable from zero during the second half (2004–2013). Section 3.2. verifies that mispricing is at least as prevalent in developed markets as in emerging markets in both subperiods. In addition, I find the same pattern for anomalies constructed solely from market data (such as momentum), for which there are no differences in data availability during any part of the sample period. In sum, these results suggest that partial differences in cross-sectional data coverage during the first half of the sample period cannot explain my key finding.

In the average country-year of my final sample, I am able to compute a valid mispricing score for about 69% (82%) of firms in emerging (developed) markets. In the second half of the sample period (2004–2013), these numbers increase to 86% (89%). The average number of individual anomalies that enters a valid composite mispricing score in the average emerging (developed) country-year is 9.72 (9.85). The difference between more and less mature stock markets is small in each year of the sample period. For instance, in 1994 (2013), the respective numbers are 9.50 and 9.68 (10.06 and 10.19). In sum, these findings suggest that the alpha difference between developed and emerging markets is not attributable to inequalities in the number of signals entering the mispricing score.

For each country, the final sample period is determined by the first and last date that satisfy the following two criteria. First, there are least 25 eligible stocks with non-missing mispricing scores at least 36 months in a row. Second, the country is classified as a developed or emerging market. These screens imply that a few countries (such as Hungary) do not enter the asset pricing tests due to a lack of data availability. It also implies that a few other countries (such as Argentina) have a shortened sample period due to market reclassification. In total, and as shown on a country-by-country basis in Panel B of Table 1, I obtain a mispricing score for about 42,600 firms in 45 markets. These numbers increase to about 44,100 firms in 57 countries once the (untabulated) frontier markets are additionally taken into account.[7]

As Panel B of Table 1 shows, firms with a valid mispricing score tend to be larger and more liquid than the unconditional firm universe. There are (potentially surprisingly) few differences between developed markets and emerging markets with respect to average idiosyncratic volatility, which is often interpreted as a proxy for limits to arbitrage (e.g., Pontiff, 2006). Similarly, there are only few differences with respect to the average fraction of days with zero return, which could be regarded as a proxy for trading activity or liquidity (e.g., Lesmond, Ogden, and Trzcinka, 1999). In Section 5.1., I use these and further country characteristics in panel regressions aimed at explaining variation in mispricing.

2.3. Return computation

Essentially, I want to benchmark potential mispricing in emerging markets against potential mispricing in developed markets. Empirically estimating this relation leaves many degrees of freedom. I therefore implement several alternative methods. In addition, I later extensively test for the robustness of these baseline specifications.

First, I compute returns in local currency. Second, and to aggregate long/short returns for developed markets on the one hand and for emerging markets on the other hand, I apply four difference specifications. These specifications differ in the way returns are aggregated within a country (equally weighted; value-weighted) as well as in the way returns are aggregated across countries (country average; country composite). “Country average” means that the final time-series of long/short returns is the arithmetic average of all eligible country-level long/short returns. Thus, the unit of observation is a country-month, and smaller countries with less firms tend to dominate the measure. “Country composite” means that I pool all (country-neutral) overpriced and underpriced stocks and construct just one time-series of long/short returns. Thus, the unit of observation is a firm-month, and larger countries with more firms tend to dominate the measure. I compute all four estimates to test for the sensitivity of the findings, as it is not clear from an empirical point of view which specification should be preferred.[8]

Third, with respect to the asset pricing model, I use both raw long/short returns (as in, for instance, Griffin, Kelly, and Nardari, 2010) and Fama and French (1993) three-factor adjusted abnormal returns (as in, for instance, Stambaugh, Yu, and Yuan, 2012). In the context of international data, this raises the additional question on how exactly the three-factor models should be constructed. As the evidence in Griffin (2002), Hou, Karolyi, and Kho (2011), and Rouwenhorst (1999) collectively points to the importance (and potential superiority) of local factors, I construct country-specific Fama and French (1993) models, and use them in the country-level regressions. In those cases in which the dependent variable is mispricing aggregated over all developed or emerging markets, I use a global three-factor model. The model comprises all developed and emerging markets, and is constructed as country average or county composite, depending on the mispricing measure under consideration. Global factor models are constructed country-neutral in that (roughly) the same number of stocks of a given country enters the long and short legs of the factors.

3. Mispricing in developed vs. emerging markets

3.1. Baseline analysis

Separately for developed and emerging markets, Table 2 (Table 3) shows raw long/short returns (three-factor alphas) based on mispricing quintiles. There is strong evidence for return predictability around the globe. This holds true for many individual anomalies and, in particular, for the aggregate mispricing measure, which stands in the focus of this paper.

In each specification, the mispricing score yields statistically significant and economically meaningful long/short

[7] In addition to prolonged sample periods for Argentina, Jordan, Morocco, and Pakistan, this leads to the inclusion of Bulgaria, Kenya, Kuwait, Nigeria, Oman, Qatar, Romania, Saudi Arabia, Slovenia, Sri Lanka, United Arab Emirates, and Vietnam.

[8] On the one hand, large stocks are economically more important and thus of particular interest. On the other hand, a few large stocks can drive long/short returns in smaller markets. Thus, both equally weighting returns (in particular after the elimination of very small stocks) and value weighting returns can have their merits. Similar arguments apply to the cross-country perspective. On the one hand, the country composite measure gives the largest weight to the economically most important markets. On the other hand, this approach is driven by a handful of markets (see, for instance, Table 1).

Table 2
Baseline results: long/short returns in developed vs. emerging markets.
The table reports the monthly raw return (in %) obtained from the quintile-based long/short portfolio of individual anomalies or aggregated mispricing [computed as in Stambaugh, Yu, and Yuan (2015) and explained in detail in Section 2]. The table also reports the difference of the raw long/short return obtained in developed markets and the raw long/short return obtained in emerging markets. In Panel A, returns in a given month are computed as the arithmetic average of all eligible country-level return estimates. In Panel B, all eligible stocks from all eligible countries are pooled before a country-neutral time-series of returns is constructed. The sample period is January 1994 to December 2013. *T*-statistics (in parentheses) are based on the heteroskedasticity-consistent standard errors of White (1980). Two-tailed statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

	Developed	Emerging	Diff	Developed	Emerging	Diff
	Equally weighted returns			Value-weighted returns		
Panel A: Country average						
Failure probability	0.690***	0.291	0.399**	0.525**	0.343	0.183
	(3.54)	(1.64)	(2.22)	(2.04)	(1.63)	(0.79)
Ohlson's O (distress)	0.452***	0.215*	0.237	0.427**	0.316**	0.110
	(3.48)	(1.66)	(1.64)	(2.51)	(1.99)	(0.53)
Net stock issues	0.727***	0.444***	0.283**	0.248**	0.207**	0.041
	(5.61)	(5.58)	(2.33)	(1.97)	(2.20)	(0.29)
Composite equity	0.802***	0.784***	0.018	0.393**	0.368***	0.025
	(5.76)	(6.44)	(0.14)	(2.37)	(2.68)	(0.14)
Total accruals	0.333***	0.352***	-0.019	0.528***	0.217	0.311*
	(4.68)	(4.22)	(-0.18)	(4.34)	(1.47)	(1.82)
Net operating assets	0.466***	0.340***	0.126	0.387***	0.303**	0.084
	(5.51)	(3.08)	(1.04)	(3.58)	(2.11)	(0.48)
Momentum	1.087***	0.501***	0.587***	0.890***	0.672***	0.218
	(5.12)	(3.08)	(3.74)	(3.38)	(3.48)	(1.02)
Gross profitability	0.392***	0.323**	0.069	0.315**	0.432***	-0.117
	(4.30)	(2.40)	(0.50)	(2.58)	(2.85)	(-0.75)
Asset growth	0.269**	0.199*	0.070	0.212*	0.135	0.077
	(2.35)	(1.91)	(0.53)	(1.72)	(1.00)	(0.48)
Return on assets	0.438***	0.061	0.377***	0.291*	0.105	0.185
	(3.83)	(0.54)	(2.85)	(1.73)	(0.81)	(1.05)
Investment-to-assets	0.250***	0.343***	-0.093	0.171	0.069	0.103
	(3.08)	(3.36)	(-0.78)	(1.64)	(0.58)	(0.68)
Mispricing score	1.090***	0.654***	0.436***	0.688***	0.433***	0.255
	(6.25)	(4.01)	(3.02)	(3.19)	(2.60)	(1.42)
Panel B: Country composite						
Failure probability	0.443*	0.368**	0.075	0.399	0.159	0.240
	(1.71)	(1.97)	(0.31)	(1.22)	(0.58)	(0.63)
Ohlson's O (distress)	0.418**	0.207	0.212	0.338	0.373*	-0.035
	(2.53)	(1.56)	(1.17)	(1.40)	(1.66)	(-0.11)
Net stock issues	0.833***	0.714***	0.119	0.421**	0.241	0.180
	(4.65)	(8.39)	(0.75)	(2.47)	(1.48)	(0.84)
Composite equity	0.693***	0.914***	-0.221	0.388**	0.250	0.138
	(3.72)	(7.84)	(-1.39)	(2.47)	(1.16)	(0.59)
Total accruals	0.306***	0.383***	-0.077	0.313**	0.240	0.073
	(5.61)	(5.60)	(-0.86)	(2.27)	(1.19)	(0.30)
Net operating assets	0.555***	0.408***	0.148	0.300***	0.290*	0.010
	(5.95)	(4.45)	(1.29)	(2.69)	(1.79)	(0.05)
Momentum	0.718**	0.343	0.375	0.299	0.244	0.055
	(2.20)	(1.55)	(1.54)	(0.81)	(0.93)	(0.15)
Gross profitability	0.490***	0.429***	0.060	0.452***	0.571***	-0.120
	(5.04)	(3.13)	(0.39)	(2.71)	(3.30)	(-0.59)
Asset growth	0.575***	0.278***	0.297**	0.264	-0.042	0.306
	(5.60)	(2.80)	(2.36)	(1.65)	(-0.19)	(1.34)
Return on assets	0.337**	0.142	0.194	0.346*	0.063	0.283
	(1.99)	(1.14)	(1.04)	(1.66)	(0.30)	(1.02)
Investment-to-assets	0.432***	0.308***	0.124	0.174	0.066	0.108
	(5.10)	(3.27)	(1.06)	(1.37)	(0.37)	(0.53)
Mispricing score	0.991***	0.748***	0.243	0.533**	0.341*	0.192
	(4.88)	(4.48)	(1.38)	(2.36)	(1.95)	(0.89)

returns as well as three-factor alphas.[9] In line with Stambaugh, Yu, and Yuan (2015), both the alpha and the Sharpe ratio of the composite anomaly ranking measure are considerably larger than the estimates obtained from averaging individual anomalies. This finding points to the ability of the approach to identify mispriced stocks.

Notably, point estimates of mispricing are higher in developed than in emerging markets. Specifically, long/short

[9] Following (Stambaugh, Yu, and Yuan, 2012; 2014; 2015), I rely on the heteroskedasticity-consistent standard errors of White (1980). Relying on Newey and West (1987) standard errors does not change inferences.

Table 3
Baseline results: Three-factor alphas in developed vs. emerging markets.
The table reports monthly alphas (in %) obtained from regressing the quintile-based long/short portfolios of individual anomalies or aggregated mispricing [computed as in Stambaugh, Yu, and Yuan (2015) and explained in detail in Section 2] on a global (Fama and French, 1993) three-factor model. The table also reports the difference of the alpha obtained in developed markets and the alpha obtained in emerging markets. In Panel A, long/short returns in a given month are computed as the arithmetic average of all eligible country-level return estimates. In Panel B, all eligible stocks from all eligible countries are pooled before a country-neutral time-series of long/short returns is constructed. The sample period is January 1994 to December 2013. *T*-statistics (in parentheses) are based on the heteroskedasticity-consistent standard errors of White (1980). Two-tailed statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

	Developed	Emerging	Diff	Developed	Emerging	Diff
	Equally weighted returns			Value-weighted returns		
Panel A: Country average						
Failure probability	0.899***	0.527***	0.373**	0.965***	0.676***	0.289
	(6.53)	(3.73)	(2.09)	(5.20)	(3.96)	(1.27)
Ohlson's O (distress)	0.485***	0.281**	0.204	0.525***	0.429***	0.096
	(4.51)	(2.27)	(1.36)	(3.99)	(2.67)	(0.45)
Net stock issues	0.682***	0.469***	0.213*	0.332***	0.221**	0.111
	(8.10)	(6.27)	(1.97)	(3.52)	(2.22)	(0.83)
Composite equity	0.866***	0.914***	-0.048	0.569***	0.391***	0.178
	(9.26)	(9.18)	(-0.37)	(4.29)	(3.10)	(1.01)
Total accruals	0.378***	0.361***	0.017	0.668***	0.362**	0.306
	(4.99)	(4.29)	(0.15)	(5.51)	(2.26)	(1.61)
Net operating assets	0.570***	0.521***	0.049	0.343***	0.550***	-0.207
	(6.27)	(5.01)	(0.39)	(2.91)	(3.99)	(-1.19)
Momentum	1.604***	0.924***	0.680***	1.573***	1.113***	0.460**
	(9.97)	(7.72)	(4.80)	(7.61)	(6.84)	(2.21)
Gross profitability	0.547***	0.543***	0.004	0.646***	0.713***	-0.067
	(6.15)	(4.68)	(0.03)	(5.74)	(5.02)	(-0.38)
Asset growth	0.149	0.251***	-0.101	0.149	0.147	0.002
	(1.64)	(2.72)	(-0.81)	(1.34)	(1.16)	(0.01)
Return on assets	0.560***	0.183*	0.377***	0.704***	0.269**	0.436**
	(5.59)	(1.71)	(2.93)	(5.50)	(2.11)	(2.57)
Investment-to-assets	0.248***	0.367***	-0.119	0.115	0.009	0.106
	(2.94)	(3.90)	(-1.01)	(1.04)	(0.07)	(0.64)
Mispricing score	1.328***	0.891***	0.436***	1.207***	0.681***	0.526***
	(11.20)	(7.25)	(3.00)	(7.96)	(5.10)	(2.81)
Panel B: Country composite						
Failure probability	0.545***	0.491***	0.054	0.742***	0.266	0.476
	(3.34)	(2.96)	(0.26)	(3.05)	(0.99)	(1.34)
Ohlson's O (distress)	0.426***	0.240*	0.187	0.385**	0.407*	-0.022
	(3.46)	(1.82)	(1.18)	(2.33)	(1.75)	(-0.07)
Net stock issues	0.770***	0.729***	0.042	0.474***	0.264*	0.210
	(8.12)	(10.33)	(0.38)	(3.68)	(1.68)	(1.01)
Composite equity issues	0.711***	0.985***	-0.273**	0.491***	0.264	0.228
	(6.77)	(10.40)	(-2.16)	(4.02)	(1.23)	(0.94)
Total accruals	0.323***	0.378***	-0.055	0.361***	0.295	0.066
	(5.79)	(5.54)	(-0.65)	(2.63)	(1.39)	(0.25)
Net operating assets	0.667***	0.483***	0.184	0.266**	0.355**	-0.089
	(7.08)	(5.42)	(1.58)	(2.39)	(2.23)	(-0.50)
Momentum	1.118***	0.546***	0.571**	0.793**	0.340	0.452
	(3.81)	(2.60)	(2.40)	(2.45)	(1.32)	(1.37)
Gross profitability	0.570***	0.555***	0.015	0.777***	0.765***	0.013
	(6.10)	(4.33)	(0.10)	(5.93)	(4.52)	(0.06)
Asset growth	0.561***	0.315***	0.247*	0.164	-0.076	0.240
	(5.66)	(3.46)	(1.91)	(1.26)	(-0.35)	(1.04)
Return on assets	0.327**	0.219*	0.107	0.627***	0.146	0.482*
	(2.59)	(1.82)	(0.67)	(3.52)	(0.71)	(1.82)
Investment-to-assets	0.479***	0.327***	0.153	0.078	0.015	0.063
	(5.69)	(3.72)	(1.29)	(0.65)	(0.09)	(0.31)
Mispricing score	1.153***	0.876***	0.277*	0.919***	0.491***	0.429**
	(8.54)	(6.23)	(1.69)	(5.81)	(3.01)	(2.11)

returns in developed (emerging) markets range from 53 bp to 109 bp (34–75 bp) per month. Alphas in developed (emerging) markets range from 92 bp to 133 bp (49–89 bp) per month. The long/short return or alpha difference between developed and emerging markets ranges from 19 bp to 53 bp, and is significant at least at the 10% level in five of the eight specifications. With respect to individual anomalies, the long/short return or alpha difference is positive (though only sporadically statistically significant) in 72 of the 88 specifications.

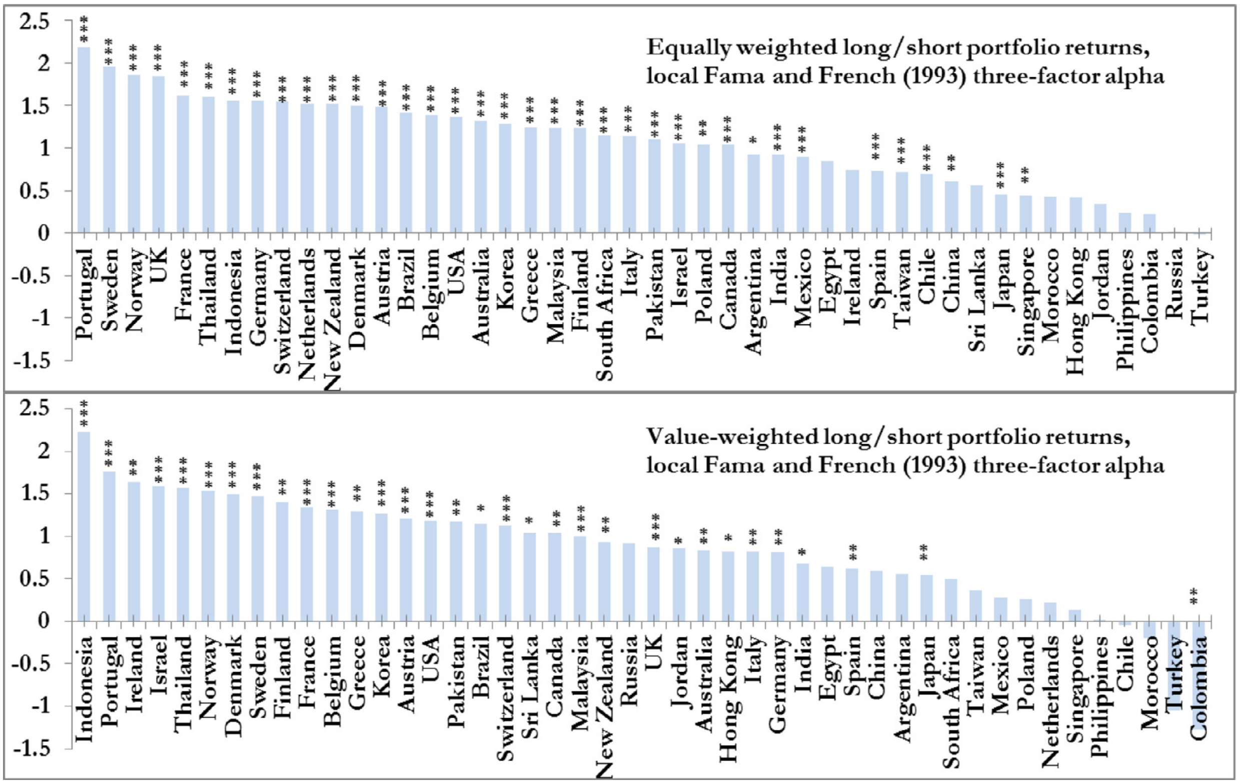

Fig. 1. Stambaugh, Yu, and Yuan (2015) mispricing per country. The figure shows country-level alphas obtained from regressing the long/short portfolio of Stambaugh, Yu, and Yuan (2015) mispricing on a local three-factor model. Each month, the portfolio goes long (short) stocks in the bottom (top) quintile of mispricing. Returns are computed in local currency and expressed in % per month. The figure considers all countries displayed in Panel A of Table 1, provided that the mispricing score can be computed for a cross-section of at least 25 firms at least 36 months in a row. The sample period is January 1994 to December 2013, or a shorter period of time depending on data availability. Statistical inference is based on the heteroskedasticity-consistent standard errors of White (1980). Two-tailed statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

In sum, judging from the aggregate mispricing score (and also from the individual anomalies), inefficiencies appear to be at least as large in developed markets as they are in emerging markets.

These insights are confirmed in Fig. 1, which displays country-level estimates of cross-sectional mispricing based on long/short quintiles. The average equally weighted (value-weighted) alpha is about 107 (84) bp per month. The alpha is positive for close to 95% of the pooled estimates. Again, mispricing seems to be particularly strong in many developed markets such as Denmark, France, Portugal, and Sweden, but rather weak in a number of emerging markets such as China, Chile, Taiwan, and Turkey. Nevertheless, there is no clear pattern. Notably, some developed markets such as Hong Kong, Japan, and Singapore yield low alphas, whereas some emerging markets such as Indonesia and Thailand generate large alphas.

3.2. Robustness checks

Table 4 shows the main insights from ten robustness checks. Unless noted otherwise, I use a global three-factor model to estimate alphas. I rely on the same four return aggregation schemes as in the baseline analysis, yielding 40 estimates in total. The main result is that inferences obtained from the baseline analysis carry over. First, virtually all alphas are economically meaningful and statistically significant. Second, the difference between the alpha obtained in developed markets and the alpha obtained in emerging markets is positive in each test, and ranges from 10 to 62 bp per month. In more than 50% of the estimates, this difference is statistically significant at least at the 5% level.

I consider three different sets of sensitivity checks. The first set deals with different sample periods. In specification 1, I use 1988 as the beginning of the sample period, which corresponds to the launching year of the MSCI Emerging Markets Index. In specifications 2 and 3, I split the baseline sample period (January 1994 to December 2013) in two halves.

The second set tests modifications of return measurement. In specification 4, I compute all returns (including the Fama/French factors) in USD. In specification five, I add a short-term reversal factor (based on the returns in month $t-1$) and a long-term reversal factor (based on the cumulative return over $t-60$ to $t-13$) to the Fama/French factors.

The third set of sensitivity checks explores changes in the country or anomaly universe. These tests are partly intended to address data mining concerns as outlined in the introduction. In specification 6, I exclude the U.S. stock market. In specification 7, I compute mispricing without the two financial distress models of Ohlson (1980) and

Table 4
Robustness tests: Mispricing in developed vs. emerging markets.
The table shows findings from sensitivity checks of the baseline analysis (see Table 3). Unless noted otherwise, the table reports alphas obtained from a global (Fama and French, 1993) three-factor model. In specification 5, factors for short-term reversal and long-term reversal are added to the model. In specification 7, mispricing is computed without the financial distress measures of Ohlson (1980) and Campbell, Hilscher, and Szilagyi (2008). In specifications 8 and 9, mispricing is computed using published anomalies only. In specification 10, I rely on an extended set of 31 individual anomalies when computing aggregate mispricing as in Stambaugh, Yu, and Yuan (2015). *T*-statistics (in parentheses) are based on the heteroskedasticity-consistent standard errors of White (1980). Two-tailed statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

		Developed	Emerging	Difference	Developed	Emerging	Difference
		Equally weighted returns			Value-weighted returns		
ID	Description	*Panel A: Country average*					
1	Using 1/1988 as start date	1.206***	0.715***	0.491***	1.096***	0.572***	0.524***
		(11.60)	(4.95)	(2.99)	(8.27)	(3.99)	(2.88)
2	Subperiod 1/1994 to 12/2003	0.858***	0.696***	0.162	0.936***	0.426**	0.510
		(3.98)	(3.32)	(0.62)	(3.21)	(1.98)	(1.48)
3	Subperiod 1/2004 to 12/2013	1.693***	1.089***	0.604***	1.378***	0.888***	0.491**
		(15.18)	(8.89)	(4.55)	(9.24)	(5.70)	(2.61)
4	All returns in USD	1.372***	0.927***	0.445***	1.247***	0.725***	0.522***
		(11.13)	(7.58)	(3.05)	(8.13)	(5.48)	(2.79)
5	Incl. reversal factors	1.292***	0.911***	0.381**	1.164***	0.698***	0.466**
		(10.86)	(7.30)	(2.54)	(8.18)	(5.08)	(2.53)
6	Excluding U.S. market	1.334***	0.891***	0.443***	1.213***	0.681***	0.533***
		(11.16)	(7.25)	(3.04)	(7.79)	(5.10)	(2.78)
7	Excluding financial distress	1.234***	0.938***	0.296**	1.083***	0.789***	0.294
		(11.04)	(8.43)	(2.25)	(7.57)	(5.92)	(1.57)
8	Only post-publication (I)	1.601***	1.161***	0.440***	1.254***	1.022***	0.232
		(13.08)	(10.09)	(3.53)	(7.86)	(5.70)	(1.01)
9	Only post-publication (II)	1.375***	0.857***	0.518***	1.386***	0.785***	0.601***
		(10.81)	(6.53)	(3.47)	(8.59)	(4.63)	(2.89)
10	Extended set of 31 anomalies	1.959***	1.344***	0.615***	1.978***	1.379***	0.598***
		(11.46)	(8.06)	(3.04)	(9.87)	(7.42)	(2.65)
ID	Description	*Panel B: Country composite*					
1	Using 1/1988 as start date	1.034***	0.618***	0.416**	0.759***	0.274	0.485**
		(8.80)	(4.01)	(2.39)	(5.40)	(1.64)	(2.32)
2	Subperiod 1/1994 to 12/2003	0.989***	0.705***	0.284	0.888***	0.466*	0.423
		(4.31)	(2.84)	(0.97)	(3.63)	(1.97)	(1.45)
3	Subperiod 1/2004 to 12/2013	1.317***	1.063***	0.253	0.939***	0.549**	0.390
		(8.68)	(8.67)	(1.61)	(4.70)	(2.58)	(1.44)
4	All returns in USD	1.159***	0.835***	0.324**	0.932***	0.428**	0.504**
		(8.45)	(5.99)	(1.99)	(5.89)	(2.48)	(2.41)
5	Incl. reversal factors	1.152***	0.876***	0.275*	0.915***	0.492***	0.423**
		(9.01)	(6.20)	(1.72)	(5.99)	(3.00)	(2.14)
6	Excluding U.S. market	1.082***	0.876***	0.206	0.789***	0.491***	0.298
		(7.97)	(6.23)	(1.28)	(4.35)	(3.01)	(1.40)
7	Excluding financial distress	1.181***	0.900***	0.282**	0.852***	0.551***	0.301
		(9.41)	(7.25)	(2.00)	(5.92)	(3.27)	(1.45)
8	Only post-publication (I)	1.242***	1.046***	0.196	0.841***	0.745***	0.096
		(7.46)	(8.64)	(1.25)	(4.09)	(3.01)	(0.31)
9	Only post-publication (II)	1.088***	0.846***	0.242	1.123***	0.728***	0.394
		(6.58)	(6.03)	(1.32)	(5.69)	(3.31)	(1.34)
10	Extended set of 31 anomalies	1.629***	1.405***	0.224	1.542***	0.917***	0.625**
		(8.26)	(6.88)	(0.96)	(6.44)	(4.33)	(2.09)

Campbell, Hilscher, and Szilagyi (2008), both of which are calibrated for the U.S. stock market. In specification 8, I require individual anomalies to have been published before I include them in the computation of (out-of-sample) mispricing. As the publication year of a given anomaly, I use the year of the oldest paper cited in Section 2.2. For instance, the investment-to-assets anomaly is considered for the years 2005–2013 only. As I require at least five valid relative ranks on individual anomalies to compute mispricing, this procedure implies a shortened sample period starting in 2005. As an alternative, specification 9 requires only three valid relative ranks, which implies that the sample period can start from 1994 again.

In specification 10, I compute aggregate mispricing based on an extended set of 31 individual anomalies, thereby including 20 cross-sectional return predictors that do not enter the original mispricing score. Examples include the post-earnings-announcement drift, the low beta anomaly, the analyst forecast dispersion anomaly, the dividend month anomaly, or the idiosyncratic risk anomaly. Reference papers and construction details are provided in the Online Appendix. The findings confirm that the positive alpha difference between developed and emerging markets is widespread, and not restricted to the 11 anomalies underlying the Stambaugh, Yu, and Yuan (2015) measure. Tests reported in the

Online Appendix provide further supportive evidence for this conjecture.

3.3. Country reclassification

There are three countries that switch from between being classified as an "emerging market" to being classified as a "developed market" between 1994 and 2013: Portugal (emerging market until November 1997), Greece (emerging market until May 2001 and again since November 2013), as well as Israel (emerging market until May 2010). MSCI (2014, p. 2) highlights that it "will only consider markets for upgrade if a change in classification status can be viewed as irreversible" and that it will communicate "its conclusions from the discussions with the investment community on the list of countries under review" every year. This within-country variation provides a conceptually different setup to explore the link between cross-sectional mispricing and market development.

For each of the three countries separately as well as for the pooled sample, I thus regress long/short mispricing on a developed market dummy that is one (zero) for the months in which the country under consideration was listed as a developed (emerging) market. I eliminate the month of the market reclassification to avoid confounding effects. To control for general time effects in the magnitude of mispricing, I subtract the average long/short mispricing return based on all available countries in a given month (excluding the country under consideration). In untabulated tests, I find that not subtracting this benchmark return, using the country composite return, or relying on the average emerging or developed market return does not change inferences.

I consider both raw returns and three-factor model adjusted abnormal returns. With respect to the latter, I implement a two-stage regression similar to Stambaugh, Yu, and Yuan (2012). In the first stage, I regress the monthly time-series of raw mispricing (adjusted for general time effects as explained above) over the whole sample period on a local Fama and French (1993) model; using a global model does not change insights. Three-factor model adjusted abnormal returns are then defined as the sum of the intercept and the fitted value of the residual. The second-stage regression is identical to the procedure for raw returns; that is, I regress the resulting time-series on a developed market dummy. As before, I report both equally weighted and value-weighted returns, which results in four regression specifications for each country.

The main findings are displayed in Table 5. The results show a pervasive picture. In all 12 specifications, the developed market dummy obtains a positive coefficient, which indicates greater mispricing after switching from being classified as an emerging market to being classified as a developed market. Country-level findings are economically meaningful, but in all but one case not significant. In the pooled country sample, estimates are statistically significant in four of the eight specifications, with the increase in mispricing ranging from 38 bp to 156 bp per month. In sum, the findings obtained from the within-country variation in market development support the results from the cross-country variation.

4. Biased beliefs as a driver of mispricing?

If the return predictability shown so far indeed represents mispricing caused by expectational errors, then anomaly spreads should be particularly large around earnings announcements. When investors have overly optimistic (pessimistic) expectations regarding the stocks in the short (long) leg of the Stambaugh, Yu, and Yuan (2015) strategy, then earnings news will force them to update their beliefs. As a consequence, the long/short portfolio should generate significantly positive abnormal returns during a narrow time window in which expected returns are close to zero. This line of reasoning underlies classical studies such as Bernard and Thomas (1990) and La Porta, Lakonishok, Shleifer, and Vishny (1997). More recently, similar tests have been employed by Edelen, Ince, and Kadlec (2016) and Engelberg, McLean, and Pontiff (2015).

To test for potential differences between developed markets and emerging markets, I gather earnings announcement dates from Worldscope (for international stock markets) and Compustat (for the U.S.) for firms that meet the criteria outlined in Section 2. If the announcement falls on a non-trading day, the date is set to the next trading day. I start by computing the abnormal announcement return for each event. To be consistent with the analysis in the previous sections, I define abnormal returns as the buy-and-hold return during the event days −1, 0, and 1 minus the expected buy-and-hold return implied by a local (Fama and French, 1993) factor model [see, for instance, DellaVigna and Pollet (2009) for a similar approach].[10] Factor loadings are estimated from rolling regressions of daily excess returns on the local market factor, the size factor, and the value factor during months $t-12$ to $t-2$.

In Panel A of Table 6, I regress the pooled abnormal announcement return on the mispricing score, standardized to be uniformly distributed over the interval (0,1]. This analysis is based on about 958,000 eligible earnings announcements, and performed separately for developed and emerging markets. Consistent with the idea of biased cashflow beliefs, the abnormal return for undervalued (overvalued) stocks is positive (negative) in both samples. In developed markets, the difference amounts to 99 bp (t-statistic 15.36), in emerging markets to 63 bp (t-statistic 11.66). The difference-in-differences (36 bp, t-statistic 4.27) indicates that market participants in developed markets are more surprised by information contained in the earnings announcements of mispriced stocks. Panel B, in which I rely on a dummy variable that is one (zero) for stocks in the short (long) leg of the mispricing portfolio, shows similar results.

As there could be cross-country differences in insider trading, information leakage, or pre-announcement speculation effects, Panel C of Table 6 shows findings from

[10] Inferences do not change if I use raw returns, if I rely on beta-adjusted abnormal returns, if I measure returns relative to a size and book-to-market matching portfolio, or if I benchmark announcement returns against the average return of the same firm during the non-announcement period. To mitigate the impact of outliers, abnormal returns are winsorized at the 0.1% and the 99.9% levels.

Table 5
Mispricing and market reclassification.
This table shows the impact of market reclassification on country-level mispricing, expressed in % per month. In Panel A, the dependent variable in the regressions is the time-series of country-level mispricing minus the time-series of average mispricing based on all other countries (excluding the country under consideration). In Panel B, I use a two-stage regression approach in order to additionally control for exposure to a local three-factor model. In both panels, the explanatory variable of interest is a *developed market dummy* that is one (zero) if a market is classified as a developed market (emerging market) in a given month. The month of the market reclassification is excluded from the analysis. *T*-statistics (in parentheses) for the country-level regressions are based on the heteroskedasticity-consistent standard errors of White (1980). *T*-statistics (in parentheses) for the pooled regressions are clustered by month. Two-tailed statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Country	Greece		Israel		Portugal		Pooled		Pooled	
Sample period	Jan 94 Dec 13	Jan 94 Dec 13	Aug 98 Dec 13	Aug 98 Dec 13	Jan 94 Dec 13	Jan 94 Dec 13	Jan 94 Dec 13	Jan 94 Dec 13	Jan 94 Dec 13	Jan 94 Dec 13
Return weighting scheme	value	equal	value	equal	value	equal	value	equal	value	equal
Panel A: Raw long/short returns										
Developed market dummy	2.369**	0.962	1.320	0.579	0.573	0.616	1.564**	0.759	1.112*	0.964**
	(2.02)	(1.04)	(0.89)	(0.86)	(0.53)	(0.75)	(2.31)	(1.50)	(1.78)	(2.11)
Constant	−1.241	−0.502	0.421	−0.225	0.012	0.483	−0.737	−0.376	−0.188	−0.196
	(−1.51)	(−0.64)	(0.70)	(−0.52)	(0.01)	(0.73)	(−1.06)	(−0.66)	(−0.43)	(−0.56)
Country fixed effects	No	No	No	No	No	No	Yes	Yes	No	No
Panel B: Three-factor alphas										
Developed market dummy	1.618	0.345	1.084	0.600	0.057	0.229	1.017	0.377	0.854	0.744*
	(1.43)	(0.38)	(0.77)	(0.92)	(0.05)	(0.27)	(1.53)	(0.75)	(1.46)	(1.71)
Constant	−0.320	0.142	0.595	−0.152	1.000	1.079	0.057	0.122	0.366	0.150
	(−0.36)	(0.17)	(1.00)	(−0.37)	(1.03)	(1.51)	(0.08)	(0.21)	(0.81)	(0.42)
Country fixed effects	No	No	No	No	No	No	Yes	Yes	No	No

a longer event window, which covers the days $t-10$ to $t+1$. Again, the abnormal return for presumably undervalued (overvalued) stocks is strongly positive (negative) in both samples. The return difference between developed and emerging markets shrinks to 10 bp, which is statistically insignificant. Nevertheless, there is again no evidence that biased beliefs are weaker in countries deemed to be more developed.

To further explore the idea that expectational errors could be the driver behind the return predictability, I implement a regression similar to the approach in Engelberg, McLean, and Pontiff (2015). The unit of observation now is a firm day. Separately for developed and emerging markets, I regress the pooled (Fama and French, 1993) model-adjusted daily stock returns on an earnings announcement window dummy, the mispricing measure, and the interaction term. Additionally, I include a set of controls consisting of daily stock returns as well as of squared daily returns (as a proxy for volatility) during days $t-10$ to $t-2$.

The biased expectations framework predicts a significant interaction term. Panel D of Table 6 shows the results, which confirm the insights of Engelberg, McLean, and Pontiff (2015). Taken together and for the case of developed (emerging) markets, the coefficients indicate that anomaly spreads are, all else equal, about 450% (200%) larger on a day of the earnings announcement window than on a regular day. Again, the highly significant difference is consistent with the idea of mispricing attributable to mistaken beliefs being particularly strong in developed markets. The positive coefficient for the earnings announcement dummy is consistent with Frazzini and Lamont (2007) and Barber, George, Lehavy, and Trueman (2013).

As a final test for the biased expectations hypothesis, I study the behavior of a specific group of important market participants. More precisely, I study whether the mispricing measure can systematically predict sell-side analysts' forecast errors. I gather analyst forecast data and earnings announcement dates from I/B/E/S, thereby excluding interim announcements due to limited data availability. I also drop firms with less than two analysts. In contrast to the unconditional sample, the final sample shows only small differences in the average number of analysts covering firms in developed markets (8.5) and emerging markets (7.3). I define the forecast error as the difference between the median analyst forecast and the actual reported earnings, scaled by the standard deviation of the forecast. In Panel E of Table 6, I regress the analyst forecast measure on the continuous version of the mispricing metric. Results show that analyst forecast errors are indeed reliably predictable in both samples. Analysts overestimate the earnings of overpriced stocks and, in developed markets only, underestimate the earnings of underpriced stocks. Most notably, the difference between the magnitude of this forecast bias in developed markets and the magnitude of the forecast bias in emerging markets is positive and statistically significant at the 1% level. The strong predictability of forecast errors in developed markets is also in line with the recent U.S. evidence of Engelberg, McLean, and Pontiff (2015).

Taken in their entirety, two conclusions can be drawn from the tests in this section. First, the findings are consistent with the hypothesis that the return predictive ability of the Stambaugh, Yu, and Yuan (2015) measure is, at least in part, caused by mispricing based on biased expectations. Second, the results are also in line with the idea that biased beliefs tend to be stronger in developed markets, which could in part explain the cross-country differences in average (Stambaugh, Yu, and Yuan, 2015) spreads.

Table 6
Biased beliefs and mispricing in developed vs. emerging markets.
This table shows the main insights obtained from regressions aimed at testing for biased beliefs as a driver of mispricing. The sample period is January 1994 to December 2013. The dependent variable in Panels A and B (C) is the pooled abnormal earnings announcement return, defined as the buy-and-hold return during the event days $t=-1$, $t=0$, and $t=+1$ (days $t=-10$ to $t=+1$), minus the expected buy-and-hold return as implied by a local (Fama and French, 1993) model. Abnormal returns are winsorized at the 0.1% level and at the 99.9% level. The dependent variable in Panel D is the pooled daily abnormal stock return (again winsorized and relative to a three-factor model). The explanatory variable in Panel A is the mispricing measure, which is uniformly distributed over the interval (0,1]. The explanatory variable in Panels B and C is a dummy variable which is one (zero) for stocks in the top (bottom) quintile of the mispricing measure. The explanatory variables in Panel D are the mispricing measure, an earnings announcement window dummy that is one for the days $t=-1$ to $t=+1$ and zero otherwise, the interaction effect, as well as (untabulated) controls for daily stock returns and squared daily stock returns for the days $t=-10$ to $t=-2$. Panel E explores the relationship between analyst forecast error and the mispricing measure. The forecast error is defined as the difference between the median analyst forecast and the actual reported earnings, scaled by the standard deviation of the forecast. Forecast errors are winsorized at the 5th and 95th percentile to mitigate the impact of outliers and to make the distribution closer to normal. In Panels A, B, C, E (D), standard errors are double-clustered by country and month (by firm and date). In all panels, corresponding *T*-statistics are reported in parentheses. Two-tailed statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Reaction around earnings announcements (days t−1 to t+1, mispricing metric (0,1))

	Developed markets		Emerging markets		Difference	
Mispricing	−0.992***	(−15.36)	−0.631***	(−11.66)	0.361***	(4.27)
Constant	0.589***	(17.93)	0.257***	(4.31)		
Observations	704,235		253,906		958,141	

Panel B: Reaction around earnings announcements (days t−1 to t+1, mispricing dummy)

	Developed markets		Emerging markets		Difference	
Mispricing	−0.829***	(−14.91)	−0.515***	(−11.39)	−0.314***	(−4.26)
Constant	0.405***	(12.17)	0.195***	(3.32)		
Observations	280,124		101,739		381,863	

Panel C: Reaction around earnings announcements (days t−10 to t+1, mispricing dummy)

	Developed markets		Emerging markets		Difference	
Mispricing	−1.501***	(−10.06)	−1.398***	(−14.84)	−0.103	(−0.59)
Constant	0.753***	(10.13)	0.647***	(5.20)		
Observations	280,124		101,739		381,863	

Panel D: Relative importance of earnings period (days t−1 to t+1, mispricing metric (0,1))

	Developed markets		Emerging markets		Difference	
Earnings period (EP)	0.173***	(19.28)	0.052***	(3.34)	0.121***	(6.98)
Mispricing	−0.057***	(−8.84)	−0.068***	(−15.73)	0.011	(1.55)
EP∗Mispricing	−0.258***	(−14.97)	−0.135***	(−7.96)	−0.123***	(−5.15)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	60,169,696		25,649,086		85,818,782	

Panel E: Analyst forecast error (mispricing metric (0,1))

	Developed markets		Emerging markets		Difference	
Mispricing	0.884***	(24.01)	0.661***	(11.36)	0.223***	(3.29)
Constant	−0.557***	(−16.67)	−0.006	(−0.16)		
N	105,493		23,907		129,400	

5. Further potential determinants of mispricing

5.1. Firm characteristics

I run panel regressions with the country-year average of the monthly local three-factor alpha as dependent variable and country-year averages of firm-level characteristics as independent variables. I start with common proxies for market frictions and limits to arbitrage. More specifically, I compute average firm size as well as average firm-level idiosyncratic volatility, defined as the standard deviation of the residual obtained from rolling regressions of daily excess returns on a local (Fama and French, 1993) model over the previous 12 months. In addition, I compute the average number of analysts providing firm-level fiscal-year-one earnings estimates. Finally, I follow Bris, Goetzmann, and Zhu (2007) and Griffin, Nardari, and Stulz (2007) in constructing a dummy variable that takes a value of one if short sales are a common practice and zero otherwise.

Next, I consider average analyst forecast dispersion, defined as the standard deviation of the forecasts scaled by the absolute value of the mean forecast, as a proxy for information uncertainty and differences of opinion. Previous work including Diether, Malloy, and Scherbina (2002) or Stambaugh, Yu, and Yuan (2012) suggests that higher dispersion of beliefs could go along with stronger mispricing.

Finally, I consider two variables for which there are competing hypotheses with respect to market efficiency, as discussed in Barberis and Thaler (2003) or Shiller (1981). First, I use the average R^2 obtained from the regressions employed to quantify idiosyncratic volatility. On the one hand, high firm-specific return variations could be interpreted as informed investors trading on valuable information. In this regard, and as argued in Morck, Yeung, and Yu (2000), lower return R^2 could reflect a higher degree of market efficiency. On the other hand, high return variations could proxy for noise trading. In this regard, and as Hou, Peng, and Xiong (2013, pp. 1–2) point out, “lower return R^2 may actually capture market inefficiency rather than efficiency." Hou, Peng, and Xiong (2013) provide theoretical and empirical support for their conjecture. Further discussions are provided in Chan and Hameed (2006), Kelly (2014), and Li, Rajgopal, and Venkatachalam (2014).

Second, I consider the average fraction of days with zero return as a proxy for trading activity. On the one hand, higher trading activity can lower transaction costs or represent arbitrage trading, which could result in less mispricing. Supportive evidence is presented in, for instance, McLean and Pontiff (2016). On the other hand, high trading activity could reflect the actions of noise traders, making it mainly a proxy for sentiment. Support for this conjecture is presented in, for instance, Baker and Stein (2004), Baker and Wurgler (2006), Daniel, Hirshleifer, and Subrahmanyam (1998), or DeVault, Sias, and Starks (2015). Thus, higher trading activity could result in more mispricing.

In all panels of Table 7, the sample period is 1994–2013. Explanatory variables are standardized. Following (Petersen, 2009), standard errors are double-clustered by country and year. The univariate findings in Panel A show that stocks in developed markets are significantly larger, have lower return R^2, and more analyst coverage. In addition, short selling is more widespread. In contrast, with respect to idiosyncratic volatility, trading activity, and analyst forecast dispersion, there are no reliable differences.

In Panel B, I explore country-level determinants of equally weighted or value-weighted mispricing. The first two regressions largely reject the idea that limits to arbitrage drive cross-country differences in alphas, which is consistent with Sun, Wei, and Xie (2015), Titman, Wei, and Xie (2013), or Watanabe, Xu, Yao, and Yu (2013). The coefficients on firm size, idiosyncratic volatility, and the short selling dummy are statistically insignificant and tend to be economically small. Analyst coverage is (as expected) negatively related to mispricing, but this finding is significant for equally weighted returns only.

The strongest impact is found for return R^2. A one standard deviation decrease in R^2 is estimated to go along with a significant alpha increase of about 55 bp. Trading activity has a strong impact as well. A one standard deviation increase in trading activity (as proxied for by a one standard deviation decrease in the fraction of days with a zero return) is estimated to go along with an increase in the long/short return of about 40 bp. Finally, stronger analyst forecast dispersion is associated with larger mispricing.

In specifications 3 to 4, I add a developed market dummy. In specifications 5 and 6, I additionally include year dummies. The firm characteristics, in particular the return R^2, explain a sizeable fraction of the difference in mispricing between developed and emerging markets, especially with respect to equally weighted returns. The alpha difference in specifications 3 to 6 is still positive (7 bp to 41 bp), yet statistically insignificant (t-statistics 0.39–1.56). Running the same regressions without the stock characteristics yields (untabulated) differences between 45 bp and 55 bp (t-statistics 2.23–2.66).

The time-series variation in most firm characteristics presents an opportunity to explore determinants of within-country changes in mispricing beyond the analysis in Section 3.3. Specifications 7 and 8, which include country and time fixed effects, indicate that the cross-country results tend to carry over to the within-country perspective. Most notably, trading activity and R^2 are reliably positively related to mispricing. The coefficient for forecast dispersion has the predicted sign, yet is statistically insignificant. In sum, anomaly spreads, both across and within countries, appear to be positively related to firm characteristics that some researchers suggest proxy for a greater presence of noise trading.

Finally, in specifications 9 and 10, I extend the cross-sectional and time variation by additionally including MSCI frontier markets. Inferences do not change.

In tests untabulated for brevity, I have additionally experimented with country characteristics gathered from the Global Competitiveness Report (GCR) of the World Economic Forum. I have thereby concentrated on explanatory factors for which differences between emerging markets and developed markets can be expected (e.g., proxies for institutional quality, the availability and affordability of financial services, or the quality of infrastructure). None of these variables exhibits robust statistical relations in the multivariate panel regressions or affects the role of the firm characteristics displayed in Table 7.

5.2. Natural experiment: mandatory IFRS adoption

When including the GCR survey variable "strength of auditing and reporting standards (1–7)" in specifications 1 and 2 of Table 7, then the corresponding t-statistics are 0.62 and 0.16, respectively. To shed further light on the apparently weak role of disclosure quality for mispricing, I follow (Hung, Li, and Wang, 2015) and exploit the switch from local financial-reporting standards to the mandatory adoption of IFRS in a number of countries in 2005. Hung, Li, and Wang (2015, pp. 1242–1244) argue that the adoption represents "one of the biggest events in the history of financial reporting" and can be viewed as an information shock associated with "exogenous and unprecedented improvement in non-U.S. firms' financial-reporting quality." Using a difference-in-differences approach, they show that the post-earnings-announcement drift (but not momentum or short-term reversal) weakens significantly for treatment countries after switching to IFRS.

I take advantage of this setting to test whether within-country variation in disclosure quality affects the magnitude of mispricing. In this test, I do not directly compare developed with emerging markets, but instead focus on the impact of one important aspect—the information environment—in which these markets are commonly believed to differ.

I compare treatment countries that mandated IFRS adoption in 2005 with control countries that did not mandate adoption around this time. For each treatment country, I focus on the subset of firms that report under local financial-reporting standards during the preshock period (2003 and 2004) and under IFRS in the post-shock period (2006 and 2007). For each control country, I concentrate on firms that report under local financial-reporting standards in both the preshock period and the postshock period. For both the treatment and the control

Table 7
Firm characteristics and mispricing.

This table shows the main insights obtained from panel regressions aimed at exploring the link between country-year averages of firm-level characteristics and mispricing. The sample period is January 1994 to December 2013. Panel A shows univariate differences between developed markets and emerging markets with respect to characteristics based on all eligible stocks in a given country, as explained in Section 2 and Panel B of Table 1. The *developed market dummy* is one (zero) if a given country-year is classified as a developed (emerging) market. Panel B shows multivariate panel regressions. The dependent variable is the country-year average of the monthly local three-factor alpha. The construction of the explanatory variables is described in detail in the text. Alphas for a given country-year are either equally weighted (*equal*) or value-weighted (*value*). The firm universe in specifications 1–8 (9–10) is developed markets and emerging markets (additionally frontier markets). In all panels, all variables (except dummy variables) are standardized to have a mean of zero and a variance of one. In both panels, standard errors are double-clustered by country and year. Corresponding *T*-statistics are reported in parentheses. Two-tailed statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Univariate differences in (normalized) firm characteristics between developed markets and emerging markets

Variable (1994–2013)	Firm size	R^2	Idiosyncratic volatility	Liquidity	Number of analysts	Analyst forecast dispersion	Short selling dummy
Developed market dummy	0.794***	−0.691***	0.004	−0.277	0.871***	−0.065	0.610***
	(4.17)	(−2.84)	(0.02)	(−1.06)	(5.40)	(−0.35)	(5.39)
Constant	−0.349**	0.366	0.010	0.081	−0.324*	0.0820	0.220**
	(−2.08)	(1.64)	(0.06)	(0.35)	(−1.80)	(0.46)	(2.46)
Observations	802	802	802	802	802	802	802
R^2	0.18	0.12	0.00	0.02	0.22	0.00	0.37

Panel B: Multivariate regressions with Stambaugh, Yu, and Yuan (2015) mispricing as dependent variable

Regression specification	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Return weighting	equal	value	equal	value	equal	value	equal	value	equal	value
Sample period	1994–2013	1994–2013	1994–2013	1994–2013	1994–2013	1994–2013	1994–2013	1994–2013	1994–2013	1994–2013
Firm universe	DM/EM	DM/EM	DM/EM	DM/EM	DM/EM	DM/EM	DM/EM	DM/EM	DM/EM/FM	DM/EM/FM
Firm size	−0.102	−0.0952	−0.118	−0.129	−0.286**	−0.161	−0.321	−0.059	−0.059	−0.048
	(−0.76)	(−0.72)	(−0.83)	(−0.98)	(−2.39)	(−1.04)	(−1.29)	(−0.20)	(−0.45)	(−0.36)
Return R^2	−0.569***	−0.547***	−0.547***	−0.498**	−0.516***	−0.424**	−0.591***	−0.458**	−0.512***	−0.434**
	(−4.01)	(−2.69)	(−3.80)	(−2.50)	(−3.53)	(−2.15)	(−3.00)	(−2.06)	(−3.98)	(−2.44)
Idiosyncratic volatility	−0.088	0.048	−0.087	0.049	−0.222**	0.037	−0.294**	0.004	−0.077	0.039
	(−1.01)	(0.31)	(−0.99)	(0.31)	(−1.98)	(0.22)	(−2.11)	(0.02)	(−0.89)	(0.30)
Fraction zero return days	−0.416***	−0.424**	−0.403***	−0.396**	−0.358***	−0.342**	−0.603***	−0.578*	−0.380***	−0.360**
	(−3.69)	(−2.35)	(−3.60)	(−2.27)	(−3.23)	(−2.04)	(−4.63)	(−1.92)	(−3.61)	(−2.10)
Number of analysts	−0.234*	−0.149	−0.264**	−0.212	0.011	−0.121	−0.278	−0.315	−0.197	−0.160
	(−1.90)	(−1.01)	(−2.04)	(−1.28)	(0.08)	(−0.58)	(−1.40)	(−0.87)	(−1.57)	(−1.06)
Analysts forecast dispersion	0.243***	0.200**	0.248***	0.210**	0.134*	0.156*	0.152	0.092	0.188**	0.091
	(2.75)	(2.06)	(2.82)	(2.18)	(1.73)	(1.71)	(1.48)	(0.61)	(2.13)	(0.98)
Short selling dummy	0.252	0.117	0.185	−0.026	0.244*	0.011	0.127	0.570	0.283	0.170
	(1.29)	(0.51)	(1.15)	(−0.12)	(1.73)	(0.05)	(0.17)	(0.47)	(1.60)	(0.79)
Developed market dummy			0.194	0.412	0.074	0.402				
			(0.96)	(1.56)	(0.39)	(1.44)				
Constant	0.966***	0.746***	0.903***	0.612***	−0.922***	−0.517***	−0.721***	0.496*	0.927***	0.747***
	(5.91)	(3.99)	(4.54)	(2.70)	(−5.66)	(−4.32)	(−3.92)	(1.68)	(6.19)	(4.19)
Time fixed effects	No	No	No	No	Yes	Yes	Yes	Yes	No	No
Country fixed effects	No	No	No	No	No	No	Yes	Yes	No	No
Observations	802	802	802	802	802	802	802	802	896	896
R^2	0.10	0.06	0.10	0.06	0.23	0.11	0.30	0.20	0.09	0.04

groups, I require a firm to have eligible data on reporting standards and aggregate mispricing at least once in both the preshock period and the postshock period. In addition, I only consider countries for which there are at least 25 eligible firms in each month of the 4-year test period defined as above. This procedure leads to 16 treatment and 12 control countries, which are displayed in Panel A of Table 8. For all countries, I compute three-factor model-adjusted monthly long/short returns (between 2003 and 2007) as before. To facilitate presentation, Table 8 reports results for abnormal returns only; using raw return does not change insights.

Panel A of Table 8 also shows the main results from country-level regressions in which I regress the mispricing measure on a dummy variable which is zero for the years 2003 and 2004 (the preshock period) and one for the years 2005 and 2006 (the postshock period). Findings do not indicate a pronounced difference between treatment and control countries. With respect to equally weighted returns, 11 of the 16 treatment countries, but also six of the 12 control countries obtain a negative coefficient on the information shock dummy. With respect to value-weighted returns, five of the 16 treatment countries, but also five of the 12 control countries obtain a negative coefficient. Only a few coefficients are significant. In sum, there are no clear patterns.

To reduce the noise of individual country estimates, Panels B to E focus on more aggregate comparisons. In Panel B (C), I pool the monthly country-level equally weighted (value-weighted) alphas. Explanatory variables

Table 8
Mispricing and mandatory IFRS adoption.

The table compares mispricing in treatment countries that mandated IFRS adoption in 2005 with control countries that did not mandate adoption around this time. For each treatment country, I focus on the subset of firms that report under local financial-reporting standards during the preshock period (2003 and 2004) and under IFRS in the postshock period (2006 and 2007). The *event dummy* is one (zero) in the postshock (preshock) period. For each control country, I concentrate on firms that report under local financial-reporting standards in both the preshock period and the postshock period. The dependent variable in Panel A (Panels B and C) is the (pooled) monthly country-level abnormal return relative to a local three-factor model. The dependent variable in Panels D and E is the monthly abnormal return relative to a global three-factor model, averaged separately for treatment and control countries. In Panels A, D, and E (B and C), standard errors are computed as in White (1980) (double-clustered by country and month). *T*-statistics are reported in parentheses. Two-tailed statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***, respectively.

Panel A: Individual country regression

Treatment group					Control group				
Return	equal Event dummy		value Event dummy		Return	equal Event dummy		value Event dummy	
Australia	−0.868	(−1.27)	−0.846	(−1.11)	Brazil	−1.735	(−0.60)	0.920	(0.50)
Belgium	−0.176	(−0.19)	0.873	(0.73)	Canada	0.0339	(0.05)	0.867	(0.90)
Denmark	−1.868**	(−2.38)	−1.264	(−1.00)	China	−0.246	(−0.42)	−0.564	(−0.86)
Finland	−0.099	(−0.12)	0.460	(0.46)	India	−1.051*	(−1.72)	−0.979	(−0.71)
France	−0.445	(−0.79)	0.528	(0.64)	Indonesia	0.144	(0.15)	−0.018	(−0.01)
Germany	−0.393	(−0.53)	−2.629*	(−2.01)	Japan	0.0704	(0.13)	0.484	(0.76)
Greece	0.182	(0.23)	0.377	(0.38)	Malaysia	−0.357	(−0.88)	0.601	(1.18)
Italy	−0.955*	(−2.01)	−0.390	(−0.46)	Mexico	0.388	(0.46)	1.974**	(2.07)
Netherlands	−0.188	(−0.21)	0.460	(0.41)	Korea	−0.0198	(−0.03)	−0.432	(−0.44)
Norway	0.159	(0.16)	2.512	(1.63)	Taiwan	1.250**	(2.30)	0.770	(0.99)
Philippines	−1.310	(−0.96)	0.535	(0.43)	Thailand	−0.393	(−0.67)	−0.274	(−0.27)
South Africa	−0.333	(−0.40)	0.530	(0.75)	USA	0.0464	(0.10)	0.491	(1.12)
Spain	0.781	(1.51)	1.127	(1.35)					
Sweden	−0.313	(−0.42)	−2.784**	(−2.29)					
Switzerland	1.186	(1.61)	0.087	(0.07)					
UK	0.006	(0.02)	0.297	(0.48)					

Panel B: Pooled country-level regression, equally weighted returns

N	Intercept		Treatment		Shock		Treatment*shock	
1,344	0.427**	(2.33)	0.586***	(3.05)	−0.156	(−0.95)	−0.134	(−0.64)

Panel C: Pooled country-level regression, value-weighted returns

N	Intercept		Treatment		Shock		Treatment*shock	
1,344	0.216*	(1.67)	0.591*	(1.74)	0.320	(1.47)	−0.328	(−0.72)

Panel D: Country-average regression, equally weighted returns

N	Intercept		Treatment		Shock		Treatment*shock	
96	0.839***	(3.33)	0.406	(1.12)	−0.279	(−0.90)	−0.075	(−0.18)

Panel E: Country-average regression, value-weighted returns

N	Intercept		Treatment		Shock		Treatment*shock	
96	0.608**	(2.58)	0.472	(1.23)	0.088	(0.28)	−0.245	(−0.48)

consist of the event dummy variable, a treatment dummy variable, and the interaction term, which yields the coefficient of interest. Equally weighted (value-weighted) returns suggest that, relative to control countries, treatment countries experience a 13 bp (33 bp) stronger drop in postshock abnormal returns. However, as already suggested by the widely fluctuating individual country results in Panel A, these estimates are not statistically significant (*t*-statistics 0.64 and 0.72, respectively). Panels D and E display the findings obtained from averaging the mispricing measure for each month and country group separately. Differences between treatment and control countries become slightly weaker.

In sum, there is little evidence that the information shock as implied by the introduction of IFRS adoption yields reliably lower mispricings. In a broader sense, these results could suggest that aggregate mispricing, as defined in this paper, is hardly affected by disclosure quality. As a consequence, differences in average disclosure quality between emerging markets and developed markets may not necessarily imply strong differences in mispricing.

6. Conclusion

Based on the Stambaugh, Yu, and Yuan (2015) mispricing score, a comprehensive international data set, and conceptually diverse tests, my findings cast doubt on the notion that the markets outside of the most developed ones are necessarily less efficient.

These findings suggest several directions for further research. First, there could be forms of mispricing that are not reflected in the Stambaugh, Yu, and Yuan (2015) score and that are particularly strong in emerging markets. Relatedly, the composite mispricing measure is based on public information only, and thus does not speak to the strong form of market efficiency. Second, the measure is purely cross-sectional and thus does not allow to draw inferences about market-wide overpricing or underpricing. Third, most of the large cross-country variation in return predictability, as indicated in Fig. 1, is currently unexplained. Fourth, it is an open question as to what extent institutional investors' sentiment-induced demand shocks, investment constraints, or agency conflicts in the sense of DeVault, Sias, and Starks (2015), Edelen, Ince, and Kadlec (2016), or Lakonishok, Shleifer, and Vishny (1994) contribute to mispricing in developed markets.

References

Akbas, F., Armstrong, W.J., Sorescu, S., Subrahmanyam, A., 2015. Smart money, dumb money, and capital market anomalies. Journal of Financial Economics 118, 355–382.

Baker, M., Stein, J., 2004. Market liquidity as a sentiment indicator. Journal of Financial Markets 7, 271–299.

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

Barber, B.M., George, E.D., Lehavy, R., Trueman, B., 2013. The earnings announcement premium around the globe. Journal of Financial Economics 108, 118–138.

Barberis, N., Thaler, R., 2003. A survey of behavioral finance. In: Harris, G.M., Stulz, M.R. (Eds.), Constantinides, Handbook of the Economics of Finance. Elsevier Science B.V., Amsterdam, pp. 1051–1121.

Bekaert, G., Harvey, C.R., 2002. Research in emerging markets finance: looking to the future. Emerging Markets Review 3, 429–448.

Bernard, V.L., Thomas, J.K., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics 13, 305–340.

Bhattacharya, U., Daouk, H., Jorgenson, B., Kehr, C.H., 2000. When an event is not an event: the curious case of an emerging market. Journal of Financial Economics 55, 69–101.

Bris, A., Goetzmann, W.N., Zhu, N., 2007. Efficiency and the bear: short sales and markets around the world. Journal of Finance 62, 1029–1079.

Busse, J.A., Goyal, A., Wahal, S., 2014. Investing in a global world. Review of Finance 18, 561–590.

Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. Journal of Finance 63, 2899–2939.

Chan, K., Hameed, A., 2006. Stock price synchronicity and analyst coverage in emerging markets. Journal of Financial Economics 80, 115–147.

Chen, L., Novy-Marx, R., Zhang, L., 2011. An alternative three-factor model. Unpublished working paper. Washington University, University of Rochester, Ohio State University.

Chui, A.C., Titman, S., Wei, K.J., 2010. Individualism and momentum around the world. Journal of Finance 65, 361–392.

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

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53, 1839–1885.

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

DellaVigna, S., Pollet, J.M., 2009. Investor inattention and Friday earnings announcements. Journal of Finance 64, 709–749.

DeVault, L.A., Sias, R.W., Starks, L.T., 2015. Who are the sentiment traders? Evidence from the cross-section of stock returns and demand. Unpublished working paper. University of Arizona, University of Texas at Austin.

Diether, K.B., Malloy, C.J., Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. Journal of Finance 57, 2113–2141.

Dyck, A., Lins, K.V., Pomorski, L., 2013. Does active management pay? New international evidence. Review of Asset Pricing Studies 3, 200–228.

Edelen, R., Ince, O., Kadlec, G.B., 2016. Institutional investors and stock return anomalies. Journal of Financial Economics 119, 472–488.

Eisdorfer, A., Goyal, A., Zhdanov, A., 2014. Distress anomaly and shareholder risk: international evidence. Unpublished working paper. University of Connecticut, University of Lausanne, Pennsylvania State University.

Eling, M., Faust, R., 2010. The performance of hedge funds and mutual funds in emerging markets. Journal of Banking and Finance 34, 1993–2009.

Engelberg, J., McLean, R.D., Pontiff, J., 2015. Anomalies and news. Unpublished working paper. University of California at San Diego, University of Alberta, Boston College.

Fama, E.F., 1991. Efficient capital markets. Journal of Finance 46, 1575–1617.

Fama, E.F., French, K., 2012. Size, value, and momentum in international stock returns. Journal of Financial Economics 105, 457–472.

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

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

Ferreira, M.A., Keswani, A., Miguel, A., Ramos, S.B., 2013. The determinants of mutual fund performance: a cross-country study. Review of Finance 17, 483–525.

Fidelity, 2014. A case for active investment in emerging markets. Whitepaper, November 2014.

Frazzini, A., Lamont, O.A., 2007. The earnings announcement premium and trading volume. NBER Working Paper No. 13090

Green, J., Hand, J.R., Zhang, X.F., 2013. The supraview of return predictive signals. Review of Accounting Studies 18, 692–730.

Green, J., Hand, J.R., Zhang, X.F., 2014. The remarkable multidimensionality in the cross-section of expected U.S. stock returns. Unpublished working paper. Pennsylvania State University, University of North Carolina, Yale University.

Griffin, J.M., 2002. Are the Fama and French factors global or country specific? Review of Financial Studies 15, 783–803.

Griffin, J.M., Hirschey, N.H., Kelly, P.J., 2011. How important is the financial media in global markets? Review of Financial Studies 24, 3941–3992.

Griffin, J.M., Kelly, P.J., Nardari, F., 2010. Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets. Review of Financial Studies 23, 3225–3277.

Griffin, J.M., Nardari, F., Stulz, R.M., 2007. Do investors trade more when stocks have performed well? Evidence from 46 countries. Review of Financial Studies 20, 905–951.

Harvey, C.R., Liu, Y., Zhu, H., 2016. ...and the cross-section of expected returns. Review of Financial Studies 29, 5–68.

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

Hou, K., Karolyi, G.A., Kho, B.-C., 2011. What factors drive global stock returns? Review of Financial Studies 24, 2527–2574.

Hou, K., Peng, L., Xiong, W., 2013. Is R^2 a measure of market inefficiency?. Ohio State University, City University of New York, Princeton University. Unpublished working paper

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: an investment approach. Review of Financial Studies 28, 650–705.

Huij, J., Post, T., 2011. On the performance of emerging market equity mutual funds. Emerging Markets Review 12, 238–249.

Hung, M., Li, X., Wang, S., 2015. Post-earnings-announcement drift in global markets: evidence from an information shock. Review of Financial Studies 28, 1242–1283.

Ince, O.S., Porter, R.B., 2006. Individual equity return data from Thomson Datastream: handle with care!. Journal of Financial Research 29, 463–479.

Jacobs, H., 2015. What explains the dynamics of 100 anomalies? Journal of Banking and Finance 57, 65–85.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.

Kang, X., Nielsen, F., Fachinotti, G., 2011. Some like it hot. MSCI Barra Research Paper No. 2011-03.

Kelly, P.J., 2014. Information efficiency and firm-specific return variation. Quarterly Journal of Finance 4, 1450018.

Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–1578.

La Porta, R., Lakonishok, J., Shleifer, A., Vishny, R.W., 1997. Good news for value stocks: further evidence on market efficiency. Journal of Finance 52, 859–874.

Lesmond, D.A., Ogden, J.P., Trzcinka, C.A., 1999. A new estimate of transaction costs. Review of Financial Studies 12, 1113–1141.

Li, B., Rajgopal, S., Venkatachalam, M., 2014. R^2 and idiosyncratic risk are not interchangeable. Accounting Review 89, 2261–2295.

Loughran, T., Ritter, J.R., 1995. The new issues puzzle. Journal of Finance 50, 23–51.

McLean, R.D., Pontiff, J., 2016. Does academic research destroy stock return predictability? Journal of Finance 71, 5–32.

McLean, R.D., Pontiff, J., Watanabe, A., 2009. Share issuance and cross-sectional returns: international evidence. Journal of Financial Economics 94, 1–17.

M&G Investments, 2015. Emerging markets in focus. Global Emerging Markets Fund brochure.
Morck, R., Yeung, B., Yu, W., 2000. The information content of stock markets: why do emerging markets have synchronous stock price movements? Journal of Financial Economics 58, 215–260.
MSCI, 2014. MSCI market classification framework June 2014.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.
Novy-Marx, R., 2013. The other side of value: the gross profitability premium. Journal of Financial Economics 108, 1–28.
Ohlson, J.A., 1980. Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research 18, 109–131.
Petersen, M.A., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Review of Financial Studies 22, 435–480.
Pontiff, J., 2006. Costly arbitrage and the myth of idiosyncratic risk. Journal of Accounting and Economics 42, 35–52.
Richardson, S., Tuna, I., Wysocki, P., 2010. Accounting anomalies and fundamental analysis: a review of recent research advances. Journal of Accounting and Economics 50, 410–454.
Ritter, J.R., 1991. The long-run performance of initial public offerings. Journal of Finance 46, 3–27.
Rouwenhorst, K.G., 1998. International momentum strategies. Journal of Finance 53, 267–284.
Rouwenhorst, K.G., 1999. Local return factors and turnover in emerging stock markets. Journal of Finance 54, 1439–1464.
Schwert, G.W., 2003. Anomalies and market efficiency. In: Harris, G.M., Stulz, M.R. (Eds.), Constantinides, Handbook of the Economics of Finance. Elsevier Science B.V., Amsterdam, pp. 939–974.
Shiller, R.J., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? American Economic Review 71, 421–436.
Sloan, R.G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review 71, 289–315.
Stambaugh, R.F., Yu, J., Yuan, Y., 2012. The short of it: investor sentiment and anomalies. Journal of Financial Economics 104, 288–302.
Stambaugh, R.F., Yu, J., Yuan, Y., 2014. The long of it: odds that investor sentiment spuriously predicts anomaly returns. Journal of Financial Economics 114, 613–619.
Stambaugh, R.F., Yu, J., Yuan, Y., 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. Journal of Finance 70, 1903–1948.
Subrahmanyam, A., 2010. The cross-section of expected stock returns: what have we learnt from the past twenty-five years of research? European Financial Management 16, 27–42.
Sun, L., Wei, K.C.J., Xie, F., 2015. On the explanations for the gross profitability effect. Shanghai University, Hong Kong University, University of Texas. Unpublished working paper
Titman, S., Wei, K.C.J., Xie, F., 2013. Market development and the asset growth effect: international evidence. Journal of Financial and Quantitative Analysis 48, 1405–1432.
Titman, S., Wei, K.J., Xie, F., 2004. Capital investments and stock returns. Journal of Financial and Quantitative Analysis 39, 677–700.
Watanabe, A., Xu, Y., Yao, T., Yu, T., 2013. The asset growth effect: insights from international equity markets. Journal of Financial Economics 108, 529–563.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.
Xing, Y., 2008. Interpreting the value effect through the q-theory: an empirical investigation. Review of Financial Studies 21, 1767–1795.