The Global Relation Between Financial Distress and Equity Returns*

Pengjie Gao

Christopher A. Parsons

Jianfeng Shen‡

This Version: September 27, 2016

Abstract

This study explores the distress risk anomaly — the tendency for stocks with high credit risk to perform poorly — among 38 countries over two decades. We find a strong, negative link between default probabilities and equity returns, concentrated among low-capitalization stocks in developed countries in North America and Europe. Although risk-based explanations provide a poor account of these patterns, several pieces of evidence point to a behavioral interpretation, suggesting that stocks of firms in financial distress are temporarily overpriced.

* We have benefited from discussions with Nick Barberis, John Campbell, Joey Engelberg, Ravi Jagannathan, Andrew Karolyi, David Ng, Terry Odean, Sheridan Titman, and Hong Yan. We thank participants at seminars of University of Notre Dame, Queensland University of Technology, and Massey University, 6th McGill Global Asset Management Conference, and 40th European Finance Association annual meeting for helpful comments. Karina Wang and Jimmy Zhu provided superb research assistance. We thank Moody's-KMV for providing us the Global Expected Default Frequency (*EDF®*) data. We are also grateful to John Wei, who provided us the individualism measure, and to Andrei Shleifer, who makes the creditor right index data available on his webpage. Part of the research was conducted during Jianfeng Shen's visit to the University of Notre Dame. Jianfeng Shen also acknowledges the funding support from UNSW Business School Research Grant scheme, and the assistance with RavenPack news data access from Malcolm Bain and Bohui Zhang.

‡ Contact: Pengjie Gao, Mendoza College of Business, University of Notre Dame, (Email) pgao@nd.edu (Tel) (1) 574-631-8048. Christopher A. Parsons, Rady School of Management, University of California at San Diego, (Email) caparsons@ucsd.edu (Tel) (1) 858-534-8782. Jianfeng Shen, School of Banking and Finance, UNSW Business School, UNSW Australia, (Email) jianfeng.shen@unsw.edu.au (Tel) (61)2-9385-4581.

I. Introduction

One of the most puzzling asset pricing irregularities is the abysmal stock performance of firms with high credit risk. The opposite result – though still constituting an anomaly – would be less surprising, since historically estimated betas might underestimate distress risk that occurs suddenly (Davydenko (2011)). Instead, when researchers have sorted firms by default risk, they find that high failure probabilities forecast low returns (Dichev (1998)). Deepening the puzzle, standard risk adjustments only strengthen the effect (Griffin and Lemmon (2002)).

Given the challenge these findings pose for rational models, the *distress risk anomaly* has attracted considerable attention in both empirical and theoretical studies. The literature is currently divided about the basic finding, as well as the correct interpretation.

There are three prevailing positions. The first view is especially straightforward: there is nothing to debate. Chava and Purnanandam (2010) argue that a wave of unexpected bankruptcies in a single country (the U.S.), and at a single point in time (the mid-1980s), is responsible for the poor returns of financially distressed firms.[1] Accordingly, when they look prior to 1980, they find no evidence that firms with poor credit risk underperform their risk benchmarks.

A second position grants the existence of the basic patterns, but on theoretical grounds, argues against it being interpreted as an anomaly. The *shareholder advantage* theory of Garlappi, Shu, and Yan (2008) and Garlappi and Yan (2011), for example, argues that if debt holders are expropriated in bankruptcy, equity holders can secure a relatively safe payoff in liquidation. As default becomes virtually certain, equity risk can actually decline. Other models capable of delivering a negative relation between credit and equity risk include George and Hwang (2010),

[1] See also Hackbarth, Haselmann and Schoenherr (2015) for a related discussion.

Johnson, Chebonenko, Cunha, D'Almeida, and Spencer (2011), Friewald, Wagner, and Zechner (2014), and Bhamra and Shim (2015).

The final position, most recently advocated by Campbell, Hilscher, and Szilagyi (2008), is that the returns of stocks of distressed firms are, in fact, too low to be reconciled within a rational framework. These authors find that distressed stocks have higher market betas, standard deviations, and other measures of risk, and yet, produce very low returns. Moreover, that the anomaly is concentrated among small, illiquid stocks lends additional credence to a mispricing interpretation (e.g., Shleifer and Vishny (1997)).[2]

In this paper, we explore a new dataset, in hopes of making progress on all three aspects of the debate. We begin by re-examining the distress risk anomaly among 38 countries over two decades (January 1992 - June 2013). Our benchmark tests ask whether throughout the world, portfolios sorted on credit risk, measured using Moody-KMV's Expected Default Frequency (*EDF*) estimates, give differential returns, after controlling for traditional risk factors.

Starting with U.S. firms for comparison, we confirm prior work that finds especially bad stock performance for firms with high credit risk. Long-short portfolios involving the 10% most and 10% least distressed firms generate significant alphas of about -50 bps per month at a one-month holding period, but taper off with longer horizons. As in Campbell et al. (2008), the results are stronger for firms below the median sized firm in the NYSE.

Among the other 37 countries in our sample, we find that a similar long-short trading strategy involving all stocks gives little evidence of underperformance for distressed firms. However, among the smaller half of firms within each country, the

[2] Other related papers include Avramov, Chordia, Jostova and Philipov (2007, 2013), who link financial distress with momentum and other return anomalies. Additionally, Avramov, Chordia, Jostova and Philipov (2012) investigate whether credit risk is a systematically priced risk factor.

evidence lines up with that found in the U.S. The returns of small, distressed firms are worse by 35-50 basis points per month, and persist up through a year after portfolio formation. Importantly, these relative results are caused by the poor performance of distressed firms, rather than good performance by firms with the least credit risk.

This small-firm distress anomaly is highly concentrated in developed markets, particularly those in Western Europe. Of the 20 developed (non-U.S.) markets in our sample, the high-low *EDF* difference in one-month stock returns is individually significant in over half (11). Aggregating European developed countries together, a strong result is observed at all horizons. In contrast, among 17 emerging markets, only Chile and South Korea show evidence of the distress anomaly, and in a few other cases (e.g., South Africa and Argentina), small distressed stocks occasionally *outperform* their non-distressed counterparts.

The remainder of the paper attempts to better understand reasons why these patterns exist, both overall and in the cross-section. We begin with a consideration of risk-based alternatives. One challenge to rational models generally is that distressed firms appear *more* risky on nearly every observable dimension, on average being smaller, and with higher betas, leverage, and volatility. Other evidence is better suited to address a specific theory related to shareholder expropriation (Garlappi and Yan (2011)). If this explanation is correct, the underperformance of high credit risk stocks should be most pronounced in countries with weak creditor protection. Yet, when we test for this using the Creditor Rights Index (*CRI*) developed by La Porta, Lopes-de-Silanes, Shleifer and Vishny (1998), the results are disappointing.

With little support for risk-based explanations, we take a closer look at behavioral alternatives. One framework capable of delivering low returns for distressed firms relies on investor overconfidence, combined with limits to arbitrage. In such a model, bad news – sufficient to plunge the firm into financial distress – would be incorporated into prices with a delay, due to investors' reluctance to fully update their posterior beliefs about the firm's poor prospects. Limits to arbitrage permit such incorrect beliefs to be temporarily reflected in prices, with the eventual collapse to fundamental value generating low returns.

This hypothesis delivers three testable implications. First, the distress risk anomaly should strengthen with limits to arbitrage, which a cross-sectional cut on size already affirms (only small distressed firms produce low returns). Second, when investor overconfidence is high, the subsequent returns of distressed firms should be especially low. Finally, because underreaction predicts temporary mispricing, returns should be worst among distressed firms with recent bad news.

We explore two proxies for overconfidence: one cross-sectional and one dynamic. Our cross-sectional proxy is borrowed from Chui, Titman, and Wei (2010), which take cross-country differences in *individualism* (Hofstede (2001)) as a measure for overconfidence, and finds a positive relation with momentum. Adapting their approach to our setting, we find that the distress anomaly is much stronger in countries with highly individualistic cultures, and virtually non-existent among those with *collectivist* backgrounds.

The next test is less about *where* the distress anomaly is most pervasive, but *when*. Theoretical work by Gervais and Odean (2001) and Daniel, Hirshleifer, and Subramanyam (1998) motivates our second measure of overconfidence. Both papers shows that traders' successes disproportionately lead them to positively update their estimates of their abilities, modeled as the precision of their privately collected signals. Consequently, prices formed during bull markets, when recent (aggregate) successes are more common, may be more impacted by overconfidence.

Consistent with this reasoning, we find that the distress anomaly is concentrated in periods directly preceded by aggregate market gains. Further underperformance is observed when up markets (price formation period) are directly followed by down markets (return measurement period). Both effects are driven by stocks with high turnover, a measure of retail trader activity.[3]

The paper concludes by testing the third empirical prediction: if underreaction is responsible for the overpricing of distressed stocks, performance should be especially poor for firms having received bad news recently. We measure news coverage in two ways: 1) directly using a novel news dataset from RavenPack, and 2) indirectly with the time a firm has been in the highest *EDF* decile. Both cases indicate that the distress anomaly is stronger firms with recent bad news, versus those having flirted with bankruptcy for an extended period of time.

The remainder of the paper is organized as follows. Section II describes the data, after which we present our main findings in Section III. We document that in a worldwide sample, stocks of small companies – particularly those in developed markets – underperform their risk benchmarks. Sections IV and V consider rational and behavioral, explanations for these results. The paper then concludes in Section VI.

II. Data and variable construction

Stock return and accounting variables. Stock returns and accounting variables for 38 countries are obtained from the CRSP/Compustat North America merged database for U.S. stocks, Compustat North America for Canadian stocks, and

[3] See Odean (1998, 1999), Statman, Thorley and Vorkink (2006), and Griffin, Nardari and Stulz (2007).

Compustat Global for the remaining countries in our analysis.[4] Our sample is limited to common stocks, those that are the primary securities of their respective companies, and those traded on major stock exchanges.[5]

Because data quality is often problematic in international studies, we apply a number of filters and/or conditions to minimize the influence of noise in our estimation. First, we drop extreme return values, following the criteria in Ince and Porter (2006).[6] Second, following Hou, Karolyi, and Kho (2011), we require a stock to have a minimum of 12 monthly observations in our sample period to be included in the sample. Third and finally, we apply stock price and market capitalization filters. Within (outside) the U.S., we drop observations with the month-end closing stock price is less than five dollars (below the 5th percentile for that country-month). Likewise for size, firms below the 5th percentile within every country-month are excluded from the analysis.[7]

We also employ filters at the country level. A country-month is admitted to the dataset only if it includes at least 50 stocks with valid values for *EDF* and market capitalization, a condition that helps ensure well-behaved portfolios. Second, we retain only countries with at least two years of data up to the end of our sample period, June 2013. Finally, we require a country to have non-missing values for La

---

[4] Appendix I contains a comparison of Compustat Global database versus Datastream/WorldScope, based on the summary statistics presented in Karolyi and Wu (2012).

[5] For most countries, there is only one major exchange on which the majority of stocks in that country are listed, except for the following countries: Canada (Toronto Stock Exchange and TSX Ventures Exchange), China (Shanghai Stock Exchange and Shenzhen Stock Exchange), India (Bombay Stock Exchange and National Stock Exchange), Japan (Osaka Securities Exchange, Tokyo Stock Exchange, and JASDAQ), Russia (Moscow Interbank Currency Exchange (MICEX) and Russian Trading System (RTS), which were later merged to form Moscow Exchange), South Korea (Korea Stock Exchange and KOSDAQ, which were later merged to form Korea Exchange but remained as separate divisions), and U.S.(NYSE, AMEX and NASDAQ).

[6] We apply two filters. First, large returns reversed in the subsequent month are set to missing. Specifically, in any month *t*, if R_t or R_{t-1} is greater than 300%, and $(1 + R_t) \times (1 + R_{t-1}) - 1 < 50\%$, then both returns are set to missing. Second, we further drop monthly returns that are above (below) the 99.9th (0.1th) percentile value in each country.

[7] See Section V.D for a discussion of our main results when microcap stocks are included in our sample.

Porta, Lopes-de-Silanes, Shleifer and Vishny's (1998) creditor rights index and Hofstede's (2001) individualism index. Our final sample consists of 44,930 unique stocks and 4,295,651 stock-month observations from 38 countries, over the period January 1992 through June 2013.

The first three columns in Table 1 describe the sample starting date, the number of unique stocks, and the number of stock-month observations for each country within each region: North America (the U.S. and Canada), European Developed Markets, Japan, Asia-Pacific Developed Markets (excluding Japan), and Emerging Markets. Unsurprisingly, more developed economies contribute disproportionately to the sample, with the U.S. (22.33%), Japan (15.81%), and the U.K. (7.31%) collectively comprising about half the stock-month sample.

Credit risk. We were provided access to Moody's KMV's complete database of monthly Expected Default Frequencies, or *EDF*s, for the 38 countries in our sample. *EDF* data for non-U.S. countries are available starting in 1992. Moody's-KMV's *EDF* database includes over 100,000 firm-year observations, and is calibrated against more than 2,000 defaults. *EDF* is a widely used estimate of financial distress.[8]

To give a sense of how *EDF* values evolve over time, the solid lines in Figures 1.1 and 1.2 plot, respectively for U.S. and non-U.S. firms, the standardized time-series mean *EDF* values. For most of the 1990s, both the typical U.S. and non-U.S. firm had a one-year default probability in the neighborhood of one percent. However, credit-tightening events like the Russian Default (1998), dot-com bust (2000/2001), and Financial Crisis of 2008 had a disproportionate impact outside the U.S., particularly in emerging countries.

---

[8] The *EDF* is an estimate of the physical probability of default for a given firm, based initially on Merton's (1974) model of credit risk, but with several additional features to account for the complexity of capital structures such as short-term liabilities, long-term liabilities, convertible debt, preferred shares, and common shares. For an overview of the EDF credit measure, see Crosbie and Bohn (2003). See also Bharath and Shumway (2008), Campbell, Hilscher and Szilagyi (2008), and Correia, Richardson, and Tuna (2012) for comparisons between structural models (Merton's (1974) model and its variants) and reduced-form models based on ad hoc accounting and market variables for predicting defaults.

More detail can be inferred from the fifth column of Table 1, which shows large variation in *EDF* values across countries. Emerging countries such as Indonesia (2.37%) and Pakistan (2.17%) having the highest failure probabilities, while at the other extreme, Switzerland has the lowest median *EDF* at 0.12%, slightly lower than those in the United States (0.32%), United Kingdom (0.24%), and Japan (0.57%). In part, these differences reflect life-cycle effects, with younger firms (more prevalent in developing economies) having greater return volatility. However, they also reflect cross-country differences in tax laws and attitudes towards leverage (Fan, Titman and Twite (2012)).

Table 1 also present country-specific summary statistics for other key firm characteristics including market capitalization (in million USD), book-to-market (*B/M*) ratio, cash flow-to-price (*C/P*) ratio, past 11-month returns (in USD), and monthly share turnover by country. Finally, note that because delisting returns are not available in Compustat, the returns of defaulting firms outside U.S. are biased upward (toward zero). This implies that the true realized performance from holding a portfolio of near-bankrupt firms would be worse than we estimate.

III. Is there a distress anomaly outside the U.S.?

Perhaps the most fundamental criticism of the distress anomaly is that is spurious, existing only within the U.S., and even then, only for relatively brief time periods (Chava and Purnanandam (2010)). Accordingly, we begin our analysis by re-examining the stock returns of high credit risk firms, but outside the U.S.

Following Campbell, Hilscher, and Szilagyi (2008), we characterize the distress anomaly with a long-short trading strategy, forming hedged portfolios long on stocks in the highest *EDF* decile and short on those in the lowest decile. At the end every month t, we form country-neutral, *EDF*-risk portfolios based on each firm's ranking within its own country. Then, we aggregate all firms within a given *EDF* decile in every month across the world. For example, in March 2003, a French firm ranking in the 4^{th} bankruptcy risk decile, relative to other French firms, would be grouped with a South African firm that also ranked in the 4^{th} decile within South Africa. This country-neutral methodology ensures that none of our results are capturing average return differences between, say, firms in developed versus developing countries.

Denoting *t* the month of portfolio formation, all tests skip month *t+1*, in order to mitigate the effects of microstructure noise and extreme return reversal (Da and Gao (2010)).[9] Our tests focus on returns over one (month *t+2*), three (*t+2* through *t+4*), and twelve (*t+2* through *t+13*) month horizons. For the latter two, we compute monthly returns based on the overlapping portfolio approach of Jegadeesh and Titman (1993).

Stocks differing in distress risk may also differ in other important ways, particularly their exposure to traditional risk factors, making it important to account for these differences in our analysis. Following Hou, Karolyi, and Kho (2011) and Karolyi and Wu (2012), which evaluate regional and country-level portfolio returns, we construct risk factors that permit flexibility in the degree to which local markets are segmented. Operationally, this involves including factors constructed at both the local and global level. In the polar case of full integration, only exposure to global risk factors will matter; in the opposite case of full segmentation, only exposure to local (say, country-level) risk factors summarizes expected returns. Including both

[9] The results in Vassalou and Xing (2004) are an example of these microstructure effects. These authors measure returns in the month following immediately portfolio formation, and find a positive relation with *EDF*. However, Da and Gao (2010) show that this is due to short-term reversals (Jegadeesh (1990), Lehmann (1990)), as most stocks with high default probabilities have experienced poor recent stock returns. Accounting for this effect results in a negative relation between credit risk and returns, both in the month following portfolio formation and thereafter.

permits partial segmentation which, given the considerable heterogeneity in development across the countries in our sample, seems appropriate.

Our basic specifications are the following:

$$\begin{aligned} R_{i,t} - R_{RF,t} = \alpha_i + \beta_i^L\left(R_{m,t}^L - R_{RF,t}\right) + \beta_i^G\left(R_{m,t}^G - R_{RF,t}\right) + c_i^L F_{C/P,t}^L + c_i^G F_{C/P,t}^G \\ + m_i^L F_{Mom,t}^L + m_i^G F_{Mom,t}^G + s_i^L F_{Size,t}^L + s_i^G F_{Size,t}^G + \varepsilon_{i,t} \end{aligned} \tag{1}$$

$$\begin{aligned} R_{i,t} - R_{RF,t} = \alpha_i + \beta_i^L\left(R_{m,t}^L - R_{RF,t}\right) + \beta_i^G\left(R_{m,t}^G - R_{RF,t}\right) + b_i^L F_{B/M,t}^L + b_i^G F_{B/M,t}^G \\ + m_i^L F_{Mom,t}^L + m_i^G F_{Mom,t}^G + s_i^L F_{Size,t}^L + s_i^G F_{Size,t}^G + \varepsilon_{i,t} \end{aligned} \tag{2}$$

We term model (1) the hybrid Hou-Karolyi-Kho four-factor (HKK-4) model, and model (2) the hybrid Fama-French-Carhart four-factor (FFC-4) model. The main difference between model (1) and model (2) concerns the construction of the value factor.[10]

In both equations, $R_{i,t} - R_{RF,t}$ is the portfolio i's return in excess of the 1-month U.S. treasury rate; $R_{m,t}$ is the market return; $F_{C/P}$ is the C/P factor-mimicking portfolio return; $F_{B/M}$ is the B/M factor-mimicking portfolio return; F_{Mom} is the momentum factor-mimicking portfolio return; and F_{size} is the size factor-mimicking portfolio return. Superscripts "L" and "G" refer, respectively, to portfolios constructed at the local and global level, with this distinction depending on the specification. In country-level analysis, "local" refers to returns within that country, and global to returns outside that country. In regressions conducted at the regional level, local (global) refers to returns within (outside of) the region of interest. When evaluating the returns of global portfolio returns, we construct global risk factors using all stocks in all countries. Further detail regarding factor construction and model specifications are provided in Appendix A.II.

[10] Hou, Karolyi, and Kho (2011) show that the value factor created based on cash flow / price (C/P) performs better than the factor created from book-to-market equity (B/M) in the international cross-sectional asset pricing tests.

a. Worldwide results

Table 2 reports the results. In Panel A, the sample is comprised of U.S.-based firms, and in Panel B, the sample contains firms in the other 37 countries listed in Table 1. In each case, we report results for three different groups: 1) all stocks, 2) stocks of large firms, and 3) stocks of small firms. Outside the U.S., a firm is characterized as small if its market capitalization falls below the median value within the relevant country-month. Within the U.S., we use monthly NYSE median market capitalization to make the large-small distinction, so as to allow our results to be compared with prior studies.

U.S. Firms. Starting with the first row in Panel A, distressed U.S. stocks appear to underperform their risk benchmarks over short and medium horizons. At one month, the return differential is -35 basis points (t=*-2.02*) according to the HKK-4 factor model, similar to that with the FFC-4 factors (-50 basis points, t=*-2.13*). The monthly realizations of these portfolios are shown in Figure 1.1 (rightmost axis). However, moving to the right in Table 2, the results weaken as the holding period lengthens. At three months, the results are borderline significant, and are insignificant at longer horizons.

The second row shows the same result, but for large firms. Although we observe slight underperformance at the one- and three-month horizon, the evidence is weak, with p-values in the range of 10% or greater. The third row, however, shows stronger results for small U.S. firms. Excess returns at the one-month horizon are -60 basis points (t=-2.77) measured against the FFC-4 model, and -62 basis points (t=-2.85) relative to HKK-4 model. Similar, though slightly weaker results are observed at the three-month horizon, with FFC-4 and HKK-4 alphas of -46 basis points (t=-2.17) and -50 basis points (t=-2.32) respectively. Results at one year are not significant.

Relative to prior studies involving U.S. firms, the magnitudes we estimate are smaller. Campbell et al. (2008), for example, using the same portfolio construction methodology estimates monthly alphas in the range of 1-2%, a factor of three to four times larger than what we find (roughly 30-50 basis points for the sample of all stocks).

One possible explanation, consistent with recent work by McLean and Pontiff (2016), is that academic work on the subject created and/or increased awareness of the anomaly, alleviating mispricing from rational arbitrageurs. While we lack direct evidence of this mechanism, it seems relevant that the earliest work on the topic surfaced in the mid-1990s (Dichev (1998)), corresponding roughly to the beginning of our sample (1992), and the end (1998) of that analyzed by Campbell et al. Moreover, that the anomaly would have almost entirely disappeared among large firms (which have the lowest barriers to arbitrage), but remained among small ones, is also consistent with this hypothesis.

Non-U.S. Firms. Panel B presents the results for firms outside the U.S. For the all-firm sample, there is little evidence that distressed stocks perform poorly, even over short horizons (see also Figure 1.2 for monthly returns). Although the point estimates are negative in every instance, in no case is the difference between high- and low-*EDF* firm returns statistically significant. Likewise, just focusing on large firms (row 2), the results are no more promising.

However, among small firms (rows 3), the patterns strengthen substantially. At the one-month holding period, we estimate a risk-adjusted return spread between high- and low-*EDF* portfolios of -41 basis points per month (t = -2.26) relative to FFC-4, and -34 basis points per month (t=-1.87) relative to HKK-4. These differences persist through one year after portfolio formation. At three months, the comparable alphas are -41 basis points (t=-2.34) and -34 basis points (t=-1.90).

Virtually identical results are observed at a one-year horizon, with monthly alphas of -46 basis points (t=-2.73) and -37 basis points (t=-2.19), respectively.

The estimates in Table 2 focus on the extreme tails of the *EDF* distribution. To paint a more complete picture, in Figure 2 we present the one-month FFC-4 alphas (Figure 2.1) and HKK-4 alphas (Figure 2.2) for all ten *EDF* decile portfolios. For large U.S. stocks (squares), large non-U.S. stocks (diamonds), small U.S. stocks (triangles), and small non-U.S. stocks (crosses), the relation between credit risk and returns is negative, being most pronounced for last two groups. Most notable is the extreme underperformance of the highest default risk decile, particularly among small stocks.

c. Country and regional analysis

Table 3 disaggregates the results by region and country. The five mutually exclusive regions we consider are North America (U.S. and Canada), European Developed Markets, Japan, Asia-Pacific Developed Markets (excluding Japan), and Emerging Markets. To save space, we report only average returns, and highlight statistical significance at the 5% level with boldface type.

The main takeaway is that the distress anomaly is heavily concentrated in North America and Europe. Of the 14 countries considered in Europe, 10 have statistically significant high-low *EDF* spreads at both the one-month and three month horizons. Most of these survive FFC-4 adjustments, but only about half remain significant when regressed against HHK-4's risk factors. The strongest individual effects are found in Sweden, Holland, France, Finland, Austria, and the U.K., with monthly alphas in the neighborhood of -1% to -2%, depending on specification.

Aggregated countries into a single European portfolio, the underperformance of small, distressed firms is clearly apparent. At every horizon, and with or without adjusting for risk, firms in the most distressed *EDF* decile perform 9-15% worse (annualized) than those in the least distressed decile.

Outside Europe however – and particularly in emerging markets – things look different. For example, Chile is the only country in South America with any evidence of the distress anomaly, and in that case, the result is weak. Likewise, emerging markets in Asia such as Indonesia, Malaysia, and Pakistan, the Philippines, and Thailand show virtually no evidence of underperformance by distressed companies. In some cases, there is evidence of *outperformance* for distressed stocks, such as in Argentina and South Africa.

Together, these findings paint a somewhat mixed picture. Aggregated worldwide, small distressed firms have poor returns. But, this is not uniformly true, being much stronger in developed (mostly Western) European markets, and in its colonies with similar cultural roots, such as the U.S. Hence, the overall result is significant not because the distress anomaly is geographically ubiquitous, but because developed markets constitute the bulk of the firm-month observations. Later analysis probes these cross-country patterns in more detail.

IV. Risk-based explanations

Our analysis to this point can be viewed primarily as an out-of-sample robustness exercise, intended to address whether the distress anomaly is a spurious event within the U.S. In this section, we start to address the mechanism. The possibility we consider here is that high-distress stocks have low systematic risk, providing a rational justification for their low returns. Section (a) takes a general approach, examining a number of traditional risk measures such as leverage, market-to-book ratios, and factor loadings, whereas section (b) uses cross-country differences in creditor rights to more precisely test the shareholder advantage theory.

a. Portfolio characteristics

Although the long-short alphas discussed above are measured relative to standard risk factors, it is useful to directly examine characteristics typically associated with risks for each distress risk portfolio. If stocks with high credit risk appear less risky on observables, it is likely that they will be less risky on unobservables as well (Altonji, Elder, and Taber (2005)). Such a finding would lend credence to risk-based explanations for the distress anomaly.

Yet, as Campbell et al. (2008) document among U.S. firms, we find that in an international sample, stocks of distressed firms appear to be more, rather than less, risky. Table 4 presents portfolio-level averages for various risk proxies. Standard deviations of equity returns and leverage are monotonically increasing in credit risk, with highly significant differences for the return difference between the first and tenth deciles. Likewise, distressed stocks are smaller and have higher book-to-market ratios, both of which previous research (e.g., Daniel and Titman (1997)) has associated with high stock returns.

The table also shows average market betas for each portfolio, calculated based on the beta estimates from a *hybrid CAPM model*.[11] Betas uniformly increase with credit risk, with a highly significant difference of 0.33 between the highest and lowest *EDF* deciles. The solid lines in Figures 3 provide a graphical representation (the long-dashed and short-dashed lines will be discussed shortly). The differences in

[11] We first calculate the beta against a value-weighted market portfolio of stocks within the same country (β_local), and simultaneously the beta against a value-weighted market portfolio of stocks worldwide constructed from global sample of stocks *outside* a stock's own country (β_global). Beta from the hybrid model, denoted as β_hybrid, is calculated as the sum of both coefficients. In untabulated analysis, we've also considered the local betas (β_local) and global betas (β_global) separately, and results are very similar.

Table 4 also allow for a richer interpretation of the return patterns in Table 2 where, recall that risk adjustments often *accentuate* the underperformance of distressed firms, rather than explain it.

Were we to observe a counterfactual pattern to Table 4 – i.e., with distressed stocks having low betas, leverage, and volatility – then reconciling the return results with a risk-based explanation would be easier. In that case, standard risk controls would attenuate the estimated underperformance of distressed firms. Even if statistically significant alphas remained, one needs only to believe that traditional proxies capture only a portion of latent equity risk. By the very same reasoning however, that standard risk controls exacerbate the poor returns of distressed firms seems difficult to square with risk-based explanations.

b. Creditor right and shareholder advantage

Whereas the prior section addresses general plausibility of risk-based explanations for the distress anomaly, this section uses cross-country comparisons to more closely examine one specific theory. In a pair of related papers, Garlappi, Shu, and Yan (2008) and Garlappi and Yan (2011) argue that the low returns of distressed stocks reflect violations of the absolute priority rule (APR). In a standard model without APR violations, equity risk increases with default risk. However, as the authors show, this can reverse when equity holders can expropriate creditors in bankruptcy. For sufficiently high levels of default risk, higher credit risk can lower the rate of return required by equity holders.

Indeed, a key prediction of the shareholder advantage theory is that equity betas are hump-shaped in default risk, unlike the monotonic relation that obtains when absolute priority is respected (e.g., Merton (1974)). Whereas we already know from Table 4 that on average, high-*EDF* stocks have higher betas than low-*EDF* stocks, these patterns are unconditional. But the theory makes a more nuanced prediction: equity risk should be particularly low (in a relative sense) for distressed firms with the weakest creditor protection, where shareholder expropriation is more likely and/or severe.[12]

Fortunately, cross-country comparisons represent a powerful test of the shareholder advantage theory, given the considerable differences in protection that creditors are offered. We draw on La Porta, Lopez-de-Silanes, Shleifer, and Vishny's (1998) seminal cross-country classification of creditor protection. LLSV measure creditor strength in four ways: 1) creditor-imposed restrictions like consent or a minimum dividend being required for reorganization, 2) the ability of creditors to immediately claim cash flow rights following default (no automatic stay), 3) absolute priority is respected, i.e., secured creditors are ranked first upon liquidation, and 4) managers do not retain administration of the firm's assets following the resolution of reorganization. Their "creditor rights index (*CRI*)" is the sum of each of these, ranging from zero (least creditor protection) to four (most creditor protection). We use the updated data and rankings in creditor rights as reported by Djankov, McLiesh and Shleifer (2007). Table 4 shows the CRI value for each country.

Our analysis of the shareholder advantage theory involves two parts. First, consider Figure 3, which plots equity betas for three samples: 1) all stocks worldwide, 2) stocks in countries below the median value for *CRI*, and 3) stocks in countries at or above the median value. As seen, in no group does equity risk measured by market beta decline with *EDF*. Rather, in all three figures, the difference in betas between the highest and lowest *EDF* portfolios is largest (and positive). Moreover,

[12] Favara, Schroth, and Valta (2012) examine the relation between equity beta and country-level bankruptcy code. However, they do not study the spread in equity beta or return between high and low distress risk firms, to which the GSY theory pertains.

the positive relation between default risk and beta is stronger in countries with weak creditor protection, opposite of that predicted by the theory.[13]

Return tests are no more promising. In Table 5, we stratify the long-short results by creditor rights, using the median value to make the designation. Regardless of horizon or adjustment for risk, there is no evidence that the distress anomaly is systematically related to measures of creditor protection. Figure 4.1 shows the average high/low *EDF* return difference as a function of each country's *CRI* index. In both the table and figure, to the extent that inferences can be made, it is among sovereignties with high creditor protection where the underperformance of distressed stocks is the most severe which, as with conditional betas, goes against the theory's prediction.[14]

V. Mispricing

In this section, we explore the possibility that the poor returns of distressed stocks is caused by valuation errors – i.e., that investors systematically overshoot the prices of firms with high failure risk, with the subsequent correction accounting for their poor returns. Perhaps the cognitive bias most studied in financial economics is *overconfidence*, or the tendency of investors to overweight their private information relative to public signals. A primary manifestation of overconfidence is under-reaction to public information, whereby prices continue to drift in the direction of a news event, such as in post-earnings-announcement drift (Ball and Brown (1968)), or simply exhibit positive serial correlation in returns (Jegadeesh and Titman (1993, 2001)).

[13] For firms in the most distressed decile (*EDF-10*), the difference in market betas between low- and high-*CRI* countries is 0.22 (t=16.51).

[14] We have experimented with alternatives (e.g., three or five groups), and find similar results.

It is possible to conceptualize the distress anomaly as another manifestation of overconfidence, particularly in the case of recently distressed firms. Specifically, if overconfidence causes the market to be sluggish when impounding news about default risk into prices, then firms with elevated default probabilities will, at least temporarily, have stock prices too high relative to their fundamentals. Provided that rational arbitrageurs are limited in their collective ability to immediately correct this mispricing, low subsequent returns result.

This simple framework suggests two additional tests. First, mispricing should increase when investors are more overconfident. Sections (a) and (b) develop two independent proxies for overconfidence, and test for its relation with the distress anomaly. Second, the mispricing of distressed stocks should be strongest among those having recently received bad news, compared to those for which (bad) news is relatively stale. Section (c) develops two proxies for bad news among distressed firms, and asks whether subsequent returns are stronger following its arrival.

a. Overconfidence measure #1: Individualism

Our first set of tests is based on the assumption that there exist long-lived differences in country-level attributes that are correlated with investor overconfidence. For example, to the extent that overconfidence has a genetic basis, one might expect to observe persistent differences between groups or people, particularly when separated by large geographic distances for protracted periods. Or, perhaps there exist long-lived cultural norms and/or traditions that foster the expression of certain behavioral traits.

In the late 1960s and 1970s, Dutch psychologist Geert Hofstede conducted a survey of IBM employees among 41 countries, in an attempt to measure cross-country differences in cultural dimensions including *individualism*. Generally, individualism refers to the extent to which members of a culture are encouraged to be unique, or otherwise distinguish themselves from others. The opposite is *collectivism*, which stresses membership of and contribution to a group. From Hofstede (2011):

> *On the individualist side we find cultures in which the ties between individuals are loose: everyone is expected to look after him/herself and his/her immediate family. On the collectivist side we find cultures in which people from birth onwards are integrated into strong, cohesive in-groups, often extended families (with uncles, aunts and grandparents) that continue protecting them in exchange for unquestioning loyalty, and oppose other in groups. Again, the issue addressed by this dimension is an extremely fundamental one, regarding all societies in the world.*

The United States, many countries in Western Europe (e.g., Italy, United Kingdom), and Australia rank highly in Hofstede's cross-country individualism index. By contrast, individualism is comparatively lower in the Far East (e.g., China, Japan, Korea) and much of Latin America (e.g., Brazil, Mexico). Hofstede's individualism index is scaled 0-100, with an average (median) value of 50.36 (51.00) across our sample of 38 countries.

Research from the psychology literature provides a linkage between individualism and overconfidence (Van den Steen (2004)). In individualistic cultures, the ego is rewarded by standing out, leading to the widespread – and, on average incorrect – belief that one's abilities are above average (Heine, Darrin, Hazel, and Shinobu (1999)). Applied to finance, if one of these abilities involves collecting and/or analyzing private signals (say, about a stock's prospects), then overconfidence, and the consequent underreaction to public signals, obtains. In a recent application to finance, Chui, Titman, and Wei (2010) appeal to the individualism-overconfidence connection to explore variation in momentum profits – which, like the distress anomaly, can be generated by investor overconfidence – across a number of countries.

Inspection of the country-level results (Table 3) suggests a positive relation between individualism and the distress anomaly, which Figure 4.2 shows graphically. For countries with individualism index values less than 51, the average high/low *EDF* spread is approximately zero, with about half positive and half negative. On the other hand, among the 17 countries with values above 51, only one (South Africa) has a positive point estimate. The slope of the line through the 38 data points (one for each country) is negative and significant at 1% level.

Table 6 presents more formal analysis. The first panel presents the high/low *EDF* return spreads for small stocks in countries with individualism values below 51 (the median). Confirming Figure 4.2, the estimates are economically tiny and statistically insignificant. Returns for the complementary set of countries, however, show much stronger effects.[15] That the distress anomaly exists only in high-individualism countries, the same setting where Chui, Titman and Wei (2010) find strong momentum profits, suggests that similar cognitive bias is likely at play for both anomalies and drives investor underreaction to public information.

b. Overconfidence measure #2: Market states

Our second proxy for overconfidence is dynamic, and accordingly, can be constructed within countries. Recent theoretical work by Daniel, Hirshleifer, and Subrahmanyam (1998) and Gervais and Odean (2001) argues that *self-attribution bias* can create positive feedback in investor overconfidence. In essence, investors treat successes and failures asymmetrically, using the former to positively update the quality of their information (i.e., become more overconfident), but not vice versa.

Combining this intuition with the fact that most investors hold long positions in the stock market, Gervais and Odean (2001) argue that aggregate investor

[15] As with the sorts on creditor rights, we find similar results by classifying countries into three or five groups based on *Individualism*.

overconfidence should increase following market gains. In an empirical test of this prediction, Cooper, Gutierrez and Hameed (2004) find, in fact, that momentum profits are higher following market booms, which they attribute to high levels of investor overconfidence.

Following this line of reasoning, we explore whether prices for distressed firms are especially inflated when aggregate overconfidence is elevated. We classify all months into "up" or "down" markets based on the most recent 12-month market return denominated in local currency. Specifically, for month *t* of country *c*, if the 12-month trailing market return from month *t-11* to *t* is both positive and above the median value for all monthly observations of the country, we classify it an "up" market, and "down" otherwise.[16] We then estimate the following pooled regression:

$$\begin{aligned} \mathrm{R}_{\mathrm{i,c,t}} = {} & \mathrm{h}_0 \times \text{High EDF}_{\mathrm{i,c,t-2}} + \mathrm{l}_0 \times \text{Low EDF}_{\mathrm{i,c,t-2}} \\ & + \mathrm{h}_1 \times \text{High EDF}_{\mathrm{i,c,t-2}} \times \text{Up Mkt}_{\mathrm{c,t-2}} + \mathrm{l}_1 \times \text{Low EDF}_{\mathrm{i,c,t-2}} \times \text{Up Mkt}_{\mathrm{c,t-2}} \\ & + \alpha_0 \times \text{Up Mkt}_{\mathrm{c,t-2}} + \sum\nolimits_{\mathrm{j}} \beta_{\mathrm{j}} \text{Control}^{\mathrm{j}}_{\mathrm{i,t-1}} + \sum\nolimits_{\mathrm{k}} \gamma_{\mathrm{j}} \text{Country}^{\mathrm{k}}_{\mathrm{i}} \\ & + \sum\nolimits_{\mathrm{m}} \tau_{\mathrm{m}} \text{Month}^{\mathrm{m}}_{\mathrm{t}} + \varepsilon_{\mathrm{i,c,t}} \end{aligned} \tag{3}$$

where *Up Mkt* is a dummy variable taking a value one for months classified as "up" markets, and zero otherwise. *High EDF* and *Low EDF* are defined as before, and firm-level *Control* variables are identical to previous regressions. We also include country and month fixed effects.

Table 7 shows the results. First considering Panel A, note that in the presence of the *Up Mkt* interactions, neither *High EDF* or *Low EDF* are significant in isolation. However, both interactions are highly significant, implying that the high-low *EDF* spread is entirely concentrated following times of high aggregate market returns.

---

[16] We choose 12-month horizon to measure market states to balance the trade-off between measurement reliability and loss of sample periods. Our results are robust to various other windows. The results are also almost identical if we use the trailing market return denominated in USD.

These findings support the specific hypothesis that investor overconfidence drives the poor returns of distressed stocks, and are consistent with general evidence provided in Stambaugh, Yu, and Yuan (2012), which shows that a variety of anomalies exist primarily during "high sentiment" states. Note also the consistency with Chava and Purnanandam's (2010) finding that the distress anomaly is particularly strong in the 1980s, a strong bull market in the U.S.

Panel B considers an extension: if stock prices of distressed firms formed during strong markets are too high, then the correction to fundamental value should be swiftest during market downturns, when overconfidence presumably evaporates more quickly.[17] To explore this implication, we consider the four possible combinations of market states in both month t-2 (portfolio formation month) and t (return measurement month) and run a similar regression as in equation (3) by interacting *High EDF* and *Low EDF* dummies with these four pairwise market states.

The results suggest that distressed stocks underperform most, with 1.16% per month below average stocks, when prices are formed during *Up Mkt* (investor confidence level is high) and returns are measured during *Down Mkt* (overconfidence dissipates and price reverts to fundamentals). As such, the *EDF* long/short strategy generates a return of 2.13% per month when *Up Mkt* is followed by a *Down Mkt*.

Further insight is gained by examining stocks with high share turnover during good market return periods. Odean (1998, 1999) shows that overconfident investors tend to trade more aggressively, and Statman, Thorley and Vorkink (2006) and Griffin, Nardari and Stulz (2007) find that share turnover tends to be higher following good market returns, which they attribute partly to increasing investor overconfidence after market gains. Accordingly, we hypothesize that investor

[17] We thank an anonymous referee for making this suggestion.

overconfidence would be especially high among heavily traded shares during booming markets, and as such, distressed stocks tend to be overvalued more in that case and subject to more underperformance in the future.

To test this implication, we run a similar regression as in equation (3) by further including three-way interactions of the *EDF* dummies, *Up Mkt* dummy, and *High Turnover*, a dummy variable taking the value one for stocks with above median share turnover in month *t*-2 and zero otherwise. The regression results, reported in Panel C, show that distressed stocks underperform most severely when market return is good and share turnover is high during the price formation month, with the estimated coefficient on the three-way interaction, *High EDF* × *Up Mkt* × *High Turnover*, being negative and highly significant, both statistically and economically.

c. News

This final section presents the most direct evidence that underreaction causes distressed stocks to be (temporarily) overpriced. Using two proxies for the arrival of bad news, we sort distressed firms into "news" and "no news" groups, and then compare their subsequent returns. Under the notion that underreaction only makes sense in the presence of news – it is difficult to react to nothing – particularly poor performance among the news group would point to underreaction as a source of mispricing. On the other hand, if the performance of these groups is similar, or even reversed, this would represent strong evidence against underreaction being the dominant reason that distressed stocks perform so poorly on average.

We obtain firm-specific news from the RavenPack News Analytics over the period January 2000 through December 2012.[18] Unfortunately, news data outside

[18] See Appendix III for further detail regarding the construction of the news data. News coverage by RavenPack is similar to other database used in prior studies of firm-specific news in the United States (see, Fang and Peress (2009)).

the U.S. is extremely sparse (Griffin, Kelly, and Nardari (2010)), even in developed markets (e.g., Western Europe). This is especially the case for small firms, as can be seen in Appendix Table AIII.1. Consequently, we limit our analysis in this section to U.S. firms.

Before proceeding, it is useful to explicitly state the key assumption underlying the tests in this section: among the most distressed firms, the average piece of news is negative. To see why this is necessary, recall that we are attempting to explain a directional pattern in stock returns. If news is relatively balanced – i.e., good and bad news cancel out – underreaction will generate price inefficiency *within* the *EDF-10* portfolio, but not a signed return prediction for the portfolio as a whole.

This assumption is testable. In Figure 5, we plot the average monthly "news sentiment score" for small U.S. firms within each *EDF* decile. The source of our news data, RavenPack, uses a proprietary algorithm to score the tone of firm-specific news stories. We standardize news tone scores to the unit interval, where zero corresponds to neutral tone, and +1 (-1) to maximal positive (negative) tone. As seen in the figure, sentiment scores are slightly positive on average.[19]

More interesting however, is that apart from the *EDF*-10 group, the tone is nearly invariant to default risk. The average sentiment scores range from the lowest level of 0.089 for *EDF*-9 to the highest level of 0.101 for *EDF*-4. No pairwise difference between any of the groups from *EDF-1* through *EDF-9* is statistically significant. However, *EDF*-10 is a clear outlier. The average sentiment score for this group is 0.069, significantly lower than that in any other decile.

The reason that *EDF-10* sticks out is due to selection bias. For any other group, say *EDF-6*, we can think about there being three types of firms: 1) those

[19] This overall upward bias in news tone does not affect our analysis, as we use the relative ranking of news sentiment in our formal analysis. We also construct an alternative news sentiment measure that is not influenced by news tone bias. The results, reported in Table A.III.2 in Appendix III, are almost identical to the baseline estimates discussed above.

having migrated downward from *EDF-5* or higher, 2) those having migrated upward from *EDF-7* or lower, and 3) long term "remainers" in *EDF-6*. Relative to the overall (slightly positive) tone, average new tone will presumably be negative for group 1, positive for group 2, and neutral for group 3. If groups 1 and 2 are similar in size, the average news tone for *EDF-6* will correspond roughly to the overall average.

However, this argument breaks down for firms near the bankruptcy threshold, since would-be upward migrants have defaulted, and thus no longer exist. Consequently, the average news is a blend of group 1 and 3, which will be negative on average. Completing the argument, this imbalance in news tone implies that underreaction to the average piece of news for firms in *EDF-10* generates stock prices that are too high, and consequently, subsequent returns that are took low.

With this framework in mind, we conduct two empirical tests. As our first proxy for bad news, we use the RavenPack sentiment score. Our second proxy is based on the length of time a firm has resided in *EDF-10*; presumably, recent migration from a less distressed decile (e.g., *EDF-9* or *EDF-8)* would have been accompanied by news, even if not present in the Ravenpack dataset. For each measure, we form terciles, and assign indicators for the "lowest" groups – i.e., a *Negative News* dummy for the third with most negative news stories, and a *Newly Distress* dummy for the third with the least time in *EDF-10.*[20] We then estimates the following regression:

$$\begin{aligned} R_{i,t} = \;& h_0 \times \text{High EDF}_{i,t-2} + h_1 \times \text{High EDF}_{i,t-2} \times \text{Negative News}_{i,t-2} \\ & + h_2 \times \text{High EDF}_{i,t-2} \times \text{Newly Distress}_{i,t-2} + l_0 \times \text{Low EDF}_{i,t-2} \\ & + \sum_j \beta_j \text{Control}^j_{i,t-1} + \varepsilon_{i,t} \end{aligned} \tag{4}$$

Table 8 shows the results. In the first column, we estimate a negative and significant interaction between the *High EDF* and *Negative News* dummy variables. Note also that the estimated coefficient on *High EDF* alone is essentially zero,

[20] On average, firms in the lowest third have been in the *EDF*-10 decile for two months.

implying that the poor returns (on average -80 basis points per month, t=-2.99) of distressed firms with recent bad news drive the overall distress effect.[21] Column 2 shows the results using the alternative news proxy, the length of time a firm has resided in the highest EDF decile. The interaction between *High EDF* and *Newly Distressed* is negative and significant (-51 basis points per month, t=-2.26) and, as in column 1, the *High EDF* indicator alone is not significant. In the third column, both interactions are included simultaneously. The coefficients are nearly unchanged, suggesting that the two approaches are complimentary in measuring news.

d. Other alternatives and robustness

In addition to the results presented formally, in the Appendix IV, we present a number of alternatives to our main regressions, addressing sample selection, specification, or alternative explanations for our results.

Skewness. Because the returns of distressed stocks are highly skewed to the right, it is possible that the high (low) returns (prices) we observe reflect preferences for these high order moments (Barberis and Huang (2008)), rather than valuation errors. Similar arguments are made in Campbell, Hilscher, and Szilagyi (2008), and Conrad, Kapadia, and Xing (2014). While this does not account for the cross-sectional or time-series (i.e., within country) evidence, it remains a possible explanation for our baseline results in Tables 2 and 3. Accordingly, in Appendix Table AIV.1, we include stock return skewness (*SKEW*), estimated from the past three months of daily returns (but skipping the portfolio formation month). While we observe a negative and significant point estimate, consistent with prior research,

[21]Note that the *Negative News* dummy does not enter the regression by itself, but only through its interaction with *High EDF*. This omission reflects our interest in measuring any news/no-news differential for firms within the *EDF-10* group, rather than the differential predictive ability of *Negative News* between *EDF* deciles. While we believe this specification is most appropriate, the result is robust to including the *Negative News* variable alone.

the underperformance of *High EDF* stocks continues to be observed, especially among small stocks, and with roughly the same magnitude as before.

Microcap stocks. Our main results exclude microcap stocks, defined at the 5th percentile of market capitalization within each country. One consideration is data quality, given that microcap stock are known to be extremely illiquid and/or suffer from stale prices, volumes, or other considerations that introduce noise into the analysis. The second concern is economic significance, given that microcap stocks comprise a tiny fraction (less than 0.04%) of overall market value.

Nonetheless, in Appendix Tables AIV.2, AIV.3, and AIV.4, we reproduce our main results in Table 3 (regional long-short profits), Tables 5 and 6 (creditor rights/individualism and the distress anomaly), and Table 7 (up/down markets) by including micro-cap stocks. The results are very similar. For example, for the value-weighted *EDF* long/short portfolio formed among small stocks in European developed markets, our benchmark monthly returns at the one-month horizon in Table 3 are -97 basis points per month (t=-5.69) and -96 basis points per month (t=-5.54) with Fama-French-Carhart and Hou-Karolyi-Kho risk adjustments, respectively. In the sample including microcaps, the comparable estimates are -111 basis points (t=-5.41) and -109 basis points (t=-5.23), respectively. Likewise, results involving creditor rights/individualism and up/down are essentially unchanged, with the inclusion of microcaps slightly strengthening the latter.

Excluding financial firms. Leverage plays an important role in the Merton (1974) model, and consequently, is strongly related to *EDF*s. Because firms in the financial sector typically have very high leverage ratios (Kalemli-Ozcan, Sorensen, and Yesiltas (2012)), but because of their ability to hedge or otherwise reduce risk, the leverage-default relation may be reduced for banks or other intermediaries (Adrian and Shin (2009, 2014)). To address this concern, Appendix Table AIV.5

repeats Table 3 (regional long-short analysis), but excludes all stocks of financial firms (those with Standard Industry Classification (SIC) Code within 6000-6999). For the small-stock sample, the results are very similar in European developed markets, but weaker in North America. In unreported results, the creditor rights/individualism (Tables 5 and 6) and market state (Table 7) results are nearly unchanged if we exclude financial firms.

Industry adjustments. Default likelihood may vary across different industries. Thus, one may be concerned that our results are driven by a small number of industries clustered within country-month *EDF* cohorts. To address this concern, we first demean *EDF* and other firm characteristics by the industry-country-month cohort, using the 1-digit Standard Industry Classification (SIC) codes, and then re-run our baseline specifications. Appendix Table AIV.6 shows the results of replicating the benchmark result (Table 3), but with this industry adjustment. As seen, the results are almost identical. Likewise, the results from the cross-sectional cuts explored in Tables 5, 6, and 7 are essentially unchanged when adjusted by industry.

VI. Conclusion

In this paper, we explore both the robustness of and mechanism underlying the distress anomaly, using a comprehensive panel of international data. Analyzing a sample of around 4.3 million firm-months, more than 44,000 stocks, in 38 countries covering over two decades 1992-2013, we document the presence of a distress anomaly, found mostly among stocks of small companies in North America and European developed markets. The magnitudes vary across specifications, but on average, financing a short position in a country's most distressed decile with a long position in its least distressed decile among small stocks earns 40-50 basis points per month, roughly on par with the returns of a plain vanilla momentum strategy during the same sample period.

While we find little evidence to support risk-based explanations for these patterns, several pieces of evidence point to a behavioral/mispricing interpretation. For example, the returns of distressed stocks are strongly linked to a country-level proxy for overconfidence (individualism), suggesting that the same type of investor under-reaction sustaining other asset pricing anomalies (e.g., momentum) is at work for the distress anomaly as well. Moreover, the distress anomaly exists mainly following periods of good market returns and is concentrated among those stocks with high share turnover during those good market states. Finally, mispricing is most severe for stocks having recently received bad news, consistent with the temporary mispricing predicted by underreaction.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber, 2005, Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools, *Journal of Political Economy* 113, 151-184.

Adrian, Tobias, and Hyun Song Shin, 2009, Money, liquidity, and monetary policy, *American Economic Review (papers and proceedings)* 99, 600-605.

Adrian, Tobias, and Hyun Song Shin, 2014, Procyclical leverage and Value-at-Risk, *Review of Financial Studies* 27, 373-403.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2007, Momentum and credit rating, *Journal of Finance* 62, 2503-2520.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2013, Anomalies and financial distress, *Journal of Financial Economics* 108, 139-159.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2012, The world price of credit risk, *The Review of Asset Pricing Studies* 2, 112-152.

Ball, Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers, J*ournal of accounting research* 6, 159-178.

Barberis, Nicholas, and Ming Huang, 2008, Stocks as lotteries: The implications of probability weighting for security prices, *American Economic Review* 98, 2066-2100.

Bhamra, Harjoat S., and Kyung Hwan Shim, 2015, Small growth and distress returns: Two sides of the same coin?, working paper, The University of New South Wales.

Bharath, Sreedhar T., and Tyler Shumway, 2008, Forecasting default with the Merton distance to default model, *Review of Financial Studies* 21, 1339-1369.

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

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

Chava, Sudheer, and Amiyatosh Purnanandam, 2010, Is default risk negatively related to stock returns?, *Review of Financial Studies* 23, 2523-2559.

Chui, Andy C.W., Sheridan Titman, and K.C. John Wei, 2010, Individualism and momentum around the world, *Journal of Finance* 65, 361–392.

Conrad, Jennifer S., Nishad Kapadia, and Yuhang Xing, 2014, Death and jackpot: Why do individual investors hold overpriced stocks?, *Journal of Financial Economics* 113, 455-475.

Cooper, Michael J., Roberto C. Gutierrez, and Allaudeen Hameed, 2004, Market states and momentum, The *Journal of Finance* 59, 1345-1365.

Correia, Maria M., Scott A. Richardson, and A. Irem Tuna, 2012, Value investing in credit markets, *Review of Accounting Studies 17*, 572-609.

Crosbie, Peter J., and Jeffrey R. Bohn, 2003, Modeling default risk, Technical Report, Moody's KMV, available at http://www.defaultrisk.com/pp model 35.htm.

Da, Zhi and Pengjie Gao, 2010, Clientele change, liquidity shock, and the return on financially distressed stocks, *Journal of Financial and Quantitative Analysis* 45, 27–48.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In search of attention, *Journal of Finance* 66, 1461-1499.

Dang, Tung L., Fariborz Moshirian, and Bohui Zhang, 2015, Commonality in news around the world, *Journal of Financial Economics* 116, 82-110.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and over-reactions, *Journal of Finance* 53, 1839-1886.

Daniel, Kent, and Sheridan Titman, 1997, Evidence on the characteristics of cross sectional variation in stock returns, *Journal of Finance* 52, 1-33.

Davydenko, Sergei A., 2011, When do firms default? A study of the default boundary, working paper, University of Toronto.

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

Djankov, Simeon, Carakee McLiesh, and Andrei Shleifer, 2007, Private credit in 129 countries, *Journal of Financial Economics* 84, 299-329.

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

Fama, Eugene F., and Kenneth R. French, 2012, Size, value, and momentum in international stock returns, *Journal of Financial Economics* 105, 457–472.

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

Fan, Joseph, Sheridan Titman, and Gary Twite, 2012, An international comparison of capital structure and debt maturity choices, *Journal of Financial and Quantitative Analysis* 47, 23-56.

Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *Journal of Finance* 64, 2023-2052.

Favara, Giovanni, Enrique Schroth, and Paolo Valta, 2012, Strategic default and equity risk across countries, *Journal of Finance* 67, 2051-2095.

Friewald, Nils, Christian Wagner, and Josef Zechner, 2014, The cross-section of credit risk premia and equity returns, *The Journal of Finance* 69, 2419-2469.

Garlappi, Lorenzo, and Hong Yan, 2011, Financial distress and the cross-section of equity-returns, *Journal of Finance* 66, 789-822.

Garlappi, Lorenzo, Tao Shu, and Hong Yan, 2008, Default risk, shareholder advantage, and stock returns, *Review of Financial Studies* 21, 2743–2778.

George, J. Thomas, and Chuan-Yang Hwang, 2010, A resolution of the distress risk and leverage puzzles in the cross section of stock returns, *Journal of Financial Economics* 96, 56-79.

Gervais, Simon, and Terrance Odean, 2001, Learning to be overconfident, *Review of Financial studies* 14, 1-27.

Griffin, John, Patrick J. Kelly, and Federico Nardari, 2010, Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets, *Review of Financial Studies* 23, 3225-3277.

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

Griffin, John M., Federico Nardari, and René M. Stulz, 2007, Do investors trade more when stocks have performed well? Evidence from 46 countries, *Review of Financial studies* 20, 905-951.

Hackbarth, Dirk, Rainer Haselmann, and David Schoenherr, 2015, Distress risk, stock returns, and the 1978 bankruptcy reform act, *Review of Financial Studies* 28, 1810-1847.

Heine, Steven J., Darrin R. Lehman, Hazel Rose Markus, and Shinobu Kitayama, 1999, Is there a universal need for positive self-regard? *Psychological Review* 106, 766–794.

Hofstede, Geert, 2001, *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations,* 2nd ed. (Sage Publication, Beverly Hills , CA).

Hou, Kewei, G. Andrew Karolyi, and Bong-Chan Kho, 2011, What factors drive global stock returns?, *Review of Financial Studies* 24, 2528-2574.

Ince, Ozgur S., and R. Burt Porter, 2006, Individual equity return data from Thomson Datastream: Handle with care! *Journal of Financial Research* 29, 463-479.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *The Journal of Finance* 45, 881-898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 56, 699-720.

Johnson, Timothy C., Tatiana Chebonenko, Igor Cunha, Fabricio D'Almeida, and Xiangyi Spencer, 2011, Endogenous leverage and expected stock returns, *Finance Research Letters* 8, 132-145.

Karolyi, G. Andrew, and Ying Wu, 2012, Size, value, and momentum in international stock returns: A partial segmentation approach, *Working Paper*, Cornell University.

Kalemli-Ozcan, Sebnem, Bent Sorensen, and Sevcan Yesiltas, 2012, Leverage across firms, banks, and countries, *Journal of International Economics* 88, 284-298.

Kolasinski, Adam C., Adam V. Reed, and Matthew C. Ringgenberg, 2013, A multiple lender approach to understanding supply and search in the equity lending market, *Journal of Finance* 68, 559-595.

La Porta, Rafael, Florencio Lopes-de-Silanes, Andrei Shleifer, and Robert Vishny, 1998, Law and finance, *Journal of Political Economy* 106, 1113-1155.

Lehmann, Bruce N., 1990, Fads, martingales, and market efficiency, *The Quarterly Journal of Economics* 105, 1-28.

Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449–70.

Odean, Terrance, 1998, Are investors reluctant to realize their losses?, *Journal of finance* 53, 1775-1798.

Odean, Terrance, 1999, Do investors trade too much?, *American Economic Review*, 1279-1298.

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5-32.

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

Statman, Meir, Steven Thorley, and Keith Vorkink, 2006, Investor overconfidence and trading volume, *Review of Financial Studies* 19, 1531-1565.

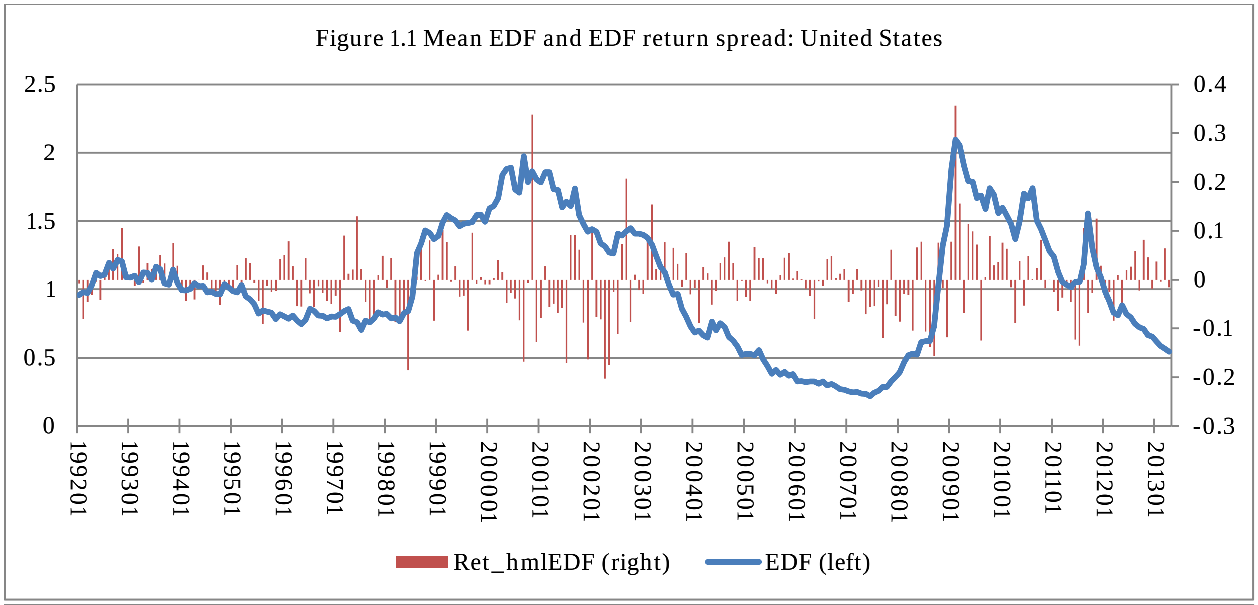
Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35-55.

Van den Steen, Eric, 2004, Rational overoptimism (and other biases), *American Economic Review* 94, 1141–1151.

Vassalou, Maria, and Yuhang Xing, 2004, Default risk in equity returns, *Journal of Finance* 59, 831–868.

Figure 1: Time-Series of the average Expected Default Frequency (*EDF*) and Returns of EDF-based Long/Short Portfolios

This figure plots the time series of monthly average values of Expected Default Frequency (EDF, in percentage) and monthly returns of EDF-based long/short portfolios for all U.S. stocks and all non-U.S. stocks in our sample respectively. The EDF-based long/short portfolios are formed as follows. At the end of each month t, we rank all stocks within every country by EDF, and then form deciles. To form EDF decile portfolios for all non-U.S. countries in the sample, we aggregate stocks across all non-U.S. countries within each EDF decile. We calculate the value-weighted returns in month t+2 (one month after portfolio formation) for a long-short strategy involving EDF-deciles 10 (long) and 1 (short).



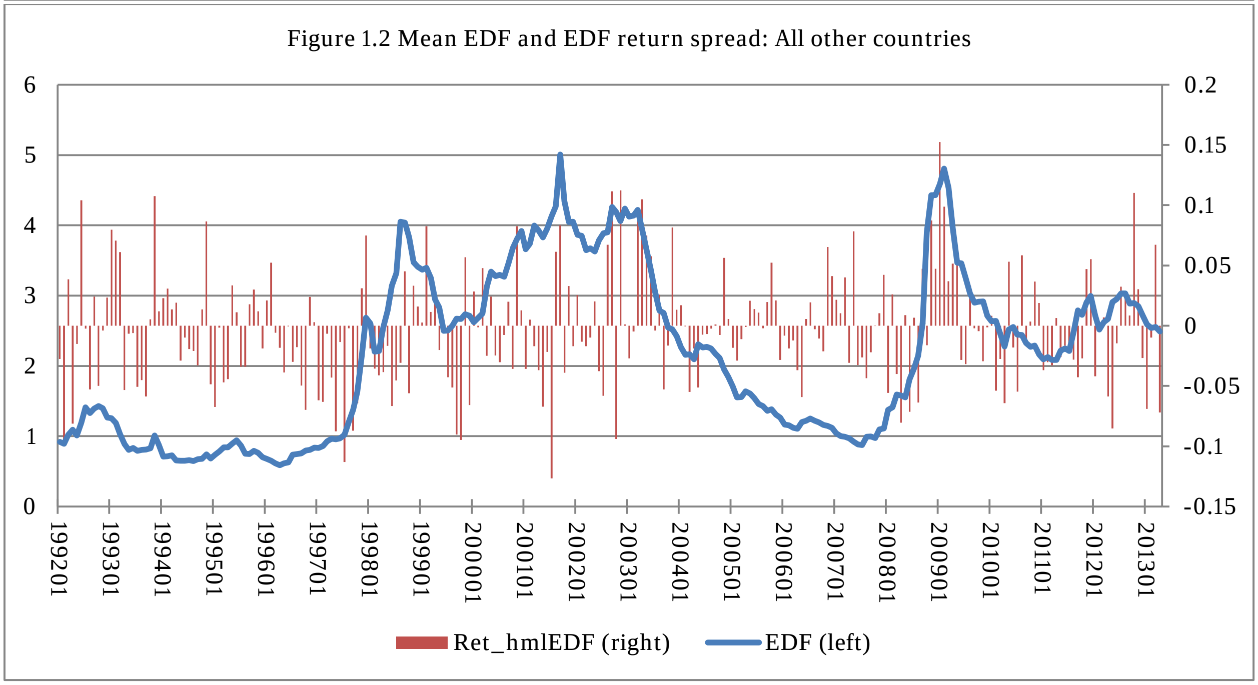


Figure 2: Factor-adjusted Returns on *EDF*-sorted Portfolios

This figure plots both Fama-French-Carhart four-factor model alpha (FFC4-α) and Hou-Karolyi-Kho four-factor model alpha (HKK4-α) of *EDF*-sorted decile portfolios for U.S. stocks and stocks in all other countries in our sample respectively. First, the *EDF*-sorted decile portfolios for a specific sample are formed as follows: 1) at the end of each month *t*, we sort all stocks in that sample into deciles by *EDF* within each country, where decile-10 is for stocks with the highest 10% *EDF* values and decile-1 is for stocks with the lowest 10% *EDF* values; 2) to form *EDF* decile portfolios for non-U.S. stocks, we then aggregate firms across all countries excluding U.S. within each *EDF* decile. Second, we calculate the value-weighted returns of *EDF*-decile portfolios in the month *t*+2, skipping the month immediately following the portfolio formation month. Finally, we calculate the alphas of portfolios from both Fama-French-Carhart and Hou-Karolyi-Kho four-factor models. The factor construction process follows Fama and French (2012) and Hou, Karolyi, and Kho (2011). Details of factor construction are provided in the appendix. We report the results both large stocks and small stocks, where the market capitalization cutoff value for large versus small stocks is the NYSE median value for the U.S. and the median value of all stocks in the country for all other countries.

Figure 2.1 FFC4-α of EDF-sorted decile portfolios

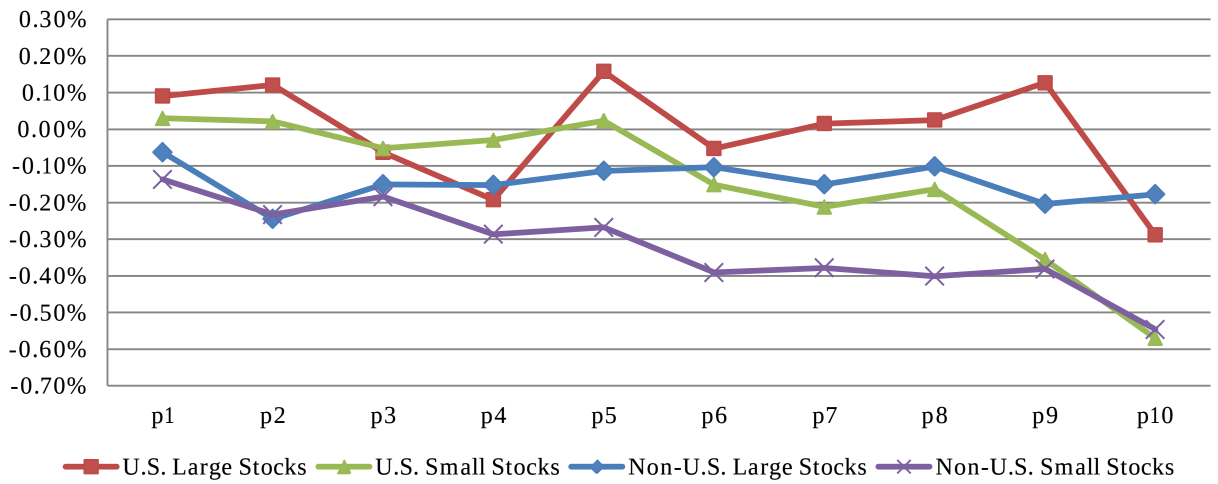


Figure 2.2 HKK4-α of EDF-sorted decile portfolios

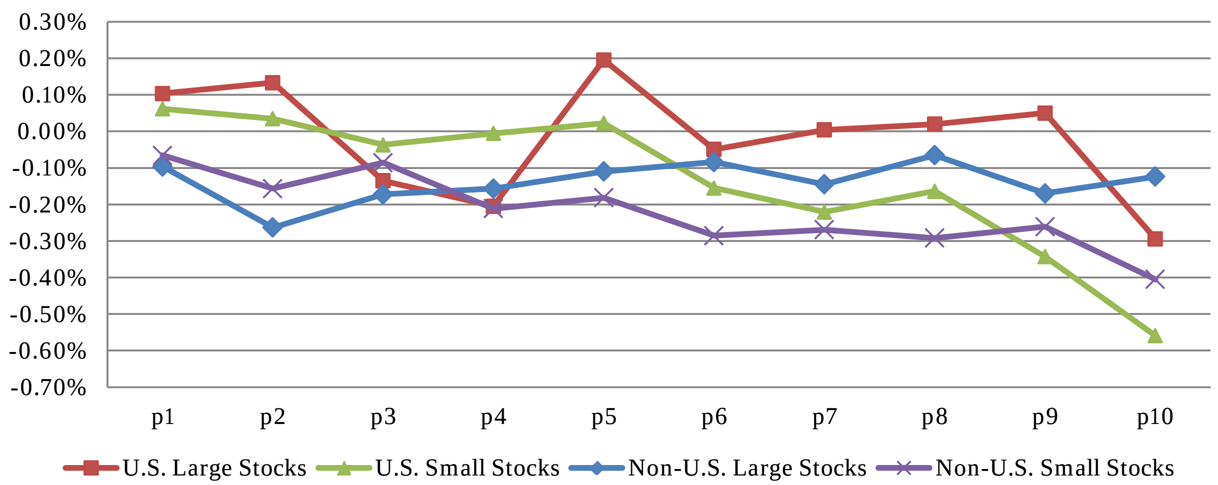


Figure 3: Average CAPM betas of *EDF*-sorted Portfolios of Global Stocks

This figure plots the average CAPM betas of *EDF*-sorted decile portfolios of all stocks in all countries, high Creditor Rights countries, and low Creditor Rights countries respectively. The *EDF*-sorted decile portfolios for all countries are formed as follows: 1) at the end of each month *t*, we sort all stocks within each country into deciles by *EDF*, where decile-10 is for stocks with the highest 10% *EDF* values and decile-1 is for stocks with the lowest 10% *EDF* values; 2) we then aggregate firms across all countries within each *EDF* decile to form global *EDF* decile portfolios. To obtain the average beta of an *EDF* decile portfolio, we first calculate the average beta of all stocks in the portfolio every month, and then take the average of monthly mean betas across all months. We form *EDF*-sorted decile portfolios for high (low) Creditor Rights countries and calculate the average betas of portfolios in a similar way, using only stocks from countries with above (below) median value of La Porta, Lopez-de-Silanes, Shleifer, and Vishny's (1998) creditor rights index. A stock's beta is the sum of estimated loadings on both local and global market returns in a hybrid version of CAPM, which includes both value-weighted average returns of all stocks in the stock's country and value-weighted average returns of all stocks in all countries other than the stock's country as factors. We use the past 60-month data and require a minimum of 24 months of valid data for the beta estimation.

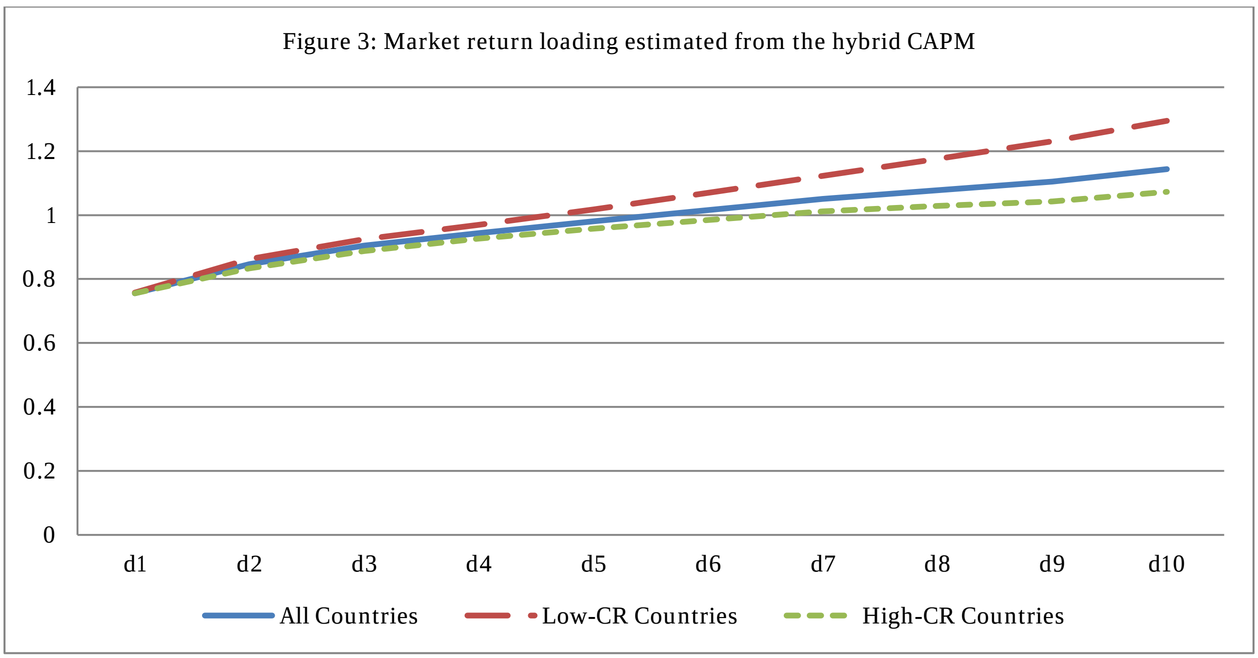
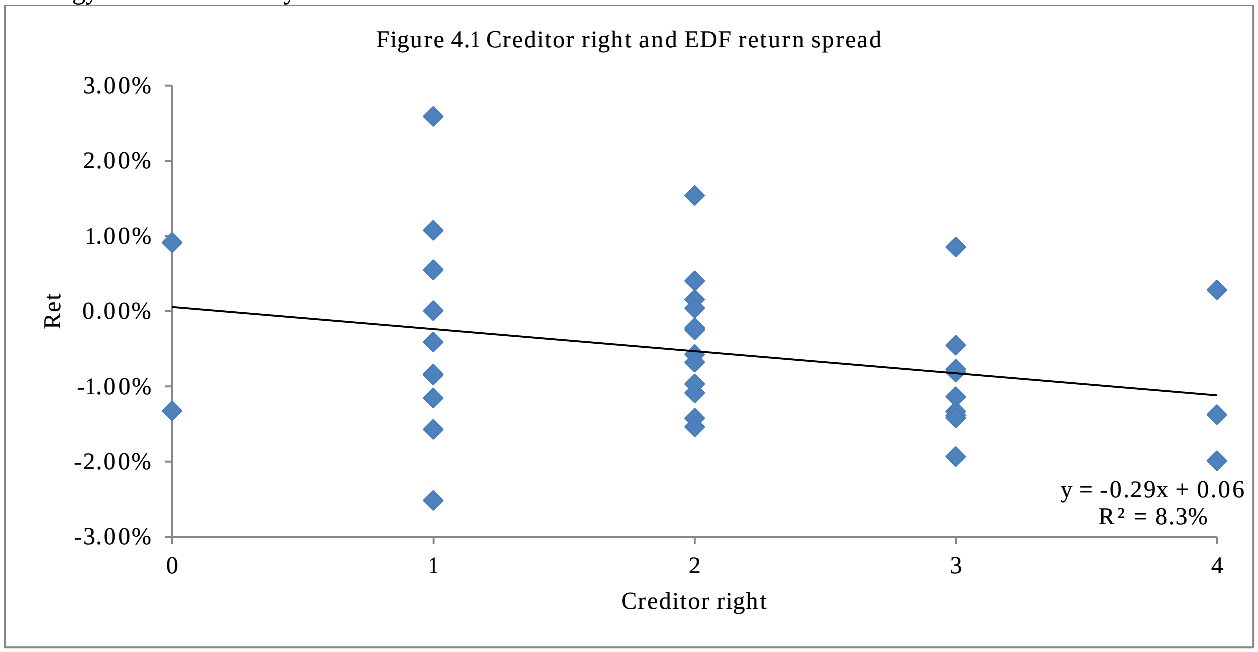


Figure 4: Individualism, Creditor Rights, and the Average Returns of Country-level Long/Short Small-stock Portfolios

This figure plots the average returns of long/short portfolios that buy high-EDF small stocks and sell low-EDF small stocks for each country against the country's Hofstede's (2001) individualism index and La Porta, Lopez-de-Silanes, Shleifer, and Vishny's (1998) creditor rights index. At the end of each month t, we rank all small stocks within each country by EDF, and then form deciles portfolios for each country. The market capitalization cutoff value for large versus small stocks is the NYSE median value for the U.S. and the median value of all stocks in the country for all other countries. We calculate the value-weighted returns in month t+2 (one month after portfolio formation) for a long-short strategy involving EDF-deciles 10 (long) and 1 (short). Finally, we calculate the time-series average return of this long/short strategy for each country.



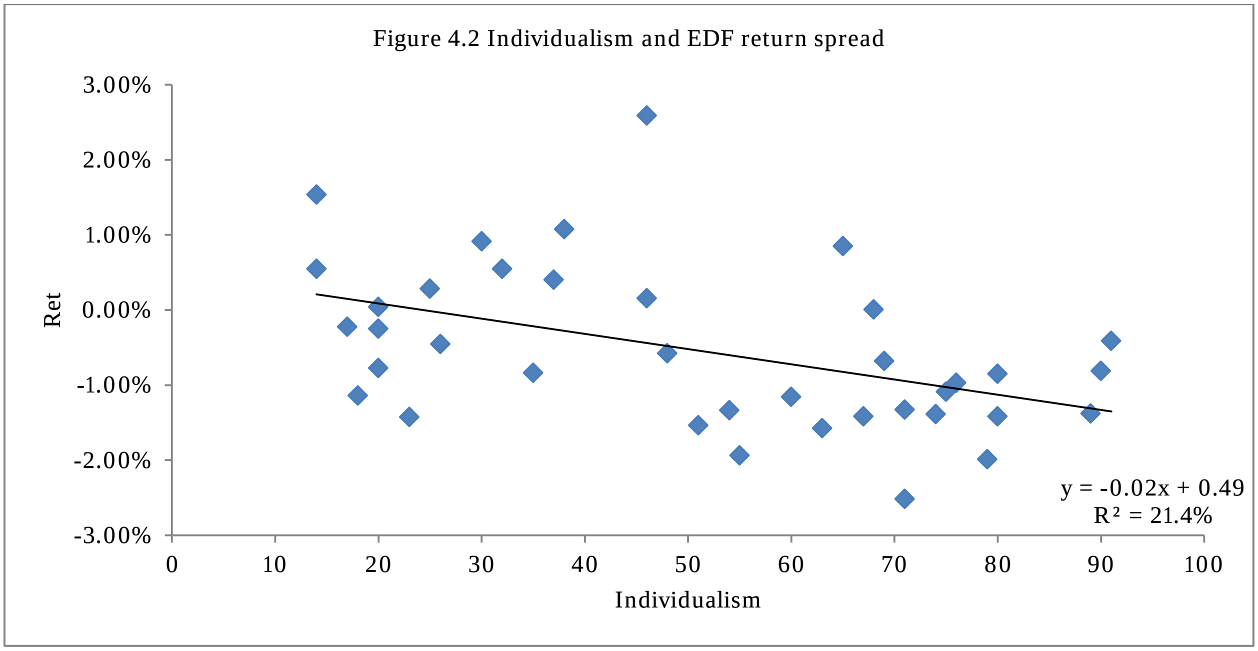


Figure 5: News Tone of U.S. Small Stocks with Different Levels of EDF

This figure plots the average news Event Sentiment Scores (ESS) for U.S. small stocks (defined as those with market capitalization below the NYSE median) different levels of EDF, where small stocks. We first take the average ESS scores of all news stories of a stock in a month as its ESS score in the month. We then rank all small stocks into ten decile groups within each month based on the EDF. Finally, for each EDF decile group, we calculate its monthly ESS scores as the average of stock-level monthly ESS scores across all member stocks, and report the time-series average.

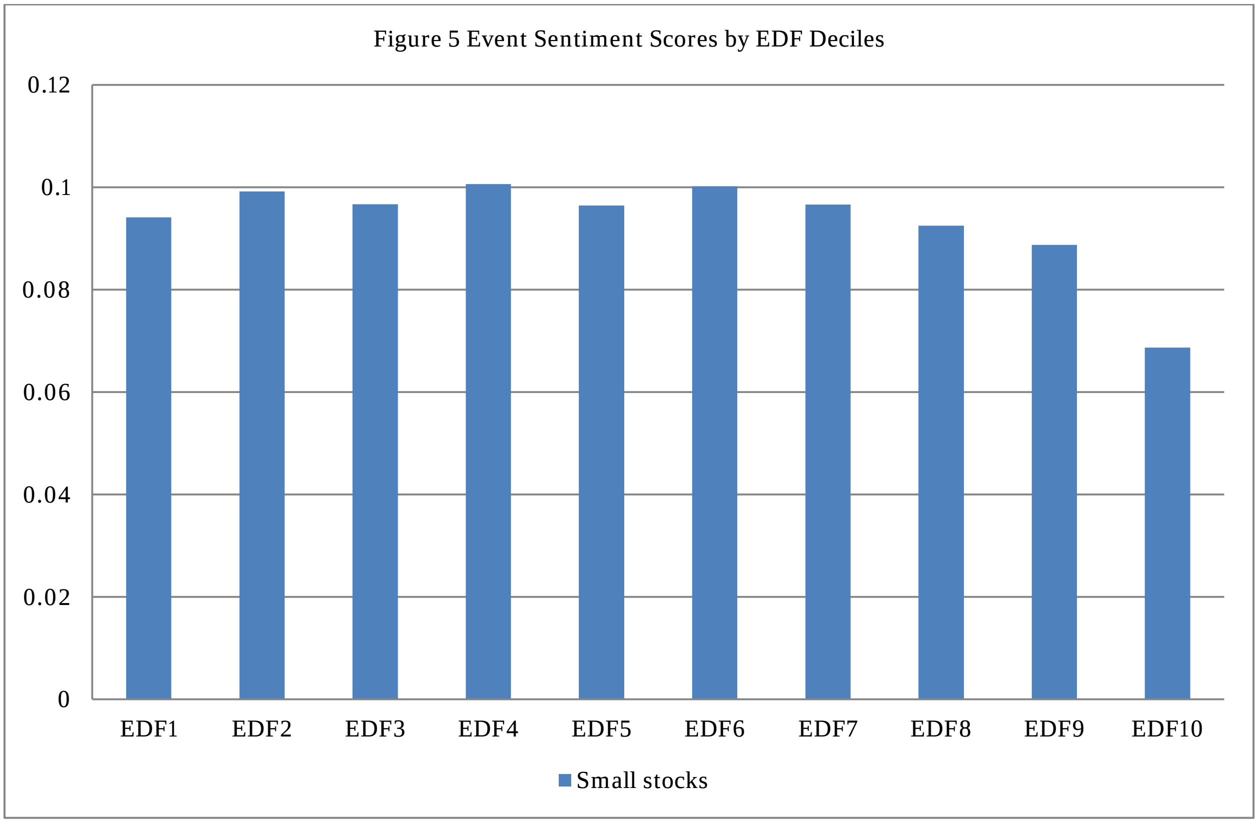


Table 1: Summary Statistics

This table presents summary statistics for 38 countries included in our sample. Countries (column 1) are classified into 5 regions based on economic and geographic proximity. Columns 2 to 4 list, respectively, the starting month, the number of unique stocks, and the number of stock-month observations for each country. Columns 5 - 10 list the time-series average value of monthly median values of the following variables: 1) EDF, the Expected Default Frequency measure produced by Moody's KMV; 2) the market capitalization in millions USD at the month end; 3) Book-to-Market ratio: the ratio of book equity of the fiscal year ending at least 6 months before to the market capitalization at the month end; and 4) Cash flow-to-Price ratio: the ratio of cash flow of the fiscal year ending at least 6 months before to the market capitalization at the month end; 5) Momentum: the most recent 11-month return denominated in USD up to the month end; and 6) Turnover: the number of shares traded in the month divided by the number of shares outstanding at the end of month. For each variable, we first calculate the median value for each country each month, and then average the monthly median values for each country to obtain the final statistics.

Country	Beginning Date	Number of Stocks	Number of observations	EDF (%)	Size (US$ millions)	Book-to-Market Ratio	Cash flow-to-Price Ratio	Momentum (%)	Turnover (%)
North America									
Canada	1998/04	2,574	181,043	0.74	92.50	0.60	0.05	4.20	2.57
United States	1992/01	12,134	959,289	0.32	442.72	0.45	0.08	10.62	9.63
European Developed Markets									
Austria	1997/03	126	12,928	0.23	320.64	0.75	0.14	8.17	1.01
Belgium	1995/12	201	22,210	0.14	236.08	0.73	0.12	8.65	0.91
Denmark	1993/08	261	30,724	0.29	107.31	0.86	0.11	8.50	1.88
Finland	1995/12	160	20,944	0.23	222.63	0.69	0.15	10.77	2.24
France	1992/01	1,156	127,100	0.32	180.16	0.63	0.11	6.85	1.05
Germany	1992/01	1,071	120,062	0.47	163.22	0.56	0.10	2.02	1.00
Greece	1996/10	369	39,163	2.07	108.21	0.97	0.10	7.52	2.78
Italy	1992/01	406	49,036	0.39	330.21	0.82	0.10	1.50	2.66
Netherlands	1992/01	223	31,448	0.22	387.01	0.58	0.13	9.64	4.29
Norway	1994/10	307	27,603	0.47	172.06	0.67	0.12	10.54	2.75
Spain	1992/01	201	26,785	0.23	691.93	0.63	0.11	7.65	3.15
Sweden	1994/08	536	50,256	0.26	173.90	0.58	0.11	10.51	3.44
Switzerland	1993/06	300	41,313	0.12	382.30	0.70	0.12	11.31	1.81
United Kingdom	1992/01	2,921	313,826	0.24	117.47	0.62	0.07	7.83	2.38

Country	Beginning Date	Number of Stocks	Number of observations	EDF (%)	Size (US$ millions)	Book-to-Market Ratio	Cash flow-to-Price Ratio	Momentum (%)	Turnover (%)
Japan	1992/01	4,422	679,086	0.57	218.34	0.91	0.10	0.60	1.98
Asia-Pacific Developed Markets (excluding Japan)									
Australia	1992/01	2,331	208,569	0.38	57.93	0.58	0.04	7.72	1.92
Hong Kong	1992/01	1,525	164,353	0.79	160.67	1.11	0.11	8.39	2.41
New Zealand	2001/10	140	11,543	0.20	102.09	0.58	0.08	17.76	0.88
Singapore	1992/04	861	88,819	0.67	113.35	0.89	0.11	8.58	1.22
Emerging Markets									
Argentina	2002/11	85	7,362	1.13	104.59	1.11	0.16	17.75	0.40
Brazil	2005/01	189	10,536	0.49	903.18	0.46	0.08	28.47	3.11
Chile	1997/01	201	20,871	0.21	351.31	0.71	0.11	11.91	0.25
China	2001/02	2,365	160,624	0.24	442.71	0.31	0.04	12.43	19.39
India	1994/07	2,348	158,041	1.35	114.70	0.77	0.14	8.71	1.77
Indonesia	1993/04	482	48,912	2.37	71.18	0.95	0.14	14.62	0.86
Israel	2001/04	436	19,887	0.61	256.04	0.75	0.11	9.25	1.65
Malaysia	1992/01	1,159	141,979	0.79	118.21	0.94	0.11	7.63	1.00
Mexico	1997/12	134	12,840	0.47	693.53	0.82	0.14	13.28	0.96
Pakistan	1997/12	287	21,039	2.17	61.43	0.86	0.19	15.86	0.88
Philippines	1996/04	255	29,056	1.63	69.86	1.25	0.10	6.16	0.61
Poland	1998/06	488	29,935	0.67	100.32	0.74	0.10	9.98	1.92
South Africa	1992/03	565	53,094	0.52	238.18	0.57	0.12	9.52	1.11
South Korea	1992/04	1,999	167,230	1.85	115.77	1.25	0.17	5.63	12.47
Taiwan	1995/04	774	101,719	0.43	246.75	0.70	0.09	1.23	12.29
Thailand	1993/04	593	71,824	1.76	54.58	0.91	0.14	9.21	1.71
Turkey	1997/04	345	34,602	1.01	122.25	0.64	0.13	17.22	8.98

Table 2: Default Risk and Stock Returns: Worldwide Evidence

Panels A and B present the average returns and alphas from both Fama-French-Carhart and Hou-Karolyi-Kho four-factor models for long/short hedging portfolios that buy high-EDF stocks and sell low-EDF stocks for the U.S. and all other countries in the sample respectively. At the end of each month t, we rank all stocks within every country by EDF, and then form deciles. We then aggregate stocks across all countries (excluding the U.S.) within each EDF decile to form portfolios for all other countries in the sample. For example, in month t, EDF-decile 7 consists of all stocks that, within each of their respective countries, ranked between the 70th and 80th percentile in terms of EDF at the end of month t. We focus on a long-short strategy involving EDF-deciles 10 (long) and 1 (short) that begins one month after portfolio formation, i.e., at the beginning of month t+2. For this long-short (L/S) portfolio, we report both the average raw monthly returns (value-weighted), alphas from the Fama-French-Carhart four-factor model (FFC4-α), and alphas from the Hou-Karolyi-Kho four-factor model (HKK4-α). The factor construction process follows Fama and French (2012) and Hou, Karolyi, and Kho (2011). Details of factor construction are provided in the appendix. Returns of EDF portfolios and factor portfolios are denominated in U.S. dollars. Results for various holding periods are shown, including one month (t+2), three-month (t+2, t+4), and one year (t+2, t+13). For holding periods greater than one month, we follow the overlapping horizon approach of Jegadeesh and Titman (1993). Finally, we report the results for three size groups: 1) all stocks, 2) large stocks and 3) small stocks, where the market capitalization cutoff value for large versus small stocks is the NYSE median value for the U.S. and the median value of all stocks in the country for all other countries.

Panel A: U.S. stocks

		Holding Periods								
		t+2			t+2, t+4			t+2, t+13		
Sample	Stat.	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α
all	Est	-0.35%	-0.50%	-0.48%	-0.24%	-0.42%	-0.43%	0.11%	-0.15%	-0.15%
	t-stat	*-0.78*	*-2.13*	*-2.02*	*-0.52*	*-1.80*	*-1.80*	*0.24*	*-0.64*	*-0.62*
large	Est	-0.12%	-0.39%	-0.40%	-0.02%	-0.34%	-0.37%	0.53%	-0.03%	-0.02%
	t-stat	*-0.27*	*-1.55*	*-1.60*	*-0.04*	*-1.38*	*-1.50*	*1.17*	*-0.10*	*-0.07*
small	Est	-0.41%	-0.60%	-0.62%	-0.30%	-0.46%	-0.50%	0.00%	-0.19%	-0.23%
	t-stat	*-1.03*	*-2.77*	*-2.85*	*-0.75*	*-2.17*	*-2.32*	*-0.01*	*-0.85*	*-1.03*

Panel B: Non-U.S. stocks

		Holding Periods								
		t+2			t+2, t+4			t+2, t+13		
Sample	Stat.	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α
all	Est	-0.34%	-0.17%	-0.06%	-0.28%	-0.15%	-0.02%	-0.31%	-0.34%	-0.22%
	t-stat	*-1.18*	*-0.71*	*-0.23*	*-0.99*	*-0.64*	*-0.08*	*-1.17*	*-1.54*	*-0.99*
large	Est	-0.22%	-0.11%	-0.03%	-0.16%	-0.09%	0.00%	-0.07%	-0.13%	-0.03%
	t-stat	*-0.80*	*-0.57*	*-0.13*	*-0.60*	*-0.46*	*-0.02*	*-0.28*	*-0.70*	*-0.16*
small	Est	-0.47%	-0.41%	-0.34%	-0.45%	-0.41%	-0.34%	-0.44%	-0.46%	-0.37%
	t-stat	*-2.36*	*-2.26*	*-1.87*	*-2.32*	*-2.34*	*-1.90*	*-2.40*	*-2.73*	*-2.19*

Table 3: Default Risk and Stock Returns: Breakdown by Country and Region

This table presents the average returns and alphas from both Fama-French-Carhart and extended Hou-Karolyi-Kho four-factor models for long/short portfolios that buy high-EDF small stocks and sell low-EDF small stocks for each country in our data set and also for each of the following regions: North America, European Developed Markets, Japan, Asia-Pacific Developed Markets (excluding Japan), and Emerging Markets. At the end of each month t, we rank all small stocks within each country by EDF, and then form deciles portfolios for each country. The market capitalization cutoff value for large versus small stocks is the NYSE median value for the U.S. and the median value of all stocks in the country for all other countries. To form regional EDF decile portfolios, we aggregate stocks across all countries in a region within each EDF decile to form portfolios for the region. For example, in month t, EDF-decile 7 consists of all stocks that, within each of their respective countries in the region, ranked between the 70th and 80th percentile in terms of EDF at the end of month t. We focus on a long-short strategy involving EDF-deciles 10 (long) and 1 (short) that begins one month after portfolio formation, i.e., at the beginning of month t+2. For this long-short portfolio, we report both the average raw monthly value-weighted returns and alphas from both Fama-French-Carhart and extended Hou-Karolyi-Kho four-factor models. Results for various holding periods are shown, including one month (t+2), three-month (t+2, t+4), and one year (t+2, t+13). For holding periods greater than one month, we follow the overlapping horizon approach of Jegadeesh and Titman (1993). Significant estimates (at 5% level) are in bold-faced and italic font.

For each country, we also report the following two country-level attributes: 1) Individualism (INDV), from Hofstede's (2001) individualism index, for which a higher value indicates a more individualistic culture; and 2) Creditor Rights (CR), corresponding to La Porta, Lopez-de-Silanes, Shleifer, and Vishny's (1998) creditor rights index, where zero represents the weakest creditor protection, and four the strongest.

			Holding Periods								
			t+2			t+2, t+4			t+2, t+13		
Region/Country	INDV	CR	Return	FF4-α	HKK4-α	Return	FF4-α	HKK4-α	Return	FF4-α	HKK4-α
North America			-0.25%	***-0.37%***	***-0.41%***	-0.20%	-0.29%	-0.34%	0.01%	-0.14%	-0.19%
Canada	80	1	-0.85%	-0.25%	-0.10%	-0.44%	0.15%	0.32%	-0.18%	0.04%	0.26%
United States	91	1	-0.41%	***-0.60%***	***-0.62%***	-0.30%	***-0.46%***	***-0.50%***	0.00%	-0.19%	-0.23%
European Developed Markets			***-1.31%***	***-0.97%***	***-0.96%***	***-1.18%***	***-0.86%***	***-0.83%***	***-0.93%***	***-0.75%***	***-0.70%***
Austria	55	3	***-1.94%***	-1.60%	***-2.07%***	***-1.95%***	-1.47%	***-1.96%***	-1.39%	-1.46%	-1.55%
Belgium	75	2	-1.09%	-0.68%	-0.39%	-1.14%	-0.69%	-0.48%	-0.85%	-0.67%	-0.52%
Denmark	74	3	***-1.39%***	-0.82%	-0.82%	***-1.25%***	-0.52%	-0.49%	-0.91%	-0.28%	-0.33%
Finland	63	1	***-1.57%***	***-1.52%***	-1.09%	***-1.30%***	***-1.27%***	-0.94%	***-2.06%***	***-2.00%***	***-1.78%***
France	71	0	***-1.33%***	***-1.16%***	***-1.13%***	***-1.18%***	***-1.02%***	***-0.95%***	***-0.96%***	***-0.83%***	***-0.75%***
Germany	67	3	***-1.42%***	-0.49%	-0.55%	***-1.40%***	-0.65%	-0.67%	***-1.00%***	-0.58%	-0.55%
Greece	35	1	-0.84%	***-1.48%***	-1.31%	-1.00%	***-1.34%***	-1.10%	-0.91%	-0.60%	-0.32%
Italy	76	2	***-0.97%***	***-0.88%***	-0.77%	***-0.89%***	***-0.79%***	-0.66%	***-0.91%***	***-0.85%***	***-0.81%***
Netherlands	80	3	***-1.42%***	***-1.25%***	***-1.30%***	***-1.09%***	-0.94%	***-1.05%***	-0.39%	-0.38%	-0.50%
Norway	69	2	-0.68%	-0.37%	-0.22%	-0.74%	-0.46%	-0.32%	-0.28%	-0.35%	-0.20%
Spain	51	2	***-1.54%***	-0.59%	-0.64%	***-1.31%***	-0.65%	-0.64%	***-1.18%***	-0.93%	-0.99%
Sweden	71	1	***-2.52%***	***-2.18%***	***-2.14%***	***-2.41%***	***-1.96%***	***-1.91%***	***-1.76%***	***-1.44%***	***-1.43%***
Switzerland	68	1	0.00%	0.31%	0.18%	-0.17%	0.05%	-0.17%	-0.19%	0.01%	-0.01%
United Kingdom	89	4	***-1.38%***	***-1.11%***	***-1.04%***	***-1.18%***	***-0.92%***	***-0.84%***	***-0.92%***	***-0.82%***	***-0.70%***
Japan	46	2	0.15%	0.00%	0.03%	0.15%	-0.05%	-0.02%	0.08%	-0.19%	-0.13%
Asia Pacific Developed Markets (Ex. Japan)			-0.18%	-0.21%	-0.04%	-0.07%	-0.11%	0.05%	0.24%	0.22%	0.33%
Australia	90	3	-0.81%	-0.38%	-0.40%	-0.52%	-0.13%	-0.09%	0.03%	0.33%	0.34%
Hong Kong	25	4	0.28%	0.27%	0.51%	0.16%	0.04%	0.22%	0.41%	0.26%	0.31%
New Zealand	79	4	***-1.99%***	***-2.37%***	***-2.56%***	-1.08%	-1.05%	-1.46%	-0.40%	-1.03%	-1.31%
Singapore	20	3	-0.77%	-0.62%	-0.79%	-0.70%	-0.63%	-0.78%	-0.38%	-0.15%	-0.37%

			Holding Periods								
			t+2			t+2, t+4			t+2, t+13		
Region/Country	INDV	CR	Return	FF4-α	HKK4-α	Return	FF4-α	HKK4-α	Return	FF4-α	HKK4-α
Emerging Markets			-0.40%	-0.41%	-0.41%	-0.39%	-0.35%	-0.33%	***-0.37%***	***-0.37%***	***-0.35%***
Argentina	46	1	2.59%	1.68%	1.26%	***2.32%***	2.08%	1.57%	***2.19%***	***2.78%***	2.36%
Brazil	38	1	1.07%	0.77%	0.96%	-0.14%	0.04%	0.14%	-0.19%	-0.20%	-0.21%
Chile	23	2	***-1.43%***	-0.99%	-1.17%	***-1.30%***	-0.98%	-1.21%	-0.88%	-0.80%	-0.91%
China	20	2	-0.25%	-0.13%	-0.20%	-0.34%	-0.24%	-0.28%	-0.17%	-0.22%	-0.26%
India	48	2	-0.58%	-0.54%	-0.49%	-0.78%	-0.71%	-0.68%	-0.58%	-0.54%	-0.54%
Indonesia	14	2	1.54%	1.89%	1.62%	0.82%	1.12%	1.17%	0.06%	0.59%	0.42%
Israel	54	3	-1.34%	-0.97%	-1.06%	***-2.01%***	-1.59%	-1.55%	-1.43%	-1.36%	-1.36%
Malaysia	26	3	-0.46%	-0.58%	-0.57%	-0.38%	-0.49%	-0.47%	-0.45%	-0.65%	-0.54%
Mexico	30	0	0.91%	-1.56%	-1.40%	2.03%	0.03%	0.23%	0.54%	-1.56%	-1.47%
Pakistan	14	1	0.55%	-0.12%	-0.25%	0.27%	0.04%	0.10%	0.98%	0.91%	0.81%
Philippines	32	1	0.55%	1.02%	1.20%	0.73%	0.97%	1.39%	1.03%	0.88%	1.33%
Poland	60	1	-1.16%	-0.15%	0.02%	-1.09%	-0.21%	-0.11%	***-1.51%***	-0.97%	-0.60%
South Africa	65	3	0.85%	1.45%	1.33%	1.11%	***1.96%***	***1.79%***	0.83%	***1.44%***	***1.39%***
South Korea	18	3	***-1.14%***	***-1.05%***	-0.46%	***-1.03%***	***-0.98%***	-0.39%	***-1.27%***	***-1.38%***	-0.83%
Taiwan	17	2	-0.22%	-0.15%	-0.16%	-0.40%	-0.42%	-0.46%	-0.47%	-0.64%	-0.67%
Thailand	20	2	0.04%	-0.51%	-0.59%	-0.17%	-0.32%	-0.24%	0.14%	0.15%	0.17%
Turkey	37	2	0.40%	-0.45%	-0.26%	-0.29%	-0.50%	-0.53%	-0.55%	-0.90%	-0.70%

Table 4: Risk Characteristics by EDF Decile for non-U.S. Stocks

This table presents the average value of the following firm characteristics at the end of portfolio formation month t for EDF decile portfolios of all non-U.S. countries: 1) standard deviation of monthly returns (SD), calculated from most recent 12 months; 2) Leverage, calculated as the sum of long term debts and short debts divided by total assets; 3) natural log of market capitalization (Size) at the month end; 4) book-to-market equity ratio (B/M), calculated as the ratio of book equity of the fiscal year ending at least 6 months before month t to the market capitalization at end of month t; 5) most recent 11-month stock return in U.S. dollars up to month t (MMT); and 6) hybrid CAPM beta (β_hybrid), the sum of loadings on local and global market returns in a hybrid version of CAPM, which includes both value-weighted average returns of all stocks in the stock's country and value-weighted average returns of all stocks in all countries other than the stock's country as factors. We use the past 60-month data and require a minimum of 24 months of valid data for beta estimation. We first obtain the average value of a firm characteristic across all stocks in a portfolio each month, and then further average across all months to obtain the final average value of the firm characteristic for the portfolio. We also report the difference in each firm characteristic between EDF decile 10 and decile 1, and the associated t-statistic.

	EDF deciles											
Variable	1 (lowest)	2	3	4	5	6	7	8	9	10 (highest)	Dif. (10 - 1)	T-stat.
N (stocks)	1183	1203	1218	1224	1217	1227	1220	1217	1214	1205		
SD	9.29%	10.46%	11.21%	11.84%	12.36%	12.97%	13.58%	14.27%	15.17%	17.05%	7.76%	68.90
Leverage	0.12	0.17	0.19	0.20	0.21	0.23	0.24	0.27	0.29	0.36	0.24	164.68
Size	6.32	6.12	5.83	5.57	5.30	5.02	4.73	4.41	4.07	3.62	-2.70	-92.26
B/M	0.81	0.83	0.95	1.04	1.05	1.06	1.12	1.23	1.28	1.50	0.69	16.85
MMT	0.24	0.24	0.22	0.20	0.17	0.14	0.11	0.07	0.02	-0.08	-0.32	-39.73
β_hybrid	0.75	0.83	0.89	0.93	0.96	0.99	1.02	1.04	1.06	1.08	0.33	35.35

Table 5: Creditor Rights and the Distress Anomaly

This table presents the average returns and alphas from both Fama-French-Carhart and Hou-Karolyi-Kho four-factor models for EDF-based long/short hedging portfolios that formed among small stocks in groups of countries with different levels of La Porta, Lopez-de-Silanes, Shleifer, and Vishny's (1998) Creditor Rights index. The procedure is as follows. First, at the end of each month *t*, we rank all small stocks within each country by EDF, and then form value-weighted decile portfolios for each country. We implement a long-short strategy involving EDF-deciles 10 (long) and 1 (short) for various holding periods that begin one month after portfolio formation: month (t+2), three-month (t+2, t+4), and one year (t+2, t+13). For holding periods greater than one month, we follow the overlapping horizon approach of Jegadeesh and Titman (1993). All countries in our sample are classified into two groups (low and high) based on the median index value. Each month, the returns of country-level long/short hedging portfolios are averaged across all countries in an index group to obtain the return of long/short hedging portfolios for the index group in that month. The final row shows the results of testing the difference between the returns of high and low index groups.

		Holding Periods								
		t+2			t+2, t+4			t+2, t+13		
Country Groups	Stat.	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α
Low CR	Est	-0.47%	-0.37%	-0.35%	-0.38%	-0.30%	-0.26%	-0.38%	-0.35%	-0.29%
	t-stat	*-1.77*	*-1.55*	-1.47	*-1.52*	*-1.34*	-1.15	*-1.74*	*-1.73*	-1.46
High CR	Est	-0.56%	-0.50%	-0.52%	-0.51%	-0.50%	-0.51%	-0.39%	-0.44%	-0.46%
	t-stat	*-2.94*	*-3.17*	-3.30	*-2.85*	*-3.37*	-3.43	*-2.38*	*-3.11*	-3.20
High – Low CR	Est	-0.09%	-0.13%	-0.16%	-0.13%	-0.19%	-0.24%	-0.01%	-0.10%	-0.16%
	t-stat	*-0.35*	*-0.49*	-0.63	*-0.59*	*-0.80*	-1.01	*-0.04*	*-0.47*	-0.80

Table 6: Investor Individualism and the Distress Anomaly

This table presents the average returns and alphas from both Fama-French-Carhart and Hou-Karolyi-Kho four-factor models for EDF-based long/short hedging portfolios that formed among small stocks in groups of countries with different levels Hofstede's (2001) Individualism index. The procedure is as follows. First, at the end of each month *t*, we rank all small stocks within each country by EDF, and then form value-weighted decile portfolios for each country. We implement a long-short strategy involving EDF-deciles 10 (long) and 1 (short) for various holding periods that begin one month after portfolio formation: month (t+2), three-month (t+2, t+4), and one year (t+2, t+13). For holding periods greater than one month, we follow the overlapping horizon approach of Jegadeesh and Titman (1993). All countries in our sample are classified into two groups (low and high) based on the median index value. Each month, the returns of country-level long/short hedging portfolios are averaged across all countries in an index group to obtain the return of long/short hedging portfolios for the index group in that month. The final row shows the results of testing the difference between the returns of high and low index groups.

		Holding Periods								
		t+2			t+2, t+4			t+2, t+13		
Country Groups	Stat.	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α
Low Indv	Est	-0.03%	-0.05%	-0.08%	-0.08%	-0.15%	-0.15%	-0.18%	-0.28%	-0.30%
	t-stat	*-0.14*	*-0.24*	*-0.41*	*-0.39*	*-0.87*	*-0.88*	*-0.96*	*-1.68*	*-1.81*
High Indv	Est	-0.92%	-0.76%	-0.74%	-0.76%	-0.63%	-0.61%	-0.54%	-0.51%	-0.48%
	t-stat	*-3.99*	*-3.85*	*-3.74*	*-3.54*	*-3.40*	*-3.29*	*-2.80*	*-2.91*	*-2.71*
High - Low Indv	Est	-0.89%	-0.72%	-0.67%	-0.68%	-0.48%	-0.46%	-0.36%	-0.23%	-0.18%
	t-stat	*-3.54*	*-2.77*	*-2.57*	*-2.93*	*-2.02*	*-1.93*	*-1.73*	*-1.08*	*-0.83*

Table 7: Market States and the Distress Anomaly

This table presents the results of time-series cross-sectional regressions of monthly stock returns on EDF and its interactions with market states and turnover, controlling for other firm characteristics. We rank stocks into EDF decile portfolios within each country-month cohort. High EDF and Low EDF are dummy variables that indicate whether a stock belongs to EDF decile 10 and decile 1 at the end of month t-2, respectively, within each country-month cohort. Up Mkt(t) is a dummy variable that equals one if a country's past 12-month cumulative return up to month t is both positive and above the median value for all months in our sample for that country, and zero otherwise. Down Mkt(t) equals one minus Up Mkt(t). In Panel A, we interact EDF dummy variables with Up Mkt(t-2), the market state indicator in our portfolio formation month. In Panel B, we interact EDF dummy variables with all four combinations of market states in both portfolio formation month t-2 and return measurement month t. In Panel C, we interact EDF dummy variables with both Up Mkt(t-2), and High Turnover, a dummy variable that equals one if a stock's turnover in month t-2 is above the median value for all sample stocks in its country, and zero otherwise. The dependent variable is the stock return in month t, denominated in U.S. dollars. Other firm characteristics include: 1) market capitalization (ME) at the end of month (t-1); 2) book-to-market equity ratio (B/M) at the end of month (t-1), calculated as the ratio of book equity of the fiscal year ending at least 6 months before month (t-1) to the market capitalization at end of month (t-1); 3) past 11-month stock return in U.S. dollars (MMT) from month (t-12) to (t-2); 4) past 1-month stock return in U.S. dollars (LRet) in month (t-1); 5) share turnover (Turnover) in month (t-1), calculated as the number of shares traded divided by the number of shares outstanding in month (t-1). All control variables enter as decile rankings (i.e., 1 for the bottom decile and 10 for the highest decile) within each country-month cohort. Country and month fixed effects are also included in the regression. Estimates of controls are not reported in Panels B for brevity. All t-statistics (indicated below in italics) are calculated based on standard errors clustered by country.

Panel A: Prior Market State	
High EDF	-0.05
	-0.76
Low EDF	-0.11
	-2.58
High EDF x Up Mkt (t-2)	-0.34
	-3.63
Low EDF x Up Mkt (t-2)	0.37
	7.79
Up Mkt (t-2)	0.28
	1.65
Controls:	
ME	0.01
	0.82
B/M	0.14
	12.98
MMT	0.18
	8.04
LRET	-0.07
	-4.81
Turnover	0.02
	1.20
Fixed effects	Yes
Adj. R2	0.13
Obs.	3,791,709

Panel B: Prior and Current Market States	
High EDF x Up Mkt (t-2) x Up Mkt (t)	-0.22
	-3.19
High EDF x Up Mkt (t-2) x Down Mkt (t)	-1.16
	-7.61
High EDF x Down Mkt (t-2) x Up Mkt (t)	0.39
	2.29
High EDF x Down Mkt (t-2) x Down Mkt (t)	-0.16
	-2.26
Low EDF x Up Mkt (t-2) x Up Mkt (t)	0.11
	2.12
Low EDF x Up Mkt (t-2) x Down Mkt (t)	0.96
	6.36
Low EDF x Down Mkt (t-2) x Up Mkt (t)	-0.62
	-6.10
Low EDF x Down Mkt (t-2) x Down Mkt (t)	0.01
	0.14
Up Mkt (t-2) x Up Mkt (t)	1.58
	5.83
Up Mkt (t-2) x Down Mkt (t)	-1.27
	-3.16
Down Mkt (t-2) x Up Mkt (t)	2.17
	4.13
Controls & Fixed effects	Yes
Adj. R2	0.13
Obs.	3,791,709

Panel C: Prior Market State and Turnover	
High EDF	0.11
	1.33
Low EDF	-0.18
	-3.42
High EDF x Up Mkt (t-2)	-0.17
	-1.99
Low EDF x Up Mkt (t-2)	0.34
	5.81
High EDF x Up Mkt (t-2) x High Turnover	-0.39
	-2.74
Low EDF x Up Mkt (t-2) x High Turnover	0.01
	0.08
High EDF x High Turnover	-0.26
	-2.22
Low EDF x High Turnover	0.08
	2.00
Up Mkt (t-2)	0.48
	2.86
Up Mkt (t-2) x High Turnover	-0.38
	-6.25
High Turnover	0.11
	2.92
Controls & Fixed effects	Yes
Adj. R2	0.13
Obs.	3,791,700

Table 8: Fama-MacBeth Regressions of Monthly Stock Returns on EDF, News Sentiment and Distress Duration for U.S. Small Stocks

This table presents the results of Fama-MacBeth regressions of monthly stock returns on the distressed dummy and its interactions with negative news dummy and newly distressed dummy, controlling for other firm characteristics, for the sample of U.S. small stocks. At the end of each month, we rank all U.S. small stocks by EDF into ten deciles, where the market capitalization cutoff value for large versus small stocks is the NYSE median value. High EDF and Low EDF are dummy variables that indicate whether a firm belongs to EDF decile 10 and decile 1 at the end of month (t-2), respectively. For each stock in EDF decile 10 in each month, we calculate its average ESS score from its news stories in the month and its distress duration (the number of months having stayed in EDF decile 10). Then at the end of each month, we rank all stocks in EDF decile 10 into three equally-sized groups based on the average ESS score and the distress duration respectively. Negative News and Newly Distressed are dummy variables that indicate whether, at the end of month t-2, a stock belongs to EDF decile 10 and is among the bottom tertile based on the average ESS score and the distress duration respectively. We interact High EDF with both Negative News and Newly Distressed. The dependent variable is the stock return in month t, denominated in U.S. dollars. Other firm characteristics include: 1) market capitalization (ME) at the end of month (t-1); 2) book-to-market equity ratio (B/M) at the end of month (t-1), calculated as the ratio of book equity of the fiscal year ending at least 6 months before month (t-1) to the market capitalization at end of month (t-1); 3) past 11-month stock return in U.S. dollars (MMT) from month (t-12) to (t-2); 4) past 1-month stock return in U.S. dollars (LRet) in month (t-1); 5) share turnover (Turnover) in month (t-1), calculated as the number of shares traded divided by the number of shares outstanding in month (t-1). All control variables enter as decile rankings (i.e., 1 for the bottom decile and 10 for the highest decile) within each month. Coefficients reported are time-series averages of the cross-sectional regressions shown above, with t-statistics calculated based on the Newey-West heteroskedasticity and autocorrelation consistent (HAC) standard errors of these estimates.

	(1)	(2)	(3)
Intercept	-0.03	-0.03	-0.03
	-0.04	*-0.04*	*-0.05*
High EDF	0.06	0.03	0.21
	0.24	*0.11*	*0.77*
Low EDF	0.08	0.08	0.08
	0.53	*0.54*	*0.55*
High EDF x Negative News	-0.80		-0.81
	-2.99		*-3.01*
High EDF x Newly Distressed		-0.51	-0.51
		-2.26	*-2.27*
ME	0.03	0.03	0.03
	0.65	*0.64*	*0.65*
B/M	0.11	0.11	0.11
	2.73	*2.73*	*2.76*
MMT	0.10	0.10	0.10
	2.32	*2.31*	*2.30*
LRET	-0.11	-0.11	-0.11
	-3.11	*-3.09*	*-3.11*
Turnover	-0.01	-0.01	-0.01
	-0.13	*-0.15*	*-0.13*
Fixed effects	Yes	Yes	Yes
Adj. R2	0.06	0.06	0.06
Avg. obs.	2,228	2,228	2,228

Appendix I
Global Equity Universe in COMPUSTAT

In this section we present the distribution of COMPUSTAT's global equity universe across regions. For ease of comparison with the typical global equity sample used in recent studies, we follow the sample selection process in Karolyi and Wu (2012) for this analysis. We restrict the sample countries to 46 countries and the sample period 1990 to 2010 as in Karolyi and Wu (2012).[22] The U.S. and Canadian stocks are sampled from COMPUSTAT North America, and the stocks in all other countries are from COMPUSTAT Global. Following filters in Karolyi and Wu (2012), a stock must satisfy the following criteria to be included: 1) being listed on its country's major exchange(s); 2) being a non-financial firm; 3) having at least 12 monthly returns over the sample period; and 4) having sufficient information to calculate either book-to-market (B/M) ratio or cash flow-to-price ratio.

Figure AI.1 presents the sample distribution of aggregate market capitalization (Panel A) and total number of stocks (Panel B) by region over the period 1990 to 2010. Based on the tabulation of total market capitalization by region, it seems that COMPUSTAT's global sample is similar to the sample studied in Karolyi and Wu (2012), drawn from Datastream/Worldscope. Overall, our sample matches with theirs closely. For example, US stocks comprise 43% in our sample versus 41% in Karolyi and Wu (2012); for Europe (25% versus 26%) and Japan (14% versus 13%), similar comparisons obtain. Figure AI.2 presents Compustat's global sample distribution by region-year in market capitalization (Panel A) and number of stocks (Panel B). The overall patterns are broadly consistent with Karolyi and Wu (2012).

In addition to comparing sample distribution by region, in untabulated analysis, we also benchmark the stock characteristics by country for our global equity universe against those reported in Karolyi and Wu (2012). Our sample achieves very close match with the benchmark

[22] Besides 38 countries listed in Table 1 in the paper, the other countries include: Portugal and Ireland in European developed markets, and Colombia, Czech Republic, Hungary, Peru, Russia, and Venezuela in emerging markets.

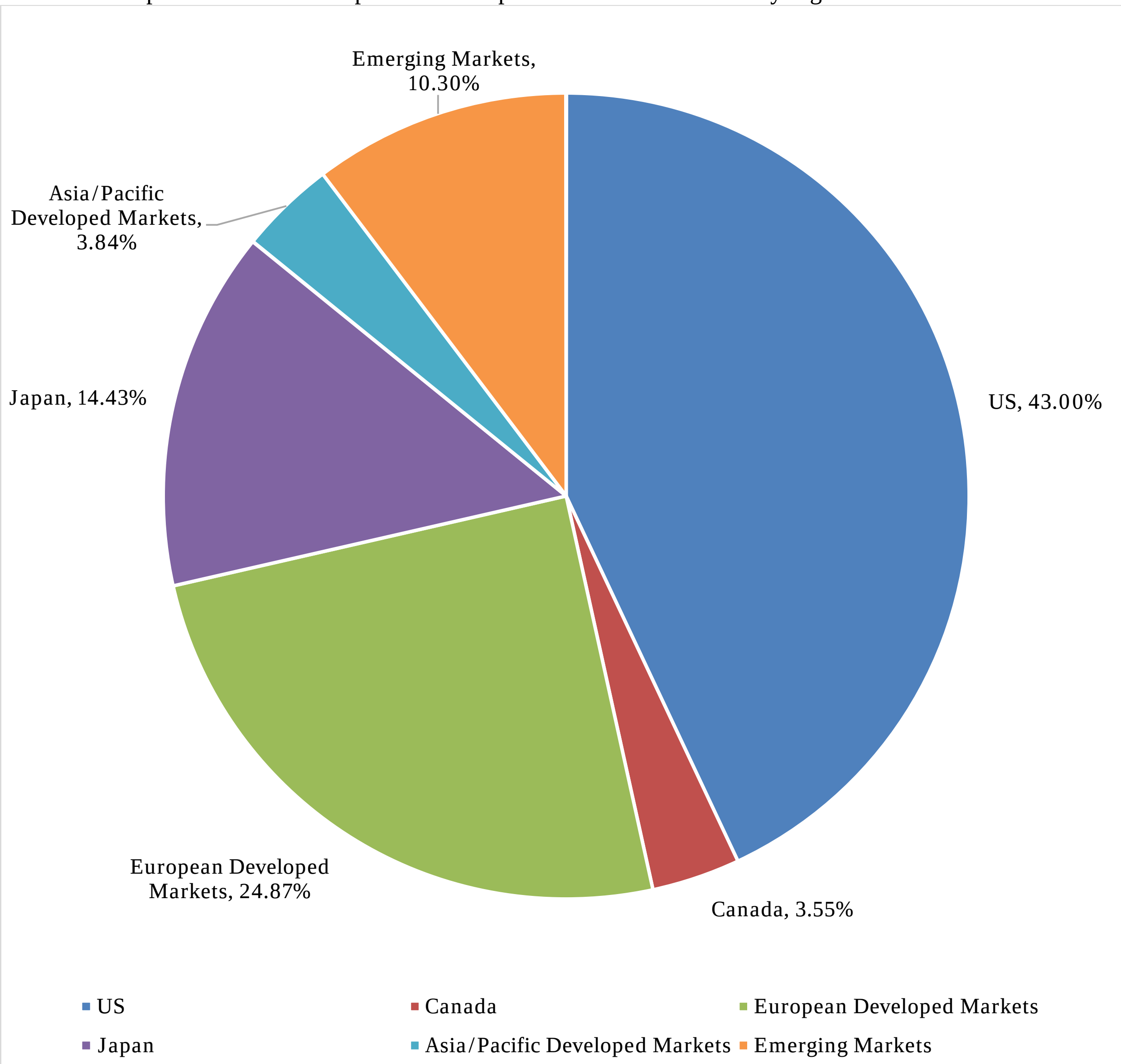
sample in Karolyi and Wu (2012) in terms of C/P ratios and B/M ratios. For example, in European developed markets, there is virtually no difference among these two samples for France, Germany, and UK in C/P ratio. Across all countries in European developed markets, the difference in C/P ratio is only about 0.02 between these two samples. The difference in B/M ratio is also small, of the similar magnitude.

There is a difference in average market capitalization and consequently momentum strategy returns. This is mainly driven by the fact that Compustat Global sample's coverage tilts towards larger firms in the earlier sample period. The coverage for smaller firms has improved dramatically post-1992, when our sample begins.

Figure AI.1: Compustat Global's sample distribution by region

This figure presents the sample distribution of market capitalization (Panel A) and number of stocks (Panel B) by region.

Panel A: Compustat Global's sample market capitalization distribution by region



Panel B: Compustat Global's sample number of stocks distribution by region

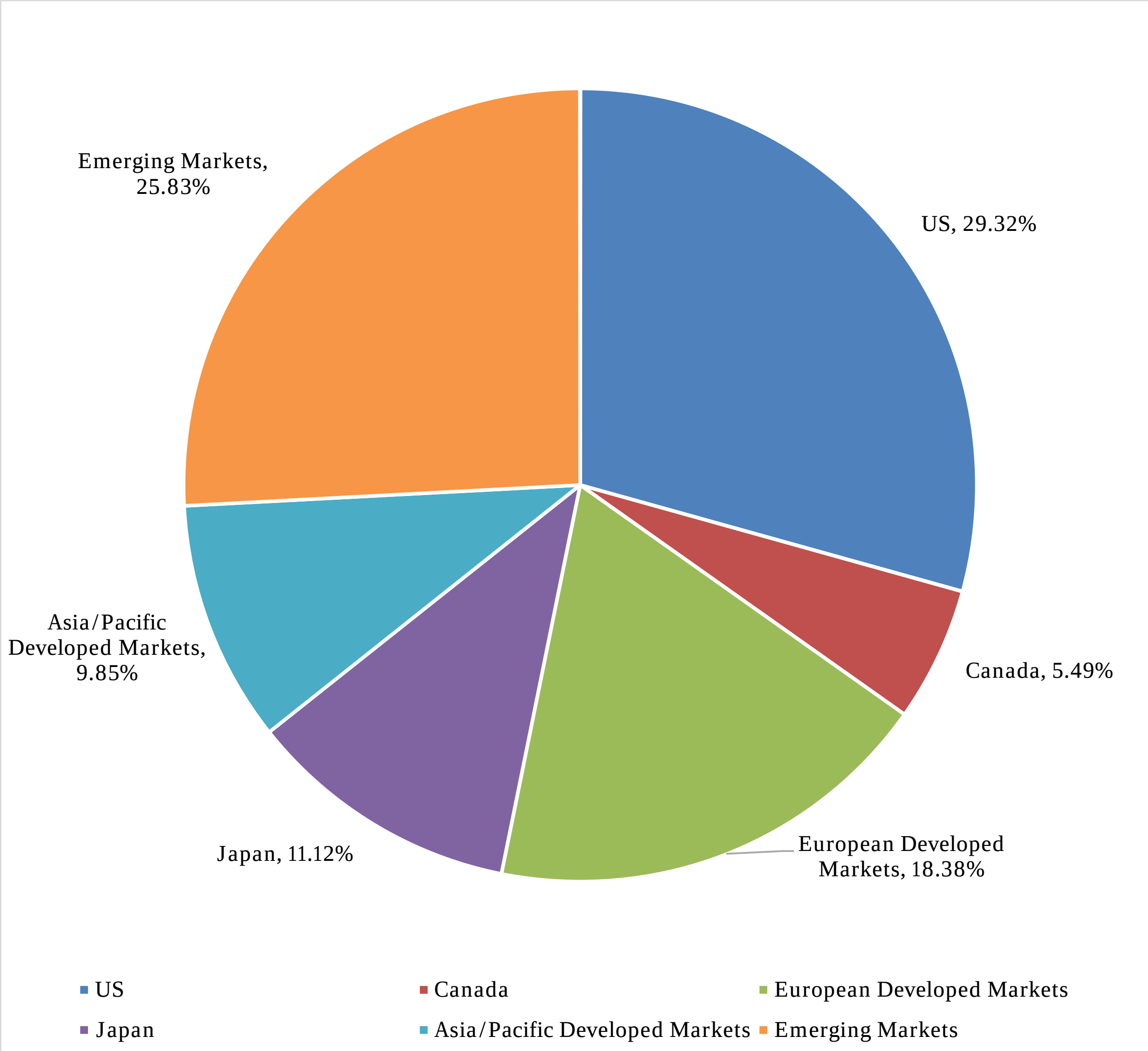
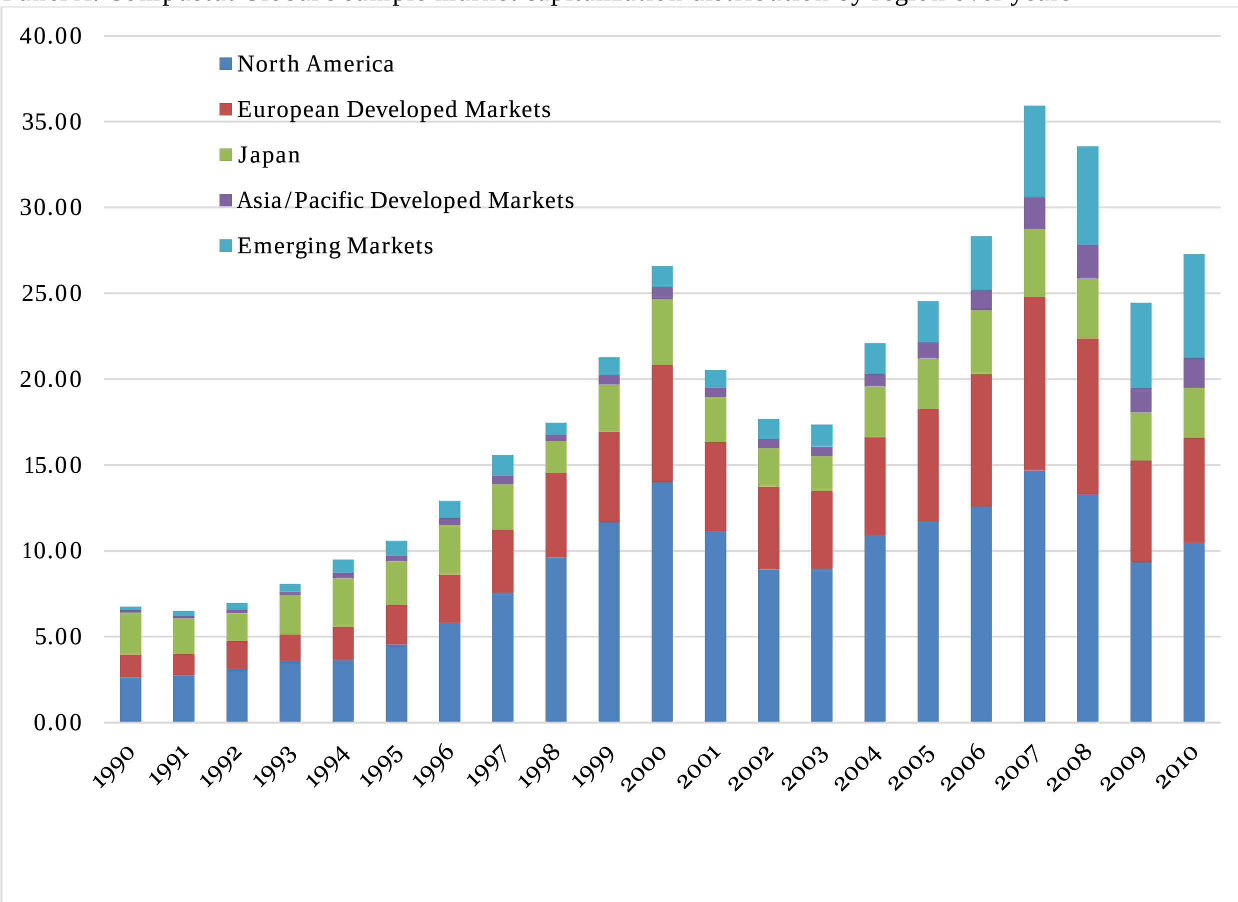


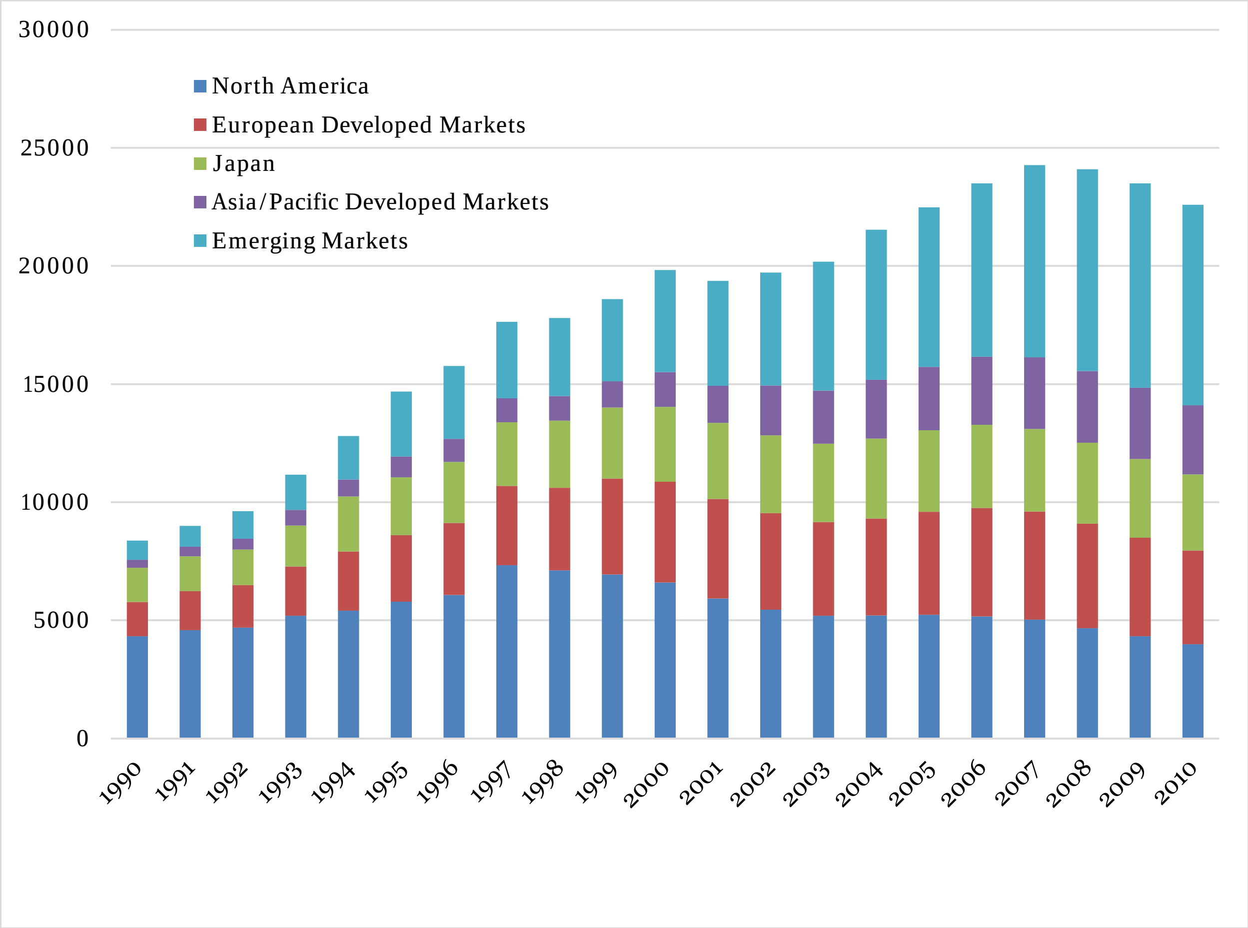
Figure AI.2: Compustat Global's sample distribution by region over years

This figure presents the sample distribution of market capitalization (Panel A) and number of stocks (Panel B) by region over the period between 1990 and 2010.

Panel A: Compustat Global's sample market capitalization distribution by region over years



Panel B: Compustat Global's sample number of stocks distribution by region over years



Appendix II

Factor Model Specifications s and Construction of Factors

We follow Fama and French (2012), Hou, Karolyi, and Kho (2011) and Karolyi and Wu (2012), to construct both global and local (at region or country level) factors. For the purpose of factor construction, our global equity universe includes all common stocks listed on the major exchanges in 46 countries (please refer to Appendix I). The returns of all factor-mimicking portfolios are denominated in USD.

The global market return factor, $R^G_{m,t} - R_{RF,t}$, is measured as month t's return on the value-weighted portfolio of all stocks in our global equity universe, in excess of the 3-month US treasury rate. Similarly, local market return factor, $R^L_{m,t} - R_{RF,t}$, is measured by the excess return on the value-weighted portfolio of all stocks in the respective local market (a region or country).

To construct factor-mimicking portfolios for the size ($F^L_{Size,t}$) and B/M ($F^L_{B/M,t}$) factors of a local market (i.e., a region or country), at the end of June of each year y, we perform a 2×3 sort within the local market based on size and book-to-market ratio. We use the market capitalization at the end of June for size classification, and the ratio of book equity at the end of the fiscal year ending in year y-1 to the market capitalization at the end of year y-1 for market-to-book construction. Stocks are classified as large versus small based on the rule that large stocks cumulatively account for at least 90% of the aggregate market capitalization of the local market.[23] The cut-off values for low and high book-to-market groupings are the 30th and 70th percentile values of book-to-market within the large stocks in the local market. The use of book-

[23] The common approach to construct the size factor (small minus big) for the U.S. market is to classify stocks based on the NYSE median market capitalization. For North America (the U.S. and Canada), the cut-off value for 90% of total market capitalization roughly corresponds to the NYSE median market capitalization according to Fama and French (2012), which is confirmed by our independent check. When applying this rule to a country with small number of stocks, we make sure that the large stock group should include at least 20% of total number of stocks in the country (if the 90% aggregate market capitalization rule results in less number of stocks) and no more than 50% of total number of stocks in the country (if the 90% aggregate market capitalization rule results in more number of stocks).

to-market breakpoints among large stocks follows Fama and French (1993, 2012), and is intended to mitigate the influence of micro-cap stocks.

We apply these size and book-to-market classification schemes for the period July in year *y* through June in year *y+1*, and form six value-weighted portfolios from the intersection of size and book-to-market sorts: SH for small/high book-to-market stocks, SM for small/medium book-to-market stocks, SL for small/low book-to-market stocks, BH for big/high book-to-market stocks, BM for big/medium book-to-market stocks, and BL for big/low book-to-market stocks. The size factor in month *t*, $F^L_{Size,t}$, is measured as the return on a hedged trading strategy involving the six size/book-to-market portfolios, (SH+SM+SL)/3 - (BH+BM+BL)/3. B/M factor ($F^L_{B/M,t}$), is measured as the return on a hedged trading strategy formed in a similar way, (SH + BH)/2 - (SL+BL)/2.

Mimicking portfolios for local C/P factor ($F^L_{C/P,t}$) is constructed in an almost identical way by sorting stocks in the local market at the end of June in each year *y* based on size and cash-to-price (C/P) ratio. We calculate C/P ratio from the cash flow of the fiscal year ending in year *y*-1 to the market capitalization at the end of year *y*-1 for market-to-book construction. C/P factor ($F^L_{C/P,t}$), is measured as the return on a hedged trading strategy long on small/high-C/P stocks and big/high-C/P stocks and short on small/low-C/P stocks and big/low-C/P stocks.

Similarly, to construct mimicking portfolios for local momentum factor ($F^L_{Mom,t}$), in each month *t*, we conduct a 2×3 sort within the local market based on size (measured in the same way as above) and past 11-month returns over months *t-12* through *t-2*. Cut-off values for losers and winners are the 30th percentile and 70th percentile values of the past returns within the large stock group. Six value-weighted intersecting portfolios are formed accordingly: SU for small/past winning stocks, SN for small/past neutral stocks, SD for small/past losing stocks, BU for big/past winning stocks, BN for big/past neutral stocks, and BD for big/past losing stocks.

Momentum factor ($F^{L}_{Mom,t}$), is measured as the return on a hedged trading strategy long on past winning stocks and short on past losing stocks, (SU + BU)/2 - (SD+BD)/2.

Mimicking portfolios for global size, B/M, C/P and momentum factors ($F^{G}_{Size,t}$, $F^{G}_{B/M,t}$, $F^{G}_{C/P,t}$, and $F^{G}_{Mom,t}$) are constructed in very similar ways as those for local factors, by performing a 2×3 sort within our global equity universe based on size and the relevant sorting variable, B/M, C/P and past return respectively. Stocks are classified as large versus small based on the rule that large stocks cumulatively account for at least 90% of the global aggregate market capitalization. The cut-off values of B/M, C/P and past return are regional break points, i.e., the 30th and 70th percentile values of the relevant sorting variable within each region. Finally, monthly global factors are measured as the return on those hedged trading strategies involving intersecting portfolios, in the same way as described in the local factor construction.

When evaluating the returns on a global trading strategy, we employ the following two factor models including global factors only:

$$R_{i,t} - R_{RF,t} = \alpha_i + \beta^{G}_{i}\left(R^{G}_{m,t} - R_{RF,t}\right) + c^{G}_{i} F^{G}_{\frac{C}{P},t} + m^{G}_{i} F^{G}_{Mom,t} + s^{G}_{i} F^{G}_{Size,t} + \varepsilon_{i,t} \tag{1}$$

$$R_{i,t} - R_{RF,t} = \alpha_i + \beta^{G}_{i}\left(R^{G}_{m,t} - R_{RF,t}\right) + b^{G}_{i} F^{G}_{\frac{B}{M},t} + m^{G}_{i} F^{G}_{Mom,t} + s^{G}_{i} F^{G}_{Size,t} + \varepsilon_{i,t} \tag{2}$$

where $R_{i,t} - R_{RF,t}$ is excess return of the trading strategy in question, and all global factors are constructed following the exact procedure described above with all stocks in our global equity universe. For simplicity, we call model (1) the global Hou-Karolyi-Kho (HKK) four-factor model, and model (2) the global Fama-French-Carhart (FFC) four-factor model.

When evaluating the returns on a trading strategy using only stocks in a local market (either a region or a country), we employ the following hybrid factor models including both global factors and local factors, in the spirit of Hou, Karolyi, and Kho (2011) and Karolyi and Wu (2012):

$$R_{i,t} - R_{RF,t} = \alpha_i + \beta^{L}_{i}\left(R^{L}_{m,t} - R_{RF,t}\right) + \beta^{G}_{i}\left(R^{G}_{m,t} - R_{RF,t}\right) + c^{L}_{i} F^{L}_{\frac{C}{P},t} + c^{G}_{i} F^{G}_{\frac{C}{P},t} \tag{3}$$

$$+m_i^L F_{Mom,t}^L + m_i^G F_{Mom,t}^G + s_i^L F_{Size,t}^L + s_i^G F_{Size,t}^G + \varepsilon_{i,t}$$

$$\begin{aligned} R_{i,t} - R_{RF,t} = \alpha_i + \beta_i^L\left(R_{m,t}^L - R_{RF,t}\right) + \beta_i^G\left(R_{m,t}^G - R_{RF,t}\right) + b_i^L F_{\frac{B}{M},t}^L + b_i^G F_{\frac{B}{M},t}^G \\ +m_i^L F_{Mom,t}^L + m_i^G F_{Mom,t}^G + s_i^L F_{Size,t}^L + s_i^G F_{Size,t}^G + \varepsilon_{i,t} \end{aligned} \quad (4)$$

The local factors are constructed following the exact procedure described above with all stocks in the local market in question, and the global factors are constructed following the procedure described above with all stocks in our global equity, excluding those stocks in the local market in question. For differentiation, we call model (3) the hybrid Hou-Karolyi-Kho (HKK) four-factor model, and model (4) the hybrid Fama-French-Carhart (FFC) four-factor model.

Appendix III
RavenPack News Data

RavenPack aggregates news from both traditional news media (newswires service) and social-media (web sites), and provides algorithm-based news analytical service. Since RavenPack is a relatively new database, we discuss our sample construction and selection process in details. We also discuss the limitation of the database in the context of studying news media in the international markets.

Linking RavenPack to Compustat

RavenPack uses its own entity code (*RP Entity ID*) to uniquely identify a company. It also provides an entity mapping file, which includes company names, country of incorporation, and third-party identifiers such as exchange ticker, International Securities Identification Number (ISIN), CUSIP, and SEDOL, among others. We map *RP Entity ID* and GVKEY from the Compustat (both North America and Global files) based on the following procedure. We first use ISIN as the primary identifier to obtain matches of *RP Entity ID* and GVKEY. We require a match to be unique for both identifiers. In the second step, for those cases we cannot find any match based on the primary identifiers, we then proceed with the similar mapping procedure by using other security identifiers in the following order: 9-digit CUSIP, SEDOL, and finally 6-digit CUSIP.

Filtering news stories

RavenPack provides two versions of data: (1) the Dow Jones Edition, which derives contents from Dow Jones Newswires, Barron's, and the Wall Street Journal; and (2) the Web Edition, which covers major publishers, government and regulatory agencies, and local and regional newspapers. For our analysis, we use the Dow Jones Edition, which starts from 2000. We do not use the Web Edition for two reasons. First, it only starts from 2007; and second, the Web Edition does not significantly improve the coverage for our sample.

A piece of news often mentions a firm that is not the primary focus of the news. To measure the relevance of the story to a firm that is mentioned in the story, RavenPack provides so-called "relevance score" that ranges between 1 and 100. According to the vendor, "usually, a relevance value of at least 90 indicates that the entity is referenced in the main title or headline of the news item, while lower values indicate references further down the story body." (RavenPack User Guide, 2012). To focus on the firm that is truly the focus of the news, we include only news stories with the relevance score of 100, which errs on the conservative side.

To ensure that news stories actually cover a company's news events rather than market movements, we exclude news stories falling in the following categories: stock-gain, stock-loss, market-close-buy-imbalance, market-close-sell-imbalance, no-market-close-imbalance, market-open-sell-imbalance, market-open-buy-imbalance, delay-imbalance, buy-imbalance, sell-imbalance, and no-imbalance. This procedure makes use of "news tag" created by RavenPack, which recognizes a particular type and property of an entity-specific news event.

To remove duplicated news for the same event of a company, we include only news stories with the Event Novelty Score (ENS) of 100, which means that no similar stories about the company has been reported in the past 24 hour window.[24]

Construction of news variables

We use RavenPack's Event Sentiment Score (ESS), which ranges between 0 and 100, to measure news sentiment. According to RavenPack, a score of 50 indicates neutral sentiment,

[24] RavenPack identifies a chain of related news about the same categorized event for the same entities by the Event Novelty Key (*ENS_KEY*). According to RavenPack's user guide: "The first story reporting a categorized event about one or more entities is considered to be the most novel and receives a score of 100. Subsequent stories about the same event for the same entities receive scores following a decay function whose values are (100 75 56 42 32 24 18 13 10 8 6 4 3 2 2 1 1 1 1 0 ...) based on the number of stories in the past 24 hour window." If a news story is published more than 24 hours after any other similar story, it will be treated as starting a separate chain of stories (with a new ENS_KEY) and receive an ENS score of 100. Overall, the intuition behind this procedure is similar to the construction of "chunky news" in Da, Engelberg and Gao (2011).

whereas a score above (below) 50 stands for positive (negative) sentiment.[25] For our baseline analysis, we adjust ESS to have a range between -1 and +1 as:

$$\text{Adjusted ESS} = (\text{ESS} - 50)/50 \qquad \text{(AIII.1)}$$

To measure the news sentiment for a company over a month, we take the average adjusted ESS across all news stories for this company in the month. If a firm does not have any news story with valid ESS in a month, then the news sentiment for the firm-month is set as *missing*. That is, the company (i)'s sentiment score in month (t) is calculated as

$$\text{Sentiment}_{i,t} = \frac{1}{N_{i,t}} \sum_{j=1}^{j=N_{i,t}} \text{Adjusted ESS}_{j,i,t} \qquad \text{(AIII.2)}$$

where Adjusted $\text{ESS}_{j,i,t}$ is based on definition above, $N_{i,t}$ is the total number of news that satisfies the inclusion criteria outlined earlier.

In our regression analysis reported in Table 8, we create a dummy variable, negative news, which takes a value of 1 if the news sentiment of a firm in a month is in the lowest tercile among all firms with sentiment score values in a country and in that month, and zero otherwise. The results are similar if a firm-month is classified into negative news portfolio if its news sentiment is below the median value or in the lowest quintile among all firms in the country in that month.

Although overall the news sentiment, gauged by RavenPack's ESS, tends to biased upward as shown in Figure 5 in the paper, please note that our regression results should not be affected by this bias, as we identify negative news firms based on relative rankings of sentiment score. Nevertheless, we also measure the company (i)'s sentiment score in month (t) based on an alternative definition that is not affected by the overall bias in news tone:

[25] Existing studies using RavenPack news data, such as Kolasinski, Reed, and Ringgenberg (2013) and Dang, Moshirian, and Zhang (2015), normally follow RavenPack's classification of positive versus negative news stories.

$$\text{Sentiment}_{i,t} = \frac{\#\text{ of positive stories} - \#\text{ of negative stories}}{\#\text{ of positive stories} + \#\text{ of negative stories}} \tag{AIII.3}$$

where we identify a news story as positive (negative) when its ESS is above 67th (below 33th) percentile value of all stories published for a country's sample firms during our sample period 2000 to 2012. This alternative sentiment measure yields almost identical regression results, as shown in Table AIV.2, as the baseline ones reported in Table 8.

Limitation of RavenPack database

One significant issue with using RavenPack news data for an international study is its very limited coverage of firms, especially small ones, outside North America, which is apparent in Table AIII.1. Even benchmarked against developed markets in Europe, data on news production for North American firms is almost an order of magnitude more likely to be recorded. Larger disparities are observed for emerging markets. Further as shown in Figure AIII.1, even though the coverage of RavenPack database improved over time, the gap between North America and other regions remained.

Part of the reason behind the low news coverage outside North America is related to the development of news intermediary markets across the globe. However, it is also related to RavenPack's data collection and analytical procedure, which are primarily based on English-language textual contents.[26] Outside U.S., the news media in English tends to cover a small and rather selected sample of stocks, such as large firms listed or cross-listed at major stock exchanges.

The sample of publicly listed stocks of Chinese firms serves as a good example. In the RavenPack data, there are 521 firms with country of incorporation identified as China. However, there are only 78 cases where the firm is only listed in one of the main exchanges in China based

[26] For example, see an interview with Jason S. Cornez, Chief Technology Officer, RavenPack, at http://www.odbms.org/blog/2016/05/on-data-analytics-for-finance-interview-with-jason-s-cornez/

on stock ticker match. In many of the remaining cases, they are listed or cross-listed on at least one stock exchange outside the mainland China. The most frequently listed or cross-listed exchange is the Stock Exchange of Hong Kong (SEHK), followed by the Stock Exchange of Singapore (SES).

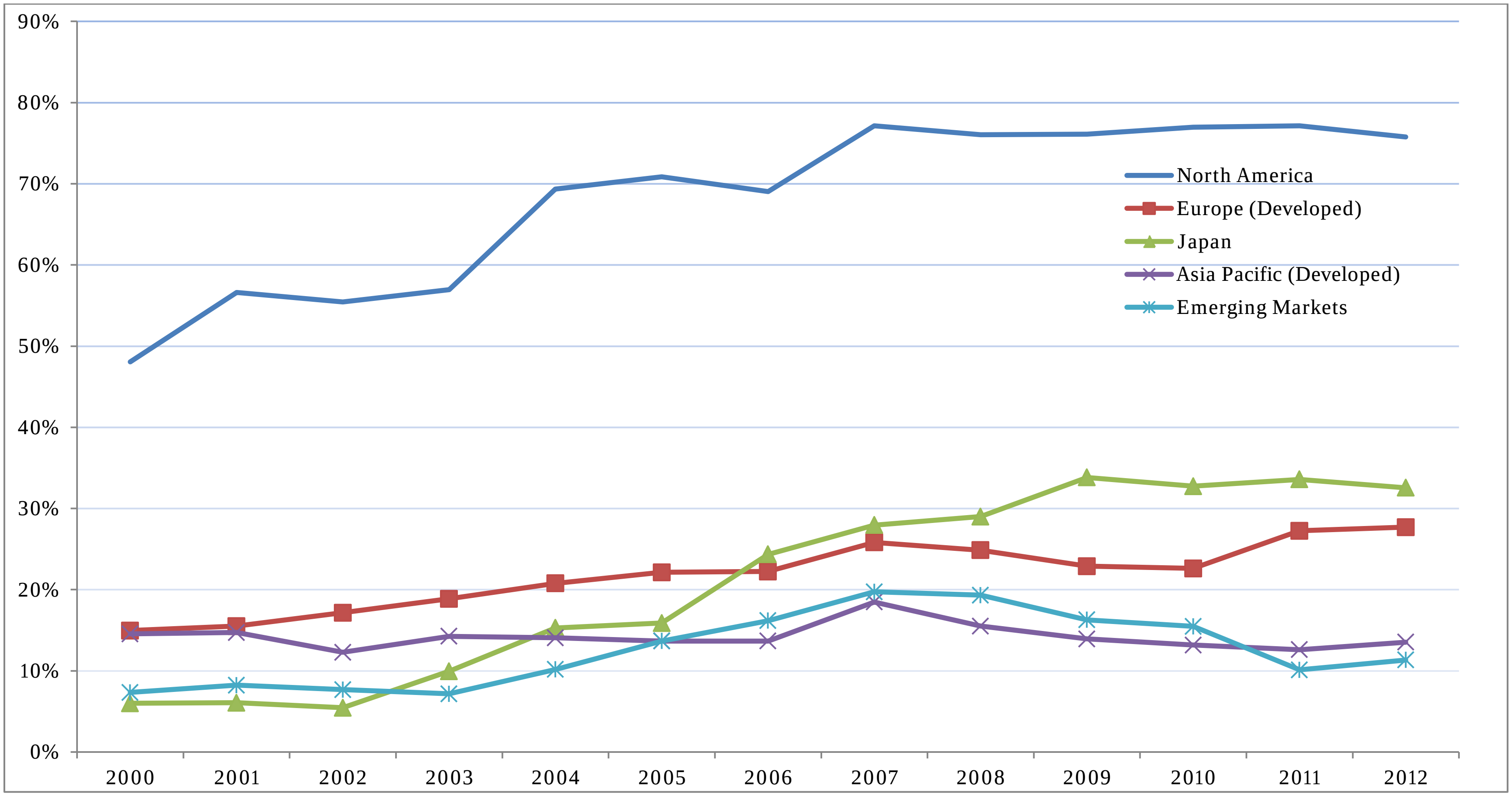
The limited coverage of Chinese firms by RavenPack is very apparent when compared with the coverage by Wiser News database. Wiser News is the Chinese counterpart of the well-known Factiva database. It systematically collect and classify news from national and local news media in China. Panel A in Figure AIII.2 plots the number of Chinese firms covered by Wiser News and RavenPack, respectively. The firm coverage by Wiser News is about four times that of RavenPack. Panel B in Figure AIII.2 makes a similar point by comparing the total number of news articles (without applying any filter mentioned previously) for Chinese firms. The comparison is even more striking. In fact, the RavenPack database covers less than 15% of all news articles in China recorded by the Wiser News database.

Overall, this case study illustrates that RavenPack's coverage of non-U.S. stocks seems to be large, cross-listed, and English language media driven. It is important to recognize such data selection issue in the empirical study of news in the international context. Exactly due to the concern of limited RavenPack's coverage outside the U.S., we limit our news analysis to the sample of U.S. stocks only.

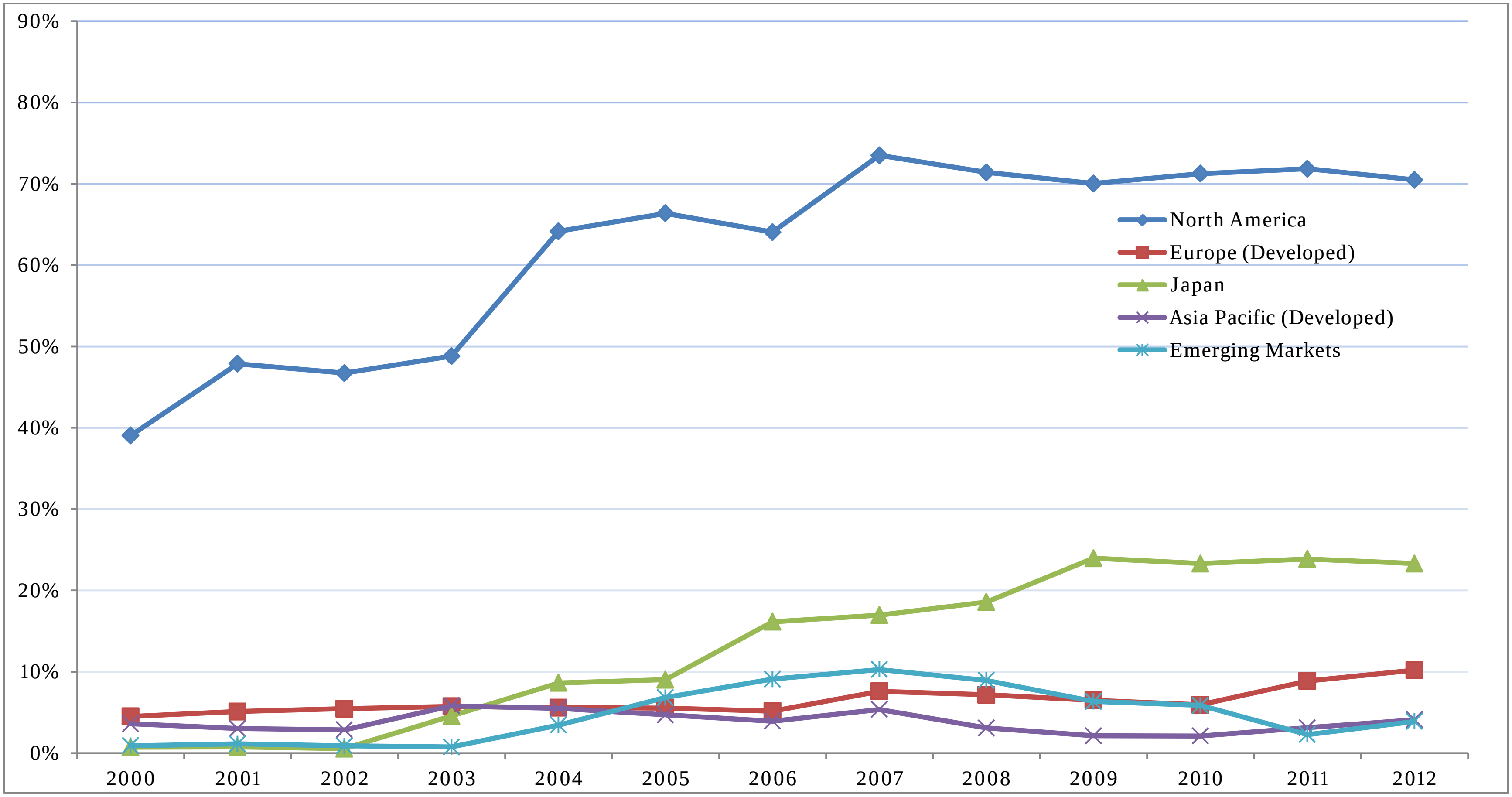
Figure AIII.1 – News Coverage by Region

This figure plots the fraction of our sample stocks with news in RavenPack database each year during the period 2000 to 2012, separately for the following regions: North America, European Developed Markets, Japan, Asia-Pacific Developed Markets (excluding Japan), and Emerging Markets. We first calculate the fraction of stocks with news each month, and then average the monthly coverage statistics in a year to obtain yearly statistics.

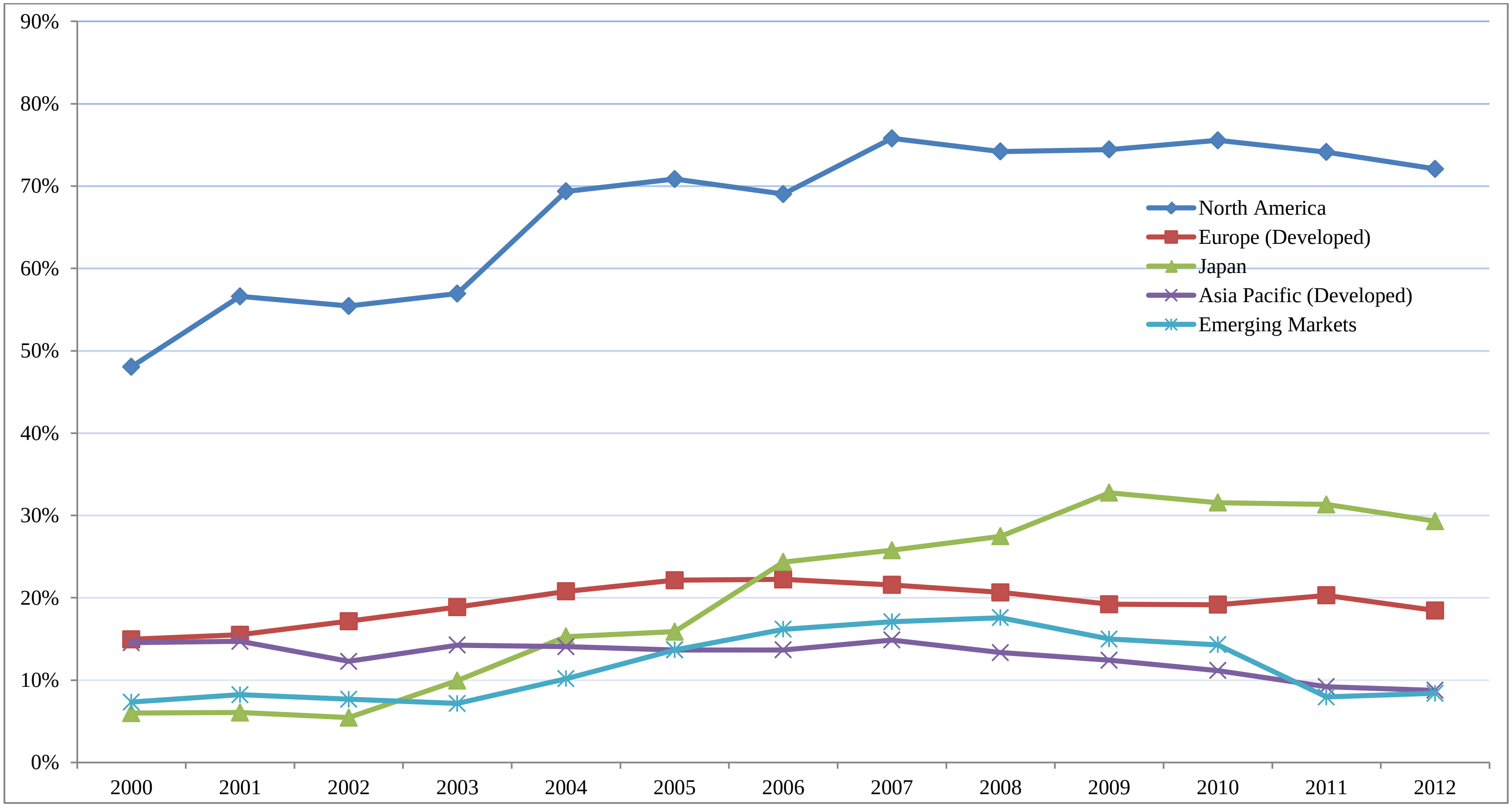
Panel A: All-stock sample, all-news edition.



Panel B: Small-stock sample, all-news edition.



Panel C: All-stock sample, Dow Jones edition.



Panel D: Small-stock sample, Dow Jones edition.

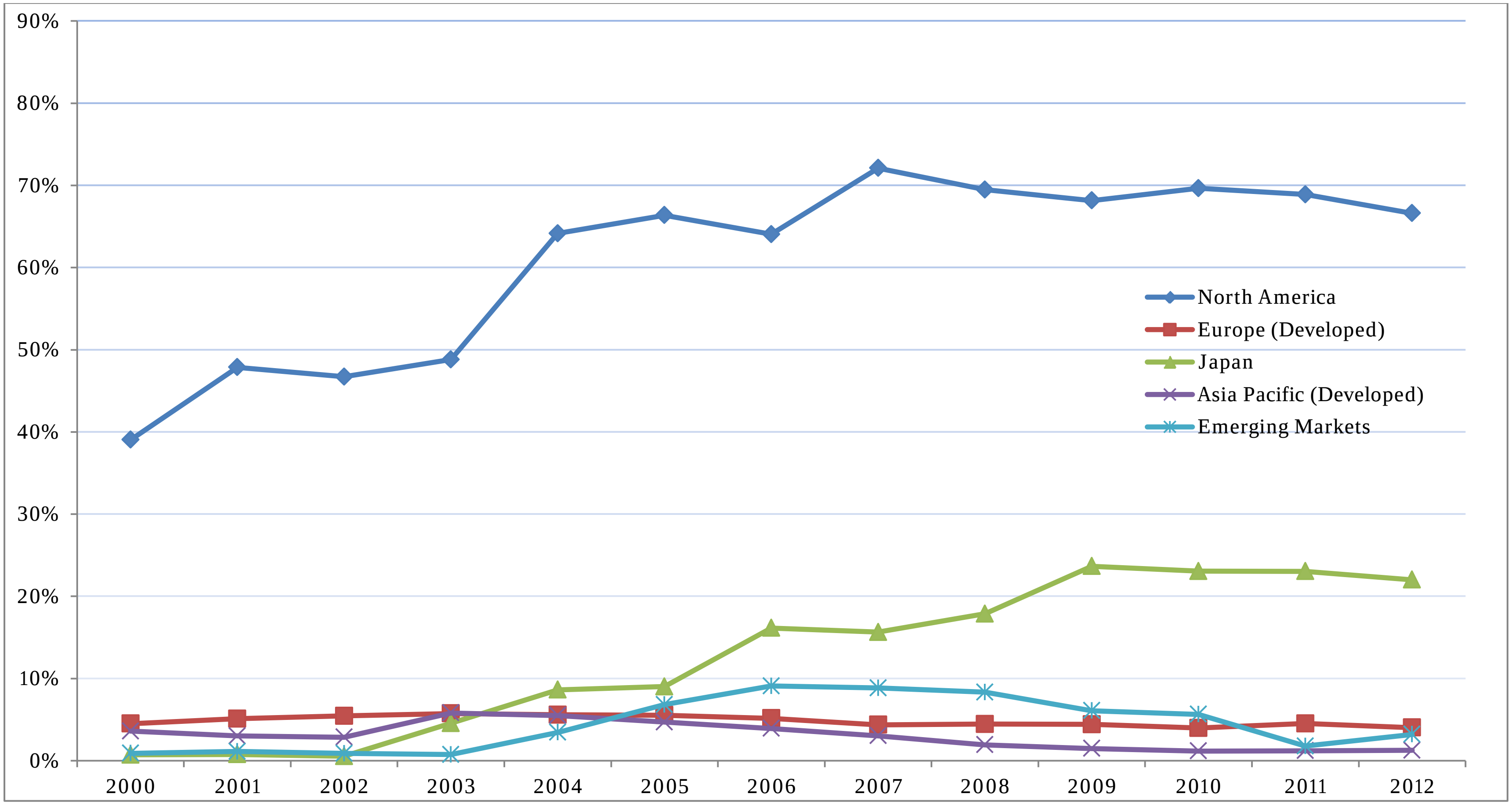
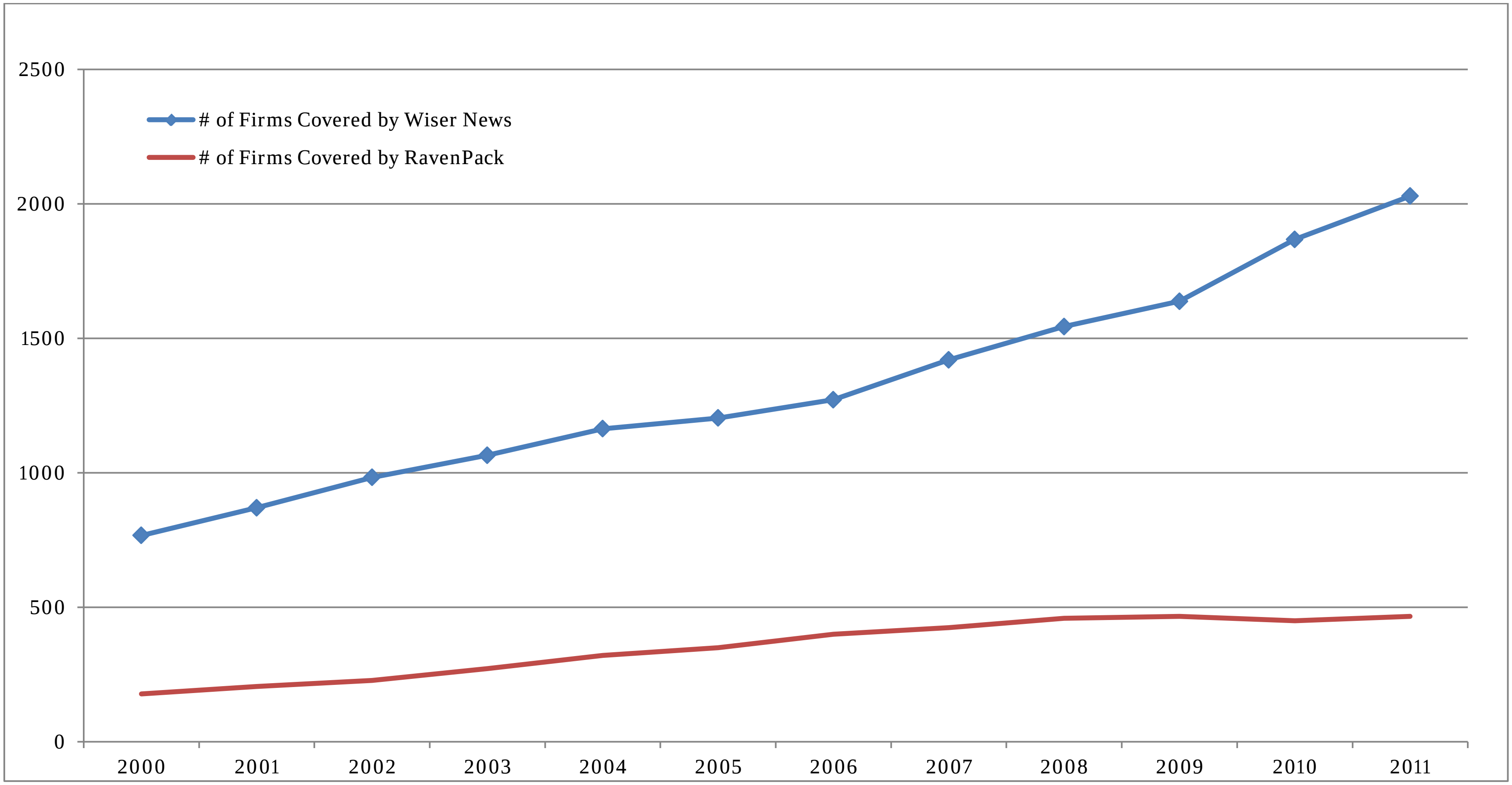


Figure AIII.2 – A Case Study: Comparison of Coverage by Wiser News vs. RavenPack of Chinese Firms

Panel A: This figure plots the number of Chinese firms covered by Wiser News database and by RavenPack each year between 2000 and 2011. Wiser News sample only include stocks traded on the Shanghai Stock Exchange and Shenzhen Stock Exchange. RavenPack sample also include stocks that are cross-listed or listed outside the Shanghai Stock Exchange and Shenzhen Stock Exchange.



Panel B: This figure plots the number of news stories of Chinese firms covered by Wiser News database (scale on the left-axis) and by RavenPack (scale on the right-axis) each year between 2000 and 2011. Wiser News sample only include stocks traded on the Shanghai Stock Exchange and Shenzhen Stock Exchange. RavenPack sample also include stocks that are cross-listed or listed outside the Shanghai Stock Exchange and Shenzhen Stock Exchange.

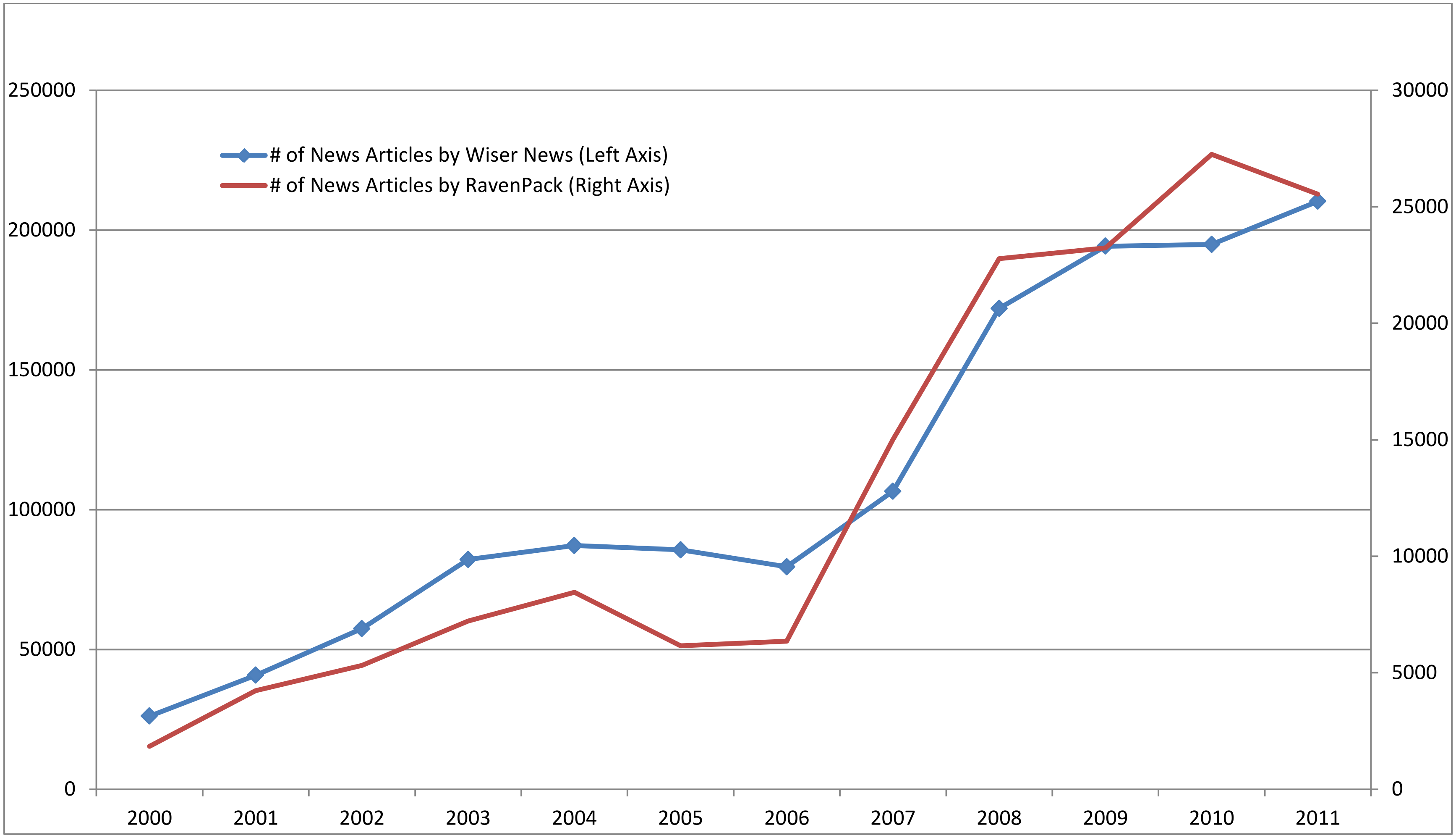


Table AIII.1 – News coverage by region and EDF deciles
This table summarizes news coverage in RavenPack by region and EDF decile during the period 2000 to 2012, based on Dow Jones Edition news (Panel A) and all sources of news (Panel B) respectively.

Panel A: Dow Jones Edition news only

	EDF deciles										
Regions	All	1 (lowest)	2	3	4	5	6	7	8	9	10 (highest)
	All stocks										
North America	67%	73%	73%	71%	69%	68%	67%	65%	64%	62%	59%
European developed markets	19%	21%	28%	27%	25%	22%	19%	17%	14%	12%	10%
Japan	20%	24%	23%	22%	21%	21%	20%	19%	19%	18%	15%
Asia Pacific developed markets	13%	26%	22%	17%	14%	12%	10%	9%	7%	6%	5%
Emerging markets	12%	17%	17%	15%	13%	12%	11%	10%	8%	7%	6%
	Large stocks										
North America	79%	77%	79%	79%	81%	80%	80%	79%	79%	78%	79%
European developed markets	34%	25%	34%	36%	35%	36%	35%	34%	34%	34%	35%
Japan	27%	28%	27%	27%	27%	27%	27%	28%	27%	28%	28%
Asia Pacific developed markets	23%	28%	30%	28%	25%	22%	21%	20%	18%	17%	17%
Emerging markets	19%	20%	21%	21%	20%	19%	19%	18%	17%	16%	17%
	Small stocks										
North America	61%	58%	62%	62%	62%	62%	62%	62%	62%	60%	59%
European developed markets	5%	2%	4%	5%	5%	5%	5%	5%	5%	5%	6%
Japan	13%	11%	12%	12%	13%	13%	13%	14%	14%	13%	13%
Asia Pacific developed markets	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%	4%
Emerging markets	4%	4%	5%	4%	4%	4%	4%	5%	5%	5%	5%

Panel B: All sources of news

Regions	All	EDF deciles 1 (lowest)	2	3	4	5	6	7	8	9	10 (highest)
					All stocks						
North America	68%	74%	74%	72%	70%	69%	68%	66%	65%	63%	60%
European developed markets	22%	22%	30%	29%	27%	24%	22%	20%	17%	15%	12%
Japan	21%	26%	24%	23%	22%	21%	21%	20%	19%	18%	16%
Asia Pacific developed markets	14%	28%	24%	19%	15%	13%	11%	10%	9%	7%	7%
Emerging markets	13%	18%	18%	16%	14%	13%	12%	10%	9%	8%	7%
					Large stocks						
North America	80%	78%	80%	80%	82%	81%	81%	80%	79%	79%	80%
European developed markets	37%	27%	37%	39%	39%	40%	39%	38%	38%	38%	39%
Japan	29%	29%	29%	28%	29%	29%	29%	29%	29%	29%	29%
Asia Pacific developed markets	25%	30%	32%	30%	27%	24%	23%	22%	20%	20%	20%
Emerging markets	20%	21%	22%	22%	21%	21%	21%	20%	19%	18%	19%
					Small stocks						
North America	62%	59%	63%	63%	63%	63%	63%	63%	63%	61%	60%
European developed markets	6%	3%	6%	6%	6%	7%	7%	7%	7%	7%	8%
Japan	13%	11%	12%	13%	13%	13%	14%	14%	14%	14%	13%
Asia Pacific developed markets	4%	3%	3%	3%	4%	4%	4%	4%	4%	4%	5%
Emerging markets	5%	4%	5%	4%	5%	5%	5%	5%	5%	5%	5%

Table AIII.2: Fama-MacBeth Regressions of Monthly Stock Returns on EDF, News Sentiment and Distress Duration for U.S. Small Stocks: Using Alternative News Sentiment Measure

This table presents the results of Fama-MacBeth regressions of monthly stock returns on the distressed dummy and its interactions with negative news dummy and newly distressed dummy, controlling for other firm characteristics, for the sample of U.S. small stocks. At the end of each month, we rank all U.S. small stocks by EDF into ten deciles, where the market capitalization cutoff value for large versus small stocks is the NYSE median value. High EDF and Low EDF are dummy variables that indicate whether a firm belongs to EDF decile 10 and decile 1 at the end of month (t-2), respectively. For each stock in EDF decile 10 in each month, we calculate its distress duration as the number of months having stayed in EDF decile 10 and its news sentiment as (# of positive stories – # of negative stories)/(# of positive stories + # of negative stories), where a news story is classified as positive (negative) if its ESS is above the 67th (below 33th) percentile value of all news stories of U.S. sample stocks during the sample period. Then at the end of each month, we rank all stocks in EDF decile 10 into three equally-sized groups based on the news sentiment and the distress duration respectively. Negative News and Newly Distressed are dummy variables that indicate whether, at the end of month t-2, a stock belongs to EDF decile 10 and is among the bottom tertile based on the news sentiment and the distress duration respectively. We interact High EDF with both Negative News and Newly Distressed. The dependent variable is the stock return in month t, denominated in U.S. dollars. Other firm characteristics include: 1) market capitalization (ME) at the end of month (t-1); 2) book-to-market equity ratio (B/M) at the end of month (t-1), calculated as the ratio of book equity of the fiscal year ending at least 6 months before month (t-1) to the market capitalization at end of month (t-1); 3) past 11-month stock return in U.S. dollars (MMT) from month (t-12) to (t-2); 4) past 1-month stock return in U.S. dollars (LRet) in month (t-1); 5) share turnover (Turnover) in month (t-1), calculated as the number of shares traded divided by the number of shares outstanding in month (t-1). All control variables enter as decile rankings (i.e., 1 for the bottom decile and 10 for the highest decile) within each month. Coefficients reported are time-series averages of the cross-sectional regressions shown above, with t-statistics calculated based on the Newey-West heteroskedasticity and autocorrelation consistent (HAC) standard errors of these estimates.

	(1)	(2)	(3)
Intercept	-0.03	-0.03	-0.03
	-0.04	*-0.04*	*-0.05*
High EDF	0.01	0.03	0.17
	0.03	*0.11*	*0.62*
Low EDF	0.08	0.08	0.08
	0.53	*0.54*	*0.54*
High EDF x Negative News	-0.64		-0.65
	-2.25		*-2.30*
High EDF x Newly Distressed		-0.51	-0.53
		-2.26	*-2.35*
ME	0.03	0.03	0.03
	0.63	*0.64*	*0.64*
B/M	0.11	0.11	0.11
	2.71	*2.73*	*2.75*
MMT	0.10	0.10	0.10
	2.32	*2.31*	*2.31*
LRET	-0.11	-0.11	-0.11
	-3.11	*-3.09*	*-3.10*
Turnover	-0.01	-0.01	-0.01
	-0.13	*-0.15*	*-0.12*
Fixed effects	Yes	Yes	Yes
Adj. R2	0.06	0.06	0.06
Avg. obs.	2,228	2,228	2,228

Appendix IV
Tables for Alternative Theories and Additional Robustness Checks

Table AIV.1: Time-Series Cross-sectional Regression of Monthly Stock Returns on EDF, Controlling Skewness of Stock Returns

This table presents the results of time-series cross-sectional regressions of monthly stock returns on EDF, controlling individual stocks' return skewness and other firm characteristics for stocks in all countries in our sample. We rank stocks in the respective sample (all, large or small stocks) into ten EDF decile groups within each country-month cohort. High EDF and Low EDF are dummy variables that indicate whether a firm belongs to EDF decile 10 and decile 1 at the end of month (t-2), respectively, within each country-month cohort. SKEW is the return skewness in month (t-2), measured from daily returns in the three-month period up to the end of month (t-2). Other firm characteristics include: 1) market capitalization (ME) at the end of month (t-1); 2) book-to-market equity ratio (B/M) at the end of month (t-1), calculated as the ratio of book equity of the fiscal year ending at least 6 months before month (t-1) to the market capitalization at end of month (t-1); 3) past 6-month stock return in U.S. dollars (MMT) from month (t-7) to (t-2); 4) past 1-month stock return in U.S. dollars (LRet) in month (t-1); 5) share turnover (Turnover) in month (t-1), calculated as the number of shares traded divided by the number of shares outstanding in month (t-1). All control variables including SKEW enter as decile rankings (i.e., 1 for the bottom decile and 10 for the highest decile) within each country- month cohort. Country and month fixed effects are also included in the regression. In addition to regression coefficients, we also report in the bottom rows of the table the implied return spreads between EDF deciles 10 and 1 based on the estimated coefficients. All t-statistics (indicated below in italics) are calculated based on standard errors with country clustering. As before, we report the results of the regressions on three samples - all stocks, large stocks, and small stocks, where the market capitalization cutoff value for large versus small stocks is the NYSE median value for the U.S. and the median value of all stocks in the country for all other countries.

	(1)	(2)	(3)
	All Stocks	Large Stocks	Small Stocks
Intercept	-7.30	-7.88	-6.83
	-5.17	*-5.47*	*-4.69*
High EDF	-0.24	-0.14	-0.30
	-4.53	*-2.76*	*-4.44*
Low EDF	0.08	0.03	0.16
	2.18	*0.69*	*3.94*
SKEW	-0.02	0.00	-0.02
	-2.79	*-0.65*	*-2.94*
ME	0.01	0.04	-0.04
	0.88	*2.27*	*-1.09*
B/M	0.14	0.12	0.16
	12.82	*7.49*	*14.36*
MMT	0.18	0.15	0.19
	8.25	*6.20*	*9.70*
LRET	-0.07	0.00	-0.11
	-4.81	*-0.24*	*-4.26*
Turnover	0.02	0.01	0.02
	1.14	*0.55*	*0.91*
Fixed effects	Yes	Yes	Yes
Adj. R2	0.13	0.18	0.11
Obs.	3,735,071	1,773,351	1,961,720
EDF L/S Return	-0.32	-0.17	-0.47
	-5.96	*-1.93*	*-5.92*

Table AIV.2: Default Risk and Stock Returns: Regional Evidence, including Microcap Stocks

This table presents the value-weighted average returns, Fama-French-Carhart four-factor model alphas, and Hou-Karolyi-Kho four-factor model alphas for long/short hedging portfolios that buy high-EDF stocks and sell low-EDF small stocks for the following regions: North America, European Developed Markets, Japan, Asia-Pacific Developed Markets (excluding Japan), and Emerging Markets. We identify small stocks as those with the market capitalization below NYSE median value for the U.S. and the median value of all stocks in the country for all other countries. At the end of each month t, we rank all small stocks within every country by EDF, and then form deciles. We then aggregate stocks across all countries in a region within each EDF decile to form portfolios for the region. We focus on a long-short strategy involving EDF-deciles 10 (long) and 1 (short) that begins one month after portfolio formation, i.e., at the beginning of month t+2. For this long-short (L/S) portfolio, we report both the average raw monthly value-weighted returns, alphas from the Fama-French-Carhart four-factor model (FFC4-α), and alphas from the Hou-Karolyi-Kho four-factor model (HKK4-α). The factor construction process follows Fama and French (2012) and Hou, Karolyi, and Kho (2011). Details of factor construction are provided in the appendix. Returns of EDF portfolios and factor portfolios are denominated in U.S. dollars. Results for various holding periods are shown, including one month (t+2), three-month (t+2, t+4), six-month (t+2, t+7), and one year (t+2, t+13). For holding periods greater than one month, we follow the overlapping horizon approach of Jegadeesh and Titman (1993).

		Holding Periods								
		t+2			t+2, t+4			t+2, t+13		
Region	Stat.	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α
North America	Est	-0.44%	-0.55%	-0.59%	-0.32%	-0.40%	-0.45%	-0.01%	-0.14%	-0.18%
	t-stat	*-1.11*	*-2.44*	*-2.60*	*-0.80*	*-1.80*	*-2.02*	*-0.03*	*-0.62*	*-0.77*
Europe DM	Est	-1.28%	-1.11%	-1.09%	-1.24%	-1.08%	-1.03%	-0.96%	-0.93%	-0.86%
	t-stat	*-5.16*	*-5.41*	*-5.23*	*-5.16*	*-5.46*	*-5.14*	*-4.36*	*-4.76*	*-4.26*
Japan	Est	0.19%	0.04%	0.06%	0.15%	-0.04%	0.00%	0.06%	-0.19%	-0.12%
	t-stat	*0.63*	*0.18*	*0.27*	*0.50*	*-0.17*	*-0.01*	*0.21*	*-0.96*	*-0.60*
Asia Pacific DM	Est	0.03%	0.07%	0.30%	-0.01%	0.01%	0.23%	0.10%	0.04%	0.17%
	t-stat	*0.07*	*0.21*	*0.79*	*-0.02*	*0.02*	*0.69*	*0.31*	*0.15*	*0.61*
Emerging	Est	-0.43%	-0.32%	-0.20%	-0.45%	-0.30%	-0.17%	-0.54%	-0.44%	-0.32%
	t-stat	*-1.42*	*-1.09*	*-0.70*	*-1.59*	*-1.07*	*-0.63*	*-2.15*	*-1.83*	*-1.39*

Table AIV.3: Creditor Rights, Individualism and the Distress Anomaly: Including Microcap Stocks

This table presents the average returns and alphas from both Fama-French-Carhart and Hou-Karolyi-Kho four-factor models for EDF-based long/short hedging portfolios that formed among small stocks in groups of countries with different levels of La Porta, Lopez-de-Silanes, Shleifer, and Vishny's (1998) Creditor Rights index (Panel A) and Hofstede's (2001) Individualism index (Panel B) respectively. The procedure is as follows. First, at the end of each month *t*, we rank all small stocks within each country by EDF, and then form value-weighted decile portfolios for each country. We implement a long-short strategy involving EDF-deciles 10 (long) and 1 (short) for various holding periods that begin one month after portfolio formation: month (t+2), three-month (t+2, t+4), and one year (t+2, t+13). For holding periods greater than one month, we follow the overlapping horizon approach of Jegadeesh and Titman (1993). All countries in our sample are classified into two groups (low and high) based on the median index value. Each month, the returns of country-level long/short hedging portfolios are averaged across all countries in an index group to obtain the return of long/short hedging portfolios for the index group in that month. The final row shows the results of testing the difference between the returns of high and low index groups.

Panel A: Creditor rights

Country Groups	Stat.	t+2 Return	t+2 FFC4-α	t+2 HKK4-α	t+2, t+4 Return	t+2, t+4 FFC4-α	t+2, t+4 HKK4-α	t+2, t+13 Return	t+2, t+13 FFC4-α	t+2, t+13 HKK4-α
Low CR	Est	-0.50%	-0.41%	-0.39%	-0.45%	-0.36%	-0.30%	-0.40%	-0.35%	-0.28%
	t-stat	*-1.81*	*-1.64*	-1.54	*-1.74*	*-1.50*	-1.29	*-1.81*	*-1.67*	-1.38
High CR	Est	-0.50%	-0.44%	-0.45%	-0.47%	-0.44%	-0.44%	-0.37%	-0.41%	-0.42%
	t-stat	*-2.65*	*-2.67*	-2.73	*-2.57*	*-2.80*	-2.84	*-2.31*	*-2.81*	-2.84
High – Low CR	Est	0.00%	-0.02%	-0.06%	-0.02%	-0.08%	-0.14%	0.03%	-0.07%	-0.13%
	t-stat	*-0.01*	*-0.09*	-0.21	*-0.08*	*-0.32*	-0.55	*0.15*	*-0.31*	-0.61

Panel B: Individualism

Country Groups	Stat.	t+2 Return	t+2 FFC4-α	t+2 HKK4-α	t+2, t+4 Return	t+2, t+4 FFC4-α	t+2, t+4 HKK4-α	t+2, t+13 Return	t+2, t+13 FFC4-α	t+2, t+13 HKK4-α
Low Indv	Est	0.11%	0.07%	0.05%	0.01%	-0.04%	-0.04%	-0.13%	-0.20%	-0.21%
	t-stat	*0.50*	*0.32*	0.25	*0.05*	*-0.23*	-0.23	*-0.72*	*-1.17*	-1.25
High Indv	Est	-0.93%	-0.76%	-0.74%	-0.78%	-0.64%	-0.62%	-0.55%	-0.53%	-0.48%
	t-stat	*-4.00*	*-3.73*	-3.63	*-3.53*	*-3.31*	-3.19	*-2.84*	*-2.90*	-2.66
High - Low Indv	Est	-1.04%	-0.83%	-0.79%	-0.79%	-0.60%	-0.58%	-0.42%	-0.33%	-0.27%
	t-stat	*-3.90*	*-3.03*	-2.91	*-3.13*	*-2.32*	-2.22	*-1.94*	*-1.46*	-1.21

Table AIV.4: Market States and the Distress Anomaly: Including Microcap Stocks

This table presents the results of time-series cross-sectional regressions of monthly stock returns on EDF and its interactions with market states and turnover, controlling for other firm characteristics. We rank stocks into EDF decile portfolios within each country-month cohort. High EDF and Low EDF are dummy variables that indicate whether a stock belongs to EDF decile 10 and decile 1 at the end of month t-2, respectively, within each country-month cohort. Up Mkt(t) is a dummy variable that equals one if a country's past 12-month cumulative return up to month t is both positive and above the median value for all months in our sample for that country, and zero otherwise. Down Mkt(t) equals one minus Up Mkt(t). In Panel A, we interact EDF dummy variables with Up Mkt(t-2), the market state indicator in our portfolio formation month. In Panel B, we interact EDF dummy variables with both Up Mkt(t-2), and High Turnover, a dummy variable that equals one if a stock's turnover in month t-2 is above the median value for all sample stocks in its country, and zero otherwise. The dependent variable is the stock return in month t, denominated in U.S. dollars. Other firm characteristics include: 1) market capitalization (ME) at the end of month (t-1); 2) book-to-market equity ratio (B/M) at the end of month (t-1), calculated as the ratio of book equity of the fiscal year ending at least 6 months before month (t-1) to the market capitalization at end of month (t-1); 3) past 11-month stock return in U.S. dollars (MMT) from month (t-12) to (t-2); 4) past 1-month stock return in U.S. dollars (LRet) in month (t-1); 5) share turnover (Turnover) in month (t-1), calculated as the number of shares traded divided by the number of shares outstanding in month (t-1). All control variables enter as decile rankings (i.e., 1 for the bottom decile and 10 for the highest decile) within each country-month cohort. Country and month fixed effects are also included in the regression. Estimates of controls are not reported in Panels B for brevity. All t-statistics (indicated below in italics) are calculated based on standard errors clustered by country.

Panel A: Prior Market State		Panel B: Prior Market State and Turnover	
High EDF	0.02	High EDF	0.25
	0.29		*2.97*
Low EDF	-0.11	Low EDF	-0.17
	-2.52		*-3.24*
High EDF x Up Mkt (t-2)	-0.36	High EDF x Up Mkt (t-2)	-0.22
	-3.64		*-3.12*
Low EDF x Up Mkt (t-2)	0.38	Low EDF x Up Mkt (t-2)	0.33
	8.29		*6.06*
Up Mkt (t-2)	0.28	High EDF x Up Mkt (t-2) x High Turnover	-0.34
	1.65		*-2.24*
		Low EDF x Up Mkt (t-2) x High Turnover	0.04
Controls:			*0.56*
ME	0.00	High EDF x High Turnover	-0.39
	0.40		*-2.94*
B/M	0.14	Low EDF x High Turnover	0.08
	12.73		*1.71*
MMT	0.17	Up Mkt (t-2)	0.49
	7.95		*2.90*
LRET	-0.08	Up Mkt (t-2) x High Turnover	-0.39
	-4.83		*-6.64*
Turnover	0.02	High Turnover	0.11
	1.02		*2.83*
Fixed effects	Yes	Controls & Fixed effects	Yes
Adj. R2	0.13	Adj. R2	0.13
Obs.	3,867,015	Obs.	3,867,006

Table AIV.5: Default Risk and Stock Returns: Regional Evidence, excluding Financial Stocks

This table presents the value-weighted average returns, Fama-French-Carhart four-factor model alphas, and Hou-Karolyi-Kho four-factor model alphas for long/short hedging portfolios that buy high-EDF stocks and sell low-EDF small stocks for the following regions: North America, European Developed Markets, Japan, Asia-Pacific Developed Markets (excluding Japan), and Emerging Markets. We identify small stocks as those with the market capitalization below NYSE median value for the U.S. and the median value of all stocks in the country for all other countries. We exclude all financial stocks from the sample. At the end of each month t, we rank all small stocks within every country by EDF, and then form deciles. We then aggregate stocks across all countries in a region within each EDF decile to form portfolios for the region. We focus on a long-short strategy involving EDF-deciles 10 (long) and 1 (short) that begins one month after portfolio formation, i.e., at the beginning of month t+2. For this long-short (L/S) portfolio, we report both the average raw monthly value-weighted returns, alphas from the Fama-French-Carhart four-factor model (FFC4-α), and alphas from the Hou-Karolyi-Kho four-factor model (HKK4-α). The factor construction process follows Fama and French (2012) and Hou, Karolyi, and Kho (2011). Details of factor construction are provided in the appendix. Returns of EDF portfolios and factor portfolios are denominated in U.S. dollars. Results for various holding periods are shown, including one month (t+2), three-month (t+2, t+4), six-month (t+2, t+7), and one year (t+2, t+13). For holding periods greater than one month, we follow the overlapping horizon approach of Jegadeesh and Titman (1993).

		Holding Periods								
		t+2			t+2, t+4			t+2, t+13		
Region	Stat.	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α
North America	Est	-0.23%	-0.32%	-0.35%	-0.20%	-0.26%	-0.31%	0.13%	0.04%	0.01%
	t-stat	*-0.54*	*-1.32*	*-1.43*	*-0.48*	*-1.09*	*-1.32*	*0.33*	*0.15*	*0.02*
Europe DM	Est	-1.01%	-0.80%	-0.83%	-1.00%	-0.82%	-0.81%	-0.77%	-0.76%	-0.72%
	t-stat	*-3.77*	*-3.60*	*-3.67*	*-3.92*	*-3.90*	*-3.75*	*-3.39*	*-3.89*	*-3.52*
Japan	Est	0.19%	0.04%	0.06%	0.19%	-0.01%	0.03%	0.08%	-0.20%	-0.13%
	t-stat	*0.66*	*0.19*	*0.30*	*0.67*	*-0.06*	*0.13*	*0.29*	*-1.03*	*-0.66*
Asia Pacific DM	Est	-0.26%	-0.25%	-0.05%	-0.26%	-0.26%	-0.11%	0.11%	0.08%	0.11%
	t-stat	*-0.68*	*-0.75*	*-0.14*	*-0.71*	*-0.83*	*-0.33*	*0.34*	*0.28*	*0.39*
Emerging	Est	-0.27%	-0.21%	-0.11%	-0.27%	-0.18%	-0.07%	-0.46%	-0.42%	-0.32%
	t-stat	*-0.90*	*-0.72*	*-0.37*	*-0.98*	*-0.65*	*-0.27*	*-1.88*	*-1.78*	*-1.40*

Table AIV.6: Default Risk and Stock Returns: Regional Evidence, Industry-Adjusted

This table presents the value-weighted average returns, Fama-French-Carhart four-factor model alphas, and Hou-Karolyi-Kho four-factor model alphas for long/short hedging portfolios that buy high-EDF stocks and sell low-EDF small stocks for the following regions: North America, European Developed Markets, Japan, Asia-Pacific Developed Markets (excluding Japan), and Emerging Markets. At the end of each month t, we first demean EDF within each country and each industry as defined by 1-digit SIC code, and then rank all small stocks within every country by the industry-adjusted EDF and form deciles. We identify small stocks as those with the market capitalization below NYSE median value for the U.S. and the median value of all stocks in the country for all other countries. We then aggregate stocks across all countries in a region within each EDF decile to form portfolios for the region. We focus on a long-short strategy involving EDF-deciles 10 (long) and 1 (short) that begins one month after portfolio formation, i.e., at the beginning of month t+2. For this long-short (L/S) portfolio, we report both the average raw monthly value-weighted returns, alphas from the Fama-French-Carhart four-factor model (FFC4-α), and alphas from the Hou-Karolyi-Kho four-factor model (HKK4-α). The factor construction process follows Fama and French (2012) and Hou, Karolyi, and Kho (2011). Details of factor construction are provided in the appendix. Returns of EDF portfolios and factor portfolios are denominated in U.S. dollars. Results for various holding periods are shown, including one month (t+2), three-month (t+2, t+4), six-month (t+2, t+7), and one year (t+2, t+13). For holding periods greater than one month, we follow the overlapping horizon approach of Jegadeesh and Titman (1993).

		Holding Periods								
		t+2			t+2, t+4			t+2, t+13		
Region	Stat.	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α	Return	FFC4-α	HKK4-α
North America	Est	-0.42%	-0.53%	-0.57%	-0.31%	-0.39%	-0.44%	-0.01%	-0.14%	-0.18%
	t-stat	*-1.06*	*-2.36*	*-2.51*	*-0.78*	*-1.76*	*-1.97*	*-0.03*	*-0.61*	*-0.77*
Europe DM	Est	-1.20%	-0.97%	-0.98%	-1.11%	-0.91%	-0.89%	-0.88%	-0.82%	-0.76%
	t-stat	*-4.96*	*-4.93*	*-4.88*	*-4.72*	*-4.80*	*-4.59*	*-4.11*	*-4.42*	*-4.00*
Japan	Est	0.15%	0.00%	0.02%	0.15%	-0.05%	-0.02%	0.08%	-0.20%	-0.13%
	t-stat	*0.49*	*-0.01*	*0.10*	*0.50*	*-0.27*	*-0.09*	*0.27*	*-1.00*	*-0.64*
Asia Pacific DM	Est	-0.28%	-0.20%	-0.02%	-0.30%	-0.27%	-0.09%	-0.11%	-0.21%	-0.10%
	t-stat	*-0.74*	*-0.65*	*-0.05*	*-0.84*	*-0.96*	*-0.31*	*-0.36*	*-0.88*	*-0.38*
Emerging	Est	-0.31%	-0.23%	-0.12%	-0.38%	-0.25%	-0.15%	-0.40%	-0.29%	-0.21%
	t-stat	*-1.12*	*-0.83*	*-0.45*	*-1.47*	*-0.98*	*-0.61*	*-1.69*	*-1.26*	*-0.95*