ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Master Thesis Economics & Business: Financial Economics

International Capital Market Efficiency and Institutional Quality

Han Lim
417025sl

Dr. Sjoerd van den Hauwe
Dr. Haikun Zhu

November 2019

Abstract

Applying a state of the art anomaly-based stock mispricing measure, constructed by Stambaugh, Yu, and Yuan (2015), this paper aims to empirically identify a relation between structural mispricing on capital markets and international differences in legal shareholder protection. A mechanism is introduced, where environments of high institutional quality promote capital market efficiency through attraction of arbitrage investors. This paper provides empirical evidence that higher institutional quality improves equity pricing efficiency. By breaking legal shareholder environments down into quantified investor policy variables, this research is successful in explaining international mispricing differences within the context of institutional quality. However, no significant relation between English, French, German or Scandinavian legal origins and mispricing has been found. This research concludes that stronger legal shareholder protection does improve capital market efficiency, although the specific policies that do so remain difficult to identify. Furthermore, this research cautiously suggests robustness of the relation between investor protection and mispricing during periods of economics downturn.

Contents

1 Introduction

Pricing inefficiency of publicly listed equities is a complex phenomenon which, after decades of research, has remained remarkably challenging to understand. Although empirical findings suggest that stock prices often fail to reflect fundamental values, asset pricing research has yet to develop credible methods to empirically explain, quantify and predict such mispricing. Generally speaking, equity pricing on capital markets results from a process, where investors react to information. They can either choose to buy (long) or sell (short) stocks, based on their evaluation of a stock's fundamental value. This is where inefficiencies arise: it appears that investors lack complete information or that they are restricted in reacting rationally to new information (Barberis and Thaler, 2003). The challenge here lies in empirically identifying all factors that restrict investors and as such create pricing inefficiencies. This research is aimed to introduce legal environments as one of these factors that influence pricing efficiency on capital markets.

Capital markets have a central position within modern societies and involve many types of stakeholders. Individual and institutional investors, management boards and governments are examples of stakeholders that have influence on capital markets, while also being affected by them. Exploring the landscape of factors that affect capital market efficiency, helps interpreting the gap between market prices and fundamental values. A proper understanding on how these markets work, can assist in improvement of investment strategies and government regulations. Both could play a large role in the development of financial markets.

The relationship between legal environments and capital market efficiency has not been explicitly discussed in previous literature. Separately, however, both fields are commonly investigated topics. This allows for a strong theoretical framework through which this relation can be explored.

Based on La Porta, Lopez-de-Silanes, Shleifer, and Vishny (1997) and Morck, Yeung, and Yu (2000), it is stated that countries that legally protect outside investors (i.e. providers of external financing) are expected to prevent managers from exhibiting self-serving and investor-harming behaviour. In this way, countries that provide strong investor protection become more attractive to investors, including arbitrage investors. As arbitrageurs actively detect and exploit mispricing, their trading activity would improve capital market pricing efficiency Ross, 1973. Through this mechanism, the countries that offer strong investor protection are hypothesised to show less mispricing, with countries that provide weak protection showing more mispricing. A categorisation of countries is made based on English, French, German and Scandinavian legal origins. Here, English origin countries offer the strongest shareholder protection, while French origin countries offer the weakest shareholder protection. German and Scandinavian origin countries offer a level of protection that is somewhere in between. To examine these legal environments in more detail, they are broken down into individual institutional variables that are related to shareholder protection. This is the first paper to describe and empirically investigate this arbitrage attraction mechanism.

The methodology of this research is based on Stambaugh et al. (2015), who recently composed a mispricing measure to empirically identify mispricing. Based on 11 anomalies, scores are calculated for each stock that should indicate the likeliness that a stock is overpriced. A country portfolio is then constructed, buying (long position) the stocks that are most likely to be underpriced and selling (short position) the stocks that are most likely to be overpriced. For each long-short portfolio, this research follows Jacobs (2016) in calculating three-factor-benchmark-adjusted alphas (excess returns) to investigate the persistence of mispricing in international markets. This method is replicated for 40 countries globally between 2008-2018, using stock price and accounting data from Datastream Worldscope. The mispricing estimates across the sample of countries are then combined with data regarding legal origins, institutional variables and macro-level control variables. Through country comparisons, this research examines whether differences across legal shareholder environments empirically explain mispricing.

The results generated from this methodology support the hypothesis that institutional factors are related to mispricing on capital markets. In most cases, stricter country-wide investor protection shows to reduce country-level stock mispricing. However, this effect is not observable for compositions of institutional factors based on legal origins, as the results do not suggest a direct effect between mispricing and legal origins. Additional regressions also suggest persistence of the arbitrage attraction mechanism during and outside periods of recessions. This research concludes that stronger shareholder protection does indeed improve pricing efficiency on capital markets.

This research offers a unique combination between the application of a recently composed mispricing metric and an original theoretical law-based efficiency mechanism. The Stambaugh et al. (2015) method for identifying and quantifying relative mispricing has not yet been applied within the context of investor protection. Their papers from the previous years are mainly focused on explaining mispricing through investor sentiment (2012 and 2014) and limits to arbitrage (2015). Jacobs (2016) has only applied the mispricing measure to investigate efficiency differences between developed and emerging markets. Papers that are focused on investor protection have mainly investigated more direct effects, such as external financing (La Porta et al., 1997), concentration of ownership (La Porta, Lopez-de-Silanes, Shleifer, and Vishny, 1998), company valuation (La Porta, Lopez-de-Silanes, Shleifer, and Vishny, 2002) and earnings management (Leuz, Nanda, and Wysocki, 2003).

As this paper sheds light on a relationship that has not been discussed before, its purpose remains completely fundamental. This paper is not written with the intention to provide any practical recommendations, but to open doors for further research within this field. Expanding research regarding the relation between legal environments and capital market efficiency might provide more pragmatic approaches regarding financial market development through investor policies.

2 Literature review

2.1 Capital asset pricing

Throughout decades of past research on asset pricing, many attempts have been made to investigate the relation between stock prices and their underlying fundamental values. Although many empirical findings suggest that stock prices often fail to reflect fundamental values, asset pricing has yet to develop credible methods to empirically quantify and explain such mispricing.

As an attempt to approach rational asset pricing, Sharpe (1964) and Lintner (1965) introduced the widely used Capital Asset Pricing Model (CAPM, equation 1), where expected asset returns are dependent on the cross-section of a single market risk exposure factor (b_i). This model assumes that investors are only rewarded with returns for increasing exposure to market-wide risk, as firm-specific idiosyncratic risk can be mitigated through diversification.

$$E(R_{i,t}) = R_{f,t} + b_i \times (E(R_{m,t}) - R_{f,t}) \tag{1}$$

Equation 1: Capital Asset Pricing Model. $E(R_i)$ represents an asset's expected return, which is dependent on the market's risk-free rate $R_{f,t}$, asset-specific market risk exposure coefficient b_i and expected market return $E(R_{m,t})$.

Although today the CAPM is still considered to be a major pillar within the theories of asset pricing, it has been recognised that market factor coefficient b_i alone is not sufficient when explaining and predicting cross-sectional returns. Violations of the CAPM, so-called "anomalies", have been found where factors unexplained by market risk appeared to be related to average returns. Examples of these factors are market capitalisation (Banz, 1981), earnings-price ratio (Basu, 1983), book-to-market equity ratio (Rosenberg, Reid, and Lanstein, 1985) and market leverage (Bhandari, 1988).

In reaction to the CAPM's apparent empirical problems, Fama and French (1993) have constructed the three-factor empirical asset pricing model (equation 2). This extension of the CAPM model was aimed to improve the accuracy of return predictions, capturing the empirically emphasized relations between average return and market capitalisation (company size) and between average return and book-to-market ratio (or other price ratios). The three factor model includes a size and value premium, representing the return difference between the smallest and largest firms and between firms with the highest (value stocks) and lowest (growth stocks) relative book-to-market measures.

$$E(R_{i,t}) = R_{f,t} + b_i \times (E(R_{m,t}) - R_{f,t}) + s_i \times (SMB_t) + h_i \times (HML_t) \tag{2}$$

Equation 2: Traditional CAPM expression, with size premium SMB (Small Minus Big) and value (book-to-market ratio) premium HML (High Minus Low) added as independent variables.

Although this three-factor model has shown to explain returns more accurately than the traditional single market factor model, many anomalies have still been found to survive adjustment for the three factors of Fama and French (1993). A selection of those anomalies are applied in the methodology described below to predict deviations from fundamental asset pricing models.

2.2 Identifying mispricing

In order to identify potential market-wide mispricing, Stambaugh et al. (2015) have constructed a mispricing score, based on the following 11 empirically proven return anomalies that the three-factor model has failed to absorb (see Appendix A.1 and A.2 for definitions, references and calculations):

1. Financial distress
2. O-Score bankruptcy probability
3. Net stock issues
4. Composite equity issues
5. Total accruals
6. Net operating assets
7. Momentum
8. Gross profitability
9. Asset growth
10. Return on assets
11. Investment-to-assets

The listed 11 anomalies do not necessarily explain mispricing, but are rather common components found across mispriced stocks. Stambaugh et al. (2015) construct one mispricing factor per stock-month by averaging stocks' rankings across the 11 anomalies. This combination method deviates from traditional methods that correspond a factor to a single anomaly. The mispricing score serves as a proxy for the likelihood that an asset's return deviates from return predictions based on systematic risk.

To empirically identify mispricing in the U.S. between 1965-2011, Stambaugh et al. (2015) have constructed value-weighted portfolios based on the calculated mispricing scores - taking long (short) positions in the stock deciles that are most likely to be underpriced (overpriced). The returns that are generated by this portfolio are consequently adjusted for the three-factor benchmark constructed by Fama and French (1993) (equation 2). After adjustment for the three market factors, consistent excess returns (so-called "alphas") should indicate mispricing. This is exactly what they found: a significant alpha spread between the two extreme anomaly score deciles of 1.48% implied a persistent market-wide mispricing in the U.S. stock market.

Jacobs (2016) investigated stock return predictability on international markets in 1994-2013, while applying the Stambaugh et al. (2015) mispricing score. To compare the estimated mispricing across the 45 investigated countries, Jacobs calculated country-level three factor alphas. These appeared to be significant and positive in most of the sampled countries. He concluded that mispricing appears to be at least as prevalent in developed markets as in emerging markets.

2.2.1 Equity market efficiency and arbitrage

One essential mechanism for price efficiency within a market containing irrational investors is called arbitrage investing (Ross, 1973). As asset prices deviate from fundamental values, opportunities emerge where arbitrage investors can generate excess returns from moving against the mispricing (e.g. buying underpriced or selling overpriced stocks). Collective activity from arbitrageurs should, in this way, drive prices towards fundamental values again. However, in reality mispricing does not always equal a free lunch opportunity, since there are also limits to arbitrage. This concerns transaction costs or constraints as well as fundamental and/or behavioural risks that cannot be simply eliminated by hedging, which can make it less attractive or impossible for arbitrageurs to react to perceived mispricing. In this paper, weak shareholder protection is hypothesised to represent one of those limits. In this way, institutional quality affects mispricing through attracting or deterring arbitrage investors.

2.2.2 Legal investor protection and arbitrage

Legislation concerning management reporting and shareholder rights becomes critical when managers of companies act in their own interest. For example, Leuz et al. (2003) explain that managers have incentives to manage reported earnings in order to mask true performance for outside investors. By managing the level and variability of reported earnings, managers attempt to reduce the likelihood of outside intervention. Leuz et al. (2003) find that stronger legal systems that protect outside investors improve the quality of earnings reports, as this weakens private control and incentives to obfuscate firm performance.

Morck et al. (2000) explain that arbitrageurs are more attracted towards economies that force firms to provide better accounting data, as this helps them making more accurate predictions of stock price movements. They also find that stronger public investor property rights successfully promote arbitrage, as this protects arbitrageurs from corporate insider problems, such as intercorporate income shifting. It appears that country-varying levels of shareholder protection can affect the presence and activity of arbitrageurs on equity markets. Considering the arbitrageurs' crucial role in correcting irrational investor behaviour, shareholder protection might therefore indirectly affect country-level price efficiency.

La Porta et al. (1997) use the distinctions made by Reynolds and Flores (1989) between countries with common law, which is English in origin, and civil law, which is also known as the Romano-Germanic law. These countries differ in shareholder protection by legal rules and quality of their enforcement against expropriation by insiders. Countries which incorporate common law are considered to provide the strongest protection towards shareholders and creditors, while civil law countries are considered to give weaker legal rights to investors. Examples of these shareholder rights are voting rights, dividend or cash flow rights and anti-director rights. According to La Porta et al. (1997), countries with strong investor protection are more successful in attracting investors' capital, resulting in relatively large capital markets compared to countries with weak shareholder rights.

Although the two law origins present clear differences concerning shareholder rights, there are still within-group differences between quality of law enforcement. A distinction can be made between English-common-law, French-civil-law, German-civil-law and Scandinavian-civil-law. The level of investor protection in German and Scandinavian origin countries lies between that of English and French origins. The origin groups will be compared in more detail in the data section, where multiple individual components of legal investor protection will be discussed. The methodology section will explain how this research aims to identify mispricing differences across the four legal origin groups.

The literature review above suggests a relationship between legal investor protection and attraction of (arbitrage) investors. Stricter accounting rules and stronger public investor property rights seem to help attracting arbitrageurs and country-level legal origins differences appear to be related to investor attraction. The reviewed literature supports the notion that strong legal environments attract arbitrage investors. As arbitrage is understood to be a major driver of pricing efficiency, legal shareholder protection might offer an explanation for international mispricing differences. It is therefore hypothesised in this paper that legal shareholder protection affects capital market mispricing. The data and methodology section will further elaborate on how this research empirically examines this relationship.

3 Data

3.1 Data collection and adjustments

The dataset contains 5,433 firms across 40 countries. These firms are part of the Datastream Market World Equity List (G#LTOTMKWD). This list does not contain all firms from the 40 countries, but a selection of large companies across these countries. The findings of this research might, as a result, be rather conditional on relatively large corporations. Similar to Jacobs (2016), this research attempts to include as many countries as possible to make the results from this paper globally relevant. The trade-off here is that inclusion of small or emerging markets could lead to data availability imbalances, e.g. small sample sizes and incomplete or inaccurate data. These imbalances are further discussed in section 3.2.1, based on the descriptive statistics in table 2.

The sample period starts in January 2008 and ends in December 2018. Although this is a somewhat arbitrarily specified time frame, various considerations have been made. Firstly, this sample period is the most recent possible, while still including data regarding the global financial crisis. This makes it possible to test for robustness of the results of this investigation throughout periods of recession. It might, for example, be that shareholder protection would play a larger role in attracting investors during times of high uncertainty or negative sentiment. Stambaugh, Yu, and Yuan (2012) find that mispricing is higher during times of high level sentiment (both positive and negative). Although it is not the main focus of this paper, a comparison of the results between recession and non-recession periods might offer some interesting insights. Secondly, a too wide time frame would have created extremely large data files that would be difficult to download and process in a limited period of time. This is another reason why the total data period has been limited to ten years (the data file also consists data from 2007, which is used for lagged variable-based anomalies).

This paper will follow La Porta et al. (1998) in labeling the 40 sampled countries according to their legal origins (English, French, German & Scandinavian). Institutional variables are also directly deducted from La Porta et al. (1998) and will be used to break down the four country origins in country-specific quantifiable components:

1. Anti-director rights
2. Efficiency of judicial system
3. Rule of law
4. Accounting standards
5. One share-one vote
6. Mandatory dividend

These variables are defined and explained in more detail in Appendix A.3. While these are obtained from relatively old data, all of the above variables are based on government policies. These institutional variables are therefore assumed to be time-invariant, making them suitable for use today. Corruption Perceptions Index data are retrieved from Transparency International and vary over time. Other variables that vary over time (GDP, GDP per capita, country market capitalisation, stocks traded as % of GDP and amount of firms listed) serve as control variables and have been retrieved from The World Bank. All time-varying country variables are updated annually.

Monthly accounting data are sourced from Datastream Worldscope, representing historical data from annual and quarterly reports of publicly traded companies. These data contain income statement and balance sheet items. Daily equity market data are obtained from Datastream, containing trading prices (quoted in local currency) and volumes. Market capitalisation is calculated per stock, combining prices and the amount of shares outstanding. Observations with a missing identifier, return index or shares outstanding are dropped, as well as all daily returns over 100% or monthly returns over 300%. All stock and accounting variables have been winsorized at the 1% top and bottom level per country. These data adjustments are done to correct for incomplete data or data outliers. The dataset consists of both active and dead stocks.

The 11 anomalies are constructed at firm-level based on the available stock market and accounting data. The formulas applied to calculate the cross-sectional anomaly metrics can be found in Appendix A.2. Similar to Jacobs (2016), composite ranks will only be constructed for firm-month combinations where at least five individual anomalies can be calculated. Section 3.2 provides an overview of the firm-month combinations that provided enough data to calculate valid mispricing scores.

For calculating long-short portfolio alphas using the three-factor CAPM of Fama and French (1993), Kenneth R. French's Data Library provides market factors and risk-free returns on global and local level. However, regarding local levels, the data only include factors for North America, Europe, Asia Pacific and Japan. This is problematic for the sampled countries that are not located within one of these geographic areas. This is a good example of the earlier mentioned data availability imbalance, where emerging markets offer less data.

Griffin (2002) points out that local factors more accurately explain time-series variation in returns with lower pricing errors compared to a world factor. However, local market factors are only available for developed markets. Applying more accurate local factors to developed markets and less accurate global factors to emerging markets seems unreasonable, as developed markets would then almost certainly produce lower mispricing estimates than emerging markets. Therefore, the global market factors are applied for all countries in the dataset. This may provide slightly less accurate three-factor alphas for all countries, but using market factors with similar levels of accuracy is crucial when comparing the estimates across countries.

3.2 Descriptive statistics

3.2.1 Data sample

Table 2 describes the data sample used in this research. The earlier mentioned imbalances in data availability across countries is already visible when observing the sampled number of firms per country. Notable is that English origin countries make up for a very large proportion of the total company sample (2,520 out of 5,433), while the data sample only contains few Scandinavian origin countries (194 out of 5,433). This is not surprising, as these differences reflect stock market size differences. Nonetheless, it should be mentioned that this data imbalance could impose some problems in the decile-based portfolio methodology used in this paper. This problem is further elaborated upon in the discussion section.

As the Stambaugh et al. (2015) anomaly-based calculations require a certain amount of stock and accounting data, only a proportion of all firm-months provided sufficient data to allow for valid mispricing scores. As can be seen in table 2, only 73% of all firm-month combinations contains sufficient data to calculate anomaly-based composite mispricing scores. Again referring to data availability imbalances, it is clear that some countries provide more data for anomaly calculations than others. However, the variation of these valid mispricing score percentages does not appear be very high across the legal origin groups.

Table 1: Descriptive statistics of the data sample, containing 40 countries categorized by legal origin (La Porta et al., 1998). Firm size (market capitalisation) is converted from local currencies to US dollars only for this table for comparison purposes. Total values (number of firms and firm-months) and averages (mean/median size and firm-months with valid mispricing scores) are displayed in the bold rows for all legal origin groups and the full sample.

Country	**Total No. Firms**	**Firm-Months**	**Mean Size (in million US$)**	**Median Size (in million US$)**	**Firm-Months with Valid Mispricing Score (%)**
Australia	148	13,012	6,420	1,796	64%
Canada	237	20,938	7,482	1,548	75%
India	184	16,700	4,840	1,521	69%
Ireland	58	4,855	41,431	1,725	62%
Israel	50	4,808	11,753	715	77%
Malaysia	86	7,696	3,264	1,278	74%
New Zealand	46	3,550	954	619	64%
Nigeria	47	3,308	552	195	57%
Pakistan	48	4,918	845	222	81%
Singapore	86	7,804	4,270	1,521	78%
South Africa	66	6,378	4,033	2,032	69%
Sri Lanka	48	4,513	272	100	72%
Thailand	45	3,910	7,689	3,247	74%
United Kingdom	468	41,836	584,326	79,911	66%

United States	903	77,622	17,100	5,550	75%
English origins	**2,520**	**221,848**	**46,349**	**6,799**	**72%**
Argentina	45	4,651	211	45	83%
Belgium	88	7,958	2,919	535	68%
Brazil	94	8,441	8,074	2,477	76%
Chile	46	4,520	3,354	2,015	80%
Colombia	48	4,194	1,918	1,056	63%
France	236	22,504	21,601	1,065	69%
Greece	49	4,854	34,231	327	80%
Indonesia	47	4,554	829	1,075	65%
Italy	145	11,111	43,320	700	67%
Jordan	37	2,537	11,566	127	67%
Mexico	86	6,644	5,924	1,096	73%
Netherlands	113	9,488	4,867	424	64%
Peru	47	4,692	13,083	767	80%
Philippines	45	4,298	3,700	2,000	81%
Portugal	48	4,685	1,308	121	78%
Spain	111	8,362	112,490	1,109	64%
Turkey	47	4,305	1,893	849	77%
French origins	**1,332**	**117,798**	**15,958**	**929**	**71%**
Austria	47	4,786	2,208	963	82%
Germany	225	20,216	6,739	1,064	75%
Japan	970	92,420	4,881	1,105	78%
Switzerland	145	13,392	49,495	1,555	68%
German origins	**1,387**	**130,814**	**15,831**	**1,172**	**77%**
Denmark	46	4,134	7,174	1,663	75%
Finland	46	4,198	6,200	1,681	76%
Norway	39	3,236	20,166	1,279	72%
Sweden	63	6,137	6,546	2,680	82%
Scandinavian origins	**194**	**17,705**	**10,022**	**1,826**	**77%**
Full Sample	**5,433**	**488,165**	**26,748**	**3,244**	**73%**

3.2.2 Institutional variables

A complete overview of the (averages of) institutional variables can be found in table 3. It shows that English origin environments offer the strictest anti-director rights, above average accounting standards and below average corruption. Countries with French origins offer the least favourable equity investor environments in terms of anti-director rights, judicial system efficiency, rule of law, accounting standards and corruption. German and Scandinavian origin countries characterise themselves by efficient judicial systems, strict rule of law and very low corruption.

Table 2: Country overview of institutional variables. Averages per legal origins group are displayed in the bold rows. All variables except corruption index are obtained from La Porta et al. (1998) and are time-invariant. The corruption index (obtained from Transparency International) is updated annually and all values from 2008-2018 are averaged by country only for this table. Definitions of the variables can be found in Appendix A.3.

Country	**Anti-Director Rights**	**Efficiency of Judicial System**	**Rule of Law**	**Accounting Standards**	**One Share-One Vote**	**Mandatory Dividend**	**Corruption Index Average**
Australia	4	10.00	10.00	75	0	0.00	8.43
Canada	5	9.25	10.00	74	0	0.00	8.48
India	5	8.00	4.17	57	0	0.00	3.50
Ireland	4	8.75	7.80	n.a.	0	0.00	7.53
Israel	3	10.00	4.82	64	0	0.00	6.03
Malaysia	4	9.00	6.78	76	1	0.00	4.80
New Zealand	4	10.00	10.00	70	0	0.00	9.18
Nigeria	3	7.25	2.73	59	0	0.00	2.55
Pakistan	5	5.00	3.03	n.a.	1	0.00	2.63
Singapore	4	10.00	8.57	78	1	0.00	8.88
South Africa	5	6.00	4.42	70	0	0.00	4.44
Sri Lanka	3	7.00	1.90	n.a.	0	0.00	3.50
Thailand	2	3.25	6.25	64	0	0.00	3.57
United Kingdom	5	10.00	8.57	78	0	0.00	7.71
United States	5	10.00	10.00	71	0	0.00	7.29
English origins	**4.07**	**8.23**	**6.60**	**69.67**	**0.20**	**0.00**	**5.90**
Argentina	4	6.00	5.35	45	0	0.00	3.15
Belgium	0	9.50	10.00	61	0	0.00	7.41
Brazil	3	5.75	6.32	54	1	0.50	3.91
Chile	5	7.25	7.02	52	1	0.30	7.08
Colombia	3	7.25	2.08	50	0	0.50	3.63
France	3	8.00	8.98	69	0	0.00	6.96

Greece	2	7.00	6.18	55	1	0.35	3.99
Indonesia	2	2.50	3.98	n.a.	0	0.00	3.08
Italy	1	6.75	8.33	62	0	0.00	4.26
Jordan	1	8.66	4.35	n.a.	1	0.00	4.85
Mexico	1	6.00	5.35	60	0	0.00	3.35
Netherlands	2	10.00	10.00	64	0	0.00	8.65
Peru	3	6.75	2.50	38	1	0.00	3.65
Philippines	3	4.75	2.73	65	0	0.00	3.01
Portugal	3	5.50	8.68	36	0	0.00	6.14
Spain	4	6.25	7.80	64	0	0.00	6.14
Turkey	2	4.00	5.18	51	0	0.00	4.53
French origins	**2.47**	**6.58**	**6.17**	**55.07**	**0.29**	**0.10**	**4.93**
Austria	2	9.50	10.00	54	0	0.00	7.54
Germany	1	9.00	9.23	62	0	0.00	7.94
Japan	4	10.00	8.98	65	1	0.00	7.59
Switzerland	2	10.00	10.00	68	0	0.00	8.73
German origins	**2.25**	**9.63**	**9.55**	**62.25**	**0.25**	**0.00**	**7.95**
Denmark	2	10.00	10.00	62	0	0.00	9.21
Finland	3	10.00	10.00	77	0	0.00	9.04
Norway	4	10.00	10.00	74	0	0.00	8.56
Sweden	3	10.00	10.00	83	0	0.00	9.04
Scandinavian origins	**3.00**	**10.00**	**10.00**	**74.00**	**0.00**	**0.00**	**8.96**
Sample average	**2.95**	**8.61**	**8.08**	**65.25**	**0.19**	**0.02**	**6.94**

4 Methodology

4.1 Constructing long-short portfolios using composite anomaly scores

The implementation of the 11 combined anomalies to quantify mispricing will mainly follow Stambaugh et al. (2015) and Jacobs (2016). Based on anomaly ranking scores, long-short portfolios will be constructed per country, after which the three-factor model will be applied to calculate risk-adjusted excess returns (alphas). Country-level alpha differences will subsequently be evaluated based on differences in law environments. The whole data merging and analysis was performed in STATA.

The 11 anomaly measures that are incorporated in the composite pricing score are described in Appendix A.1. As described per anomaly, the measures predict stock returns based on certain accounting or stock measures. These anomalies often represent lagged effects, because the market usually needs a certain amount of time to react to new public information, depending on the anomaly.

After calculating the anomaly measures for each firm-month based on accounting and stock data, all stocks are ranked for each anomaly-country-month combination, with the most overpriced (underpriced) stocks receiving the highest (lowest) ranking. The ranks are uniformly distributed over the interval (0,1], which would ideally (when all data are available) result in all stocks being assigned eleven different anomaly ranks per month. Arithmetic monthly averages are then calculated per stock based on its individual anomaly ranking score, resulting in one computed anomaly score per stock-month. Based on these scores, all stocks are again ranked per country-month, uniformly distributed over the (0,1] interval.

Each stock will consequently have a mispricing score, with high (low) values indicating potential overpricing (underpricing). It is important to note that these are only relative metrics, indicating relative over- or underpricing within the cross-section. A stock being assigned the lowest score within a country-month combination will therefore only mean that this stock is the least likely to be overpriced within its group, which does not necessarily imply underpricing.

Next, the calculated scores are used to compose two groups, consisting of the highest (most overpriced) and lowest (most underpriced) deciles of mispricing per country-month combination. This will result in two equally-sized groups of overpriced and underpriced groups per country for each month. Equal-weighted and value-weighted portfolios will then be constructed per country using a long-short strategy, taking short positions in the top decile and long positions in the bottom decile of the mispricing scores. The weighted portfolios are constructed based on market capitalisation, with larger firms receiving a higher weight. This portfolio construction process will be repeated every month within every country, with stocks entering and leaving the portfolios based on their anomaly scores. The monthly excess returns resulting from this strategy will then be calculated per portfolio using the Fama and French (1993) three-factor CAPM.

4.2 Quantifying mispricing through construction of long-short anomaly portfolios

To examine whether this investment strategy succeeds to consistently generate abnormal returns, the three-factor model from Fama and French (1993) will be applied, using the following regression model for each country-month combination:

$$R_{i,t} - R_{f,t} = a_i + b_i \times MKT_t + s_i \times SMB_t + h_i \times HML_t + u_{i,t} \tag{3}$$

With $R_{i,t}$ as the daily returns generated by the country-specific long-short portfolios, R_f as risk-free rate (30-day U.S. T-bill return), MKT as market risk premium (market return minus risk-free rate), SMB (Small Minus Big) as size premium and HML (High Minus Low) as value (B/M) premium.

The regressions will estimate coefficients for the three market factors and an intercept value for each country-month combination. As Stambaugh et al. (2015) estimate mispricing based on portfolio alphas, the intercept (a_i) will be the most important estimate from the regression output. The intercept is the expected mean value of Y when all independent variables equal zero, which will be interpreted as risk-adjusted excess returns. The objective here is to determine whether these risk-adjusted returns are non-zero values, which would imply persistent mispricing. This will be statistically tested. The regressions calculate standard errors for all parameter estimates, which indicate how much uncertainty is associated with a coefficient estimate. Z-scores of the three-factor alphas will then be calculated as:

$$Z - score = \frac{Alpha\ coefficient}{Standard\ error} \tag{4}$$

The z-scores serve as a measure of standard deviation. Assuming a normal distribution of the alphas, p-values are derived using the calculated z-scores. Low p-values indicate that it is unlikely that alpha deviations from zero are caused by a random pattern or noise. Therefore, when low p-values (below the 10% level) are observed for non-zero alphas, these will be considered to be statistically significantly different from zero.

The statistical significance, sign and weight of each alpha can be observed to see whether the mispricing-based portfolios persistently generate positive risk-adjusted returns. A long-short portfolio consistently generating excess returns, suggests a predictability of stock mispricing based on the anomaly scores. The height of the alphas represent the extent to which of the country's capital market is sensitive to the 11 anomalies. Portfolios can be constructed for any country and any group of countries, which will be done in the following analyses.

4.3 Interpreting mispricing scores

4.3.1 Regression model specification

Statistically significant and positive alphas suggest structural country-wide mispricing. As this mispricing is now quantified, the focus can shift towards explaining the mispricing. Using country-wide variables, this research aims to explain the mispricing alphas within the context of legal shareholder protection. Based on La Porta et al. (1997), the lowest alphas are expected to be observed in the countries that provide stronger shareholder protection and highest alphas where shareholder protection is the weakest.

Following La Porta et al. (1998), regressions will be used to examine and test the specific determinants of capital market mispricing. Since the dataset of this research also contains variables that vary over time, a panel data regression will be applied. The time-invariant legal origin dummies and institutional variables are treated as fixed effects and are allowed to be correlated with other included explanatory variables. In order to test for the relationship between mispricing and institutional factors, the following two models will be applied in a Pooled Ordinary Least Squares (POLS) regression:

$$\begin{aligned} Alpha_{i,t} = a + b \times French_i + c \times German_i + d \\ \times Scandinavian_i + \sum_{z=1}^{z=5}(k_z \times Control_{i,t,z}) + u_{i,t} \end{aligned} \tag{5}$$

$$\begin{aligned} Alpha_{i,t} = a + b \times French_i + c \times German_i + d \times Scandinavian_i + e \\ \times Efficiency_i + f \times Rule_i + g \times Accounting_i + h \times Vote_i + i \\ \times Dividend_i + j \times Corruption_{i,t} + \sum_{z=1}^{z=5}(k_z \times Control_{i,t,z}) + u_{i,t} \end{aligned} \tag{6}$$

With monthly country alphas as dependent variables. $French_i$, $German_i$ and $Scandinavian_i$ are dummy variables for legal origins (English origins when all dummies are zero). $Efficiency_i$, $Rule_i$, $Accounting_i$, $Vote_i$ and $Dividend_i$ (see data section for scaling per variable) are time-invariant variables. Inclusion of these variables prohibits controlling for other non-specified fixed effects, which potentially results in the used models ignoring other forms of country-level variation (omitted variable bias). Specifying time-invariant regression variables is therefore done at the risk of generating biased estimators. $Corruption_i$ is based on an annually updated index and therefore does vary over time. The control variables included in the models are relative trading volume (as % of GDP), log of listed companies, log of country stock market capitalisation, log of GDP and log of GDP per capita. These are all time-varying. $u_{i,t}$ represents the error term.

Given the time-invariant and macro-economic nature of institutional variables that are established by governments, reversed causality seems unlikely. It also seems unlikely that mispricing affects corruption, trading volume, the number of listed companies, total country market capitalisation, GDP or GDP per capita, since these are variables that are mostly affected by external macro-level factors.

Model 5 includes legal origin dummies and several control variables. This model is aimed to test the direct effect that legal origin has on the estimated mispricing per country after controlling for country size, wealth and trading volume. Significant dummy coefficients would imply that legal origin affects capital market efficiency. Model 6 adds individual institutional variables to the equation. Coefficient significance of these variables suggests capital market sensitivity to these specific institutional factors.

Should there be any significance of legal origin dummies found in model 5, then addition of the institutional variables should reduce the explanatory role of legal origins. This is expected, because the institutional variables are supposed to be part of the effect that legal environment has on mispricing. For example, if the coefficient for the German origins dummy turns insignificant after adding a rule of law variable, a conclusion could be that German origin countries reduce mispricing through a strong rule of law.

4.3.2 Regression diagnostics

Before running the regressions, it is important to test for the assumptions that the regression model makes. One of the assumptions that P-OLS makes, is homoskedasticity of the error terms. This means that the variance of all error terms would be constant across the plot. A violation of this assumption would result in the regression estimates becoming less efficient, increasing the likelihood that the coefficient estimates are further from their correct values. However, heteroskedasticity would still leave the parameter estimates unbiased. Table 3 shows the Breusch-Pagan test results for heteroskedasticity. As p-values are high, the null hypothesis can not be rejected for any of the models and portfolios. This suggests that the variances of the error terms are constant, meaning that homoskedasticity holds. It is therefore concluded that heteroskedasticity does not form a problem in the used panel dataset. For visualisation purposes, the corresponding scatter plots of the error terms can be found in Appendix A.4.

Table 3: Breusch-Pagan test for heteroskedasticity. Test statistics are shown for model 5 (legal origin dummies and controls as independent variables) and model 6 (institutional variables added). This is done for both the equal-weighted and value-weighted portfolio three-factor alphas. Under the null hypothesis, homoskedasticity is assumed. P-values (Prob > χ^2) under 0.01, 0.05 and 0.1 suggest rejection of the null hypothesis at the 1%, 5% and 10% level respectively.

	Equal-weighted alphas		Value-weighted alphas	
Breusch-Pagan test	Model 5	Model 6	Model 5	Model 6
Chi square (χ^2) statistic	0.030	0.840	2.700	1.320
Prob > χ^2	0.872	0.360	0.100	0.251

Ho: constant variance - homoskedasticity

Another potential problem that should be tested for is serial correlation. This means that the error terms ($u_{i,t}$) are directly correlated with error terms in their ensuing years ($u_{i,t+1}$). The Wooldridge test will be applied to test for first-order serial correlation (table 4). As all p-values are below 0.01, the null hypothesis is strongly rejected for both models and both portfolios. Similar to heteroskedasticity, this problem will not cause the coefficient estimates to become biased. Nonetheless, serial correlation is still considered a problem, as it does affect the efficiency of the estimators. As the regression will underestimate standard errors, it will conclude that the estimates are more precise than they really are. This might lead to coefficient estimates being falsely considered as significantly different from zero. To prevent this from happening, robust standard errors of White et al. (1980) are used to account for this diagnosed serial correlation. This does not change the estimation results, but only the assessment of standard errors and therefore also the statistical significance of the variable coefficients.

Table 4: Wooldridge test for serial correlation. Test statistics are shown for model 5 (legal origin dummies and controls as independent variables) and model 6 (institutional variables added). This is done for both the equal-weighted and value-weighted portfolio three-factor alphas. Under the null hypothesis, no first-order serial correlation is assumed. P-values (Prob > F) under 0.01, 0.05 and 0.1 suggest rejection of the null hypothesis at the 1%, 5% and 10% level respectively.

	Equal-weighted alphas		Value-weighted alphas	
Wooldridge test	Model 5	Model 6	Model 5	Model 6
Degrees of freedom	34	29	34	29
F-statistic	29.753	12.502	16.060	9.845
Prob > F	0.000	0.001	0.000	0.004

H0: no first-order serial correlation

4.3.3 Robustness check: coefficient interaction with the global financial crisis

As briefly discussed in the data section, the data sample period also offers an opportunity to explore the suggested mispricing mechanism within and outside the global financial crisis. A robustness check of the regression coefficients during the global financial crisis will therefore be performed. To do so, a recession dummy will be created, which equals 1 during the recession (2008-2009) and 0 after the recession (2010-2018). This dummy will be added to regression model 6 to test for interaction effects between the recession dummy and all independent variables. Again, this will be done for the equal-weighted and value-weighted portfolios. The regression model is formulated as follows:

$$\begin{aligned}Alpha_{i,t} = a &+ b \times French_i + c \times German_i + d \times Scandinavian_i + e \\ &\times Efficiency_i + f \times Rule_i + g \times Accounting_i + h \times Vote_i + i \\ &\times Dividend_i + j \times Corruption_{i,t} + \textstyle\sum_{z=1}^{z=5}(k_z \times Control_{i,t,z}) + l \\ &\times Recession_t + \textstyle\sum_{z=1}^{z=9}(m_z \times IndepInvariant_{i,z} \times Recession_t) \\ &+ \textstyle\sum_{z=1}^{z=6}(n_z \times IndepVariant_{i,t,z} \times Recession_t) \\ &+ u_{i,t}\end{aligned} \tag{7}$$

With $IndepInvariant_{i,z}$ representing all the time-invariant independent variables from model 6. $IndepVariant_{i,z}$ represents all time-variant variables. Model 7 includes interaction effects with all variables from model 6, including the intercept. Coefficients a-k_z will display the variable coefficients in periods outside the recession. Coefficients l, m_z and n_z will describe how these coefficients change during the recession. These changes will also be tested for statistical significance. This examination of the interaction effects could provide insights on how the sensitivity of mispricing to institutional quality changes in times of economic downturn. However, given that this is an extension rather than the main purpose of this research, the results from this analysis should be interpreted with discretion.

5 Results

5.1 Cross-sectional analysis: three-factor alphas across groups

Table 5 shows the portfolio alphas for each individual anomaly and for the composite anomaly score. For quick comparison purposes, anomaly portfolios of common and civil law countries are constructed. The alphas indicate levels of mispricing for the portfolios of either common or civil law countries.

For the equal-weighted long-short portfolios, significant three-factor alphas for all 11 individual anomalies and for the composite anomaly measure are generated. For the value-weighted portfolios, most anomalies also generate significant excess returns portfolios, although these appear to be smaller and less significant than the equal-weighted alphas. The strongly positive significant alphas across most anomalies suggests a generally high sensitivity of stock prices to the 11 anomalies and the composite mispricing score from Stambaugh et al. (2015).

The long-short portfolio strategy is successful in persistently generating positive risk-adjusted returns when invested in a global portfolio of stocks. However, the anomaly alphas show little to no consistent alpha differences between the two origin groups across the 11 individual anomaly measures. The composite mispricing score does not seem to do a better job in pointing out any mispricing differences between these portfolios. The results from table 5 do, therefore, not provide convincing evidence for any structural differences in capital market mispricing between countries with common and civil law origins.

The anomaly three-factor portfolio alphas per country are displayed on the bar charts of figure 1. The country portfolio alphas are a good representation of the alphas that are used in the panel data regression. The equal-weighted (value-weighted) portfolios generate significantly positive excess returns for 12 (10) countries out of the 40 sampled countries. The average equal-weighted (value-weighted) three-factor alpha is about 28 (40) basis points per month. Although these bar charts cannot be used to make any definitive conclusions, the equal-weighted portfolio seems to generate higher risk-adjusted returns for developing countries, such as Indonesia, Nigeria, Thailand, Philippines, Peru and Argentina. More developed countries, such as the U.K., Canada, Singapore, the U.S., the Netherlands, Austria, Sweden and Germany tend to show low and statistically insignificant alphas. The bar chart might imply that mispricing is correlated to country wealth, which corresponds with the notion that emerging markets are less efficient. Jacobs (2016), however, has cast doubt on this notion through application of the Stambaugh et al., 2015 mispricing score. The value-weighted portfolio alphas do not seem to expose a similar pattern, which corresponds more to the findings of Jacobs (2016).

The cross-sectional analyses confirm that positive alphas can be consistently generated through a long-short strategy based on anomaly measures, including the Stambaugh et al. (2015) composite score. Mispricing differences between countries with different law origins have not been found, which is inconsistent with the expectation that capital market efficiency would be related to law origins. Hence, only comparing mispricing across countries based on their common or civil legal origins does not explain mispricing differences. The panel data regressions in the following section will include more specific legal origin and institutional variables, while also controlling for other non-institutional variables, such as size, wealth and trading volume.

Table 5: Three-factor alphas for equal-weighted and value-weighted long-short portfolios per anomaly measure. The long-short portfolios are constructed by taking long (short) positions in the stocks (top deciles) that are most likely to be underpriced (overpriced), according to the applied anomaly measure. Portfolios are constructed under common and civil law groups. See Appendix A.1 and A.2 for definitions and formulas for the 11 individual anomalies. "Mispricing score" refers to the composite anomaly mispricing metric composed by Stambaugh, Yu, and Yuan (2015).

	Equal-weighted alphas		Value-weighted alphas	
	Common	Civil	Common	Civil
Financial distress	0.0122***	0.0055***	0.0014	0.002
	(5.91)	(3.00)	(0.85)	(1.37)
O-score bankruptcy probability	0.0058***	0.0072***	0.0030***	0.0024***
	(4.10)	(6.29)	(3.58)	(4.05)
Net stock issues	0.0069***	0.0049***	0.0042***	0.0021***
	(7.98)	(5.65)	(5.58)	(3.10)
Composite equity issues	0.0027**	0.0038***	0.0022**	0.0002
	(2.42)	(3.59)	(2.18)	(0.18)
Total accruals	0.0040***	0.0045***	0.0024**	0.0018**
	(3.03)	(3.92)	(2.53)	(2.37)
Net operating assets	0.0067***	0.0044***	0.0028***	0.001
	(4.68)	(3.66)	(3.14)	(1.25)
Momentum	0.0074***	0.0050***	0.0034***	0.0017*
	(5.64)	(4.10)	(3.02)	(1.86)
Gross profitability	0.0040***	0.0026**	0.0035***	0.0012
	(3.61)	(2.35)	(3.22)	(1.37)
Asset growth	0.0033***	0.0030***	0.0022***	0.0018***
	(2.92)	(2.94)	(3.04)	(2.90)
Return on assets	0.0076***	0.0075***	0.0033**	0.0024***
	(5.29)	(6.85)	(2.41)	(2.66)
Investment to assets	0.0053***	0.0035***	0.0030**	0.0015**
	(3.77)	(3.14)	(2.57)	(2.28)
Mispricing score	0.0048***	0.0050***	0.0021**	0.0021***
	(4.07)	(5.00)	(2.30)	(2.72)

T-statistics in parentheses, based on robust standard errors.

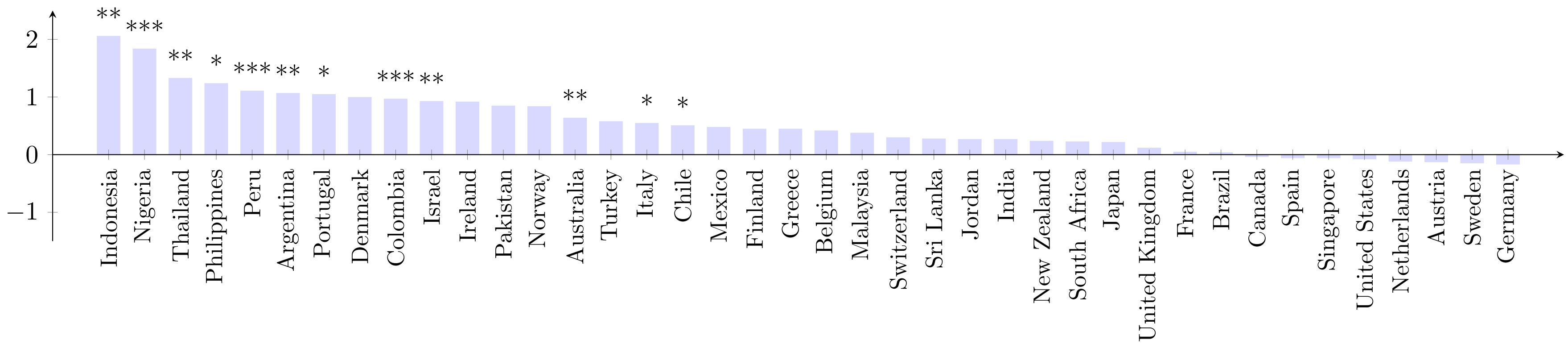
*** Significant at the 1 percent level.

** Significant at the 5 percent level.

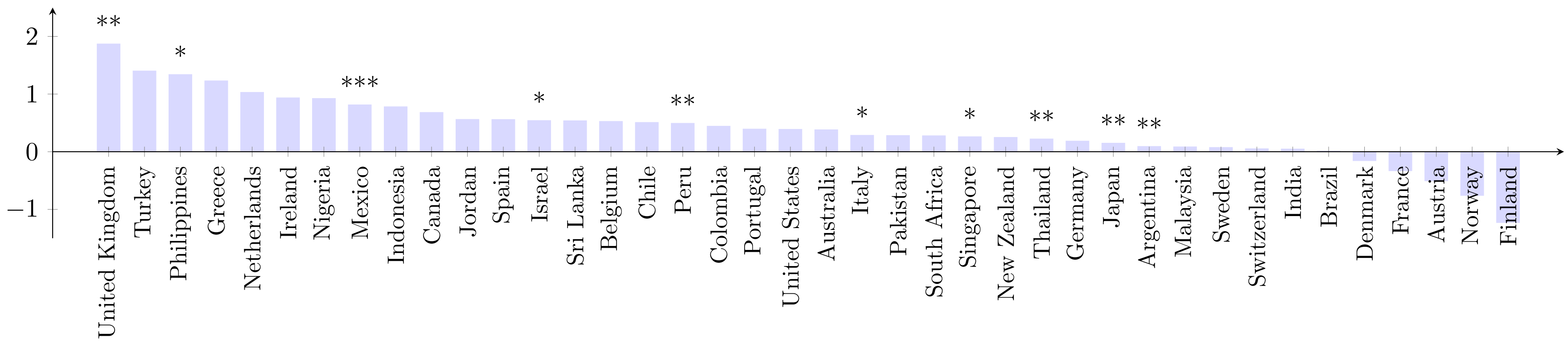
* Significant at the 10 percent level.

Figure 1: Mispricing across countries. The bar charts display equal-weighted and value-weighted portfolio alphas, as obtained from regressing Stambaugh et al. (2015) long-short portfolios using the Fama and French (1993) three-factor model. Portfolios have been been constructed monthly for each country. The sample period is January 2008 to December 2018, or a shorter time period depending on data availability. *, ** and *** respectively indicate 10%, 5% and 1% (two-tailed) significance levels. The alphas are displayed in percentages.

Equal-weighted long-short portfolio three-factor alphas



Value-weighted long-short portfolio three-factor alphas



5.2 Panel Data Regressions

5.2.1 Full sample regressions

Table 6 and table 7 show the regression coefficients for equal-weighted and value-weighted portfolio alphas respectively. The "Basic regression" column (model 5) only consists of legal origin dummies and of several control variables. The control variables are relative trading volume (as % of GDP), log of listed companies, log of country stock market capitalisation, log of GDP and log of GDP per capita. These control variables should account for size, wealth and trading volume effects within the sampled countries. "Incl. investor rights" (model 6) includes specific policy-related variables that constitute the institutional environments. The legal origins, anti-director rights, efficiency of judicial system and rule of law are deducted from La Porta et al. (1998) and are assumed be invariant over time. The corruption index, relative trading volume (as % of GDP), log of listed companies, log of market capitalisation, log of country GDP and log of country GDP per capita are time-varying variables.

Both basic regressions indicate higher mispricing for Scandinavian civil-law countries, as both the equal-weighted and value-weighted long-short portfolios yield strongly significant positive excess returns for this group. As the individual institutional variables are added, the Scandinavian origins dummy remains significant. This indicates that the added institutional factors do not manage to explain the significance of the Scandinavian dummy. This is inconsistent with the theoretical framework, as these institutional variables are assumed to at least partially constitute the legal environment of the origin group. However, this is not a surprising finding, as Denmark and Finland have been excluded from both regressions. The only remaining Scandinavian countries in the regression are Norway and Sweden. It seems therefore highly likely that the higher mispricing in Scandinavian countries is related to a shortage of data availability, as already discussed in the data section.

The regression tables also show information regarding individual institutional variables (Incl. investor rights). Anti-director rights and efficiency of the judicial system have significant and negative effects on portfolio excess returns for both equal-weighted and value-weighted portfolios. This implies that stricter anti-director rights and more efficient judicial systems can reduce capital market pricing inefficiencies. For value-weighted returns, rule of law also significantly decreases excess portfolio returns, which suggests that stronger enforcement of investor protection policies reduces mispricing. Furthermore, value-weighted portfolio alphas are significantly and positively related to the corruption index. This is a rather surprising finding, as this implies that more corrupt countries would have, all else equal, more efficient markets.

Accounting standards, one share-one vote policies and mandatory dividends do not seem to affect pricing efficiency on capital markets. Especially the insignificant coefficient for accounting standards is surprising, given that Morck et al. (2000) have explicitly found that strong accounting standards attract arbitrage investors.

Another remarkable finding is that countries with higher trading volumes experience more mispricing. This contradicts the findings of Chordia, Roll, and Subrahmanyam (2011). They find an upward trend in trading volume, particularly in stocks with greater levels of institutional holdings. Combining this with another upward trend in pricing efficiency, they conclude that market efficiency has increased in response to increased information-based institutional trading. Trading volume, however, only serves as a control variable. Hence, this peculiar finding does not affect any conclusion made in this paper.

Country size (GDP) reduces alphas for only value-weighted portfolios. Country wealth (GDP per capita) reduces mispricing for equal-weighted and value-weighted portfolios. Consistent with the bar chart of the equal-weighted portfolio alphas in the previous subsection, country wealth is significantly related to capital market efficiency. The significantly negative coefficient suggests that more developed countries have more efficient markets. This finding opposes Jacobs (2016), who finds no significant relation between country wealth and mispricing.

Table 6: Panel data regressions with equal-weighted portfolio three-factor alphas as dependent variable. The ”Basic regression” column displays the regression coefficients from model 5, containing only legal origin dummies (English origins when all three equal zero) and control variables. The ”Incl. investor rights” column represents model 6, where individual institutional variables are added. Denmark, Finland and the U.K. are excluded from ”Basic regression” due to missing control variable data. Additionally, Indonesia, Ireland, Jordan, Pakistan and Sri Lanka are excluded from ”Incl. investor rights” due to missing institutional variables from La Porta et al. (1998).

Independent variable	Basic regression	Incl. investor rights
French origins	0.000	-0.001
	(0.005)	(0.004)
German origins	-0.000	0.001
	(0.005)	(0.004)
Scandinavian origins	0.018***	0.025***
	(0.006)	(0.003)
Anti-director rights		-0.003**
		(0.001)
Efficiency of judicial system		-0.002**
		(0.001)
Rule of law		0.001
		(0.001)
Accounting standards		-0.000
		(0.000)
One share-one vote		-0.001
		(0.003)
Mandatory dividend		0.004
		(0.011)
Corruption index		0.002
		(0.002)
Stocks traded (% of GDP)	0.009	0.015**
	(0.007)	(0.006)
Log of listed companies	-0.002	0.004*
	(0.002)	(0.002)
Log of market capitalisation	0.002	-0.003
	(0.004)	(0.004)
Log of country GDP	-0.003	-0.004
	(0.003)	(0.003)
Log of country GDP per capita	-0.005*	-0.013***
	(0.003)	(0.004)
Constant	0.096	0.311***
	(0.065)	(0.064)
Country-months (observations)	276	235
Number of countries	37	32
R-squared	0.024	0.088
Adjusted R-squared	-0.005	0.026

Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Table 7: Panel data regressions with value-weighted portfolio three-factor alphas as dependent variable. The "Basic regression" column displays the regression coefficients from model 5, containing only legal origin dummies (English origins when all three equal zero) and control variables. The "Incl. investor rights" column represents model 6, where individual institutional variables are added. Denmark, Finland and the U.K. are excluded from "Basic regression" due to missing control variable data. Additionally, Indonesia, Ireland, Jordan, Pakistan and Sri Lanka are excluded from "Incl. investor rights" due to missing institutional variables from La Porta et al. (1998).

Independent variable	Basic regression	Incl. investor rights
French origins	-0.001	-0.003
	(0.006)	(0.007)
German origins	0.000	-0.001
	(0.005)	(0.007)
Scandinavian origins	0.020***	0.028***
	(0.006)	(0.004)
Anti-director rights		-0.005***
		(0.001)
Efficiency of judicial system		-0.002**
		(0.001)
Rule of law		-0.002**
		(0.001)
Accounting standards		-0.000
		(0.000)
One share-one vote		-0.003
		(0.004)
Mandatory dividend		0.003
		(0.010)
Corruption index		0.004**
		(0.002)
Stocks traded (% of GDP)	0.010	0.016***
	(0.008)	(0.006)
Log of listed companies	-0.001	0.006**
	(0.003)	(0.003)
Log of market capitalisation	0.003	-0.002
	(0.005)	(0.004)
Log of country GDP	-0.006*	-0.007**
	(0.003)	(0.003)
Log of country GDP per capita	-0.005	-0.007*
	(0.003)	(0.004)
Constant	0.132*	0.297***
	(0.075)	(0.061)
Country-months (observations)	276	235
Number of countries	37	32
R-squared	0.030	0.093
Adjusted R-squared	0.001	0.031

Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

5.2.2 Robustness check: mispricing during and after the global financial crisis

As mentioned in earlier sections, a reasonable suggestion might be that the relation between mispricing and institutional quality is conditional on time-specific factors, such as economic cycles or market sentiment. No specific recession or sentiment data has been gathered for this research. Nevertheless, an examination of the coefficients across specific time periods could provide rough insights on how sensitive this relation is to economic circumstances. Regression model 7 is performed to test for interactions between the coefficients from model 6 and the global financial crisis (2008-2009). The results are shown in table 8 and table 9. "No recession" shows the institutional variables under non-recession (2010-2018) conditions and "Recession interaction" shows how the coefficients change during periods of recession. The interaction effects are also tested for statistical significance.

Both table 8 and table 9 show few statistically significant coefficient changes during the recession. This implies that most coefficients of the independent variables tend to remain identical during and outside the recession. The few significant differences indicate an insensitivity of the coefficients to periods of economic downturn. Efficiency of the judicial system reduces mispricing both during and outside times of recession. Strong accounting standards now also significantly reduce mispricing.

Anti-director rights and mandatory dividends are the only institutional factors that are significantly different during the recession compared to non-recession periods. Stricter anti-director rights appear to be even more successful in reducing mispricing during a recession, for both equal-weighted and value-weighted portfolios. Mandatory dividends are found to increase mispricing during the recession for equal-weighted portfolios, which contradicts the notion that stronger investor protection reduces mispricing. However, the institutional variables generally seem to reduce mispricing, both during and after the recession.

Another noteworthy finding, is the positive effect that the relative amount of stocks traded has on mispricing only during the crisis, for both equal-weighted and value-weighted portfolios. This implies an interaction effect between economic stress and trading activity, which could provide an explanation for the earlier found positive relation between mispricing and trading volume in table 6 and table 7. This is consistent with Brown (1999), who finds that both trading volume and the proportion of irrational (noise) traders increase in times of strong sentiment. Within this context, higher trading volumes in times of recession could point to more noise trading activity, which increases mispricing. Outside of recessions, no significant effect of trading volume on mispricing has been found.

GDP per capita only decreases mispricing during the crisis (for equal-weighted portfolios), which suggests a negative country wealth effect on mispricing that is conditional on times of economic downturn. Outside of recessions, the results of this research are more consistent with Jacobs (2016), who find no relation between country wealth and mispricing.

Also remarkable is that the constant component of the alphas is substantially higher during the crisis for equal-weighted and value-weighted portfolios. This suggests that crisis-specific factors increase mispricing globally during 2008-2009, that are not covered by the model used in this research. An example of one of these factors is investor sentiment: Stambaugh et al. (2012) find that their long-short portfolio returns are higher during times of high-level sentiment. However, this paper cannot provide any evidence that the detected mispricing is in any way related to sentiment.

In summary, table 8 and table 9 suggest that institutional factors still affect mispricing during recessions, although it is difficult to predict which factors do so specifically and how large the effects are. Anti-director rights even seem to play a larger role in reducing mispricing during recessions. It seems that recessions can explain some of the surprising results from table 6 and table 7, for control variables such as trading volume and GDP per capita. The effects that these factors have on mispricing show a tendency to contradict previous literature under circumstances of economic stress. Furthermore, other recession-specific factors that are unexplained by the models used in this research appear to be reflected in a recession-specific alpha component.

Table 8: Robustness check: recession interaction effect on the sensitivity of equal-weighted portfolio alphas to institutional factors. The table displays the coefficients outside the recession (2010-2018) under column "No recession". "Recession interaction" represents the extent to which the coefficients differ during the recession (2008-2009). Denmark, Finland, Indonesia, Ireland, Jordan, Pakistan, Sri Lanka and Sweden are excluded from the regression.

Independent variable	No recession	Recession interaction
French origins	-0.004	-0.003
	(0.002)	(0.016)
German origins	0.002	0.013
	(0.002)	(0.018)
Scandinavian origins	0.020***	-0.015
	(0.002)	(0.014)
Anti-director rights	0.000	-0.009*
	(0.001)	(0.005)
Efficiency of judicial system	-0.002**	-0.001
	(0.001)	(0.003)
Rule of law	0.001	0.005
	(0.001)	(0.003)
Accounting standards	-0.000***	0.000
	(0.000)	(0.001)
One share-one vote	0.000	-0.004
	(0.002)	(0.009)
Mandatory dividend	-0.009	0.050*
	(0.006)	(0.026)
Corruption index	-0.001	0.004
	(0.001)	(0.004)
Stocks traded (% of GDP)	-0.003	0.044***
	(0.003)	(0.014)
Log of listed companies	0.001	0.010
	(0.002)	(0.008)
Log of market capitalisation	0.002	-0.010
	(0.002)	(0.010)
Log of country GDP	-0.005**	-0.005
	(0.002)	(0.011)
Log of country GDP per capita	-0.003	-0.028**
	(0.002)	(0.013)
Constant	0.120***	0.545**
	(0.043)	(0.225)
Country-months (observations)	235	
Number of countries	32	
R-squared	0.183	
Adjusted R-squared	0.059	

Robust standard errors in parentheses.
*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table 9: Robustness check: recession interaction effect on the sensitivity of value-weighted portfolio alphas to institutional factors. The table displays the coefficients outside the recession (2010-2018) under column "No recession". "Recession interaction" represents the extent to which the coefficients differ during the recession (2008-2009). Denmark, Finland, Indonesia, Ireland, Jordan, Pakistan, Sri Lanka and Sweden are excluded from the regression.

Independent variable	No recession	Recession interaction
French origins	-0.009	-0.014
	(0.007)	(0.014)
German origins	-0.005	-0.007
	(0.006)	(0.019)
Scandinavian origins	0.022***	-0.018
	(0.003)	(0.016)
Anti-director rights	-0.003***	-0.006*
	(0.001)	(0.004)
Efficiency of judicial system	-0.002**	0.000
	(0.001)	(0.004)
Rule of law	-0.002	0.002
	(0.001)	(0.004)
Accounting standards	-0.000*	0.001
	(0.000)	(0.001)
One share-one vote	-0.003	-0.003
	(0.004)	(0.015)
Mandatory dividend	0.000	0.008
	(0.007)	(0.034)
Corruption index	0.002	0.003
	(0.002)	(0.006)
Stocks traded (% of GDP)	0.000	0.033**
	(0.005)	(0.016)
Log of listed companies	0.004	0.009
	(0.003)	(0.009)
Log of market capitalisation	0.003	-0.005
	(0.003)	(0.015)
Log of country GDP	-0.006	-0.008
	(0.004)	(0.015)
Log of country GDP per capita	0.003	-0.023
	(0.006)	(0.014)
Constant	0.116	0.439**
	(0.073)	(0.200)
Country-months (observations)	235	
Number of countries	32	
R-squared	0.141	
Adjusted R-squared	0.010	

Robust standard errors in parentheses.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

6 Discussion

Before drawing any conclusions based on the findings in the result section, it should be recognised that both the data collection and the methodology contain limitations that should be taken into account when interpreting the results. These limitations are discussed in the section below. Besides, this discussion section is also intended to provide suggestions for further research.

The limited data availability in this research comes clear in the stock mispricing that has been found on Scandinavian markets. The question can be raised whether this effect is the result of specific Scandinavian factors or simply the lack of data. As discussed in the data section, Scandinavian countries have a strong disadvantage when it comes to data availability. Firstly, Scandinavian countries have by far the smallest amount of companies in the data sample. This could be problematic, combined with the chosen methodology. By definition, a decile-based long-short strategy results in more diversified portfolios for larger countries. Larger countries contain more stocks per decile group and as such minimise the role of idiosyncratic risk in their corresponding portfolios. With less idiosyncratic risk, portfolio returns are more likely to reflect market returns. Secondly (and probably most importantly), the regression analyses drop Denmark, Finland, and later Sweden when institutional variables are included, due to missing data. This means that the Scandinavian dummy coefficient is mostly based on Norway, which strongly reduces the meaning of this coefficient. It is highly likely that incomplete data explain the mispricing found in Scandinavian countries, rather than Scandinavia-specific factors.

As described in the data section, Kenneth R. French's Data Library lacks market factors for geographic areas outside North America, Europe, Asia Pacific and Japan. Since this research aims to investigate mispricing on a global level, this data limitation forced the use of global market factors. The use of these global market factors may have led to less accurate mispricing estimates, as previous research has found that local factors explain more time-series variation in returns and have lower pricing errors than the world factor model (Griffin, 2002). The most accurate method would be to manually calculate the market factors per country. Another option might have been to exclude countries that are outside the areas where Kenneth R. French provides market factors. This is a sub-optimal solution, as this option would result in this research being conditional on mostly wealthy countries.

Another shortcoming of the data sample used is that the country-specific institutional variables are relatively old (based on La Porta et al., 1998). The data are gathered in a period from 1980 to 1995. These variables are assumed to be invariable over time, concluding that these values would be similar to the actual variables during the investigated period. However, this assumption becomes harder to make as the data become older. Therefore, performing this research with updated data could generate results that are more representative of current world markets.

Furthermore, the dataset used in this research lacked industry information for each company. As a result, it was not possible to filter forms of non-common equity, such as mutual funds, index funds or Real Estate Investment Trusts (RE-

ITs). The role of managers and boards of directors might differ significantly from that of common equities. These types of listed entities might therefore present different agent-principal relationship dynamics than the targeted common equities. Following Jacobs (2016) in dropping non-common entities would have created a dataset that is more likely to represent behaviour of common equities.

Although section 5.2.2 provides brief insights on the arbitrage attraction mechanism during times of recession, this has clearly not been the main focus of this paper. The results have indicated that institutional factors still play a role in explaining mispricing, but the usefulness of the information that table 8 and 9 provide is rather limited. For example, a significantly higher constant component has been found for the alphas during the recession, indicating higher mispricing in times of economic downturn. However, in order to explain this recession-specific mispricing, factors that are not part of this investigation will need to be included. Addition of variables such as the investor sentiment index from Baker and Wurgler (2006) might provide a more detailed look at mispricing during recessions. This would also allow for analysis of positive sentiment periods, which has not been possible with the data used for this research.

In the context of further research, it is important to emphasize that this paper is only aimed to point out the existence of a relation between institutional quality and equity mispricing. This focus on institutional factors leaves room for many other explanations, extensions and methods of measure. Examples of plausible explanations for the mispricing found in this research are limits to arbitrage and investor sentiment. Both have been backed by previous research. Shleifer and Vishny (1997) explain that, in reality, arbitrageurs often bear significant idiosyncratic risk and Stambaugh et al. (2015) emphasize the high financial costs and limitations of shorting stocks compared to buying stocks. Corredor, Ferrer, and Santamaria (2013) have shown that investor sentiment has a significant influence on returns, varying in intensity across markets. It would be reasonable to argue that an interaction effect between these factors and legal investor protection exists. However, measuring limits to arbitrage and sentiment would not serve the purpose of this research. It could be interesting for following research to examine whether the mispricing found in this research would survive after correction for these specific effects or whether any interaction effects would appear.

Lastly, it remains important to remember that there is no perfect method to estimate mispricing. A combination of creating anomaly-based returns and calculating their three-factor alphas potentially involves some fundamental problems. Barberis and Thaler, 2003 explain that anomaly returns sometimes also contain a rational component besides a behavioural (irrational) component. Therefore, anomalies do not always fully reflect inefficiencies or investor biases, but also rational behaviour at times. Furthermore, three-factor alphas are no perfect proxies for mispricing. Positive alphas simply imply a proportion of returns that cannot be explained by a simple model of risk and return. These are fundamental problems that are difficult to tackle, but nevertheless should be kept in mind when interpreting the results.

7 Conclusion

This paper proposes a capital market efficiency mechanism, where countries with better institutional quality attract arbitrage investors who reduce stock mispricing. A recently developed composite anomaly score constructed by Stambaugh et al. (2015) is used to quantify and compare stock mispricing across international capital markets. The empirical results in this paper provide supporting evidence for the suggested mechanism, showing a positive relationship between institutional quality and capital market efficiency.

Legal origins seem to be ineffective when it comes to explaining mispricing. No consistent mispricing differences have been found between groups of common and civil law origins. Regressions based on English, French, German and Scandinavian law origins also fail to explain persistent mispricing. Mispricing differences have been found on country-level, with more mispricing particularly in Scandinavian origin countries. It seems that this geography-specific mispricing is not related to investor protection, as individual institutional variables do not explain any of the mispricing in Scandinavian countries. This finding is most likely to be related to data limitations for this group of countries.

More interesting evidence is offered when separately examining the specific institutional components that constitute the legal environments. After correcting for size, wealth and trading volume, some individual institutional factors appear to be directly related to capital market efficiency. High quality anti-director rights, efficiency of the judicial system and a strong rule of law manage to reduce country-wide mispricing.

Institutional factors remain relevant determinants of mispricing during times of recession. Some institutional and macroeconomic factors are found to be conditional on times of recession. Stricter anti-director rights more strongly reduce mispricing during times of recession. A significant alpha component during the recession suggests existence of recession-specific effects that are not explained by the factors examined in this research.

To conclude, institutional quality has some effect on mispricing, although it is difficult to assign any mispricing differences to legal origins. It is more reasonable to break the mispricing effect down to specific institutional components, of which some are found to significantly affect capital market efficiency. Capital market efficiency therefore appears to depend on country-specifically composed sets of institutional factors, regardless of legal origins. Examination of the global financial crisis shows that institutional and macroeconomic factors still affect mispricing during recessions, albeit in a less predictable manner.

This paper introduces the relationship between mispricing and shareholder protection and as such aims to affect and inspire future research. Nevertheless, it should be evident that this paper has not been written with any normative intent, nor has its purpose been to initiate a new field of study. The findings of this paper will most probably serve best within the context of investigations that more widely explore capital market efficiency factors. Although it is uncertain to which extent shareholder protection drives equity prices, this paper demonstrates that legal shareholder protection is an important factor when it comes to understanding capital market efficiency.

References

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The journal of Finance*, *61*(4), 1645–1680.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, *9*(1), 3–18.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, *1*, 1053–1128.

Basu, S. (1983). The relationship between earnings' yield, market value and return for nyse common stocks: Further evidence. *Journal of financial economics*, *12*(1), 129–156.

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *The journal of finance*, *43*(2), 507–528.

Brown, G. W. (1999). Volatility, sentiment, and noise traders. *Financial Analysts Journal*, *55*(2), 82–90.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, *63*(6), 2899–2939.

Chen, L., Novy-Marx, R. et al. (2011). An alternative three-factor model. *Available at SSRN 1418117*.

Chordia, T., Roll, R., & Subrahmanyam, A. (2011). Recent trends in trading activity and market quality. *Journal of Financial Economics*, *101*(2), 243–263.

Cooper, M. J., Gulen, H., & Schill, M. J. (2008). Asset growth and the cross-section of stock returns. *The Journal of Finance*, *63*(4), 1609–1651.

Corredor, P., Ferrer, E., & Santamaria, R. (2013). Investor sentiment effect in stock markets: Stock characteristics or country-specific factors? *International Review of Economics & Finance*, *27*, 572–591.

Daniel, K., & Titman, S. (2006). Market reactions to tangible and intangible information. *The Journal of Finance*, *61*(4), 1605–1643.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, *33*(1), 3–56.

Fama, E. F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of financial economics*, *82*(3), 491–518.

Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *The Journal of Finance*, *63*(4), 1653–1678.

Griffin, J. M. (2002). Are the fama and french factors global or country specific? *The Review of Financial Studies*, *15*(3), 783–803.

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

Jacobs, H. (2016). Market maturity and mispricing. *Journal of Financial Economics*, *122*(2), 270–287.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, *48*(1), 65–91.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (2002). Investor protection and corporate valuation. *The journal of finance*, *57*(3), 1147–1170.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1997). Legal determinants of external finance. *The journal of finance*, *52*(3), 1131–1150.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1998). Law and finance. *Journal of political economy*, *106*(6), 1113–1155.

Leuz, C., Nanda, D., & Wysocki, P. D. (2003). Earnings management and investor protection: An international comparison. *Journal of financial economics*, *69*(3), 505–527.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The journal of finance*, *20*(4), 587–615.

Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *The Journal of finance*, *50*(1), 23–51.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of financial economics*, *58*(1-2), 215–260.

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

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 109–131.

Reynolds, T., & Flores, A. (1989). Foreign law: Current sources of basic legislation in jurisdictions of the world. *Rothman and Co., Littleton, Colorado*.

Ritter, J. R. (1991). The long-run performance of initial public offerings. *The journal of finance*, *46*(1), 3–27.

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, *11*(3), 9–16.

Ross, S. A. a. o. (1973). *Return, risk and arbitrage*. Rodney L. White Center for Financial Research, The Wharton School.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, *19*(3), 425–442.

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of finance*, *52*(1), 35–55.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting review*, 289–315.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, *104*(2), 288–302.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2015). Arbitrage asymmetry and the idiosyncratic volatility puzzle. *The Journal of Finance*, *70*(5), 1903–1948.

Titman, S., Wei, K. J., & Xie, F. (2004). Capital investments and stock returns. *Journal of financial and Quantitative Analysis*, *39*(4), 677–700.

White, H. et al. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *econometrica*, *48*(4), 817–838.

Xing, Y. (2007). Interpreting the value effect through the q-theory: An empirical investigation. *The Review of Financial Studies*, *21*(4), 1767–1795.

A Appendix

A.1 Definitions of 11 anomalies

1. Financial distress: higher failure probability is related to lower returns (Campbell, Hilscher, and Szilagyi, 2008).

2. O-Score bankruptcy probability: bankruptcy probability based on accounting variables predicts lower returns (Ohlson, 1980).

3. Net stock issues: equity issuers underperform non-issuers with matching characteristics (Ritter (1991), Loughran and Ritter (1995), Fama and French (2008)).

4. Composite equity issues: equity issuers underperform non-issuers with matching characteristics. Equity issuance measured using a composite metric developed by Daniel and Titman (2006).

5. Total accruals: firms with high accruals underperform firms with low accruals (Sloan, 1996).

6. Net operating assets: firms with high net operating assets are predicted to earn lower long-run stock returns (Hirshleifer, Hou, Teoh, and Zhang, 2004).

7. Momentum: high (low) past recent returns predict high (low) future returns (Jegadeesh and Titman, 1993). Note: this measure predicts underpricing instead of overpricing.

8. Gross profitability: firms with higher gross profits yield higher returns (Novy-Marx, 2013). Note: this measure predicts underpricing instead of overpricing.

9. Asset growth: growth of total assets is negatively related to returns (Cooper, Gulen, and Schill, 2008).

10. Return on assets: more profitable firms have higher expected returns (Fama and French (2006), Chen, Novy-Marx, et al. (2011)). Note: this measure predicts underpricing instead of overpricing.

11. Investment-to-assets: higher past investment predicts lower future returns (Titman, Wei, and Xie (2004), Xing (2007)).

A.2 Formulas for construction of 11 anomaly metrics

1. Financial distress: See equations below. NIMTA is net income divided by firm scale (sum of total liabilities and market equity capitalisation). EXRET is the stock's monthly log return minus the country-specific market's log return. TLMTA is total liabilities divided by firm scale. SIGMA is the stock's annualised daily standard deviation for the most recent three months. RSIZE is the log of the ratio of the stock's market capitalisation to that of its country-specific market. CASHMTA is cash and short-term investment divided by firm scale. MB is the market-to-book ratio. PRICE is the log of the share price.

   The CRSP-based variables (EXRETAVG, SIGMA, RSIZE and PRICE) are for month t-1 and and income statement variables (NIQ) and balance sheet variables (LTQ, CHEQ, MB) are for month t-3 (previous quarter).

$$\begin{aligned}\pi = &-20.26NIMTAAVG + 1.42TLMTA \\ &- 7.13EXRETAVG + 1.41SIGMA - 0.045RSIZE \\ &- 2.13CASHMTA + 0.075MB - 0.058PRICE - 9.16\end{aligned} \tag{8}$$

   with

$$\begin{aligned}NIMTAAVG_{t-1,t-12} = \frac{1-\phi^3}{1-\phi^{12}}(&NIMTA_{t-1,t-3} + \cdots \\ &+ \phi^9 NIMTA_{t-10,t-12})\end{aligned} \tag{9}$$

$$EXRETAVG_{t-1,t-12} = \frac{1-\phi}{1-\phi^{12}}(EXRET_{t-1} + \cdots + \phi^{11} EXRET_{t-12}) \tag{10}$$

2. O-Score bankruptcy probability: See equations below. SIZE is the log of total assets. TLTA is the book value of debt divided by total assets. WCTA is working capital divided by total assets. CLCA equals current liabilities divided by current assets. ONEEG is a dummy variable; 1 if total liabilities ¿ total assets and 0 otherwise. NITA is net income divided by total assets. FUTL is funds provided by operations divided by total liabilities. INTWO is a dummy variable; 1 if net income is negative for the last 2 years and 0 otherwise. NI equals net income for year j, which is the most recent reporting year that ends at least four months before the end of t-1.

$$\begin{aligned}O = &-0.407SIZE + 6.03TLTA - 1.43WCTA \\ &+ 0.076CLCA - 1.72ONEEG - 2.37NITA \\ &- 1.83FUTL + 0.285INTWO - 0.521CHIN - 1.32\end{aligned} \tag{11}$$

with

$$CHIN = \frac{(NI_j - NI_{j-1})}{|NI_j| - |NI_{j-1}|} \tag{12}$$

3. Net stock issues: annual log change in split-adjusted shares outstanding. The most recent reporting year used is the one that ends at least four months before the end of month t-1.

4. Composite equity issues: (12-month) growth in the firm's total market value of equity minus the stock's (12-month cumulative) rate of return. This measure is lagged four months, to align its timing with that of the above measure of net stock issues.

5. Total accruals: annual change in non-cash working capital (= change in current assets minus the change in cash and short-term investment, minus the change in current liabilities, plus the change in debt included in current liabilities, plus the change in income taxes payable) minus depreciation and amortisation expense, divided by average total assets for the previous two fiscal years. The most recent reporting year used is the one that ends at least four months before the end of month t-1.

6. Net operating assets: total operating assets (total assets minus cash and short-term investment) minus operating liabilities (total assets minus debt included in current liabilities, minus long-term debt, minus common equity, minus minority interests, minus preferred stocks), divided by lagged total assets. The most recent reporting year used is the one that ends at least four months before the end of month t-1.

7. Momentum: cumulative returns from month t-12 to t-2.

8. Gross profitability: total revenue minus the cost of goods sold, divided by current total assets. The most recent reporting year used is the one that ends at least four months before the end of month t-1.

9. Asset growth: most recent year-over-year annual growth rate of total assets. The most recent reporting year used is the one that ends at least four months before the end of month t-1.

10. Return on assets: income before extraordinary items divided by the previous quarter's total assets. Income is for the most recent quarter for which the reporting date provided by Compustat precedes the end of month t-1. Assets are lagged one quarter.

11. Investment-to-assets: changes in gross property, plant and equipment plus changes in inventory, divided by lagged total assets. The most recent reporting year used is the one that ends at least four months before the end of month t-1.

A.3 Definitions of institutional variables (La Porta et al., 1998)

1. Anti-director rights: aggregated index of shareholder rights that are labeled as "anti-director rights." The index is formed by adding 1 when (1) the country allows shareholders to mail their proxy vote to the firm, (2) shareholders are not required to deposit their shares prior to the general shareholders' meeting, (3) cumulative voting or proportional representation of minorities in the board of directors is allowed, (4) an oppressed minorities mechanism is in place, (5) the minimum percentage of share capital that entitles a shareholder to call for an extraordinary shareholders' meeting is less than or equal to 10 percent, or (6) shareholders have preemptive rights that can be waived only by a shareholders' vote. The index ranges from zero to six.

2. Efficiency of the judicial system: assessment of the efficiency and integrity of the legal environment as it affects business (particularly foreign firms). Data provided by the country risk rating agency Business International Corp. The score represents investors' assessments of conditions in the country in question. Average between 1980 and 1983. Scale from zero to 10. Lower scores indicate lower efficiency levels.

3. Rule of law: assessment of the country law and order tradition, in particular the quality of contract enforcement, property rights and the courts. Data provided by the country risk rating agency International Country Risk. Average of the months of April and October of the monthly index between 1982 and 1995. Scale from zero to 10, with lower scores indicating less tradition for law and order.

4. Accounting standards: index created by examining and rating companies' 1990 annual reports on their inclusion/omission of 90 items. These items are general information, income statements, balance sheets, funds flow statement, accounting standards, stock data, and special items. The companies represent a cross section of industrial companies (70 percent) and financial companies (30 percent).

5. One share–one vote: equals one if the country legally requires ordinary shares to carry one vote per share (zero otherwise). This variable also equals one when the law prohibits the existence of both multiple-voting and nonvoting ordinary shares and does not allow firms to set a maximum number of votes per shareholder irrespective of the number of shares owned (zero otherwise).

6. Mandatory dividend: percentage of net income that the country legally requires firms to distribute as dividends among ordinary stockholders.

7. Corruption perceptions index: ranks countries based on the extent to which corruption is perceived to exist among public officials and politicians. Obtained from Transparency International. Lower scores indicate higher corruption.

A.4 Heteroskedasticity check: variance of error terms

Figure 2: Scatter plots of error terms across fitted regression values. The plots represent model 5 and model 6 for equal-weighted and value-weighted portfolios. A line of best fit is drawn through each of the scatter plots. Based on the Breusch-Pagan test, none of the residual terms are considered heteroskedastic.