The Cross Section of Country Equity Returns: A Review of Empirical Literature

Adam Zaremba 1,2

1 Department of Investment and Capital Markets, Institute of Finance, Poznan University of Economics and Business, al. Niepodległości 10, 61-875 Poznań, Poland; adam.zaremba@ue.poznan.pl
2 Dubai Business School, University of Dubai, Academic City, P.O. Box 14143 Dubai, UAE

Received: 27 September 2019; Accepted: 23 October 2019; Published: 28 October 2019

Abstract: The last three decades brought mounting evidence regarding the cross-sectional predictability of country equity returns. The studies not only documented country-level counterparts of well-established stock-level anomalies, such as size, value, or momentum, but also demonstrated some unique return-predicting signals such as fund flows or political regimes. Nonetheless, the different studies vary remarkably in terms of their dataset and methods employed. This study aims to provide a comprehensive review of the current literature on the cross-section of country equity returns. We focus on three particular aspects of the asset pricing literature. First, we study the choice of dataset and sample preparation methods. Second, we survey different aspects of the methodological approaches. Last but not least, we review the country-level equity anomalies discovered so far. The discussed cross-sectional return patterns not only provide new insights into international asset pricing but can also be potentially translated into effective country allocation strategies.

Keywords: cross section of country equity returns; country-level stock market anomalies; empirical asset pricing; international equity markets; return predictability

JEL Classification: G12; G14; G15

1. Introduction

The last three decades brought an unprecedented growth of exchange traded funds (ETFs) and index funds, which enable investors to quickly move their capital around the world. Currently, more easily than ever before, investors can relocate their equity allocation from Germany to Brazil or from Japan to South Africa. Not surprisingly, the ETF industry has been growing very rapidly. Already in 2017, the assets under management of ETFs exceeded five trillion U.S. dollars, and the compound annual growth rate over the past four years amounted to almost 19% (Lord 2018). The growth of ETFs coincides with a structural change in asset management and a shift from active investing to passive investing. As of December 2017, passive funds accounted for 45% of the aggregate assets under management in U.S. equity funds, compared to less than 5% in 1995 (Anadu et al. 2018). This profound revolution requires a whole new set of tools for equity investors, who now focus much less on which stocks to choose than on which countries to allocate money in.

The asset pricing literature produced a preponderance of trading signals, which help to predict the cross-section of individual stock returns. Recent surveys documented literally hundreds of different equity anomalies (e.g., Harvey et al. 2016; Hou et al. 2018). Notably, many of these cross-sectional patterns, such as value, momentum, or seasonality, have their parallels at the inter-market level and could be potentially used for country allocation. The last 30 years of asset pricing research produced mounting evidence regarding the cross-sectional predictability of country equity returns. The studies documenting numerous country-level equity anomalies not only provide new insights...
into international asset pricing but can also be translated into efficient country allocation strategies. Moreover, they are invaluable to practical investors.

The studies of the cross section of country equity returns not only examined different return patterns but also employed different methodologies and data sources. Issues such as choice of the index provider, return computation methodology, or portfolio formation can visibly influence the results. The diversity of empirical design and data sources and preparation methods calls for systematic review and for introducing a structure into the methodological choices in the field of country-level asset pricing.

The major objective of this article is to provide a comprehensive review of the current state of literature on the cross section of country equity returns. In particular, our survey considers data sources and preparation, research methods, and, last but not least, the cross-sectional return patterns documented in the country-level equity returns. The cross-section of stock-level returns is summarized in many excellent surveys, concerning both the anomalies themselves (e.g., Nagel 2013; Harvey et al. 2016; Hou et al. 2018; Bali et al. 2016), as well as methodological and data choices (Jagannathan et al. 2010; Waszczuk 2014a, 2014b). For country-level cross-sectional asset pricing, such surveys are clearly missing. To the best of our knowledge, any such review has not been yet presented. This work aims to fill this gap. We not only review, but we also structure and introduce some order into the current state of country-level asset pricing literature.

The article reviews three aspects of the studies of cross-section of country equity returns. First, we focus on the choice of data and the underlying asset universe as well as on dataset preparation. At the same time, we review the approaches regarding the country coverage, study period, return measurement, currency unit, and asset universe. Second, we survey some common methodological choices in the asset pricing literature, such as the number of portfolios, return calculation, and portfolio weighting scheme. Finally, we examine the current state of knowledge on country-level cross-sectional return patterns. We review the most prominent of such patterns, such as momentum, value, long-run reversal, size, seasonality, and price and non-price risk, as well as a basket of minor anomalies. We also discuss several additional aspects of these return patterns, including their fundamental sources and implementation details. Finally, we also consider additional practical aspects of country-level return patterns: the role of trading costs and strategy timing.

The remainder of the article proceeds as follows. Section 2 focuses on datasets, data preparation, and asset universe. Section 3 focuses on some specific methodological choices. Section 4 reviews the documented cross-sectional patterns in country equity returns. Finally, Section 5 concludes the article.

2. Datasets and Sample Preparation

This section concentrates on the choice of dataset representing country equity returns and preparation of the sample. We survey the approaches to selection of country coverage, study period, return measurement period, currency unit, and asset universe.

2.1. Country Coverage

The datasets used in examinations of the cross-sectional patterns in country index returns are obviously smaller than in the stock-level studies, which often encompass several thousand companies. Naturally, the scope in this case is limited to the countries with operating stock markets. The early studies usually focused on less than 20 developed markets. For example, Keppler (1991a), Ferson and Harvey (1994a), and Richards (1995) considered 18 developed markets. Modern studies usually concentrate on about 40 countries selected on the basis of classification into developed and emerging by one of the major index providers. For instance, Clare et al. (2016) investigate 40 markets, and Fisher et al. (2017) examine 37. The broadest studies take into account also less tradeable frontier markets, and their sample size can exceed 70. The article by Avramov et al. (2012), investigating 75 equity markets, may serve as an example of such an approach. Perhaps one of the broadest studies...
was conducted by Suleman et al. (2017), who took into consideration 83 countries. The detailed outline of the sample size in selected studies is presented in Table 1.

### Table 1. Research samples in the studies of the country equity returns.

<table>
<thead>
<tr>
<th>Article</th>
<th>Number of Countries in the Sample</th>
<th>Sample Period</th>
<th>Asset Universe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asness et al. (2013)</td>
<td>18</td>
<td>1978–2011</td>
<td>MSCI indices</td>
</tr>
<tr>
<td>Avramov et al. (2012)</td>
<td>75</td>
<td>1989–2009</td>
<td>MSCI and Datastream indices</td>
</tr>
<tr>
<td>Baltussen et al. (2019a)</td>
<td>14</td>
<td>1799–2016</td>
<td>Local country indices</td>
</tr>
<tr>
<td>Cenedese et al. (2016)</td>
<td>42</td>
<td>1983–2011</td>
<td>MSCI indices</td>
</tr>
<tr>
<td>Clare et al. (2016)</td>
<td>40</td>
<td>1993–2011</td>
<td>MSCI indices</td>
</tr>
<tr>
<td>Daniel and Moskowitz (2016)</td>
<td>18</td>
<td>1978–2013</td>
<td>Index futures</td>
</tr>
<tr>
<td>Ferson and Harvey (1994a)</td>
<td>18</td>
<td>1970–1989</td>
<td>MSCI indices</td>
</tr>
<tr>
<td>Fisher et al. (2017)</td>
<td>37</td>
<td>1990–2015</td>
<td>MSCI indices</td>
</tr>
<tr>
<td>Hurst et al. (2017)</td>
<td>11</td>
<td>1880–2016</td>
<td>Index futures</td>
</tr>
<tr>
<td>Keloharju et al. (2019)</td>
<td>15</td>
<td>1987–2016</td>
<td>MSCI indices</td>
</tr>
<tr>
<td>Moskowitz et al. (2012)</td>
<td>9</td>
<td>1965–2009</td>
<td>Index futures</td>
</tr>
</tbody>
</table>

Note. The table summarizes the data samples used in selected studies of the cross-section of country equity returns, indicating the number of countries covered, the length of the study period, and the asset universe.
In general, the larger sample size increases the power of statistical tests and allows additional insights on the examined return pattern. However, some studies may deliberately limit the sample of considered countries. One reason for that may be the focus put on some particular geographical region. For example, Grobys (2016) concentrates solely on the European Monetary Union. Another motivation to limit the number of countries in the sample may be alignment of the study with research practice. Since very liquid futures or ETFs cover only a small number of countries, the examinations may be reduced to just 10–20 of the most tradable markets. For example, the highly influential studies of Asness et al. (2013) or Keloharju et al. (2016) examined the samples of only 18 and 16 equity indices, respectively. Furthermore, the studies utilizing early security data tend to be sometimes quite narrow due to data unavailability. For example, Hurst et al. (2017), who investigated more than a century of evidence of trend following profits, constrained their scope to only 11 countries.

2.2. Study Period

The study period is usually dictated by the index data availability. Thus, numerous studies focusing on the most prominent cross-sectional patterns start in the years 1969 or 1970, when the coverage of many developed markets by MSCI begins (e.g., Balvers and Wu 2006; Bhojraj and Swaminathan 2006; Muller and Ward 2010; ap Gwilym et al. 2010). Consequently, the research period encompasses usually three to four decades. If the study period is shorter, this is usually due to inability to collect some sort of additional data for the 1970s or 1980s. For example, Berkman and Yang (2019), who focus on country-level analysts’ recommendations, reduced their study period to the years 1994–2015. Finally, as an alternative to equity indices, some studies proxy the equity markets with respective ETFs. In such cases, the price availability is, naturally, shorter. Smith and Pantilei (2015), who test the “Dogs of the World” strategy in ETFs, examine their returns for the years 1997–2012.

A separate and rapidly growing field encompasses studies of early security data that allow insights into the long-run nature of the financial market phenomena. In asset pricing studies in particular, examinations of the close-to-century long datasets make it possible to check the true robustness of the return patterns and secure against the risk of false discoveries and data mining. Some data providers, like Global Financial Data, offer their own proprietary indices going back to the 19th or even 18th century. A representative study of this type could be Geczy and Samonov (2017), who examine the momentum effect in the returns on major asset classes for the years 1800–2014. Baltussen et al. (2019a) research several major anomalies for the years 1799–2016. Other studies researching similar long-run data sets include Imanen et al. (2019), Hurst et al. (2017), or Spierdijk et al. (2012). For a detailed outline of different lengths of the study period, see Table 1.

2.3. Return Measurement Periods

The most common choice in individual stock studies is to use monthly returns. The motivation is that this choice forms a consensus that allows the large number of observations necessary for statistical tests to be accumulated and, at the same time, mitigates the influence of microstructure effects (Waszczuk 2014a). Nearly all of country-level studies take a similar approach and utilize monthly returns (e.g., Richards 1997; Chan et al. 2000; Blitz and van Vliet 2008). This refers, in particular, to the studies of early security data (e.g., Geczy and Samonov 2017; Baltussen et al. 2019a), where the more frequent observations are hardly available.

The use of different return intervals is rather infrequent and usually limited to examinations of alternative holding periods, as in Andreu et al. (2013) or Kasa (1992). On the other hand, Vu (2012) is one of the very few studies that relies on weekly returns to amass a bigger quantity of observations.

2.4. Currency Unit

Asset pricing studies of firm-level data frequently focus on single countries (e.g., Fama and French 2015) or replicate analyses in multiple individual markets (e.g., Chui et al. 2010). Therefore, the role of currency unit is of lesser importance and the calculations oftentimes rely on local currencies.
On the other hand, in the cross-country analysis the volatile foreign exchange rates and inflation rates—especially in emerging and frontier markets—play a significant role. Consequently, the majority of cross-country studies set a common currency as a unit of calculations, and the most obvious and common choice is the U.S. dollar. Dobrynskaya (2015), Clare et al. (2016), Keppler and Encinosa (2011), or Smith and Pantilei (2015) may serve as examples of papers that denominate all the prices in U.S. dollars. This currency is also a default choice in the studies that utilize futures or ETFs as representation of the country exposure, as it directly expresses the perspective of a U.S. investor (e.g., Andreu et al. 2013; Daniel and Moskowitz 2016; Smith and Pantilei 2015).

The use of returns calculated on the basis of local currency prices is rather rare. Usually such a framework has a sort of robustness check applied or is examined explicitly to evaluate the role of currencies in return predictability (Chan et al. 2000; Bhojraj and Swaminathan 2006). For example, Jordan et al. (2015) examine empirically the importance of the currency numeraire for the stock return predictability. They argue that, for instance, the presence (absence) of predictability for an American investor does not need to imply the existence (absence) of predictability for other international investors. Sometimes the local currency returns are also used in the studies of early security data to alleviate the problem of reliability of more than a century-old foreign exchange rate (Geczy and Samonov 2017). Nonetheless, even for the early asset prices studies the U.S. dollar is the very common choice (Baltussen et al. 2019a).

2.5. Asset Universe

Examination of cross-sectional patterns in country equity markets requires some representation of the market return. In country-level studies the asset universe comprises usually one of two types of instruments: either equity indices or some real investable instruments.

The major benefit of equity indices is that they provide a broad and accurate representation of the local equity markets. The articles investigating samples of international stock market indices basically follow one option. Most commonly, the studies are based on indices from a single provider. Alternatively, a study can rely on an amalgamation of local indices computed by national stock exchanges or local companies.

The use of indices from a single provider certainly has some benefits. They include calculation transparency, result comparability, and consistency in index calculation across many countries. Indices provided by MSCI are the most popular choice in country-level asset pricing. MSCI indices represent value-weighted equity portfolios covering approximately 85% of the largest and most liquid companies in each country. They also form the basis for multiple investment products, including popular iShares ETFs. MSCI estimated that more than 7 trillion US dollars were benchmarked to MSCI indices as of June 2011 (Cenedese et al. 2016). Furthermore, importantly from a practitioner’s perspective, MSCI usually does not apply any retroactive changes to the reported returns of its indices, so it reduces the risk of potential biases.

The current coverage encompasses 85 countries, including developed, emerging, frontier, and so-called standalone markets. The data period dates back to December 1969. The additional benefit of the MSCI indices is that they are calculated in several different ways, including different currencies, controlling for taxes, accounting for dividends, etc. For example, the MSCI indices were used by Dobrynskaya (2015), Clare et al. (2016), Fisher et al. (2017), Keppler and Encinosa (2011), Richards (1997), Balvers and Wu (2006), Keimling (2016), Keloharju et al. (2016), Malin and Bornholt (2013), Ferson and Harvey (1994b), and many others.

Datastream Global Equity Indices are the second most popular index choice. These cover currently 64 countries and go back in time to January 1973. Notably, the Datastream indices also assure a broad and consistent international representation, and at certain periods in the past their coverage may be better than in the case of MSCI. This index provider was selected, for example, by Bali and Cakici (2010), Umutlu (2015, 2019), and Zaremba (2019).
The studies of more than a century long dataset usually take advantage of indices computed by Global Financial Data (GFD). This provides time-series going back to the 19th century for numerous developed and emerging markets. Obviously, such long-run datasets are not free from different biases or omissions, and the index portfolios frequently contain very few securities, but certainly they provide a unique look into the past data. The GFD indices were employed by Geczy and Samonov (2017) and Baltussen et al. (2019a), among others.

One of the drawbacks of the indices obtained from different providers is that their coverage may differ; some countries may be taken into account by one provider but not considered by others. Consequently, to maximize the size of the research sample, some studies merge indices from different sources. Erb et al. (1995), in one of the first studies of this type, represent the developed markets by the MSCI indices and the emerging ones by the portfolios calculated by the International Finance Corporation (IFC). Avramov et al. (2012) use MSCI indices and supplement the coverage of missing countries with Datastream portfolios. Geczy and Samonov (2017) blend Bloomberg and GFD indices. Finally, Baltussen et al. (2019b) collect data from Bloomberg, with gaps filled in by Datastream data, spliced with index-level data, as in Baltussen et al. (2019b), and, eventually, backfilled data downloaded from Global Financial Data.

Besides using the indices from acknowledged providers, there are also several other options. Ellahie et al. (2019) use aggregated stock-level data from CRSP and Compustat. In other words, they calculate the country portfolios themselves instead of obtaining them from external sources. The final variant is to use national indices computed by local providers such as DAX, NIKKEI, or S&P. This approach is employed by Chan et al. (2000) and Vu (2012), among others. This approach has two major benefits. First, it may help to increase the dataset, because these local indices may have a longer history available than their counterparts offered by MSCI or Datastream. Second, the most liquid equity index futures are oftentimes linked with local indices rather than with international ones. For instance, in Poland the most liquid equity index future is based on the WIG20 index computed by the Warsaw Stock Exchange. Consequently, the use of local indices may be more aligned with investment practice. Nevertheless, on the other hand, the major shortcoming in relying on the local indices is the lack of computational consistency. Different indexes rely on different selection and weighting methods, so the study outcomes may be potentially influenced by the index calculation methodology, resulting in misleading conclusions. For example, better performance of an index in a certain market may stem from its bigger exposure to small-cap companies rather than from a true factor examined by the researcher.

Instead of investigating “paper” equity indices, some studies focus on actual investment instruments providing exposure to international markets. While this framework may potentially limit the size of the dataset, certainly it reflects most closely the investor’s practical perspective. Following this reasoning, Daniel and Moskowitz (2016), Moskowitz et al. (2012), and Hurst et al. (2017) base their computations on futures markets. Alternatively, Andreu et al. (2013), Breloer et al. (2014), and Smith and Pantilei (2015) focus on single-country ETFs.

For more examples of different asset universes used by the studies of the cross-section of country equity returns, see Table 1.

3. Methodological Choices

The country-level asset pricing studies strongly rely on econometric and statistical toolsets very similar to those used in the regular studies applied to the individual firms. The two most common approaches are cross-sectional (or panel) regressions and portfolio sorts. These two complementary approaches are frequently used jointly, as recommended by Fama (2015), and their benefits and shortcomings are discussed in detail by Fama and French (2008).

In the most typical applications of the cross-sectional regressions following Fama and MacBeth (1973), the future returns are regressed against a number of return-predicting variables, i.e., characteristics. Cross-sectional regressions are used, for instance, by Bali and Cakici (2010), Fisher et al. (2017), and
Sometimes this approach is supplemented with different types of panel regressions, as in Hjalmarsson (2010), Lawrenz and Zorn (2017), and Bali and Cakici (2010). Wisniewski and Jackson (2018) apply pooled ordinary least squares and two-way fixed-effects regressions.

Portfolio sorts are the second most popular tool. In this framework all the considered assets—which are in this case country equity markets—are ranked based on certain empirical characteristics, such as past returns or valuation ratios. Subsequently, they are grouped into subsets, and portfolios are formed. Finally, the performance of the cross-sectional portfolios is evaluated on the basis of mean returns, volatilities, Sharpe ratios, and with factor pricing models and monotonicity checks in the style of Patton and Timmermann (2010). The portfolio sorts reduce the cross-sectional dimension of the joint distribution of returns and also help to reduce the impact of measurement error (Waszczuk 2014a).

In the evaluation of the portfolios from one-way sorts, called also single-sorts, there is also a common practice to calculate the returns on a differential portfolio (or spread portfolio, long-short portfolio, zero-investment portfolio), which takes long and short positions in the two most extreme quantiles of assets from one-way sorts. The performance of such portfolios is then subsequently evaluated. Importantly, it should be noted that frequently this exercise serves as a quick check of monotonicity rather than a reflection of actual investment performance. Due to tradability and short-sale limitations, forming and rebalancing zero-investment portfolios across many countries is not always possible, unless they are made of liquid futures, as in Daniel and Moskowitz (2016) or Moskowitz et al. (2012). Some further discussion of the details of sorting methods in asset pricing studies is provided in Bali et al. (2016), Vaihekoski (2004), Van Dijk (2011), and Waszczuk (2014a).

The outcomes of the cross-sectional analysis based on portfolio sorts is sensitive to several methodological choices made by the researcher. Importantly, some of the country-level practices may differ from stock-level studies due to different number of assets, data availability, liquidity considerations, etc. I will focus then on several of the most important methodological choices.

3.1. Number of Portfolios

The studies of the cross-section of returns on common stocks rely on datasets of hundreds or thousands of companies. Therefore, decile (e.g., Jegadeesh and Titman 1993, 2001; Lakonishok et al. 1994) or quintile (Banz 1981; Chan et al. 1998) groupings belong to the most common choices. At the country level the number of assets is more limited, so this type of study requires also a smaller number of portfolios. Otherwise, the grouping could result in portfolios containing only a few—or even one—markets, hence being susceptible to the noise in returns. The most popular choices include tertiles (e.g., Daniel and Moskowitz 2016; Geczy and Samonov 2017; Asness et al. 2013; Atilgan et al. 2019), quartiles (Richards 1997; Blitz and van Vliet 2008; Macedo 1995a; Malin and Bornholt 2013; Erb et al. 1995), or quintiles (Clare et al. 2016). Alternatively, some studies which assume different portfolio formation methodologies, consider only two portfolios—long and short (e.g., Moskowitz et al. 2012). Bali and Cakici (2010) consider portfolio groupings including 30%, 40%, and 30% of the markets.

Finally, a number of studies, instead of assuming a certain quantile cut-off point, focus only on the extreme portfolios from single-sorts and assume a fixed number of countries included. For instance, Kortas et al. (2005) include 11 of the most extreme countries in each portfolio. On the other hand, Keloharju et al. (2016) test cross-sectional seasonality based on portfolios including the three equity indices with the highest or lowest average return in the past.

3.2. Portfolio Weighting Scheme

Once the portfolios are formed, the next important step is the selection of the weighting scheme. The most common choice is between the value-weighted portfolios and equal-weighted portfolios. In the first framework, the returns are weighted according to the market capitalization. On the other hand, in the equal-weighted approach, all the returns are assigned an equal-dollar value. At the stock level, the value-weighting approach is markedly more popular, and there are several reasons for that. The equal-weighted portfolios may tend to assume very large positions in small and micro companies,
which would be unrealistic in practice, due to liquidity or market capacity issues, for example. In addition, the equal-weighted portfolios have a built-in rebalancing assumption, which may distort the results (Willenbrock 2011). Finally, value-weighting deemphasizes observations that are more likely to suffer from the data errors, thus reducing the variation in average returns. Nevertheless, at the country level the choice is not that obvious. Indeed, the equal-weighted portfolios may gravitate towards small and illiquid frontier markets, where any large exposure of frequent share purchases may be unrealistic. However, on the other hand, in the case of a limited sample size of just 30–40 countries, the value-weighted portfolios may be strongly dominated by only a few of the largest countries. Furthermore, the aggregate market value may not always be available, or it may not have any intuitive equivalent, as is the case with the futures or ETFs. Consequently, the equal-weighted portfolios are much more common, or at least used along with the value-weighted portfolios. The equal-weighted portfolios are used, for example, by Geczy and Samonov (2017), Clare et al. (2016), Hurst et al. (2017), and Balvers and Wu (2006). The value-weighted strategies, on the other hand, are analyzed by Chan et al. (2000) and Rikala (2017).

Besides the classical value- or equal-weighted portfolios, some articles pursue alternative frameworks. Clare et al. (2016) and Moskowitz et al. (2012) use so-called risk-parity, i.e., they weight the portfolio components on their inverse volatility. On the other hand, Ilmanen et al. (2019) and Asness et al. (2013) link the weight with the value or rank of the underlying characteristic so that the absolute weight increases when the sorting variables take more extreme values.

3.3. Return Calculation: The Treatment of Dividends and Taxes

The index-level return calculations face two major methodological choices. The first issue refers to the treatment of dividends. Most of the studies are based on total return indices, which include reinvested dividends, regardless of the particular index provider (e.g., Richards 1997; Balvers and Wu 2006; Bali and Cakici 2010). Accounting for dividends reflects the investor’s perspective well; nonetheless, sometimes the coverage and the length of the time-series may be bigger for the price returns. Therefore, price indices, which do not account for dividends, are employed by Keppler (1991a), for instance. On the other hand, ap Gwilym et al. (2010) and Geczy and Samonov (2017) use both price and return indices. Finally, some examinations use the two types of measures in combination as different inputs. For example, Clare et al. (2016) measure portfolio performance with the total return indices but compute return predictive signals based on price indices.

The total returns indices include dividends, which are taxed in various ways in the majority of countries. Importantly, the dividend tax rates may vary both across time and countries, affecting the net portfolio performance. Some groups of investors, like mutual funds, may be exempted from taxation on dividends in many countries. Nonetheless, this is not true for all the countries, at all times, and for all the groups of investors. Consequently, the taxes may still potentially affect the cross-section of country equity returns. The majority of the country-level asset pricing studies use gross returns, not accounting for taxation. On the other hand, Zaremba (2016) also use MSCI Net Return indices, which account for dividend tax rates within the particular countries.

4. Cross-Sectional Patterns in Country-Level Returns

We now turn to the review of patterns demonstrated in the cross-section of country equity returns. We begin by focusing on the most prominent and best-established ones, such as momentum, size, and value and, subsequently, carry on with more minor return regularities. In addition, we consider different types of risk that influence future index-level returns. To introduce some order, we arbitrarily classify these risks into the ones that can be derived from prices (price based), and others, that is, non-price risks such as credit or political risks. Eventually, we survey the studies’ treatment of some additional aspects of the country-level anomalies, such as factor timing and the role of trading costs.
4.1. Momentum

The momentum effect, which is the tendency of assets with high (low) past returns to continue to overperform (underperform) in the future, is one the most robust and pervasive asset pricing anomalies ever documented. It has been demonstrated in U.S. and international stocks, including developed, emerging, and frontier, markets, commodities, bonds, currencies, and also in equity market indices.

Index-level evidence. The first empirical evidence for the momentum effect in country equity indices may be found in Ferson and Harvey (1994b), Macedo (1995a, 1995b), Richards (1997), and Asness et al. (1997). Other researchers have continued the examinations of country-level momentum in the subsequent years. Balvers and Wu (2006) investigate a Jegadeesh and Titman (1993)-style portfolio based on stock market indices from 18 developed equity markets within the years 1969–1999. They demonstrate strong momentum effects, which worked particularly well in combination with the mean-reversion patterns. In the same year, Bhojraj and Swaminathan (2006) published a paper which examined a broader sample of 38 country indices within the same period. The authors document that the quintile of the best performing countries over the previous 6 months continued to significantly outperform the laggard indices during the next three quarters. The mean return on the long/short portfolio within a year after its formation amounted to 7.65%.

The following years saw further examinations of the momentum effect that extended the study sample both in terms of number of countries and the length of the study period. Muller and Ward (2010) investigated 70 countries and Zaremba (2016) researched 74. In terms of the sample length, several studies extended the time-series back to the 19th century and researched approximately 200 years of returns (Geczy and Samonov 2017; Hurst et al. 2017; Baltussen et al. 2019a). The momentum effect is robust to many considerations and could be successfully implemented with the use of ETFs (Andreu et al. 2013). Angelidis and Tessaromatis (2018) argue that “country-based factor portfolios offer a viable alternative implementation of factor investing in a world of illiquidity, transaction costs, and capacity constraints.” Some other studies that investigated the momentum effect at the country level are Chan et al. (2000), Daniel and Moskowitz (2016), Grobys (2016), Guilmin (2015), Ilmanen et al. (2019), Breloer et al. (2014), Nijman et al. (2004), Clare et al. (2017), L’Her et al. (2004), Vu (2012), Kortas et al. (2005), and Shen et al. (2005).

Formation and holding periods. The seminal study of Jegadeesh and Titman (1993) considered 3–12-month-long sorting and holding periods. Numerous country-level studies, including the early ones, take a similar approach (e.g., Balvers and Wu 2006; Andreu et al. 2013). Later studies frequently used the approach advertised by Fama and French (1996), i.e., 1-month holding period and 12-month sorting period with the most recent month skipped (e.g., Dobrynskaya 2015; Blitz and van Vliet 2008; Asness et al. 2013). The 1-month skip period is usually applied in order to disentangle the short-term reversal effect discovered by Rosenberg et al. (1985), Jegadeesh (1990), and Lehmann (1990). Nonetheless, at the country level no similar one-month reversal effect has been documented, and Zaremba et al. (2019) argue that the returns display rather a short-term continuation. Consequently, the country-level studies do not always assume the one-month skip period, and if they do, this is usually motivated by liquidity and implementation issues (Asness et al. 2013; Baltussen et al. 2019b). For this reason, Geczy and Samonov (2017), who study early security data, decided to skip even two months in part of their tests.

Momentum improvements and alternative implementations. While the classical momentum assumes sorting the indices on raw past returns, a number of studies offer alternative, but closely related approaches. Notably, while some are conceptually very close to momentum, more detailed tests show that they provide incremental information about future returns. Moskowitz et al. (2012) and Hurst et al. (2017) evaluate so called “time-series momentum”. This strategy assumes including markets into long or short portfolio depending on whether the excess return in the sorting period was positive or negative. ap Gwilym et al. (2010), Clare et al. (2017), and Baltussen et al. (2019b) test trends following strategies that focus on whether the most recent index value is above or below its moving average. Bornholt and Malin (2010, 2011) research the 52-week high strategy, whereby the
return-predicting signal is the distance to the 52-week maximum index value. Avramov et al. (2018) concentrate on the distance between short- and long-run moving averages of prices. Finally, several studies demonstrate that the momentum effect could be efficiently combined with long-run reversal to augment the performance of the strategy (Balvers and Wu 2006; Asness et al. 2013; Bornholt and Malin 2014).

**Sources of the momentum effect.** The stock-level momentum studies highlight a number of different explanations of the momentum effect, such as risk premium, behavioral underreaction or overreaction, herding, or confirmation. For example, Bhojraj and Swaminathan (2006) highlight the potential overreaction to news about macroeconomic conditions. In addition, Cenedese et al. (2016) link the momentum effect with the tendency of investors to increase their holdings in markets that have recently outperformed (Froot et al. 1992; Bohn and Tesar 1996; Griffin et al. 2004; Chabot et al. 2014). Other studies offer some alternative explanations. Balvers and Wu (2006) link the momentum effect with production-based asset pricing concepts. From the risk-based perspective, Asness et al. (2013) argue that global funding liquidity risk is a partial source of the momentum pattern. Cooper et al. (2019) demonstrate that momentum returns are explained by the portfolio loadings on global macroeconomic risk factors. Eventually, Evans and Schmitz (2015) link the global momentum effect with data mining for anomalies, calling it a likely example of a selection bias.

### 4.2. Size Effect

The country-level size effect is a phenomenon parallel to the firm-level size effect discovered by Banz (1981). Kepper and Traub (1993) were the first to demonstrate that low-capitalization equity markets outperform large equity markets. The authors found that the smaller national equity markets in the MSCI Developed Markets universe produced an average annual return of 19.19% within the years 1975–1992. This outcome compared favorably with the 12.67% total compound return on the MSCI World Index. Furthermore, the small markets displayed lower downside characteristics. The outperformance of the small firms was later confirmed also by Asness et al. (1997) and by Kepper and Encinosa (2011). The size, measured with market capitalization, also belonged to the risk attributes examined by Harvey (2000).

The size effect was further demonstrated in several more recent studies. Fisher et al. (2017) show that stocks from small equity markets tend to have higher average returns than stocks from large countries. Notably, they accentuate that the country size effect is largely independent of the firm size effect and other country quantitative factors such as the momentum or value effects. Zaremba and Umutlu (2018) also demonstrate the size effect in large international sample, and Li and Pritamani (2015) show that it drives the returns on emerging and frontier markets. Similarly, Pungulescu (2014) points out that the market size effects account for up to 1% per year in terms of expected returns in emerging countries. Finally, Rikala (2017) focuses solely on European markets and finds no consistent evidence that small countries outperform large ones.

**Sources of the country size premium.** The firm-level size effect is frequently linked to additional risk factors, such as liquidity, information risk (see Norges Bank (2012) for a comprehensive review of the sources of small firm effect). While Fisher et al. (2017) provide evidence that the country-size effect is not simply a firm-size effect “in disguise” (the effect does not arise because smaller markets are populated by smaller firms), the potential explanations usually oscillate around the concept of risk. Rikala (2017) writes that “Intuitively, small countries producing higher returns is logical because of the widely acknowledged return profile of small stocks; investing in small firms produces higher returns in exchange for greater volatility and possibly even a return premium; a return in excess of the required compensation for additional risk.” Fisher et al. (2017) conjecture that the small-country effect is due to home bias, but they provide mixed evidence in support of this conjecture. They also demonstrate that the country size effect does not simply stem from lower analysts’ coverage. Zaremba (2016) shows that accounting for country-specific risks (sovereign, political, etc.) can largely explain the abnormal returns for small markets. Finally, Pungulescu (2014), similarly to Zaremba (2016), demonstrates
that the size effect is more pronounced in emerging countries than in developed countries, and the size premium exists independently of the segmentation premium documented in the literature. Finally, Zaremba and Umutlu (2018) provide evidence that the country size premium is strongly concentrated in January, as in the case of the firm size effect (Keim 1983; Lamoureux and Sanger 1989; Daniel and Titman 1997). Last but not least, a white paper by Evans and Schmitz (2015) argues that the cross-sectional pattern related to market capitalization may be simply a statistical artifact, which cannot be confirmed in the recent data.

4.3. Value Effect

The value effect refers to the tendency of stocks with low valuation ratios, such as the price-to-earnings ratio or price-to-book ratio, to outperform stocks with high valuation ratios. For individual stocks, this phenomenon has been well known for about six decades now (Nicholson 1960; Basu 1975, 1977, 1983; Reinganum 1981), but in the equity indices it has been documented only in the 1990s (Kepler 1991a, 1991b). In one of the earliest studies, Macedo (1995a, 1995b, 1995c) researched the performance of country portfolios based on 18 country equity indices. She forms quartile portfolios from sorts on three different indicators, the book-to-market ratio, dividend yield, and earnings yield, and tests their performance within an almost 20-year period. She concludes that the “cheap” countries outperformed the “expensive” markets, and the differential annual return between the countries with the lowest and highest valuation ratios ranged from 1.25% to 8.54%, depending on the ratio selection, rebalancing frequency, and hedging approach.

The valuation effect was also confirmed in more recent studies that use broader data samples and longer timespans. For example, Angelidis and Tessaromatis (2018) investigated the performance of 23 developed markets within the 1980–2014 period. They found that value portfolios vividly outperformed market portfolios, delivering information ratios ranging from 0.27 to 0.39, depending on the weighting scheme. Further evidence for the value effect across countries was provided by Faber (2012), Klement (2012), Angelini et al. (2012), Ellahie et al. (2019), Novotny and Gupta (2015), Keimling (2016), Kim (2012), Heckman et al. (1996), Ferson and Harvey (1994b, 1998), Kortas et al. (2005), Lawrenz and Zorn (2017), Ferreira and Santa-Clara (2011), Desrosiers et al. (2007), Zaremba and Szczygielski (2019), Asness et al. (1997), and, finally, L’Her et al. (2004). Furthermore, Baltussen et al. (2019b) included the value effect in their two-century study, confirming its pervasive and robust character. However, Kim (2012) and Zaremba (2016) show that the effect is stronger among the emerging markets rather than in developed countries.

Valuation ratios. The value effect in country equity indices can be examined with different valuation ratios. The majority of them are parallels of similar ratios or techniques used at the firm level. The most popular include price-to-earnings (P/E) ratio (e.g., Ellahie et al. 2019; Kim 2012; Keimling 2016), price-to-book (P/B) ratio (Ellahie et al. 2019; Angelidis and Tessaromatis 2018; Kortas et al. 2005), or dividend yield (Keimling 2016; Hjalmarsson 2010; Kepler 1991a). Some articles focus also on modified versions of these valuation ratios. For instance, Kortas et al. (2005) use forward P/E ratios and Lawrenz and Zorn (2017) concentrate on conditional price-to-fundamental ratios. The other utilized ratios encompass price-to-cash flow ratio (e.g., Kepler 1991a; Keimling 2016). Desrosiers et al. (2007) offer an alternative framework based on residual income. Zaremba and Szczygielski (2019) review several popular valuation ratios to conclude that the EBITDA-to-EV signal seems to be the most effective predictor of future cross-sectional returns. In addition, Ferreira and Santa-Clara (2011) show that several ratios can be combined to obtain superior performance.

Finally, there is one specific valuation ratio, which was designed purportedly for the country-level predictions: the cyclically adjusted price-to-earnings ratio, abbreviated CAPE. This technique could be traced back to the seminal work “Security Analysis” by Graham and Dodd (1940). The authors put forward an idea of smoothing earnings over the previous few years in order to calculate valuation ratios. Nevertheless, the true father of the application of CAPE to equity premium predictions is Robert Shiller, the Nobel laureate of 2013. In his 1988 study (Campbell and Shiller 1988, p. 675) he
demonstrated that “a long moving average of real earnings helps to forecast future real dividends.” Consequently, it might be also used to predict future returns. CAPE, called also Shiller P/E, is computed as an index value divided by the average of trailing 10-year earnings adjusted for inflation. Numerous studies demonstrate that CAPE could be also successfully applied to country selection. For example, Faber (2012) examines the role of CAPE in a sample of 30 countries’ equity markets for the years 1980–2011. Faber (2012) provides evidence that an equal-weighted quarter portfolio of the countries with the lowest CAPE produces a mean yearly return of 13.5%, whereas the most expensive markets deliver only 4.3% per year. At the same time, the equal-weighted portfolio of all of the countries in the sample returned 9.4% per year. Klement (2012) demonstrates that CAPE can predict returns even within a five- to ten-year horizon. The efficiency of CAPE as the predictor of future returns was later verified and confirmed also by Angelini et al. (2012), Novotny and Gupta (2015), Keimling (2016), and Ilmanen et al. (2019).

Sources of the value effect across countries. The common reasoning regarding the value effect is similar to the parallel effect at the firm level, linking it either to behavioral mispricing or to some risk factors not captured by the established asset pricing models. Nonetheless, the catalogue of risks may be slightly different due to differences in the nature of the asset class. Ellahie et al. (2019) find that low P/B countries face temporarily depressed current earnings, and their recovery in future earnings growth is uncertain. Moreover, the markets with low P/B also exhibit greater downside sensitivity to global earnings growth. Ferson and Harvey (1998) argue, for instance, that the P/B ratio has cross-sectional explanatory power at the global level, mainly because it contains information about global market risk exposures. Zaremba (2016) also shows that the country specific risk explains a large part of the country-level value premium.

4.4. Seasonality

Cross-sectional seasonality is a relatively new phenomenon described by Heston and Sadka (2008) and later confirmed by several other authors in international markets (Heston and Sadka 2010; Keloharju et al. 2016). What Heston and Sadka (2008) found is that the stocks with a high same-month average return in the past tend to outperform stocks with a low same-month return in the past. Notably, Keloharju et al. (2016) extend this evidence to country equity indices. They find that this seasonal return pattern is admittedly weaker than in other asset classes but still visible. The tertile of countries with the highest same-month return outperform the tertile of the markets with the lowest same-month return by 0.48% (t-stat = 2.20). Notably, the markets with the highest average return in the remaining months underperformed the markets with the lowest other-month return by −0.36% (−1.66). Consistent findings were also presented in a later paper by the same authors (Keloharju et al. 2019), but, again, the statistical significance was low. The phenomenon has also been verified in early data samples by Baltussen et al. (2019b).

4.5. Long-Run Reversal

The long-term reversal at the firm level dates back to the seminal study of De Bondt and Thaler (1985), who provided convincing evidence that stocks with a poor (good) performance over the previous 3–5 years tend to produce high (low) returns in the future. Further studies demonstrated that the effect is not only robust, but also pervasive, driving the returns on individual stocks globally (Baytas and Cakici 1999; Blackburn and Cakici 2017), futures (Lubnau and Todorova 2015), currencies (Chan 2013), and commodities (Bianchi et al. 2015; Chaves and Viswanathan 2016). Notably, the effect is also present in country equity indices.

1 Note that this article focuses only on cross-sectional seasonality. Apart from this, there is some evidence that the equity indices demonstrate some monthly calendar patterns in the time series, for example in Kepler and Xue (2003) or Bouman and Jacobsen (2002).
Index-level evidence. The first evidence of the long-term reversal effect was provided by Kasa (1992) and Richards (1995, 1997). These authors based their research usually on limited samples of developed markets and demonstrated that indices with low (high) long-term performance significantly outperform in the future. The results were later confirmed with larger and longer samples by Kortas et al. (2005), Balvers et al. (2000), Balvers and Wu (2006), and Shen et al. (2005). Gharaibeh (2015) corroborated the long-term reversal phenomenon in the Middle East market indices, and Spierdijk et al. (2012) confirmed the findings in study periods exceeding a century. The strategy works well for 36–60-month sorting periods, and Malin and Bornholt (2013), who develop so-called “late stage” contrarian strategies, experiment also with skipping the most recent 12 months as in Fama and French (1996). Finally, Smith and Pantilei (2015) develop a simple mean-reversion-based strategy, which they called “Dogs of the Word”. The technique assumes buying in five countries with the worst performance over the last year and holding them for five years. The strategy proves profitable both in indices and single-country ETFs. Smith and Pantilei (2015) argue that “assuming a five-year holding period, such a portfolio would have produced compounded annual returns of 10.39%”, exceeding the profits on the global passive equity portfolios. In the years 1997–2012 their strategy implemented with the ETFs of the worst-performing countries outperforms the MSCI All Country World Index (MSCI ACWI) by 246 bps, delivering a higher Sharpe ratio and net of ETF expenses.

Sources of long-run reversal. Although there is no consensus on the source of long-run reversals, the existing studies offered some potential explanations. Richards (1997) considers whether the contrarian profits may stem from risk-differentials but finds no support for this hypothesis. He argues that no evidence suggests that loser-index returns are riskier in terms of their volatility or exposure to the world equity market returns. Cooper et al. (2019) link some similar patterns to global macroeconomic risks.

The winner-loser reversals profits are larger among the smaller countries than in the larger markets, so there may be an element of a “small-country effect”, but still this phenomenon does not fully explain the long-term reversal effect (Zaremba and Umutlu 2018).

Another option is that the long-run reversal is just a statistical artifact and that its returns were purely period specific. Indeed, country-level long-term reversal tends to be very volatile and unstable over time, but its robustness over very long periods casts doubt on such an explanation (Spierdijk et al. 2012). Furthermore, Malin and Bornholt (2013), who employ longitudinal analysis, argue that the mean-reversion effect is present even in the post-1989 sample despite the absence of visible contrarian profits for the developed markets.

Further explanations point to behavioral mispricing that cannot be arbitraged away for many reasons, including cross-border flow limitations. The behavioral overreaction hypothesis is also consistent with the link to the momentum effect (Richards 1997; Balvers and Wu 2006; Malin and Bornholt 2013).

4.6. Price Risk

The relationship between the risk measures calculated on the basis of prices and future returns on stocks has been a controversial and intensively researched topic in recent years. On the one hand, early theoretical models suggest that systematic risk should positively correlate with future returns in the cross section, and some early studies seem to produce consistent evidence (Sharpe 1964; Black et al. 1972; Fama and MacBeth 1973; Blume 1970; Miller and Scholes 1972; Blume and Friend 1973). Similarly, the stock-specific risk should also be either positively correlated or unrelated, depending on market integration (Levy 1978; Tinic and West 1986; Merton 1987; Malkiel and Xu 1997, 2004). However, the empirical evidence mounting over the past two decades documents a contrary phenomenon—the so called low-risk anomaly. The high-risk firms tend to underperform the low-risk firms on the risk adjusted basis, both when the risk is understood as a systematic risk or an idiosyncratic risk (Frazzini and Pedersen 2014; Ang et al. 2006, 2009). The effect is usually explained with the combination of behavioral biases and limits to arbitrage (Blitz et al. 2019). Notably, some other measures of
price-based risk, such as value at risk, display a rather positive than negative relationship with future returns in the cross section (Bali and Cakici 2004).

**Market beta.** The risk-return relationship at the country level is also far from obvious and depends strongly on risk measures. The first studies bring weak evidence on the pricing on systematic risks, especially in emerging markets (Harvey 1991, 1995; Harvey and Zhou 1993). In one of the first studies, Harvey (1995) finds no relationship between beta and future returns across 20 emerging markets. In addition, more recent studies by Estrada (2000) and Bali and Cakici (2010) lead to similar conclusions. Nonetheless, the seminal study of Frazzini and Pedersen (2014) demonstrates that on a risk-adjusted basis, low-beta indices outperform high-beta indices, and the effect is confirmed by Berrada et al. (2015). Hedegaard (2018) also corroborates the low-beta effect in developed and emerging market indices, demonstrating additionally that it is partially predictable by past market returns.

**Idiosyncratic risk.** The country-level examinations display no evidence of the low-idiosyncratic risk anomaly, which is similar at the firm level. The majority of the studies find either a positive relationship or no significant relationship between idiosyncratic (or total) volatility and expected country returns in the cross section. Bali and Cakici (2010) compute total and idiosyncratic volatility measures of different asset pricing models based on estimation periods ranging from one to six months and find a positive relationship. On the other hand, articles by Umutlu (2015, 2019), Liang and Wei (2019), and Hueng and Ruey (2013) show either very weak or unreliable links between idiosyncratic or total volatility and future returns in the cross section. The pricing of similar measures of price risk has been also considered by Bekaert and Harvey (1995), Estrada (2000), and Hueng (2014).

**Other definitions of risk.** Several studies examined other definitions of price risk. Some of them documented significant relationships, while others were less successful. Hollstein et al. (2019) investigate the pricing of tail risk in international stock markets. They find that both local and our newly computed global tail risk strongly predict global equity index excess returns. Sorting equity market countries into portfolios by their tail risk generates sizable excess returns across various holding periods. Arouri et al. (2019) examine the role of jump risk. Umutlu and Bengitöz (2017) offer a similar metric based on return range. Finally, Atilgan et al. (2019) test the forecasting power of several measures of downside risk, i.e., downside beta, tail beta, value at risk, and expected shortfall, but find no consistent evidence of return predictability.

### 4.7. Non-Price Risks

Besides the measures of risk derived from price behavior discussed in the previous section, numerous studies explore the role of alternative definitions and source of risk. The logic behind these studies is the following: if the country-specific risk matters for country-level asset pricing, what actually is this country-specific risk? Can it be conceptualized and captured more precisely with some alternative measures?

Examinations of the country-specific risks as determinants of future market-level performance are found in the earliest studies of cross section of country returns and date back to the 1990s (Ferson and Harvey 1994a, 1994b; Erb et al. 1995, 1996a; Bekaert et al. 1996). Some of these studies focus on just one type of risk, such as credit risk or political risk, while others examine several categories or exposures to them (Ferson and Harvey 1994a; Erb et al. 1996a; Harvey 2004). The types of considered country specific risks could be categorized into several broad classes.

**Credit risk.** Country credit risk (sovereign risk, default risk) belongs among the best-established predictors of future returns. Not only has it been extensively documented by practitioners, it is also widely employed by practitioners in models of cost of equity. A widely used database in Damodaran (2019) advocates using country risk premia based on local sovereign ratings. Erb et al. (1995) employ measures of credit risk calculated on the basis of the Institutional Investor Semiannual Survey of Bankers and demonstrate that the credit risk is priced in the country equity premium. In a later study, the same authors show how the credit risk could be used to estimate risk premia for 135 different countries—even those without developed stock markets (Erb et al. 1996b). More recent research
confirms these early findings with different measures of credit risk. Avramov et al. (2012) use quantified credit ratings for 75 countries in the period 1989–2009. They show that the high credit risk tercile outperforms the stocks in the countries in the low credit risk tercile by 0.57% monthly. Zaremba (2016) further corroborates these findings by using the Economist Intelligence Unit sovereign risk indicator calculated by its Country Risk Service. Having examined 74 countries for the years 1999–2015, Zaremba arrives at a qualitatively similar return on a tertile differential portfolio of 0.50% per month.

**Political risk.** The political risk is another category of risk that has been examined since the beginning of studies of the cross section of country returns (Erb et al. 1996a; Diamonte et al. 1996). The political risk is most frequently measured with the Political Risk Index, which constitutes a component of the International Country Risk Guide calculated by the PRS Group.\(^2\) In general, the studies find that the political risk is positively related to the expected returns in the cross section. (Erb et al. 1996a; Dimic et al. 2015; Lehkonen and Heimonen 2015; Vortelinos and Saha 2016). Bilson et al. (2002) show that the political risk is more strongly priced in emerging markets rather than in developed ones. Consistently with this, Diamonte et al. (1996) concentrate on changes in political risk and demonstrate that average emerging market returns in countries experiencing declining political risk exceed those of emerging markets experiencing growing political risk by approximately 11% per quarter. In contrast, the analogous return for developed markets amounted to only 2.5%. Furthermore, Zaremba (2016) show that country-risk pricing is stronger in emerging and—in particular—frontier markets. Dimic et al. (2015) explore this difference further and show that while composite political risk is priced in all the types of stock markets (i.e., developed, emerging, and frontier), the role of individual components varies across countries. For example, government action is a common source of risk in all market categories, but the impact of government stability is unique to frontier equities.

Recent studies offer some further insights into the effect of political risk. Pagliardi et al. (2019) propose an international capital asset pricing model that accounts for the political risk. The model explains up to 77% of cross-sectional returns, outperforms some other benchmark models, and has a good predictive power. Gala et al. (2019) offer two new politics and policy risk factors and demonstrate that markets with lower politics and policy rankings produce higher average returns. They also offer some long-short strategies, which are argued to produce returns exceeding 12% per year with a corresponding Sharpe ratio of 0.59.

**Other non-price risks.** While credit risk and political risk seem to be the most intensively researched categories, other studies also consider alternate types of risks, such as economic and financial risks (Erb et al. 1996a), macroeconomic and political risks and uncertainty (Chang et al. 2017; Rapach et al. 2005), or expropriation risk (Dahlquist and Bansal 2002). Lee (2011) empirically tests the liquidity-adjusted asset pricing model of Acharya and Pedersen (2005) at the global level. The latter provide evidence that liquidity risk is priced internationally, independently of other risks. Additional analyses of country-level risk are also performed by Suleman et al. (2017).

### 4.8. Other Predictors

In this section, we review an array of less known predictors that have been discovered and examined in recent years.

**Fund flows.** Srimurthy et al. (2019) offer a new country asset allocation approach based on fund flows. The authors find reliable positive returns on a strategy that goes long in the countries that have attracted indirect investment via equity fund flows and short in the countries that have not. The effect is independent of some other well-established return predictors, such as size or momentum.

**Economic freedom.** Several studies explore the role of economic freedom for the future stock market returns. Stocker (2005) was, most probably, the first to try to examine this relationship. Having examined the returns on developed and emerging markets in the years 1975–2002, he demonstrates

---

that the rate of increase in economic freedom is directly related to equity returns. He also develops an investment strategy based on this phenomenon, which earns attractive investment returns. Similar evidence is provided by Smimou and Karabegovic (2010), who concentrate on MENA markets. Finally, Stocker (2016) corroborates his own earlier results. He documents that the index of economic freedom provides incremental information about future returns that have low correlation with value, momentum, and size factors. Stocker (2016) christens the abnormal returns from investing in low economic freedom countries “the price of freedom”.

**News.** Calomiris and Mamaysky (2019) develop a new classification methodology for using the content and context of news to forecast the performance of 51 equity markets. They consider issues such as topic-specific sentiment, frequency, and unusualness (entropy) of word flow. They demonstrate significant predictive abilities of the news flow for returns, volatilities, and drawdowns, particularly for longer (one-year) horizons. The effect is more pronounced in emerging markets.

**Analyst recommendations.** There are numerous studies of the predictive power of analysts’ recommendations for individual stock returns (Kothari et al. 2016), but Berkman and Yang (2019) are the first to consider a country-level parallel. The authors digest analysts’ reports from 30 countries for the years 1994–2015 to demonstrate that the aggregate recommendation score helps to predict international stock market returns. The country-level recommendations make it possible to predict future aggregate cash flow and returns. A country-allocation strategy based on the insights of Berkman and Yang (2019) yields an approximate abnormal return of 1% per month.

**Asset growth.** The role of asset growth for future returns on individual stocks has become well known since Cooper et al. (2008); it was even incorporated in some popular recent factor pricing models (Fama and French 2015; Hou et al. 2015). Wen (2019) checked whether any similar effect exist at the country level. The author provides convincing evidence that aggregate asset growth constructed from bottom-up data negatively predicts future market returns across the G7 countries. This information about future performance is not captured by other measures of investment growth and macroeconomic variables.

**Growth of government debt.** Using a set of 77 countries and data from World Development Indicators, Wisniewski and Jackson (2018) document a negative association between increases in the central government debt-to-GDP ratio and stock index returns, expressed in U.S. dollars. The authors estimate that raising the debt ratio by one percentage point decreases the stock returns by between 39 to 95 basis points. Wisniewski and Jackson (2018) explain this phenomenon with an upward pressure on private interest rates, which appears to signal a greater tax burden in the future.

**Democracy.** Lei and Wisniewski (2018) explore the role of democracy, proxied with the Political Right Index calculated by the Freedom House. Having researched a sample of 74 countries for the years 1975–2015, they conclude that, compared with autocracies, democratic states are characterized by higher returns despite displaying lower volatility risk. Lei and Wisniewski (2018, p. 1) offer three potential explanations of this effect: “First, the strength of investor protection under authoritarian leaders is relatively weak, making capital holders more vulnerable to expropriation. Second, our findings appear to be partly attributable to investors’ sentiment that is driven by media reports. Last but not least, autocracies appear to hinder the development of pension funds, suppressing thereby the demand for stocks.”

**Gravity.** Bae (2017) documents an interesting linkage between performance of different countries, namely, large countries lead returns of small countries, and this predictability decreases with geographical distance of the two countries. The effect could be translated into a long-short strategy producing about 10% risk-adjusted return per annum, which is not explained by the well-established return predictors.

**Interest rates.** Hjalmarsson (2010) investigates several potential predictors of future stock returns. The empirical results demonstrate that short-term interest rate and the term spread are fairly robust predictors of stock returns in developed markets. In contrast, Hjalmarsson (2010) finds no robust and
consistent evidence of predictability by earnings or dividend yields. Consistent evidence is provided by Charles et al. (2017).

4.9. Further Investment Considerations

Besides discovering, testing, and explaining different cross-sectional patterns, separate strains of literature examine different practical aspects of country allocation based on market-level cross-sectional patterns. From the practitioners’ perspective, two issues seem of particular importance: (1) the influence of trading costs and (2) the timing and selection of different country-allocation strategies.

Transaction costs. At the individual stock level, Novy-Marx and Velikov (2016) and Chen and Velikov (2019) demonstrate that transaction costs may have a detrimental impact on the profitability of anomaly-based quantitative strategies, in particular in the case of high turnover anomalies. At the inter-market level, the effect could be potentially even worse due to the necessity to move capital across countries. The results may also strongly depend on the implementation method chosen. Nonetheless, several studies document that when implemented with the use of ETFs, the most prominent country allocation strategies may remain profitable. Andreu et al. (2013) examine the momentum effect in single-country ETFs. They find that investors are potentially able to exploit the country momentum strategies with an excess return of about 5% per year. They note that the bid-ask spreads on ETFs are markedly below the implied break-even transaction cost levels, so the momentum effect could be profitable even after accounting for the trading costs. Finally, Blitz and van Vliet (2008) provide similar evidence extending the asset universe to additional asset classes, and Angelidis and Tessaromatis (2018) put forward analogous arguments also for value and size effects.

Factor timing and selection. The large number of different potential factor strategies that could be used to allocate money across countries raises the question of factor timing and factor selection. In other words, which strategies could be selected at a given time and how can we predict their performance? Several studies demonstrate significant time-series variation in country-level strategy returns, which can be linked, for example, to macroeconomic variables, sentiment, or liquidity and arbitrage constraints (Asness et al. 2013; Cooper et al. 2019; Ilmanen et al. 2019). Indeed, some papers provide evidence that country-level strategies could be timed. Yara et al. (2018) argue that value spreads, i.e., the differences in valuations of long and short sides of the spread portfolios, help to predict their performance. Finally, Ilmanen et al. (2019) compare several well-known anomaly selection strategies from the firm-level universe. In particular, they investigate 12 different timing signals. In general, they find weak and inconsistent evidence of factor timing. The strongest results are found for timing based on inverse volatility and valuation spreads.

5. Concluding Remarks

Over the last 30 years the asset pricing literature has accumulated remarkable evidence on the predictability of the country equity returns in the cross section. The empirical findings demonstrate numerous cross-sectional patterns in country equity indices. Some of them resemble their stock-level counterparts, such as value, momentum, or seasonality. Others, such as fund flows or political risk, are strictly characteristic for country-level return patterns.

The studies of the cross section of country equity returns use various data sets and differing methodologies. Such a situation may lead to inconclusive results and inconsistencies across papers. This highlights the need, therefore, of further standardization of country-level asset pricing studies.

In this article, we attempted to capture possibly the broadest survey of the studies of the country-level returns. Nevertheless, we acknowledge that due to possible omissions the presented picture may be incomplete. Furthermore, an additional limitation of this paper is the reliance on previously published research without any replication or verification of the accuracy of their outcomes.

The current landscape of cross sections of market index returns is growing in sophistication. The number of documented patterns is increasing. Meanwhile, the sources of this massive mispricing remain still largely unknown or not commonly agreed upon. The future studies of the topics discussed
in this paper should focus on pan-anomaly examinations that will try to bring some order into the factor structure of the country equity returns. Perhaps the multiple return predictive signals could be summarized in only several variables. Furthermore, we still need to improve our understanding of the economic mechanisms behind these patterns. Studies focused on the sources of the cross-sectional patterns in country returns would be very valuable. Finally, future investigations should also consider a practical investor’s perspective. The questions of implementability, transaction costs, or potential improvements of the trading strategies based on cross-sectional patterns would be highly valuable for market practitioners.

**Funding:** Adam Zaremba acknowledges support from the National Science Centre of Poland. This paper is a part of project No. 2016/23/B/HS4/00731 of the National Science Centre of Poland.

**Conflicts of Interest:** The author declares no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

**References**


Harvey, Campbell R., Yan Liu, and Heqing Zhu. 2016. ... and the Cross-Section of Expected Returns. *Review of Financial Studies* 29: 5–68. [CrossRef]


© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).