Erasmus School of Economics

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

International Bachelor Economics and Business Economics (IBEB)

Bachelor Thesis

The relationship between the Buffett Indicator and subsequent long-term equity returns: an analysis of 12 developed countries.

Author:
Kadir Bssila
506230

Supervised by:
L.A.P. Swinkels
Second assessor:
R.M. Spigt

Date final version: 6th of August 2021

Abstract

In 1999, Warren Buffett claimed that the ratio of the market value of equity scaled by economic output could predict subsequent long-term stock price corrections. Recent working papers have analyzed the relationship between these two variables in the context of the US and at an international aggregate level. Through the use of first-differencing, OLS regressions and Pearson correlation, this research examines the relationship between the MVE/GDP ratio, described as "Buffett indicator", and subsequent 8-year and 10-year average returns at an individual level across 12 developed countries. The paper finds empirical evidence that the indicator possesses reliable and statistically significant explanatory power for forward returns in all the 12 countries analyzed, especially between 1973 and 2020. Historical data demonstrates a significant negative correlation between MVE/GDP ratios and subsequent equity returns. This research indicates that the Buffett Indicator is a valid supplementary metric for stock market valuation and the formulation of long-term return prospects in developed countries.

Contents

1. Introduction 4

2. Literature Review 7

3. Data 9

 Short-term sample 10

 Long-term sample 12

 Country-specific data 17

4. Methodology 29

5. Results 31

 Short-term 31

 Long-term 33

6. Discussion 35

7. Conclusion 38

 Conclusive remarks on the research questions 38

 Limitations and suggestions for future research 39

8. References 41

9. Appendix 44

1 Introduction

Real capital gains of 7% per year have been earned by US equities throughout the last 100 years: no other asset class (bonds, cash, gold, or real estate) had a return potential that was close to that number (Blackrock, 2021). Nonetheless, equity markets are susceptible to very high volatility and to numerous shocks, and the returns that investors may earn are highly dependent on the time of their investment (Baele, 2005). As a result, a challenge for investors is how to predict long-term equity market performance with the greatest accuracy.

Warren Buffett, American investor and chairman of Berkshire Hathaway, publicly expressed his opinions concerning an indicator of stock market valuation and its predictive power for long-term future stock returns on several occasions between 1999 and 2001, including the Allen & Company Sun Valley Conference in 1999 and a speech to a group of Buffett's friends in 2001 (Buffett & Loomis, 2001). This indicator, which was later nicknamed the "Buffett Indicator", consists of the market capitalization of all publicly traded stock as a percentage of GDP, or the "Market Value of Equity to GDP" ratio. Financial journalist Carol Loomis summarized Buffett's views in two articles published on Fortune magazine in 1999 and 2001. Buffett portrayed the indicator as "probably the best single measure of where valuations stand at any given moment" (Buffett & Loomis, 1999).

Buffett also suggested that the market value of equity scaled by GDP would have successfully predicted the higher returns for the period going from 1981 to 1998 by indicating an undervalued stock market in the antecedent years. This low level of equity value relative to GDP was, according to Buffett (1999), a natural consequence of investors' sentiment, as investors leading to 1981 valued stocks poorly due to their expectations for the future. This, in turn, followed the widespread approach of mainly looking at how interest rates and profits had been in the past instead of projecting likely future scenarios (Buffett, 1999). As this paper finds, extraordinarily high returns did happen internationally starting in 1982, led by the lowering of US government's interest rates (Buffett & Loomis, 1999) and accompanied by an unprecedentedly sharp and extended rise in the Buffett Indicator that subsequently did not reverse to previous levels.

Fama and French (1998) claimed that meaningful return forecasts are only possible for medium- and long-term horizons, as the non-random component of the variance of returns increases in importance when the forecasting horizon widens. While validating the impossibility of forecasting equity prices over the short run with significant accuracy because of the prevalence of day-to-day noise over fundamentals, Umlauft (2020) found that fundamentals become increasingly determinant for equity prices as the forecast horizon widens. After an analysis on the US market from 1951 to 2019, the author concluded that the Buffett Indicator is a strong and consistent instrument for determining whether the stock market is under- or overpriced at any point in time. In fact, according to the results of Umflauft (2020) for the US, market valuations relative to GDP have tended to mean-revert to a temporary long-term equilibrium at least from 1952 onwards. The Buffett Indicator can then be considered a gauge of investor sentiment towards stock markets and thus towards financial assets scaled by the economy. As a growing amount of money is invested on capital markets, equity prices tend to rise without a proportionate growth in "real" economic activity: as a result, expected returns tend to drop.

However, Rajan and Zingales (2003) previously pointed out that the discrepancies of economies' equilibrium fraction of stock market capitalization to GDP are pronounced and that there are important cross-sectional differences in the real economies of different countries, warning against unsighted extrapolations to the markets of other countries or regions of the findings obtained for the US market. Therefore, the main question I investigate in this paper concerns the validity of the results found by Umlauft (2020) for the US in other developed countries with different contexts regarding financial markets and the real economy.

In particular, I aim to assess the Buffett Indicator as a predictor of medium-to-long-term returns in the context of 12 individual countries with similar levels of economic and industrial development: Australia, Belgium, Canada, France, Germany, Italy, Japan, the Netherlands, Norway, Sweden, Switzerland, and the United Kingdom. Despite their capital markets being several times smaller in size than the US market, these countries have relatively developed financial markets (Volosovych, 2011). The following main question thus needs to be assessed empirically:

Is it possible to see a relationship between the market value of equity scaled by GDP and subsequent long-term returns in developed countries other than the US, and how does the magnitude of this relationship vary among countries?

A correlated question is also examined in this paper:

Can the Buffett Indicator be considered a valid valuation indicator for the purpose of equity market analysis and the formation of long-term return expectations in individual developed countries?

As explained in detail in the Data and the Methodology section, this paper analyzes short-term and long-term samples through descriptive tables after sorting them in valuation buckets. However, the main research method implemented by this paper to assess both research questions is represented by univariate OLS regressions ran over the samples and over their subsamples, comparing the time series of the MVE/GDP ratio with 8-year and 10-year returns.

In addition to an analysis of the BI's predictive power over all samples, this paper puts additional focus on each individual country analyzed. This is in order to not misattribute the BI's properties that hold at the aggregate level to individual countries, disregarding idiosyncrasies in a similar way to the group-to-individual generalization found in studies of human behavior (Fisher, Medaglia & Jeronimus, 2018).

My first major finding is that throughout history, the MVE/GDP and subsequent long-term returns have a negative and significant Pearson correlation across all 12 developed countries analyzed in this paper, after accounting for non-stationarity and for serial correlation of the time series. Second, the paper finds especially high explanatory power of the MVE/GDP ratio during the most recent decades, demonstrating the current validity and relevance of this indicator for long-term equity market analysis.

The remainder of this paper proceeds as follows. Section 2 delivers an overview of the existent literature on the Buffett Indicator compared to the main fundamentals and metrics used in market valuation analysis. Section 3 describes the data and defines the variables used to perform this investigation. Section 4 outlines the methodology of this paper and illustrates the reasoning behind it. Section 5 provides an extensive report of the results for each individual country, providing further country-specific information relative to the compiling of datasets. I compare my results discussing their applicability in section 6 and conclude in section 7.

2 Literature Review

Despite this strong claim by arguably one of the worlds' most successful investors, the predictive power of the so-called Buffett Indicator for the whole stock market has drawn surprisingly minimal academic scrutiny in comparison to other fundamentals-scaled valuation indicators.

Campbell and Shiller (1988) implemented the first research on the applicability of the value effect of individual stocks to entire equity markets. Through scaling the value of the S&P 500 by the cumulative earnings of all the firms in the index, obtaining the Price-to-earnings ratio (PE), the two Yale professors demonstrated that high market valuations often precede years of low equity returns. The value effect for individual stocks, in fact, implies that undervalued stocks tend to return to their intrinsic value realizing much greater capital growth than overvalued stocks (Asness, Moskowitz, Pedersen, 2013; Fama & French, 1992; Lakonishok, Shleifer & Vishny, 1994).

The traditional PE, on the other hand, has two main drawbacks. The first is that corporate earnings are of very volatile nature. For example, from 2009 to 2010, the earnings of the S&P 500 companies swung between 7 and 77 points (Keimling, 2016). P/E ratios are disproportionally impacted by earnings variations, which are in turn heavily influenced by the phases of the business cycle (Molodovsky 1953). Therefore, the current level of returns is not always revealing their future growth. Furthermore, P/E ratios are especially unappetizing in times of crisis, when low or negative corporate earnings offer advantageous purchasing opportunities. The PE does not take into consideration the possibility for profits growth following the crisis at such periods. As a result, during the start of the economic cycle, when the economy is rebounding from contractionary conditions associated with low earnings, price-earnings ratios may look unreasonably high. On the other side, as the economic cycle progresses and profit margins rise, P/E may be too small due to excessively high transitory earnings components (Molodovsky, 1953).

To avoid the distortionary effect of temporary earnings on the P/E ratio, Graham, Dodd and Cottle (1934) had famously advised smoothing the denominator by taking the mean of earnings over the previous 5 to 10 years:

"A conservative valuation of a stock issue must bear a reasonable relation to the average earnings. […] This approach shifts the original point of departure, or basis of computation, from the current earnings to the average earnings, which should cover a period of not less than five years, and preferably seven to ten years." (Graham, Dodd & Cottle, 1934)

Based on this suggestion, Campbell and Shiller (2001) recommended taking a mean of gains over the previous ten years adjusted for inflation to smooth earnings fluctuations over the course of about 2 business cycles, obtaining the famous Cyclically-Adjusted Price to Earnings (CAPE) ratio. This smoothing of profits makes the CAPE ratio, also named P/E10 or Shiller P/E, more suitable for measuring earnings that are sustainable for longer periods. At the same time, adjusting for inflation guarantees the comparability of profits even during periods of high inflation. The CAPE ratio successfully predicted low US equity returns after the year 2000 and has been demonstrated to hold strong predictive power across 17 MSCI countries in subsequent research (Keimling, 2015). Several papers (Bunn and Shiller, 2014, Keimling, 2005, 2016; Klement, 2012) researched the association between m CAPE and subsequent long-term returns in the US and at an international level, demonstrating the consistency of the Shiller PE as a long-term indicator for both developed and emerging markets. However, the same has not yet been done at a similar completeness for the MVE/GDP ratio.

As mentioned in the introduction, in a recent working paper, Umlauft (2020) researches the BI and forward returns in the context of the US, using a similar methodology to the one used here. In another working paper, Kuvshinov & Zimmermann (2018) analyze the historical movements of the MVE/GDP ratio across countries and its predictive power for subsequent equity returns, but their analysis aggregates 17 economies and does not consider countries individually. On the other hand, my paper highlights localized trends that are not visible at an aggregate level and provide more precise findings from which to extract insight for future academic research and for long-term investment decisions.

According to the European Central Bank, the average business cycle in the Euro Area spans between 6 and 9 years (Lenza, Reichlin & Giannone, 2009). Therefore, given the limited length of the observation period, I set the forward average return period for the short-term analysis to 8 years. For the long-term analysis, given a larger observation period, I utilize 10-year average returns, considering their prevalent use in the literature (Keimling, 2016; Campbell & Shiller, 2001; Wilcox, 1984).

3 Data

The variables used in the regressions for each individual country are: real returns ("RR", "returns"), obtained by adjusting total equity market returns ("Total Returns". "TR") for inflation; the market value of the equity or total market capitalization ("MVE", "market capitalization"); the national Gross Domestic Product ("GDP"). The MVE is then divided by GDP to compute the "Buffett Indicator" ("BI", "MVE/GDP ratio", "MVE/GDP").

This research paper implements two analyses: a short-term and a long-term analysis. This is because to run predictive models, it is necessary to match observations of the same frequency for all three variables. The highest frequency at which GDP data is available is quarterly. However, data is available at this frequency only for relatively recent years – roughly 5 decades - whereas complete annual data is available for some countries starting in the XIX century. Therefore, to investigate historical predictability of forward returns based on the Buffett Indicator, it is ideal to examine the largest timeframe possible, to comprise different societal and economic settings. On the other hand, it is important to include the highest data frequency available so that the analysis is as precise as possible. Therefore, where possible, two different samples are created for each country analyzed, using all consecutive periods with an observation for all 3 variables: one sample with quarterly data, used for the short-term analysis, and one with annual data, used for the long-term analysis. Two additional variables are used to obtain Real Returns and transform GDP data of the long-term samples, namely inflation measured by CPI and the countries' domestic currency exchange rate measured against the United States Dollar. A challenge when using stock market capitalization indices is that the measure of the Market Value of Equity of a country should total the value of the shares listed on all domestic stock exchanges. For most countries in this paper's data, the bias from excluding smaller exchanges is insignificant or non-existent due to the centralization of stock markets and to many securities exchanged on minor markets being also listed on the countries' major stock exchange (Kuvshinov & Zimmermann, 2018). However, a greater potential bias arises for early Canada data, where the only historical data I was able to retrieve regarded the Toronto Stock Exchange exclusively. Multiple large stock exchanges and an active curb market – in Montreal (Jorgensen, Kavajecz & Swisher, 2011) - were in operation in the early XX century, making it unfeasible to use the Toronto SE as a proxy for the Canadian market in the long-term (Sylla, Wilson & Wright, 2006). Therefore, Canada is excluded from the long-term analysis of this paper.

Short-term sample

Market value of equity (MVE) and real equity returns (RR)

To obtain data on the total market capitalization and total equity returns, Thomson Reuters Datastream (DS) is used for all countries analyzed. This paper uses Datastream's own Total Market indices, representing all the stocks trading in a country's stock market. Datastream's indices are available at yearly and quarterly frequency for most of the countries analyzed starting in 1973. Therefore, the largest available observation periods for data-types RI (Returns Index) and MV (Market Value) are downloaded from Datastream. The MV data-type contains data on the Market Value of Equity ("MVE") expressed in domestic currency, while the RI data-type is expressed in one-quarter percentage changes to obtain the Total Returns ("TR") of the national stock market. Table 2 summarizes the sources of the quarterly data for each country. This investigation focuses on Real Returns (RR), i.e. nominal total returns adjusted for inflation. To convert TR into RR, quarterly data on the Consumer Price Index (CPI) are downloaded from Global Financial Data. Inflation is measured by the annual rate of change of the CPI. Then, Real Returns (RR) were obtained through the following calculation:

$$RR = \frac{1 + TR}{1 + \Delta CPI} - 1$$

Here, ΔCPI is the annual percentage change of the Consumer Price Index. The data on quarterly real returns is annualized to account for the period-on-period compounding, which makes the results of this paper usable by investors to gauge potential long-term performance of potential equity investments. Indeed, the volume of investment lost or earned each year is dependent from the amounts earned/lost in previous years under investigation because of compounding. This is done by taking the geometric average of returns according to the following formula:

$$forward\ return_n = [(1 + R_1) * (1 + R_2) \ldots * (1 + R_n)]^{\frac{1}{n}} - 1$$

Gross Domestic Product (GDP)

GDP data is obtained either through the International Monetary Fund database ("IMF") or through the database of Global Financial Data ("GFD" or "GFDatabase"). As it is managed by an official financial institution, the IMF database is the preferred source for GDP data. However, GFDatabase is used where IMF 's observation period for quarterly data starts later than 1973, the beginning date of the observation period available on Datastream. In fact, the availability of data for all three variables determines the final sample length; e.g. if quarterly data for two variables are available starting in Q1 of 1973 but quarterly observations for the third variable only start in Q1 of 1981, the final sample for data at quarterly frequency selected for the analysis must start in 1981. All quarterly GDP data from both IMF and GFD are seasonally adjusted. While all quarterly GDP data from the IMF database is expressed in EUR for countries the Euro area, quarterly GDP obtained from GFD for Belgium, Italy, Germany and the Netherlands is expressed in Euros starting only in 1999. Therefore, data preceding this year is converted into Euros to maintain consistency in the unit of measurement. This paper implements the same conversion method used by the IMF's World Economic Outlook (WEO) Reports, a survey published by the IMF twice every year. To convert the data prior to 1999, the WEO used the fixed conversion rates between the Euro and the countries' local currencies, adopted by the Council of the European Union on January 1, 1999. The rates used by the WEO and in this paper are the following:

Table 1. *Local currency per Euro, fixed conversion rates for data preceding 1999.*

Belgian francs	**German marks**	**Italian lire**	**Dutch guilders**
40.3399	1.9558	1936.2700	2.2037

Table 2. *Short-term sample: start of the available observation period and data sources for all 12 countries and the 3 variables analyzed.*

Variable	**GDP**		**MVE**		**TOTAL RETURNS**		**Final**
Source	**Start**	**Source**	**Start**	**Source**	**Start**	**Source**	**Sample start**
Australia	1960	IMF	1973	DS	1973	DS	1973
Belgium	1980	GFD	1973	DS	1973	DS	1980
Canada	1961	IMF	1973	DS	1973	DS	1973
France	1949	IMF	1973	DS	1973	DS	1973
Germany	1961	GFD	1973	DS	1973	DS	1973
Italy	1961	GFD	1973	DS	1973	DS	1973
Japan	1955	IMF	1973	DS	1973	DS	1973
Netherlands	1977	GFD	1973	DS	1973	DS	1977
Norway	1978	IMF	1980	DS	1981	DS	1981
Sweden	1981	GFD	1982	DS	1982	DS	1982
Switzerland	1970	IMF	1973	DS	1973	DS	1973
United Kingdom	1955	IMF	1965	DS	1965	DS	1965

Long-term sample

Table **3** summarizes the sources for the annual data used in the long-term analysis and the start of the observation period for each of the 3 variables (GDP, MVE, and TR). Column 4 of Table **3** indicates the start of the sample used in the regressions for the long-term analysis.

Market Value of Equity (MVE)

For the long-term analysis, data on the total market capitalization of the companies listed on each country's national stock exchange (MVE) are downloaded from Global Financial Data and are expressed in USD. GFD uses different sources to compile its datasets. In fact, data from the interwar period are downloaded from "Statistisches Jahrbuch fur das Deutsche Reich" (Statistisches Reichsamt, various issues). Among the sources are: the International Federation of Stock Exchanges (FIBV), International Finance Corporation, Marcel Garneau, World Directory of Stock Exchanges, W.I.S.E.R. Research, 1995, World Stock Exchange Fact Book, the World Federation of Exchanges website, and annual reports issued by national stock exchanges and statistical agencies. Column 2 of Table **3** displays the start of the observation period for MVE data available at yearly frequency on GFD. Following 30 years of decennial and quinquennial observations, annual market capitalization data for the Italian Stock Exchange on Global Financial Data starts only in 1975 - later than the quarterly data used for the short-term analysis. Therefore, a long-term analysis is not implemented for Italy. The same holds for Norway, as consistent annual observations on the Oslo stock exchange market capitalization start only in 1974.

Gross Domestic Product

Because of the need to maximize the observation period for the long-term analysis, annual data on the Gross Domestic Product data and Total Returns index are downloaded from the Jordà-Schularick-Taylor "Macrohistory" dataset. For some countries, Macrohistory has an observation period of 127 years (1870-2017). Macrohistory's GDP data are expressed in nominal terms and in local currency. Given that the market capitalization data downloaded from Global Financial Data are expressed in United States Dollars (USD), the GDP data are converted from local currency into USD using the USD exchange rate (local currency/USD) provided by the same Macrohistory dataset. For 2018, 2019, and 2020, years for which Macrohistory has no observations, official GDP data is downloaded from the International Monetary Fund database and converted into USD using the USD/local currency exchange rate published on GFD.

Total Returns and Real Returns

Total Returns data were downloaded from the Jordà-Schularick-Taylor Macrohistory dataset due to its larger observation period for some countries compared to GFD. Total Returns observed in 2018, 2019, and 2020 – for a few countries also in 2015 and 2016 - are from annual DataStream's market indices. TR are converted in Real Returns using – where possible – Consumer Price Index (CPI) data taken from the Macrohistory dataset and converted into a percentage change from the previous year to measure inflation. Where CPI data is not available on Macrohistory, i.e. the years 1870, 2018,2019, and 2020, inflation data measured in CPI percentage changes are obtained from Global Financial Data. More details on the quarterly and annual data used for each country are provided in each country's Results subsection.

Table 3. *Long-term sample: start of the observation period and data source for all the 12 countries and the 3 variables analyzed.*

Variable	**GDP**	**MVE**	**TR**	**FINAL**
Source	MH	GFD	MH	MH
Australia	1870	1825	1870	1870
Belgium	1870	1956	1870	1956
Canada	-	-	-	-
France	1870	1801	1870	1870
Germany	1901	1900	1870	1901
Italy	-	-	-	-
Japan	1875	1901	1886	1901
Netherlands	1870	1899	1900	1900
Norway	-	-	-	-
Sweden	1870	1870	1871	1871
Switzerland	1870	1899	1900	1900
United Kingdom	1870	1689	1871	1871

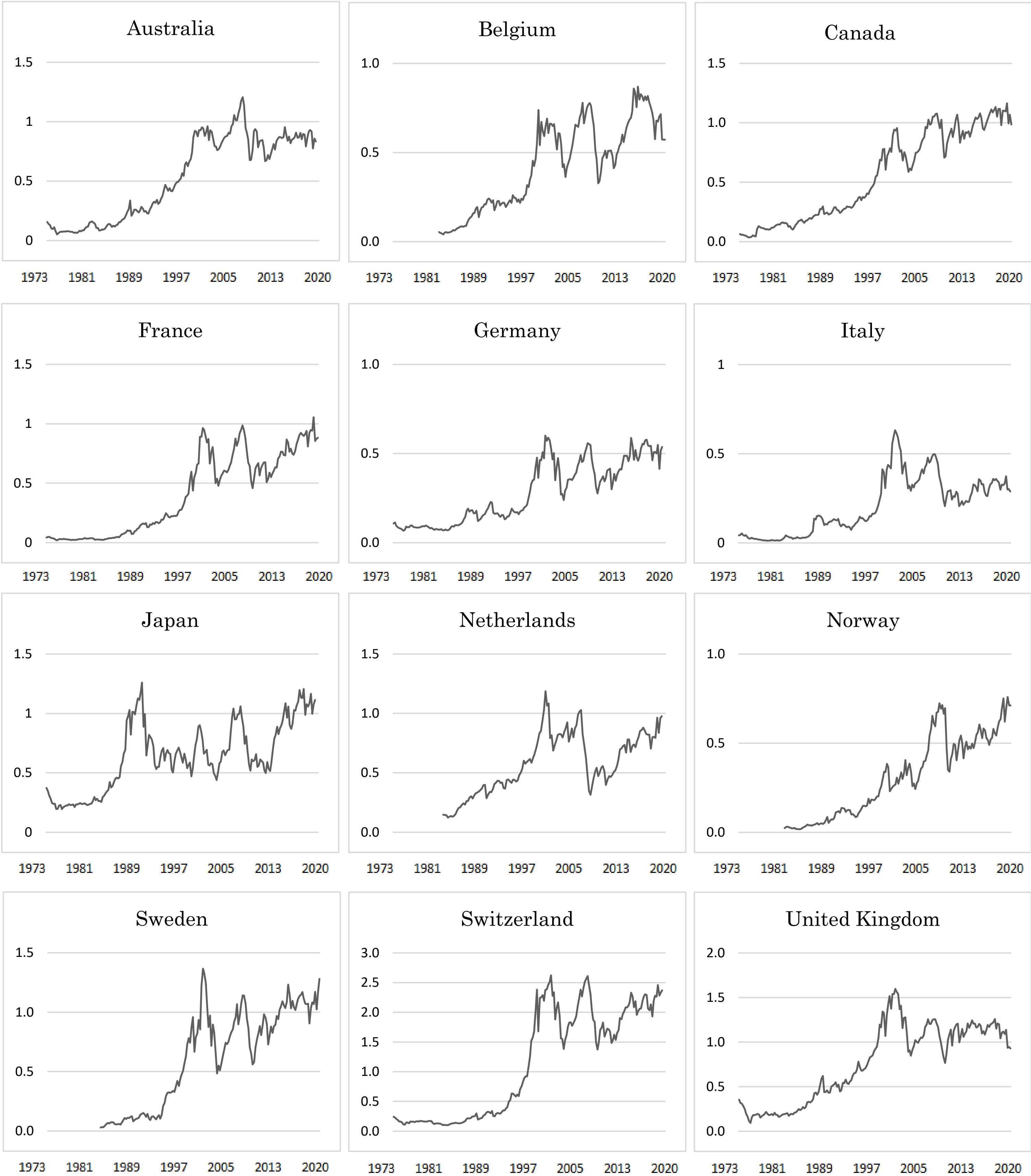

Figure 1. *A visual cross-country comparison of the short-term MVE/GDP time series through a line plot that covers the chosen observation period, 1973:Q1-2020:Q4.*

Short-term line plots

Figure 1 shows how over the most recent four decades, the stock market has experienced a historically unprecedented growth. The median market capitalization scaled by GDP remained well below 0.5 throughout the 1970s for all the analyzed countries and started increasing only in the second half of the 1980s, increasing sharply in the 1990s and reaching a new equilibrium level for all countries analyzed. For Japan, this sharp increase of the MVE/GDP ratio happens sharply earlier than for the other countries, in the 1980s, with a peak value of around 1.26 reached in Q1 1990. For Italy and Canada, on the other hand, the steep increase occurred only in the late 1990s. All countries reached a new peak right before the drop in market valuations at the time of the dot-com market crash – around 2001. percentage decrease of the Australian BI following the dot-com crisis is also not as steep as that of all other countries analyzed: that MVE/GDP ratio lost less than 20%. of its previous value and then started to increase again until the 2008 financial crisis (Ofek & Richardson, 2003). This rise in the level stock market capitalization in absolute terms and relative to GDP seems to be persistent and to have resisted the sharp equity price corrections resulting from the Internet bubble in the early 2000s and from the Global Financial Crisis of 2008-2009. In fact, Buffett Indicators today remain from four to over ten times greater than their usual level throughout the XX century until the '70s, which appears to indicate a new equilibrium for all countries analyzed. In particular, the Swiss equity market reached a peak value of over 2.5 times the size of the country's GDP in correspondence of the dot-com bubble and the 2008 Global Financial Crisis. Figure 1 attests that the Big Bang, as Kuvshinov and Zimmermann (2018) call the steep increase in the MVE/GDP ratio at the end of the XX century, is very much a cross-country phenomenon. The sharp equity market growth of the 1980s and 1990s is evident in all countries of this paper's dataset. For France, the United Kingdom, and in smaller measure for Germany and Australia, this dramatic shift of equities can be interpreted as a mean reversion to a higher level of market valuation preceding WWI and WWII, as shown by the long-term line plots in the Appendix that are described hereunder.

Long-term line plots

The long-term line plot for Australia displayed in Figure A1 shows a low equilibrium level of the BI is evident around 0.1 up to 1900 and around 0.2 until 1930. The graph shows a sharp and unprecedented increase of the BI in 1930, leading to an apparent new equilibrium level around 0.6 until the early 1970s. Here, a steep decrease brought the BI to a new median level of about 0.2 which lasted until the Big Bang. The long-term plot of Belgium displayed in Figure A2 only adds 20 years of yearly data to the short-term plot, but it hints at a decrease of the BI before the Big Bang, throughout the '60s and '70s, beginning at 0.36 in 1961 and ending below 0.1 in 1981. The MVE/GDP ratio of France and of the UK follows a U-shape over their long-term time plot in figure A3 and A9, reflecting high financial market development preceding the World Wars (Volosovych, 2011). The BI reached a minimum following the second World War, remaining at a low equilibrium level until the 1980s, right before the Big Bang and the subsequent mean-reversions to around 1. This new level is similar to the pre-World Wars equilibrium for both countries. The long-term line plot for Germany shows a higher level of the BI's equilibrium- between 0.2 and 0.4 - from 1870 to the end of the 1930s. Following an exceptional temporary peak of over 1 in 1947 corresponding to a severe reduction in the German GDP after WWII, the BI appears to stabilize around 0.1 until the 1980s. After remaining relatively stable around 0.15 for the first three decades of the XX century, Japan's MVE/GDP ratio sharply increases in the 1930s, returning to its previous level during the 1940s - data is missing due to WWII – and then starting a slow increase until the Big Bang. The long-term time plot of the Netherlands' BI in Figure A6 indicates that right before their unprecedented increase, market valuations had reached a general minimum within the sample of 0.07 in 1981, after undergoing a decreasing trend throughout the XX century. The long-term line plot of Sweden's MVE/GDP ratio shows a sharp increase of 4 decimal points of the indicator between 1905 and 1907, followed by a sharp price correction to previous levels in correspondence of WWI. The BI's equilibrium level then appears to remain stable around 0.2 for several decades until the 1980s.

The long-term line plots of the MVE/GDP ratio reported in the Appendix show some homogeneity in the time series of the MVE/GDP across all countries: for many countries, the level of the BI reached either a local or a general minimum at the end of the 1970s or at the beginning of the 1980s. However, as previously showed by Rajan and Zingales (2003), there are visible cross-country differences in the magnitude and timing of the BI fluctuations, other than in the time-varying steady state of the MVE/GDP. In the case of Japan, for example, the BI had not decreased a minimum value at the beginning of the Big Bang, but it had remained at a level that was relatively close level to its equilibrium throughout the XIX and XX centuries.

Country-specific data

The following section of the paper includes more specific information regarding the data used in this paper, but it also investigates to some extent the predictive power of the Buffett Indicator for long-term equity returns. Hereunder, the samples used for each country's short-term and long-term analysis are summarized in descriptive tables, and the data within them is described to provide insight on the country-specific relationships between the MVE/GDP ratio and forward returns. Within a descriptive table, each row represents a valuation bucket. The first two columns of each table report the two extreme BI values for each bucket of observations: each bucket spans two decimal points. The number of observations for each bucket are displayed in the rightmost column of each table. Column 3, 4, 5, 6, 7 of each descriptive table report the mean forward returns, median forward returns, the standard deviation of forward returns, the minimum and the maximum returns, respectively. A description of each sample's noticeable relationships with respect to the BI follows every table.

Australia

Short-term

The short-term sample for Australia goes from Q1 of 1973 to Q4 of 2020. All observations of GDP and MVE are expressed in Australian Dollars.

Table 4. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of Australia, 1973:Q1-2020:Q4* .

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	2.0%	2.2%	0.8%	-0.5%	3.3%	56
0.2	0.4	2.5%	2.7%	0.7%	0.8%	3.6%	27
0.4	0.6	1.9%	1.9%	0.2%	1.7%	2.2%	15
0.6	0.8	1.9%	1.8%	0.3%	1.0%	2.4%	21
0.8	1	1.2%	1.1%	0.5%	0.4%	2.5%	33
1	-	0.3%	0.4%	0.3%	-0.2%	0.7%	8

As Table 4 above reports, for Australia's short-term sample, except for the lowest valuation bucket mean and median 8-year returns monotonically decrease as the BI increases. Mean quarterly forward equity returns amount to 2.0% when MVE/GDP ratios are below 0.2 and to 0.3% when the ratio was above 1. Standard deviations of returns are relatively low and tend to decrease with the BI. Minimum and maximum returns also tend to decrease with higher BIs.

Long-term

The long-term sample of Australia begins in 1871 and ends in 2020.

Table 5. *Buffett Indicator buckets and subsequent 10-year annual real returns of Australia, 1871-2020.*

MVE/GDP BUCKET		10-YEAR ANNUAL REAL RETURNS					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	9.2%	9.6%	2.8%	1.7%	15.2%	51
0.2	0.4	8.5%	9.0%	2.7%	3.0%	12.9%	24
0.4	0.6	5.9%	6.0%	3.8%	0.1%	12.2%	26
0.6	0.8	4.1%	4.8%	4.4%	-3.8%	10.3%	20
0.8	1	4.8%	6.2%	3.9%	-2.9%	10.7%	9
1	-	2.9%	3.8%	3.0%	-4.6%	5.5%	10

Table 5 above shows some distinct trends for Australia's long-term sample: mean and median 10-year annual forward returns both almost monotonically decrease as the BI increases, interrupting their trend in the 0.8-1 bucket. Mean 10-year annual returns amount to 9.2% when the BI is smaller than 0.2, while they average 0.2% when the ratio is above 1.2. While maximum annual forward returns monotonically decrease as the Buffett Indicator increases, the decrease of minimum forward returns is non-monotonic. Standard deviations do not show clear trends but are always relatively small.

Belgium

Short-term

The short-term sample for Belgium covers a period going from Q1 of 1981 to Q4 of 2020. All observations of GDP and MVE are expressed in Euros. Data from 1981 to 1995 were computed by GFD using the Chow-Lin method of interpolation and extrapolation of national accounts' indicators.

Table 6. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of Belgium, 1981:Q1-2020:Q4* .

MVE/GDP BUCKET		8-YEAR QUARTERLY REAL RETURNS					
Low	**High**	**Mean**	**Median**	**St.Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	3.0%	2.9%	1.4%	0.8%	6.0%	34
0.2	0.4	2.4%	2.3%	0.8%	1.2%	4.1%	36
0.4	0.6	1.5%	1.5%	0.7%	-0.3%	2.7%	31
0.6	0.8	0.6%	0.6%	0.6%	-1.2%	1.6%	27

In Belgium's short-term sample, mean and median 8-year quarterly forward returns both monotonically decrease as MVE/GDP increases, as displayed in Table 6. Mean forward returns amount to 3.0% when the BI is below 0.2, and to 0.6% when the indicator is above 1.2. Standard deviations of returns and maximum returns present a monotonic decrease that excludes the lowest valuation bucket for minimum returns.

Long-term

The long-term sample of Belgium extends from 1961 to 2020: this observation period is limited in size compared to those of other countries, only including 50 observations of 10-year forward returns.

Table 7. *Buffett Indicator buckets and subsequent 10-year annual real returns of Belgium, 1961-2020.*

MVE/GDP BUCKET		10-YEAR ANNUAL REAL RETURNS					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	12.4%	13.1%	3.5%	5.3%	15.9%	11
0.2	0.4	4.1%	3.3%	5.4%	-3.4%	12.0%	24
0.4	0.6	11.3%	9.1%	6.4%	6.0%	22.6%	7
0.6	-	0.6%	0.8%	2.8%	-5.1%	4.4%	8

Table 7 displays no distinct trends for median and minimum returns in the Belgian long-term sample. Mean and maximum returns non-monotonically decrease as the BI increases. Mean 10-year annual returns measure 11.3% for values of the BI smaller than 0.2 and 0.6% when the BI is above 1.2.

Canada

Short-term

The short-term sample for Canada extends from Q2 1973 to Q4 2020. GDP and MVE are expressed in Canadian Dollars.

Table 8. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of Canada, 1973:Q2-2020:Q4* .

MVE/GDP BUCKET		8-YEAR QUARTERLY REAL RETURNS					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	1.3%	1.2%	0.8%	-0.7%	2.8%	51
0.2	0.4	2.5%	2.6%	1.0%	0.6%	4.5%	38
0.4	0.6	2.3%	2.3%	0.2%	2.0%	2.5%	9
0.6	0.8	1.8%	1.9%	0.5%	0.6%	2.7%	26
0.8	1	1.1%	1.1%	0.4%	0.3%	1.7%	26
1	-	0.5%	0.5%	0.2%	0.2%	0.8%	9

As displayed in Table 8, Canada’s mean and median 8-year quarterly real returns monotonically decrease as the MVE/GDP ratio increases, except for the lowest and largest valuation bucket. Forward returns average 2.5% when the BI is larger than 0.2 and lower than 0.4 and 0.5% when the BI is above 1. Minimum and maximum forward returns show a monotonic decline starting in the valuation buckets where the BI is higher than 0.4 and 0.2, respectively. Standard deviations are small, and they tend to decrease as the BI is higher.

France

Short-term

The observation period of the French short-term sample extends from Q1 of 1973 to Q4 of 2020. All observations of GDP and MVE are expressed in Euros.

Table 9. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of France, 1973:Q1-2020:Q4 .*

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	2.7%	3.0%	1.5%	-1.5%	5.3%	83
0.2	0.4	2.3%	2.3%	0.4%	1.7%	3.1%	17
0.4	0.6	1.9%	2.0%	0.7%	0.5%	3.0%	22
0.6	0.8	1.2%	1.0%	0.7%	0.3%	2.3%	22
0.8	-	0.0%	0.3%	0.6%	-1.1%	0.6%	16

In France's short-term sample, mean and median 8-year quarterly forward returns both monotonically decrease as the MVE/GDP increases, as Table 9 shows. Mean forward returns amount to 2.7% when MVE/GDP ratios are below 0.2, whereas the mean of returns is 0.0% for BIs above 0.8. Standard deviations of returns tend to decrease in higher MVE/GDP buckets. Except for the lowest MVE/GDP bucket, minimum and maximum returns monotonically decrease with a higher MVE/GDP.

Long-term

The long-term sample of France extends goes 1870 to 2019. Data from 1933 to 1945 and before 1872 are estimates based on data from GFD's own Indices.

Table 10. *Buffett indicator buckets and subsequent 10-year annual real returns of France, 1870-2019.*

MVE/GDP BUCKET		**10-YEAR ANNUAL REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	-0.3%	2.3%	11.2%	-26.2%	13.3%	36
0.2	0.4	1.8%	5.4%	9.5%	-24.5%	15.5%	37
0.4	0.6	2.8%	3.7%	3.7%	-7.7%	8.4%	35
0.6	0.8	-1.0%	-0.2%	5.3%	-10.6%	7.8%	25
0.8	-	-0.4%	-1.0%	3.0%	-5.3%	3.0%	7

The long-term sample of France does not show clear trends for mean returns, whereas except for the lowest valuation bucket median forward returns decrease as the MVE/GDP ratio grows. The standard deviation of forward returns tends to decline as the BI increases, while minimum and maximum forward returns almost monotonically increase with the BI.

Germany

Short-term

The short-term sample for Germany goes from Q1 of 1973 to Q4 of 2020. All observations of GDP and MVE are expressed in Euros.

Table 11. *Buffett indicator buckets and subsequent 8-year quarterly real returns of Germany, 1973:Q1-2020:Q4.*

MVE/GDP BUCKET		8-YEAR QUARTERLY REAL RETURNS					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	1.9%	2.0%	1.1%	-0.6%	4.1%	91
0.2	0.4	1.6%	1.6%	0.7%	0.5%	3.1%	39
0.4	-	0.6%	0.8%	0.7%	-1.0%	1.4%	30

Table 11 shows that in the short-term sample of Germany, both mean and median quarterly 8-year returns monotonically decrease as the BI increases. Average forward returns amount to 1.9% for BIs below 0.2 and to 0.6% when the BI is larger than 0.4. Standard deviations of returns decrease with higher BIs. While minimum returns do not show any trends relative to the BI buckets, maximum returns monotonically decrease as the MVE increases relative to GDP.

Long-term

The long-term sample of Germany starts in 1881 and ends in 2020. For the GDP of 1945, as observations on Macrohistory were missing, GFD data was used.

Table 12. *Buffett indicator buckets and subsequent 10-year annual real returns of Germany, 1872-2020.*

MVE/GDP BUCKET		10-YEAR ANNUAL REAL RETURNS					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	14.2%	12.0%	15.3%	-20.1%	37.4%	12
0.2	0.4	4.4%	7.8%	7.7%	-15.4%	12.8%	34
0.4	0.6	3.8%	5.1%	6.9%	-18.6%	35.7%	71
0.6	-	0.6%	1.6%	7.0%	-14.1%	12.3%	22

Table 12 shows a monotonic decline of median forward returns corresponding to an increase in the BI of Germany's long-term sample. Exhibiting a monotonic decline, average annual forward returns amount to 14.2% when the MVE/GDP ratio is lower than 0.1 and to 0.6% in the highest valuation bucket. Minimum returns tend to increase with the BI, whereas maximum returns do not display unambiguous trends relative to MVE/GDP.

Italy

Short-term

The short-term sample for Italy begins in Q1 of 1973 and ends in Q4 of 2020. All observations of GDP and MVE are expressed in Euros.

Table 13. *Buffett indicator buckets and subsequent 8-year quarterly real returns of Italy, 1973:Q1-2020:Q4.*

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.1	2.5%	2.6%	1.9%	-0.7%	6.4%	62
0.1	0.2	1.0%	1.4%	1.3%	-1.3%	3.0%	36
0.2	0.4	0.7%	1.0%	1.0%	-1.2%	2.3%	37
0.4	-	-0.5%	-0.8%	0.9%	-1.8%	1.1%	25

Table 13 indicates that, for Italy's short-term sample, both mean and median 8-year quarterly real returns monotonically decline as the BI increases. Average forward returns amount to 2.5% for MVE/GDP ratios below 0.1 and to -0.5% when the BI is larger than 0.4. Standard deviations and maximum returns present a monotonic decline corresponding to an increase in the BI. Minimum returns do present any clear trends w.r.t the BI.

Japan

Short-term

The short-term sample for Japan covers a period that starts in Q1 of 1973 and ends in Q4 of 2020. GDP and MVE data are expressed in Euros.

Table 14. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of Japan, 1973:Q1-2020:Q4.*

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	1.7%	1.8%	0.5%	1.2%	2.3%	3
0.2	0.4	2.9%	3.1%	1.7%	-1.0%	5.1%	44
0.4	0.6	1.1%	1.1%	1.1%	-0.9%	3.0%	41
0.6	0.8	0.0%	-0.4%	1.1%	-1.4%	2.6%	43
0.8	1	-0.5%	-0.3%	0.8%	-1.9%	0.7%	19
1	-	-1.3%	-1.5%	0.7%	-2.6%	0.2%	10

In the short-term sample of Japan, mean and median forward real returns decline nearly monotonically as the BI increases, apart from the lowest valuation bucket. The standard deviation of forward returns also decreases with higher valuation buckets. The mean of observed forward returns is 2.9% when the BI is between 0.2 and 0.4 and -1.3% when the BI is above 1. While maximum returns monotonically decrease with higher BIs, minimum returns follow nearly monotonic declining trend relative to the BI.

Long-term

The long-term sample of Japan extends from 1901 to 2020, excluding the decade 1941-1950 due to missing data on both Macrohistory and GFD in 1945 (most likely because of World War II). Data for the Japanese MVE were compiled by GFD using the Tokyo Stock Exchange Capitalization dataset. Data from 1900 to 1925 are from Lyndon Moore's "World Financial Markets 1900-1925" (Moore, 2012). Data from 1926 to 1929 are GFD estimates. Data from 1930 to 1940 and from 1950 to date were compiled using the "Tokyo Stock Exchange's Monthly Statistical Handbook" (Tokyo SE, various issues).

Table 15. *Buffett Indicator buckets and subsequent 10-year annual real returns of Japan, 1870-2019.*

MVE/GDP BUCKET		**10-YEAR ANNUAL REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	9.1%	7.6%	6.3%	1.6%	28.4%	35
0.2	0.4	9.1%	8.9%	5.2%	-0.7%	24.2%	33
0.4	0.6	4.4%	3.4%	6.2%	-3.8%	18.7%	10
0.6	0.8	2.0%	0.0%	6.0%	-3.9%	15.4%	12
0.8	1	-0.6%	1.0%	3.5%	-6.3%	2.4%	6
1	-	-0.9%	-0.4%	5.2%	-6.5%	3.7%	4

The long-term sample of Japan presents a monotonic decrease of average forward returns corresponding to an increase of the BI, whereas a non-monotonic fall of median returns is observed as the BI grows. 10-year returns average 9.1% when the BI is lower than 0.2 and -0.9% for MVE/GDP ratios above 1. There is no definite trend for standard deviations, while both minimum and maximum returns monotonically decrease as the BI increases.

Netherlands

Short-term

The short-term sample for the Netherlands covers a period including observations from Q1 of 1981 to Q4 of 2020. All observations of GDP and MVE are expressed in Euros.

Table 16. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of the Netherlands, 1981:Q1-2020:Q4 .*

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	4.3%	4.4%	0.5%	3.4%	5.0%	10
0.2	0.4	2.7%	2.7%	0.7%	1.8%	4.8%	29
0.4	0.6	3.2%	3.2%	1.3%	1.2%	5.1%	33
0.6	0.8	1.1%	1.2%	0.8%	0.1%	2.9%	16
0.8	1	0.3%	0.3%	0.4%	-0.1%	1.1%	16
1	1.2	0.1%	-0.1%	0.6%	-0.3%	1.1%	6
1.2	-	-0.3%	-0.1%	1.0%	-2.3%	0.8%	18

As Table 16 shows, the mean of observed forward returns of the Netherlands' short-term sample is 4.3% when MVE/GDP ratios are below 0.2. After a non-monotonic decrease in mean and median forward returns, in the highest valuation bucket forward returns average -0.3%. Standard deviations do not follow any clear trends with respect to the BI, whereas minimum and maximum forward returns display, respectively, a monotonic and non-monotonic decrease as MVE/GDP increases.

Long-term

The observation period of the Netherlands' long-term sample starts in 1901 and includes annual observations up to and including 2020.

Table 17. *Buffett indicator buckets and subsequent 10-year annual real returns of the Netherlands, 1901-2020.*

MVE/GDP BUCKET		**10-YEAR ANNUAL REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	10.4%	12.7%	5.9%	-1.8%	17.8%	27
0.2	0.4	5.2%	3.0%	6.8%	-4.8%	19.3%	40
0.4	0.6	1.9%	2.6%	4.3%	-6.0%	7.8%	20
0.6	-	1.6%	2.8%	4.5%	-5.6%	8.2%	12

In the Netherlands' long-term sample, the mean and median forward returns monotonically decline as MVE/GDP increases. Mean returns amount to 10.4% and 1.6% for the lowest and the highest valuation bucket, respectively. Minimum and maximum returns become lower as the BI increases, although not in a constant way.

Norway

Short-term

The short-term sample for Norway starts in Q1 of 1981 and finishes in Q4 of 2020. All observations of GDP and MVE are expressed in Norwegian Crowns.

Table 18. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of Norway, 1981:Q1-2020:Q4.*

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	2.9%	3.0%	1.3%	0.5%	6.1%	65
0.2	0.4	2.5%	2.4%	0.8%	0.7%	4.2%	35
0.4	0.6	1.6%	1.6%	0.5%	0.8%	2.5%	19
0.6	-	0.1%	0.2%	0.4%	-0.3%	0.9%	9

As illustrated in Table 18, in the short-term sample of Norway both mean and median quarterly 8-year quarterly real returns monotonically decrease with higher BIs. Forward returns average 2.9% for BIs below 0.2 and 0.1% in the highest valuation bucket. While minimum returns do not show any clear trends w.r.t the BI, maximum returns and standard deviations monotonically decrease as the MVE increases relative to GDP.

Sweden

Short-term

The short-term sample for Sweden starts in Q2 of 1982 and ends in Q4 of 2020. All observations of GDP and MVE are expressed in Swedish Crowns.

Table 19. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of Sweden, 1982:Q2-2020:Q4.*

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	3.0%	2.5%	1.7%	0.6%	7.5%	44
0.2	0.4	3.0%	2.9%	1.1%	1.8%	4.8%	11
0.4	0.6	2.9%	3.0%	0.7%	1.9%	3.8%	11
0.6	0.8	2.2%	2.0%	0.6%	1.5%	3.3%	18
0.8	1	1.9%	1.9%	0.7%	0.1%	3.1%	29
1	1.2	0.4%	0.8%	1.1%	-0.9%	1.6%	10
1.2	-	-0.5%	-0.6%	0.6%	-0.9%	0.3%	4

Table 19 indicates that, for Sweden's short-term sample, mean 8-year quarterly real returns monotonically decline as the BI increases. The same holds for the median returns when the BI is higher than 0.2. Mean forward returns amount to 3.0% when MVE/GDP ratios are below 0.2 and to -0.5% when the BI is above 1.2. Standard deviations present a non-monotonic decline corresponding to an increase in MVE/GDP. Maximum returns monotonically decrease with an increase in the BI; the minimum returns' monotonic decline excludes the bucket with the lowest BIs.

Long-term

The observation period of Sweden's long-term sample goes from 1870 to 2019. In the single years with no CPI data, 1910 and 1924, MH assumed constant inflation from the last observed value. The effect of this approximation is likely not as impactful on inferences as an interpolation of GDP, but it could nonetheless lead to minor distortions of real returns in those 2 country-year observations.

Table 20. *Buffett indicator buckets and subsequent 10-year annual real returns of Sweden, 1870-2019.*

MVE/GDP BUCKET		**10-YEAR ANNUAL REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	7.9%	7.5%	4.1%	-1.3%	22.7%	60
0.2	0.4	6.2%	5.9%	3.8%	-0.3%	17.7%	48
0.4	0.6	-0.8%	-1.3%	10.0%	-14.5%	14.3%	15
0.6	-	5.9%	7.6%	7.2%	-14.9%	13.9%	17

Table 20 shows that mean and median 10-year annual forward returns tend to decrease as MVE/GDP increases. The standard deviations of returns are particularly high in the two highest valuation buckets. While maximum annual forward returns monotonically decrease as the Buffett Indicator increases, the declining trend for minimum forward returns is non-monotonic.

Switzerland

Short-term

The short-term sample for Switzerland goes from Q1 of 1973 to Q4 of 2020. All observations of GDP and MVE are expressed in Swiss Francs.

Table 21. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of Switzerland, 1973:Q1-2020:Q4 .*

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	1.8%	2.0%	1.1%	-1.1%	3.8%	50
0.2	0.4	3.1%	3.5%	1.7%	-1.2%	5.0%	31
0.4	0.8	2.5%	2.6%	0.8%	1.6%	3.9%	11
0.8	1.2	1.6%	1.6%	0.1%	1.6%	1.8%	4
1.2	-	1.2%	1.1%	0.8%	-0.8%	2.8%	63

Table 21 above shows that for the Swiss short-term sample, mean and median forward returns monotonically decrease with an increase in valuations, except for the lowest valuation bucket where the BI goes from 0 to 0.2. Standard deviations of returns, together with minimum and maximum forward returns, do not present any consistent trends based on valuation buckets.

Long-term

The long-term sample of Switzerland extends from 1900 to 2019. Data from 1900 to 1920 are from Lyndon Moore's "World Financial Markets 1900-1925" (Moore, 2012).

Table 22. *Buffett Indicator buckets and subsequent 10-year annual real returns of Switzerland, 1900-2019.*

MVE/GDP BUCKET		**10-YEAR ANNUAL REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	10.3%	10.7%	3.0%	5.0%	14.5%	8
0.2	0.4	4.1%	4.5%	5.5%	-8.3%	16.2%	47
0.4	0.6	3.6%	3.2%	5.9%	-8.6%	17.9%	25
0.6	0.8	7.5%	7.3%	6.1%	0.0%	14.5%	8
0.8	1	5.7%	5.6%	2.4%	2.9%	8.5%	4
1	-	4.5%	4.7%	3.6%	-0.9%	9.8%	18

The long-term sample of Switzerland presents no monotonic trends of forward returns relative to the BI. Mean and median returns monotonically decrease with an increase in MVE/GDP for 73% of observations, i.e. for buckets containing observations with a BI between 0 and 0.6. Neither standard deviations nor minimum and maximum returns display clear trends with respect to the MVE/GDP ratio.

United Kingdom

Short-term

The short-term sample for the UK spans from Q1 1973 to Q4 2020. All observations of GDP and MVE are expressed in Pounds Sterling.

Table 23. *Buffett Indicator buckets and subsequent 8-year quarterly real returns of the UK, 1973:Q1-2020:Q4.*

MVE/GDP BUCKET		**8-YEAR QUARTERLY REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	3.2%	3.3%	0.8%	1.2%	4.6%	31
0.2	0.4	0.8%	-0.2%	1.7%	-1.6%	3.8%	51
0.4	0.6	2.3%	2.6%	1.5%	-2.0%	3.7%	29
0.6	0.8	1.6%	1.6%	0.7%	0.7%	2.8%	14
0.8	1	1.3%	1.5%	0.5%	0.8%	1.9%	17
1	1.2	0.8%	0.8%	0.3%	0.2%	1.5%	27
1.2	-	0.3%	0.5%	0.5%	-0.8%	1.0%	23

In the short-term UK sample, as shown in Table 23, both mean and median forward real returns non-monotonically drop as the value of the BI soars. The mean of observed forward returns is 3.2% when the BI is less than 0.2 and 0.3% when the BI is above 1.2, While maximum returns monotonically fall with greater BIs, for minimum returns the lowest valuation bucket is the exception to an otherwise declining tendency.

Long-term

The long-term sample of the UK goes from 1871 to 2019. MVE data from 1871 to 1948 are GFD estimates. Data from Lyndon Moore's "World Financial Markets 1900-1925" (Moore, 2012) are used from 1900 to 1925. Estimates based on GFD's London Stock Exchange database are utilized from 1926 to 1938, and actual LSE values from the Stock Exchange Official Yearbook are used in 1933 and from 1939 onwards.

Table 24. *Buffett Indicator buckets and subsequent 10-year annual real returns of the United Kingdom, 1871-2019.*

MVE/GDP BUCKET		**10-YEAR ANNUALIZED REAL RETURNS**					
Low	**High**	**Mean**	**Median**	**St. Dev.**	**Min.**	**Max.**	**Obs.**
0	0.2	5.8%	5.8%	-	5.8%	5.8%	1
0.2	0.4	14.0%	13.0%	2.5%	11.2%	17.8%	6
0.4	0.6	8.3%	9.9%	6.4%	-0.6%	15.0%	7
0.6	0.8	7.3%	9.5%	5.1%	-1.3%	13.4%	19
0.8	1	5.6%	7.2%	4.5%	-5.4%	12.0%	33
1	1.2	5.7%	5.8%	1.5%	1.1%	9.2%	34
1.2	-	2.2%	2.6%	4.2%	-8.0%	9.4%	39

Table 24 above shows that, for the UK's long-term sample, both mean and median returns monotonically decrease with a higher BI, if we ignore the only outlier with a BI below 0.2

and an abnormally low forward return of 5.8%. Annual forward returns average 14.0% for MVE/GDP ratios between 0.2 and 0.4 and 2.2% when the BI is above 1.2. While maximum forward returns monotonically fall as the BI increases (with the outlier being an exception), there is no definite tendency by minimum returns relative to the BI.

4 Methodology

To quantify the predictive power of MVE/GDP with respect to forward returns and the relationship between these two variables, I follow Umlauft (2020) in defining a simple Ordinary Least Square (OLS) regression model of the following form:

$$r_h = \alpha_t + \beta_t * \frac{MVE}{GDP} + \varepsilon \quad (1)$$

In Equation (1), the dependent variable is r_h, the average forward return over 8-year and 10-year forecast horizons); α_t is the intercept; β_t is the regression coefficient for MVE/GDP, the Buffett Indicator observed at the beginning of the period used for computing the average forward return; finally, ε is a random error term. As this paper is interested in the explanatory power of MVE/GDP for forward returns, the focus of the regression analysis is on the R^2 or coefficient of determination. The R^2 measures the percentage of the dependent variable's variance explained by the model (Glantz & Slinker, 2001), giving a measure of how well the observed forward returns are reproduced by my univariate linear regression model. The Pearson correlation coefficient ("r") is used to measure the sign and magnitude of the linear relationship between the two continuous variables. This paper does not assume normality of distribution, considering the evidence on the insensitivity of the Pearson r to extreme violations of this assumption (Havlicek & Peterson, 1977; Nefzger & Drasgow, 1957). Interpretations of β_t, the regression coefficient of the regressor MVE/GDP, are not relevant for forecasting purposes and are thus not explored by this research paper.

To some extent, the predictive power of the BI was also examined in the data section: through descriptive tables that categorize observations in valuation buckets, i.e. groups of observations based on values of the MVE/GDP ratio, the data within both samples for each country were analyzed to obtain insight on the country-specific relationships between the indicator and forward real returns. The purpose of this division was to highlight time-independent relations between the Buffet Indicator and forward returns. To examine the coefficient of determination and the r through time, a separate OLS regression is performed every 8 or 10 years of observed BI, using the subsequent returns to calculate the forward return corresponding to every MVE/GDP observation. This creates a series of subperiods that have fixed start and end dates to allow cross-country comparisons of results. Each subperiod contains an independent subsample where the BI only uses the first half of observations, while average forward returns use also the second half. As previously mentioned, in the short-term analysis the geometric average of real returns observed during the previous 32 quarters is taken to compute each 8-year return. This means that, for example, the MVE/GDP ratio of 1980:Q4 is paired with the geometric average return from 1981:Q1 to 1988:Q4. This mechanism, similar to the rolling window regression, is used also for the long-term analysis where forward returns equal the 10-year average of returns following the observed BI. For instance, the forward return corresponding to the BI of 1993 equals the average of annual real returns from 1994 to 1993. This rolling computation creates a sort of overlapping structure, where every forward return in year t matches the forward return at year t-1 in all but one of its components. This overlap has two main consequences: the first is the autocorrelation of geometric forward returns, confirmed by the Ljung-Box test for time series and for all samples; the second is that the geometric forward returns' time series presents a feature that is normally not found in returns time series: non-stationarity. Through an Augmented Dickey-Fuller (ADF) tests, I determined that both the time series, of the Buffett Indicator and of the forward returns, have a unit root for the short-term and long-term samples of all countries. The data is transformed using first differencing to perform meaningful OLS regressions and to enable the testing of the significance of Pearson correlation coefficients between the two time series. After taking the first difference of both series and running the regressions again, the autocorrelation of the forward returns and the BI is removed from all samples according to Ljung-Box test and the Durbin-Watson test. After differencing, both time series are stationary for all samples according to the ADF test. With stationary time series, the R^2 estimated by the regressions are valid, and I can perform a t-test with regular standard errors to assess the Pearson correlation coefficient's significance. Given that the variables in this paper's model are in levels, and the transformation into first differences is just for estimation purposes, the results are interpreted using the levels model in (1).

5 Results

The findings of each regression analysis that are essential to this investigation are reported in Table 25, and in the tables 27 and 28. Each row these tables refers to a single OLS regression analysis that took as sample a subperiod of 16 years for the short-term analysis and of 20 years for the long-term analysis. Column 1 displays the dates of the first and the last value of the BI included in the subperiod, while column 2 reports the end of the period including the 8- or 10-year geometric returns that follow each observed BI. An inverse relationship between the Buffett Indicator and subsequent long-term returns can be observed in all 12 countries analyzed. However, the significance of this relationship between the two variables analyzed and the size of the BI's R^2 for forward returns varies between subperiods and from the short-term to the long-term analysis. Therefore, an overall assessment of the differences in results between subperiods and countries is necessary to provide an answer to the research question.

Short-term

In all short-term samples examined, a clear decreasing trend of mean forward returns is visible when sorting observations by valuation buckets. The descriptive tables in the Data section show that MVE/GDP ratios below 0.2 are always followed by larger mean and median 8-year real returns than the highest MVE/GDP ratios. In several countries, the lowest valuation bucket with BIs between 0 and 0.2 represents an exception to an otherwise monotonic decrease of mean and median forward returns. Nonetheless, this decrease is visible in the short-term sample of all countries. The results for the short-term sample of all countries validate the descriptive tables' hints pointing towards an inverse monotonic relationship. In fact, a high R^2 and a statistically significant and negative Pearson r between MVE/GDP and 8-year quarterly real returns are found across all countries analyzed and throughout nearly all subsamples. Additionally, a negative Pearson r significant at 1% is found for the overall short-term sample of all countries examined. Nevertheless, there are 8 subsamples across all countries in the short-term analysis where the BI performed poorly as a predictor of annualized real returns over the next 8 years. These are only 8 subsamples where the Pearson correlation coefficient is insignificant at the 10%, and they all present an R^2 smaller than 8%. Six of these subsamples contain observations from 1989:Q1 to 2004:Q4, while the remining two subsamples start in 1973:Q1 and end in 1988:Q4. Therefore, the only subperiod where we can see an insignificant correlation coefficient for many countries is the one starting in Q1 of 1989 and ending in Q4 of 2004. In fact, for Belgium, Canada, France, Germany, the

Netherlands, and Switzerland, the BI from 1989 to 1996 is not significantly linearly correlated to 8-year average real returns whose last observations are from 1997 to 2004. During this subperiod, as illustrated in Table 25, the R^2 of the BI is on average 22.89% across all 12 countries. As Table 26 reports, the BI is a particularly strong predictor of forward returns internationally during the last 23 years, from 1997:Q1 to 2020:Q4. Over the last two subperiods, the BI explains on average 49.75% and 62.94% across all countries. The mean correlation coefficient across all subperiods is highest for the UK and lowest for Germany. Alongside Switzerland, Germany is also the country with the lowest number of subperiods where the correlation coefficient is significant.

Table 25. *Short-term results by subperiod for each individual sample, 1973:Q1-2020:Q4.*

MVE/GDP Period	Forward Real Return	Australia		Belgium		Canada		France	
		r	**R^2**	**r**	**R^2**	**r**	**R^2**	**r**	**R^2**
1973:Q1-1980:Q4	1981:Q1-1988:Q4	-0.696***	48.37%	-	-	-0.464***	21.50%	-0.581***	33.75%
1981:Q1-1988:Q4	1989:Q1-1996:Q4	-0.912***	83.12%	-0.754***	56.84%	-0.899***	80.78%	-0.846***	71.63%
1989:Q1-1996:Q4	1997:Q1-2004:Q4	-0.607***	36.85%	-0.060	0.36%	-0.142	2.02%	0.003	0.00%
1997:Q1-2004:Q4	2005:Q1-2012:Q4	-0.455***	20.71%	-0.725***	52.55%	-0.651***	42.40%	-0.805***	64.79%
2005:Q1-2012:Q4	2013:Q1-2020:Q4	-0.800***	64.06%	-0.789***	62.27%	-0.803***	64.54%	-0.767***	58.76%
1973:Q1-2012:Q4	**1981:Q1-2020:Q4**	**-0.685**	**49.82%**	**-0.6324*****	**39.99%**	**-0.658 *****	**32.22 %**	**-0.688*****	**47.35%**

MVE/GDP Period	Forward Real Return	Germany		Italy		Japan		Netherlands	
		r	**R^2**	**r**	**R^2**	**r**	**R^2**	**r**	**R^2**
1973:Q1-1980:Q4	1981:Q1-1988:Q4	-0.271	7.32%	-0.519***	26.91%	-0.511***	26.11%	-	-
1981:Q1-1988:Q4	1989:Q1-1996:Q4	-0.726***	52.71%	-0.791***	62.56%	-0.524***	27.44%	-0.839***	70.45%
1989:Q1-1996:Q4	1997:Q1-2004:Q4	-0.192	3.68%	-0.497***	24.68%	-0.725***	52.57%	-0.262	6.84%
1997:Q1-2004:Q4	2005:Q1-2012:Q4	-0.754***	56.91%	-0.746***	55.68%	-0.686***	47.10%	-0.719***	51.62%
2005:Q1-2012:Q4	2013:Q1-2020:Q4	-0.768***	59.01%	-0.698***	48.75%	-0.786***	61.71%	-0.798***	63.60%
1973:Q1-2012:Q4	**1981:Q1-2020:Q4**	**-0.609*****	**37.10%**	**-0.658*****	**43.30%**	**-0.674*****	**45.40%**	**-0.694*****	**48.16%**

MVE/GDP Period	Forward Real Return	Norway		Sweden		Switzerland		UK	
		r	**R^2**	**r**	**R^2**	**r**	**R^2**	**r**	**R^2**
1973:Q1-1980:Q4	1981:Q1-1988:Q4	-	-	-	-	-0.249	6.18%	-0.805***	64.72%c-
1981:Q1-1988:Q4	1989:Q1-1996:Q4	-0.850***	72.19%	-0.393**	15.47%	-0.767***	58.80%	-0.804***	64.66%
1989:Q1-1996:Q4	1997:Q1-2004:Q4	-0.548***	29.98%	-0.667***	44.34%	-0.193	3.71%	-0.444**	19.72%
1997:Q1-2004:Q4	2005:Q1-2012:Q4	-0.519***	26.91%	-0.871***	75.93%	-0.846***	71.60%	-0.743***	55.18%
2005:Q1-2012:Q4	2013:Q1-2020:Q4	-0.875***	76.55%	-0.796***	63.32%	-0.791***	62.59%	-0.792***	62.64%
1973:Q1-2012:Q4	**1981:Q1-2020:Q4**	**-0.624*****	**56.67%**	**-0.667*****	**44.49%**	**-0.599*****	**35.84%**	**-0.750*****	**56.27%**

Notes: For each country, the leftmost column reports the Pearson's bivariate correlation coefficient (r), while the rightmost column displays the R^2 of the OLS regression. The values of r are rounded to 3 decimals, while the values of the R-squared are expressed as a percentage and rounded to 2 decimals.

*** = Significant at the 1% level; ** = Significant at the 5% level; * = Significant at the 10% level.

Table 26. *Short-term results by subperiod for all countries, 1973:Q1-2020:Q4.*

MVE/GDP Period	Forward Real Return	All samples	
		r	R^2
1973:Q1-1980:Q4	1981:Q1-1988:Q4	**-0.512**	**29.36%**
1981:Q1-1988:Q4	1989:Q1-1996:Q4	**-0.759**	**58.38%**
1989:Q1-1996:Q4	1997:Q1-2004:Q4	**-0.361**	**22.89%**
1997:Q1-2004:Q4	2005:Q1-2012:Q4	**-0.706**	**49.75%**
2005:Q1-2012:Q4	2013:Q1-2020:Q4	**-0.789**	**62.94%**

Notes: Column 3 and 4 display the average r and the average R^2 for each subperiod over all countries, respectively. The r is rounded to 3 decimals. The R^2 is expressed as a percentage and rounded to 2 decimals.

Long-term

Looking at the results for the long-term analysis reported in the tables 27 and 28 we see that for most subsamples analyzed, the bivariate Pearson correlation coefficient is not significant at the 10% level. This is likely due to the annual frequency of the data: 20-year subsamples have only 10 observations of forward annual returns, over 3 times less than the 32 observations of 8-year quarterly returns included in the subsamples of the short-term analysis. A low sample size could impede a correct rejection of the null hypothesis of no Pearson correlation between the two variables analyzed (David, 1938). For this reason, I have also computed the R-squared and the Pearson r using the entire observation period for each country. For Japan and the Netherlands, I used all observations from 1951 onwards, given that the Pearson r measures correlation between continuous variables and there is an interruption in the data before this year. The number of observations of 10-year returns in the entire sample of each country is several times higher than that of a single subsample. Moreover, the number of observations used for the latter regression analysis is always larger than 25, the minimum sample size recommended by David (1938) to utilize Pearson correlations. As the last row of both Table 27 and Table 28 shows, the bivariate Pearson r is negative for all countries' long-term sample. The correlation coefficient is significant at the 1% level for all countries except for Japan and Germany, where it is significant at the 5% and 10% level, respectively.

Despite the low frequency of the data, in the second half of the XX century tables 27 and 28 show at least one significant and negative average bivariate Pearson r in all countries analyzed. While from 1870 to 1950 there are only 6 subsamples with an r that is significant at 10% across all countries, the tables show 24 of these instances from 1950 to 2020. When we omit Belgium, whose sample starts in 1960, there are still 22 subsamples with significant negative correlation. Obviously, since we are dealing with a univariate regression, the explanatory power of the BI is also noticeably higher in the second half of the XX century. Australia presents 4 subsamples with a significant (at 5%) correlation coefficient from the 1960s to the 2020s, the most across all countries.

These results hinted at a possible overall insignificance – or lower significance - of the relationship between the two variables before the 1970s, "hidden" when performing regressions over the whole sample. Therefore, I ran further regressions to determine the reason why the short-term analysis' subsamples showed a widespread high significance and an explanatory power that is several times larger compared. Table A4 of the Appendix shows the results obtained by regressing the BI and subsequent 10-year returns before and after 1973, the start of the short-term sample. This created two subsamples for each country where annual data was continuous throughout the long-term sample. The significance of the Pearson r is reduced to 5% or 10% in all subsamples before 1972, and it is nullified for Germany. Furthermore, the magnitude of R^2 is several times lower in the subsample before 1972 than in the subsample corresponding to the short-term analysis. This demonstrates that in the long-term analysis, it was not just the frequency of the data to cause less significant correlations across countries, but also an actually weaker relationship between the BI and subsequent 10-year returns. Finally, across all countries' long-term samples, only the MVE/GDP ratio of the 1920s is positively correlated with subsequent 10-year returns, being the only subsample with a significantly positive correlation (α=1%).

Table 27. *Long-term results by subperiod for individual countries, 1871-2010.*

MVE/GDP Period	**Forward Real Return**	**Australia**		**Belgium**		**Germany**		**Japan**		**Netherlands**	
		r	**R^2**	**r**	**R^2**	**r**	**R^2**	**r**	**R^2**	**r**	**R^2**
1871-1880	1881-1890	-0.501	25.13%	-	-	-0.512	26.16%	-	-	-	-
1881-1890	1891-1900	0.133	1.77%	-	-	-0.153	2.34%	-	-	-	-
1891-1900	1901-1910	-0.661*	43.70%	-	-	0.467	21.76%	-	-	-	-
1901-1910	1911-1920	0.150	2.24%	-	-	0.421	17.70%	-0.419	17.57%	0.328	10.73%
1911-1920	1921-1930	-0.299	8.94%	-	-	-0.257	6.60%	-0.062	0.38%	0.292	8.54%
1921-1930	1931-1940	-0.015	0.02%	-	-	0.732**	53.63%	-0.065	0.42%	0.073	0.53%
1931-1940	1941-1950	-0.234	5.47%	-	-	0.050	0.25%	-	-	-	-
1941-1950	1951-1960	0.504	25.35%	-	-	-0.348	12.13%	-	-	-	-
1951-1960	1961-1970	-0.529	28.03%	-	-	-0.358	12.79%	-0.772**	59.65%	0.500	25.02%
1961-1970	1971-1980	-0.669**	44.75%	-0.476	22.69%	-0.846***	71.55%	0.267	7.15%	-0.425	18.04%
1971-1980	1981-1990	-0.830***	68.82%	-0.480	23.05%	-0.352	12.39%	-0.649*	42.06%	-0.439	19.24%
1981-1990	1991-2000	-0.784**	61.50%	-0.715**	51.13%	-0.776**	60.21%	-0.526	27.62%	-0.509	25.86%
1991-2000	2001-2010	-0.416	17.28%	-0.828***	68.48%	-0.093	0.86%	-0.706**	49.89%	0.187	3.49%
2001-2010	2011-2020	-0.948***	89.81%	-0.249	6.22%	-0.938***	87.95%	0.044	0.20%	-0.876***	76.71%

Observation period	**Australia**		**Belgium**		**Germany**		**Japan**		**Netherlands**	
	r	**R^2**	**r**	**R^2**	**r**	**R^2**	**r**	**R^2**	**r**	**R^2**
1871-2020	**-0.530*** **	**28.18%**	**-0.508*** **	**25.77%**	**-0.148* **	**2.18%**	**-0.238** **	**5.66%**	**-0.318*** **	**10.13%**

Table 28. *Long-term results by subperiod for individual countries, 1870-2009.*

MVE/GDP Period	Forward Real Return	France		Sweden		Switzerland		UK	
		r	R^2	r	R^2	r	R^2	r	R^2
1870-1879	1880-1889	0.481	23.13%	-0.560	31.39%	-	-	-0.626*	39.23%
1880-1889	1890-1899	-0.514	26.41%	0.223	4.96%	-	-	0.055	0.30%
1890-1899	1900-1909	0.234	5.48%	-0.342	11.67%	-	-	-0.553	30.60%
1900-1909	1910-1919	-0.483	23.37%	0.168	2.84%	-0.038	0.15%	-0.552	30.49%
1910-1919	1920-1929	0.258	6.67%	-0.731**	53.41%	-0.835***	69.79%	-0.166	2.75%
1920-1929	1930-1939	-0.492	24.20%	-0.150	2.24%	-0.075	0.57%	0.074	0.55%
1930-1939	1940-1949	-0.378	14.31%	-0.910***	82.84%	-0.135	1.82%	-0.557	0.3102
1940-1949	1950-1959	-0.013	0.02%	0.473	22.36%	-0.159	2.52%	-0.286	0.0816
1950-1959	1960-1969	-0.593*	35.16%	-0.543	29.45%	-0.275	7.56%	-0.458	20.95%
1960-1969	1970-1979	-0.180	3.24%	-0.679**	46.04%	0.203	4.11%	-0.443	19.60%
1970-1979	1980-1989	-0.574*	32.92%	-0.097	0.94%	0.064	0.41%	-0.937***	87.80%
1980-1989	1990-1999	-0.697**	48.53%	-0.750**	56.21%	-0.862***	74.27%	-0.623*	38.81%
1990-1999	2000-2009	-0.325	10.55%	0.133	1.76%	-0.338	11.42%	-0.361	13.02%
2000-2009	2010-2019	-0.885***	78.35%	-0.927***	85.95%	-0.657*	43.16%	-0.881***	77.61%

Observation period	France		Sweden		Switzerland		UK	
	r	R^2	r	R^2	r	R^2	r	R^2
1870-2019	**-0.295*** **	**8.72%**	**-0.465*** **	**21.61%**	**--0.396 *** **	**15.66%**	**-0.553*** **	**30.62%**

Notes for Table 27 and Table 28: For each country, the leftmost column reports the Pearson's bivariate correlation coefficient (r), the other displays the R^2 of the OLS regression. The last row displays the r and the R-squared computed through an OLS regression covering the whole observation period for each country. The values of r are rounded to 3 decimals, while the values of the R-squared are expressed as a percentage and rounded to 2 decimals.

*** = Significant at the 1% level; ** = Significant at the 5% level; * = Significant at the 10% level.

6 Discussion

In this section the results found in this research are discussed and connected to the existing literature. Furthermore, a discussion on the validity of these results for the consideration of the Buffett Indicator as a predictor of long-term returns is provided to give insight into the usefulness of this measurement of market valuations.

In the long-term analysis, the magnitude of R^2 and the Pearson r differs greatly between the first and the second half of the twentieth century. The main purpose of the subsample division in the long-term analysis was to verify whether the predictive power of the BI and its negative correlation with forward returns are observed consistently throughout the history of financial markets. As shown in the last rows of Table 27 and Table 28, the bivariate Pearson correlation coefficient is negative and significant (at least α=10%) over the whole long-term sample of all countries, although it is insignificant in most subsamples. However, the results obtained using the subsamples of the long-term analysis cannot be considered robust in terms of the correlation significance or of the R^2.

In fact, 25 is the minimum sample size recommended by David (1938) to utilize Pearson correlations, while the long-term analysis' subsamples have only 10 observations – 9 after differencing – because of the low frequency of annual data. This is a cause behind many subsamples recording an insignificant correlation coefficient in the long-term analysis. For example, if we run linear regressions using annual data and 8-year forward returns to compare the significance with those of Australia's short-term analysis, for the subsample going from 1981:Q1 to 1996:Q4 we get a Pearson r of -0.613. In the regressions on quarterly data subsamples with 32 observations, this coefficient would be significant at 1%, but here it is insignificant at 10% even if the estimated R^2 is 37.53%. It is thus evident that it would not be sensible to arrive at any definitive conclusions from annual data about the variation of the predictive power of the BI for forward returns based in specific decades. Nevertheless, when comparing annual data, the inverse relationship between the BI and 10-year returns is historically weaker in the XIX century and in the first half of the XX century, as table A4 shows. This result proves that while the BI has had a historically reliable explanatory power for long-term returns, the ratio's relationship with subsequent long-term returns has substantially increased in strength at the end of last century. Similarly, in the subsamples of the long-term analysis, the last 7 decades record the largest number of significant correlation coefficients over all countries. These results demonstrate a particularly high reliability of the Buffett Indicator as a predictor of 8-to-10-year returns since the Big Bang (Kuvshinov & Zimmermann, 2018). The results found by the short-term analysis complete a clear picture with respect to the consistency of the Buffett Indicator throughout the last 5 decades, from 1973 to 2020. Indeed, the Pearson correlation between the two variables under examination is significantly negative at the 1% level across most countries and subperiods, with the R^2 of the BI being relatively high. These results are especially robust for the last 23 years, 1997:Q1 to 2020:Q4. Here, in every subsample across all 12 countries analyzed the Pearson r is significant at the 1% level, and the BI always explains more than 20% of the 8-year returns' variance. Another interesting result is that the subperiod starting in Q1 of 1989 and ending in Q4 of 2004 is the only one where an insignificant correlation coefficient is persistent across the subsamples of many countries. In fact, for Belgium, Canada, France, Germany, the Netherlands, and Switzerland, the MVE/GDP ratio from 1989:Q1 to 1996:Q4 is not significantly linearly correlated to 8-year real returns computed from 1997:Q1 to 2004:Q4. It is worth noting that the late 1980s mark the beginning of dramatic upward shifts in the values of the equity market for most countries, as emphasized in the Data section of this paper. The descriptive tables have also illustrated that, for most countries, minimum forward equity returns tend to decrease as valuations increase. Column 3 of Table 29 below reports the average of minimum returns obtained for each valuation bucket across all countries' short-term samples. A decreasing trend is visible in this table that applies to all countries, suggesting an internationally comparable relationship between market valuation and risk exposure. However, more research is needed to assess this affiliation.

Table 29. *The relationship between MVE/GDP and 8-year minimum real returns in all countries' short-term samples.*

MVE/GDP BUCKET		
Low	**High**	**Minimum returns**
0	0.2	0.2%
0.2	0.4	0.3%
0.4	0.6	0.3%
0.6	0.8	0.1%
0.8	1	-0.2%
1	1.2	-0.5%
1.2	-	-1.3%

The short-term analysis of this paper found that that standard deviations of forward returns often decrease with higher valuation buckets, meaning that observations tend to be less widely scattered around their (lower) mean as the BI increases. This in turn indicates that a decrease of forward quarterly returns is increasingly predictable as the Buffett Indicator is higher, in accordance with the higher correlations found in recent decades. However, this phenomenon does not occur in all countries. In Sweden, for example, standard deviations of returns are particularly high in the two highest valuation buckets. When considering recent papers running similar analyses to predict long-term returns, the numeric results of this paper can be considered as relatively reliable. The method applied by Keimling (2016), while finding similar or higher R^2 values of the CAPE ratio for 17 MSCI countries, is less precise than this paper. In fact, his investigation evaluates the predictive power for average forward returns of "10-15 years depending on each country". Moreover, while he often obtains higher coefficients, his paper uses monthly data: as this paper has shown, a higher frequency of historical data makes it more likely to obtain higher coefficients by increasing the sample size for regressions. Moreover, his paper does not try to correct for autocorrelation or unit roots, thus potentially incurring partially spurious results. Umlauft (2020) experiences the same issue, as it does not account for autocorrelation and of non-stationarity while it assesses the explanatory power of the BI in the US. I replicated and analyzed the data used by the

Austrian researcher: both the US time series analyzed have a unit root and autocorrelation, implying that Umlauft (2020) found partially spurious results for the US market. Table A1 of the Appendix shows the results for the US where the two variables were transformed with first differencing before performing OLS. The correlation coefficients are still negative, on average smaller in magnitude but all significant at 1%. Nevertheless, the R^2 is still high and the previously mentioned conclusions reached by Umlauft with respect to the validity of the BI as a predictor of forward returns can still be deemed valid. OLS regressions for 8-year geometric quarterly returns of the US after first-differencing the time series produce an average R^2 of over 40% over the whole sample, with negative and statistically significant correlation coefficients for all subsamples, as Table A3 of the Appendix reports. The analysis of the US reveals that the relationship found by Umlauft (2020) between MVE/GDP and subsequent average returns are partially spurious. In fact, as shown in Table A1 of the Appendix, the coefficients are noticeably different than the ones in Umlauft (2020). Additionally, while the Austrian researcher does not test the significance of the Pearson r for the US, he finds a correlation between BIs of the 1980s and 10-year forward returns that is undoubtedly insignificant (-.033). On the other hand, after differencing both series, the overall r found by this paper for the same period measures -0.727 and is significant at 1%, as displayed in table A1.

7 Conclusion

Conclusive remarks on the research questions

This paper investigated whether there is a discernible relationship between the Buffett Indicator and subsequent long-term returns across developed countries, and it compared cross-country differences with regards to this relationship. Moreover, this paper asked whether the Buffett Indicator can be considered a valid valuation indicator to analyze the equity market and predict long-term returns in individual developed countries.

This paper has demonstrated a strong and statistically significant inverse relationship between the Buffett Indicator and subsequent long-term returns in the US and in other 12 developed economies. Even though there are some cross-country differences in the strength of this relationship, the BI historically explains a large portion of the variance of subsequent real returns for all the countries analyzed, averaging an especially high R^2 from 1973 to 2020 for both annual data and quarterly data. Additionally, the analysis of standard deviations through descriptive tables suggested that forward returns are more predictably low for higher observed values of the BI. This investigation thus validates for

12 individual developed countries the claims by Umlauft (2020) and by Kuvshinov & Zimmermann (2918) that the Buffett Indicator is a useful indicator to determine investor sentiment and mispricing of the stock market at any moment in time, just as Warren Buffett had suggested (Buffett & Loomis, 1999). Although part of the relationship estimated by Umlauft (2020) was found to be spurious because of autocorrelation and non-stationarity of the “rolling 10-year returns”, this paper has proven that the results by Umlauft (2020) for the US hold internationally across some of the largest developed economies in the world. While its relationship with forward returns was weaker over the XIX century and the first half of 1900, the Buffett Indicator showed strong and consistent predictive power for long-term returns in all the last 5 decades, probably due to the fact that it seemingly mean-reverts to different time-varying long-term “equilibrium” levels over the last 150 years. The dramatic shift in the BI’s “equilibrium” level of the last 40 years has not hindered the efficacy of the ratio, which instead looks as relevant as ever for subsequent equity returns. Since investment decisions ought not be made only following a single statistic, the Buffett Indicator appears to be a valid supplementary metric for current stock market valuation and the formulation of long-term return prospects. Because it adds significant non-earnings-based information, MVE/GDP can complement other long-term returns indicators such as the CAPE ratio.

Limitations and suggestions for future research

One limitation of this paper relates to foreign stocks. For all market capitalization data, the share of foreign stocks is either not specified in the source - in recent data – or insignificant, as for most countries data in the mid-20th century during the period of low stock market development (Volosovych, 2011). Therefore, this limitation concerns international stock exchanges in the early years of the long-term sample, in particular the UK’s stock exchange, and all other stock exchanges from the 1970s-1980s onwards. The data sources do not provide information on foreign stocks, so it is impossible to determine whether this issue leads this investigation to understate or overstate domestic stock market capitalization. Additionally, this research is limited in that it does not control, as Kuvshinov and Zimmermann (2018) do at the aggregate level, for credit growth and for dividends. However, since this paper is concerned with the explanatory power of the MVE/GDP ratio for predictive purposes, omitted variables are negligible.

Outliers represent a potential limitation of this research paper as they can significantly distort data. When removing them, one can significantly alter what the data is showing. However, given that the data used in this paper was compiled by reputable sources, I assumed that the outliers of these datasets have not occurred because of a sampling error.

This implies that their removal would compromise the validity of the results. The low frequency of data prevents a precise assessment via subsamples of the historical consistency of the correlation between MVE/GDP and subsequent 10-year equity returns. This limitation was already considered in the Discussion section: this issue would probably be solved only by improving the frequency of historical data. The feasibility of this solution is however questionable, as such improvement would be attainable through further assumption-based estimations that could cause large distortions of results. Another potential limitation that is hereunder analyzed concerns the choice of "benchmark years" selected for every subperiod of the analysis, i.e. the first and the last quarter (or year) used in each regression analysis. In fact, when keeping the length of the forward returns period constant, it is intuitive that results might change as the subperiod is shifted e.g. by two years. An example is thus analyzed hereunder to discuss the impact of this arbitrary choice on results. As Table A2 of the appendix shows, shifting the start of each subperiod backwards by 2 years only increases the average explanatory power of the BI for 8-year real returns from 53.38%% to 55.43% over the UK sample. Moreover, the significance and sign of the correlation coefficients are not altered. This issue thus appears to be irrelevant for the accuracy and the validity of results found in this paper.
An additional practical constraint for the implementation of this paper is that the substantial delay with which the data necessary for calculating the MVE/GDP ratio is published is a practical constraint. However, the lagged publications of the data have a negligible practical impact due to the long-term nature of the Buffett Indicator. Moreover, if timely snapshots were needed, investors could approximate MVE/GDP almost in real-time via large-cap market indexes and GDP “nowcasts” instead of retrospective publications. Finally, some interesting insight for future research is provided by the fact that for Belgium, Canada, France, Germany, the Netherlands, and Switzerland, the bivariate Pearson correlation between the BI and subsequent 8-year real returns is insignificant for the subperiod starting in 1989:Q1 and ending in 2004:Q4. What caused this result has not been thoroughly investigated by this paper, although the visual analysis of the BI time plots suggest that the initial shock of the Big Bang (Kuvshinov & Zimmermann, 2018) could be the cause behind the independence of the two time series. Therefore, further investigation of the contributory causes this phenomenon is advised, together with an analysis of the Buffett Indicator’s relationship with subsequent returns in emerging countries.

8 References

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *The Journal of Finance*, *68*(3), 929-985.

Baele, L. (2005). Volatility spillover effects in European equity markets. *Journal of Financial and Quantitative Analysis*, *40*(2), 373-401.

BlackRock. (n.d.). *BlackRock return Map: Blackrock Investment Institute*. BlackRock. https://www.blackrock.com/corporate/insights/blackrock-investment-institute/interactive-charts/return-map.

Campbell, J. Y., & Shiller, R. J. (2001). *Valuation ratios and the long-run stock market outlook: An update* (No. w8221). National bureau of economic research.

David, F. N. (1938). *Tables of the ordinates and probability integral of the distribution of the correlation coefficient in small samples*. Cambridge University Press.

Eisenhammer, J., & Rhodes, M. (2012). The politics of public sector steel in Italy: From the 'economic miracle' to the crisis of the eighties. In *The Politics of Steel* (pp. 416-475). De Gruyter.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of financial economics*, *22*(1), 3-25.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, *47*(2), 427-465.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, *115*(27), E6106-E6115.

Glantz, S. A., & Slinker, B. K. (2001). *Primer of Applied Regression & Analysis of Variance, ed*. McGraw-Hill, Inc., New York.

Graham, B., Dodd, D. L., & Cottle, S. (1934). Security analysis: Principles and technique.

Havlicek, L. L., & Peterson, N. L. (1977). Effect of the violation of assumptions upon significance levels of the Pearson r. *Psychological Bulletin*, *84*(2), 373.

Johnson, J. S. (1935). History and organization of the Montreal Stock Exchange.

Jordà, Ò., Knoll K., Kuvshinov, D., Schularick, M., & Taylor, A.M. (2019). The Rate of Return on Everything, 1870–2015. *Quarterly Journal of Economics. 134* (2019), 1225–1298. Retrieved on April 16, 2021 from https://www.macrohistory.net/

Jordà, Ò., Schularick, M., & Taylor, A.M. (2017). Macrofinancial History and the New Business Cycle Facts. *NBER Macroeconomics Annual 2016, volume 31*, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press. Retrieved on April 16, 2021 from https://www.macrohistory.net/

Jorgensen, B. N., Kavajecz, K. A., & Swisher, S. (2011). The Historical Evolution of Financial Exchanges. *available on http://papers. ssrn. com/sol3/papers. cfm*.

London Stock Exchange (1992). *Stock Exchange Official Yearbook, 1991-1992.* Macmillan Pub., under license from the International Stock Exchange.

Keimling, N. (2016). Predicting Stock Market Returns Using the Shiller CAPE—An Improvement Towards Traditional Value Indicators?. *Available at SSRN 2736423*.

Kiljunen, K. (2004). *The European Constitution in the Making*. Centre for European Policy Studies. pp. 21–26. ISBN 978-92-9079-493-6.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The journal of finance*, *49*(5), 1541-1578.

Lenza, M., Reichlin, L., & Giannone, D. (2009). *Business cycles in the euro area*. Centre for Economic Policy Research.

Nefzger, M. D., & Drasgow, J. (1957). The needless assumption of normality in Pearson's r. *American Psychologist*, *12*(10), 623.

Molodovsky, N. (1953). A theory of price-earnings ratios. *Financial Analysts Journal*, *9*(5), 65-80.

Moore, L. (2012). World Financial Markets, 1900-1925. Unpublished Manuscript, University of Melbourne.

Nefzger, M. D., & Drasgow, J. (1957). The needless assumption of normality in Pearson's r. *American Psychologist*, *12*(10), 623.

Ofek, E., & Richardson, M. (2003). Dotcom mania: The rise and fall of internet stock prices. *The Journal of Finance*, *58*(3), 1113-1137.

Rajan, R. G., & Zingales, L. (2003). The great reversals: the politics of financial development in the twentieth century. *Journal of financial economics*, *69*(1), 5-50.

Statistisches Reichsamt (2016). *Statistisches jahrbuch für das deutsche reich*. Hansebooks GmbH. Retrieved on May 12, 2021 from https://finaeon.globalfinancialdata.com/

Sylla, R., Wilson, J. W., & Wright, R. E. (2006). Trans-Atlantic Capital Market Integration, 1790–1845. *Review of Finance*, *10*(4), 613-644.

Tokyo Stock Exchange (various issues). *Tokyo Stock Exchange Fact Book*. Retrieved on May 12, 2021 from https://finaeon.globalfinancialdata.com/

Umlauft, T. (2020) "The market value of equity-to-gross domestic product ratio as a predictor of long-term equity returns: evidence from the U.S. market, 1951-2019", *SSRN working paper*

Volosovych, V. (2011). Measuring financial market integration over the long run: Is there a U-shape?. *Journal of International Money and Finance, 30(7),* 1535-1561.

Wright, W., & Bax, L. (2016). What do EU capital markets look like post-Brexit. *New Financial, September* 2016.

Wilcox, J. W. (1984). The P/B-roe valuation model. *Financial Analysts Journal*, *40*(1), 58-66.

World Federation of Exchanges (various issues), *Annual Report and Statistics*.

9 Appendix

Table A1. Results for the US: 10-year average returns, 1951:Q1 - 2020:Q4.

MVE/GDP Period	Forward Return	Correlation Coefficient	R^2
1951:Q4-1959:Q4	1960:Q1-1969:Q4	-0.634***	42.89%
1960:Q1-1969:Q4	1970:Q1-1979:Q4	-0.559***	31.29%
1970:Q1-1979:Q4	1980:Q1-1989:Q4	-0.682***	46.51%
1980:Q1-1989:Q4	1990:Q1-1999:Q4	-0.727***	52.81%
1990:Q1-1999:Q4	2000:Q1-2008:Q4	-0.599***	35.99%
2000:Q1-2009:Q4	2010:Q1-2019:Q4	-0.692***	52.78%
1951:Q1-2009:Q4	**1960:Q1:2019:Q4**	**-0.649**	**43.71%**

Table A2. *Short-term results for the UK: forward returns regressed on ln(MVE/GDP) based on subperiods, 1971:Q1 - 2018:Q4.*

MVE/GDP Period	Forward Return	Correlation Coefficient	R^2
1971:Q1-1974:Q4	1979:Q1-1986:Q4	-0.885***	78.25%
1979:Q1-1986:Q4	1987:Q1-1993:Q4	-0.490***	23.96%
1987:Q1-1993:Q4	1994:Q1-2001:Q4	-0.699***	48.88%
1994:Q1-2001:Q4	2002:Q1-2009:Q4	-0.753***	56.68%
2002:Q1-2009:Q4	2010:Q1-2018:Q4	-0.833***	69.37%
1971:Q1-2009:Q4	**1979:Q4-2018:Q4**	**-0.732**	**55.43%**

Notes: *** = significant at the 1% level.

Table A3. R*esults for the US using 8-year average returns, 1957:Q1 - 2020:Q4.*

MVE/GDP Period	Forward Real Return	Correlation	R^2
1957:Q1-1964:Q4	1965:Q1-1972:Q4	-0.440**	19.37%
1965:Q1-1972:Q4	1973:Q1-1980:Q4	-0.608***	36.93%
1973:Q1-1980:Q4	1981:Q1-1988:Q4	-0.748***	55.92%
1981:Q1-1988:Q4	1989:Q1-1995:Q4	-0.849***	72.09%
1989:Q1-1995:Q4	1996:Q1-2003:Q4	-0.297*	8.83%
1996:Q1-2003:Q4	2004:Q1-2011:Q4	-0.429***	18.41%
2004:Q1-2011:Q4	2012:Q1-2020:Q4	-0.809***	65.50%
1957:Q1-2011:Q4	**1965:Q4-2020:Q4**	**-0.647*** **	**41.89%**

Table A4. *Comparison of results for annual data before and after 1973.*

Country	Observation period	Results		Observation period	Results	
		r	R^2		r	R^2
Australia	1870-1972	-0.232**	5.39%	1973-2019	-0.739***	54.66%
Belgium	1961-1972	-	-	-		
France	1870-1972	-0.184*	3.37%	1973-2019	-0.654***	42.81%
Germany	1872-1972	-0.095	0.89%		-0.581***	33.78%
Italy	-	-	-	-	-	-
Japan	-	-	-	1973-2020	-	-
Netherlands	-	-	-	1973-2020	-	-
Norway	-	-	-	-	-	-
Sweden	1871-1972	-0.365***	13.32%	1973-2020	-0.617***	38.02%
Switzerland	1900-1972	-0.378***	14.30%	1973-2019	-0.637***	40.55%
UK	1871-1972	-0.261**	6.83%	1973-2020	-0.8486***	72.01%

Notes: * = significant at the 10% level;** = significant at the 5% level; *** = significant at the 1% level.

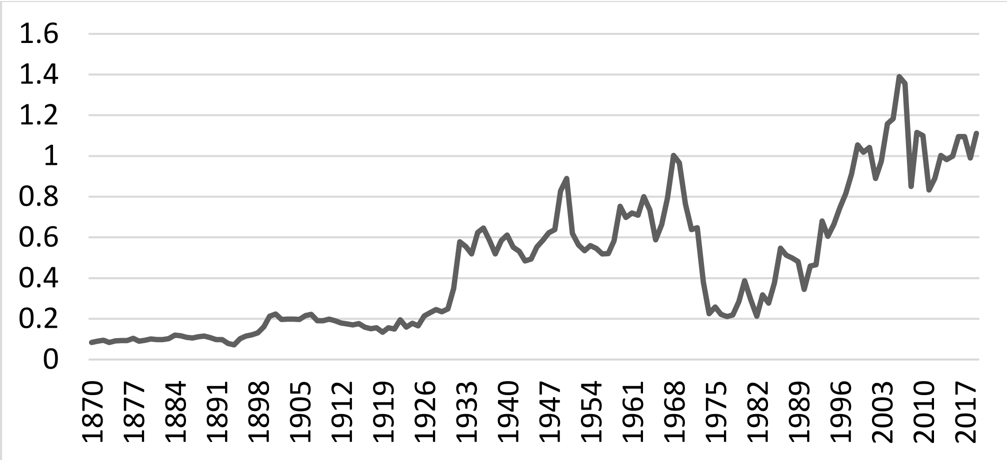

Figure A1. *Time plot of the annual MVE/GDP ratio of Australia, 1871-2020.*

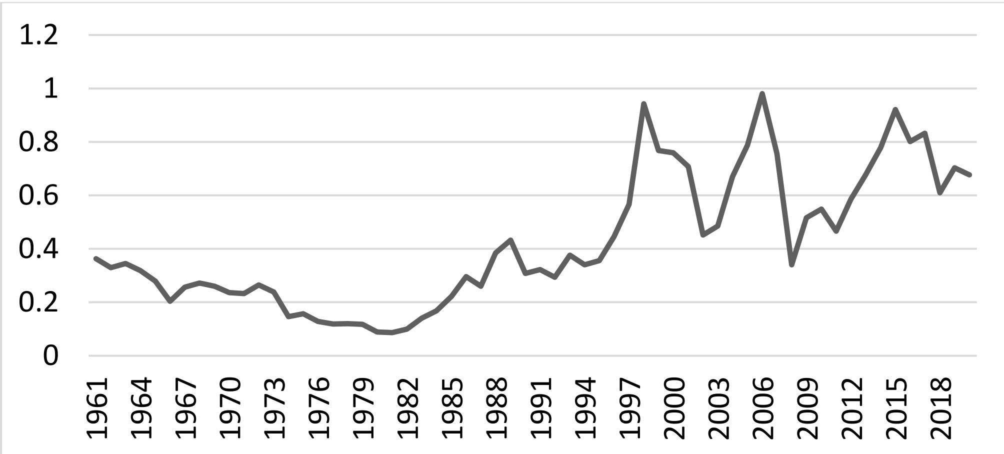

Figure A2. *Time plot of the annual MVE/GDP ratio of Belgium, 1961-2020.*

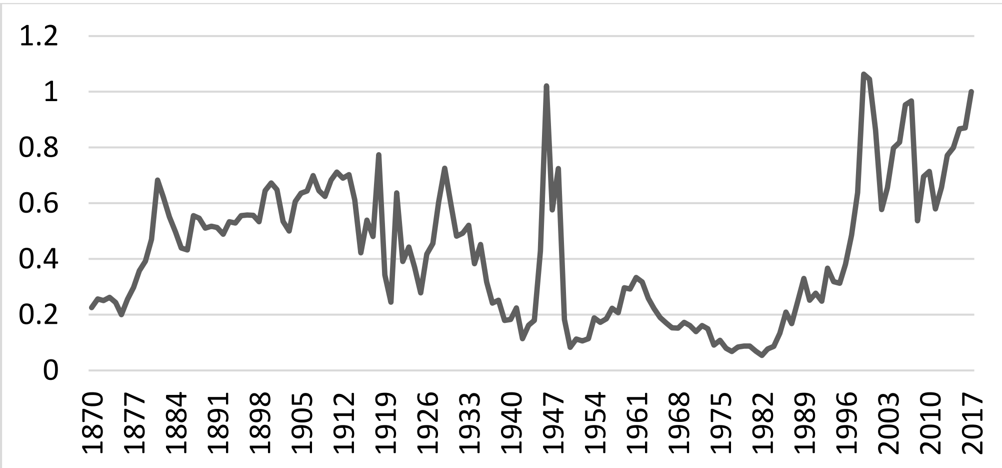

Figure A3. *Time plot of the annual MVE/GDP ratio of France, 1870-2019.*

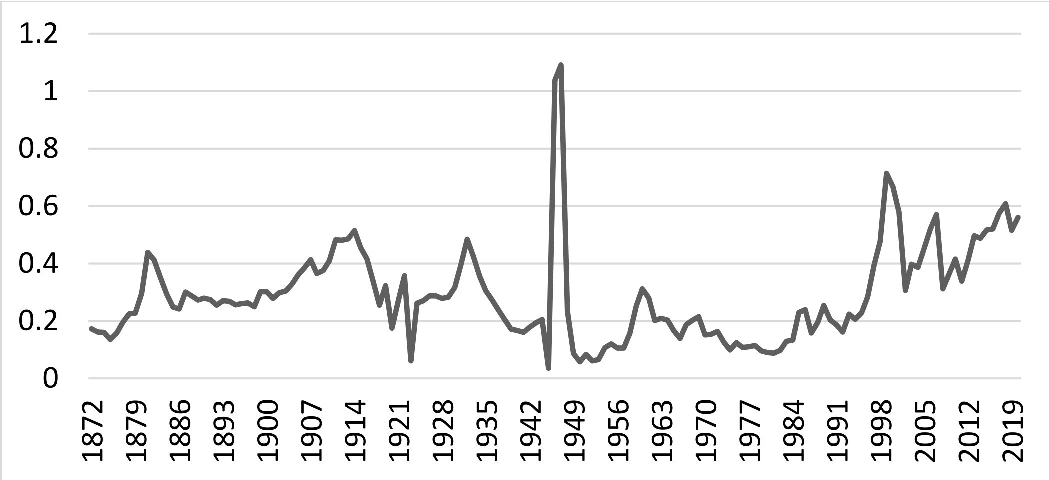

Figure A4. *Time plot of the annual MVE/GDP ratio of Germany, 1872-2020.*

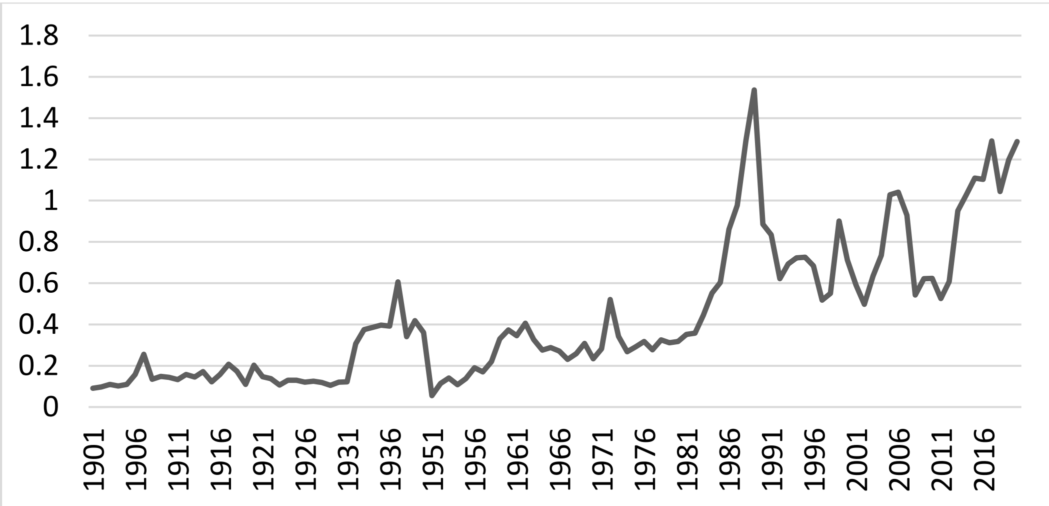

Figure A5. *Time plot of the annual MVE/GDP ratio of Japan, 1901-2020.*

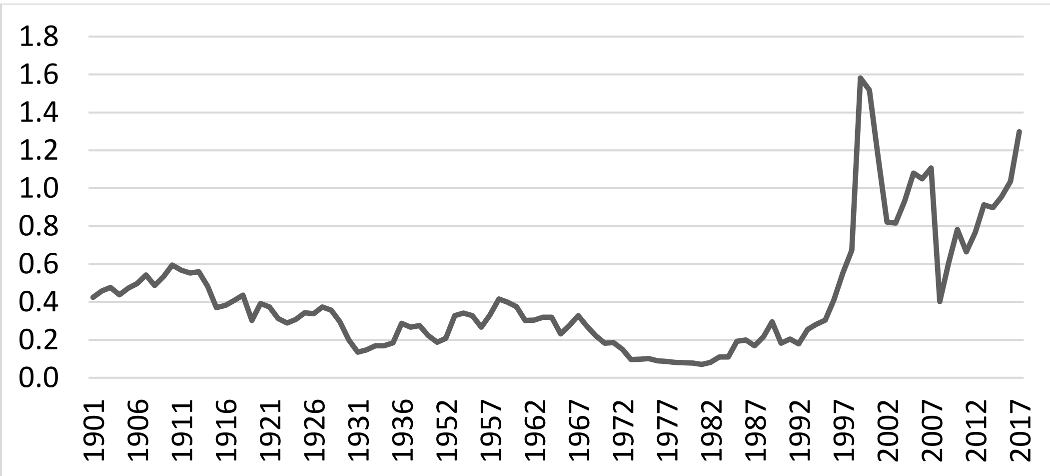

Figure A6. *Time plot of the annual MVE/GDP ratio of the Netherlands, 1901-2020.*

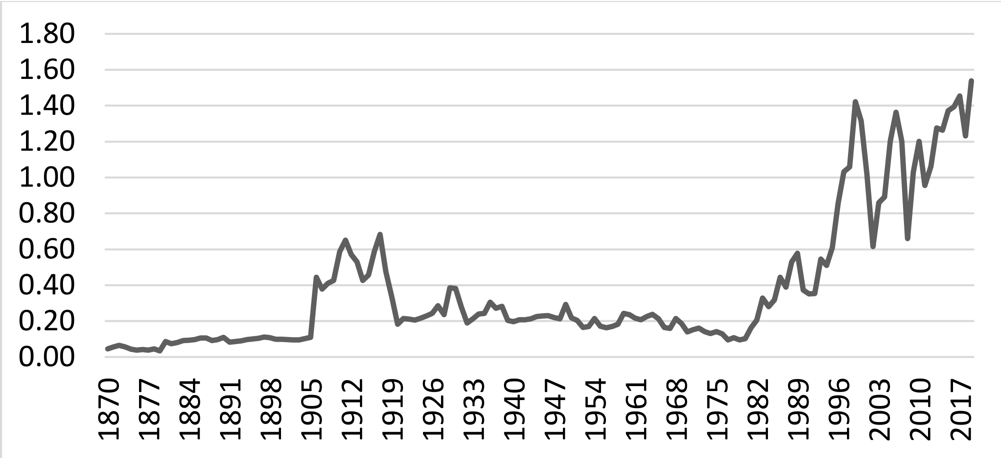

Figure A7. *Time plot of the annual MVE/GDP ratio of Sweden, 1870-2019.*

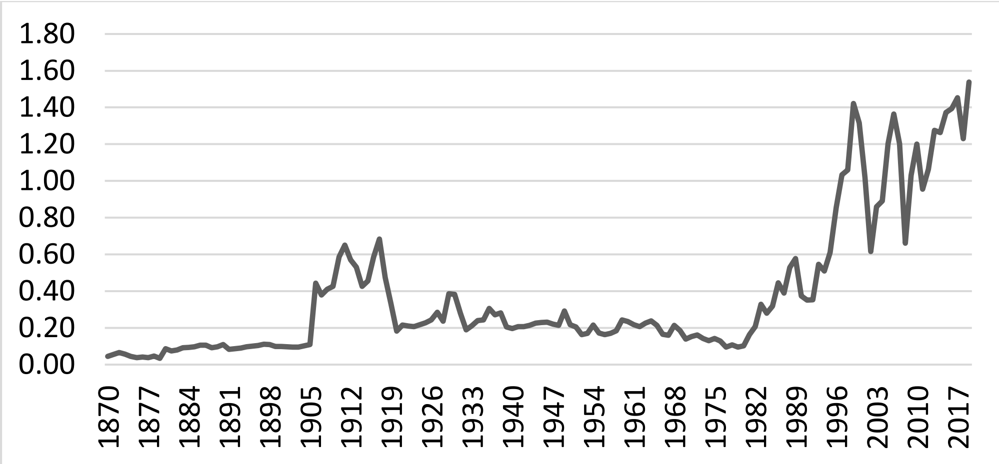

Figure A8. *Time plot of the annual MVE/GDP ratio of Switzerland, 1870-2019.*

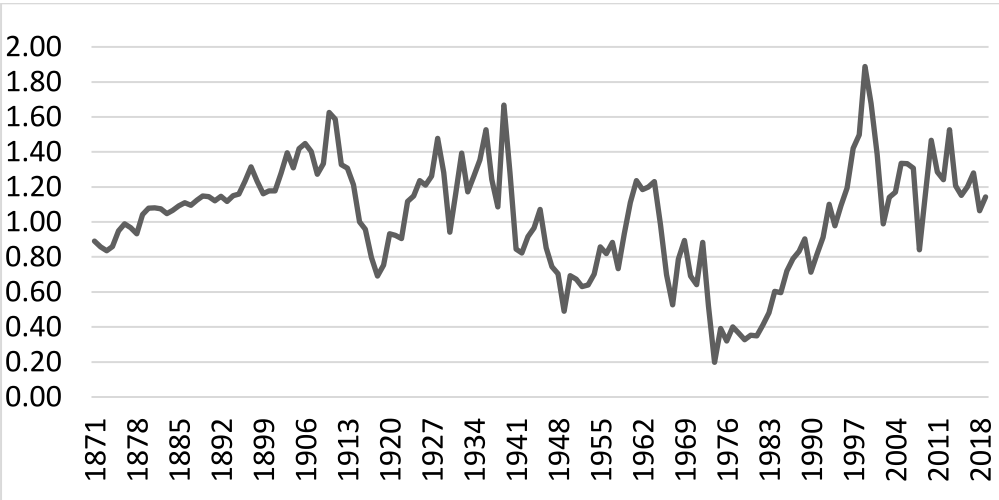

Figure A9. *Time plot of the annual MVE/GDP ratio of the UK, 1871-2019.*