QUANTITATIVE FINANCE RESEARCH

QUANTITATIVE FINANCE RESEARCH

Research Paper 386 August 2017

Market Efficiency and the Growth Optimal Portfolio

Eckhard Platen and Renata Rendek

ISSN 1441-8010 www.qfrc.uts.edu.au

Market Efficiency and the Growth Optimal Portfolio

Eckhard Platen[1,2,3] and **Renata Rendek**[4]

August 8, 2017

Abstract

The paper predicts an Efficient Market Property for the equity market, where stocks, when denominated in units of the growth optimal portfolio (GP), have zero instantaneous expected returns. Well-diversified equity portfolios are shown to approximate the GP, which explains the well-documented good performance of equally weighted portfolios. Our proposed hierarchically weighted index (HWI) is shown to be an even better proxy of the GP. It sets weights equal within industrial and geographical groupings of stocks. When using the HWI as a proxy for the GP the Efficient Market Property cannot be easily rejected and appears to be robust.

JEL Classification: G10, G11

Key words and phrases: efficient market property, growth optimal portfolio, long term growth rate, naive diversification, benchmark pricing theory, hierarchical diversification.

Acknowledgements: The authors like to express their gratitude for receiving valuable suggestions on the research as it evolved by Yacine Ait-Sahalia, Jerome Detemple, Harry Markowitz, Jean-Charles Rochet, Johan Walden and many colleagues and friends, as well as, participants of workshops and conferences, including the World Finance Conference in Cyprus, Bachelier Finance Society World Congresses and the Quantitative Methods in Finance conference series.

[1]University of Technology Sydney, Finance Discipline Group
[2]University of Technology Sydney, School of Mathematical and Physical Sciences
[3]University of Cape Town, Department of Actuarial Science
[4]UNSW Sydney, School of Mathematics and Statistics,
Corresponding author: Eckhard Platen (eckhard.platen@uts.edu.au)

1 Introduction

For many decades the question whether markets are 'efficient', e.g., as described in Fama (1970, 1991, 1998), has been widely discussed but never conclusively answered in the literature. The current paper aims to answer this question for developed equity markets. This answer is deeply linked to the respective growth optimal portfolio (GP). We show, when using the GP (also called benchmark) as a denominator, the benchmarked (denominated in units of the GP) value of any portfolio has zero instantaneous expected return. Thus, its current benchmarked value is the best forecast of its instantaneously next benchmarked value. We call this fundamental fact the *Efficient Market Property*. It resembles various types of efficient market hypotheses discussed in Fama (1970, 1991, 1998) and a subsequent rich stream of literature. We emphasize that this is an objectively given fact which holds true under extremely general model assumptions. The GP and, thus, the *Efficient Market Property* are unique for a given investment universe and the respective available evolving information. This property is robust, since we do not need to involve any particular model assumptions or estimation for constructing an excellent proxy for the GP for developed stock markets. We use this proxy as a benchmark and show for equities in developed markets that the Efficient Market Property cannot be easily rejected. The findings provide a new understanding of market efficiency, which theoretically exists for any reasonable market that has a GP. When the GP does not exist for a market model then its candidate explodes in finite time, which means that there is economically meaningful arbitrage in such a market and the model has to be dismissed.

The GP is the expected log-utility maximizing portfolio. It was discovered in Kelly (1956) and has been widely studied; see MacLean et al. (2011) for an edited collection of papers. It has the fascinating property that it maximizes pathwise in the long run the long-term growth rate (GR). The GR is the logarithm of the portfolio value normalized by the length of the observation window.

The key question we answer in the current paper is how to construct an excellent proxy of the GP. Despite decades of research on how to construct optimal portfolios for stocks, the simple, model-independent, equally-weighted approach seems to do at least as well as more complicated and theoretically grounded approaches. This stylized empirical fact has been established in DeMiguel et al. (2009), where it has been shown that naive diversification (equal-weighting of stocks) outperforms most known portfolio strategies. The current paper goes beyond naive diversification and makes use of readily available information, capturing key economic dependencies of stocks. This information does not involve any estimation and is obtained through classification of economic activities of respective companies. Companies which belong to the same industrial and geographical group are exposed to similar uncertainties and their own specific uncertainty. The hierarchical groupings provided by such classification remain quite stable over time and persist in periods of extreme market moves.

By naive diversification within each group and at each level of the hierarchy a new proxy for the GP is constructed, which we call the hierarchically weighted index (HWI). To illustrate the excellent performance of the HWI for stocks of developed markets we show in Figure 1.1 its trajectory, together with those of the respective market capitalization weighted index (MCI), and the equally-weighted index (EWI) in US dollar denomination. All three indexes are normalized such that they start in 1984 at an initial value of one. In Figure 1.2 we display the HWI's and the EWI's observed (annualized) long term growth rates (GRs) in percent as functions of the end date of the observation period. For the longest available observation window the long-term growth rate (GR) amounts for the EWI to 11.6 and for the HWI to 14.5, whereas the GR for the MCI reaches only 9.4; see also Table 4.1. Theoretically, it is the GP that achieves pathwise the highest GR in the long-run. This makes the GR for the longest available observation window a realistic performance measure to distinguish between proxies of the GP.

By using the HWI as a benchmark and then studying more than 30 million benchmarked stock returns we show that the theoretically predicted *Efficient Market Property* cannot be easily rejected for stocks in developed markets. When choosing the MCI or EWI as a benchmark, the Efficient Market Property can be easily rejected at typical significance levels.

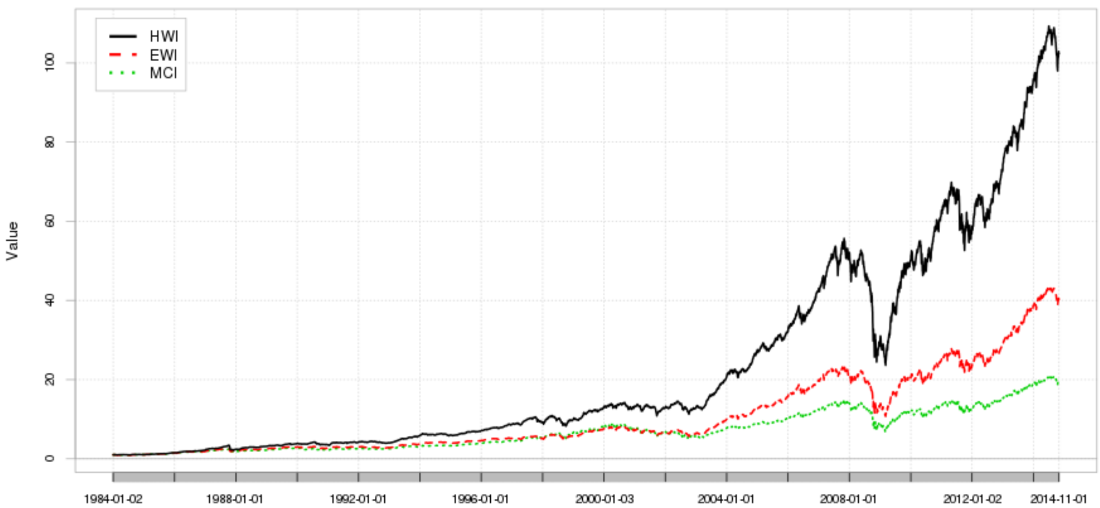

Figure 1.1: The trajectories of the HWI, EWI and MCI.

With the HWI we also propose a new benchmark for long-term equity fund management. It is far more difficult to outperform in the long run than the traditional benchmark, the MCI. Moreover, we propose with the HWI a practically feasible approximation of a theoretically optimal portfolio, the GP. Once the GP is constructed, it is straightforward to obtain other optimal portfolios with desired risk characteristics by involving the riskless asset in an appropriate manner; see e.g. Chapter 11 in Platen and Heath (2010). This allows one

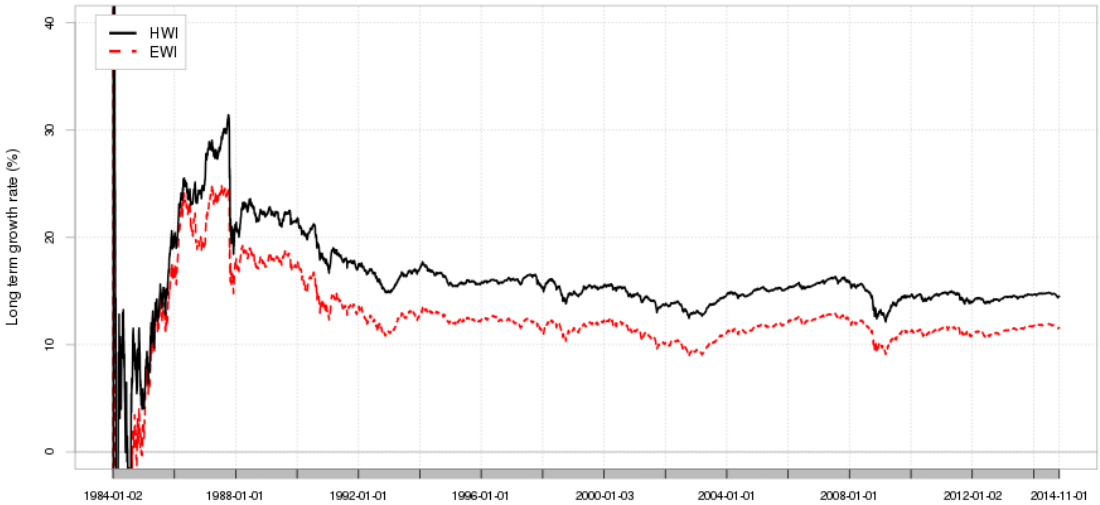

Figure 1.2: Observed long term-growth rate (GR) (in percent) with dependence on the final observation date for the HWI and the EWI.

to overcome the main practical obstacle in portfolio optimization, identified in DeMiguel, Garlappi & Uppal (2009). It results from the fact that the standard sample-based mean-variance methodology of modern portfolio theory, originated by Markowitz (1959), is not, in general, accurate enough to provide useful proxies for targeted optimal portfolios. The impossibility of implementing sample-based portfolio optimization for large equity markets has also been noted, e.g., in Best & Grauer (1991), Chopra & Ziemba (1993), Bai & Ng (2002), Ludvigson & Ng (2007), Plyakha et al. (2014), Kan & Zhou (2007), Kan, Wang & Zhou (2016) and Okhrin & Schmid (2006). The dilemma is that the available observation windows are too short for estimating the most likely moving expected returns.

Building on Markowitz's mean-variance approach, Sharpe (1964) introduced the capital asset pricing model (CAPM), which became generalized in a stream of literature. A consequence of the classical assumptions underpinning the CAPM is that the market capitalization weighted index (MCI) should maximize the Sharpe ratio. However, empirical evidence suggests that the MCI may, in reality, not yield the highest Sharpe ratio; see e.g. Harvey, Liu & Zhu (2016). This fact is also confirmed by our observations, where we report in Table 4.1 a Sharpe ratio of 0.90 for the HWI, which is more than double that of 0.43 for the MCI.

Since market capitalization weighted indexes seem unlikely to be mean-variance optimal, a wide range of rule-based investment strategies has emerged, aiming mostly for higher Sharpe ratios. One group of rule-based strategies uses stock characteristics, known as factors, which have been found in historical data to have an effect on the expected returns of stocks; see e.g. Rosenberg & Marathe (1976), Fama & French (1992), Carhart (1997) and Arnott et al. (2005). If investors systematically exploit such factors on a large scale, e.g. fundamental value, size or momentum, then most of these effects are likely to weaken or even vanish over time, as argued in Van Dijk (2011). Harvey et al. (2016) find that some 'observed' factor premia resulted from just 'mining' the available finite data set and should be dismissed. In summary, factor-based strategies may not provide sustained high Sharpe ratios in the long run.

Several rule-based strategies have emerged which exploit the inverse of estimated covariance matrices of returns. These strategies aim at, e.g., the minimum variance portfolio, as in Clarke et al. (2011); the risk parity portfolio, as in Maillard et al. (2010); or the maximum diversification portfolio, as in Choueifaty & Coignard (2008). As revealed in Plyakha et al. (2014), even for a rather small number of stocks the arising estimation errors may easily offset the benefits of such theoretically optimal portfolio strategies.

Despite an abundance of empirical and theoretical work; see e.g. Fernholz (2002), Chow et al. (2011), Leote et al. (2012), Gander et al. (2013) and Oderda (2015), no authors seem to have managed to extract convincingly the theoretical reason why various rule-based strategies outperform the respective MCI. The current paper reveals this reason through its *Diversification Theorem*, which states that well-diversified portfolios asymptotically approximate the GP for increasing number of constituents. Rule based portfolios are usually better diversified than the MCI.

The paper is organized as follows: Section 2 summarizes several facts about the GP and the Efficient Market Property. In Section 3 we construct the HWI. Section 4 analyzes the performance of the HWI, while Section 5 empirically explores the Efficient Market Property. Appendix A describes the data used, whereas Appendix B proves the Efficient Market Property. Appendix C presents the Diversification Theorem. Finally, Appendix D demonstrates that the HWI equals the GP under a stylized hierarchical stock market model.

2 Efficient Market Property

In this section we summarize several properties related to the *growth optimal portfolio* (GP), including the Efficient Market Property.

Let V_t^π denote the value of a strictly positive portfolio at time t with strategy π, which in turn denotes the weights invested. We can then write the respective *long-term growth rate* (GR) at time $T > 0$ in the form

$$G_T^\pi = \frac{1}{T} \ln \left(\frac{V_T^\pi}{V_0^\pi} \right). \tag{2.1}$$

The value of the GP, at time $t \geq 0$ is denoted by $V_t^{\pi^*}$, where we set $V_0^\pi = 1$. The GP is the portfolio that maximizes the expected GR, that is, $E(G_T^\pi)$ for all $T \geq 0$. We assume that $G_T^{\pi^*} < \infty$ for all $T > 0$, and have the following fundamental asymptotic relation

$$\lim_{T\to\infty} G_T^{\pi} \leq \lim_{T\to\infty} G_T^{\pi^*} \tag{2.2}$$

for any strictly positive portfolio V^{π}; see Theorem 3.3 in Platen (2011).

One notes in Figure 1.2 that the GRs of the HWI and the EWI fluctuate similarly, which is a consequence of the fact that both well-diversified indexes are driven mostly by the non-diversifiable uncertainty of the equity market. This observation supports our search for the 'best' proxy for the GP among various competing, well-diversified portfolios. It leads us to aim for the largest GR in the longest available observation window.

As previously mentioned, the GP is also called the benchmark, and values denominated in units of the GP are called benchmarked values. We will prove in Appendix B the following fundamental fact:

Theorem 2.1 (Efficient Market Property) *In a continuous market, nonnegative benchmarked portfolios have zero instantaneous expected returns and zero or negative expected returns over any time period, as long as the respective GP exists.*

For easier presentation, we prove this result for continuous markets. However, it holds far more generally, even for semimartingale markets, see Karatzas & Kardaras (2007). The GP is in many ways the 'best' performing portfolio. The Efficient Market Property captures a vital property where it is the 'best' portfolio. Any other portfolio, when used as a benchmark, would violate the statements of Theorem 2.1 for some security and some time. The theorem states that, in the very short term, the current value of any nonnegative benchmarked portfolio is equal to or greater than its future benchmarked values. Note that this statement refers to the given investment universe and the information captured by the state variables of the model that determine the respective GP. In this precise sense the market is efficient.

Since instantaneous expected returns are zero in the denomination of the GP, there is no possibility of maximizing expected returns. Note that only when replacing the denominator, e.g. by the risk-free asset, do nonzero instantaneous expected returns appear for the stock and portfolio dynamics in the new denomination as a consequence of Itô calculus. The mean-variance approach to portfolio optimization has often aimed at maximizing instantaneous expected returns in the denomination of the risk-free asset. The Efficient Market Property identifies the GP as the only benchmark such that for all benchmarked portfolios their current value is the 'best' forecast of their 'next' value over vanishing time periods. The conditional expectation underpinning this statement is taken with respect to the real world probability measure and the information employed is the one captured by the evolution of all state variables of the underlying market model up to the current time.

3 Hierarchically Weighted Index

This section explains the construction of the *hierarchically weighted index* (HWI), which we implement for stocks in developed markets. Details on the Industry Classification Benchmark (ICB), see Reuters (2008), which we use when forming hierarchical groupings, together with information about the data, are provided in Appendix A.

Portfolio Construction

We assume four levels in our hierarchy for developed stock markets. This appears to be reasonable but other numbers of levels are possible. At time $t \geq 0$ we denote by M_t the number of geographical regions, $M_t^{j_1}$ the number of countries in the j_1-th region, $M_t^{j_1,j_2}$ the number of industrial groupings in the j_2-th country of the j_1-th region and $M_t^{j_1,j_2,j_3}$ the number of stocks in the j_3-th industrial grouping of the j_2-th country of the j_1-th region.

Let S_t^j denote the cum-dividend price of the j-th stock (denominated in US dollars) at time $t \geq 0$, $j = (j_1, j_2, j_3, j_4)$. The portfolio weight for the investment in the j-th stock at time t is denoted by π_t^j. The vector process of weights of a strictly positive, self-financing portfolio with value V_t^π at time $t \geq 0$ is (with slight abuse of notation) denoted by $\pi = \{\pi_t = (\pi_t^1, \pi_t^2, \ldots, \pi_t^{N_t})^\top, t \geq 0\}$, where

$$N_t = \sum_{j_1=1}^{M_t} \sum_{j_2=1}^{M_t^{j_1}} \sum_{j_3=1}^{M_t^{j_1,j_2}} M_t^{j_1,j_2,j_3} \tag{3.1}$$

denotes the number of stocks in our investment universe at time $t \geq 0$. We denote here by $x^\top$ the transpose of a vector or matrix x.

We introduce by $0 = t_0 < t_1 < \cdots < t_i < t_{i+1} < \ldots$ the reallocation times for a portfolio V^π. Its value $V_{t_i}^\pi$ at time t_i is calculated, recursively, via the relation

$$V_{t_i}^\pi = V_{t_{i-1}}^\pi \left(1 + \sum_{j=1}^{N_{t_{i-1}}} \pi_{t_{i-1}}^j \frac{S_{t_i}^j - S_{t_{i-1}}^j}{S_{t_{i-1}}^j} \right) \tag{3.2}$$

for $i \in \{1, 2, \ldots\}$ with $V_{t_0}^\pi = 1$. For all portfolios constructed in this paper the rebalancing frequency is quarterly, which is adequate, but can be easily changed. Additional rebalancing is performed when a stock 'dies' between the quarterly rebalancing times. More frequent rebalancing, e.g. monthly, shows similar results to what we report.

The 'traditional' benchmark for fund management is the *market capitalization weighted index* (MCI) with the weight $\pi_{t_i}^{MCI,j}$ invested in the j-th stock at time t_i. This weight is determined by the reported respective market value $MV_{t_i}^j$ via the formula

$$\pi_{t_i}^{MCI,j} = \frac{MV_{t_i}^j}{\sum_{k=1}^{N_{t_i}} MV_{t_i}^k} \tag{3.3}$$

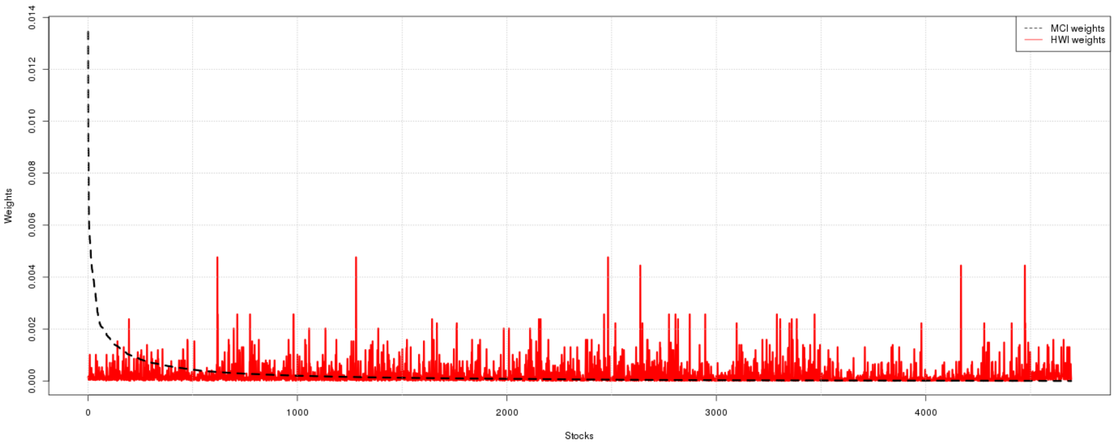

Figure 3.1: The HWI and MCI weights of the stocks ordered by market capitalization.

for $j \in \{1, 2, \ldots, N_{t_i}\}$, $i \in \{0, 1, \ldots\}$.

The weights $\pi_{t_i}^{EWI,j}$ for $j \in \{1, 2, \ldots, N_{t_i}\}$ of the *equally-weighted index* (EWI) are set to be equal at each time t_i,

$$\pi_{t_i}^{EWI,j} = \frac{1}{N_{t_i}}, \tag{3.4}$$

for $i \in \{0, 1, \ldots\}$.

HWI

The proposed *hierarchically weighted index* (HWI) uses region, country and industry groupings of stocks in a hierarchical manner to determine their respective weights. It invests equal fractions of wealth in the constituents of each group. These constituents are themselves equally-weighted indexes or (on the lowest level) stocks. The three geographical regions we distinguish are Europe, Asia-Pacific and North-America. In the first column of Table 3.1 we list the 23 developed markets (countries) considered in this paper. These were chosen by using the ICB classification. The base dates for the start of investments made in each respective country are listed in column two. These are chosen by taking into account the reported number of dead stocks at a given time. Due to the downgrading of Greece to the status of an emerging market in 2013, we ignore Greece. We include Israel, since it was acknowledged as a developed market in 2010.

Due to these (and other) deviations of our set of stocks when compared to those of the MSCI Developed Markets, we form the MCI, which allows us to make a fairer comparison with other constructed indexes. This constructed MCI deviates only marginally from the MSCI, as we show later on. The total number of stocks considered is over $40,000$. About $4,810$ stocks are selected dynamically at a given time from these stocks by the following rules, which make sure that there is, on average, a reasonable number of stocks in the industrial groups formed within each country: The rule for the selection of subsectors (in the ICB sense) as an industrial grouping, is that the country needs to have more than 900 stocks available. For countries with a number of stocks between 80 and 900 we employ the sector grouping. For countries with fewer than 80 stocks the supersector grouping is used as an industrial grouping. For instance, in the case of the United States, we choose the 998 largest (by market value) stocks that are alive at a rebalancing date. The list of the number of stocks chosen for a given country is recorded in column three of Table 3.1. Column four in Table 3.1 indicates the level of ICB grouping we use as an industrial grouping of stocks in the given country.

Recall, for the j-th stock with $j = (j_1, j_2, j_3, j_4)$ we denote by $M_t^{j_1,j_2,j_3}$ the number of stocks in the industrial grouping in its country. $M_t^{j_1,j_2}$ is then the number of respective industrial groupings (subsectors, sectors or supersectors) in the country the j-th stock belongs to. $M_t^{j_1}$ denotes the number of countries in the region of the j-th stock and M_t counts the number of regions considered at time $t \geq 0$.

We equally-weight in each group of the hierarchy the constituents we have formed at the next lower level. The weight for the j-th stock, with $j = (j_1, j_2, j_3, j_4)$, is then of the form

$$\pi_t^{HWI,j} = \frac{1}{M_t} \frac{1}{M_t^{j_1}} \frac{1}{M_t^{j_1,j_2}} \frac{1}{M_t^{j_1,j_2,j_3}} \tag{3.5}$$

for $t \geq 0$. As a result, the weights of industrial groupings and countries in the HWI are rather different to those of the MCI and EWI, see Table 3.2. Another illustration of how different the weights of the HWI are compared to those of the MCI, is given in Figure 3.1. It shows the weights of stocks for the MCI and the HWI, where the stocks are ordered by their market capitalization at the end of our observation period. The HWI employs information about stock groups with exposure to similar uncertainties, whereas the EWI uses no information.

Other Hierarchical Groupings

When forming the HWI we arguably use the most natural hierarchical groupings of stocks, starting with an industry grouping located in a country, which is then part of a region. Hierarchically weighted indexes can also be constructed in other ways, e.g., by using only industrial or only geographical classifications of stocks. Furthermore, one may first use in the lower levels of the hierarchy both geographical and industrial groupings and at higher levels only industrial groupings, as illustrated in Platen & Rendek (2012). This would still group stocks according to exposure to similar uncertainties.

An important question is, does it matter significantly if we have in our hierarchy at the lower level industrial groupings or geographical groupings? As we will see below, the performance of the resulting index appears to be better when

Country	HWI Country Base Date	No. of Stocks	Industrial Grouping
CANADA	01/01/1990	245	sector
UNITED STATES	02/01/1984	998	subsector
HONG KONG	01/01/1986	127	sector
JAPAN	01/01/1990	1000	subsector
UNITED KINGDOM	01/01/1985	539	sector
SPAIN	03/01/2000	116	sector
NETHERLANDS	01/01/1990	107	sector
AUSTRALIA	01/01/1988	160	sector
SWITZERLAND	01/01/1992	146	sector
BELGIUM	02/01/1984	90	sector
FRANCE	01/01/1993	247	sector
GERMANY	01/01/1990	235	sector
ITALY	01/01/1986	150	sector
SINGAPORE	03/01/2005	100	sector
NORWAY	01/01/1990	50	supersector
IRELAND	01/01/1991	37	supersector
SWEDEN	01/01/1991	62	supersector
FINLAND	01/01/1996	47	supersector
AUSTRIA	01/01/1992	49	supersector
PORTUGAL	01/01/1993	48	supersector
DENMARK	01/01/1993	47	supersector
NEW ZEALAND	03/01/2000	50	supersector
ISRAEL	03/01/2000	49	supersector

Table 3.1: Base dates, number of stocks and type of industrial grouping used in a country for the construction of the HWI.

ICB Supersector	HWI	EWI	MCI	Country	HWI	EWI	MCI
Industrial Goods & Services	13.95	16.19	11.42	CANADA	16.67	5.219	4.243
Personal & Household Goods	6.889	4.729	5.272	UNITED STATES	16.67	21.26	48.51
Real Estate	6.586	8.649	3.859	HONG KONG	6.667	2.706	4.499
Technology	6.051	6.966	9.503	JAPAN	6.667	21.28	9.917
Retail	5.932	5.944	5.492	AUSTRALIA	6.667	3.409	2.819
Basic Resources	5.917	3.622	2.244	SINGAPORE	6.667	2.13	1.27
Oil & Gas	5.482	6.221	8.978	NEW ZEALAND	6.667	1.065	0.1381
Food & Beverage	5.418	4.325	4.587	UNITED KINGDOM	2.083	11.48	7.865
Financial Services	5.146	7.861	4.23	SPAIN	2.083	2.45	1.807
Insurance	5.033	2.897	4.137	NETHERLANDS	2.083	2.28	1.238
Health Care	5.01	6.242	10.02	SWITZERLAND	2.083	3.11	3.293
Utilities	4.662	3.409	3.685	BELGIUM	2.083	1.917	0.8247
Telecommunications	4.491	1.513	4.183	FRANCE	2.083	5.262	4.531
Travel & Leisure	4.393	4.325	3.029	GERMANY	2.083	4.985	3.753
Construction & Materials	3.624	4.154	1.434	ITALY	2.083	3.174	1.397
Media	3.237	2.983	2.6	NORWAY	2.083	1.065	0.6474
Chemicals	3.107	2.812	2.699	IRELAND	2.083	0.7882	0.1599
Banks	2.956	4.708	9.696	SWEDEN	2.083	1.321	1.261
Automobiles & Parts	2.111	2.45	2.934	FINLAND	2.083	1.001	0.4259
				AUSTRIA	2.083	1.044	0.2295
				PORTUGAL	2.083	1.001	0.1515
				DENMARK	2.083	1.001	0.6733
				ISRAEL	2.083	1.044	0.3431

Table 3.2: Supersector and country weights for the HWI with comparison to the EWI and MCI.

having an industrial grouping as lowest level and a geographical grouping at the second lowest level.

To decide whether there may exist significantly better performing hierarchical portfolios than the proposed HWI we studied various alternative hierarchical groupings to those we mention below. To illustrate our findings, we report observed long-term growth rates (GRs) for selected hierarchically weighted indexes and compare these to the HWI and EWI. First, we mention two illustrative examples of indexes that are diversified by their geographical origin at the stock level. These are the hierarchically weighted index diversified by region only (HWI.r)

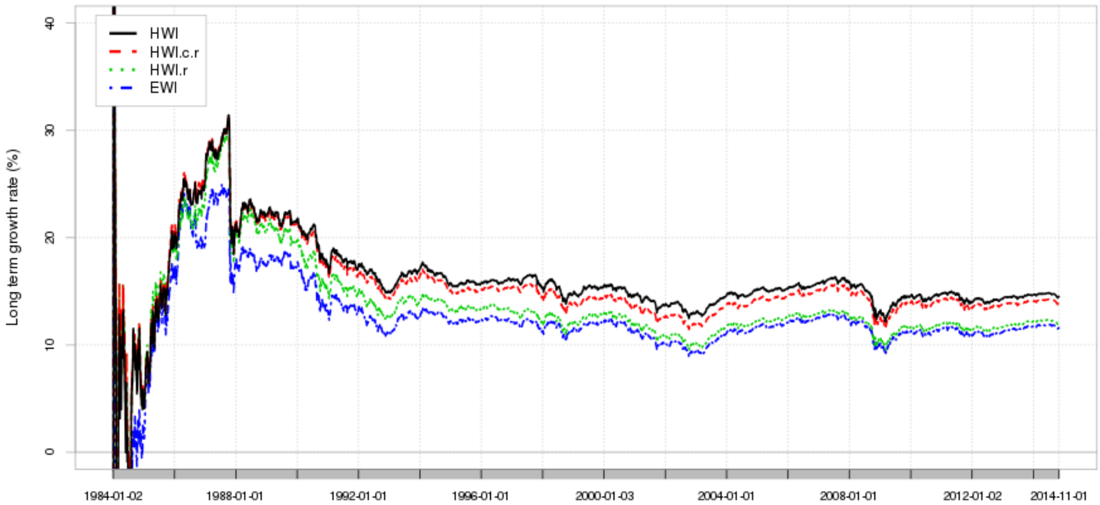

Figure 3.2: The observed long-term growth rates (GRs) (in percent) for the EWI and the hierarchically weighted indexes HWI.r and HWI.c.r, diversified by geographical origin, together with those of the HWI.

and the hierarchically weighted index diversified by country and then by region (HWI.c.r). The GRs of the EWI and these two indexes are compared in Figure 3.2 to the GR of the HWI. We observe that the HWI outperforms the HWI.c.r, which outperforms the HWI.r, and then the EWI. This figure illustrates that the addition of an extra hierarchical level provides consistently also an improvement in the GR. Note in Figure 3.2 that the use of information about the region already provides a slight improvement over the performance of the EWI. Introducing the country level in the hierarchy visually produces the main improvement in the long run. The GR improves further by introducing appropriate industrial groupings of stocks in each country, which leads to the HWI.

We emphasize that the GRs of all these indexes fluctuate similarly, which already allows us to distinguish between their GRs after a few years of observation. The index with the highest GR at the end of the observation period is the HWI, as is also shown in Table 4.2 in the next section.

For further illustration, we mention three other examples of indexes grouped by their industrial origin on the stock level. These are: the hierarchically weighted index HWI.s diversified by global industry sector only, the hierarchically weighted index HWI.c.g diversified by country industry sector first and global industry sector second and the hierarchically weighted index HWI.c.r.g, diversified first by country sector, then by regional sector and, finally, by global sector. In comparison to these indexes and all similar indexes studied we find that our proposed HWI generates the highest observed GR. For this reason, and by the theoretical underpinning we give in Appendix C and Appendix D, we consider the HWI to be the best proxy for the GP for the stocks in developed markets compared to the indexes studied in this paper.

4 Further Empirical Results

This section presents further empirical results concerning the performance of the HWI, EWI and MCI. In column two of Table 4.1 we report the observed (annualized) long-term growth rate (GR) over the available observation window of $T = 31$ years, calculated according to formula (2.1). We re-emphasize that this has been our key performance measure throughout the paper, since it directly targets the GP when maximized. The traditionally used benchmark in equity fund management, the MSCI total return index for developed markets, is included in Tables 4.1-4.5 for comparison. This index draws stocks from the same 23 developed markets we consider. However, it is based only on approximately 1,700 stocks, while the HWI, the MCI and our other indexes are based on over 4,700 stocks. Since the investment universe of the MSCI captures fewer sources of uncertainty than that of the MCI, one should expect a lower GR for the GP of the constituents of the first, which most likely leads also to the lower GR observed for the MSCI compared to that of the MCI.

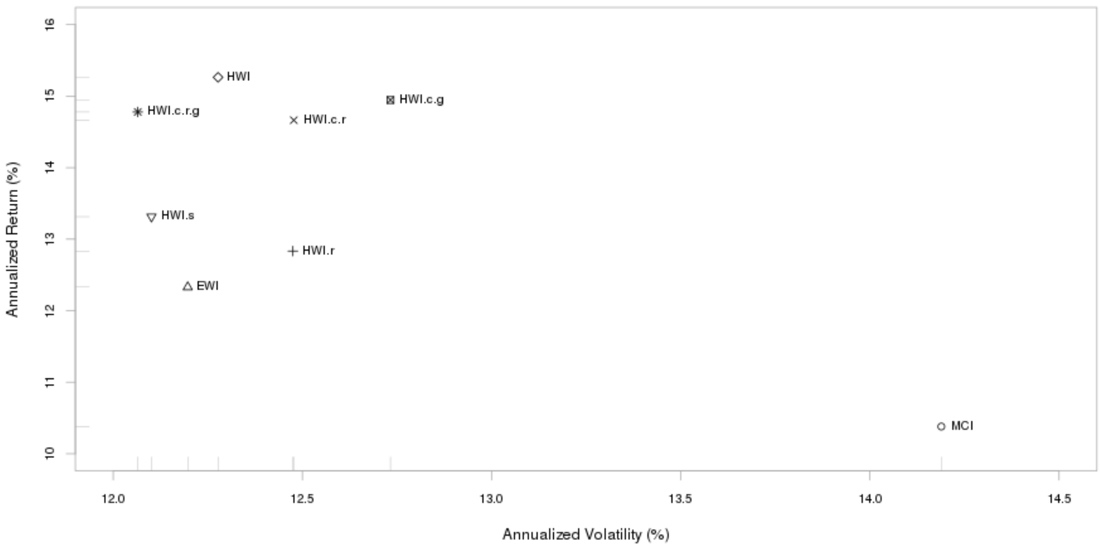

Figure 4.1: Average return versus average volatility of selected indexes.

In Table 4.2 we show the average (with 95% confidence intervals) of the difference between the GRs of the HWI and those of the MSCI, MCI, EWI, HWI.r and HWI.c.r, respectively, estimated from all available observation windows of length 1, 2, 3, 4, 5, 6, 7 and 8 years. Note that the HWI performs best for all observation windows. As we have seen in Figure 1.1 and observe now in Table 4.2, due to

Index	GR	Average Return	Risk Premium	Volatility	Sharpe Ratio
MSCI	9.23	10.48	6.26	15.79	0.3963
MCI	9.37	10.38	6.16	14.19	0.4343
EWI	11.58	12.33	8.11	12.20	0.6650
HWI	14.50	15.26	11.05	12.28	0.8997

Table 4.1: Long-term growth rate (GR), average return, risk premium, volatility and Sharpe ratio for the MSCI, MCI, EWI and HWI.

Period (years)	MSCI	MCI	EWI	HWI.r	HWI.c.r
1	5.418 (5.220,5.616)	5.216 (5.040,5.392)	2.783 (2.680,2.885)	2.536 (2.438,2.635)	0.6297 (0.5989,0.6606)
2	5.518 (5.364,5.673)	5.350 (5.223,5.477)	2.938 (2.877,2.999)	2.633 (2.573,2.693)	0.6329 (0.6080,0.6578)
3	5.631 (5.494,5.769)	5.351 (5.240,5.461)	3.005 (2.963,3.047)	2.654 (2.611,2.697)	0.6290 (0.6066,0.6514)
4	5.771 (5.642,5.900)	5.323 (5.221,5.425)	2.989 (2.957,3.022)	2.728 (2.695,2.761)	0.6312 (0.6102,0.6523)
5	5.914 (5.793,6.035)	5.303 (5.208,5.399)	3.012 (2.986,3.037)	2.811 (2.786,2.836)	0.6413 (0.6212,0.6614)
6	5.973 (5.859,6.086)	5.248 (5.160,5.336)	3.002 (2.979,3.025)	2.847 (2.825,2.869)	0.6504 (0.6312,0.6696)
7	5.964 (5.856,6.073)	5.172 (5.090,5.254)	2.990 (2.968,3.012)	2.870 (2.849,2.891)	0.6639 (0.6452,0.6826)
8	5.986 (5.880,6.091)	5.154 (5.076,5.231)	2.992 (2.972,3.011)	2.904 (2.886,2.921)	0.6696 (0.6515,0.6877)

Table 4.2: Difference in average percentage long-term growth rate (GR) over observation windows reaching from one to eight years with 95% confidence interval between HWI and MSCI, MCI, EWI, HWI.r and HWI.c.r, respectively.

	EWI	HWI
Daily	0.526793	0.539724
Monthly	0.557688	0.605034
Quarterly	0.596067	0.654559
Half-yearly	0.618356	0.689812
Yearly	0.670045	0.747530
2 Yearly	0.748774	0.866499
3 Yearly	0.801810	0.885338
5 Yearly	0.777516	0.886695

Table 4.3: Relative frequency of outperforming the MCI over a given period length for the EWI and HWI.

Index	VaR (95%)	ES (95%)
MCI	-0.012834	-0.021029
EWI	-0.011252	-0.018382
HWI	-0.010636	-0.018680

Table 4.4: Value at Risk (VaR) and Expected Shortfall (ES) for MCI, EWI and HWI at a 95% level.

the similar fluctuations of the well-diversified portfolios considered it seems to take an observation window of only about one year to distinguish reasonably well between the GRs of the HWI, the MCI and the EWI. We emphasize that in Table 4.2 one can observe with 95% confidence that after one year the typical difference between the GRs of the HWI and MCI is already at least about 5%. This is a substantial difference in performance, which only decreases by about 0.35% when taking a realistic 40 basis points proportional transaction cost into account, as

Index	Av. Drawdown	Av. Recovery
MCI	0.0199	14.7533
EWI	0.0187	11.9556
HWI	0.0169	9.4541

Table 4.5: Average relative drawdown and average recovery time (in days) for the MCI, EWI and HWI.

shown below in Table 4.6.

Despite our strategic focus on maximizing the GR, in this section we also provide some popular short term performance and risk measures. The annualized percentage average returns are displayed in column three of Table 4.1 and the risk premium is estimated in column four. The risk premium of the HWI is the highest and reaches approximately 11% compared to 6% for the MCI. The annualized percentage volatility is recorded in column five, and equals 12% for the HWI, which is close to the volatility of the EWI, whereas the volatility of the MCI is higher at about 14%.

Figure 4.1 plots the annualized average daily return against the annualized average volatility for selected indexes, including those with the alternative hierarchical groupings mentioned in the previous section. The HWI exhibits the most favorable annualized average return, and has the highest Sharpe ratio, as shown in column six of Table 4.1. Furthermore, in Table 4.1 it can be seen that the improvement of the HWI in its average return on that of the MCI is about 4.9%, which is above that of the EWI, which is about 2.9%. As indicated in the introduction, the Sharpe ratio of the MCI is about 0.43, that of the EWI is 0.67, and that of the HWI about 0.90. The latter is the highest Sharpe ratio observed in our study. One has to conclude that the MCI is poorly positioned on the respective mean-variance efficient frontier. As discussed, this observation empirically defies classical theory, as developed in Markowitz (1959), Sharpe (1964), and a related stream of literature.

Table 4.3 provides us with observed relative frequencies for outperforming the MCI over a given period length. Note that the largest relative frequencies are observed for the HWI. In Table 4.4 the HWI shows about the smallest absolute values for daily Value at Risk (VaR) and Expected Shortfall (ES), respectively, on a 95% quantile level when compared to those of the MCI and the EWI.

In Table 4.5 we summarize the average drawdown relative to the running maximum and the average time (in days) of recovery back to the level of the running maximum for the different indexes. Again, the HWI performs best in comparison to the MCI and the EWI. This is consistent with the theoretical prediction in Kardaras & Platen (2010), where the GP is shown to require the shortest expected 'market time' to reach a target level.

From an equity fund management perspective one needs to query the impact of transaction costs. In Table 4.6 we are imposing 40 basis points proportional transaction cost on every transaction, leading from the HWI, EWI and MCI to

Index	GR	Average Return	Risk Premium	Volatility	Sharpe Ratio
MCI-TC	9.20	10.22	6.00	14.19	0.4228
EWI-TC	11.30	12.05	7.83	12.20	0.6423
HWI-TC	14.15	14.92	10.70	12.28	0.8714

Table 4.6: Long-term growth rate (GR) and other common statistics for the HWI-TC, EWI-TC, MCI-TC constructed with 40 basis points proportional transaction costs.

the HWI-TC, EWI-TC and MCI-TC, respectively. The turnover for the HWI is surprisingly low. Table 4.6 shows only minor changes in performance estimates that should be compared to the respective values in Table 4.1. Most important is that we observe only a minor decrease of the GR for the HWI from 14.50% shown in Table 4.1 to the GR of the HWI-TC of 14.15% shown in Table 4.6. This makes the HWI-TC a valuable long-term investment security that can be efficiently implemented in practice.

5 Efficient Market Property

We are now ready to employ the constructed indexes to empirically demonstrate that the Efficient Market Property cannot be easily rejected. As shown in Theorem 2.1, a crucial property of the GP is that, when it is used as a benchmark, it causes the expected returns of nonnegative benchmarked portfolios to be negative or zero, but never strictly positive. Thus, one can reject any candidate proxy for the GP by showing that the mean of returns of benchmarked securities is strictly positive at a respective significance level. To attempt this for any single stock would not work, since the available observation window, here 31 years, is clearly too short to provide any reasonable level of significance. However, we can gain sufficient evidence by combining all available daily returns of all benchmarked stocks in a large sample of $31,472,596$ daily returns, where we employ the HWI-TC as the benchmark. For comparison, we also employ the MCI-TC, EWI-TC, HWI.c.r-TC, HWI.c.g-TC and HWI.c.r.g-TC as benchmarks to see whether the theoretical Efficient Market Property, in the sense of Theorem 2.1, cannot be rejected for several of these indexes.

Benchmark	Sample mean	Standard Error	99% LCI	99% UCI	Z-test	p-value
MCI-TC	3.504079	0.142278	3.137594	3.870563	24.628	0
EWI-TC	0.936921	0.141376	0.572761	1.301081	6.627	0
HWI-TC	-1.671584	0.141828	-2.036909	-1.306259	-11.786	1
HWI.c.r-TC	-1.072318	0.141815	-1.437608	-0.707027	-7.561	1
HWI.c.g-TC	-1.238168	0.141928	-1.603750	-0.872587	-8.724	1
HWI.c.r.g-TC	-1.364769	0.141681	-1.729714	-0.999823	-9.633	1

Table 5.1: Test for the mean daily annualized percentage returns of all benchmarked stocks.

Benchmark	Bootstrap mean	99% LCI	99% UCI	Test statistic	p-value
MCI-TC	3.502956	3.140459	3.910323	23.133	0
EWI-TC	0.936228	0.571123	1.398584	6.256	0
HWI-TC	-1.664403	-2.018709	-1.277940	-11.529	1
HWI.c.r-TC	-1.074673	-1.419532	-0.674604	-7.392	1
HWI.c.g-TC	-1.243531	-1.617380	-0.893065	-8.177	1
HWI.c.r.g-TC	-1.361862	-1.725689	-1.001110	-8.968	1

Table 5.2: Bootstrap test for the mean daily annualized percentage returns of benchmarked stocks.

We consider all available daily, annualized percentage returns of all stocks (that constitute the HWI) when benchmarked by the index shown in column one of Table 5.1, and produce the respective sample mean displayed in column two of Table 5.1. In a first step, one could argue that the returns are reasonably independent when observed on different days, and that the returns of different benchmarked stocks on the same day are also reasonably independent because they are mainly driven by their idiosyncratic or specific uncertainties. Therefore, initially assuming independent and identically distributed returns, appears to be acceptable. The Central Limit Theorem determines then the length of the respective confidence intervals. In column three of Table 5.1 we show the resulting standard error and in columns four and five the lower level (LCI) and the upper level (UCI), respectively, of the 99% confidence interval for the 'true' expected daily return of benchmarked stocks. The reader may be surprised to see some confidence intervals covering only negative values. However, this is what Theorem 2.1 predicts when considering returns over some positive time period, here about one day. Theorem 2.1 explains this phenomenon via the supermartingale property of benchmarked securities, see Platen & Heath (2010), Karatzas & Kardaras (2007), Loewenstein & Willard (2000) and Heston, Loewenstein & Willard (2007) for further details in this direction.

To be precise, we denote by μ (in line with Theorem 2.1 and (B.11)), the 'true' expected average of all daily returns of all benchmarked stocks and test the hypothesis:

$$H_0 : \mu \leq 0 \qquad \text{versus} \qquad H_1 : \mu > 0. \tag{5.1}$$

We display the corresponding test statistic of the well-known Z-test (see e.g. Mode (1966)) in column six, and the respective one-sided p-values in column seven of Table 5.1. On a 1% level of significance we can clearly reject H_0 for the MCI-TC and EWI-TC when used as benchmarks. However, we cannot reject H_0 for the HWI-TC, HWI.c.r-TC, HWI.c.g-TC and HWI.c.r.g-TC. This means, these hierarchically constructed indexes are closer to the GP than the MCI-TC and EWI-TC.

One may argue that the assumption of independent and identically distributed returns in the first step of our analysis is too strong and should be relaxed. Therefore, in the second step of our study we remove this assumption and report in Table 5.2 the block bootstrap percentile 99% confidence intervals, with the respective test statistics and p-values for the hypothesis (5.1)[5]. We observe that the results in Table 5.2 resemble those in Table 5.1.

We emphasize in Table 5.1 and Table 5.2 that for the not rejected indexes the 99% confidence intervals for the mean μ do not include zero in all cases, and cover slightly negative values. As already explained, this is a consequence of the theoretically predicted supermartingale property of GP benchmarked stocks. We note that when using the HWI-TC as a proxy for the GP, the 99% confidence interval for the mean of the daily returns of benchmarked stocks is the 'most negative', which supports our choice of the HWI-TC as the best proxy of the GP among the considered indexes.

Benchmark	Sample mean	99 % LCI	99 % UCI	Z-test	p-value
MCI-TC	-5.239304	-8.707531	-1.771077	-3.89	1
EWI-TC	-2.943474	-4.914147	-0.972799	-3.85	1
HWI.c.r-TC	-0.587549	-1.219394	0.044296	-2.4	0.99
HWI.c.g-TC	-0.440414	-2.217511	1.336683	-0.64	0.74
HWI.c.r.g-TC	-0.477921	-1.354879	0.399037	-1.4	0.92

Table 5.3: Test for the mean daily annualized returns of selected HWI-TC benchmarked portfolios.

We already mentioned that well-diversified candidate proxies for the GP are driven by similar uncertainties and can be empirically compared with each other, even over relatively short observation windows. In Table 5.3 we test the average daily returns of the HWI-TC benchmarked MCI-TC, EWI-TC, HWI.c.r-TC, HWI.c.g-TC and HWI.c.r.g-TC according to the hypothesis (5.1). We note that the hypothesis H_0 in (5.1) cannot be rejected for the HWI-TC when used as a benchmark for each of the in Table 5.3 considered indexes. Since all the p-values are above 0.7, we also cannot reject the Efficient Market Property at the 1% level of significance. We repeated this study using the block bootstrap methodology and obtained very similar results.

Our deliberately simply designed empirical study indicates that it is difficult to reject the theoretically predicted Efficient Market Property for stocks in developed markets by employing the HWI-TC as a proxy for the respective GP. In Fama (1970, 1991, 1998) and subsequent literature various efficient market hypotheses have been proposed and empirically studied. What is important for these forms of market efficiency is the degree of information exploited. We have now seen that such information is not highly relevant. We only need to take into account information about the natural industrial and geographical grouping of stocks to construct the HWI-TC, our best proxy for the GP for which the Efficient Market Property has not been rejected. This means that due to the well-diversified nature of the GP for stocks in developed markets, the Efficient Market Property of this market is quite robust.

[5]The block bootstrap replicates of the sample mean are obtained with the tsboot() function in the boot package in R, see e.g. Davison & Hinkley (2007), using block resampling with block lengths having a geometric distribution.

It is beyond the range of this paper to go any further in empirically studying the Efficient Market Property. Our aim here is to provide a new understanding of the objectively present market efficiency and to open a new direction for empirical research.

6 Conclusion

This paper reveals a deep connection between the Efficient Market Property and the growth optimal portfolio (GP), since stocks denominated in units of the GP display zero or negative expected returns. Due to the impossibility of estimating means and covariances of stock returns, theoretically optimal stock portfolios, including the GP, cannot be implemented accurately enough for larger stock markets to be useful. The paper approximates the GP for stocks in developed markets by a hierarchically weighted index (HWI), which does not rely on any estimation and sets equal weights within industrial and geographical groupings. A Diversification Theorem explains why naive diversification works well and why the HWI performs even better. The HWI is that portfolio which has the highest observed long-term growth rate among the other well-diversified portfolios considered. The Efficient Market Property is difficult to reject empirically when using the HWI as a proxy for the GP. This indicates that the GP is rather close to the HWI. Since no information is needed beyond that encapsulated in the prices of stocks and their industrial and geographical hierarchical groupings, the Efficient Market Property appears to be robust. These findings provide a new understanding of market efficiency. By constructing an excellent proxy for the GP for a particular market one can empirically test for market efficiency. All that is required for such a test is a sufficiently accurate proxy for the GP, which is an optimal portfolio that gives access to the efficient construction of many other optimal portfolios. The HWI, as a proxy for the GP of developed markets, is in itself useful in equity fund management and can serve as a building block in portfolio optimization and risk management. The GP plays a central role as the numéraire portfolio for pricing and hedging under the real world probability measure in the benchmark pricing theory, which goes beyond classical finance with its richer modeling world and coverage of new phenomena. The HWI, as an excellent proxy for the GP, allows one to practically demonstrate and exploit such new phenomena that can potentially explain various puzzles in classical finance, as forthcoming work will demonstrate.

Appendix A: Data

In this appendix we describe our data. Thomson Reuters Datastream (TRD) provides a range of TRD calculated country, region and sector indexes together with their current constituents and lists of dead stocks. For developed stock markets the current paper builds global stock indexes from the data available in the TRD database. The 23 developed countries included in this study are given in the first column of Table A.1. These developed countries are identified based on the FTSE/ICB country classification. For each of these developed markets Datastream uses a sample of stocks covering a minimum of 75 - 80% of total market capitalization by choosing the largest stocks by market value. Column two of Table A.1 lists the corresponding approximate number of stocks in the Datastream index for the selected 23 countries and in column three the relevant base date from which the index is available; see also Reuters (2008).

The respective list of constituents of Datastream country indexes was obtained from the TRD database by quoting the mnemonic for each country list. Table A.2 lists the mnemonics used for the active and dead stocks in the included markets. We also provide in this table the number of active and dead stocks present on both mentioned lists for each country. Note that we only consider those stocks in the lists whose "GEOLN"= "GEOGN"= "Country Name". Additionally, companies with datatype "MAJOR"= "Y" are included, which means that for companies with more than one equity security the one with the largest market capitalization is chosen.

Country	Approx. no. of stocks	Base Date
CANADA	250	Jan 1973
UNITED STATES	1000	Jan 1973
HONG KONG	130	Jan 1973
JAPAN	1000	Jan 1973
UNITED KINGDOM	550	Jan 1965
SPAIN	120	Jan 1986
NETHERLANDS	130	Jan 1973
AUSTRALIA	160	Jan 1973
SWITZERLAND	150	Jan 1973
BELGIUM	90	Jan 1973
FRANCE	250	Jan 1973
GERMANY	250	Jan 1973
ITALY	160	Jan 1973
SINGAPORE	100	Jan 1973
NORWAY	50	Jan 1980
IRELAND	50	Jan 1973
SWEDEN	70	Jan 1982
FINLAND	50	Mar 1988
AUSTRIA	50	Jan 1973
PORTUGAL	50	Jan 1990
DENMARK	50	Jan 1973
NEW ZEALAND	50	Jan 1988
ISRAEL	50	Jan 1992

Table A.1: Base dates and number of constituents for Datastream indexes

TRD classifies equities according to the previously mentioned Industry Classification Benchmark (ICB). We only consider those stocks that are classified into one of the subsectors and exclude all stocks that are unclassified (UNCLAS) or classified as one of the following: unqouted equities (UQEQS), exchange traded funds (NEINV), suspended equities (SUSEQ), and other equities (OTHEQ). The number of such removed securities is recorded in the last column of Table A.2.

We use the ICB classification of stocks into subsectors, sectors and supersectors, downloaded from Thomson Reuters Datastream (TRD) with mnemonics FTAG3, FTAG4, FTAG5. The country where the given stock originates is also recorded by TRD. Finally, the countries are grouped into three regions: Americas, Europe (EMEA) and Asia-Pacific.

Country	Active	No. Active	Downl.	Dead	No. Dead	Downl.	Removed.
CANADA	LTOTMKCN	250	245	DEADCN1-2	6814	5310	598
UNITED STATES	LTOTMKUS	999	998	DEADUS1-6	22189	17009	467
HONG KONG	LTOTMKHK	130	127	DEADHK	248	200	8
JAPAN	LTOTMKJP	1000	1000	DEADJP	1681	1569	14
UNITED KINGDOM	LTOTMKUK	549	539	DEADUK	5625	5263	1037
SPAIN	LTOTMKES	120	116	DEADES	264	180	9
NETHERLANDS	LTOTMKNL	117	107	DEADNL	429	343	39
AUSTRALIA	LTOTMKAU	160	160	DEADAU	1784	1531	32
SWITZERLAND	LTOTMKSW	150	146	DEADSW	360	246	12
BELGIUM	LTOTMKBG	90	90	DEADBG	271	245	23
FRANCE	LTOTMKFR	250	247	DEADFR	1534	1400	243
GERMANY	LTOTMKBD	250	235	DEADBD	3000	2229	21
ITALY	LTOTMKIT	160	150	DEADIT	422	339	24
SINGAPORE	LTOTMKSG	100	100	DEADSG	409	391	4
NORWAY	LTOTMKNW	50	50	DEADNW	415	400	34
IRELAND	LTOTMKIR	37	37	DEADIR	129	108	22
SWEDEN	LTOTMKSD	70	62	DEADSD	819	709	72
FINLAND	LTOTMKFN	50	47	DEADFN	149	124	17
AUSTRIA	LTOTMKOE	50	49	DEADOE	196	160	8
PORTUGAL	LTOTMKPT	50	48	DEADPT	239	165	51
DENMARK	LTOTMKDK	50	47	DEADDK	277	254	15
NEW ZEALAND	LTOTMKNZ	50	50	DEADNZ	252	224	4
ISRAEL	LTOTMKIS	50	49	DEADIS	505	413	1

Table A.2: Datastream stock lists and number of equity data obtained

The number of stocks for each active and dead list is recorded in Table A.2 in the fourth and seventh columns, respectively. For the downloaded stocks we obtained the total return prices and the market capitalization. The market value is understood here to be the reported number of ordinary shares in the market multiplied by the stock price. Note that we have accounted for the fact that TRD repeats the last valid stock price or market capitalization for delisted stocks after a delisting. We found it necessary to remove this zero return from the end of the time-series. With the downloaded data prepared in this manner we perform our study, where we recover well the by TRD historically formed market capitalization weighted indexes for the 23 developed markets. We then use this data set for constructing the EWI, MCI and hierarchically weighted indexes.

Appendix B: Efficient Market Property

Market Setting

We prove in this appendix several theoretical results which underpin our derivation of the Efficient Market Property and construction of well-diversified portfolios to approximate the GP. To avoid technicalities, we consider in our proofs and derivations a continuous financial market. However, the Efficient Market Property can be shown to hold for general semimartingale markets, see Karatzas & Kardaras (2007) and Du & Platen (2016). The traded uncertainty of stocks is modeled by an n-dimensional standard Brownian motion $W = \{W_t = (W_t^1, \dots, W_t^n)^\top, t \in [0, \infty)\}$, $n \in \{2, 3, \dots\}$ on a filtered probability space $(\Omega, \mathcal{F}, \underline{\mathcal{F}}, P)$ satisfying the usual conditions, see e.g. Karatzas & Shreve (1998), where the filtration $\underline{\mathcal{F}} = (\mathcal{F}_t)_{t \geq 0}$ models the evolution of information, generated by the quantities constituting the model. This information is characterized at time $t \geq 0$ by the sigma algebra $\mathcal{F}_t$. Concerning the Efficient Market Property, information is, thus, 'quickly' and 'correctly' expressed in prices and other quantities constituting the market model. Again, $x^\top$ denotes the transpose of x. For matrices x and y we write $x \cdot y$ for the matrix product of x and y. Moreover, $\mathbf{1} = (1, \ldots, 1)^\top$ is a vector, and we write 0 for a zero matrix or vector, where the dimensions follow from the context.

Consider m nonnegative, stocks with vector value process $S = \{S_t = (S_t^1, \ldots, S_t^m)^\top, t \in [0, \infty)\}$, denominated in units of the domestic currency, which satisfies the stochastic differential equation (SDE)

$$\frac{dS_t}{S_t} = a_t dt + b_t \cdot dW_t, \tag{B.1}$$

$t \in [0, \infty)$, $S_0^i > 0$ for $i = 1, \ldots, m$. Note that all dividends are reinvested. The instantaneous expected return vector process $a = \{a_t = (a_t^1, \ldots, a_t^m)^\top, t \in [0, \infty)\}$ and the volatility matrix process $b = \{b_t = [b_t^{j,k}]_{j,k=1}^{m,n}, t \in (0, \infty]\}$ are assumed to be adapted and such that there exists a unique strong solution of the SDE (B.1); see e.g. Section 7.7 in Platen & Heath (2010) for respective sufficient conditions.

A strictly positive, self-financing, portfolio process V^π is characterized by the weights or fractions of wealth $\pi_t = (\pi_t^1, \ldots, \pi_t^m)^\top, t \in [0, \infty)$, invested in the stocks, together with its positive initial value $V_0^\pi > 0$, where

$$\pi_t^\top \cdot \mathbf{1} = 1. \tag{B.2}$$

The portfolio value V_t^π at time t then satisfies the SDE

$$\frac{dV_t^\pi}{V_t^\pi} = \pi_t^\top \cdot \frac{dS_t}{S_t} = \pi_t^\top \cdot a_t dt + \pi_t^\top \cdot b_t \cdot dW_t \tag{B.3}$$

for $t \in [0, \infty)$.

Growth Optimal Portfolio

We characterize the growth optimal portfolio (GP) by the following result, which follows directly from Theorem 3.1 in Filipović & Platen (2009), where it has been shown that the GP is equivalent to the expected logarithmic utility maximizing portfolio, also called the Kelly portfolio, see Kelly (1956), and studied in a stream of literature; see MacLean et al. (2011).

Theorem B.1 (Growth Optimal Portfolio Theorem) *If a GP exists in a given continuous market, then the process π^* of GP weights may not be unique. However, the GP value process $V^{\pi_*} = \{V_t^{\pi_*}, t \in [0, \infty)\}$ is unique for some fixed initial portfolio value, which we set as $V_0^{\pi_*} = 1$, and the SDE of the GP is of the form*

$$\frac{dV_t^{\pi_*}}{V_t^{\pi_*}} = \lambda_t dt + \theta_t^\top \cdot (\theta_t dt + dW_t) \tag{B.4}$$

for $t \in [0,\infty)$. Here we set

$$\theta_t = b_t^\top \cdot \pi_t^*, \tag{B.5}$$

with π_t^* and λ_t representing the components of the solution of the matrix equation

$$\begin{pmatrix} b_t b_t^\top & \mathbf{1} \\ \mathbf{1}^\top & 0 \end{pmatrix} \begin{pmatrix} \pi_t^* \\ \lambda_t \end{pmatrix} = \begin{pmatrix} a_t \\ 1 \end{pmatrix} \tag{B.6}$$

for all $t \in [0,\infty)$. A sufficient condition for the existence of a solution of (B.6) is the invertibility of the covariance matrix $b_t \cdot b_t^\top$ for all $t \in [0,\infty)$. In the case when the risk-free asset is included in the investment universe λ_t equals the risk-free rate and θ_t represents the vector of market prices of risk.

Assuming that the GP exists, it follows from the above Growth Optimal Portfolio Theorem that the Lagrange multiplier λ_t and the GP volatility vector θ_t are uniquely determined through a_t and b_t. Moreover, by (B.5) and (B.6) the vector of instantaneous expected returns has the form

$$a_t = \lambda_t \mathbf{1} + b_t \cdot \theta_t. \tag{B.7}$$

Consequently, for any self-financing portfolio V_t^π, the SDE (B.3) takes the form

$$\frac{dV_t^\pi}{V_t^\pi} = \lambda_t dt + \pi_t^\top \cdot b_t \cdot (\theta_t dt + dW_t). \tag{B.8}$$

To correctly identify the optimal strategy for the GP one would need accurate information about a_t and b_t, which, as we argue in this paper, seems impossible to extract sufficiently precisely to be useful in portfolio optimization for large equity markets. However, some reliable information is available through the hierarchical industrial and geographical grouping of stocks. As we show in this paper, this information is sufficient to approximate the GP well enough so that the following Efficient Market Property cannot be easily rejected.

Efficient Market Property

In the benchmark pricing theory of Platen & Heath (2010), the GP of a given set of constituents is called the benchmark. Any security or portfolio V_t^π is called benchmarked when denominated in units of the benchmark. By applying the Itô formula to the benchmarked portfolio value $\hat{V}_t^\pi = \frac{V_t^\pi}{V_t^{\pi *}}$ we obtain the SDE for its return process $\hat{Q}^\pi = \{\hat{Q}_t^\pi, t \in [0,\infty)\}$ from (B.8) and (B.4) as

$$d\hat{Q}_t^\pi = \frac{d\hat{V}_t^\pi}{\hat{V}_t^\pi} = \left(\pi_t^\top \cdot b_t - \theta_t^\top\right) \cdot dW_t \tag{B.9}$$

for $t \in [0,\infty)$. The key observation is here that this SDE is driftless. Consequently, the following fundamental fact emerges:

Theorem B.2 *In a continuous market the instantaneous expected returns of GP-benchmarked, self-financing portfolios equal zero.*

To make the Efficient Market Property testable we need to acknowledge the fact that we can only observe returns over a nonvanishing strictly positive time period. To characterize the situation where we have a strictly positive time period over which we observe returns, we note that driftless benchmarked portfolio values form, so called, local martingales. By Fatou's lemma any nonnegative local martingale is a supermartingale, which means that any nonnegative benchmarked portfolio $\hat{V}^\pi$ satisfies the inequality

$$\hat{V}_t^\pi \geq E_t(\hat{V}_{t+h}^\pi) \tag{B.10}$$

for all $0 \leq t \leq t+h < \infty$; see e.g. Platen & Heath (2010). Here E_t denotes the conditional expectation under the real-world probability measure P given the information $\mathcal{F}_t$ at time t. As indicated earlier, the supermartingale property (B.10) holds generally in semimartingale markets, see Karatzas & Kardaras (2007) and Du & Platen (2016).

By setting the benchmark equal to the GP one has only negative or zero (but never strictly positive) expected returns for nonnegative benchmarked securities, since by (B.10) we have

$$E_t\left(\frac{\hat{V}_{t+h}^\pi - \hat{V}_t^\pi}{\hat{V}_t^\pi}\right) \leq 0 \tag{B.11}$$

for $0 \leq t \leq t+h < \infty$. This result, together with Theorem B.2, proves the statement of Theorem 2.1.

Appendix C: Hierarchical Diversification

Hierarchical Grouping of Stocks

To formulate the Diversification Theorem we make a few assumptions that avoid technicalities in its formulation and proof. These assumptions can be significantly relaxed in an obvious manner. We assume that the stocks can be classified into hierarchical groupings with a fixed number $H \in \{1, 2, \dots\}$ of hierarchical levels. For example, in the construction of the HWI, shown in Figure 1.1, we choose $H = 4$. In the asymptotics of the Diversification Theorem, we let a number $M \in \{2, 3, \dots\}$ tend to infinity, whereas H remains fixed. We assume that in each group of the hierarchy we have at least $\underline{K}M$ and at most $\bar{K}M$ next lower level subgroups, with fixed integers $\underline{K}$ and $\bar{K}$, $0 < \underline{K} \leq \bar{K} < \infty$. This means, for given M we have at least $(\underline{K}M)^H$ and at most $(\bar{K}M)^H$ stocks in our investment universe.

We denote by $W^{k_1} = \{W_t^{k_1}, t \in [0, \infty)\}$ the k_1-th independent, standard Brownian motion that primarily drives the k_1-th group on the highest level of the hierarchy, $k_1 \in \{1, 2, \dots, \bar{K}M\}$. Furthermore, for $k_1, k_2 \in \{1, 2, \dots, \bar{K}M\}$

we let $W^{k_1,k_2} = \{W_t^{k_1,k_2}, t \in [0,\infty)\}$ denote the (k_1,k_2)-th independent standard Brownian motion that primarily models the uncertainty driving the k_2-th group on the second highest level in the k_1-th group of the highest level. In an analogous manner we introduce independent standard Brownian motions for next lower level groups until we reach at the lowest level the stocks. Here $W^{k_1,k_2,\ldots,k_H}$ denotes the $(k_1,k_2,\ldots,k_H)$-th independent standard Brownian motion that primarily drives the k_H-th stock in the k_{H-1}-th lowest level group, of the k_{H-2}-th second lowest level group, etc. For the j_H-th benchmarked stock in the j_{H-1}-th lowest level group of the j_{H-2}-th second lowest level group, etc., we write $\hat{S}_t^j$, where $j = (j_1, j_2, \ldots, j_H) \in \Gamma_M = (1, 2, \ldots, \bar{K}M)^H$. By using (B.1) and (B.7) we capture the hierarchical structure of the stock market dynamics for the j-th benchmarked stock price $\hat{S}_t^j$ by assuming the SDE

$$\begin{aligned} \frac{d\hat{S}_t^j}{\hat{S}_t^j} &= \sum_{k_1=1}^{\bar{K}M} (\psi_t^{j,k_1} dW_t^{k_1} + \sum_{k_2=1}^{\bar{K}M} (\psi_t^{j,k_1,k_2} dW_t^{k_1,k_2} \\ &+ \cdots + \sum_{k_H=1}^{\bar{K}M} \psi_t^{j,k_1,k_2,\ldots,k_H} dW_t^{k_1,k_2,\ldots,k_H})). \end{aligned} \tag{C.1}$$

Note that by setting the respective volatility coefficient to zero for some $j \in \Gamma_M$ we can conveniently model groups that have less than $\bar{K}M$ next lower level subgroups.

According to (C.1) the hierarchical groupings allow us to capture in each group the subgroups that have exposure to similar industrial and geographical uncertainties. The typical sources of uncertainty for members of a group are only assumed to be of significance to other members of the group, which seems to be reasonable.

Diversification Theorem

For given $M \in \{2, 3, \ldots\}$ the SDE for the return process $\hat{Q}_t^{\pi_M}$ of a given benchmarked portfolio $\hat{V}_t^{\pi_M}$, with fraction $\pi_{M,t}^j$ invested in the jth stock, $j \in \Gamma_M$, is by (B.9) of the form

$$d\hat{Q}_t^{\pi_M} = \frac{d\hat{V}_t^{\pi_M}}{\hat{V}_t^{\pi_M}} = \sum_{j\in\Gamma_M} \pi_{M,t}^j \frac{d\hat{S}_t^j}{\hat{S}_t^j} \tag{C.2}$$

$$= \sum_{k_1=1}^{\bar{K}M} \sum_{j\in\Gamma_M} \pi_{M,t}^j \psi_t^{j,k_1} dW_t^{k_1} \tag{C.3}$$

$$+ \sum_{k_1=1}^{\bar{K}M} \sum_{k_2=1}^{\bar{K}M} \sum_{j\in\Gamma_M} \pi_{M,t}^j \psi_t^{j,k_1,k_2} dW_t^{k_1,k_2} \tag{C.4}$$

$$+ \cdots + \sum_{k_1=1}^{\bar{K}M} \sum_{k_2=1}^{\bar{K}M} \cdots \sum_{k_H=1}^{\bar{K}M} \sum_{j\in\Gamma_M} \pi_{M,t}^j \psi_t^{j,k_1,k_2,\ldots,k_H} dW_t^{k_1,k_2,\ldots,k_H}. \tag{C.5}$$

Note that the benchmarked GP is trivially one. Therefore, the time derivative of the quadratic variation, see e.g. Platen & Heath (2010), of its return process is zero. Since we aim to identify proxies for the GP, we identify these asymptotically as follows (where the time derivative of the quadratic variation of the return process of the benchmarked proxy vanishes as M tends to infinity):

Definition C.1 *We call a sequence of benchmarked portfolio processes* $(\hat{V}^{\pi_M})_{M\in\{2,3,\dots\}}$, *each with return process* $\hat{Q}^{\pi_M}$, *a sequence of benchmarked approximate GP processes if for all* $\varepsilon > 0$ *and* $t \in [0, \infty)$ *the limit in probability*

$$\lim_{M\to\infty} P\left(\frac{d[\hat{Q}^{\pi_M}_{\cdot}]_t}{dt} > \varepsilon\right) = 0 \tag{C.6}$$

holds.

For diversification to be possible, we need a condition which ensures that not all benchmarked stocks are driven to a significant extent by the same uncertainties. Therefore, we make the following (rather reasonable) assumption:

Assumption C.2 *For given* $k_1, k_2, \dots, k_h \in \{1, 2, \dots, \bar{K}M\}$ *we assume that for all* $M \in \{2, 3, \dots\}$ *and all* $h \in \{1, 2, \dots, H\}$

$$\sum_{j\in\Gamma_M} |\psi_t^{j,k_1,k_2,\dots,k_h}| \leq (\bar{K}M)^{H-h}\sigma_t, \tag{C.7}$$

where the adapted stochastic process $\sigma = \{\sigma_t, t \geq 0\}$ *satisfies the square integrability condition*

$$E((\sigma_t)^2) \leq \bar{\sigma}^2 < \infty \tag{C.8}$$

for all $t \in [0, \infty)$.

Assumption C.2 covers an extremely wide range of hierarchical market models. Note that the particular form of the volatilities $\psi_t^{j,\cdot}$ is not relevant here. What is limited by (C.7) is the sum of the absolute values of volatilities with respect to the same source of uncertainty. The above property secures some convergence towards the GP if the weights of the constituents vanish 'fast enough' for increasing M, as we specify in the Diversification Theorem below:

Theorem C.3 (Diversification Theorem) *A sequence of benchmarked portfolios* $(\hat{V}^{\pi_M})_{M\in\{2,3,\dots\}}$, *is a sequence of benchmarked approximate GPs, if for each* $M \in \{2, 3, \dots\}$ *the maximum of the weights satisfies the relation*

$$\max_{j\in\Gamma_M} |\pi^j_{M,t}| \leq CM^{\xi-H} \tag{C.9}$$

for some parameter $\xi \in [0, \frac{1}{2})$, *some constant* $C \in (0, \infty)$, *and for all* $t \in [0, \infty)$.

Obviously, relation (C.9) is satisfied for the EWI and the HWI, since in these cases we have by (3.1), (3.4) and (3.5) that $\pi^j_{M,t} \leq \underline{K}^{-H} M^{-H}$ with $C = \underline{K}^{-H}$ and $\xi = 0$ in (C.9). Therefore, both indexes form sequences of approximate GPs. We will see below that the proof of the above Diversification Theorem crucially exploits the Efficient Market Property.

The Diversification Theorem can also be intuitively interpreted as a consequence of a generalized version of the Law of Large Numbers for returns of benchmarked proxies for the GP. To see this in an illustrative situation, consider benchmarked constituents of the GP that have independent, square integrable returns. We know from the Efficient Market Property that these returns have asymptotically zero mean over vanishing time periods. When we form an equally-weighted index (EWI), the total return of the benchmarked EWI becomes the average of the independent returns. Thus, by the Law of Large Numbers this asymptotically yields zero returns for the benchmarked EWI for increasing number of constituents. Consequently, the benchmarked EWI equals in the limit the constant one. By multiplying the limiting benchmarked EWI, that is the constant value one, with the GP in domestic currency denomination we obtain the GP in domestic currency denomination. Thus, the GP asymptotically equals the limit of the EWI. This illustration explains intuitively that, for an increasing number of stocks, naive diversification asymptotically approximates the GP. The above Diversification Theorem identifies not only the EWI as a good proxy for the GP but also the HWI and many other well-diversified portfolios as good proxies.

Proof of Theorem C.3

According to Definition C.1 we need to examine the time derivative of the quadratic variation of the return process $\hat{Q}^{\pi_M}$ of the benchmarked portfolio $\hat{V}^{\pi_M}$. The time derivative of the quadratic variation of the above return process $\hat{Q}^{\pi_M}$ equals

$$\begin{aligned} \frac{d[\hat{Q}^{\pi_M}_{\cdot}]_t}{dt} &= \sum_{k_1=1}^{\bar{K}M} \Big(\sum_{j \in \Gamma_M} \pi^j_{M,t} \psi^{j,k_1}_t \Big)^2 \\ &+ \sum_{k_1=1}^{\bar{K}M} \sum_{k_2=1}^{\bar{K}M} \Big(\sum_{j \in \Gamma_M} \pi^j_{M,t} \psi^{j,k_1,k_2}_t \Big)^2 \\ &+ \dots + \sum_{k_1=1}^{\bar{K}M} \sum_{k_2=1}^{\bar{K}M} \cdots \sum_{k_H=1}^{\bar{K}M} \Big(\sum_{j \in \Gamma_M} \pi^j_{M,t} \psi^{j,k_1,k_2,\dots,k_H}_t \Big)^2 \end{aligned} \tag{C.10}$$

for $t \in [0, \infty)$. By (C.7) we obtain from (C.10) that

$$
\begin{aligned}
\frac{d[\hat{Q}^{\pi_M}_{\cdot}]_t}{dt} &\leq (\max_{j\in\Gamma_M}|\pi^j_{M,t}|)^2\Bigg(\sum_{k_1=1}^{\bar{K}M}(\sum_{j\in\Gamma_M}|\psi_t^{j,k_1}|)^2 \\
&+ \sum_{k_1=1}^{\bar{K}M}\sum_{k_2=1}^{\bar{K}M}(\sum_{j\in\Gamma_M}|\psi_t^{j,k_1,k_2}|)^2 \\
&+ \cdots + \sum_{k_1=1}^{\bar{K}M}\sum_{k_2=1}^{\bar{K}M}\cdots\sum_{k_H=1}^{\bar{K}M}(\sum_{j\in\Gamma_M}|\psi_t^{j,k_1,k_2,\ldots,k_H}|)^2\Bigg) \\
&\leq (\max_{j\in\Gamma_M}|\pi^j_{M,t}|)^2\sigma_t^2\sum_{h=1}^{H}(\bar{K}M)^h(\bar{K}M^{H-h})^2 \\
&\leq (\max_{j\in\Gamma_M}|\pi^j_{M,t}|)^2\sigma_t^2\bar{K}^{2H}\sum_{h=1}^{H}(M)^h(M^{H-h})^2 \\
&= \sigma_t^2(\max_{j\in\Gamma_M}|\pi^j_{M,t}|)^2\bar{K}^{2H}M^{2H-1}\sum_{h=1}^{H}\left(\frac{1}{M}\right)^{h-1} \\
&\leq \sigma_t^2\bar{K}^{2H}(\max_{j\in\Gamma_M}|\pi^j_{M,t}|)^2\frac{M^{2H-1}}{1-\frac{1}{M}}.
\end{aligned}
$$

The last estimate follows from the well-known formula for the limit of the sum of a geometric series. Thus, by using (C.9) we arrive at

$$
\begin{aligned}
\frac{d[\hat{Q}^{\pi_M}_{\cdot}]_t}{dt} &\leq C^2\sigma_t^2\bar{K}^{2H-h}M^{2(\xi-H)}M^{2H-1}(1-M^{-1})^{-1} \\
&\leq C^2\sigma_t^2\bar{K}^{2H-h}M^{2\xi-1}(1-M^{-1})^{-1} \\
&\leq 2C^2\sigma_t^2\bar{K}^{2H-h}M^{2\xi-1}.
\end{aligned}
\tag{C.11}
$$

Since $\xi\in[0,\frac{1}{2})$, it follows by the Markov inequality with (C.8) for each $\varepsilon>0$ and $t\in[0,T]$ that

$$
\lim_{M\to\infty}P(\frac{d[\hat{Q}^{\pi_M}_{\cdot}]_t}{dt}>\varepsilon)\leq\lim_{M\to\infty}\frac{1}{\varepsilon}E(\frac{d[\hat{Q}^{\pi_M}_{\cdot}]}{dt})\leq\frac{2C^2}{\varepsilon}\bar{\sigma}^2\bar{K}^{2H}\lim_{M\to\infty}M^{2\xi-1}=0,
\tag{C.12}
$$

which proves the Diversification Theorem. □

Appendix D: Stylized Hierarchical Market Model

To illustrate why the HWI performs so well, even though in reality market prices of risk and other quantities are most likely changing rapidly and cannot be estimated, we describe here a realistic but stylized hierarchical stock market model that is covered by the Diversification Theorem. The key observation is that under this model the HWI coincides exactly with the GP. The somewhat unexpected insight is that the GP strategy does not require any quantity as input that has to be estimated. The stylized model exploits our assumed hierarchical stock market structure. It then sets various stochastic quantities equal, which for economic reasons are most likely similar, but have no chance of being estimated with any useful accuracy. The particular values of the randomly changing quantities turn out to be irrelevant for the optimal weights of the GP. We stress that under the following stylized model various key quantities can form wide fluctuating stochastic processes.

Quantities that characterize the mean-variance optimal wealth evolution of a particular company are extremely difficult to estimate. Our stylized model assumes that each company achieves a mean-variance optimal wealth evolution. Thus, by setting its total issued stock value equal to its wealth, the uncertainties of its economic activities become securitized in the stock. The level of risk aversion applied by the management of the company is a key quantity and cannot be easily dismissed when characterizing the company's mean-variance optimal wealth evolution. This risk aversion level most likely changes over time and most likely changes similarly for most companies. Therefore, the stylized model assumes a common risk aversion process for all companies, denoted by $\gamma = \{\gamma_t > 0, t \geq 0\}$.

The other crucial input that plays a significant role in characterizing the mean-variance optimal wealth evolution of a company is the vector of market price of risk processes for the various sources of uncertainty faced by the companies. Also these processes most likely change over time and are difficult to estimate. All companies are exposed to similar prices for raw materials, energy, labor, etc. Consequently, common market price of risk processes can be assumed in the stylized model. Additionally, one could argue that capital is most likely flowing to those business opportunities facing slightly higher market prices of risk and is avoiding investments with lower market prices of risk, thus, reducing through 'demand pressure' higher market prices of risk. Therefore, market price of risk processes can be assumed to be equal in our stylized model, but to fluctuate over time. The common market price of risk process $\theta = \{\theta_t, t \geq 0\}$ is then assumed to denote the market price of risk for each of the independent sources of uncertainty in the stylized model.

We then assume that, in its mean-variance wealth optimization, the j-th company applies the risk aversion γ_t at time t to invest through its management activities a fraction $\frac{1}{\gamma_t}$ in its 'own' GP, denoted by $S_t^{j,GP}$. This is the GP for the 'investment universe' determined by the business opportunities and activities of the j-th company. As is well-known, the company then holds the fraction $1 - \frac{1}{\gamma_t}$ of its wealth in units of the risk-free asset; see e.g., Campbell & Viceira (2002) or Theorem 11.1.3 in Platen & Heath (2010). The stochastic differential equation (SDE) for the stock price S_t^j is then obtained in the form

$$\frac{dS_t^j}{S_t^j} = \frac{1}{\gamma_t} \frac{dS_t^{j,GP}}{S_t^{j,GP}} + \left(1 - \frac{1}{\gamma_t}\right) r_t dt, \tag{D.1}$$

where r_t denotes the risk free rate.

To model the uncertainties faced by the j-th company, $j = (j_1, j_2, j_3, j_4)$, its wealth is driven by the specific uncertainty W^{j_1,j_2,j_3,j_4}; the uncertainty W^{j_1,j_2,j_3}, typical for the industrial grouping the company belongs to; the uncertainty W^{j_1,j_2}, specific for the country where the company is located; and the uncertainty W^{j_1}, typical for the region of the company's country. Here W^{j_1,j_2,j_3,j_4}, W^{j_1,j_2,j_3}, W^{j_1,j_2} and W^{j_1} are independent standard Brownian motions. The SDE for the 'own' GP of the j-th company follows then from Theorem B.1 in Appendix B in the form

$$\begin{aligned}\frac{dS_t^{j,GP}}{S_t^{j,GP}} &= r_t dt + \theta_t(\theta_t dt + dW_t^{j_1}) + \theta_t(\theta_t dt + dW_t^{j_1,j_2}) \\ &+ \theta_t(\theta_t dt + dW_t^{j_1,j_2,j_3}) + \theta_t(\theta_t dt + dW_t^{j_1,j_2,j_3,j_4}),\end{aligned} \tag{D.2}$$

for $t \geq 0$. We highlight that the risk-free asset is included here in the 'investment universe' of the j-th company when forming its 'own' GP. When we form the GP of a given set of stocks, we do not include the risk-free asset in this investment universe, since we aim in this case to identify the GP for the stocks only.

Following (D.1) and (D.2) the j-th cum-dividend stock value, with $j = (j_1, j_2, j_3, j_4)$, satisfies the SDE

$$\begin{aligned}\frac{dS_t^{j}}{S_t^{j}} &= r_t dt + \frac{1}{\gamma_t}\Big(\theta_t(\theta_t dt + dW_t^{j_1}) + \theta_t(\theta_t dt + dW_t^{j_1,j_2})) \\ &+ \theta_t(\theta_t dt + dW_t^{j_1,j_2,j_3}) + \theta_t(\theta_t dt + dW^{j_1,j_2,j_3,j_4})\Big).\end{aligned} \tag{D.3}$$

Based on the SDE (D.3) and the weights (3.5), the HWI under the stylized model has the return process with stochastic differential

$$\begin{aligned}\frac{dS_t^{HWI}}{S_t^{HWI}} &= r_t dt + \frac{1}{\gamma_t}\frac{1}{M_t}\sum_{j_1=1}^{M_t}\Bigg(\theta_t(\theta_t dt + dW_t^{j_1}) \\ &+ \frac{1}{M_t^{j_1}}\sum_{j_2=1}^{M_t^{j_1}}\Bigg(\theta_t(\theta_t dt + dW_t^{j_1,j_2}) \\ &+ \frac{1}{M_t^{j_1,j_2}}\sum_{j_3=1}^{M_t^{j_1,j_2}}\Bigg(\theta_t(\theta_t dt + dW_t^{j_1,j_2,j_3}) \\ &+ \frac{1}{M_t^{j_1,j_2,j_3}}\sum_{j_4=1}^{M_t^{j_1,j_2,j_3}}\theta_t(\theta_t dt + dW_t^{j_1,j_2,j_3,j_4})\Bigg)\Bigg)\Bigg)\end{aligned} \tag{D.4}$$

for $t \geq 0$ with $S_0^{HWI} > 0$.

By application of the Itô formula it is straightforward to show that the j-th stock, when denominated in units of the HWI, has zero drift. That is, we have for the benchmarked j-th stock $\hat{S}_t^j = \frac{S_t^j}{S_t^{HWI}}$ the SDE

$$\begin{aligned} \frac{d\hat{S}_t^j}{\hat{S}_t^j} &= \sum_{k_1=1}^{M_t} (\psi_t^{j,k_1} dW_t^{k_1} + \sum_{k_2=1}^{M_t^{k_1}} (\psi_t^{j,k_1,k_2} dW_t^{k_1,k_2} \\ &+ \sum_{k_3=1}^{M_t^{k_1,k_2}} (\psi_t^{j,k_1,k_2,k_3} dW_t^{k_1,k_2,k_3} + \sum_{k_4=1}^{M_t^{k_1,k_2,k_3}} \psi_t^{j,k_1,k_2,k_3,k_4} dW_t^{k_1,k_2,k_3,k_4}))) \end{aligned} \tag{D.5}$$

with

$$\psi_t^{j,k_1,\ldots,k_n} = \begin{cases} \frac{1}{\gamma_t}\theta_t \Big(1 - \frac{1}{M_t^{j_1} \ldots M_t^{j_1,\ldots,j_{n-1}}}\Big) & \text{for } k_i = j_i \text{ for all } i \in \{1,\ldots,n\} \\ -\frac{1}{\gamma_t}\theta_t \frac{1}{M_t^{j_1} \ldots M_t^{j_1,\ldots,j_{n-1}}} & \text{otherwise} \end{cases} \tag{D.6}$$

for $n \in \{1, 2, 3, 4\}$ with $M_t^{j_0} = M_t$, where we again use the previous notation. Since the SDE for the benchmarked j-th stock is driftless, it is a local martingale. This means that according to Theorem 3.1 and Theorem 4.1 in Filipović & Platen (2009), the HWI is the GP of the given stylized hierarchical market model when setting $S_0^{HWI} = S_0^{GP} = 1$. This remarkable fact provides us with some extra intuition in understanding why hierarchical diversification works well in practice despite the fact that the market price of risk processes and the risk aversion processes may fluctuate significantly and are difficult to estimate.

References

Arnott, R., J. Hsu, & P. Moore (2005). Fundamental indexation. *Financial Analyst Journal* **61**(2), 83–99.

Bai, J. & S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica* **70**(1), 191–221.

Best, M. J. & R. R. Grauer (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results. *Review of Financial Studies* **4**(2), 315–342.

Campbell, J. & L. Viceira (2002). *Strategic Asset Allocation, Portfolio Choice for Long Term Investors.* Oxford, Oxford University Press.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance* **52**(1), 57–82.

Chopra, V. K. & W. Ziemba (1993). The effect of error in means, variances and covariances in optimal portfolio choice. *Journal of Portfolio Management* **19**(1), 6–11.

Choueifaty, Y. & Y. Coignard (2008). Towards maximum diversification. *Journal of Portfolio Management* **35**(1), 40–51.

Chow, T., J. Hsu, V. Kalesnik, & B. Little (2011). A survey of alternative equity index strategies. *Financial Analyst Journal* **5**(67), 37–57.

Clarke, R., H. Silva, & S. Thorley (2011). Minimum-variance portfolio composition. *Journal of Portfolio Management* **37**(2), 31–45.

Davison, A. & D. Hinkley (2007). *Bootstrap Methods and their Applications*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

DeMiguel, V., L. Garlappi, & R. Uppal (2009). Optimal versus naive diversification: How inefficient is the $1/n$ portfolio strategy? *Review of Financial Studies* **22**(5), 1915–1953.

Du, K. & E. Platen (2016). Benchmarked risk minimization. *Mathematical Finance* **26**(3), 617–637.

Fama, E. F. (1970). Efficient capital markets: a review of theory and empirical work. *The Journal of Finance* **25**(2), 383–417.

Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance* **46**(5), 1575–1617.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* **49**(3), 283–306.

Fama, E. F. & K. R. French (1992). The cross-selection of expected stock returns. *The Journal of Finance* **XLVII**(2).

Fernholz, E. R. (2002). *Stochastic Portfolio Theory*. Springer.

Filipović, D. & E. Platen (2009). Consistent market extensions under the benchmark approach. *Mathematical Finance* **19**(1), 41–52.

Gander, P., D. Leveau, & T. Pfiffner (2013). Diversification - a multi-facetted concept. *1741 Asset Management Research Note Series 1/2013*.

Harvey, C., Y. Liu, & H. Zhu (2016). ... and the cross-section of expected returns. *Review of Financial Studies* **29**(1), 5–68.

Heston, S. L., M. Loewenstein, & G. A. Willard (2007). Options and bubbles. *Review of Financial Studies* **20**(2), 359–390.

Kan, R., X. Wang, & G. Zhou (2016). On the value of portfolio optimization in the presence of estimation risk: The case with and without risk-free asset. *Working paper. University of Toronto*.

Kan, R. & G. Zhou (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis* **42**, 621–656.

Karatzas, I. & C. Kardaras (2007). The numeraire portfolio in semimartingale financial models. *Finance and Stochastics* **11**(4), 447–493.

Karatzas, I. & S. E. Shreve (1998). *Methods of Mathematical Finance*. Springer.

Kardaras, C. & E. Platen (2010). Minimizing the expected market time to reach a certain wealth level. *SIAM Journal of Financial Mathematics* **1**(1), 16–29.

Kelly, J. R. (1956). A new interpretation of information rate. *Bell Systems Technology Journal* **35**, 917–926.

Leote, R., X. Lu, & P. Moulin (2012). Demystifying equity risk-based strategies: A simple alpha plus beta description. *Journal of Portfolio Management* **38**, 56–70.

Loewenstein, M. & G. A. Willard (2000). Local martingales, arbitrage, and viability: Free snacks and cheap thrills. *Econometric Theory* **16**(1), 135–161.

Ludvigson, S. & S. Ng (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics* **83**(1), 171–222.

MacLean, L., E. Thorp, & W. Ziemba (2011). *The Kelly Capital Growth Investment Criterion*. World Scientific.

Maillard, S., T. Roncalli, & J. Teiletche (2010). The properties of equally weighted risk contribution portfolios. *Journal of Portfolio Management* **36**(4), 60–70.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investment*. Wiley, New York.

Mode, E. (1966). *Elements of Probability and Statistics*. Prentice-Hall, Englewood Cliffs, N.J.

Oderda, G. (2015). Stochastic portfolio theory optimization and the origin of rule-based investing. *Quantitative Finance* **15**(8), 1259–1266.

Okhrin, Y. & W. Schmid (2006). Distributional properties of portfolio weights. *Journal of Economics* **134**, 235–256.

Platen, E. (2011). A benchmark approach to investing and pricing. in:. *MacLean, L.C. and Thorp, E. O. and Ziemba, W. (2011), The Kelly Capital Growth Investment Criterion. World Scientific*, 409–425.

Platen, E. & D. Heath (2010). *A Benchmark Approach to Quantitative Finance*. Springer Finance. Springer.

Platen, E. & R. Rendek (2012). Approximating the numéraire portfolio by naive diversification. *Journal of Asset Management* **13**(1), 34–50.

Plyakha, Y., R. Uppal, & G. Vilkov (2014). Equal or value weighting? Implications for asset pricing tests. Technical report, SSRN 1787045.

Reuters (2008). Datastream global equity indices. User Guide. Issue 5. *Thomson Reuters*.

Rosenberg, B. & V. Marathe (1976). Common factors in security returns: Microeconomic determinants and macroeconomic correlates. *University of*

California Institute of Business and Economic Research, Research Program in Finance, Working paper No. 44.

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

Van Dijk, M. (2011). Is size dead? A review of the size effect in equity returns. *Journal of Banking and Finance* **35**(12), 3263–3274.