Discrete hierarchy of sizes and performances in the exchange-traded fund universe

B. Vandermarliere\textsuperscript{a,b,*}, J. Ryckebusch\textsuperscript{b}, K. Schoors\textsuperscript{b}, P. Cauwels\textsuperscript{c}, D. Sornette\textsuperscript{c}

\textsuperscript{a} Department of General Economics, Ghent University, Belgium
\textsuperscript{b} Department of Physics and Astronomy, Ghent University, Belgium
\textsuperscript{c} Department of Management, Technology and Economics, ETH Zürich, Switzerland

**Highlights**
- The size distribution of exchange-traded funds exhibits a discrete hierarchy of sizes.
- This is shown with a spectral analysis of residuals and by studying the generalized derivative.
- Several explanatory mechanisms are offered for this discrete scale invariance.
- Larger ETFs exhibit a significantly stronger similarity compared with smaller ETFs.

**Abstract**
Using detailed statistical analyses of the size distribution of a universe of equity exchange-traded funds (ETFs), we discover a discrete hierarchy of sizes, which imprints a log-periodic structure on the probability distribution of ETF sizes that dominates the details of the asymptotic tail. This allows us to propose a classification of the studied universe of ETFs into seven size layers approximately organized according to a multiplicative ratio of 3.5 in their total market capitalization. Introducing a similarity metric generalizing the Herfindahl index, we find that the largest ETFs exhibit a significantly stronger intra-layer and inter-layer similarity compared with the smaller ETFs. Comparing the performance across the seven discerned ETF size layers, we find an inverse size effect, namely large ETFs perform significantly better than the small ones both in 2014 and 2015.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction
An exchange-traded fund (ETF) can be thought of as a portfolio of stocks, commodities, or bonds, which is traded like stocks on stock exchanges. Exchange-traded funds have been made available as investment funds in the US in the early nineties and in Europe in the late nineties. Ever since, ETFs have emerged as a very important investment vehicle attracting ever increasing volumes of capital. Its attractiveness is partly due to the relatively low management and transaction costs involved, an element that is particularly important in times of low yields and low interest rates. Exchange-traded funds represent an increasingly important investment vehicle with potential hazards for systemic risk and possible dangerous menaces for the financial system [1–3]. For example, it has been shown that arbitrageurs can contribute to cross-sectional return co-movement via ETF arbitrage. The presence of a stock in ETFs increases return co-movement at both the fund and...
the stock levels, where the effect is strongest among small and illiquid stocks [4]. These days, ETFs come in many different types of flavours [5]. For example, the degree of active management varies very much from one ETF to another.

The focus of this paper is on establishing a taxonomy of the equity ETF landscape on the basis of their size. From our discussion we exclude leveraged ETFs and ETFs holding bonds and commodities, mainly to not overly complicate the analysis. As our focus is on determining the robust and stylized features of the equity ETF landscape using size, we do not segregate by types of ETFs, for example in terms of managed versus active versus passive, or index tracking ETFs.

Size distributions often carry information about the underlying dynamics of a system. The analysis of the distribution of the equity ETF sizes described below discloses some features that suggest departures from a simple power-like tail. The occurrence of a fat tail in the distribution of ETF sizes does not really come as a surprise given the well-documented approximate Zipf-law distribution of firm capitalizations [6]. The fact, however, that there are strong indications that the tail is decorated with some log-periodic structure is remarkable. As this structure is connected with discrete scale invariance, one can infer some interesting constraints on the underlying dynamics of the equity ETF universe. Accordingly, we consider the disclosed log-periodic structure in the size distribution as a natural tool for classification of the universe of ETFs. The inferred classification of the ETFs in several size layers is used to study various economic indicators. We address questions like: ‘How similar are the various kinds of ETFs?’; ‘How do ETFs distribute their holdings over the wide landscape of possible holdings?’; and ‘Is there a connection between the ETF size and their performance?’? These questions are naturally motivated by the existence of the size effect, exploited in the famous Fama–French 3 factor model [7] that also addresses the fundamental issues of the relationships between diversification and performance.

The remainder of this paper is organized as follows. In Section 2 we present our empirical analysis of the equity ETF size distribution. We start off (Section 2.1) with providing details of the ETF size data used and with performing a maximum-likelihood fit to their distribution. This reveals indications for an interesting discrete hierarchical structure in the ETF size distribution that is discussed in more detail in Section 2.2. In order to put this structure on more solid grounds and to get better hold on the disclosed periodicity in the size distribution, in Section 2.3 we pursue a detailed analysis of the ETF size distribution using kernel density estimation and Lomb periodograms. In Section 2.4 we sketch some dynamical features of the ETF universe that may give rise to the observed hierarchical structure. We work out in detail how a model based on nonextensive (or, Tsallis) statistical mechanics, a current generalization of Boltzmann–Gibbs (BG) statistical mechanics, can give rise to the discerned oscillatory structures in the ETF size distribution. The basic premises of the proposed model is that the system consisting of all ETFs operates as an open system in a capital reservoir. The size of the ETF system is subject to capital exchange with the reservoir, whereby there is a mechanism of both preferential attachment and growth. In Section 3 we introduce a classification into seven layers of the equity ETFs based on the discerned log-periodic hierarchy. We also explore how the economic properties vary over the various size layers. Thereby, we investigate the intra-layer and inter-layer similarities (Section 3.1), the variations in the stock holding ubiquity and capitalization over the different layers (Section 3.2), and the connection between layer and performance (Section 3.3). Our conclusions are drawn in Section 4.

2. Analysis of the distribution of ETF sizes

2.1. Distribution of total net asset values of ETFs

At the end of 2014, we collected data for all exchange-traded funds (ETFs) labelled as equity ETFs from Thomson Reuters Eikon. This resulted in a set of 479 ETFs for which we obtained the total net assets and the entire composition of their portfolios. In total, this comprised 11,643 different assets and about 100,000 positions, for a total net assets over all ETFs of $1.399 \times 10^{12}$. Fig. 1 includes the complementary cumulative distribution function (CCDF) of the total net assets of ETFs, i.e., the fraction of ETFs of total net assets larger than or equal to $S$. Also shown is the CCDF of the log-normal that best fits the data, as obtained by the maximum-likelihood method. The probability density function (PDF) of the log-normal law $\ln N$ reads

$$\ln N(\mu_L, \sigma_L^2) = \frac{1}{\sqrt{2\pi \sigma_L^2}} e^{-\frac{(\ln x - \mu_L)^2}{2\sigma_L^2}},$$

with $\mu_L$ the location and $\sigma_L$ the scale parameter whose maximum-likelihood estimates are $\hat{\mu}_L = 18.7$ and $\hat{\sigma}_L = 2.24$. This corresponds to the mode (or most probable) ETF size of approximately $130 \times 10^{10}$ US$ and a mean ETF size of $1.6 \times 10^{10}$ US$. The much larger value of the mean compared to the mode reflects the existence of a very strong “fat tail” quantified by $\hat{\sigma}_L$.

When referring to fat tails, it is often convenient to use power law distributions. The tail of a log-normal distribution with large variance (as found here) is difficult to distinguish from a power law distribution (see e.g. Ref. [8] and Section 4.1.3 of Ref. [9]). Indeed, visually, the tail of the empirical CCDF shown in Fig. 1 seems roughly compatible with an asymptotic power law with an exponent of about 1 (Zipf’s law). Such an approximate asymptotic Zipf’s law has been documented for the distribution of firm sizes [6]. The fact that a similar approximate behaviour in the asymptotic tail is observed for the distribution of ETF sizes is not really a surprise as it can be expected from the presence of two joint and mutually reinforcing mechanisms. First, it is well known that the size of individual firms approximately obeys Zipf’s law [6,10–13]. This result is robust [14] and has been confirmed for different countries [10] and for several measures of firm size including...
دریافت فوری
متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات