Fractality in market risk structure: Dow Jones Industrial components case
Ladislav Kristoufek
Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodárenskou vezi 4, Prague 8 182 08, Czech Republic

1. Introduction

The Global Financial Crisis and its aftermath have shown that the notion of systematic (and also systemic) risk is not vain. During the critical periods, practically all stocks kept losing their value, and the losses and risk could not have been diversified away. Dating back to Markowitz [1], diversification, i.e. lowering the portfolio risk by its enlarging, is tightly connected to the correlation structure of the market. If the assets are all strongly correlated, they will rise and fall together. Only a single asset moving against the market can lower the portfolio risk markedly. Connection between the market risk and individual assets’ risk is nicely captured by the capital asset pricing model (CAPM), which has become one of cornerstones of the modern financial economics since its introduction in the 1960s [2–4]. The model describes the relationship between an asset and market in a simple linear manner. Regardless its simplicity, the model has several intuitive but important implications. The most important one from the portfolio construction perspective is the existence of the market (systematic) risk that cannot be diversified away. In words, as most assets are at least somehow connected to the global market movements, this principal component cannot be gotten rid of as it is common to all said assets. Another appealing outcome of the model’s simplicity is that it is described by only two parameters one of which – \( \beta \) – identifies the asset as an aggressive one, a defensive one, or a market-following one. However, if the past decade has taught the financial theorist and practitioners anything, market participants can perceive an asset behavior differently. There are different types of investors with different trading strategies and different investment horizons and it is hard to believe they all agree on risk specifics of a given asset as called for by the efficient market hypothesis [5,6]. Quite the contrary, it is more realistic to assume that the market participants differ as well as their expectations as asserted by the fractal market hypothesis [7,8]. Our main motivation is thus to inspect the stock markets via the capital asset pricing model with a special focus on scale specifics of the model. To do so, we utilize the quite newly proposed regression frameworks build on the fractal methods, specifically the detrended cross-correlation analysis and the detrending moving-average cross-correlation analysis. In addition, we provide a novel approach towards the statistical significance of the scale variability.

The paper is organized as follows. The next section describes the capital asset pricing model in detail and focuses on the fractal methods and how to approach statistical inference in the CAPM setting. The following section introduces the analyzed data and explains the specific choices. The last section presents the results, provides economic interpretation and sketches some further venues into the topic.
2. Methods

2.1. Capital asset pricing model

The capital asset pricing model (CAPM) is one of the important building blocks of the modern financial economics as it describes the relationship between risk and return in the rational equilibrium market. Building on the Markowitz modern portfolio theory [19], the CAPM was developed by several authors independently of one another [2–4]. For individual assets, the model is stated as

$$E(R_i) = R_f + \beta_i (E(R_m) - R_f)$$

(1)

where $E(R_i)$ is the expected return of asset $i$, $R_f$ is the risk-free rate, and $E(R_m)$ is the expected market return. $\beta_i$ is the crucial parameter of the model and it can be interpreted as a sensitivity of the asset return to the market return (both cleared by the risk-free rate). With respect to Refs. [2–4], it can be shown that

$$\beta_i = \frac{\rho_{im} \sigma_i \sigma_m}{\text{cov}(R_i, R_m)} = \frac{\text{cov}(R_i, R_m)}{\text{var}(R_m)}$$

(2)

where $\rho_{im}$ is the correlation between $R_i$ and $R_m$, $\sigma_i$ is the standard deviation of $R_i$, and $\sigma_m$ is the standard deviation of $R_m$. Note that the representation of $\beta_i$ on the right-hand side of Eq. (2) is the same as the least squares estimator of the simple regression

$$R_i - R_f = \alpha_i + \beta_i (R_m - R_f) + u_i$$

(3)

where $u_i$ is the error term and $\alpha_i$ is a deviation from the equilibrium return. The $\beta$ parameter can be thus easily estimated using the least squares methodology. In general, $\beta$ can attain any value but cases when $\beta \leq 0$ are rare (assets moving against the market, or short positions). Apart from this unlikely case, there are three interesting cases:

- $0 < \beta < 1$: defensive assets, which move in the same direction as the market but have lower volatility
- $\beta = 1$: assets following the market, e.g. market-index-based assets, or strong contributors to a market index
- $\beta > 1$: aggressive assets, which move in the same direction as the market but with higher volatility

The definition of $\beta$ and the CAPM construction imply that the market return $R_m$ and error term $u$ in Eq. (3) are uncorrelated. This allows to split the asset variance (risk) $\sigma_i^2$ into two orthogonal components as

$$\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma_{u_i}^2$$

(4)

where $\sigma_m^2$ is the market variance and $\sigma_{u_i}^2$ is the error term variance [10]. The component $\beta_i^2 \sigma_m^2$ is called the systematic risk associated with the market and it cannot be eliminated (i.e. by diversification). The error term variance is usually referred to as the idiosyncratic risk (or specific risk or unsystematic risk) and this one can be eliminated (or at least mitigated). High $\beta$ assets can thus return high profits in the growing market but they do not contribute to risk optimization. Therefore, a high $\beta$ portfolio is possibly very profitable but also very risky. Low $\beta$ assets thus help diversifying the risk.

The capital asset pricing model is connected to an understanding of a market as an efficient one with respect to the efficient market hypothesis (EMH) [5,6,11–14], specifically to one of its assumptions that the investors are homogeneous in their expectations and have a common investment horizon [15]. However, observing reality suggests that investors are far from homogeneous and they differ in their investment horizons, ranging from algorithmic and noise trading (with very short horizons in a span of second fractions) to pension funds (with long investment horizons of several years or even decades). Specifically, we want to examine whether an asset can be seen in a different perspective (in the CAPM sense) by a short-term investor and a long-term investor, i.e. whether the asset $\beta$s can be different for different investment horizons. For this purpose, we utilize the regression frameworks build with scaling and fractality in mind – fractal regressions based on the detrended cross-correlation analysis and the detrending moving-average cross-correlation analysis.

2.2. Fractal regressions

The capital asset pricing model is based on a bivariate relationship between an asset’s return (corrected by the risk-free rate) and a market return (also corrected by the risk-free rate). The model can be thus ideally studied by quite recently proposed regression frameworks based on the detrended fluctuation analysis (DFA) and detrending moving average (DMA) procedures [16,17]. Here, the methods are not only useful due to their robustness to persistence, short-range correlations and heavy tails [18–20], but specifically for their ability to study the relationship between series at different scales so that we can distinguish between short-term and long-term investment horizons. This leads to possible findings such that a specific stock is considered to be an aggressive investment for short-term investors but a defensive (safe) investment for long-term investors. Such results would support claims of the fractal markets hypothesis (FMH) [7,8,21–23] as opposed to the efficient market hypothesis (EMH) [5,6,11–14], which assumes that all investors agree on the riskiness of a specific asset.

The two fractal regression frameworks are based on the methods usually used for detecting fractal structure and long-range dependence properties of analyzed series – specifically the detrended fluctuation analysis (DFA) [24] and the detrending moving average (DMA) [25,26]. Both methods have been generalized for analysis of bivariate properties of the series which has given rise to the detrended cross-correlation analysis (DCCA) [27–29] and the detrending moving-average cross-correlation analysis (DMCA) [30,31]. Combination of DFA and DCCA allowed for an introduction of the DCCA-based correlation coefficient which describes correlations between series at different scales [32]. In the same logic, DMA and DMCA have been combined to form the DMCA-based correlation coefficient [33], which surpasses the original DCCA-based method under various specifications of long-range dependence [34]. These scale-specific correlation coefficients have been extensively used in empirical studies across disciplines [28,35–46]. The regression frameworks are only a step away from the correlation analysis.

The DCCA and DMCA-based correlation coefficients are based on a simple idea of substituting the covariance and variances (standard deviations) in the definition of correlation coefficient with the scale-specific covariances and variances obtained during the DFA/DCCA and DMA/DMCA procedures. Without a need to eventually arrive at the Hurst exponent given by DFA and DMA, we can use the fluctuation functions obtained during the procedures. Specifically for the DFA procedure, we select the scale $s$ and split the profile series (integrated demeaned original series) into boxes of given length. In each box, a polynomial trend (usually linear as in our application) is fitted, residuals are obtained and mean squared error is calculated. The mean squared errors are then averaged over all boxes of size $s$ and to get $F_{DFA}(s)$. For the bivariate

---

1 The $\alpha$ parameter can be used for investment decisions as $\alpha > 0$ suggests overpricing of the asset and $\alpha < 0$ suggest underpricing of the asset. However, we focus primarily on the $\beta$ parameter here and leave possible $\alpha$ discussions for future research.

2 If the series is not divisible by $s$, we divide the series into boxes from the beginning and from the end, i.e. obtained twice as many boxes compared to the divisible case.
case and DCCA, the residuals in each box are obtained in the same manner only for each series and an average of their product is retrieved for each box. This is again averaged over all boxes of size $s$ so that we get $F_{XY}^2$ (DCCA)($s$). The DMA/DMCA procedures are more straightforward as we only construct a moving average of size $\lambda$ for the analyzed series and calculate the mean squared error from this moving average to get $F_{XY}^2$ (DMA). For the bivariate case, we calculate the average product of the deviations from the moving averages of the two series (with the same window size $\lambda$) to get $F_{XY}^2$ (DMCA). It has been shown that the centered moving average delivers the best results [47]. The correlation coefficients are then simply [32,34]
\[
\hat{p}_{XY}^{DFA} (s) \sim \frac{F_{XY, DCCA}^2 (s)}{\sqrt{F_{X, DFE}^2 (s) F_{Y, DFE}^2 (s)}},
\]
\[
\hat{p}_{XY}^{DMA} (\lambda) \sim \frac{F_{XY, DMA}^2 (\lambda)}{\sqrt{F_{X, DMA}^2 (\lambda) F_{Y, DMA}^2 (\lambda)}}.
\]

Even though the correlation coefficient is useful in uncovering whether the two series are related or not, it does not quantify the actual effect. This holds for the scale-specific correlation coefficients $\hat{p}_{XY}^{DFA} (s)$ and $\hat{p}_{XY}^{DMA} (\lambda)$ as well. The regression framework takes one variable, say $Y$, as a response (dependent) variable, and one variable, say $X$, as an impulse (independent) variable. Their relationship is described by a model
\[
Y = \alpha + \beta X + u
\]
where $u$ is the error term, $\alpha$ is the model intercept (the expected value of $Y$ when $X = 0$) and $\beta$ is the expected change in $Y$ when $X$ changes by one (given the set of basic assumptions). In this simple regression case, the $\beta$ parameter is estimated using the (ordinary) least squares procedure as
\[
\hat{\beta} = \frac{\sum_{i=1}^{T} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{T} (x_i - \bar{x})^2} \sim \frac{\text{cov}(X, Y)}{\text{var}(X)}
\]
for time series of length $T$ with $\bar{x}$, $\bar{y}$ being averages and $\sim$ symbolizing asymptotic equivalence. The least squares estimator of the correlation coefficient is thus simply the ratio between covariance of $X$ and $Y$, and variance of $X$, and it can be thus seen as a rescaled correlation coefficient. This simple representation of the estimator has been utilized to introduce the estimators of scale-specific effects of $X$ on $Y$ based on DFA/DCCA and DMA/DMCA in a similar logic as $\hat{p}_{XY}^{DFA} (s)$ and $\hat{p}_{XY}^{DMA} (\lambda)$. Specifically, the estimators for given scales $s$ or $\lambda$ are given as [16,17]
\[
\hat{\beta}_{XY}^{DFA} (s) \sim \frac{F_{XY, DCCA}^2 (s)}{\sqrt{F_{X, DFE}^2 (s) F_{Y, DFE}^2 (s)}},
\]
\[
\hat{\beta}_{XY}^{DMA} (\lambda) \sim \frac{F_{XY, DMA}^2 (\lambda)}{\sqrt{F_{X, DMA}^2 (\lambda) F_{Y, DMA}^2 (\lambda)}}.
\]

Compared to the fractal correlation coefficients $\hat{\beta}_{XY}^{DFA} (s)$ and $\hat{\beta}_{XY}^{DMA} (\lambda)$, which tell the correlation at a specific scale (i.e. between $-1$ and $1$), the fractal regression estimates give the actual effects between $X$ and $Y$. Even though it might seem as a negligible difference or even not a very useful upgrade of the correlation coefficients, in specific case, we are interested in the actual effect between $X$ and $Y$. This is the case of the capital asset pricing model investigated here as it is important to know whether $\beta = 1$ or $\beta > 1$, or $0 < \beta < 1$ as each of these cases has its own interpretation, which could not be uncovered using the correlation coefficients.

2.3. Statistical inference and parameters selection

The above described methods provide a range of $\beta$ estimates for a selection of scales. However, it is not straightforward to identify whether the variability of such estimates is high enough so

that we could claim that one global parameter is not enough to describe the relationship between the variables of interest, in our case the asset return and the market return. To be able to do so, we apply the following procedure:

1. Estimate Eq. (3) using the (ordinary) least squares estimator and obtain estimates $\hat{\alpha}$, $\hat{\beta}$, and residuals $\hat{u}_i$.
2. Create a surrogate series of $\hat{u}_i$ using Theiler’s phase randomization [48], specifically by randomizing the phases of the Fourier coefficients while keeping the amplitude of the series. This gives us a surrogate series $\hat{u}_i^{\text{sr}}$ with the same spectrum and the same distribution as the original series.
3. Reconstruct the returns series as $(R_i - R_f) = \hat{u}_i + \hat{\beta}_i(R_m - R_f) + \hat{\alpha}_i(R_m - R_f)$
4. Estimate $\hat{\beta}_{XY}^{DFA} (s)$ and $\hat{\beta}_{XY}^{DMA} (\lambda)$ on the surrogate return series $(R_i - R_f)$ from the previous step for a given range of $s$ and $\lambda$.
5. Repeat steps 2, 3, and 4 many times and keep the results.
6. Based on the previous step, construct confidence intervals for the null hypothesis of a single global $\beta$ with no scale differences.
7. Compare the original $\hat{\beta}_{XY}^{DFA} (s)$ and $\hat{\beta}_{XY}^{DMA} (\lambda)$ estimates with the constructed confidence intervals.

Such procedure gives the approach statistical validity. If the estimated $\hat{\beta}_{XY}^{DFA} (s)$ and $\hat{\beta}_{XY}^{DMA} (\lambda)$ fall outside of the confidence intervals, we detect heterogeneity in investors’ perception of risk with respect to investment horizons.

There are several parameters in the estimation procedures that need to be specified. For the DCCA-based regression, we estimate $\beta$s from $s_{\text{min}} = 10$ to $s_{\text{max}} = 500$ with a step of 10. For the DMCA-based regression, we use similar setting with $\lambda_{\text{min}} = 11$ and $\lambda_{\text{max}} = 501$ with a step of 10 due to the central moving averages. For the construction of confidence intervals, we repeat the procedure 333 times for each setting and construct 95% confidence intervals, i.e. the 2.5% and 97.5% percentiles.

3. Data

We study scaling of the capital asset pricing model $\beta$ parameter on a sample of the Dow Jones Industrial Average (DJIA) index components. The index comprises 30 important stocks3 traded on the New York Stock Exchange (NYSE) and the NASDAQ. The stocks are analyzed between 2 January 2009 and 30 June 2017 (2139 daily observations).4 The specific period has been selected for two reasons. First, all the index components as of 30 June 2017 have existed and have been traded from 2 January 2009 onwards. And second, the beginning of 2009 marks the stock market turning point after the Global Financial Crisis (we wanted to avoid the results being influenced by such a structural break).

The capital asset pricing model for each stock uses three different variables. Apart from the stock returns, we need the market return and the risk-free rate. For the market return, we use the DJIA (ticker DJI) returns, and for the risk-free rate, we use the CBOE (Chicago Board Options Exchange) Interest Rate 10 Year Treasury Note Index (ticker TNX), which is based on the yield-to-maturity on the most recently auctioned 10-year Treasury note. This interest rate is quoted in annual interest rate percentage

3 By the end of June 2017, the index comprised of 3M (MMM), American Express (AXP), Apple (AAPL), Boeing (BA), Caterpillar (VAR), Chevron (CVX),Cisco (CSCO), Coca-Cola (KO), Disney (DIS), E I du Pont de Nemours and Co (DD), Exxon Mobil (XOM), General Electric (GE), Goldman Sachs (GS), Home Depot (HD), IBM (IBM), Intel (INTC), Johnson & Johnson (JNJ), JPMorgan Chase (JPM), McDonald’s (MCD), Merck (MRK), Microsoft (MSFT), Nike (NKE), Pfizer (PFE), Procter & Gamble (PG), Travelers Companies Inc (TRV), United Technologies (UTX), UnitedHealth (UNH), Verizon (VZ), Visa (V) and Wal-Mart (WMT).
4 All time series were downloaded from the finance.yahoo.com site.
points and as we use daily stock and index data, the interest rates need to be transformed to daily yields. Counting with the 250-day trading year, the (compounded) daily yields are obtained as $R_f = (1 + AY/100)^{1/250} - 1$, where $AY$ is the annual yield in percentage points. All necessary variables for Eq. (3) are thus well defined.

4. Results and discussion

The capital asset pricing model is estimated for all 30 components of the Dow Jones Industrial Average index between 2 January 2009 and 30 June 2017. For each stock, we estimate the model using the DCCA and DMCA-based regression for a range of scales between 10 and 500 trading days (11 and 501 for the DMCA method). To test whether the potential variability of the CAPM $\beta$ across scales is statistically valid, we construct confidence intervals with the null hypothesis of a single global parameter $\beta$. The procedure specifics are described in the Methods section.

The results are summarized in Figs. 1–3. Recall that $\beta > 1$ signifies an aggressive stock, more volatile with hard to diversify risk, whereas $0 < \beta < 1$ is a defensive, less volatile stock that can contribute to portfolio diversification. In the figures, we present the results based on DCCA on the left and the ones based on DMCA on the right for all 30 stocks. The estimated scale-specific $\beta$s are marked by a red line and the confidence intervals are represented by the shaded area. If the red line crosses outside of the shaded area, $\beta^{DCCA}(s)$ or $\beta^{DMCA}(\lambda)$ differs for scales $s$ or $\lambda$, respectively. If the red line remains inside of the shaded area for all the analyzed scales, it suggests that a simple least squares regression is enough for the analysis. From the interpretational side, if the red curve remains inside the shaded are, the risk is perceived the same way for investors of various investment horizons, which is in hand with the efficient market hypothesis. If the red curve goes outside, investors with different investment horizons weigh the stock risk differently, which is more in hand with the fractal market hypothesis.

From the technical perspective, both methods provide very similar results. However, the DMCA-based regression yields much smoother estimates and confidence intervals across scales compared to the DCCA-based method. This is well in hand with some previous results in the literature using and comparing the methods [33,34,49]. Further, we observe that the confidence intervals widen with an increasing scale. This is not surprising given the fact that the estimates are based on less observations for higher scales. Nevertheless, it explains why some results suggest rather erratic behavior at high scales which can easily be due to statistical fluctuations rather than an actual scale dependence of the parameters of interest.

We can see examples of aggressive stocks and defensive stocks as well as market stocks. The market stocks are represented by

![Fig. 1. Scale-specific estimates of the CAPM $\beta$ parameter L. Results for the DCCA-based regression (left) and DMCA-based regression (right) are shown. For the DCCA method, the estimates are presented for scales between 10 and 500 with a step of 10. For the DMCA method, the estimates are shown for scales between 11 and 501 (due to central moving averages). The estimates are represented by the red line. 95% confidence intervals are marked by the shaded area and these are based on 333 surrogate series. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)](image-url)
large companies such as AAPL, CSCO, CVX, HD, IBM, INTC, MMM, MSFT, TRV, UNH, UTX, and XOM. All these share a common $\beta$ around one and this is majorly the case for all studied scales. From the aggressive titles, we have AXP, BA, CAT, DD, DIS, GE, GS, and JPM, i.e. mainly financial institutions but not universally. And as examples of the defensive stocks, we have JNJ, KO, MCD, MRK, PFE, PG, VZ and WMT, i.e. healthcare, consumer goods and services companies. The market following, riskiness and diversification opportunities thus seem quite sector-specific.

Looking at the variability of $\beta$ across scales, we find some variability for practically all the studied titles. However, when these are compared with the confidence intervals, we find statistically significant scale-dependence of $\beta$ only sparsely. For a majority of stocks, the red curve representing the scale-specific $\beta$ estimates remains within the shaded area of the simulated confidence intervals. Therefore, for most stocks, it is enough to describe its riskiness in the sense of the CAPM using only a single global $\beta$ parameter. Nevertheless, there are exceptions. The McDonald’s Corporation (MCD) is a strongly defensive stock with $\beta \approx 0.5$ for low scales. However, for the longer investment horizons, the parameter approaches 0. This deviation is statistically significant for high scales for both utilized methods. In the long-term, MCD does not react to the market trends practically at all and it is thus a strong candidate for the long-term risk diversification in a portfolio. Similar pattern is observed for NIKE, Inc. (NKE), which is a market following asset ($\beta \approx 1$) for short scales but becomes quite defensive ($\beta \approx 0.5$) for higher scales. In the short term, NKE does not provide much diversification potential but in the long term, it does quite significantly. Then we have E. I. du Pont de Nemours and Company (DD) and General Electric Company (GE), two aggressive stocks with a long-term $\beta$ around 1.25 which, however, statistically significantly deviates upwards for the lower scales, specifically between 1.4 and 1.75. These companies thus seem as a good short-term risky investment. From the perspective of portfolio diversification, either short-term or long-term, these stocks are not good candidates. For several other companies, there are episodes of red curves slightly escaping the shaded area but no strong patterns emerge. (Fig. 2)

In summary, we have examined the Dow Jones Industrial Average index components with respect to the capital asset pricing model, specifically its scaling properties in the sense of different investment horizons. To do so, we have used the novel methods of fractal regressions based on the detrended fluctuation analysis and the detrending moving average. We report three standard groups of stocks – aggressive, defensive and market-following – which are rather uniformly represented (as expected due to the model and mainly market return specification). For most of the stocks, the $\beta$ parameter of the CAPM does not vary significantly across scales. There are two groups of exceptions. One of aggressive stocks which
are even more aggressive for short investment horizons. These do not provide portfolio diversification potential but allow for high profits above the market returns and even more so for short investment horizons. And the other group of more defensive stocks which become very defensive in the long term. These stocks do not deliver short term profits but can serve as strong risk diversifiers. Apart from these direct results, our analysis opens several interesting topics and future research directions, both technical and experimental. First, the DMCA-based regression provides more stable and smoother results than the DCCA-based one. Even though this is not a completely new result (see Refs. [33,34,49]), it highlights the need for further comparison of statistical properties of the methods. Second, the constructed confidence intervals are very wide for some stocks. This stresses the need for proper statistical analysis of experimental studies utilizing fractal regressions and/or DCCA-based (and other) correlation coefficients. Even though the estimates by themselves might seem to vary strongly across scales, it does not say much without properly specified critical values of the opposite case. And third, the fact that most of the stocks report very stable $\beta$ across scales is likely connected to the analyzed period of 2009–2017, which corresponds to a strongly bullish market, i.e. strongly growing. Such stability is well in hand with both the efficient market hypothesis but also with the fractal market hypothesis, which suggests that the growing market is characteristic by no dominant investment horizon and rather uniformly distributed traded activity across the horizons. It will be interesting to study whether the stability is present also during different market states, specifically before and during various recent crises. Our study thus serves as a starting point towards studying the capital asset pricing model from a new perspective.

Acknowledgement

Financial support of the Czech Science Foundation (project 17-12386Y) is gratefully acknowledged.

References