Review of Economics & Finance
Submitted on 14/09/2017
Article ID: 1923-7529-2018-02-16-13
Natascia Angelini, Giacomo Bormetti, Stefano Marmi, and Franco Nardini

Value Matters: The Long-run Behavior of Stock Index Returns[1]

Dr. ***Natascia Angelini***
School of Economics, Management and Statistics, University of Bologna
Via Angherà 22, 40127 Rimini, ITALY
E-mail: natascia.angelini@unibo.it

Prof. Dr. ***Giacomo Bormetti***
Department of Mathematics, University of Bologna
Viale Filopanti 5, 40126 Bologna, ITALY
E-mail: giacomo.bormetti@unibo.it

Prof. Dr. ***Stefano Marmi***
Scuola Normale Superiore and C.N.R.S. UMI 3483, Laboratorio Fibonacci
Piazza dei Cavalieri 7, 56126 Pisa, ITALY
E-mail: stefano.marmi@sns.it

Prof. Dr. ***Franco Nardini*** (Corresponding author)
Department of Mathematics, University of Bologna
Viale Filopanti 5, 40126 Bologna, ITALY
E-mail: franco.nardini@unibo.it

Abstract: We present a simple dynamical model of stock index returns grounded on the ability of the Cyclically Adjusted Price Earning valuation ratio devised by Robert Shiller to predict long-horizon performances of the market. Specifically, within the model returns are driven by a fundamental term and an autoregressive component perturbed by external random disturbances. The autoregressive component arises from the agents' belief that expected returns are higher in bullish markets than in bearish ones. The fundamental value, towards which fundamentalists expect that the current price should revert, varies in time and depends on the initial averaged Price-to-Earnings ratio. We demonstrate both analytically and by means of numerical experiments that the long-run behavior of the stylized dynamics agrees with empirical evidences reported in literature.

Keywords: Fundamental and momentum strategies; Valuation ratios; CAPE; Long-run stock market returns; Value investing

JEL Classifications: D53, G12, G17

[1] NA and FN acknowledge partial support of a "PRIN 2009" grant provided by MIUR (Italian Ministry of Education, University and Research) as part of the research project "Local interactions and global dynamics in economics and finance: models and tools".

1. Introduction

Since the path breaking works by Fama (1965) and Mandelbrot (1963), it is widely accepted that stock prices and stock market indexes behave like random walks. Such a long-lived popularity is supported by two different arguments. The first one is the argument put forth by Fama that financial markets are "informationally efficient". One cannot achieve returns in excess of average market returns on a risk-adjusted basis, given the information available at the time the investment is made. The instantaneous adjustment property of an efficient market implies that successive price changes in individual securities may be assumed independent for any practical purpose; see Fama (1965) and Samuelson (1965). The second argument is the possibility, within the formal framework of stochastic processes, to develop pricing models.

The bubble burst in 1987, the exceptional price boom in the late 1990s, and the subsequent crash lead more and more researchers to cast doubts on the hypothesis' truth. As early as 1988, Campbell and Shiller (1988b) found statistical evidence that "the present value of future dividends is, for each year, roughly a weighted average of moving-average earnings and current real price". This fact has strong implication for the present-value model of stock prices and for recent results that long-horizon stock returns are highly predictable. At the very beginning of 2000 Robert Shiller wrote: "We do not know whether the market level makes any sense, or whether they are indeed the result of some human tendency that might be called irrational exuberance", Shiller (2000).

He reached his conclusion through an innovative test of the appropriateness of prices in the stock market: the Cyclically Adjusted Price Earning (CAPE) ratio, which he proved to be a powerful predictor of future long-run performances of the market. Shiller builds on a key idea remounting to the famous 1934 book by Graham and Dodd, where the authors strongly advocated for the fundamental approach to investment valuation and recommended to "shift(s) the original point of departure, or basis of computation, from the current earnings to the average earnings, which should cover a period of not less than five years, and preferably seven to ten years" Graham, *et al*. (1962). The performance of the test is quite satisfactory in the case of the US market from the end of 19th century till today. The market crash soon followed the appearance of the book has been the first strong confirmation of this performance[2].

To cope with this evidence a model of stock market price dynamics should be able to determine whether the trajectory is wandering far from the fundamentals. To do so, the model should explicitly take into account macroeconomic variables such as the CAPE, but surprisingly enough to our knowledge no such models have ever been put forth[3].

A possible explanation for this is the above mentioned focus on option pricing, for which "the relevant time scales for our purpose range between several days and several months" (see Cont and Tankov (2004) p.3), time scales for which what matters is how "the full effects of new information on intrinsic values to be reflected "instantaneously" in actual prices" (see Fama (1965) p.56).

It is clear that modeling a Shiller-type price dynamics requires a completely different time scale. On the other hand, introducing in the model some mean-reverting mechanism would not be enough to generate stock prices which "have a life of their own; they are not simply responding to

[2] For a discussion of the technical aspects of the model, please refer to Campbell and Yogo (2006), and literature cited therein.

[3] A model of stock market price dynamic accounting for macroeconomic variables different from the valuation ratios we are interested in is the so-called Fed Model, which presumes a simple relationship between earnings yields and yields on government and high-grade corporate bonds (see Lander, *et al*. (1997)).

earnings or dividends. Nor does it appear that they are determined only by information about future" earnings or dividends, see Shiller (2000) p.183 and Zhong, *et al*. (2003). A dynamical model able to generate a significant transitory component around the equilibrium which reflects the rationally expected value of the asset requires the action of at least two different contrasting forces: one pushing the price towards its equilibrium and the other pointing the opposite direction. This is what Chiarella (1992) has pioneered introducing the first model of financial market with heterogeneous agents, where the different trading strategies of fundamentalists and chartists generate complex dynamics of the price, which nevertheless gravitates around its fundamental value. Unfortunately, estimating heterogeneous agents models is challenging. It is very difficult to obtain any quantitative measure of preferences, which determines the demand functions of each type of agent. To avoid this, we follow the approach suggested in a similar context by Biagini, *et al*. (2013), who describe at an aggregate level the effects of the interaction at a micro level of different types of agents. In particular, they assume that "the perceived fundamental value" shifts in time because of the varying share of optimists in the market. Differently from all the above-cited papers, we do not try to a priori guess how the mood of the market dictates "the perceived fundamental value". Instead, we allow the fundamental value, towards which fundamentalists expects that the current price should revert, to vary in time and to depend on the initial averaged Price-to-Earnings ratio as on an initial anchor (see Tversky and Kahneman (1974)).

In our model the price growth depends on three components

1. An autoregressive component, naturally justified in terms of agents' expectation that expected returns are higher in bullish markets than in bearish ones;
2. A fundamental component, proportional to the level of the logarithmic averaged Earnings-to-Price ratio[4] and the perceived fundamental value;
3. A stochastic component ensuring the diffusive behavior of stock prices.

Estimating the parameters of the model on Standard and Poor Composite Stock Price Index (S&P) historical series, we find that the assumptions of Lengnick and Wohltmann (2010) are in some sense corroborated by the model. Initially, the fundamentalists' perception of the fundamental value is biased in the direction of the most recent performance of the market. If prices are high (low) the fundamental stock price is perceived to lie above (below) its true counterpart. However, optimism (pessimism) does not last forever, as in Biagini, *et al*. (2013) (see p.10), and within approximately 11 or 12 years it reverts to a value independent of the initial level and compatible with the long-run mean observed by Shiller.

Beside this empirical result, we are able to prove that, if we consider a sufficiently long horizon, the expected rate of return and the expected gross return are linear in the initial value of the logarithmic Earnings-to-Price ratio. Moreover, their variance converges to zero with a rate consistent with a diffusive behavior. This means that, in our model, the stock prices dynamics may exhibit significant and persistent upwards and downwards deviations from the long-run mean value of the averaged Earnings-to-Price ratio. Nevertheless, the averaged Earnings-to-Price ratio is a good predictor of future long-run returns, as claimed by Campbell and Shiller (1988a), Shiller (2000), Lander *et al*. (1997). The result holds for both returns and gross returns. In the latter case, we assume that the Dividends-to-Price ratio follows a stationary process as in Campbell and Shiller (1988a,b). Our results are also in keeping with Hodrick (1992), who "demonstrates that a relatively large amount of long-run predictability is consistent with only a small amount of short-run predictability".

[4] To shorten the writing, we will refer to logarithmic Earnings-to-Price and Dividends-to-Price ratios as to Earnings-to-Price and Dividends-to-Price ratios.

2. The Model

We refer to the real (inflation adjusted) price of the stock index measured at the beginning of time period t with P_t, while D_t denotes the real dividend paid between t and t+1. Accordingly, the real logarithmic gross return on the index held from time t until time t+1 reads as $H_t = \log(P_{t+1} + D_t) - \log P_t$.

The description of the return dynamics is on a monthly basis. The notation $t + 1$ refers to time t increased by one month. The real gross yield over a period of h months corresponds to

$$y_{t,h} = \frac{1}{h}\sum_{i=0}^{h-1} H_{t+i} \tag{1}$$

We also introduce the index logarithmic price $p_t = \log P_t$, in terms of which the gross yield can be rewritten as

$$y_{t,h} = \frac{1}{h}\sum_{i=0}^{h-1}(p_{t+i+1} - p_{t+i}) + \frac{1}{h}\sum_{i=0}^{h-1}\log(1 + \frac{D_{t+i}}{P_{t+1+i}}) \tag{2}$$

where the telescopic sum on the right hand side is equivalent to $(p_{t+h} - p_t)/h$. The latter term on the right hand side represents a non linear function of the Dividends-to-Price ratio. Campbell and Shiller argue that the ratio $d_t - p_{t+1} \doteq \log D_t - \log P_{t+1}$ follows a stationary stochastic process. Specifically, Campbell and Shiller (1988a) (see Table 3) test the hypothesis that the time series has a unit root[5], rejecting the null of an integrated process at 1% significance level. In light of this evidence the dynamics of the Dividends-to-Price is given by

$$\Delta(d_{t-1} - p_t) = -\theta(d_{t-1} - p_t - \log \mathcal{G}) + \sigma_d W_t^d \tag{3}$$

with initial condition equal to $d_{-1} = \log D_{-1}$. The AR(1) coefficient is given by $1 - \theta$, σ_d is a positive volatility constant, $\{W_t^d\}$ are independent and identically distributed (i.i.d.) Gaussian increments with zero-mean and unit variance, and $\log \mathcal{G}$ is the fixed mean that can be used as an expansion point. By means of a first-order Taylor expansion, the quantity $\log(1 + D_t/P_{t+1})$ appearing in equation (2) can be replaced by a linear function of the Dividends-to-Price ratio.

The dependent variable dealt with throughout the paper is the gross return of the stock index, while as a predictive quantity we consider the Earnings-to-Price ratio $\log\langle e\rangle_t^{10} - p_t$. The symbol $\langle e\rangle_t^{10}$ refers to the moving average of real earnings over a time window of ten years. At variance with Campbell and Shiller (1988b), where the analysis is based on the geometric average, we use the arithmetic average. This choiche allows to extend the framework to the case of negative earnings, for which the former case is troublesome. Switching from the geometric to the arithmetic average has a minor impact on the overall picture. The use of average earnings in computing the price ratios has been strongly advocated by the literature in recognition of the cyclical variability of earnings. Graham *et al*. (1962) recommend an approach that "Shifts the original point of departure, or basis of computation, from the current earnings to the average earnings, which should cover a period of not less than five years, and preferably seven to ten years."

In Campbell and Shiller (1988a,b), the regression of real and excess stock returns on explanatory variables which are known at the start of the year t shows that the Dividends-to-Price ratio and the Earnings-to-Price ratio have good predictive capabilities. The ratio variables are used as indicators of fundamental value relative to price. The basic idea is that if stocks are underpriced

[5] By adopting the same notation, the statistic is formed from the F-statistic in the regression $\Delta y_t = \mu + \beta t + \alpha y_{t-1}$, where y_t is the logarithmic Dividends-to-Price ratio corrected for serial correlation in the equation error using a fourth order Newey-West correction.

relative to fundamental value, returns tend to be high subsequently, while the converse holds if stocks are overpriced. A realistic model for the price dynamics should take into account such evidence. Thus, we describe the dynamics of the price assuming the existence of an exogenous fundamental component given by a mean-reverting term whose long-run target level depends linearly on the current value of the Earnings-to-Price ratio.

The following linear system of stochastic difference equations drives the dynamics of log prices

$$\begin{cases} p_{t+1} = p_t + \mu_t + \xi_t \\ \mu_{t+1} = \gamma\mu_t + \kappa(\log\langle e\rangle_t^{10} - p_t + \mathcal{H} + g\mathcal{F}t) + \sigma_\mu W_t^\mu \\ \xi_{t+1} = \xi p_t + \sigma_\xi W_t^\xi \end{cases} \tag{4}$$

with initial time conditions equal to $p_0 = \log P_0$, and μ_0. The quantities $\{W_t^\mu\}$, and $\{W_t^\xi\}$ for $t = 0, \ldots, h$ are i.i.d. Gaussian increments with zero-mean and unit variance. The coefficients σ_μ, and σ_ξ are positive volatility constants. The system of equations (4) determines the evolution of prices as a superposition of a local drift μ_t and a noise component ξ_t. The latter is a zero-mean process originating from ξ_0 which ensures the diffusive behavior of stock prices. The most relevant component corresponds to the equation driving the local drift

$$\mu_{t+1} = \gamma\mu_t + \kappa(\log\langle e\rangle_t^{10} - p_t + \mathcal{H} + g\mathcal{F}t) + \sigma_\mu W_t^\mu \tag{5}$$

The future level of μ_{t+1} depends on the value μ_t prevailing at the previous time step. The influence of the autoregressive component is determined by the agents' sensitivity to the market trend γ. This effect can be justified in terms of the expectation that returns are higher in bullish markets than in bearish markets. This is true provided that γ turns out to be positive, which is the result of our estimation (see section 3). This comes at not surprise. It is well known that, while successive monthly returns of single stocks are negatively auto-correlated, the auto-correlation is positive for the market index. Competing with the autoregressive component, a second mechanism affects the drift from a fundamental perspective.

The second term in the right hand side of (5) represents an exogenous "fundamental" component expressed by a mean-reverting term. The actual stock price may deviate from the long-run behavior as a combined effect of both random external disturbances and the autoregressive component. Investors reallocate assets in response to this disequilibrium causing stock prices to move in the direction that reduces the deviation. In modeling the fundamental effect, we bear in mind that "in reality it is very difficult (if not impossible) to identify the true fundamental value of any stock", see Lengnick and Wohltmann (2010). We, therefore, allow the mean reversion target to vary with time and depend on the quantities $\mathcal{F}$ and $\mathcal{H}$. The latter are linear functions of the level of the averaged Earnings-to-Price ratio at time zero. Specifically, supported by the results of Campbell and Shiller (1988a,b), we assume that the quantities $\mathcal{F}$ and $\mathcal{H}$ have the form

$$\begin{gathered} g(1+\mathcal{F}) = \alpha_\mathcal{F} + \beta_\mathcal{F} \log EP_0 \\ \mathcal{H} = \alpha_\mathcal{H} + \beta_\mathcal{H} \log EP_0 \end{gathered} \tag{6}$$

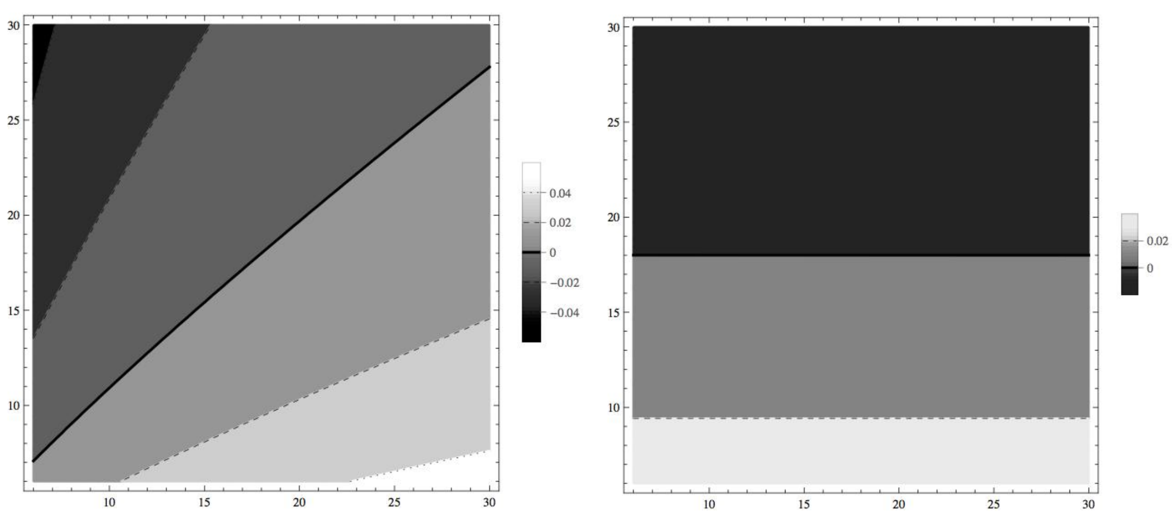
In a similar fashion, we also assume that $\mathcal{G}$ (see equation (3)) depends linearly on the initial Earnings-to-Price ratio

$$\mathcal{G} = \alpha_\mathcal{G} + \beta_\mathcal{G} \log EP_0 \tag{7}$$

The value of the coefficients $\alpha_\mathcal{F}$, $\alpha_\mathcal{G}$, $\beta_\mathcal{F}$, and $\beta_\mathcal{G}$ is fixed exploiting the bias corrected approach to the predictive regressions of Campbell and Shiller mentioned in the next section. It is worth to stress that these values enter the dynamical model of equations (3) and (4) exogenously. In section 3 we discuss estimation of the model parameters, i.e. γ, κ, g, θ, and the three volatility constants. In the current setting, we do not model explicitly the earning dynamics. The evolution of averaged earnings is exogenous and follows an exponential law, i.e. $\langle e\rangle_t^{10} = \langle e\rangle_0^{10}\exp(gt)$.

It is interesting to comment how the perceived fundamental value $\mathcal{H} + g\mathcal{F}t$ evolves in time. Since we find that $\beta_{\mathcal{F}} > 0$ and $\beta_{\mathcal{H}} < 0$ (please refer to section 3), its initial value is smaller [larger] the higher [lower] CAPE_0, but it gradually reverts toward larger and larger [smaller and smaller] values as time elapses. Within $-\beta_{\mathcal{H}}/\beta_{\mathcal{F}}$ months, it reaches a value independent of CAPE_0. Since $-\frac{\beta_{\mathcal{H}}}{\beta_{\mathcal{F}}} \cong 133$ and the full adjustment is not immediate, but the price needs a finite time of order $\kappa^{-1} \cong 33$ to recover the fundamental level, the Cyclically Adjusted Price Earning ratio is expected to mean revert to its long-run value, independent of the initial one, within approximately 14 years [6]. In figure 1 we draw the level curves of the increment of the drift μ ascribed to the fundamental component $\kappa(\log\langle e\rangle_t^{10} - p_t + \mathcal{H} + g\mathcal{F}t)$ in the second equation in (4) as a function of CAPE_0 and CAPE_t for t = 1 and t = 133. Both plots are drawn for values within the 95% confidence region estimated in section 3: specifically, for $\alpha_{\mathcal{F}} = 0.236$, $\beta_{\mathcal{F}} = 0.2531$, and $\alpha_{\mathcal{H}} = 0.236$, $\beta_{\mathcal{H}} = -0.85$. The left panel reports that the perceived fundamental value (i.e. the value of CAPE_t for which this increment is null) increases from 7 when CAPE_0 = 6 to 28 when CAPE_0 = 30 (Biagini *et al*. (2013) and Lengnick and Wohltmann (2010)).

The right panel shows that for t = 133 the perceived fundamental value is constant and equal to 18, irrespectively of CAPE_0. These findings confirm that momentum strategies are more likely to be profitable in the short than in the medium/long-run. The reverse is true for fundamental strategies. It is to be stressed that an analogous remark holds for the dividend component of the gross return. Since $\beta_{\mathcal{G}} > 0$, "no movement of U.S. aggregate stock prices beyond the trend growth of prices has ever been subsequently justified by dividend movements (see Shiller (1981))" (see Shiller (2000) p.183), since a high [low] CAPE_0 is followed by a low [high] expected dividend yield.



Time (months)

Figure 1. Left panel: Increment of μ ascribed to the fundamental component as a function of CAPE_0 (horizontal axis) and CAPE_t (vertical axis) for t = 1. Right panel: Increment of μ for t = 133.

[6] These values are near those considered by Campbell and Shiller (1988a,b) to prove the predictability of long-term stock returns.

Now, we are ready to state the main theoretical result of the paper, which characterises the asymptotic behavior of the first and second moment of the logarithmic price gross returns.

Proposition 1. The expected gross yield over *h* months is asymptotically linear in $\mathcal{F}$ and $\mathcal{G}$.

$$\mathbb{E}_0[y_{0,h}] = g(1+\mathcal{F}) + \mathcal{G} + O(1/h) \tag{8}$$

while the variance converges to zero as predicted by a diffusive model

$$\mathrm{Var}_0[y_{0,h}] = \frac{\sigma_p^2}{h} + \frac{1}{h}\frac{\mathcal{G}^2\sigma_p^2}{\theta(2-\theta)} + o(\frac{1}{h}) \tag{9}$$

with $\sigma_p = \sigma_\xi(1-\gamma)/\kappa$ [7].

The coefficient g determines the exponential growth of the cyclically adjusted earnings. If it is equal to zero, then the real net yield grows sub-exponentially, and the main contribution to the growth comes from the dividend component. Independently of g, both components affect the variance of the gross yield, which asymptotically converges to zero with a rate consistent with the diffusive behavior of stock returns. Finer effects, like possible non zero autocorrelation induced by business cycles over long horizons, can be included in the model modifying the covariance structure of the $\{W_t^\xi\}$ noise. The results summarized in Proposition 1 do not depend crucially on the Gaussian assumption about the noise increments. It can be relaxed at any time, to take into consideration additional features of the empirical time series.

Remark. It is worth to notice that the contribution to the long-term yield whose scaling over time is more persistent can be determined explicitely. This term ultimately determines convergence towards the limit $1 + g(1+\mathcal{F}) + \mathcal{G}$. The expression of the leading correction, proportional to 1/*h*, reads

$$\mathcal{H} - g(1+\mathcal{F})\left[1 + \kappa\frac{1-\lambda_-\lambda_+}{(1-\lambda_-)^2(1-\lambda_+)^2}\right] + \log EP_0 + \mathcal{G}\left(1-\frac{1}{\theta}\right)(\log DP_0 - \log\mathcal{G}) \tag{10}$$

The coefficients λ_- and λ_+ read $\frac{\gamma+1}{2} + \frac{1}{2}\sqrt{(1-\gamma)^2 - 4\kappa}$ and $\frac{\gamma+1}{2} - \frac{1}{2}\sqrt{(1-\gamma)^2 - 4\kappa}$, respectively. Assuming $\theta \ll 1$, the quantity $1/\theta$ fixes the typical time scale of the mean-reverting process. Then, the last term in (10) contributes only marginally to the leading correction, since neither an extremely large time scale for the process nor an extreme discrepancy between $\log DP_0$ and $\log\mathcal{G}$ are expected. Another interesting point to note is that the choice of μ_0 plays a minor role in relation to the speed of convergence of the process to the long-term expected value. Indeed, μ_0 does not appear in expression (10). It is possible to show that its effect is exponentially damped by the *h* power of the lambda coefficients, that is by λ^h_- and λ^h_+. Since $0 < \lambda_- < \lambda_+ < 1$, the dominating contribution for $h \gg 1$ is given by λ^h_+.

3. Data Set and Parameter Estimation

The data set analyzed in this paper consists of records on a monthly basis of prices, earnings, and dividends for the Standard and Poor Composite Stock Price Index (S&P). The data are discussed in Campbell and Shiller (1987, 1988a,b), and are freely available from Robert J. Shiller's webpage http://www.econ.yale.edu/. The time series cover the entire period from January 1871 until December 2012.

[7] The details of the proof are available from authors upon request.

Following the approach of Amihud and Hurvich (2004) [8], we fix the exogenous coefficients $\alpha_{\mathcal{F}}$, $\alpha_{\mathcal{G}}$, $\beta_{\mathcal{F}}$, and $\beta_{\mathcal{G}}$. Results are reported in Table 1. We also report the prediction coefficients of the gross returns on the initial logarithmic CAPE, α and β. The statistical significance of the β coefficients is assessed by means of the reduced-biased standard errors discussed in Amihud and Hurvich (2004)[9]. If we compute the *t*-statistics for the β parameters, we find t = 2.29, which corresponds to a significant evidence of predictability. The predictive coefficient values can be disaggregated in two components due to the net return growth and dividends. Data in Table 1 show that the long-run return predictability is largely due to the highly predictable behavior of the Dividends-to-Price ratio. Nonetheless, the contribution of net returns to the predictive regression is relevant. This evidence confirms empirical results discussed in Table 4 of Campbell and Yogo (2006). In their case it is important to mention that the gross return exceeding the risk-free rate is considered. Finally, Table 1 confirms that the bootstrap approach rejects the null hypothesis β = 0 at 5% significance level.

Table 1. All coefficient values and associated errors are expressed on a yearly basis and in basis points. The *p*-values obtained by means of ten thousands bootstrap samples are between parentheses.

Parameters	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\alpha_{\mathcal{F}}}$	$\boldsymbol{\beta_{\mathcal{F}}}$	$\boldsymbol{\alpha_{\mathcal{G}}}$	$\boldsymbol{\beta_{\mathcal{G}}}$
S&P	3667	1023 ±445	2531	767 ±459	1527	393 ±19
***p*-value**		(0.048)		(0.091)		(< 0.001)

Moving to the estimation of the dynamical model parameters, we preliminary estimate the rate of growth of the averaged earnings, g. We regress a vector of 192 monthly logarithmic averages on the time horizon h. Then, we repeat the procedure over rolling windows from January 1881 and compute 68% confidence intervals. In Table 2, we report the results for the S&P index. Similarly, we estimate the value of θ and σ_d regressing $\Delta(d_{t-1} - p_t)$ on $(d_{t-1} - p_t - \log \mathcal{G})$. We find θ = 0.0271, a value implying a scale of mean-reversion of order 37 months. Concerning γ and κ, we regress μ_{t+1} on μ_t and $(\log\langle e\rangle_t^{10} - p_t + \mathcal{H} + g\mathcal{F}t)$. In order to do so, we need to introduce a proxy for the unobservable variable μ_t, and to fix a value for the unknown quantity $\mathcal{H}$. The former point is solved by means of $p_t - p_{t-1}$, while coefficients $\alpha_{\mathcal{H}}$ and $\beta_{\mathcal{H}}$ are choosen in order to minimize the correction term (10). For $\alpha_{\mathcal{H}} = 0.85$ and $\beta_{\mathcal{H}} = -0.85$, we find $\gamma = 0.25$ and $\kappa = 0.03$. These values satisfy the constraint $4\kappa < (1-\gamma)^2$ required to ensure the stability of the dynamical system. As far as the estimate of σ_p is concerned, equation (9) shows that the variance of $p_h - p_0$ scales linearly with h with a coefficient equal to σ_p^2 . Fitting the empirical curve with a linear relation, we obtain the value 0.182% reported in Table 2.

[8] As pointed out by many authors, when the regressors in predictive regressions correspond to financial ratios the Ordinary Least-Squares estimator is strongly biased in finite samples, see for instance Mankiw and Shapiro (1986), Stambaugh (1986), and Stambaugh (1999).

[9] Furher details concerning the bootstrap approach used to assess statistical significance are available from the authors upon request.

Table 2. Parameter values for the S&P time series with associated 68% confidence level

Parameter with scale	Mean and range
g ($\times 10^{-4}$ month^{-1})	12 (-3, 31)
θ ($\times 10^{-4}$ month^{-1})	271 (111, 430)
κ ($\times 10^{-4}$ month^{-1})	323 (81, 597)
γ	0.25 (0.18, 0.33)
σ_d^2 ($\times 10^{-4}$ month)	13 (10, 20)
σ_μ^2 ($\times 10^{-4}$ month)	12 (9, 18)
σ_p^2 ($\times 10^{-4}$ month)	18.2 (18.1, 18.3)

Figure 2 reports a Monte Carlo simulation of the model (3) and (4). Each point corresponds to a single realization of $y_{t,h}$ with h = 24, …, 192 months with t starting from January 1881. The initial time values of Y_0 , μ_0, and $\log DP_0$ are fixed equal to the empirical ones. We plot the linear relation between yields and Earnings-to-Price as predicted by equation (8) using the coefficient values of Table 1 rescaled to a monthly level. Boundaries of the 95% confidence region are provided too. All analytical predictions are in full agreement with Monte Carlo numerical results.

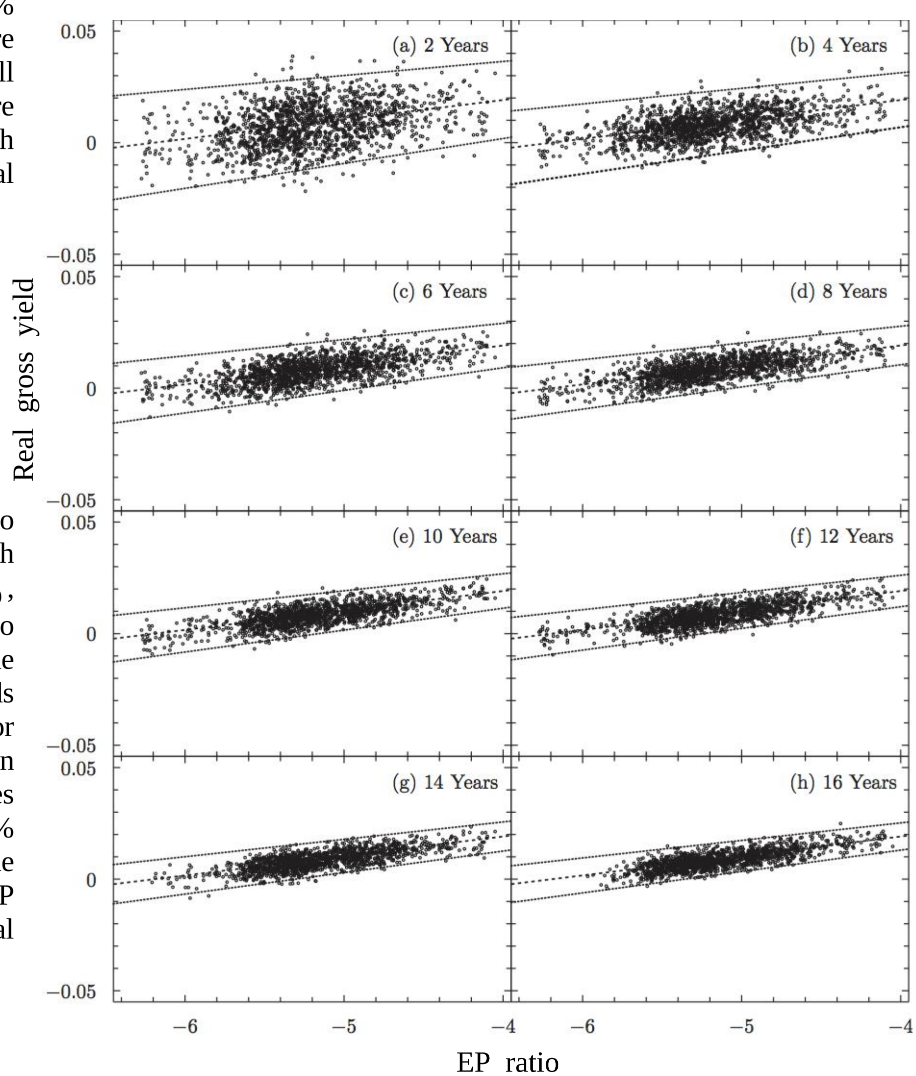


Figure 2. Monte Carlo scenarios generated with initial conditions Y_0, μ_0, and $\log DP_0$ equal to empirical values. The dashed line corresponds to the long-run behavior predicted by equation (8). The dotted lines delimit the 95% confidence region. The horizontal axis is EP ratio, and the vertical axis is Real gross yield.

Finally, we test the goodness of the estimated parameters and the effectiveness of our dynamical system. In figures 3 and 4, we plot the prediction of equation (8) with the associated confidence band on the net returns and on the gross yields, respectively, for time horizons ranging from 24 to 192 months. As expected, the model captures the shrinking of the historical data cloud with a scaling exponent dominated by the diffusive component of the price dynamics. As far as the central value is concerned, the consistency is very good for the short-time horizons, whilst it worsens for longer maturities. This effect is partially expected, since the linear coefficients of the predictive regressions are constants quantities which are exogenous to the dynamical model and do not evolve with time[10].

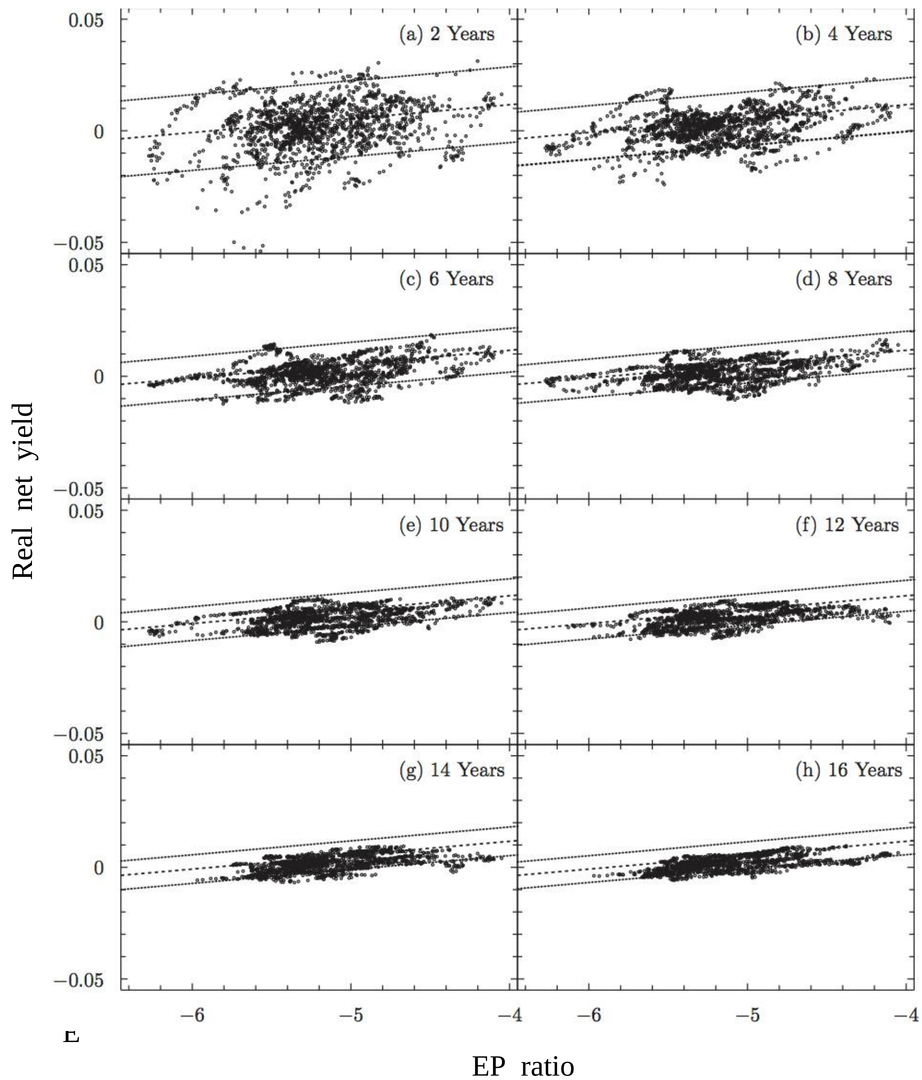


Figure 3. Regression of the yield $(p_{t+h} - p_t)/h$ on the explanatory EP ratio. Solid points: empirical data; dashed line: model prediction; dotted line: 95% confidence region from model prediction. The vertical axis is Real net yield.

[10] We are currently exploring an econometric approach extending the low-bias procedure of Amihud and Hurvich (2004) to long-run horizons. Preliminary results suggest that the β coefficient rescales geometrically over time.

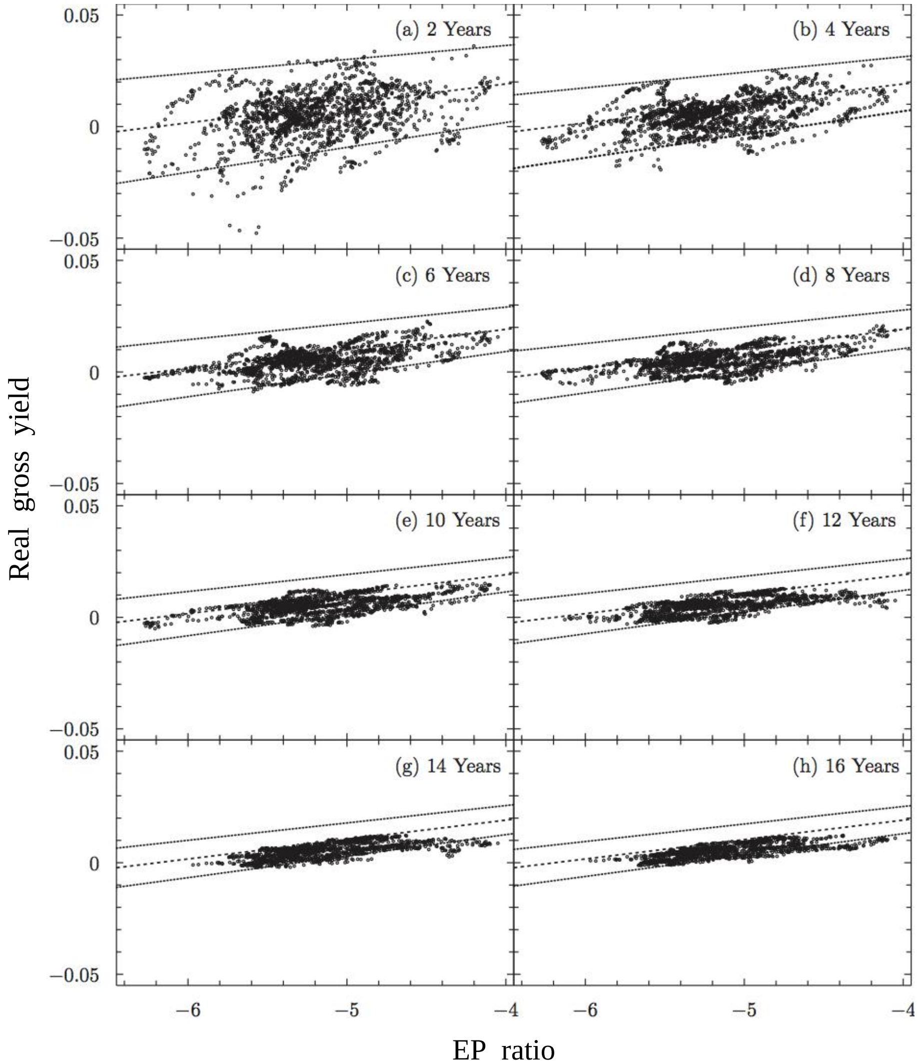


Figure 4. Regression of yield (1) on the explanatory EP ratio. Solid points: empirical data; dashed line: model prediction; dotted line: 95% confidence region from model prediction. The vertical axis is Real gross yield.

4. Conclusions and Perspectives

This paper proposes a simple dynamical model for the long-run behavior of stock index returns for the U.S. market. The log price dynamics depends on two components: An autoregressive component typical for stock index returns and a mean-reverting component whose long-run level is fixed by the level of Shiller's CAPE.

Substantial evidence of the importance of fundamentals in the valuation of international stock markets has been accumulated by the proponents of fundamental indexation e.g. Arnott, *et al*. (2005). Practitioners and academicians alike have been using several valuation measures for estimating the intrinsic value of a stock index. For instance, in Table 2 of Poterba and Samwick (1995) the ratio of market value of corporate stock to GDP, the year-end Price-to-Earnings ratio, the year-end Price-to-Dividends ratio, and Tobin's q are reported from 1947 to 1995 in an effort of alerting the reader on the possible overvaluation of the index[11]. In particular Tobin's q has been proposed as another efficient method of measuring the value of the stock market, with efficiency comparable to the CAPE (see Smithers (2009)). The *q* ratio is the ratio of price to net worth at replacement cost rather than the historic or book cost of companies. It therefore allows for the impact of inflation, much alike the CAPE, which averages real earnings over a ten-year span. It would be interesting to carry out an empirical analysis of the relationship between Tobin's q and future stock index returns as far as to extend the present approach to countries other than the U.S. Both perspectives are worth to be followed but require high-quality long-term time series, which at the moment we have not been able to find. Longer time series would also be important so as to investigate possible "evolutionary" phenomena like the end of the relationship between market valuation and interest rates, which may perhaps be interpreted as an example of the adaptive market hypothesis of Farmer and Lo (1999), Lo (2004). As a possible future extension to model the emergence of explosive bubbles, we plan to relax the assumption of stationarity of the Dividends-to-Price ratio process following the approach recently investigated by Engsted, *et al*. (2012).

References

[1] Amihud, Y. , and C. M. Hurvich (2004). "Predictive regressions: A reduced-bias estimation method", *Journal of Financial and Quantitative Analysis*, 39(4): 813-841.

[2] Arnott, R. D., J. Hsu, and P. Moore (2005). "Fundamental indexation", *Financial Analysts Journal*, 61(2): 83-99.

[3] Biagini, F. , H. Follmer, and S. Nedelcu (2013). "Shifting martingale measures and the birth of a bubble as a submartingale", *Finance and Stochastics*, 18(2): 297-326.

[4] Campbell, J. Y., and R. J. Shiller(1987). "Cointegration and tests of present value models", *Journal of Political Economy*, 95(5): 1062-1088.

[5] Campbell, J. Y. and R. J. Shiller (1988a). "The dividend-price ratio and expectations of future dividends and discount factors", *Review of Financial Studies*, 1(3): 195-228.

[6] Campbell, J. Y., and R. J. Shiller (1988b). "Stock prices, earnings and expected dividends", *Journal of Finance*, 43(3): 661-676.

[7] Campbell, J. Y., and M. Yogo (2006). "Efficient tests of stock return predictability", *Journal of Financial Economics*, 81(1): 27-60.

[8] Chiarella, C. (1992). "The dynamics of speculative behaviour", *Annals of Operations Research*, 37(1): 101-123.

[9] Cont, R., P. Tankov (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC.

[11] "It is difficult to distill a simple conclusion from table 2. While price-to-earnings ratios are not unusually high at present, other measures of stock price valuation are at, or near, historical highs. (…) Table 2 does suggest, however, that in assessing the macroeconomic consequences of stock price movements, it may be important to distinguish between stock price fluctuations that are associated with movements in the price-to-earnings or price-to-dividends ratios and those that are not. A number of recent studies suggest that variations in the earnings-to-price ratio are correlated with prospective stock market returns, (…). Sharp increases in either the price-to-earnings or the price-to-dividends ratio, other things equal, are associated with lower prospective returns", see Poterba and Samwick (1995).

[10] Engsted, T., T. Q. Pedersen, and C. Tanggaard (2012). "The log-linear return approximation, bubbles, and predictability", *Journal of Financial and Quantitative Analysis*, 47(3): 643-665.

[11] Fama, E. F. (1965). "The behavior of stock-market prices", *The Journal of Business*, 38(1): 34-105.

[12] Farmer, D., and A. W. Lo (1999). "Frontiers of finance: Evolution and efficient markets", *Proceedings of the National Academy of Sciences*, 96(18): 9991-9992.

[13] Graham, B., D. L. Dodd, and S. Cottle (1962). *Security Analysis*. McGraw-Hill, New York.

[14] Hodrick, R. J. (1992). "Dividend yields and expected stock returns: Alternative procedures for inference and measurement", *Review of Financial studies*, 5(3): 357-386.

[15] Lander, J., A. Orphanides, and M. Douvogiannis (1997). *Earnings forecasts and the predictability of stock returns: Evidence from trading the S&P*. Technical report, Board of Governors of the Federal Reserve System, 1997.

[16] Lengnick, M., and H.-W. Wohltmann (2013). "Agent-based financial markets and New Keynesian macroeconomics: A synthesis", *Journal of Economic Interaction and Coordination*, 8(1): 1-32.

[17] Lo, A. W.(2004). "The adaptive markets hypothesis: Market efficiency from an evolutionary perspective", *Journal of Portfolio Management*, 30(5): 15-29.

[18] Mandelbrot, B. (1963). "The variation of certain speculative prices", *The Journal of Business*, 36(4): 394-419.

[19] Mankiw, G. N., and M. D. Shapiro (1986). "Do we reject too often? : Small sample properties of tests of rational expectations models", *Economics Letters*, 20(2): 139 -145.

[20] Poterba, J. M., and A. A. Samwick (1995). "Stock ownership patterns, stock market fluctuations, and consumption", *Brookings Papers on Economic Activity*, 1995(2): 295-372.

[21] Samuelson, P. (1965). "Proof that properly anticipated prices fluctuate randomly", *Industrial Management Review*, 6(2): 41-49.

[22] Shiller, R. J.(2000). *Irrational Exuberance*. Princeton, N.J.: Princeton University Press.

[23] Smithers, A. (2009). *Wall Street Revalued: Imperfect Markets and Inept Central Bankers*. John Wiley and Sons, Ltd..

[24] Stambaugh, R. F. (1986). *Bias in regressions with lagged stochastic regressors*. Center for Research in Security Prices, Graduate School of Business, University of Chicago, 1986.

[25] Stambaugh, R. F. (1999). "Predictive regressions", *Journal of Financial Economics*, 54(3): 375 - 421.

[26] Tversky, A., and D. Kahneman (1974). "Judgment under uncertainty: Heuristics and biases", *Science*, 185(4157):1124-1131.

[27] Zhong, M., A. F. Darrat, and D. C. Anderson (2003). "Do US stock prices deviate from their fundamental values? Some new evidence", *Journal of Banking & Finance*, 27(4): 673-697.