Quantitative Finance and Economics

http://www.aimspress.com/journal/QFE

QFE, 5(2): 287–310.
DOI: 10.3934/QFE.2021013
Received: 31 January 2021
Accepted: 06 April 2021
Published: 16 April 2021

Research article

Estimating market index valuation from macroeconomic trends

Andrea Afify and Hector Eduardo Roman*

Department of Physics, University of Milano-Bicocca, Piazza della Scienza 3, 20126 Milan, Italy

* **Correspondence:** Email: eduardo.roman@mib.infn.it; Tel: +390264482334.

Abstract: We discuss USA stock market data from 1789 until 2020, focusing our attention on the S&P 500 index (1957–2020). We find that the data can be split into two periods, (1789–1948) and (1948–2020), displaying roughly 2% and 7% growth rates, respectively. The index variations from each trend appear similar, suggesting some degree of stationarity in market fluctuations. We then correlate market behavior to macroeconomic data, such as world (and USA) population growth and gross domestic product (GDP), on different time horizons. The analysis signals that the S&P 500 might be overvalued, possibly undergoing a series of bubbles, since the 1990s. To understand this behavior, we introduce a model for bubbles, showing that they can be caused by a lack of correlations between stock prices and a virtual market index, the latter calculated self-consistently from the stock prices. We argue that variations, $\Delta\gamma$, in the "bubble parameter" (or decoupling factor γ), are anticorrelated to variations of the Federal Funds Rate (FFR), which may trigger a bubble phenomenon ($\gamma \to 1$) when persistent rate cuts become too pronounced. The FFR are confronted with the consumer price index (CPI) in the period (1955–2020) as an attempt to complete the picture. Our analyses suggest that the strong departure of the S&P 500 from historical fundamental trends within (1990–2020) may reflect the development of financial anomalies, in part related to monetary policies, which should be carefully addressed in the near future.

Keywords: S&P 500 index; fluctuations and trends; market bubbles; world and USA population growth rates; GDP growth rates; Federal Funds Rate

JEL Codes: C01, C02, E01, E02, J00

1. Introduction

Determining the degree of a market index valuation is essential to assess any investment position. Very popular indeed are exchange traded funds (ETF), the latter designed to replicate market behavior, in keeping with a lower degree of risk than focusing on single assest positions. In the case of medium time horizons, over periods of few years, the valuation may be done empirically by looking at the historical data of the market index of interest, in an attempt to find suitable entry-exit points. A complementary and more elaborated approach requires the consideration of macroeconomic scenarios and the scrutiny of economic fundamentals (see e.g. Graham (1965); Malkiel (1999); Lynch and Rothchild (2000)); Kahneman (2011); Bogle (2017)), which may be useful on longer time horizons.

It is widely accepted that market fluctuations are not stationary over sufficiently long periods of time, although they may appear so within restricted circumstances. The most important issue associated to non-stationarity is the emergence of bubbles which have a detrimental effect on both the market and the economy as well. There is a great deal of discussions about the origin of bubbles, and eventually how to detect them (see e.g. Shiller (2003); Malkiel (2010); Phillips and Yu (2011); Wu (2012); Caspi et al. (2014); Zhang et al. (2016); Sornette et al. (2018); Demirer et al. (2019)).

Not all analysts and financial experts agree on the very existence of bubbles though. One can argue that they are the result of rather "typical" departures from stationarity yielding large amounts of excess volatility in asset prices. However, since global stock markets are correlated to each other to some extent, bubbles can develop almost simultaneously everywhere. Indeed, their presence is not limited to just a small economic domain, but they can affect world markets as a whole. Those who support the concept of bubbles affirm that their very existence actually invalidates the efficient market hypothesis, which should serve as the underlying and fundamental rule of stock markets. While the second group just claims that bubbles do not exist and the strong departures of market volatility we see can be well fit within the available information hitting the market.

In this work, we discuss the issue of overvaluation of a market index on long-time horizons, and suggest a simple stochastic model of stock market bubbles. Regarding the former, we discuss a mixed type of index valuation obtained by combining both historical market and macroeconomic data, the latter based on world and USA population growth, complemented by GDP data from the USA. Related analysis have been considered in several publications dealing with demography, GDP, oil prices, and other macroeconomic factors (see e.g. Flannery and Protopapadakis (2002); Geanakoplos et al. (2004); Balcilar et al. (2017); Alexius and Spång (2018)). These studies, however, deal with daily variations in the time series considered, while we focus on the long-time behavior of trends.

We apply this approach to the *Standard and Poor's 500 Index* (or S&P 500) by using historical data from 1789, on essentially a daily basis. This large set of data allows us to show the emergence of two distinct growth periods, (1789–1948) and (1948–2020), characterized by 2% and 7% growth rates, respectively. The separation point in 1948 is obtained by using an unbiased search algorithm. We study each period separately by analyzing the index fluctuations from each respective trend, finding a similar behavior for both periods. We see a significant degree of departure from the trend during 1929 and 2000, both considered to be classical cases of bubbles. However, as we will discuss in detail, the 2000 bubble may be just the first "member" of a complex series of bubbles extending over two decades. The latter are interpreted using the market model specifically introduced here for that purpose. Finally, we discuss effects of FFR on the behavior of the S&P 500, and its relation to the CPI.

The paper is organized as follows. We start in Section 2 with a discussion of the historical data used in this work, and the method of analysis employed. In Section 3, we review world and USA population data, and market capitalization to GDP growth rate ratios. Section 4 contains a summary of our results on growth rates over different time periods. Section 5 describes our model for bubbles within a stock market. Section 6, is devoted to the discussion of the results based on a detailed comparison of the S&P 500 with Federal Funds Rate data, and their relation to CPI. Finally, our concluding remarks are summarized in Section 7.

2. Historical background and scaling analysis

The existence of stock markets, in a modern sense, dates back to the 17th century. Here, we provide a brief historical summary taken from an article by Mcgee J (2017)*, concerning the London Stock Exchange (LSE) and the New York Stock Exchange (NYSE): *"The largest European stock exchange today is the LSE. During the trading of the East India companies, investors would post their items for sale at coffee shops such as that of John Castaing, who posted a list of stock prices in 1698. In 1761, 150 men formed a club to buy and sell stock. In 1773, they built The Stock Exchange building in Sweeting's Alley. The official LSE was established on March 3, 1801. At this point, the brokers begin regulating stock exchange. America's stock exchange history began 18 years after the LSE group was formed. In May 1789, the Philadelphia Stock Exchange (PSE) was established. Three years later, on May 17, 1792, 24 merchants from New York met under a Buttonwood tree to form the NYSE. The stock exchange building was located at 11 Wall St. in New York City. Because of its prominent position, the NYSE quickly replaced the PSE as the most powerful stock market in the United States"*.

In 1906, Luther Lee Blake founded the Standard Statistics Bureau, with the goal of providing financial information on a wide variety of companies. In 1923, under the name Standard Statistics Co., it released its first stock market indicator, containing about 230 companies. It became the present Standard & Poor's after its merger with Poor's Publishing in 1941. The merger also boosted the stock index to more than 400 companies, reaching its present number 500 in 1957, yielding what is now known as the Standard & Poor's 500 Index. Although the history of the S&P 500 is quite recent, we will use USA historical data from 1789† by considering a longer database denoted simply as the S&P 500 Index. The data we use is based on: (1) Monthly close values (from May/1st/1789 until February/1st/1885), which have been converted into a daily time series by linearly interpolating monthly values over 20 trading days; and (2) Daily close values since March/1/1885 until today (November 10, 2020).

In the following, we analize the S&P 500 index data by studying its growth rate as a function of trading days, from inception in 1789 until today. We study the data on different time periods, by considering for simplicity the first two cases: (1) a single economic period, and (2) a two-period approximation. This in keeping with our goal of studying long-time horizons. Later below, we relax this condition and deal with trends on shorter-time scales too.

2.1. Single economic period

We present this scenario for the purpose of introducing the notation and methods of analysis. The full S&P 500 index database is plotted in a semilogarithmic scale in Figure 1, where the index, here denoted as $S(t)$, grows roughly exponentially with the trading days t. Therefore, in a first approximation, the index follows a single exponential "trend",

$$\langle S(t)\rangle \simeq S(0)\,\exp(t/T) \tag{1}$$

where T is a characteristic time for growth. The parameters $S(0)$ and T are obtained by performing a least-square fit to the real data. In this first scenario, we use Equation (1) within the whole time interval, 1 day $\leq t < 60000$ days. We find $T = 8338.6$ days, yielding an annual growth rate of about 3%. In other words, the market value has grown, in a first approximation, 1000 times in its 235 years of existence. In

* Joshua Mcgee (2017), https://pocketsense.com/stock-market-created-5261293.html.
† Historical data since 1789 from: S&P 500 - U.S. (SPX), at https://stooq.com.

Figure 1(b), we plot the "detrended" index, i.e. $S(t)/\langle S(t)\rangle$, from which one can appreciate more clearly the index departures from the trend in semilogarithmic scale.

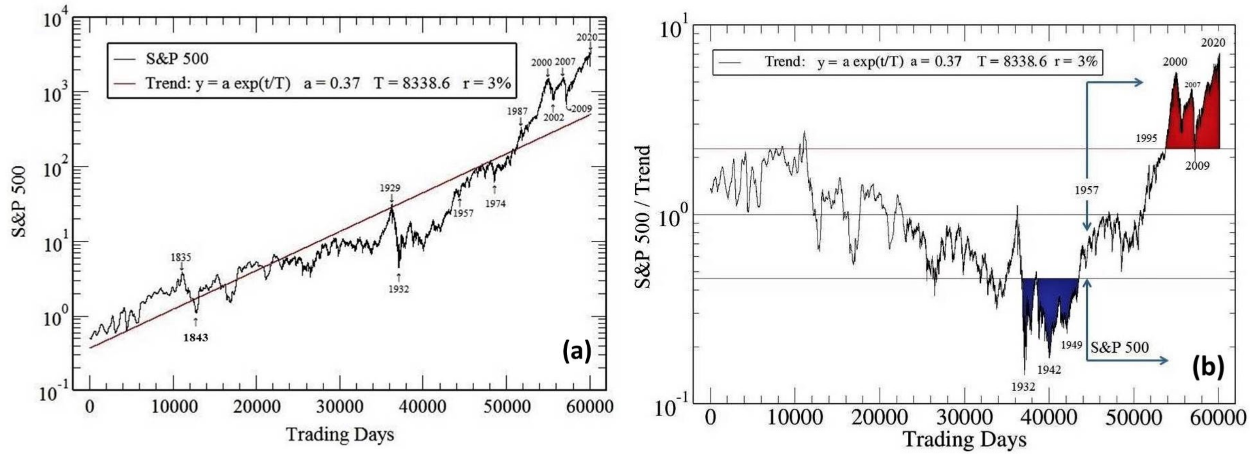

Figure 1. Single growth period (1789–2020). **(a)** S&P 500 vs trading days. Using equation (1), a least-squares fit (red line) is performed over the whole history of the index (black line), yielding $S(0) = a = 0.37$, and $T = 8338.6$ days, corresponding to an annual growth rate $r \simeq 3\%$. The latter is calculated as $r = \exp(D/T) - 1 = 0.031$, where we have used $D = 255$ trading days in a year. Few prominent dates are indicated by the arrows. **(b)** Detrended S&P 500 index, $S(t)/\langle S(t)\rangle$ vs trading days, using the trend given by the red line fit in Figure 1(a). The horizontal lines correspond to the ratios: $S(t)/\langle S(t)\rangle = 2.2$ and 0.45, respectively. The two extreme periods out of those lines are: (1) (1929–1950) (Blue color); and (2) (1995–2020) (Red color). The official S&P 500 started trading in 1957 (arrows).

The result shown in Figure 1(a) covers the whole USA stock market history. Had we restricted ourselves, as it is commonly done in financial analysis, to the period starting around the 1920's, we would had missed about 2/3 of the data. The first 150 years in the history of the index are quite relevant in the sense they give us a hint about the role played by the long time behavior of fundamentals in determining the total market value, a central issue in our present study.

Before considering the latter, let us stress the presence of two clear extreme market periods, the blue and red zones depicted in Figure 1(b). The former corresponds to the market meltdown during the 1930s as a consequence of the great depression (including a bull market from 1933 to 1937); while the latter to the 2000 Internet bubble and to the great recession during (2007–2009), followed by a strongly growing market (2009–2020). In a statitiscal sense, both periods mirror each other's shape but having opposite effects. We study next the two-period case.

2.2. *Two-period case*

In the two-period case (or second scenario), we regard the behavior of the index in Figure 1(a) as composed of two economic periods: The initial one roughly consisting of the first 2/3 of the data, followed by the remaining 1/3 of the data at modern times. Thus, we expect to deal with two different trends in this case, represented by the relations,

$$\langle S_1(t)\rangle \simeq S_1(0)\ \exp(t/T_1), \qquad \text{for } t \leq t_c \tag{2}$$

and

$$\langle S_2(t)\rangle \simeq S_1(t_c)\,\exp[(t-t_c)/T_2], \quad \text{for} \quad t \geq t_c \tag{3}$$

Now, we are dealing with four unknown parameters, i.e. $S_1(0)$, T_1, T_2 and the crossover time t_c. Using the above equations, the least-square fit is obtained by assuming a guess value for t_c, and solving the resulting equations to obtain the other three parameters. The total error is then calculated, and the value of t_c is varied (in a discrete fashion) until the minimum total error is found. The result of this double minimization scheme is shown in Figure 2, yielding $t_c = 41787$ days, corresponding to the post World-War-II year, 1948. This suggests the start of a "new" economic growth period in USA markets, governed by a growth rate ~7%, much higher than the first period one of about 2%.

It is useful to look at the detrended index values, after using Equations (2,3) for each period separately. The result is displayed in Figure 3, where now the fluctuations around the trends look more homogeneously distributed, behaving "stationary" on time scales say, larger than (5000–10000) days. However, one can still distinguish the two bubbles from the 1920's and early 2000, highlighting again their anomalous origin. We have added a third member to this group from 1835, when the market also reached very high quotations.

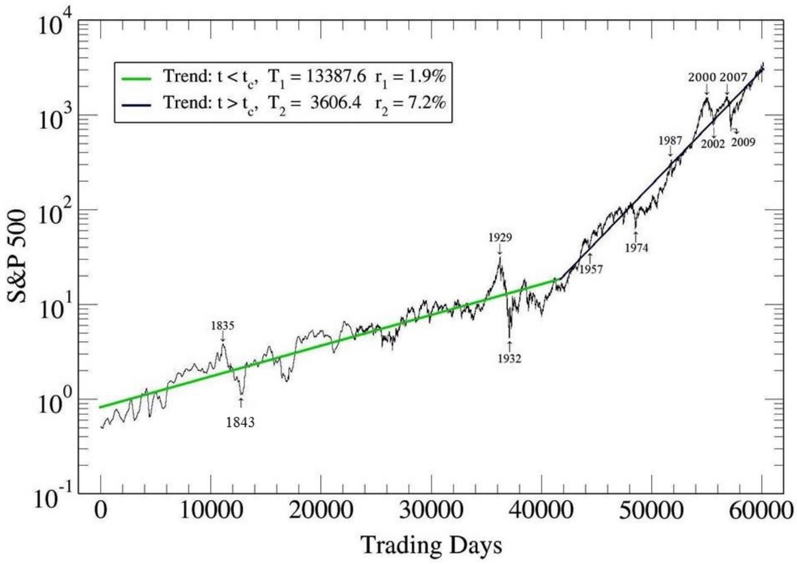

Figure 2. Two-period case: The fit obtained using Equations (2,3), yields $t_c = 41787$ days (1948). The first period (1789–1948) has $S_1(0) = 0.816$, $T_1 = 13387.6$ days, and $r_1 = 1.9\%$, and the second one (1948–2020), $T_2 = 3606.4$ days and $r_2 \simeq 7.2\%$.

The data representation shown in Figure 2 regarding the second period (1948–2020), is in agreement with the widespread approach commonly applied to financial data when discussing market returns over several decades. Here, we do not study higher-order period cases in general, but will concentrate on particular situations. To this end, we go on by studying "low frequency" (long time) phenomena such as population growth and GDP.

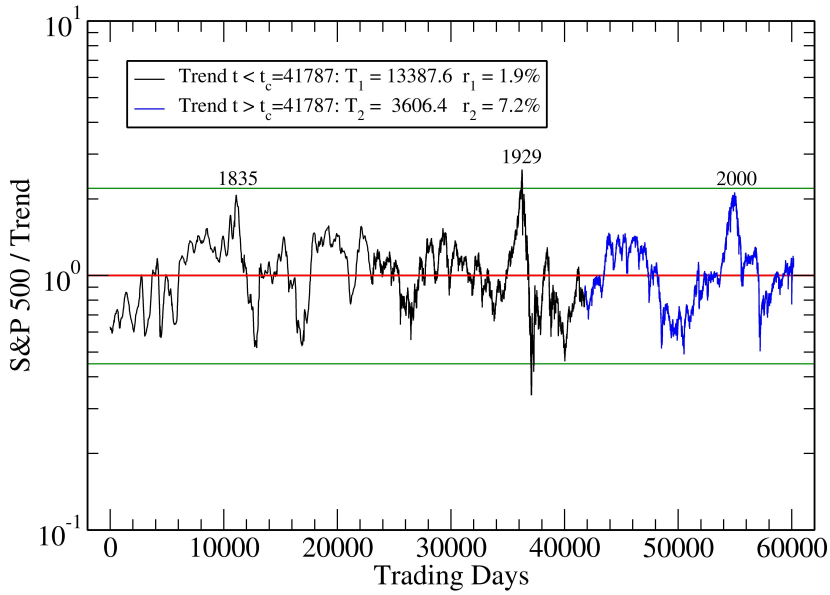

Figure 3. Detrended S&P 500 index, $S(t)/\langle S_i(t)\rangle$ vs trading days, using the trends $\langle S_{1,2}(t)\rangle$ obtained from Figure 2. Few years, corresponding to prominent events, are indicated. The horizontal lines correspond to the ratios $S(t)/\langle S(t)\rangle = 2.2$ and 0.45, as in Figure 1(b). Notice the apparent regular behavior of the data at the present time (60000 days).

3. Population growth and GDP data revisited

Time horizons of several decades are of central importance as we can correlate variations of fundamental quantities, such as population and GDP, to market outcomes on long time scales. Considering that the financial data we are analysing is barely 250 years long, 50 or 70 years represent a sensible fraction of the data, and their historical significance can be well expected to play a role in shaping the medium-long time scenarios we observe.

3.1. World and USA population data

Let us start by considering the world population growth from 1750 until today. The data is shown in a semi-logarithmic scale in Figure 4, which suggest the occurrence of well defined periods characterized by a constant growth rate, such that

$$P(t) \simeq P(t_0)\, \exp(t/T_{\rm P}) \tag{4}$$

where $T_{\rm P}$ is the characteristic time scale for population growth within the relevant time interval. The smoothness of the data allows us to identify three main growth periods: (1750–1942), (1942–1990) and (1990–2020), characterized by the growth rates, $r = 0.59\,\%$, $1.94\,\%$ and $1.23\,\%$, respectively. The second period lasted for about 50 years.

Clearly, the resulting global economic activity depends not only on the number of people, $P(t)$, but also on the wealth distribution among the population. Yet, $P(t)$ contains useful information to analize market behavior as discussed in Section 4. Since we are interested in the USA markets, we report in

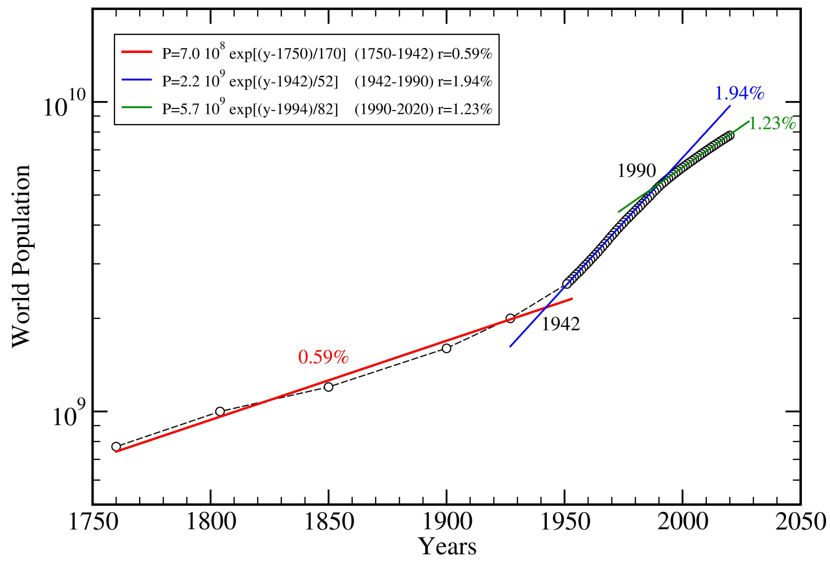

Figure 4. World population from 1750 to 2020. Three periods have been identified: (1750–1942) with a growth rate $r \simeq 0.59\,\%$; (1942–1990) with $r \simeq 1.94\,\%$, and (1990–2020) with $r \simeq 1.23\,\%$. Data from: *https://www.worldometers.info/world-population/world-population-by-year/*.

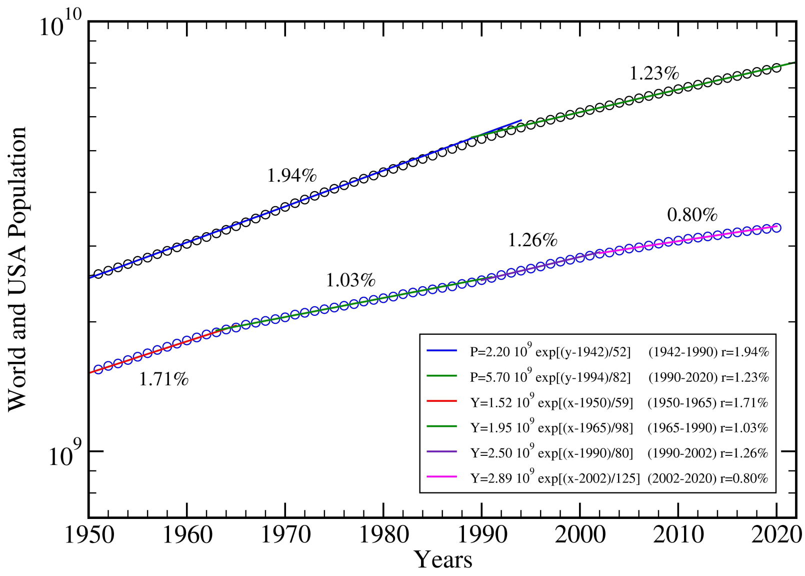

Figure 5. World and USA population from 1950 to 2020. The world data was taken from Figure 4. USA data has been multiplied by 10. For the latter, four periods have been identified: (1950–1965) with a growth rate $r \simeq 1.71\,\%$; (1965-1990) $r \simeq 1.03\,\%$; (1990–2002) $r \simeq 1.26\,\%$, and (2002–2020) $r \simeq 0.80\,\%$. Data from: *https://www.worldometers.info/world-population/world-population-by-year/*.

Figure 5 the corresponding population data within (1950–2020), which is most relevant to us. To be noted, is that the population growth in the USA is slower than in the world.

3.2. Market capitalization and GDP

Let us consider the second issue at hand, that is the interplay between the total market capitalization and the GDP. The basic quantity for the former is the so-called enterprise value, E_v, of a company, which is the amount of money required for its takeover. The enterprise value is calculated as,

$$E_v = M_s + M_d - C \tag{5}$$

where $M_s = N_s P_s$ is the market value of the company, obtained as the product of the number of outstanding shares, N_s, times the current share price, P_s. The second term, M_d, represents the total (both short- and long-term) debt of the company, and the last term, C, is the total cash, or cash equivalents, of the firm. It is taken with negative sign because the acquiring company keeps the cash of the target firm, which can be used to pay part of the debt.

To proceed further, we define the total debt simply as $M_D = M_d - C$, and the ratio of total debt, M_D, to entreprise value, as

$$r_D = \frac{M_D}{E_v} = \frac{M_D}{M_s + M_D} \tag{6}$$

with $0 \leq r_D \leq 1$, from which follows that $M_D = [r_D/(1 - r_D)]M_s$, so that we can express the enterprise value in terms of the market capitalization as follows,

$$E_v = \frac{1}{(1 - r_D)} M_s \tag{7}$$

Next, we relate the total enterprise value of all companies in a country, $\bar{E}_v$, to the total financial position, T_v, of the country, i.e.

$$\bar{E}_v = r_v\, T_v \tag{8}$$

so that the total market value becomes,

$$\bar{M}_s = (1 - r_D)\, r_v\, T_v = r_s\, T_v \tag{9}$$

Similarly, the GDP, denoted for convenience as G_v, is related to the total financial position as,

$$G_v = r_g\, T_v \tag{10}$$

Combining equations (9,10), we finally find $\bar{M}_s = (r_s/r_g)\, G_v$, so that the total market capitalization to GDP ratio, $R = \bar{M}_s/G_v$, is given by

$$R = (1 - r_D)\frac{r_v}{r_g} \tag{11}$$

Let us put some numbers to get an idea of the relative amount of the quantities involved. First, let us assume that $\bar{E}_v \approx G_v$, then $r_v \cong r_g$. In addition, we assume that, on average, each company has total debt $M_D \approx M_s/2$, yielding $r_D = 1/3$. Hence, we find $R = 2/3 \cong 0.66 < 1$. Clearly, the ratio becomes $R > 1$ when $r_v > r_g/(1 - r_D)$. The threshold $R = 1$ occurs when $r_v = r_g/(1 - r_D)$, which yields $r_v = (3/2) r_g$ in the case $r_D = 1/3$.

For illustration, we report in Table 1 few representative figures on the USA economy for the year 2014. As one can see, the ratio $R = 1.31$ exceeded the value one in 2014. We also include the notional

Table 1. USA financial data from 2014. First row: total financial position of USA (total worth, total debt and net worth). Second row: GDP and r_g ratio. Third row: $\bar{M}_s$ and r_s ratio. Fourth row: R ratio. Fifth row: the notional value of USA markets for the day (2014/12/26), total trade in 2014, and r_T ratio to total worth. Notional value (NV) is calculated by multiplying the execution price of each transaction by the total number of shares executed in each transaction (Data from *https://markets.cboe.com/us/equities/*). The total year Trade is obtained by multiplying NV by $D = 255$ trading days. All figures are in trillions USD.

USA (2014)	Worth = 270 T	Debt = 146 T	Net Worth = 124 T
GDP	17.52 T	$r_g = 0.065$	-
$\bar{M}_s$	23.00 T	$r_s = 0.085$	-
$R = r_s/r_g$	1.31	-	-
Notional Value	0.29 T	Trade = 74 T	r_T=0.274

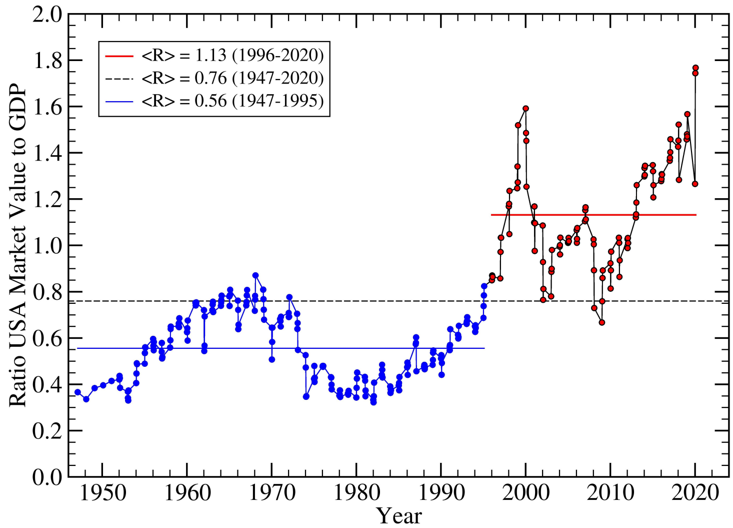

Figure 6. Ratio of USA market capitalization (corporate equities) to GDP vs Year. The horizontal lines represent the mean ratios within the period indicated in the inset. Data from *https://www.longtermtrends.net/market-cap-to-gdp-the-buffett-indicator/*.

value of USA markets, taken from the Chicago Board Options Exchange (CBOE) website. The historical data reported in Figure 6 show that values $R > 1$ have been reached only recently.

The mean ratios R found in Figure 6 can be obtained, for illustration purposes, by assuming again $\bar{E}_v \approx G_v$, yielding $R = 0.56$ when $r_D = 0.44$, while values $R > 1$ necessarily correspond to the cases $\bar{E}_v > G_v$. For instance, if $r_D = 0.1$, then $\bar{E}_v = 1.25\,G_v$ would yield $R = 1.13$; while $R = 1.31$ would correspond to $\bar{E}_v = 1.46\,G_v$, that is the total enterprise value would become 46% higher than the GDP. Similar behavior of R for other countries are shown in Figure 7, using data taken from Table 2 (2017–2018).

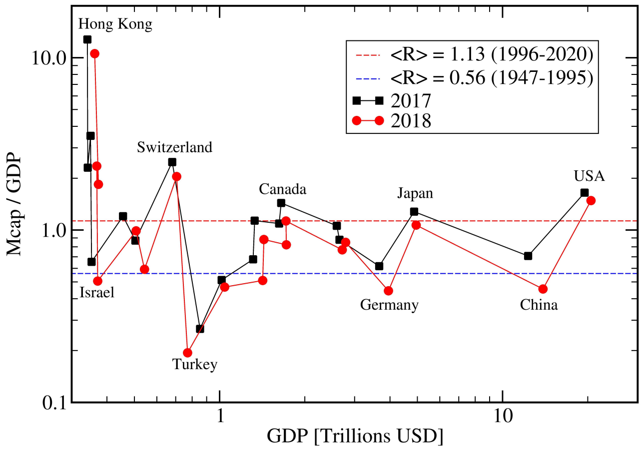

Figure 7. Ratios of total market capitalization to GDP vs GDP during 2017 and 2018, taken from Table 2. The mean ratios over the periods (1947–1995) $\langle R\rangle$=0.56, and (1996–2020) $\langle R\rangle$=1.13 (see Figure 6), are indicated by the dashed lines. Few countries names are displayed for convenience.

Table 2. GDP and total market capitalization (Mcap) for the years 2017 and 2018.

Country	GDP		Mcap		Mcap/GDP	
	2017	2018	2017	2018	2017	2018
United States	19.485	20.529	32.121	30.436	1.648	1.483
China	12.310	13.895	8.711	6.325	0.708	0.455
Japan	4.867	4.950	6.223	5.297	1.279	1.069
Germany	3.666	3.950	2.262	1.755	0.617	0.444
India	2.653	2.713	2.332	2.083	0.879	0.768
France	2.595	2.789	2.749	2.366	1.059	0.849
Canada	1.650	1.716	2.367	1.938	1.435	1.129
Korea	1.624	1.721	1.772	1.414	1.091	0.822
Australia	1.330	1.434	1.508	1.263	1.134	0.881
Spain	1.313	1.420	0.889	0.724	0.677	0.510
Indonesia	1.017	1.042	0.521	0.487	0.513	0.467
Turkey	0.853	0.771	0.228	0.149	0.267	0.194
Switzerland	0.680	0.705	1.686	1.441	2.480	2.044
Belgium	0.504	0.543	0.438	0.321	0.869	0.592
Thailand	0.456	0.507	0.549	0.501	1.203	0.989
Israel	0.353	0.371	0.231	0.187	0.654	0.506
South Africa	0.350	0.368	1.231	0.865	3.522	2.350
Singapore	0.342	0.373	0.787	0.687	2.303	1.841
Honk Kong	0.341	0.362	4.351	3.819	12.749	10.559
Vietnam	0.224	0.245	0.125	0.133	0.560	0.541

In Table 2, the mean ratios, R = Mcap/GDP, and standard deviations, σ, are: $\langle R \rangle = 1.78$, $\sigma = 2.70$ (2017); and $\langle R \rangle = 1.43$, $\sigma = 2.23$ (2018)[‡]. As a matter of fact, the total market capitalization can become larger than the GDP, even for several decades in a row (see Figure 6). This long persistance for $R > 1$ might, eventually, be considered the "rule", inducing investors to expect higher equity yields in the future.

It is clear that market activity is much more volatile than the country economic activity represented by the GDP. This can be already seen on the amount of daily trading assets: the notional value reported for (2014/12/26) in Table 1 is about 0.29 T, yielding an estimated total yearly trade of 27% of the whole USA financial worth of 270 T in 2014. In other words, total trading activity can exceed GDP by several multiples, in this example by a quadruple amount. Therefore, the resulting total market capitalization can vary conspicuously from year to year. Indeed, the GDP is a more slowly varying variable which is easier to estimate. Considering also that GDP represents actual trades of goods, while market cash flows are more proned to higher speculative behavior, the GDP can be taken as reference for determining the economic evolution over time.

We have thus retrieved GDP (and GDP/Capita) data from 1790 until 2020 and displayed them in Figure 8(a). One can distinguish a first period (1790–1950) during which the GDP has grown with an yearly rate of about 4.4%, while the GDP/Capita rate was about 1.9%. From 1950 until 2020, the growth rate was much higher, reflecting the pronounced economic activity after WWII. Since we are interested in the GDP growth during the second period (1950–2020), we show the corresponding GDP data in Figure 8(b). One can distinguish at least four periods within which the rate has remained pretty constant.

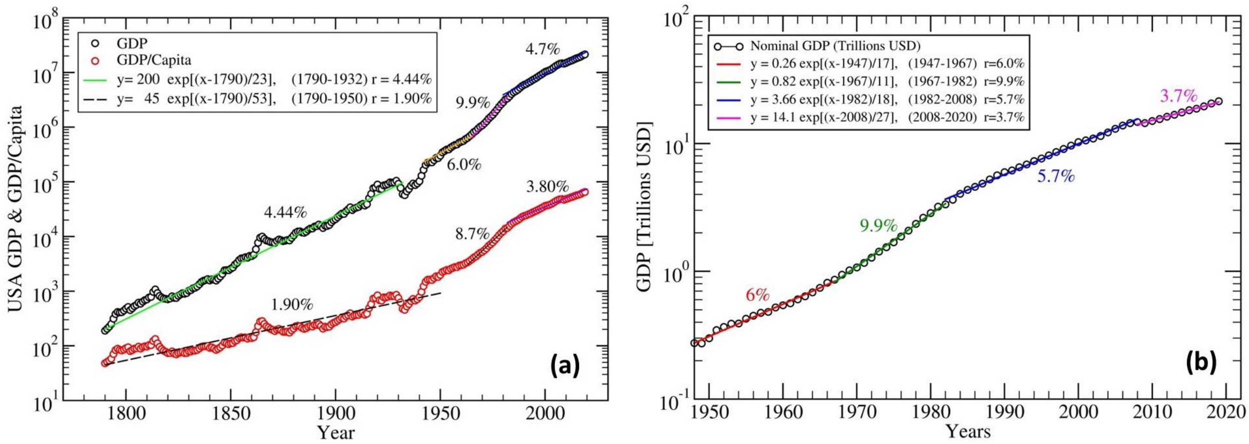

Figure 8. **(a)** USA GDP [USD] vs year (black open circles). And GDP/Capita [USD] vs year (red open circles). **(b)** GDP [Trillions] vs year in the period (1950–2020). The growth rates are indicated along the straight lines. The GDP and GDP/Capita data are a courtesy of *https://www.measuringworth.com/datasets/usgdp/result.php*.

These results are the basis for analyzing market returns displayed by the S&P 500 index, as we discuss next.

‡ Data from: *https://databank.worldbank.org*.

4. Historical growth rates for the S&P 500 index, World and USA population, and USA GDP

In order to discuss the issue of valuation of the S&P 500, we compare first its corresponding growth rates with those of the population and GDP data, within selected periods of time for which the latter display approximately constant rates. By looking back to the historical data (Figure 2, Figure 4, Figure 5 and Figure 8), we realize that the three main figures, S&P 500, World population and USA GDP, all display approximate single growth rates within the period (1790–1950), suggesting a kind of stationarity in the economic activity. The rates, however, have changed from 1950 until 2020 (see Table 3), ostensibly for debatable reasons.

Table 3. Empirically determined growth rates [%] for: S&P 500 (Figure 2), World population (Figure 4), USA population (Figure 5), USA GDP (Figure 8), and their ratios indicated within square brackets.

	(1790–1942)	(1942–1990)	(1990–2020)		
S&P 500	1.90 [3.22]	6.20 [3.20]	7.33 [5.96]		
World POP	0.59	1.94	1.23		
	(1900–1965)	(1965–1990)	(1990–2002)	(2002–2020)	
S&P 500	10.7 [6.26]	5.35 [5.19]	13.3 [10.6]	7.73 [9.66]	
USA POP	1.71	1.03	1.26	0.80	
	(1790–1932)	(1947–1967)	(1967–1982)	(1982–2008)	(2008–2020)
S&P 500	1.90 [0.43]	10.5 [1.75]	1.60 [0.16]	10.1 [1.77]	11.2 [3.03]
USA GDP	4.44	6.00	9.90	5.70	3.70

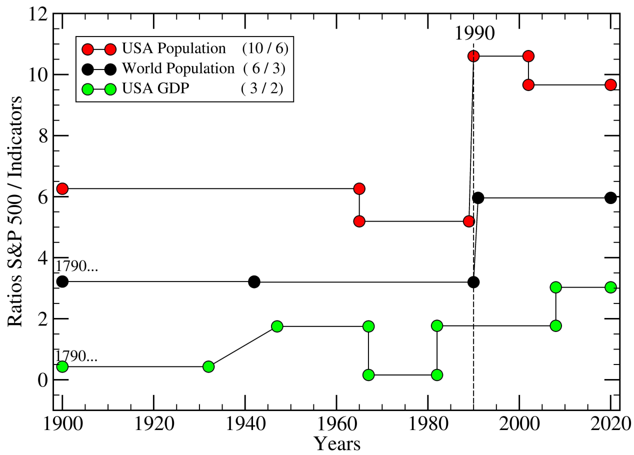

Figure 9. Ratios of S&P 500 to macroeconomic indicators from Table 3. In the inset, we display the rounded mean values for the historical ratios calculated before and after 1990. For instance, for the S&P 500-USA population, we find a mean ratio R_{USApop}~6 before 1990, and R_{USApop}~10 after 1990, yielding a mean increment of ~1.7 from the first period. In the same way, we obtain increments of ~2 and ~1.5 for the S&P 500-World population and S&P 500-USA GDP, respectively.

Based on the results displayed in Table 3 for the S&P 500 and World population data, we have replotted the former as shown in Figure 10. This approach suggests that the post-WWII growth rates were consistent, until 1990, with the previous long (1790–1945) period found in Figure 2 and Figure 4. For the sake of illustration, we extended the rate $r_2 = 6.2\,\%$ until 2020[§] (the red dashed line in Figure 10), which can be compared with the fitted rate, $r_2 \cong 7.2\%$, obtained in Figure 2.

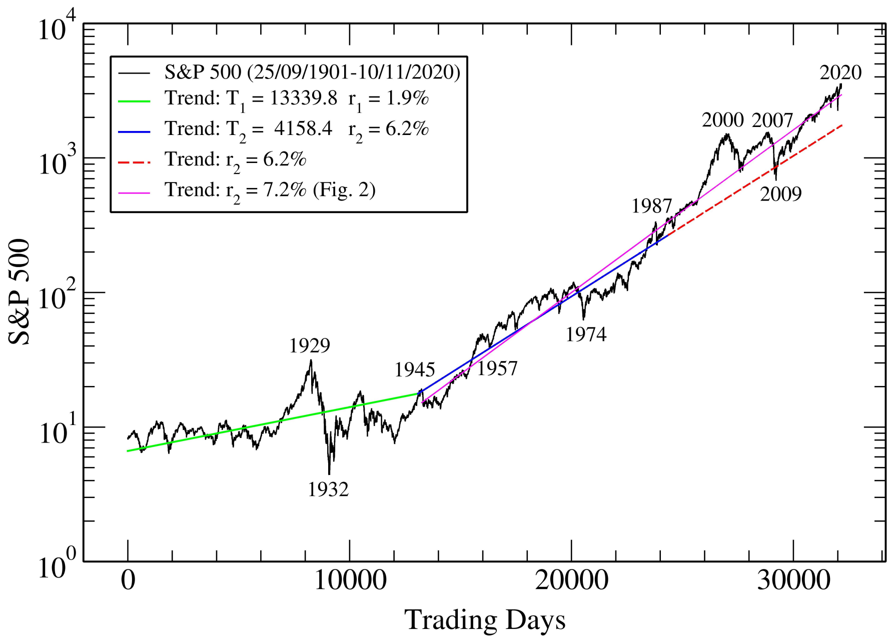

Figure 10. S&P 500 fits until 1990, as suggested by the world population growth rate (Figure 4), where now the crossover year is found to be 1945 (cf. Figure 2): (green line) growth rate $r_1 = 1.9\,\%$ (1900–1945), and (blue line) $r_2 = 6.2\,\%$ (1945–1990) (see Table 3). The value r_2 has been extended to the range (1990-2020) as illustration (dashed red line), which can be compared with the rate, $r_2 = 7.2\,\%$ (thin magenta line), expected from Figure 2.

The departure observed for the latter suggests that after 1990 the economy, and financial activity, turned into a new regime, distinct from the historical trend observed until 1990. Whether this departure from the historical trend is due to a kind of market bubble, or similar phenomenon, needs to be considered in some detail. We proceed therefore with the issue of modeling market bubbles in the next section.

5. Modeling market bubbles

In the following, we discuss a model of a stock market bubble as an attempt to understand the origin of assets overvaluation quantitatively. To this end, we introduce the concept of virtual market index, as shown below. To deal with a stock price time series, it is convenient to use the logarithm, $\log P_i(t)$, of the price $P_i(t)$ of stock i at time t, where $1 \leq i \leq N$, and N is the total number of different stocks considered. We take $N = 500$, as for the S&P 500. For simplicity, we use a discrete time process, $t = nt_0$, where $n \geq 0$ is an integer, and $t_0 = 1$ day, yielding $t = 0, 1, 2, 3, \ldots, T$, with $T = 20000$ (~78 years). The

§ A conservative lower bound would be $r_2 \cong 3.2 \times 1.23 \cong 4\,\%$ (see Table 3).

daily price variations of stock i are then given by the log-returns,

$$\Delta \log P_i(t) = \log P_i(t) - \log P_i(t-1) = x_i(t) + \beta_i \tag{12}$$

where $x_i(t)$ are i.i.d. random numbers having zero mean and standard deviation σ_i (the "stochastic" component of the price), while β_i is a constant daily growth rate of the price (the "deterministic" part). Hereafter, we consider a purely stochastic situation and take $\beta_i = 0$.

We model the variations $x_i(t)$ using a simple ARCH(1) model, originally introduced by Engle (1982), which are drawn from a Gaussian distribution with variable variance, $\sigma_i^2(t) = a + b\, x_i^2(t-1)$, such that $a > 0$ and $0 \le b < 1$. This process has therefore a mean variance $\langle \sigma^2 \rangle = a/(1-b)$. The model reproduces the phenomenon of clustering of volatility and, in addition, it has the advantage of yielding fat tails probability distribution functions, as observed empirically, i.e. $\mathcal{P}(x) \sim |x|^{-(1+\alpha)}$ for $|x| \gg 1$, where the asymptotic exponent $\alpha > 2$ depends on b in an exact analytical form (see e.g. Dose et al. (2003)). It is also known that simple ARCH models do not describe the long-time correlations of volatility, since they lack of long-range memory, features which need to be introduced explicitely (see e.g. Roman and Porto (2008)). To describe more accurately the behavior of stock prices, additional features need to be implemented (see e.g. Roman et al. (2008)), which are beyond the scope of the present work.

From Equation (12), it is easy to show that the price of stock i at time t, is given by

$$P_i(t) = A_i \, \exp\left(\sum_{n=1}^{t} x_i(n) \right), \quad t \ge 0 \tag{13}$$

where $P_i(0) = A_i$, and we take $A_i = 1$ for all stocks[¶]. In a real stock market, correlations among stocks, and between stocks and market index are present, which are not explicitely expressed in Equation (13). The former can be simulated by choosing the stock price log-returns, $x_i(t)$, to be dependent on each other. This implies the use of several parameters in the model, a complication that we want to avoid at this stage. Therefore, we are left with modeling only correlations between stock prices and the index. Notice that in this way, there will also result, a posteriori, correlations among stocks "mediated" by their individual correlations with the index.

Historically, the capital asset pricing model (CAPM) was introduced to describe stock price returns, based on the efficient market hypothesis (EMH) (see e.g, Markowitz (1959); Sharpe (1964); Black (1972); Fortune (1991)), which are a function of market returns[‖]. More advanced models were proposed soon afterwards (see e.g. Mandelbrot (1997)). Recent developments regarding multifractality and Elliott waves features have been discussed (see, e.g. Drożdż et al. (2003, 2018); Gündüz (2021)). The CAPM is in agreement with the empirical observation that, in our terminology, stock log-returns are correlated to the market index log-returns (see e.g. Bonanno et al. (2003); Roman et al. (2006)).

¶ Similarly as done for the admittance of a stock into the S&P 500, we use a simple criterion to include a time series into our market index. The condition for the latter is that the sum, $\sum_{n=1}^{T} x_i(n) > 0$, meaning that $P_i(T) > 1$. This introduces a bias into our market index, but it seems reasonable for our purposes.

‖ In the CAPM, price returns R_i of stock i are a function of market return, R_m, plus a free-risk return R_f, such that, $R_i(t) = (1-\beta_i)R_f + \beta_i R_m(t) + \varepsilon_i(t)$, where $0 \le \beta_i \le 1$ is the covariance factor of stock i and market fluctuations, and ε_i is an uncorrelated random number with zero mean.

For the purpose of describing a bubble phenomenon, we introduce first a "virtual" market index, here denoted as $S(t)$, constructed by adding the values of P_i, according to,

$$S(t) = \sum_{i=1}^{N} \alpha_i P_i(t) \tag{14}$$

where α_i are the weights of each stock within the index. For simplicity, and without loss of generality, we simply take $\alpha_i = 1/N$, i.e. the same weight for all stocks. Therefore, the virtual market log-returns are given by,

$$\Delta \log S(t) = \log S(t) - \log S(t-1) \tag{15}$$

In oder to correlate the stock prices, Equation (13), to the virtual market log-returns, Equation (15), we add an additional parameter, $\gamma(t)$, which we denote as the decoupling factor for log-returns, and it is assumed, for the sake of simplicity, to be the same for all stocks. According to this new parameter, Equation (13) is now changed to the form,

$$P_i(t) = \exp\left(\gamma(t) \sum_{n=1}^{t} x_i(n) + (1-\gamma(t))\, \Delta \log S(t)\right), \quad t \geq 0 \tag{16}$$

where $\gamma_0 \leq \gamma(t) \leq 1$ is assumed to be given, and γ_0 needs to be determined. In this way, the stock price is coupled to the virtual index variation, Equation (15), via the factor $(1-\gamma)$ in Equation (16). In more involved applications, γ can depend on the stock, possibly drawn from a distribution function $\mathcal{P}(\gamma)$. This extension of the model goes beyond the scope of the present work and it will be considered elsewhere.

Notice that Equation (16) reduces to Equation (13) when $\gamma = 1$, i.e. in the case of maximum decoupling factor, and the prices become decoupled from the virtual index variations. To be noted is that the price log-returns obtained from Equation (16) take a quite complicated form. However, the system of Equations (14,15,16) admits a simple solution. Substituting Equation (16) into Equation (14), we obtain,

$$S(t) = S(t-1)\left(\frac{S_0(t)}{S(t-1)}\right)^{1/\gamma(t)}, \quad \text{for } t \geq 1 \tag{17}$$

where

$$S_0(t) = \sum_{i=1}^{N} \alpha_i \exp\left(\gamma(t) \sum_{n=1}^{t} x_i(n)\right) \tag{18}$$

which can be solved iteratively for $S(t)$ with the condition $S(0) = 1$. Notice that $S(t)$ coincides with $S_0(t)$ in the uncorrelated case, i.e. $\gamma = 1$. The exact expression for $S(t)$ can be obtained in close form by iterating Equation (17). This is done for the purpose of determining the lower bound γ_0 analytically. In practical applications, it is more convenient to work directly with Equation (17). After a little algebra we find,

$$S(t) = \prod_{m=0}^{t-1} [S_0(t-m)]^{\theta(m,\gamma)} \tag{19}$$

where $\theta(m, \gamma(t)) = (-)^m z_{t-m} \prod_{n=0}^{m-1}(z_{t-n} - 1)$, for $m \geq 1$, and $\theta(0, \gamma(t)) = z_t$, with $z_t = 1/\gamma(t)$. It is clear that in order to keep $S(t)$ bounded as t grows, the exponent θ should satisfy the condition, $|\theta(m, \gamma)| \leq 1$, which yields $z_t \leq 2$**, or equivalently, $\gamma(t) \geq \gamma_0 = 1/2$. This can be seen more clearly in the case of constant $z = 1/\gamma$, yielding $|\theta(m,\gamma)| = z(z-1)^m$, implying that, $z - 1 \leq 1$.

As one can see from Equation (16), stock price variations diminish in amplitude (lower volatility) when $\gamma(t) < 1$, becoming closer to the change of the virtual market index, $\Delta \log S(t)$. We may interpret this "collective behavior" as a self-consistent stock price adjustment process acting during a trading day. In the extreme case $\gamma = 1$, no price adjustments take place and stocks behave independently of each other. We expect that a "regular" market behavior corresponds to small values of γ (i.e. $\gamma_0 \leq \gamma \ll 1$), while bubble like phenomena should occur when $\gamma \to 1$.

Illustrative examples of bubble behavior of $\gamma(t)$ are displayed in Figure 11. In case (1), a "peak" around $t = 1.1\,10^4$ occurs, spanning an anomalous period of about 12 years, while outside it the market behaves regularly with a constant decoupling factor $\gamma = 0.6$ (black line). In case (2), the decoupling factor changes from its regular value, $\gamma = 0.6$, to the extreme one with $\gamma = 1$, at longer times (dashed green line).

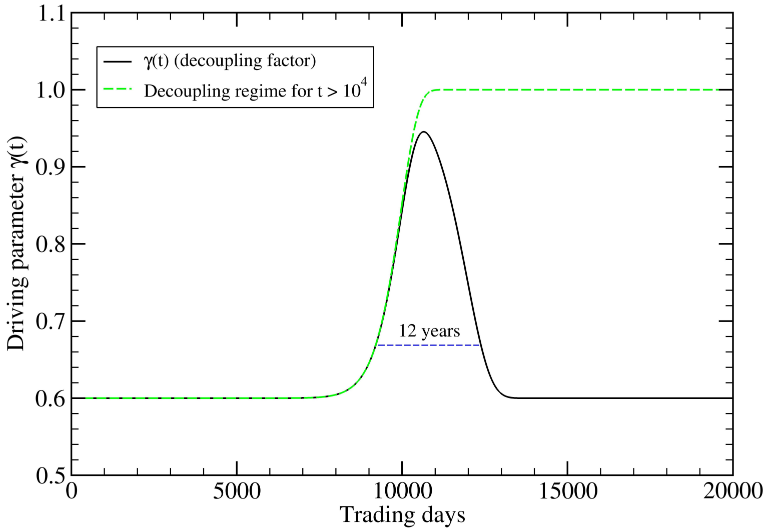

Figure 11. Market driving parameter $\gamma(t)$ vs trading days. Case (1): (Black line) A decoupling factor peak, described by the analytical form $\gamma(t) = \gamma_1 + \Delta\gamma f(t) \exp[-(t/\tau_1)^{20}]$, where $f(t) = [1 - \exp(-(t/\tau_0)^{20})]$, with $\gamma_1 = 0.6$, $\Delta\gamma = 0.4$, $\tau_0 = 10^4$ and $\tau_1 = 1.2\,10^4$; Case (2) (Green dashed line) A single step function for $\gamma(t)$, changing between the value $\gamma(t) = 0.6$ at smaller times, to $\gamma(t) = 1$ (decoupling regime) for $t > 1.1\,10^4$, obtained by setting $\tau_1 \to \infty$.

To make contact with real market data, we define a real market index, $S_r(t)$, as a weighted sum of stock prices, $P_i^{(r)}(t)$, according to,

$$S_r(t) = \sum_{i=1}^{N} w_i\, P_i^{(r)}(t) \tag{20}$$

where w_i are the stock weights, such as those defining the S&P 500 index. The form of the stock prices $P_i^{(r)}(t)$ can be obtained by making contact with the CAPM, to which our model should reduce when

** Actually, z_t can become larger than 2 temporarily. In such cases, $S(t)$ displays a stronger oscillating behavior than in the case $z_t \leq 2$.

stocks are not correlated with the virtual market index, that is when $\gamma = 1$. To this end, we define,

$$P_i^{(r)}(t) = P_i(t)\, S(t)^{\gamma_r} \tag{21}$$

where γ_r is a constant, and $P_i(t)$ is, according to Equations (16,17), given by,

$$P_i(t) = P_i(0) \left(\frac{S_0(t)}{S(t-1)}\right)^{[1-\gamma(t)]/\gamma(t)} \exp\left(\gamma(t) \sum_{n=1}^{t} x_i(n)\right) \tag{22}$$

Now, in the case $\gamma = 1$, we have $S(t) = S_0(t)$ (Equation (18) with $\gamma = 1$), and

$$P_i^{(r)}(t) = P_i(0) \exp\left(\sum_{n=1}^{t} x_i(n) + \gamma_r \log S_0(t)\right) \tag{23}$$

yielding the log-returns,

$$\Delta \log P_i^{(r)}(t) = x_i(t) + \gamma_r \Delta \log S_0(t) \tag{24}$$

similarly as for the CAPM, where γ_r plays the role of the covariance factor β_i in the CAPM. In the following, we use Equations (20,21) as our market model for stocks, generalizing the CAPM in terms of a self-consistent virtual market process driven by the parameter $\gamma(t)$. For illustration, we consider $N = 500$ surrogate time series as described in Figure 12.

In Figure 12(a), we display the "regular" market behavior with $\gamma = 0.6$, yielding a good "mix" of prices and virtual index variations, and we use a constant $\gamma_r = 0.4$. The example corresponds to a full time span of about 78 years, and 4.3% yearly growth rate. The case in Figure 12(b) corresponds to full decoupling of prices and index, i.e. $\gamma(t) = 1$. Notice the dramatic grow of the index (8.2% yearly growth rate). In Figure 12(c), we display the effect of a single bubble produced by using the time dependence for $\gamma(t)$ reported in Figure 11 (black line), in the case $\Delta\gamma = 0.3$. During the bubble period, the index attempts to reproduce the behavior observed in Figure 12(b), turning back to the regular behavior after the bubble has burst. In Figure 12(d), we show the (self-consistent) market log-returns for the regular market situation (black line), and the excess bubble volatility (red line) confined to the bubble period.

This model suggests a mechanism for the emergence of a bubble in a stock market, governed by the decoupling or mix parameter γ. When $\gamma_0 < \gamma < 1$, stock prices are self-consistently coupled to the virtual market index, and as a result, their variations are "attenuated" by the market index. Bubbles can occur when $\gamma \to 1$, i.e. when prices get uncorrelated from the virtual market index.

The case $\gamma = 0.6$, discussed in Figure 12, may represent a sort of "regular" behavior of the market in which prices evolve in a self-consistent fashion. In other words, one can argue that such self-consistent price adjustments have the effect of keeping the market value under a more uniformly distributed influence of the different economic sectors, represented by the stocks composing the index. This, in turn, would result in more accurate valuations of single stocks based on global economic scenarios closely related to real economy data such as the GDP. In contrast, when $\gamma \to 1$, stocks decouple from the mean market behavior producing exacerbated price variations not conforming to real macroeconomic data. In these cases, the problem is that also the whole market grows outside a realistic range of variation, leading to the formation of a bubble. So the question is: What can cause the decoupling factor to increase so dramatically to generate a bubble? This question is considered in the next section in relation to real market data.

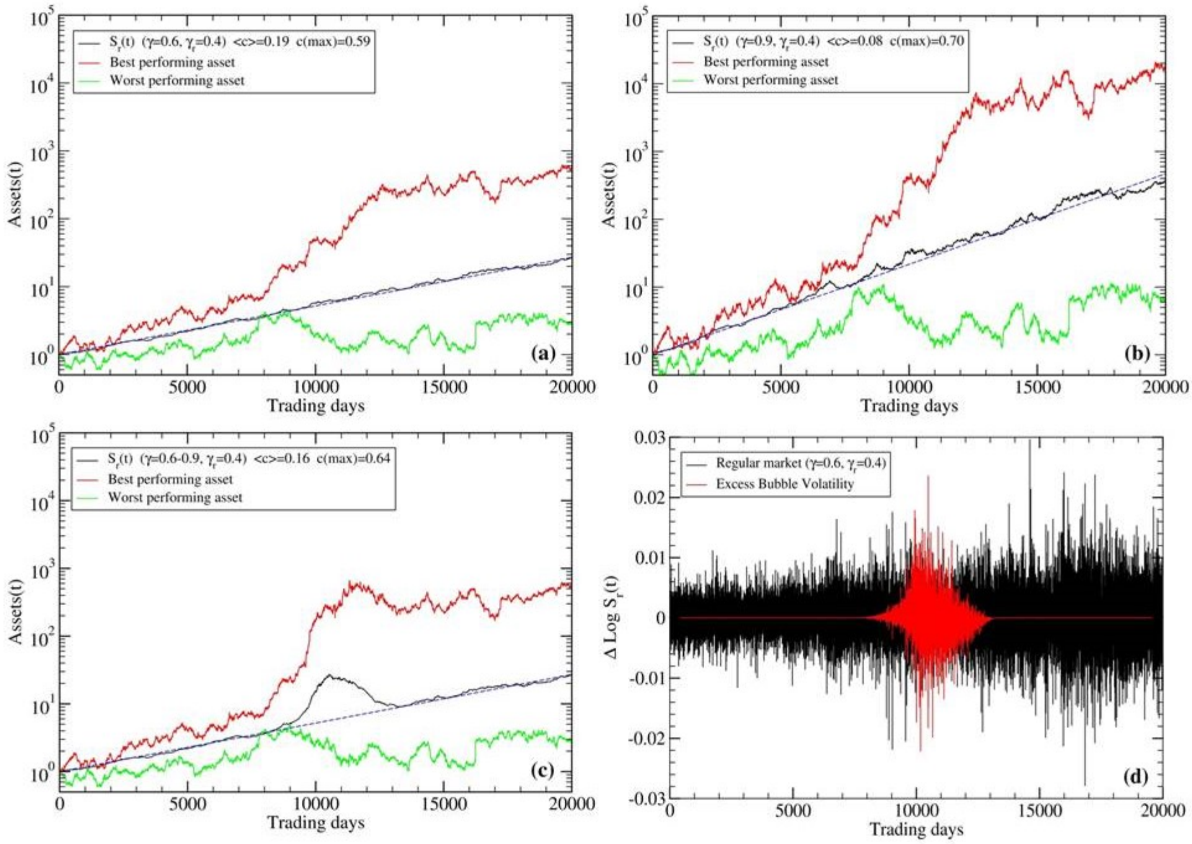

Figure 12. An S&P 500 market, $S_r(t)$ vs trading days t (black lines), from 500 surrogate time series with bubble volatility effects. The ARCH parameters are: $a = 2 \times 10^{-4}$, $b = 0.6$, yielding $\sigma = 0.022$. (**a**) Regular market with $\gamma = 0.6$ and $\gamma_r = 0.4$. The weights w_i are those typical of the S&P 500. The red line displays the best performing stock, and the green line the worst. The dashed line is a fit ($y = \exp(ax)$, $a = 1.65\ 10^{-4}$) representing 4.3% yearly growth rate. (**b**) Same as in (a) for $\gamma = 0.9$, the fit yields $a = 3.07\ 10^{-4}$, i.e. a higher yearly growth rate of about 8.2%. (**c**) The single bubble case with $\gamma(t)$ ($\Delta\gamma = 0.3$) taken from Figure 11 (black line). The dashed line is the same as in (a). (**d**) The market log-returns, $\Delta \log S_r(t)$ vs time, for the regular market in (a) (black line). The red line yields the excess market log-returns within the bubble period from (c).

6. Discussion

The analysis of historical data from Section 3 and the results presented in Section 4, suggest that the S&P 500 might be overvalued since the 1990s. This apparent overvaluation can be interpreted as a form of a "persistent" bubble regime, a situation that can be suitably explained by the model of market bubbles presented in Section 5. Indeed, the market behavior from 1790 until 1990 seems be well described by two periods (1790–1945) and (1945–1990) characterized by different yearly growth rates of the S&P 500, but consistent with the growth of world population during those epochs. The situation after 1990 appears less predictable in terms of the past returns, and additional information entering the market needs to be taken into account.

In order to investigate the overvaluation issue in detail, we resume data from 1956 until 2020 in Figure 13, by considering the role played by the USA Federal Reserve (FED), created in 1913. Notably, the latter has at its disposal, among others, the powerful instrument of the Federal Funds Rate (FFR) which can be adjusted to the current economic scenario. The FFR is the target interest rate set by the Federal Open Market Committee (FOMC) at which commercial banks borrow and lend their excess reserves to each other overnight. The FOMC cannot force banks to charge the exact FFR, but it is considered as a guidepost. Besides the FFR, the Federal Reserve also sets a direct discount rate (DDR), which is the interest rate charged to a bank that borrows directly from the FED. This rate is typically higher than the target FFR, in part to encourage banks to borrow from other banks at the (lower) FFR. The Federal Funds Rate can influence short-term rates on consumer loans and credit cards, as well as having an impact on the stock market.

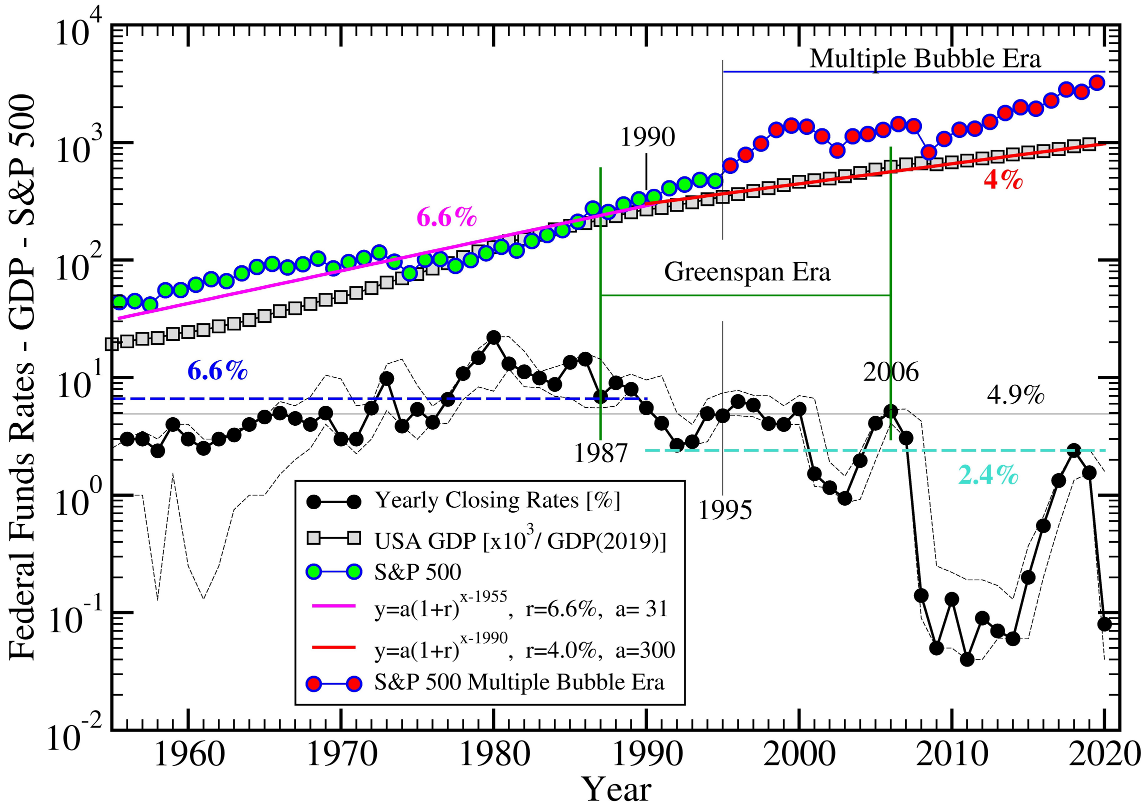

Figure 13. Federal Funds Rates (full black circles) and S&P 500 index (full green circles) vs year. The mean rate 4.9% over the period (1956–2020) is indicated by the continuous line. The thin black dashed lines display the maximum and minimum rates for the year. The mean rate within (1956–1990) is 6.6% (blue dashed line), and 2.4% within (1990–2020) (turquoise dashed line). The GDP is indicated by the shaded squares, and the data have been shifted for convenience. We have included two lines corresponding to growth rate 6.6% (magenta line) and 4% (red line), as a guide. For comparison, the period (1987–2006), during Greenspan's chair at the FED, is displayed. The dubbed "Multiple Bubble Era" (1995–2020, red circles) is highlighted.

We display in Figure 13 the time track of FFRs over the period (1956–2020). We take the year 1990 as the reference date separating the two subperiods, (1956–1990) and (1990–2020), within which we calculate the mean FFR, yielding 6.6% and 2.4%, respectively, while the mean FFR over both periods amounts to 4.9%. To be noted is that, within the first period (1956–1990), the GDP and S&P 500 had a similar growth rate, consistent with the FFR mean growth rate of about 6.6%, the latter reaching a peak of about 22% in 1980 due to the huge inflation rate at that time (see below). Indeed, inflation and GDP growth go hand in hand, in the sense that lack of inflation can be an indication of poor economic growth.

During the late 1980s and early 1990s (second subperiod), the GDP growth rate starts to decrease, going from a peak at 9.9% until 1985, down to a slower growth rate of about 5.7% until 2008 (cf. Figure 8(b)). The GDP is shown in Figure 13, clearly manisfesting its slowdown around 1987. Coincidentally, Greenspan's chair at the FED (1987–2006) started its duties few months before the market crash from October 1987. A closer inspection to Figure 13 suggests a possible long-term cause for the market meltdown, that is, the incipient GDP slowdown, in addition to the short-term market overvaluation. To be noticed is that the FFR went from ≈10% in 1986 down to 7% in 1987.

What happened next? We entered the Greenspan's era. The 1990s were characterized by the technology boom, similarly to what we are experiencing today. In contrast, the whole economy was slowing down, and the GDP was growing at a lower 4% annual rate. Despite this still high rate, the FED decided to cut the FFR, reaching a minimum of 2.4% in 1992. Thus, the market kept increasing, and the S&P 500 reached its all times high in 2000, at the peak of the internet bubble. The latter was followed by a second bubble, as a result of a second round of drastic FFR cuts in 2002, which burst in 2008, initiating the so-called great recession. Hence, the FFR was cut down close to zero around (2010–2011) (see Figure 13), provoking the birth of the third bubble in the series, starting in 2009. If we look back at our prediction in Figure 10(b), we can interpret each bubble burst as an attempt to reach the expected growth rate of about 4%. This was not possible due to the strong actions taken by the FED to timely and agressively cutting the FFR[††].

Theoretically, we may ask whether the above picture is consistent with bubble phenomena as described by the model in Section 5. From a mathematical perspective, one can model the occurrence of consecutive "bubbles", such as the set dubbed multiple bubble era in Figure 13. It is possible to find the corresponding decoupling factor $\gamma(t)$ to get a sequence of (in this case) three bubbles. For our present purposes, we may argue that γ can be very sensitive to FFR variations. In particular, we expect that, to first order, $\Delta\gamma(t) \approx -\Delta r_{\mathrm{FFR}}(t)$, suggesting that decreasing FFRs (increasing γ) may result in a temporary decorrelation of stock prices from their "virtual market" mean, leading to the issues discussed in Section 5. This will be discussed in detail elsewhere.

Finally, we consider the interplay between FFR and CPI, in an attempt to complete our dicussions from a different point of view. The question remains whether the index departures from historical trends shown in Figure 10 and Figure 13 can be attributed to some extent to inflation. We have therefore plotted in Figure 14(a) values of CPI together with the FFR. One can see that CPI have undergone strong variations during the 1970s, stabilizing since 1980. This behavior can not explain the 2000 bubble and later anomalies. One can also observe that FFR have been typically larger than CPI, except in particular periods (1970s and 2010s). To appreciate the behavior more clearly, we have plotted in Figure 14(b) the ratios CPI/FFR over the years.

[††] We do not discuss the reasons argued by the FED to cut rates.

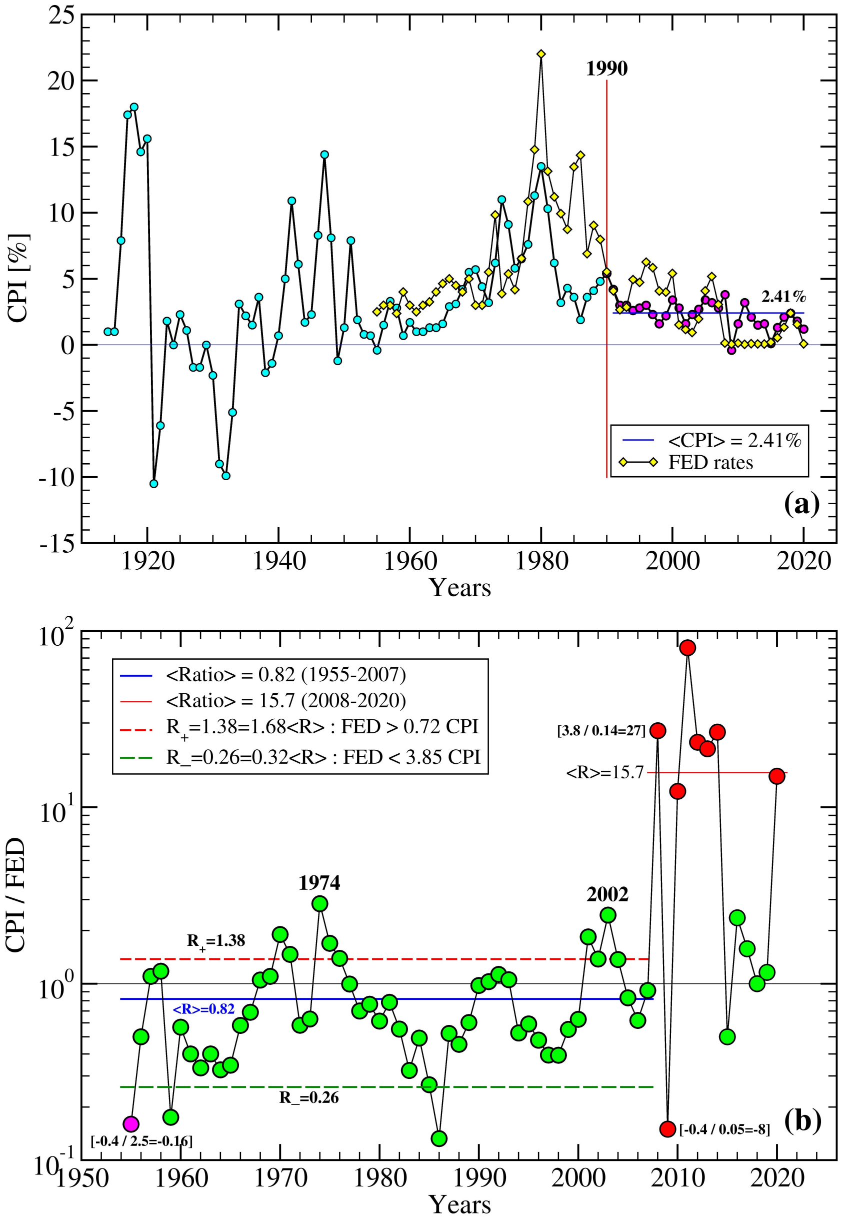

Figure 14. (a) Inflation data representing the CPI within (1914–2020) and FFR within (1955–2020) vs year. The mean CPI in the range (1990–2020) is about 2.41 % (blue line). **(b)** Ratios CPI/FED from (a). The mean ratios are given by: $\langle R \rangle = 0.82$ (1955–2007), and $\langle R \rangle = 15.7$ (2008–2020). The S.D. of R within the first period is $\sigma = 0.52$, from which we obtain $R_{\pm} = \langle R \rangle \pm \sigma$, yielding $R_{+} = 1.38$ and $R_{-} = 0.26$. Negative ratios, in 1955 and 2009, are indicated together with the associated values of CPI and FFR above 0.1 for illustration purposes. The green circles can be expected to represent "appropriate" ratio values, while the red ones, extremely anomalous cases.

The latter show two distinct periods: The first one (1955–2007) displays a mean ratio $\langle R \rangle \cong 0.82$ and standard deviation (S.D.) $\sigma \cong 0.52$; while the second period (2008–2020) has $\langle R \rangle \cong 15.7$, i.e. about

20 times larger. Notice that around the 2000 bubble, the FFR moved above 1σ, but turned down soon afterwards. The advent of the great recession was detrimental to the economy and the FED responded aggressively. Regarding the period (1990–2020) of interest to us, we may conclude that inflation was playing a minor role and the causes for a possible overvaluation of the market can not be attributed to the behavior of the CPI. Rather, the monetary policy adopted during the latest crises had a major impact.

From a technical point of view, one can consider to define a ratio, $R = \text{CPI/FFR}$, such that it is always positive. This would imply that when the CPI turns negative, so should the FFR do. In the particular case of 2009, when the CPI was −0.4 %, one could have considered the use of a negative FFR. Indeed, negative rates are not unusual at present, and in this particular example a positive ratio could have been obtained for FFR in the range, $-1.55\% < \text{FFR} < -0.29\%$, in order to bring R within its historical 1σ range, $0.26 < R < 1.38$. Once the CPI turns positive again then the use of negative rates can be considered to be over. The so-called "repurchase agreements" (Repo overnight rates), for instance, have indeed turned negative quite often in recent weeks.

7. Conclusions

We have analyzed USA financial data from 1789 until 2020, mainly focused on the S&P 500 index. We discussed two different scenarios describing the evolution of the market index over long time scales. One in which the whole history is regarded, in a first order approximation, as a single economic period having a mean yearly growth rate of about 3%, and a second order one, split into two periods (1789–1948) and (1948–2020), characterized by 2% and 7% growth rates, respectively. We then made a long tour to correlate, within the second scenario, long-time market behavior to macroeconomic data, such as GDP and population growth.

From the discussions in Section 3 and Section 4, we conclude that the S&P 500 might be considered to be overvalued since the 1990s, having entered a period of successive bubbles. To understand the origin of bubbles, we introduced a model in Section 5 showing that strong market upturns might be the result of a lack of correlations between stock prices and a virtual market index, the latter calculated self-consistently from the stock prices. We introduced a parameter γ, denoted as the decoupling factor, which leads to a bubble in the limit $\gamma \to 1$, while a regular market behavior is obtained for lower values of γ, somewhere in the range $1/2 \lesssim \gamma < 1$. In Section 6, we argued that variations $\Delta\gamma$ are possibly anticorrelated to variations in the Federal Funds Rate, i.e. $\Delta\gamma(t) \approx -\Delta r_{\text{FFR}}(t)$, eventually triggering a bubble phenomenon when rate cuts become too pronounced and persist over long times. We have provided examples of such events from the early 1990s, suggesting that warning signs were present against the proliferation of too low FFRs. The estimations performed in Section 4 suggest that at the end of 2020 the S&P 500 might have been conspicuously overvalued with respect to historical fundamental trends.

Acknowledgments

We would like to thank M. Bonesini and R. Benocci for illuminating discussions. The authors would like to thank the reviewers for their useful and constructive comments.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

Alexius A, Spång D (2018) Stock prices and GDP in the long run. *J Appl Financ Bank* 8: 107–126.

Balcilar M, Gupta R, Wohar ME (2017) Common cycles and common trends in the stock and oil markets: Evidence from more than 150 years of data. *Energy Econ* 61: 72–86.

Black F (1972) Capital market equilibrium with restricted borrowing. *J Bus* 45: 444–455.

Bogle JC (2017) *The Little Book of Common Sense Investing: The Only Way to Guarantee Your Fair Share of Stock Market Returns*, Prabhat Prakashan.

Bonanno G, Caldarelli G, Lillo F, et al. (2003) Topology of correlation-based minimal spanning trees in real and model markets. *Phys Rev E* 68: 046130.

Caspi I, Katzke N, Gupta R (2014) Date stamping historical oil price bubbles: 1876–2014. Stellenbosch Economic Working Papers: 20/14.

Demirer R, Demos G, Gupta R, et al. (2019) On the predictability of stock market bubbles: evidence from LPPLS confidence multi-scale indicators. *Quant Financ* 19: 843–858.

Dose Ch, Porto M, Roman HE (2003) Autoregressive processes with anomalous scaling behavior: Applications to high-frequency variations of a stock market index. *Phys Rev E* 67: 063103.

Drożdż S, Grümmer F, Ruf F, et al. (2003) Log-periodic self-similarity: An emerging financial law? *Phys A* 324: 174–182.

Drożdż S, Kowalski R, Oświecimka P, et al. (2018) Dynamical variety of shapes in financial multifractality. *Complexity* (Hindawi) 2018: 1–13.

Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50: 987.

Flannery MJ, Protopapadakis AA (2002) Macroeconomic factors do influence aggregate stock returns. *Rev Financ Stud* 15: 751–782.

Fortune P (1991) Stock market efficiency: an autopsy? *New England Econ Rev* 3/4: 17–40.

Geanakoplos J, Magill M, Quinzii M (2004) Demography and the long-run predictability of the stock market. *Brookings Pap Econ Act* 2004: 241–325.

Graham B (1965) *The Intelligent Investor*, Prabhat Prakashan.

Gündüz G (2021) Physical approach to elucidate stability and instability issues, and Elliott waves in financial systems: S&P-500 index as case study. *Quant Financ Econ* 5: 163–197.

Kahneman D (2011) *Thinking, Fast and Slow*, Macmillan.

Lynch P, Rothchild J (2000) *One Up On Wall Street: How To Use What You Already Know To Make Money In The Market*, Simon and Schuster.

Malkiel BG (1999) *A Random Walk Down Wall Street: Including a Life-cycle Guide To Personal Investing*, WW Norton & Company.

Malkiel BG (2010) Bubbles in asset prices, *The Oxford Handbook Of Capitalism*, John Wiley & Sons.

Mandelbrot BB (1997) The variation of certain speculative prices, *In: Fractals And Scaling In Finance*, Springer, New York, NY, 371–418.

Markowitz HM (1959) *Efficient Diversification of Investments–New York*, John Wiley and Sons, New York.

Phillips PCB, Yu J (2011) Dating the timeline of financial bubbles during the subprime crisis. *Quant Econ* 2: 455–491.

Roman HE, Porto M (2008) Fractional derivatives of random walks: Time series with long-time memory. *Phys Rev E* 78: 031127.

Roman HE, Porto M, Dose Ch (2008) Skewness, long-time memory, and non-stationarity: Application to leverage effect in financial time series. *Europhys Lett* 84: 28001.

Roman HE, Albergante M, Colombo M, et al. (2006) Modeling cross correlations within a many-assets market. *Phys Rev E* 73: 036129.

Sharpe WF (1964) Capital asset prices: A theory of market equilibrium under conditions of risk. *J Financ* 19: 425–442.

Shiller RJ (2003) From efficient markets theory to behavioral finance. *J Econ Perspect* 17: 83–104.

Sornette D, Cauwels P, Smilyanov G (2018) Can we use volatility to diagnose financial bubbles? Lessons from 40 historical bubbles. *Quant Financ Econ* 2: 486–594.

Wu TK (2012) Critical phenomena with renormalization group analysis of a hierarchical model of financial crashes. Master of Science Thesis, Simon Fraser University.

Zhang Q, Sornette D, Balcilar M, et al. (2016) LPPLS bubble indicators over two centuries of the S&P 500 index. *Phys A* 458: 126–139.