Using Valuation Ratios in Detecting Stock Market Bubbles, Crises and Exuberance in the S&P Composite Index

Hannu Kahra*

January 2020

Abstract

This paper applies an econometric bubble detection method using a recursive unit root test pioneered by Phillips, Wu, and Yu (2011) and refined by Phillips, Shi, and Yu (2015), commonly referred to the PWY (Phillips-Wu-Yu) method. The method is applied in testing and dating periods of explosive dynamics (exuberance) in Robert Shiller's monthly Standard & Poor's data from January 1881 to December 2019. While previous studies detect bubbles and crises in price-dividend ratios, this paper detects them also in two price-earnings ratios. The procedure applies three valuation ratios: the price-dividend ratio, Shiller's cyclically adjusted price-earnings ratio (CAPE) and Shiller's cyclically adjusted total return price-earnings ratio (TRCAPE) that takes into account the changes in corporate payout policy (i.e. share repurchases rather than dividends). All three valuation ratios exhibit the two well-known stock market bubbles of the 20th century and several crises and exuberance periods in the data. Present high valuation ratios do not imply a bubble – instead they predict many years of poor stock returns.

Keywords: Asset pricing, bubbles, crises, exuberance.

*Estonian Business School and Index Varainhoito, email: hannu.kahra@ebs.ee. The first draft of this paper was prepared for the Finnish Economic Association 42nd Annual Meeting, February 6–7, 2020, in Tampere, Finland. I thank Anustup Kundu for valuable comments.

1 Introduction

On December 3, 1996, John Campbell and Robert Shiller gave their testimony before the board of governors of the Federal Reserve, just before Alan Greenspan's irrational exuberance speech. In the previous month Shiller's cyclically adjusted price-earnings ratio (CAPE) hit 31.4, close to its all-time high 32.6 in September 1929, just before the stock market crash on Black Thursday, October 24, 1929. In their testimony Campbell and Shiller argue that despite all the evidence that stock returns are hard to forecast in the short-run, the simple theory of mean reversion is basically right and does indeed imply a poor long-run stock market outlook. In their paper Campbell and Shiller (1998) continued to assert their pessimistic long-run scenario.

Two of the three 2013 Nobel laureates in Economics, Robert Shiller and Eugene Fama, have very different opinions on how prices are set in financial markets. This difference in opinion extends to the notion of a "bubble". In his Nobel Price Lecture *Speculative Asset Price*s (reprinted in Shiller (2015)) on December 8, 2013, Shiller defines that "a speculative bubble is a peculiar kind of fad or social epidemic that is regularly seen in speculative markets; not a wild orgy of delusions but a natural consequence of the principles of social psychology coupled with imperfect news media and information channels." Shiller offers a definition of bubble that he thinks represents the term's best use:

> A situation in which news of price increases spurs investor enthusiasm which spreads by psychological contagion from person to person, in the process amplifying stories that might justify the price increase and bringing in a larger and larger class of investors, who, despite doubts about the real value of the investment, are drawn to it partly through envy of others' successes and partly through a gambler's excitement.

While Shiller argues that irrational bubbles account for a substantial part of movements in financial markets, and irrational mispricing can extend over several years – Fama claims that prices rationally reflect available information such that markets are "informationally efficient". Fama's great insight was that, because profit-seeking investors quickly incorporate new information into asset prices, the movements of the prices are not predictable in the short-run. He completely rejects the idea that speculative bubbles are the main feature of asset price movements. Sometimes Fama rejects the whole notion of a bubble. While Fama clearly expresses his discontent with the notion of an "irrational bubble", he has never publicly expressed his opinion on "rational bubbles" (Engsted (2016)).

Cooper (2008) (pp. 124–125) argues that in bubble spotting credit growth is the key factor: "... if credit creation is running substantially ahead of economic growth then that growth is likely itself to be supported by by the credit creation, and will not be sustained once the credit expansion ends." Signals of unsustainable credit expansion can be detected indirectly through the behavior of asset prices.

If asset price inflation is unusually high, compared to the income generated by those assets, then the assets may be overvalued. Cooper asks "what will happen to the income levels in the event of a credit contraction." This is particularly important in the stock markets, since credit creation flows so directly into both the earnings and price side of the price-earnings ratio.

Although speculative bubbles and their respective crashes have long been studied, they remain among the most poorly understood and mysterious of economic phenomena. Wöckl (2019) provides a recent survey on theoretical bubble models and an overview of empirical methods for the detection of rational bubbles. Gürkaynak (2008) is an earlier survey on econometric tests of asset price bubbles.

Recent literature of bubble detection has focused on real-time monitoring techniques rather than ex-post identification strategies. Phillips, Wu, and Yu (2011) and Phillips, Shi, and Yu (2015) have proposed a practical real-time bubble detection procedure (PWY hereafter) that has been successfully employed as an early warning alert system for exuberance in a wide variety financial, commodity, and real estate markets. For applications, see Phillips and Shi (2020).

The PWY procedure serves as an early warning device for bubbles and crises. The procedure employs the augmented Dickey-Fuller (ADF) model specification and a recursive evolving algorithm. The algorithm relies on historical information and permits a time-varying model structure.

The authors have applied the method to analyze the sovereign risk in European Union countries (Phillips and Shi (2017)) and Phillips and Shi (2020), the S&P 500 index (Phillips, Shi, and Yu (2015)), (Phillips and Shi (2017)) and (Phillips and Shi (2020)), the NASDAQ index (Phillips, Wu, and Yu (2011)) and financial bubbles during the subprime crisis (Phillips and Yu (2011)).

The potential of the PWY method has been recognized by central bank economists and financial regulators, as well as more widely in the financial industry and financial press (Phillips and Shi (2020)). The procedure is now employed by the Federal Reserve Bank of Dallas providing an exuberance indicator for 23 international housing markets.

The PWY procedure is now a standard item in Matlab, Eviews, and R software programs. R has two packages that apply the PWY procedure, *exuber* (Vasilopoulos et al. (2019)) and *psymonitor* (Caspi, Phillips, and Shi (2018)) that are employed in this paper.

The data are Shiller's monthly S&P Composite Index from January 1881 to December 2019. While previous studies detect bubbles and crises in price-dividend ratios, this paper detects them also in two price-earnings ratios. The bubbles, crises and exuberance periods are identified from the price-dividend ratio, Shiller's cyclically adjusted price-earnings ratio (CAPE) and Shiller's cyclically adjusted total return price-earnings ratio (TRCAPE). The price-dividend ratio is the inverse of dividend yield and the price-earnings ratio is the inverse of earnings yield.

2 Price bubbles

Garber (2001) offers market-fundamental explanations for the three most famous bubbles: the Dutch Tulipmania (1634–1637), the Mississippi Bubble (1719–1720), and the closely connected South Sea Bubble (1720). Galbraith (1993) reviews the major speculative episodes of the last three centuries. Aliber and Kindleberger (2015) provide an updated history of financial bubbles and crises from Tulipmania to the collapse of Lehman Brothers in 2008.

The word "crash" was not commonly attached to stock market movements before 1929. Shiller (2015) reports six S&P Composite Index crashes in the 20th century, notably 1907, 1929, 1973–1974, 1987, 1989 that can be associated with potential bubble or crisis episodes.

What has caused the unusual surge and fall in stock prices, whether there were bubbles and, if so, whether they were rational or behavioral are among the most actively debated issues in macroeconomics and finance in recent years. The occurrence of financial market crashes may be due to significant news events or reported changes in fundamentals, e.g., dividends and earnings. However, many of the dramatic crashes in the U.S. stock market – most notably the 1929 and 1987 crashes – are documented to occur without any particular significant news events or fundamental changes (Shiller (2015)).

Shiller (2019) associates bubbles with narratives. Narratives about stock market bubbles are stories about excitement and risk taking. The Dot-Com bubble was a historic period of excessive speculation mainly in the United States that occurred roughly from 1994 to 2000. The underlying narrative was the "New Economy". The Bitcoin narrative involves stories about inspired cosmopolitan young people, contrasting with uninspired bureaucrats; a story of riches, inequality, advanced information technology, and involving mysterious impenetrable jargon (Shiller (2019), p. 4).

Akerlof and Shiller (2009) refer to price-to-price feedback. This leads to a vicious circle, causing a continuation of the cycle, at least for a while. Eventually an upward price movement, a bubble, must burst, since price is supported only by expectations of further price increases. This cannot continue forever.

Brunnermeier (2008) defines that a bubbles refer to asset prices that exceed an asset's fundamental value because current owners believe they can resell the asset at an even higher price. Aliber and Kindleberger (2015) (pp. 58, 63) document that the term "the greater fool" has been used to suggest the last buyer was always counting on finding someone else to whom the asset could be sold.

Brunnermeier (2008) provides four main strands of bubble models: (1) all investors have rational expectations and identical information, (2) investors are asymmetrically informed and bubbles can emerge because their existence need not be commonly known, (3) rational traders interact with behavioral traders and bubbles persist since limits of arbitrage prevent rational investors from eradicating the price impact of behavioral traders, and (4) investors hold heterogeneous beliefs, potentially due to psychological biases, and agree to disagree about the fundamental value.

Rosser (2000) defines that a speculative bubble exists when the price of something does not equal its market fundamental for some period of time for reasons other than random shocks. When P is price, F is the market fundamental, B is the bubble, and ε is a random process over time, then

$$P_t = F_t + B_t + \varepsilon_t. \tag{1}$$

The most fundamental is determining what is the fundamental. It is usually argued to be a long-run equilibrium consistent with general equilibrium. Equation (1) may represent a temporary equilibrium with demand temporarily not at its long-run position due to the speculative dynamics. The fundamental should reflect an expected value of the long-run equilibrium which is frequently unobservable with any certainty. It is usually assumed that expectation is rational.

For assets with returns it is generally thought that the fundamental should be the unique present discounted value of the expected future returns. One cannot, however. impute definitely from returns in one period what rationally expected returns in later periods will be. This is the key to the difficulty in reality of separating out the fundamental from the bubble, even in a simple case. This is known as the misspecified fundamental problem proposed by Flood and Garber (1980).

Shiller (1981) is a good example. His pioneering paper looked at the relationship between the prices of shares and their intrinsic value – the cashflows that shareholders will eventually receive. Shiller found that prices were much more volatile than their intrinsic value would suggest, something that is hard to square with the idea of efficient markets. In the long-run the valuation of assets are assumed to revert to the mean, and therefore the market movements are eventually predictable.

To motivate the presence of a price bubble Hurn, Martin, Phillips, and Yu (2021, forthcoming) consider the following model

$$P_t(1+R) = E_t[P_{t+1} + D_{t+1}], \tag{2}$$

where P_t is the price of an asset, R is the risk-free rate assumed to be constant, D_t is the dividend and $E_t[\cdot]$ denotes conditional expectation. Equation (2) proposes two types of investment strategies. First, the left-hand side provides a strategy which involves investing in a risk-free asset at time t yielding a risk-free payoff of $P_t(1+R)$ in the next period. Second, the right hand-side demonstrates that by holding the asset the investor expects to earn the capital gain from owning the security with a higher price the next period plus a dividend payment.

In equilibrium there are no arbitrage opportunities, hence the two types of investments are equal to each other. Writing equation (2) as

$$P_t = \beta E_t[P_{t+1} + D_{t+1}], \tag{3}$$

where $\beta = (1+R)^{-1}$ is the discount factor, shows that the price equals the expected discounted payoff. Writing (3) at $t+1$ provides

$$P_{t+1} = \beta E_t[P_{t+2} + D_{t+2}],$$

which is used to substitute out P_{t+1} in (3) yielding

$$P_t = \beta E_t[\beta E_t[P_{t+2} + D_{t+2}] + D_{t+1}] = \beta E_t[D_{t+1}] + \beta^2 E_t[D_{t+2}] + \beta^2 E_t[P_{t+2}].$$

When this approach is repeated N times the principle gives the price of the asset in terms of two components

$$P_t = \sum_{j=1}^{N} \beta^j E_t\left[D_{t+j}\right] + \beta^N E_t[P_{t+N}]. \tag{4}$$

In the model the first on the right-hand side of (4) is the standard present value of the asset whereby the price of the asset equals the discounted present value stream of expected dividends. The second term represents the price bubble

$$B_t = \beta^N E_t[P_{t+N}], \tag{5}$$

since it is an explosive nonstationary process.

Next, consider the conditional expectation of the bubble the next period discounted by β and using the property that $E_t[E_{t+1}[\cdot]] = E_t[\cdot]$ yielding

$$\beta E_t[B_{t+1}] = \beta E_t\left[\beta^N E_{t+1}|P_{t+N+1}\right] = \beta^{N+1} E_t[P_{t+N+1}]. \tag{6}$$

Expression (6) would also correspond to the definition of bubble in (5) if the N forward iterations that produced (4) actually went for $N+1$ iterations. In this case

$$B_t = \beta E_t[B_{t+1}]$$

or, since $\beta = (1+R)^{-1}$

$$E_t[B_{t+1}] = (1+R)B_t,$$

representing a random walk in B_t but with an explosing parameter $1+R$.

3 Empirical tests for rational bubbles

The debate over the existence of bubbles is an inherently empirical question. Wöckl (2019) stresses that the goal of empirical bubble literature is to develop mechanisms to identify the origination, termination, and extent of explosive behavior in asset prices based on explicit quantitative measures. The aim is to empirically separate the contribution of rational bubbles and market fundamentals to exuberance detected in the data. Empirical bubble detection is a difficult task, since the determinants of the fundamental value are generally unobservable, and therefore it is a challenging task to determine an asset's fundamental value.

Gürkaynak (2008) provides a throughout survey on econometric tests for rational bubbles in the context of the present value of dividends model. Wöckl (2019) provides a more recent survey of econometric methods for detecting rational bubbles. She divides the literature on empirical bubble detection into two broad categories: early econometric methods and advanced econometric methods. The first category includes: variance bounds tests (i.e. excess volatility tests), the two-step test of West (1987) and standard stationarity and cointegration based tests. The second category includes: recursive unit root tests, fractional integration tests and regime-swithching tests.

3.1 Variance bounds tests

One method for judging whether there is evidence supporting the basic validity of the efficient markets hypothesis is to see whether the very volatility of stock prices can be justified by the variability of dividends over long interval of time. If the stock price movements are to be justified in terms of the future dividends that firms pay out, then under efficient markets it is not possible to have volatile prices without subsequently volatile dividends.

The present value model is in the core of the Classical Period (1950–1980)[1] of financial economics. Rubinstein (2006) notes that since about 1980, the foundations laid down during the Classical Period have come under increasing strain.

Shiller (1981) rejected the the present value model. He concluded that no movement of the U.S. aggregate stock prices beyond the trend growth of prices has ever been subsequently justified by dividend movements, as the dividend present value with constant discount rate exhibits an extraordinary smooth growth path.

Summers (1984) reports that this conclusion was savagely attacked both orally and in print by several financial economists. He asks, "why has work on volatility testing generated such hostility from mainstream finance researchers." Summers deduces that there are two reasons. First, because users of volatility tests have interpreted their rejections of the null hypothesis in controversial ways. Second, because volatility tests raise some interesting and complex methodological issues. Summers concludes that a large part of the answer must be a deep distrust of research purporting to explore fundamental valuations.

The excess volatility literature starting with Shiller (1981) and LeRoy and Porter (1981) coincides with the regime switch from the Classical Period to Modern Period of financial economics. Shiller (1981) and Grossman and Shiller (1981) developed variance bound tests to empirically explain excess volatility. Gürkaynak (2008) points out, however, that these authors do not explicitly

[1] Rubinstein (2006) divides the history of financial economics into three periods: (1) the Ancient Period before 1950, (2) the Classical Period about 1950 to 1980, and (3) the Modern Period post–1980.

design their tests as bubble detection tests or even refer to bubbles as a potential explanation for excess volatility.

Long-term predictability of asset returns became a major research issue in the 1980s. The seminal excess volatility literature contribution starting with Shiller (1981) points out that "excess volatility" is exactly the same thing as return predictability (Cochrane (2005), p. 396). The conclusion is that when asset prices are "too high" or "too low" necessarily implies that subsequent returns will be too low or too high as prices rebound to their correct levels.

Cochrane (2005) notes that when valuation ratios are high – when prices are high relative to dividends (earnings, cashflows, book values or some other divisor) one of three things must be true: (1) investors expect dividends to rise in the future, (2) investors expect returns to be low in the future, or (3) investors expect prices to rise forever. This conclusion is based on the approximate present value identity provided by Campbell and Shiller (1988). Cochrane (2005) summarizes that price-dividend ratios can only move at all if they forecast future returns, if they forecast future dividend growth, or if there is a bubble – if the price-dividend ratio is nonstationary and is expected to grow explosively.

Cochrane (2011) (p. 1050) explores the present value identity empirically and summarizes that all price-dividend ratio volatility corresponds to variation in expected returns. None corresponds to variation in expected dividend growth, and none to "rational bubbles."

3.2 Recursive unit root tests

Several recent papers place empirical tests for bubbles and rational exuberance applying new development in the field of unit root testing. Instead of concentrating on performing a test of a unit root against the alternative of stationary, essentially using a one-sided test where the critical region is defined in the left-hand side tail of the distribution of the unit root test statistic; the process having an explosive unit root is appropriate for asset prices exhibiting price bubbles. In the latter explosive unit root can be detected from the right tail of the distribution.

The original testing procedures for unit roots were developed by Dickey and Fuller (1979) and Dickey and Fuller (1981) and this framework remains one of the most popular methods to test for nonstationarity in financial time series. The Dickey-Fuller testing framework for unit root testing is for the null hypothesis that a time series y_t is nonstationary or $I(1)$. Another popular test commonly known as the KPSS test, after Kwiatkowski et al. (1992) has a null hypothesis of stationarity or $I(0)$. The PWY method proposed by Phillips, Wu, and Yu (2011) and refined by Phillips, Shi, and Yu (2015) is a recursive method applying the Dickey-Fuller framework in detecting exuberance in asset prices during an inflationary phase.

Table 1 in Appendix B compares the hypotheses in the augmented Dickey-Fuller test (ADF), the KPSS test and the PWY procedure in the simple AR(1) regression equation

$$y_t = \alpha + \rho y_{t-1} + \upsilon_t,$$

where y_t is a valuation ratio.

The null hypotheses in the ADF test and the PSY procedure are the same, the variable is nonstationaty, while the null hypothesis in the KPSS test is stationarity. The alternative hypothesis in the ADF test is stationarity, the alternative hypothesis in the KPSS test is nonstationarity, while in the PWY under the alternative the variable is explosive and there is a price bubble.

The mechanisms developed in Phillips, Shi, and Yu (2015) extend those of Phillips and Yu (2011) by allowing for flexible window widths in the recursive regressions on which the test procedures are based. The approach adopted in Phillips, Wu, and Yu (2011) uses a sup ADF (SADF) to test for the presence of a bubble based on a sequence of forward recursive right-tailed ADF tests. Thereafter Phillips, Wu, and Yu (2011) proposed a dating strategy, which identifies points of origination and termination of a bubble based on a backward regression technique.

As noted in Phillips, Shi, and Yu (2015), when the sample period includes multiple episodes of exuberance and collapse, the PWY procedures may suffer from reduced power and can be inconsistent, failing to reveal the existence of bubbles. To overcome this weakness Phillips, Shi, and Yu (2015) propose a generalized sup ADF (GSADF) method to test for the presence of bubbles as well as a recursive backward regression technique to time-stamp the origination and termination of the bubble. The proposed new methods also rely on recursive right-tailed ADF tests but they use flexible window widths in their implementation. The dating strategy is an ex ante procedure extending the dating strategy proposed by Phillips, Wu, and Yu (2011) by changing the start point in the real-time analysis.

4 Empirical application of the PSY procedure

4.1 Data

Next, the PSY function is applied to three S&P Composited Index valuation ratios. The function is available in the *psymonitor* (Caspi, Phillips, and Shi (2018)) R package. It implements the real time bubble detection procedure of Phillips, Shi, and Yu (2015).

I apply Shiller's monthly data[2] from January 1881 to December 2019. This data set consists of monthly stock price, dividend, and earnings data and the consumer price index. I compute the price-dividend ratio (the inverse of dividend yield) using the available data. The ratio is denoted by PD in the analysis.

The second valuation ratio is Shiller's cyclically adjusted price-earnings ratio (CAPE). It is the real, inflation-corrected, S&P Composite Index divided by the ten-year moving average of real earnings of the index. The price-earnings ratio is a measure of how expensive the market is relative to an objective measure of the ability of corporations to earn profits. The CAPE uses the ten-year average of earnings, along lines proposed by Graham and Dodd (1934).

Both the price-dividend ratio and the price-earnings ratio are biased valuation measures since they do not take into account changes in corporate payout policy. Share repurchases rather than dividends have now become a dominant approach in the United States for cash distribution to shareholders. A stock buyback, also known as a share repurchase, occurs when a company buys back its shares from the marketplace with its accumulated cash. A stock buyback is a way for a company to re-invest in itself. The repurchased shares are absorbed by the company, and the number of outstanding shares on the market is reduced.

Table 2 in Appendix B displays the dividend and buyback yields covering three years, from 2016 to 2018. In these years the buyback yield has exceeded the dividend yield.

Shiller's TRCAPE is an alternative version of CAPE. Changes in corporate payout policy may affect the level of the CAPE ratio through changing the growth rate of earnings per share. This

[2] http://www.econ.yale.edu/~shiller/data.htm.

subsequently may affect the average of the real earnings per share used in the CAPE ratio. A total return CAPE (TRCAPE) corrects for this bias through reinvesting dividends into the price index and appropriately scaling the earnings per share. However, it does not reduce the bias completely.

Finance theory suggests that there are theoretical links between the behavior of prices, dividends and earnings. This can be tested using cointegration tests, e.g. the test proposed by Phillips and Ouliaris (1990). Cointegration implies a dynamic long-run equilibrium or equilibria between the variables. In this case the hypothesis of no cointegration between log prices and log dividends is rejected at $p < 0.01$ significance level. Likewise, the hypothesis of no cointegration between log prices and log earnings is rejected at $p < 0.01$ significance level.

Table 3 displays the results from two popular unit root tests applied to the price-dividend ratio (PD), price-earnings ratio (CAPE) and total return price-earnings ratio (TRCAPE). The tests are the ADF test and the KPSS test. The table also shows the GSADF test statistics of the PWY procedure with simulates critical values for 90, 95 and 95 percent significance levels. The ADF test does not reject the null hypothesis that the valuation ratios are nonstationary, while the KPSS test rejects the null hypothesis that the valuation ratios are stationary. The GSADF test rejects the null hypothesis that the valuation ratios have unit roots, e.g. are nonstationary, supporting the alternative hypothesis that the valuation ratios have explosive dynamics. Since the tests indicate that the valuation ratios are nonstationary, the ratios are not mean reverting.

4.2 Empirical application with R

The PSY procedure was originally designed to identify and date stamp explosive periods in asset prices. Research by Phillips and Shi (2017) shows that the method has detective power against both speculative bubbles and market collapses, including flash crashes. The method is based on a recursive evolving algorithm calculating the ADF statistic from a backward expanding sample sequence. The recursive evolving algorithm enables real-time identification of bubbles and crises while allowing for the presence of multiple structural breaks within the sample period.

Phillips and Shi (2020) illustrate the effectiveness of the PSY procedure with an application to the S&P 500 index. Their monthly data are taken from Datastream International and the sample period runs from January 1973 to July 2018.

They conclude that in the case of a speculative bubble, asset prices typically deviate in an explosive way from fundamentals (dividends), representing exuberance in the speculative behavior driving the market. In this case, this deviation implies that the log price-dividend ratio is expected to follow an explosive process over the expansive phase of the bubble. During crisis periods, the price-dividend ratio is expected to follow a random (downward) drift martingale process. While under normal market conditions the process is a small (local to zero) constant drift martingale process.

Figures 1–3 in Appendix A display the identified bubble and crisis periods using PD, CAPE and TRCAPE as valuation measures. As evident in the figures, the procedure detects two bubble episodes and several crisis episodes in the data. The bubbles are the months preceding the 1929 stock market crash on Black Monday, October 28, 1929, and the Millennium Bubble in the 1990s, followed by a long-run stock market meltdown.

Tables 4–6 documents the beginning and ending months of the bubble and crisis periods in the PD, CAPE and TRCAPE series. CAPE and TRCAPE detect the first crisis period in 1917, from November to December. The U.S. had entered the Great War in April. The start of WW1 lead to stock markets around the world closing down for months, or years. The Dow Jones Industrial Average would fell more than 31 percent by the end of the year.

When the procedure is applied to PD it gives an early warning almost one year before the Black Monday on October 28, 1929, with a few breaks before the crash. When applied to CAPE and TRCAPE the warning signal becomes almost two years before the stock market crash.

The method also detects the bull market of the 1950s in the valuation ratios, associated with "new era" thinking of the 1950s and 1960s (Shiller (2015)). Between September 1953 and December 1955 the market increased 94.3% in real, inflation-corrected, terms. Shiller reports that investors were terribly optimistic and confident of the market and that was itself part of the new era thinking.

The 1973–1974 bear market was a global bear market that was part of the overall U.S. economic recession of the 1970s. The U.S. stock market lost nearly half its value in less than two years, making this particular bear market one of the worst in U.S. history. In the 694 days between 11 January 1973 and 6 December 1974, the New York Stock Exchange's Dow Jones Industrial Average benchmark suffered the seventh-worst bear market in its history, losing over 45% of its value. The market then practically doubled in just over 3 months. The procedure detects this episode in 1974, from September to December.

The procedure provides a warning signal in PD two months before the forthcoming stock market crash of October 19, 1987, while CAPE and TRCAPE provide signals 1.5 years before the crash. The crash is commonly associated with the "portfolio insurance" strategy.

The second boom and bust episode is the well-known Dot-Com bubble from 1996 to May 2001 or September 2000, with several breaks in between.

5 Robustness check

Next, I split the data into four sub-samples and apply the ADF and GSADF tests in the sub-samples using the *radf* function in the *exuber* R package Vasilopoulos et al. (2019). The GSADF test is the same as the one applied to the full sample using the *psymonitor* package of Caspi, Phillips, and Shi (2018). In the application I set the minimum window size to six. This restricts the minimum length of a bubble or a crisis period to be six months and excludes shorter breaks. The null and alternative hypotheses are displayed in Table 1. Tables 7–10 display the ADF and GSADF test statistics and their simulated critical values for 90, 95 and 99 significance level for the GSADF statistic.

In the first sub-sample (1881:01–1915:12, Table 7) ADF cannot reject H_0 (nonstationarity) and GSADF cannot reject H_0 (nonstationarity) in all three valuation ratios. In the second sub-sample (1916:01–1950:12, Table 8) ADF cannot reject H_0 (nonstationarity) and GSADF rejects H_0 (nonstationarity) in all three valuation ratios. In the third sub-sample (1951:01–1985:12, Table 9) ADF cannot reject H_0 (nonstationarity) and GSADF again rejects H_0 (nonstationarity) in all three valuation ratios. Finally, in the fourth sub-sample (1986:01–2019:12, Table 10) ADF cannot reject H_0 (nonstationarity) and GSADF rejects H_0 (nonstationarity) in all three valuation ratios.

The procedure identifies two bubble and one crisis episode reported in Table 11. The bubbles are the Great Crash episode in the 1920s and the Dot-Com bubble in the 1990s. The crisis is the Subprime Mortgage crisis in 2008 and 2009. The full sample analysis was unable to detect this. The third sub-sample has detected bubble/crisis periods, but their lengths are shorter than the minimum length six months and therefore they are not reported. Figures 4–6 display the bubble and crash periods in the last three subsamples.

The procedure was unable to detect the stock market crash of 1907 neither in the full sample nor in the sub-sample, followed by the dramatic slide of stock values between 1907 and 1920. The crash coincided with a banking panic.

6 Conclusion

This paper applies the recursive evolving test algorithm developed by Phillips, Shi, and Yu (2015) that provides a real-time empirical device for detecting speculative bubbles, crises and ballooning credit risks that can foreshadow impending damage to the real economy.

Following Phillips and Shi (2020) the procedure is applied to Shiller's S&P Composite index using a sample period from January 1881 to December 2019. The valuation ratios applied are the price-dividend ratio, as in Phillips and Shi (2020), Shiller's cyclically adjusted price-earnings ratio (CAPE) and Shiller's cyclically adjusted total return price-earnings ratio (TRCAPE) that takes into account stock repurchases.

The procedure is able to detect the two well-known 20th century bubbles in the valuation ratios: the great boom and bust in the 1920s and the Dot-Com bubble in the 1990s. The procedure also detects the "new era" thinking in the 1950s, the stock market crash of 1974, explosive dynamics several months before the stock market crash of October 19, 1987. However, the procedure was unable to detect the subprime mortgage crisis between 2007 and 2010, detected by Phillips, Shi, and Yu (2015) using a sample from January 1871 to December 2010. When analysis is applied to sub-samples, the procedure detects the subprime mortgage crisis period from 2008 to 2009.

Bubbles are associated with high valuation ratios, but a high valuation ratio does not necessarily imply the existence of a bubble. Presently, valuation ratios are extremely high, partly due to quantitative easing (QE) that has inject liquidity into the economy. By providing liquidity in the banking sector, QE has made it easier and cheaper for banks to extend loans to companies and households, thus stimulating credit growth.

Cochrane (2005) concludes that the expected or required return – the risk premium – on individual securities as well as the market as a whole varies slowly over time. It is possible to track market expectations of returns by watching valuation ratios, e.g., price-dividend, price-earnings, or book-market ratios. High prices, relative to dividends and earnings, have reliably preceded many years of poor stock returns. That is the present outlook.

Appendix A Figures

Figure 1: Bubble and crisis periods in the S&P stock market index using the S&P price-dividend ratio, January 1881–December 2019. The solid line is the prices-to-dividends ratio and the shaded areas are the periods where the PSY statistic exceeds its 95% bootstrapped critical value.

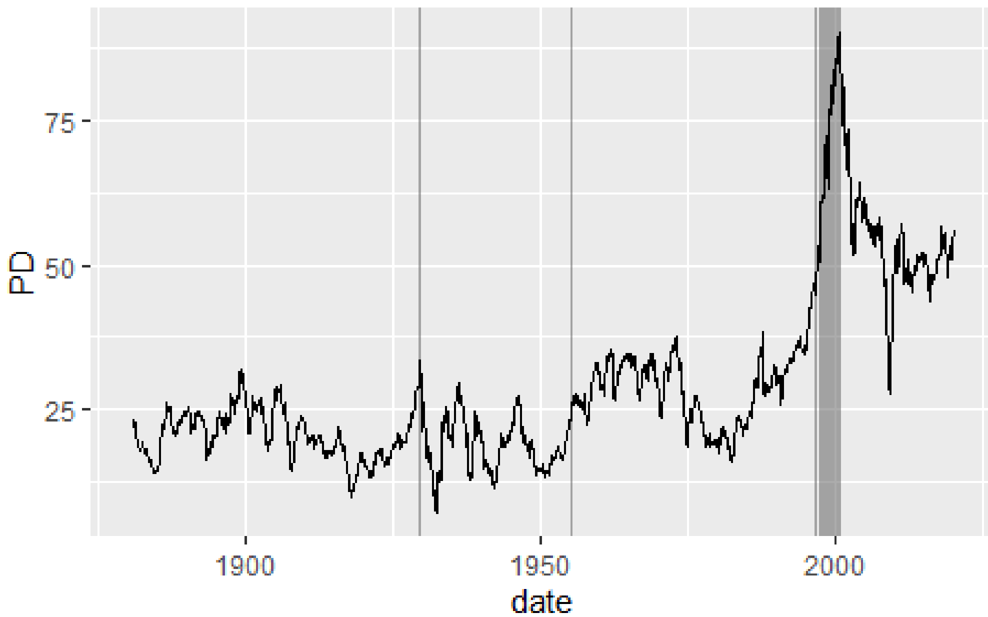


Figure 2: Bubble and crisis periods in the S&P stock market index using the Shiller's cyclically adjusted price-earnings ratio (CAPE), January 1881–December 2019. The solid line is the prices-to-earnings ratio and the shaded areas are the periods where the PSY statistic exceeds its 95% bootstrapped critical value.

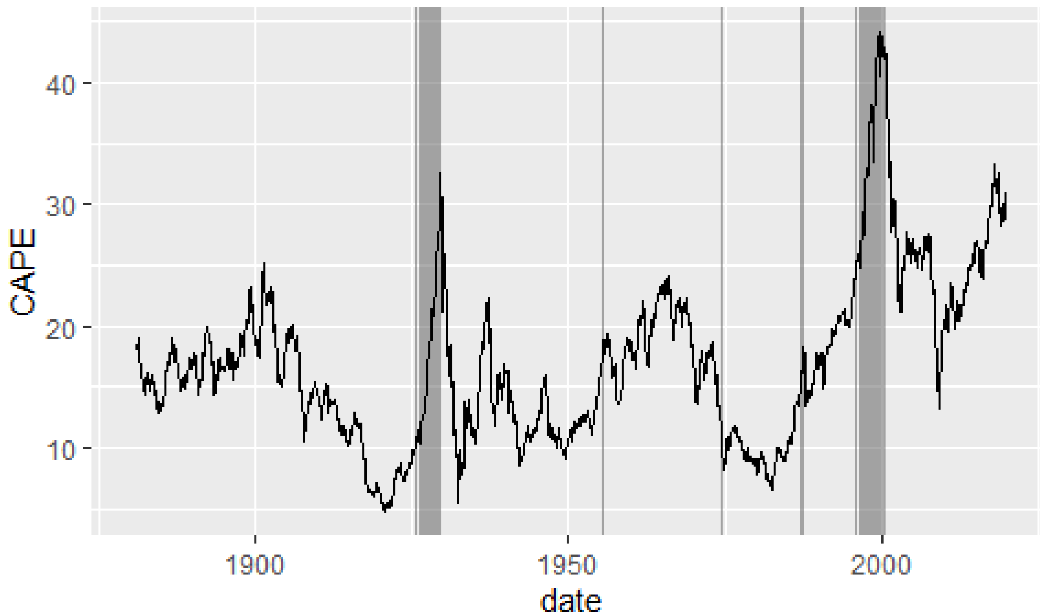


Figure 3: Bubble and crisis periods in the S&P stock market index using Shiller's total return cyclically adjusted price-earnings ratio (TRCAPE), January 1881–December 2019. The solid line is the prices-to-earnings ratio and the shaded areas are the periods where the PSY statistic exceeds its 95% bootstrapped critical value.

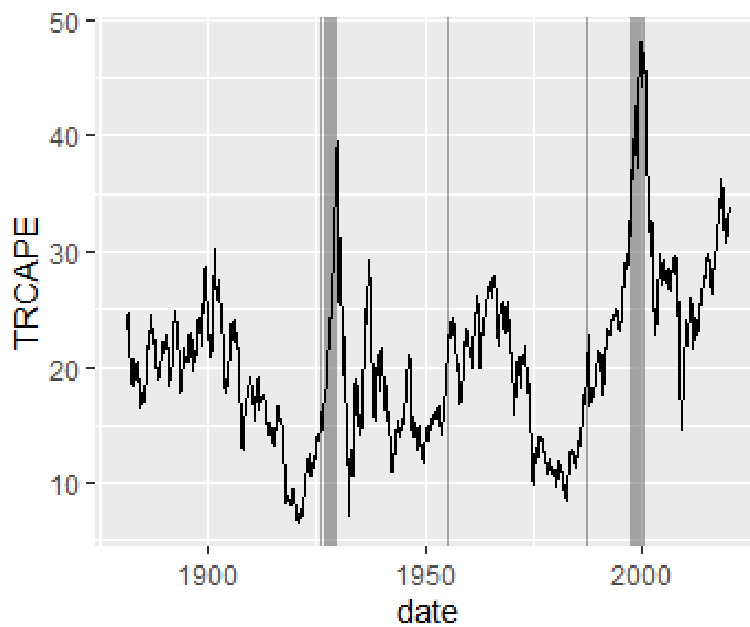


Figure 4: Bubble and crisis periods in the S&P index 1916:01–1950:12

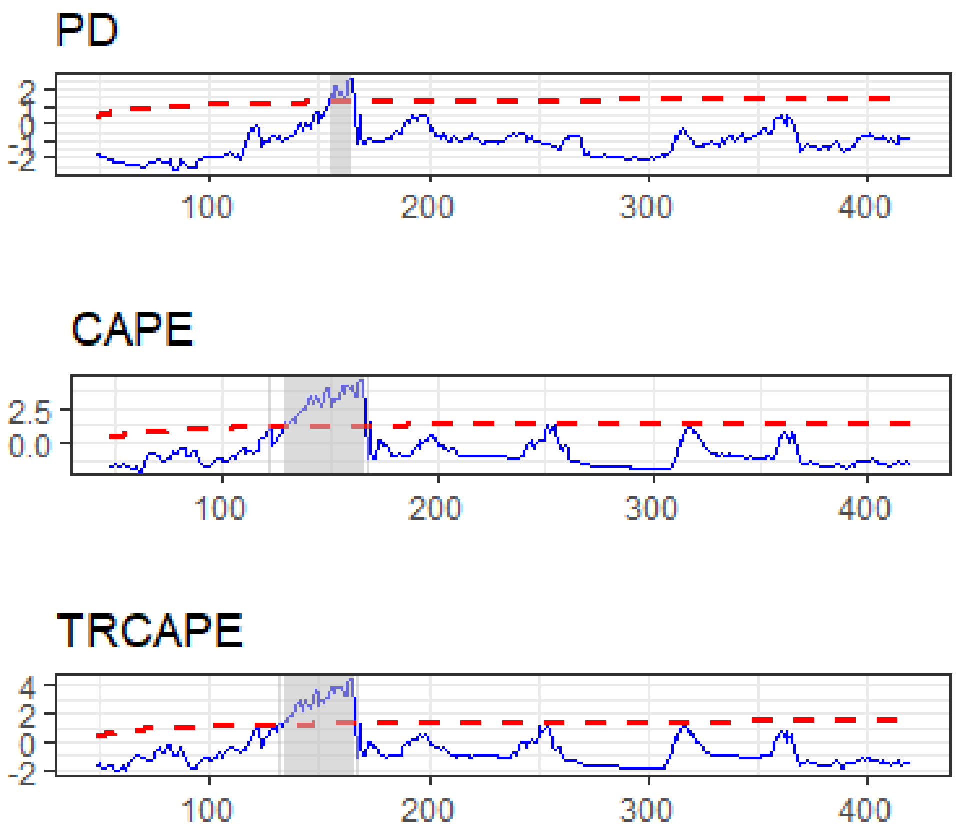


Figure 5: Bubble and crisis periods in the S&P index 1951:01–1985:12

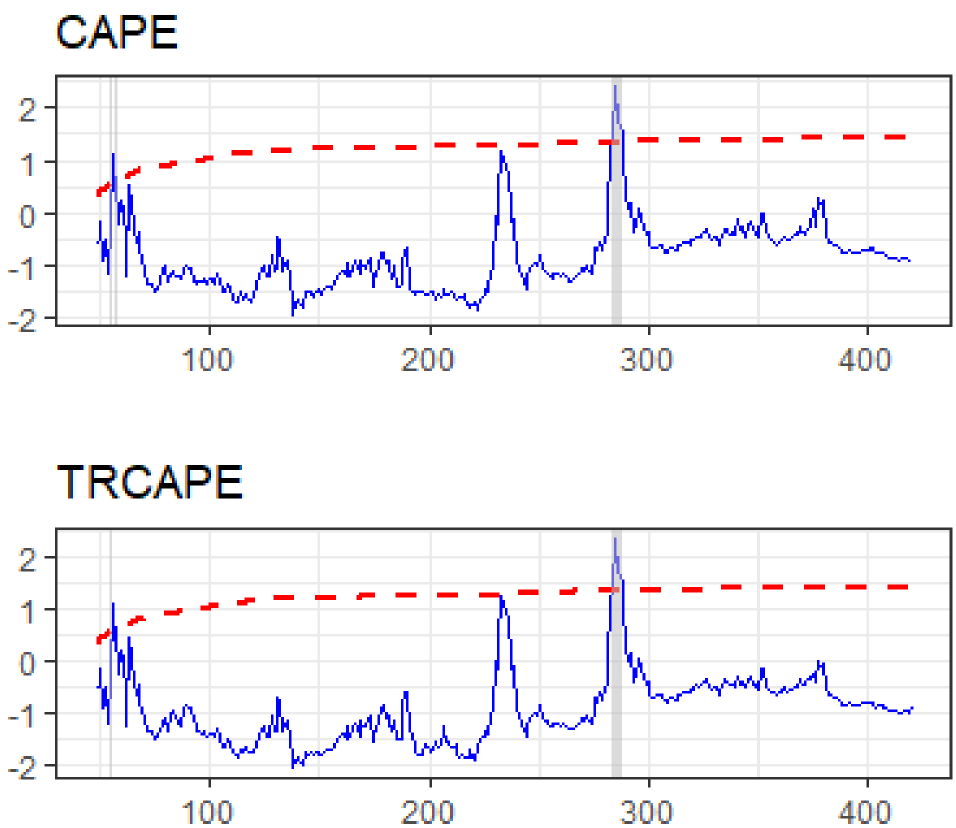
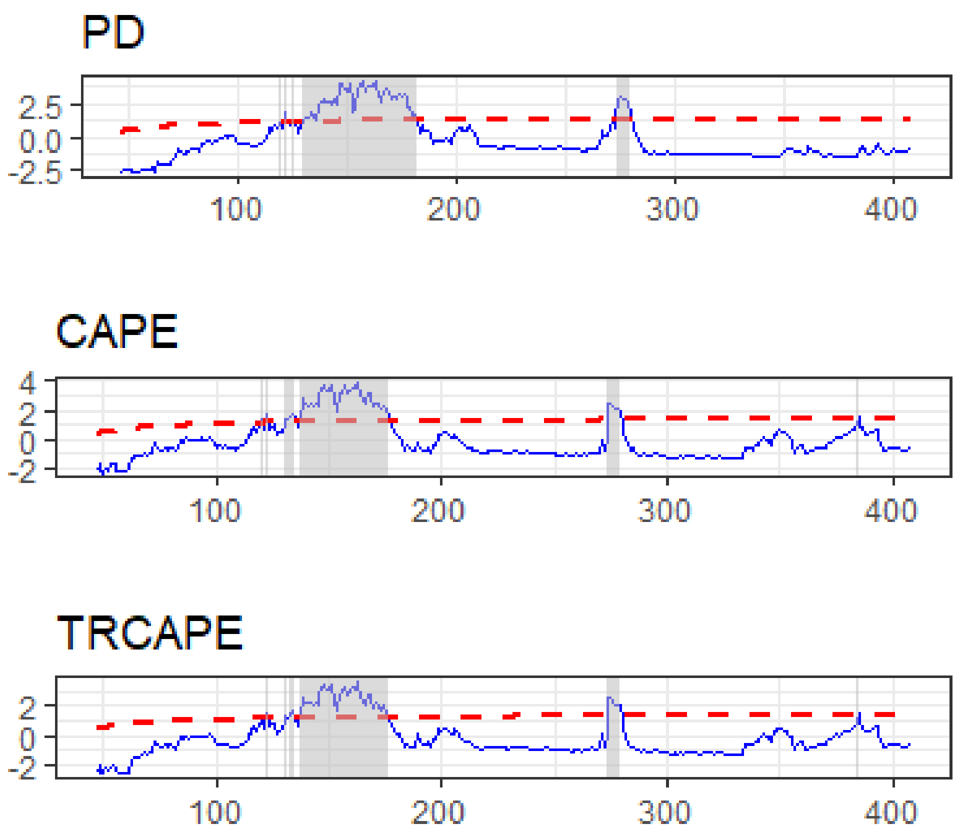


Figure 6: Bubble and crisis periods in the S&P index 1986:01–2019:12



Appendix B Tables

Table 1: The ADF, KPSS and PWY hypotheses

	ADF	KPSS	PWY
H_0	$\rho = 1$	$\rho < 1$	$\rho = 1$
H_1	$\rho < 1$	$\rho = 1$	$\rho > 1$

Table 2: Dividend, buyback and combined S&P 500 yields

	Yields		
Year	Dividend	Buyback	Combined
2016	2.06%	2.78%	4.85%
2017	1.84%	2.28%	4.12%
2018	2.17$	3.84%	6.01%

Source: `https://www.prnewswire.com/news-releases/sp-500-buybacks-reverse-declines-in-q3-2` `html.`

Table 3: Unit root test statistics, p-values and simulated critical values for 90, 95 and 99 percent significance level

	ADF	p-value	KPSS	p-value	GSADF	90%	95%	99%
PD	-2.97	0.17	9.64	< 0.01	4.25	0.70	1.25	2.05
CAPE	-3.02	0.15	4.44	< 0.01	4.60	0.80	1.07	1.78
TRCAPE	-3.35	0.06	3.09	< 0.01	4.15	0.77	1.10	1.74

Table 4: Bubble and crisis periods in the S&P index using the price-dividend ratio

Start	End
1928:11	1928:11
1929:01	1929:03
1929:07	1929:10
1955:06	1955:09
1955:11	1955:12
1987:08	1987:09
1996:02	1996:02
1996:11	1997:03
1997:05	2001:02
2001:05	2001:05

Table 5: Bubble and crisis periods in the S&P index using Shiller's cyclically adjusted price-earnings ratio (CAPE)

Start	End
1917:11	1917:12
1925:11	1926:02
1926:07	1929:10
1929:12	1929:12
1955:02	1955:02
1955:04	1955:04
1955:06	1955:12
1956:03	1956:04
1974:09	1974:12
1986:04	1986:06
1986:08	1986:08
1987:01	1987:09
1996:03	1996:08
1996:09	1998:08
1998:10	2000:09

Table 6: Bubble and crisis periods in the S&P index using Shiller's totel return cyclically adjusted price-earnings ratio (TRCAPE)

Start	End
1917:11	1917:12
1925:12	1926:02
1926:08	1926:09
1926:11	1929:10
1929:12	1929:12
1955:02	1955:02
1955:04	1955:04
1955:06	1955:09
1955:11	1955:12
1956:03	1956:04
1974:09	1974:10
1974:12	1974:12
1986:04	1986:06
1986:08	1986:08
1987:01	1987:09
1997:02	1997:02
1997:05	2000:09

Table 7: ADF and GSADF tests applied to sub-sample 1881:01–1915:12 (N = 420)

	t-stat	90%	95%	99%	
			PD		
ADF	-3.431				Cannot reject H_0, $p = 0.05$
GSADF	1.165	1.925	2.118	2.641	Cannot reject H_0
			CAPE		
ADF	-2.930				Cannot reject H_0, $p = 0.18$
GSADF	1.760	1.925	2.118	2.641	Cannot reject H_0
			TRCAPE		
ADF	-3.318				Cannot reject H_0, $p = 0.07$
GSADF	1.453	1.925	2.118	2.641	Cannot reject H_0

Table 8: ADF and GSADF tests applied to sub-sample 1916:01–1950:12 (N = 420)

	t-stat	90%	95%	99%	
			PD		
ADF	-3.346				Cannot reject H_0, $p = 0.06$
GSADF	2.663	1.925	2.118	2.641	Rejects H_0 for significance level 99%
			CAPE		
ADF	-2.339				Cannot reject H_0, $p = 0.42$
GSADF	4.603	1.925	2.118	2.641	Rejects H_0 for significance level 99%
			TRCAPE		
ADF	-2.382				Cannot reject H_0,$p = 0.41$
GSADF	4.372	1.925	2.118	2.641	Rejects H_0 for significance level 99%

Table 9: ADF and GSADF tests applied to sub-sample 1951:01–1985:12 (N = 420)

	t-stat	90%	95%	99%	
			PD		
ADF	-2.731				Cannot reject H_0, $p = 0.21$
GSADF	2.101	1.925	2.118	2.641	Rejects H_0 for significance level 90%
			CAPE		
ADF	-1.565				Cannot reject H_0, $p = 0.44$
GSADF	2.385	1.925	2.118	2.641	Rejects H_0 for significance level 95%
			TRCAPE		
ADF	-1.641				Cannot reject H_0, $p = 0.37$
GSADF	2.3413	1.925	2.1175	2.6413	Rejects H_0 for significance level 95%

Table 10: ADF and GSADF tests applied to sub-sample 1986:01–2019:12 ($N = 408$)

	t-stat	90%	95%	99%	
				PD	
ADF	-1.784				Cannot reject H_0, $p = 0.71$
GSADF	4.254	1.928	2.115	2.610	Rejects H_0 for significance level 99%
				CAPE	
ADF	-1.850				Cannot reject H_0, $p = 0.65$
GSADF	3.989	1.928	2.115	2.610	Rejects H_0 for significance level 99%
				TRCAPE	
ADF	-1.860				Cannot reject H_0,$p = 0.64$
GSADF	3.599	1.9281	2.115	2.610	Rejects H_0 for significance level 99%

Table 11: Bubble and crisis periods in the S&P Composite index 1881:01–2019:12

Start	End	Length	Explanation
		PD	
1928:11	1929:09	11	Great crash episode
1996:10	2001:02	53	Dot-Com bubble
2008:10	2009:04	7	Subprime mortgage crisis
		CAPE	
1926:08	1929:10	39	Great crash episode
1997:05	2000:09	41	Dot-Com bubble
2008:10	2009:03	6	Subprime mortgage crisis
		TRCAPE	
1927:02	1929:10	33	Great crash episode
1997:05	2000:09	41	Dot-Com bubble
2008:10	2009:03	6	Subprime mortgage crisis

References

Akerlof, George A. and Shiller, Robert J. (2009). *Animal Spirits: How Human Psychology Drives the Economy and Why it Matters for Global Capitalism*. Princeton University Press.

Aliber, Robert Z. and Kindleberger, Charles P. (2015). *Manias, Panics, and Crashes: A History of Financial Crises*. 7th Edition. Palgrave Macmillan.

Brunnermeier, Markus K. (2008). “The New Palgrave Dictionary of Economics”. In: ed. by Steven N. Durlauf and Lawrence E. Blume. Vol. 1–8. Palgrave Macmillan. Chap. Bubbles, pp. 578–583.

Campbell, John Y. and Shiller, Robert J. (1988). “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors”. In: *The Review of Financial Studies* 1.3. Papers and Proceedings of the Forty-Third Annual Meeting American Finance Association, Dallas, Texas, December 28-30, 1984, pp. 195–228.

— (1998). “Valuation Ratios and the Long-Run Stock Market Outlook”. In: *Journal of Portfolio Management* 24, pp. 11–26.

Caspi, Itamar, Phillips, Peter C. B., and Shi, Shuping (2018). *psymonitor: Real Time Monitoring of Asset Markets*. R package version 0.0.1. URL: `https://itamarcaspi.github.io/psymonitor/`.

Cochrane, John Y. (2005). *Asset Pricing*. Revised edition. Princeton University Press.

— (2011). “Presidential Address: Discount Rates”. In: *The Journal of Finance* 66.4, pp. 1047–1108.

Cooper, George (2008). *The Origins of Financial Crises: Central Banks, Credit Bubbles and the Efficient Market Fallacy*. Harriman House Ltd.

Dickey, David A. and Fuller, Wayne A. (1979). “Distribution of the Estimators for Autoregressive Time Series with a Unit Root,” in: *Journal of the American Statistical Association,* 74.366a, pp. 427–431.

— (1981). “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root”. In: *Econometrica* 49.4, pp. 1057–1072.

Engsted, Tom (2016). “Fama on Bubbles”. In: *Journal of Economic Surveys* 30.2, pp. 370–376.

Flood, Robert and Garber, Peter (1980). “Market Fundamentals versus Price-Level Bubbles: The First Tests”. In: *Journal of Political Economy* 88.4, pp. 745–70.

Galbraith, John Kenneth (1993). *A Short History of Financial Euphoria*. Penguin.

Garber, Peter M. (2001). *Famous First Bubbles: The Fundamentals of Early Manias*. The MIT Press.

Graham, Benjamin and Dodd, David (1934). *Securities Analysis*. McGraw-Hill.

Grossman, Sanford J. and Shiller, Robert J. (1981). “The Determinants of the Variability of Stock Market Prices”. In: *The American Economic Review* 71.2. Papers and Proceedings of the Ninety-Third Annual Meeting of the American Economic Association (May, 1981), pp. 222–227.

Gürkaynak, Refet S. (2008). “Econometric Tests of Asset Price Tests: Taking Stock”. In: *Journal of Economic Surveys* 22.1, pp. 166–186.

Hurn, Stan et al. (2021, forthcoming). *Financial Econometric Modelling*. Oxford University Press.

Kwiatkowski, Denis et al. (1992). “Testing the null hypothesis of stationarity against the alternative of a unit root”. In: *Journal of Econometrics* 54, pp. 159–178.

LeRoy, Stephen F. and Porter, Richard D. (1981). “The Present-Value Relation: Tests Based on Implied Variance Bounds”. In: *Econometrica* 49.3, pp. 555–574.

Phillips, Peter C. B. and Ouliaris, Sam (1990). “Asymptotic Properties of Residual Based Tests for Cointegration”. In: *Econometrica* 58.1, pp. 65–193.

Phillips, Peter C. B. and Shi, Shupin (2017). *Detecting Financial Collapse and Ballooning Sovereign Risk*. Cowles Foundation Discussion Paper No. 3010.

Phillips, Peter C. B. and Shi, Shuping (2020). “Financial, Macro and Micro Econometrics Using R”. In: ed. by Hrishikesh Vinod and R. C. Rao. 1st edition. Vol. 42. Handbook of Statistics. Also available: Cowles Foundation Discussion Paper No. 2152, 2018. North Holland. Chap. Real time monitoring of asset markets: Bubbles and crises.

Phillips, Peter C. B., Shi, Shuping, and Yu, Jun (2015). “Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapses in the S&P 500”. In: *International Economic Review* 56.4, pp. 1034–1078.

Phillips, Peter C. B., Wu, Yangru, and Yu, Jun (2011). “Explosive Behavior in the 1990s Nasdaq: When Did Exuberance Escaoete Asset Values?” In: *International Economic Review* 52.1, pp. 201–226.

Phillips, Peter C. B. and Yu, Jun (2011). “Dating the timeline of financial bubbles during the subprime crisis”. In: *Quantitative Economics* 2.3, pp. 455–491.

Rosser, J. B. (2000). “Catastrophe to Chaos: A General Theory of Economic Discontinuities”. In: Springer. Chap. Speculative Bubbles and Crashes I: Irrational, pp. 57–71.

Rubinstein, Mark (2006). *A History of the Theory of Investments: My Annotated Bibliography*. Wiley.

Shiller, Robert J. (1981). “Do Stock Prices Move Too Much to Be Justified by Subsequent Movements in Dividends?” In: *American Economic Review* 71.3, pp. 421–436.

— (2015). *Irrational Exuberance*. Revised and expanded 3rd edition. Princeton University Press.

— (2019). *Narrative Economics: How Stories Go Viral & Drive Major Economic Events*. Princeton University Press.

Summers, Lawrence H. (1984). “On Economics and Finance”. In: *The Journal of Finance* 40.3, pp. 633–635.

Vasilopoulos, Kostas et al. (2019). *exuber: Testing and Simulating Explosive Periods*. R package version 0.3.0. URL: `https://CRAN.R-project.org/package=exuber`.

West, Kenneth D. (1987). “A Specification Test for Speculative Bubbles”. In: *The Quarterly Journal of Economics* 102.3, pp. 553–580.

Wöckl, Ines (2019). *Bubble Detection in Financial Markets - A Survey of Theoretical Bubble Models and Empirical Bubble Detection Tests*. Available at SSRN: https://ssrn.com/abstract=3460430 or http://dx.doi.org/10.2139/ssrn.3460430.