Testing for Speculative Bubbles using Spot and Forward Prices*

Efthymios G. Pavlidis,[†] Ivan Paya, and David A. Peel

Department of Economics, Lancaster University Management School, UK

Abstract

The probabilistic structure of periodically collapsing bubbles creates a gap between future spot and forward (or futures) asset prices in small samples. By exploiting this fact, we use two econometric methods, namely, the recursive unit root method of Phillips, Shi, and Yu (2015a,b) and the rolling regression method of Fama (1984), for detecting bubbles. Both methods do not rely on a particular model of asset price determination, they are robust to an explosive root in the process for market fundamentals, and are accompanied by a date-stamping strategy. By applying these methods to the German mark-US dollar and British pound-US dollar exchange rates, we provide evidence in favor of speculative bubbles in the foreign exchange market during the interwar German hyperinflation, but not during the recent floating-rate period. A further application to the S&P 500 index supports the existence of speculative bubbles in the US equity market.

Keywords: Rational bubbles, Forward exchange rates, Futures prices, Explosive root, Hyperinflation.
JEL Classification: C12, C22, F31.

1 Introduction

A rational bubble reflects a self-confirming belief that an asset's price depends on a variable (or a combination of variables) that is intrinsically irrelevant - that is, not part of market fundamentals (Diba and Grossman, 1988, p. 520). Since the widespread adoption of rational expectations into macroeconomics, a large number of studies have attempted to empirically separate the contribution of rational bubbles and market fundamentals to asset price movements (e.g., Balke and Wohar, 2009; Phillips et al., 2011).[1] A variety of econometric tests have been developed in the literature that can be used for this purpose. The most popular include volatility, Hausman-type, and (co-)integration tests (Cochrane, 1992; Diba and Grossman, 1988; Engsted and Nielsen, 2012; LeRoy and Porter, 1981; Phillips et al., 2015a; Taylor and Peel, 1998; West, 1987, 1988).

A common feature of all tests for bubbles is that they actually examine a composite hypothesis of no bubbles and a model for market fundamentals. Because rejection of the null may be due to the presence of bubbles or model misspecification, or to both of these reasons, results are deemed to be inconclusive (Flood et al., 1994; Gürkaynak, 2008; Hamilton and Whiteman, 1985). To make matters worse, even if the specification of the true model for market fundamentals is known, econometric tests may still fail to detect bubbles if these processes periodically crash (Blanchard, 1979; Blanchard and Watson, 1983; Evans, 1991; van Norden, 1996; van Norden and Schaller, 1993). Following Evans (1991), several studies have shown that, in finite samples, periodically collapsing bubbles can make asset prices look non-explosive and co-integrated with market fundamentals, thus leading to a severe power loss of standard statistical tests.

In this paper, we propose two novel methods for detecting bubbles that periodically collapse. The methods do not rely on a particular asset price determination model, impose minimal assumptions on the process for market fundamentals, and, thus, alleviate the misspecification problem. The methods have two further appealing features. First, because they are recursive they allow the identification of the starting and ending points of bubble episodes. Second, they only require data on spot and forward (or futures) prices, which contrary to market fundamentals, are readily available at high frequencies and over long time periods. This latter feature is especially important in light of the low power of statistical tests (Evans, 1991; West, 1994).

Although we focus on the foreign exchange market, the methods are of general interest since they are applicable to any market for which spot and forward (or futures) prices are available. The underlying idea is simple: when a bubble is occurring, both the forward exchange rate and the future spot rate incorporate it but with different weights. This is due to the fact that rational agents, when forming expectations, correctly attach a

---

[1]In line with most of the empirical literature, we focus on rational extrinsic bubbles. Froot and Obstfeld (1992) and Driffill and Sola (1998) examine intrinsic bubbles, i.e., bubbles that depend on market fundamentals. There are also studies that relax the rational expectations assumption and incorporate behavioral aspects into the analysis, such as overconfidence, and differences in priors and beliefs (Abreu and Brunnermeier, 2003; Brunnermeier, 2008; Lansing, 2010; Scheinkman and Xiong, 2003). For a survey on speculative bubbles see LeRoy (2004).

nonzero probability to the bubble bursting. As a consequence, the forward rate becomes a downward biased predictor of the future spot rate, with the difference between the two rates (i.e., the degree of the bias) depending on the bubble process and, as such, being explosive. As long as the forecast errors for fundamentals are not explosive, the presence of explosive dynamics in the difference between the two rates can be attributed to the existence of bubbles.

On this basis, rather than applying (co-)integration tests to the spot rate and a set of observed market fundamentals, which is the common practice in the literature, we advocate the application of right-tailed unit root tests to the difference between future spot and forward rates. Of the various unit root tests proposed in the literature, we employ the Generalized Supremum Augmented Dickey Fuller (GSADF) recently developed by Phillips et al. (2015a). This choice is based on Monte Carlo simulation results that indicate that the GSADF exhibits good power properties in our setting.

The second method we propose is based on the failure of the unbiasedness hypothesis in small samples and consists of estimating rolling Fama regressions. The fact that the variables that enter the Fama regression are highly persistent (potentially explosive) implies that conventional tests may be invalid (Cavanagh et al., 1995; Engsted and Nielsen, 2012; Jansson and Moreira, 2006; Nielsen, 2010; Phillips and Magdalinos, 2008). To circumvent this obstacle, we adopt the IVX instrumentation procedure of Phillips and Magdalinos (2009), Phillips and Lee (2013), and Kostakis et al. (2015).

We provide two empirical applications of the proposed bubble detection methods to US dollar exchange rates. In the first application, we examine one of the most studied events in monetary history: the interwar German hyperinflation. The existence of speculative bubbles in hyperinflation Germany has been the subject of a considerable amount of research (see, e.g., Burmeister and Wall, 1982, 1987; Casella, 1989; Christiano, 1987; Durlauf and Hooker, 1994; Engsted, 1993, 1994; Hooker, 2000; Sargent, 1977; Taylor, 1991). However, no general consensus has been reached in the literature. Previous studies, that focus on prices and money, find mixed results that depend crucially on the validity of a variety of different assumptions. For instance, Burmeister and Wall (1982) find that bubbles cannot be rejected if the money supply process depends on expected inflation. Casella (1989) shows that the presence of bubbles in prices can be rejected if the money supply process is assumed exogenous to current inflation but not when there is a feedback rule. Hooker (2000) shows through Monte Carlo simulations that the small sample properties of tests for bubbles may differ substantially from their asymptotic counterparts. By using finite-sample critical values, he concludes that there is evidence of misspecification in the Cagan money demand model but no evidence of bubbles. Engsted (2003) points out that Hooker's methodology neglects the magnitude of misspecification noise in Cagan's model. Taking the magnitude of noise into account, he also doesn't find any evidence of bubbles.

Unlike existing studies on the German hyperinflation, we examine the presence of speculative bubbles in the foreign exchange market. The lack of research in this area is surprising given the wide use of the German mark as a vehicle for speculation at that

time. Keynes, for instance, notes that foreign exchange speculation during the German hyperinflation was "on a tremendous scale and was, in fact, the greatest ever known".[2] Our results reveal two bubble episodes, one in July-August 1922, and another a year later in the summer of 1923, almost a month before forward trading was suspended by the German authorities. These two bubble periods coincide with large and persistent speculative movements in the foreign exchange market (Bresciani-Turroni, 1937).

Our second empirical application focuses on the recent floating-rate period. This period is particularly interesting because, following the collapse of the Bretton Woods system, exchange rates among major currencies became volatile and highly persistent, moving in one direction over long time spans (see, e.g., Engel and Hamilton, 1990). A notable example of such movements is the large appreciation of the US dollar against the British pound in the early 1980s. The magnitude of this appreciation motivated several researchers to investigate whether exchange rate changes only reflect changes in fundamentals, or speculative bubbles are also present in the foreign exchange market (De Grauwe and Rovira-Kaltwasser, 2012; Sarno and Taylor, 2002). For instance, Evans (1986) develops a nonparametric bubble-detection test, where the null hypothesis is that the median of excess foreign exchange returns is zero. By applying the test to data on the British pound-US dollar rate for the period 1981 to 1984, he rejects the null hypothesis, and concludes that speculative bubbles were present over the sample period examined. In another study, Meese (1986) employs an econometric methodology based on Hausman's specification test. The joint null hypothesis of the test is that there are no bubbles and that the exchange rate is determined in a hybrid monetary model. By using data for the period 1973 to 1982, Meese (1986) also rejects the null hypothesis. Wu (1995), on the other hand, uses the Kalman filter to decompose the British pound-US dollar exchange rate into two parts: a part that depends on monetary fundamentals and another that depends on the unobserved bubble process. Contrary to the previous studies, Wu (1995) finds that the unobserved bubble component is not statistically different from zero throughout the period 1974 to 1988. In line with the findings of the latter study, our empirical analysis of the British pound-US dollar spot and forward rates provides no evidence in favor of speculative bubbles during the recent float.

To demonstrate the applicability of our methods to markets other than the foreign exchange, as a final application, we investigate the behavior of the Standard & Poor's 500 (S&P 500) index. The long history of the S&P 500 is characterized by several episodes of large price run-ups as well as severe market crashes, such as the boom and bust of the 1920s, *Black Monday* in October 1987, and the *millennium* boom. Due to its rich history, the index is one of the most commonly examined series for speculative bubbles. Early studies related to this topic show that the variance of the actual price index exceeds by a substantial margin the bound imposed by the variance of the estimated ex-post rational price (see, e.g., LeRoy and Porter, 1981; Shiller, 1981). As noted by Blanchard and Watson (1983) and Tirole (1985), a potential cause of such excess asset price volatility is the presence of speculative bubbles. Cochrane (1992), however, using a decomposition of

---

[2]This quote is taken from Ferguson (1995).

the variance of the S&P 500 price-to-dividend ratio, provides evidence against this hypothesis. West (1987) also examines the S&P 500 index. Contrary to Cochrane (1992), he rejects the null of no bubbles by employing a novel Hausman-type misspecification test. Following Diba and Grossman (1988), a large literature has developed that explores the presence of explosive bubbles in the US equity market by adopting a (co-)integration framework. The results of this literature are mixed. In general, standard (co-)integration tests applied to long spans of data fail to provide evidence in favor of bubbles (Diba and Grossman, 1988; Taylor and Peel, 1998). On the other hand, recursive unit root tests, and tests that allow for either discrete changes in the integration properties of the time series under investigation, or for markov switching regimes produce much more supportive results (Homm and Breitung, 2012; Phillips et al., 2015a; van Norden and Vigfusson, 1998). Our findings suggest that using information from the futures market to mitigate the potential misspecification problem produces somewhat more conservative results than using observed market fundamentals. Nonetheless, the overall conclusion remains that episodes of exuberance exist.

The rest of the paper is structured as follows. Section 2 describes the proposed bubble detection methods based on recursive unit root tests and rolling Fama regressions. The following section investigates the performance of the proposed methods in detecting and dating bubble episodes by conducting a battery of Monte Carlo simulations. Section 4 presents the empirical applications. The final section concludes.

2   Rational Bubbles in the Foreign Exchange Market

We focus on the foreign exchange market, and consider rational expectations models that relate the log exchange rate s_t to economic fundamentals v_t, and to a bubble process B_t (Engel and West, 2005; Sarno and Taylor, 2002, Ch. 2)[3]

$$s_t = v_t + B_t. \tag{1}$$

There are many models of the exchange rate that take this general form (Engel et al., 2007). The most popular is the simple monetary model in which the fundamental value of the exchange rate v_t can be expressed as the discounted sum of current and expected future relative money supplies and relative incomes (Mark, 1995).

For our analysis, a precise definition of economic fundamentals is not required. We assume for simplicity that v_t follows an autoregressive (AR) process of order one

$$v_t = \phi v_{t-1} + \theta_t, \tag{2}$$

where $\theta_t \sim iid(0, \sigma_\theta^2)$. We also assume that the bubble term has the simplest form of a rational speculative bubble that pops, the type proposed by Blanchard (1979). (As it will become evident below, the main implications of the analysis go through for a much broader set of processes for fundamentals and more complicated forms of the bubble

---

[3]This expression assumes absence of the transversality condition.

term.) Blanchard's bubble process has two regimes, which occur with probabilities π and $1 - \pi$. In the first regime, the bubble grows exponentially at the rate $(1 + r)/\pi$, while, in the second regime, the bubble collapses to a white noise

$$B_{t+1} = \begin{cases} \frac{(1+r)}{\pi} B_t + \epsilon_{t+1}, & \text{with prob. } \pi \\ \epsilon_{t+1}, & \text{with prob. } 1 - \pi, \end{cases} \tag{3}$$

where r is a positive constant derived from the structural model determining the exchange rate, and $\epsilon_t \sim iid(0, \sigma_\epsilon^2)$. Taking expectations of both sides yields

$$E_t[B_{t+1}] = (1 + r)B_t, \tag{4}$$

where $E_t[\cdot]$ is the expectation operator. The expected growth rate of the bubble is $1 + r$ and, therefore, Equation (3) satisfies the condition for a rational bubble (see Diba and Grossman, 1988).

2.1 Bubble Detection I: Recursive Unit Root Tests

Conditioning on the first regime, the bubble term follows an explosive AR(1) process. It follows from Equation (1) that the spot rate, the sum of the bubble process and fundamentals, also has an explosive root. On this basis, it is typical in the literature to test for bubbles by running right-tailed unit root tests on asset prices. Although widely used, unit root tests have an important drawback; they are inconclusive. This is due to the fact that asset prices may be explosive even in the absence of bubbles as long as their other component, fundamentals, exhibits explosive dynamics, i.e. $\phi > 1$. Since explosive fundamentals cannot be ruled out *a priori*, especially during periods of hyperinflation, unit root tests may lead to false inference.

In order to address this issue, researchers have incorporated fundamentals into the analysis by, among other methods, running unit root tests on ratios of asset prices-to-fundamentals (e.g., stock price-to-dividend and house price-to-income ratios), testing for co-integration between asset prices and fundamentals, and comparing the variance of asset prices to that implied by fundamentals (Craine, 1993; Diba and Grossman, 1988; Phillips et al., 2014; West, 1987). A common factor in all these methods is that they rely on an economic model of fundamentals (Gürkaynak, 2008). Uncertainty about the true economic model raises concerns about misspecification problems. As we show next, information from derivative markets can be useful in this context.

Let n denote the periods to maturity. Under risk neutrality, the log of the forward rate at time t for delivery at time $t + n$ is defined as

$$f_{t,n} = E_t[s_{t+n}] = E_t[v_{t+n}] + E_t[B_{t+n}]. \tag{5}$$

Using Equations (2) and (4), we can rewrite the above expression as

$$f_{t,n} = \phi^n v_t + (1 + r)^n B_t. \tag{6}$$

Leading Equation (1) n periods, the future spot rate at $t + n$ is given by

$$s_{t+n} = v_{t+n} + B_{t+n}. \tag{7}$$

Given that the bubble is ongoing, substituting recursively yields

$$s_{t+n} = \phi^n v_t + \left(\frac{1+r}{\pi}\right)^n B_t + \epsilon^\star_{t+n}, \tag{8}$$

where $\epsilon^\star_{t+n}$ consists of two moving average processes, $\sum_{i=1}^{n} \phi^{n-i} \theta_{t+i}$ and $\sum_{i=1}^{n} \left(\frac{1+r}{\pi}\right)^{n-i} \epsilon_{t+i}$. A comparison of Equations (6) and (8) reveals that the actual future spot rate is in expectation greater than the forward rate. The difference arises because rational agents at time t are not certain that the bubble will continue and attach a nonzero probability to the bubble bursting. As a consequence, the expected growth rate $(1 + r)^n$ of the bubble component is lower than the actual rate of $(1 + r)^n / \pi^n$. Subtracting (6) from (8) yields

$$
\begin{aligned}
s_{t+n} - f_{t,n} &= \left(\phi^n v_t + \frac{(1+r)^n}{\pi^n} B_t + \epsilon^\star_{t+n}\right) - \left(\phi^n v_t + (1+r)^n B_t\right) = \\
&= (1+r)^n \left(\frac{1}{\pi^n} - 1\right) B_t + \epsilon^\star_{t+n}.
\end{aligned} \tag{9}
$$

According to the above expression, $s_{t+n} - f_{t,n}$ is a linear function of the bubble process and the sum of two moving average processes and therefore it is explosive. From an empirical perspective, this implies that researchers can test for bubbles by simply running unit root tests on $s_{t+n} - f_{t,n}$. These tests have the major advantage over tests on s_t (or $f_{t,n}$) of not depending on fundamentals. Hence, rejection of the null provides conclusive evidence in favor of bubbles.

Several unit root tests have been proposed in the literature to examine whether a time series exhibits periods of explosive behavior. In principle, any of these tests can be applied to $s_{t+n} - f_{t,n}$ to draw inference about bubbles. We employ the GSADF unit root test proposed recently by Phillips et al. (2015a) for three reasons. First, the test is consistent with multiple changes in regime. This feature is very appealing since, in the presence of nonlinear dynamics, such as those displayed by periodically collapsing bubbles, standard unit root tests frequently indicate stationarity even though the process under examination is inherently explosive (Evans, 1991). Second, simulation evidence in Homm and Breitung (2012) and Phillips et al. (2015a) suggests that the test has accurate size and good power properties, and in many cases is superior to alternative tests for speculative bubbles. In simulation experiments (not reported due to space considerations), we investigate the performance of four alternative bubble-detection tests, proposed by Homm and Breitung (2012), in our setting. The first three are modified versions of the tests derived by Bhargava (1986), Busetti and Taylor (2004), and Kim (2000), and the last is a new Chow-type unit root test. According to the simulation results, apart from Kim's test which exhibits power properties similar to those of Phillips et al. (2015a), the rest of the tests have no power to detect bubbles that burst in sample.[4] This is in line

---

[4]The simulation results for the tests of Homm and Breitung (2012) are available upon request.

with the finding of Homm and Breitung (2012) that the Phillips et al. (2011) procedure is more robust to multiple breaks. Finally, the recursive nature of the GSADF test allows the identification of the periods, if any, during which the series under examination displays explosive dynamics (date-stamping). Appendix A provides a detailed description of the GSADF test and the date-stamping strategy.

The above analysis assumes that agents are risk-neutral and that market fundamentals follow a simple AR(1) process. These assumptions are sufficient but not necessary for valid statistical inference based on unit root tests on $s_{t+n} - f_{t,n}$. Suppose that agents are risk-averse, so that they demand compensation for the risk of holding foreign currency. The definition of the forward rate becomes $f_{t,n} = E_t[s_{t+n}] + p_{t,n} = E_t[v_{t+n}] + E_t[B_{t+n}] + p_{t,n}$ where $p_{t,n}$ denotes a time-varying risk premium (Engel, 1996). The difference between the future spot rate and the forward rate is now given by $s_{t+n} - f_{t,n} = (1+r)^n \left(\frac{1}{\pi^n} - 1\right) B_t + \epsilon_{t+n}^{\star} - p_{t,n}$. It is easy to see that if the risk premium is non-explosive (which is the typical assumption made in the literature, see, e.g., Taylor, 1987; Corbae et al., 1992; Engel, 1996; Carriero, 2006; and Lettau et al., 2008) this expression, like (9), is explosive only in the presence of a bubble. Therefore, even under risk aversion, unit root tests should provide conclusive evidence about bubbles.

Regarding the process for market fundamentals, a necessary condition for valid statistical inference is that the forecast error $v_{t+n} - E_t[v_{t+n}]$ does not increase exponentially in sample. This condition is met by the majority of processes suggested for v_t since it allows for higher-order AR terms, nonlinear dynamics and heteroskedastic errors.[5] A case worth noting that violates this condition is a process with probabilistic regimes, i.e., a 'peso' process, where one of the regimes has an explosive root. Although not typically considered in asset price determination models, such a process may be relevant when studying hyperinflation periods because of the probability of a monetary reform.

In the case of the German interwar hyperinflation, the possibility of an expected structural change in the money supply has been thoroughly studied (see, e.g., Flood and Garber, 1980, 1983; LaHaye, 1985). The findings in the literature suggest that the probability of reform was near zero until the middle of August 1923. Consequently, employing data prior to the middle of August 1923, as has been common in other studies (see, e.g., Casella, 1989; Hooker, 2000), should yield unbiased results.

In the next section we show that when a bubble is ongoing the difference between the future spot rate and the forward rate also gives rise to deviations of the slope coefficient in the Fama regression from its efficient market hypothesis value. We exploit this fact and propose a testing procedure that is complementary to unit root tests on $s_{t+n} - f_{t,n}$.

---

[5]For instance, threshold and smooth transition autoregressive processes, which arise in theoretical models of exchange rate determination with market frictions (Dumas, 1992; O'Connell and Wei, 2002), fall in this category.

2.2 Bubble Detection II: Rolling Fama Regressions

Fama (1984) develops the following regression to test for efficiency in the foreign exchange market

$$s_{t+n} - f_{t,n} = \alpha_n + \beta_n(f_{t,n} - s_t) + u_{t+n}, \tag{10}$$

where α_n and β_n are regression coefficients, and u_{t+n} denotes the error term. In a standard framework with rational expectations and risk neutrality, the forward rate is an unbiased predictor of the future spot and the efficient market hypothesis (EMH) corresponds to $H_0 : \beta_n = 0 (= \beta^{\text{EMH}})$. Following Fama (1984) and Hodrick and Srivastava (1986), a number of studies have reported least squares estimates for the slope coefficient in regression (10),

$$\widehat{\beta}_n = \frac{\widehat{cov}(s_{t+n} - f_{t,n}, f_{t,n} - s_t)}{\widehat{var}(f_{t,n} - s_t)}, \tag{11}$$

that are significantly less than zero and, in many cases, less than minus one. Perhaps the most common explanation for the strong rejection of the null hypothesis is the presence of a risk premium which, if negatively correlated with the forward premium, could cause a downward bias in $\widehat{\beta}_n$ (see, e.g., Maynard, 2003).

An ongoing bubble, on the other hand, has the opposite effect. It drives $\widehat{\beta}_n$ above rather than below the EMH value. To simplify the analysis, let fundamentals follow a random walk, $\phi = 1$. We have already obtained an expression for the regressand in the Fama regression in Equation (9). Using (1) and (6), we can also obtain an expression for the forward premium [6]

$$f_{t,n} - s_t = \left(v_t + (1+r)^n B_t\right) - (v_t + B_t) = \left((1+r)^n - 1\right) B_t. \tag{12}$$

Substituting (9) and (12) into the expression for the least squares slope coefficient yields

$$\widehat{\beta}_n = \frac{(1+r)^n \left(\frac{1}{\pi^n} - 1\right)}{(1+r)^n - 1} + \frac{\widehat{cov}(\epsilon_{t+n}^\star, B_t)}{[(1+r)^n - 1]\widehat{var}(B_t)}.$$

Because the bubble term is uncorrelated with future errors the plim of $\frac{\widehat{cov}(\epsilon_{t+n}^\star, B_t)}{[(1+r)^n - 1]\widehat{var}(B_t)}$ is zero (for a proof see Rubin, 1950), and

$$\text{plim } \widehat{\beta}_n = \frac{(1+r)^n \left(\frac{1}{\pi^n} - 1\right)}{(1+r)^n - 1} > 0 = \beta^{\text{EMH}}. \tag{13}$$

Hence, when the bubble is ongoing the value of the slope coefficient exceeds the EMH value.

---

[6]Notice that the variables that enter the Fama regression are explosive, which implies that regression coefficients follow non-standard limiting distributions and results based on conventional inference methods can be misleading (Cavanagh et al., 1995; Engsted and Nielsen, 2012; Jansson and Moreira, 2006; Phillips and Magdalinos, 2008). We show that the variance of the least squares estimator (11) depends on the parameters π and r in a simulation experiment below and return to the issue of statistical inference at the end of this section.

Table 1 shows how the theoretical value of the slope coefficient (plim $\widehat{\beta}_n$), computed using expression (13), changes with values of $r = \{0.05, 0.10\}$, $\pi = \{0.3, \ldots, 0.9\}$ and $n = \{1, 4\}$. We observe that the size of the deviation of plim $\widehat{\beta}_n$ from the EMH value of zero depends crucially on the probability of the bubble regime π. This is to be expected since when a bubble develops but the probability of such an event occurring is small (e.g., $\pi = 0.3$), there is a large difference between the realized future states of the economy (s_{t+n}) and the expectations of rational agents ($f_{t,n}$). In these rare occasions, the deviation of the slope coefficient from its EMH value is large. As the probability of the bubble regime increases, the difference between the future spot and the current forward rate narrows and the slope of the Fama regression approaches the EMH value.

–Table 1–

In many empirical applications, tests for bubbles are applied to small samples of data. A question that naturally arises is what are the small-sample properties of the least squares estimator (11) when the regression variables contain a bubble. To answer this question, we run a simulation experiment. In the experiment, we set the number of repetitions equal to 20,000, the sample size equal to 25, $\theta \sim \mathcal{N}(0,1)$, $\epsilon \sim \mathcal{N}(0,0.09)$, $n \in \{1,4\}$, $r \in \{0.05, 0.1\}$, and $\pi \in \{0.3, \ldots, 0.9\}$. The mean and standard deviation of the estimates are reported in Table 1 next to the theoretical values. Importantly, there is no evidence of bias with mean estimates being almost identical to theoretical values. Regarding the standard deviation, we observe that this statistic decreases with r and increases with π, i.e., it decreases as the regression variables become more persistent, with very few exceptions.

We have established that the slope coefficient in the Fama regression exceeds the EMH value when a bubble is ongoing and fundamentals follow a random walk. We note that this result cannot be generated by explosive fundamentals. Suppose that fundamentals follow an AR(1) process with $\phi > 1$ and there are no bubbles. The forward premium is given by

$$f_{t,n} - s_t = (\phi^n - 1)v_t, \tag{14}$$

and

$$s_{t+n} - f_{t,n} = \epsilon'_{t+n}, \tag{15}$$

where $\epsilon'_{t+n} = \sum_{i=1}^{n} \phi^{n-i}\theta_{t+i}$. Substituting (14) and (15) into the expression for the least squares slope coefficient yields

$$\widehat{\beta}'_n = \frac{\widehat{cov}(\epsilon'_{t+n}, v_t)}{(\phi^n - 1)\widehat{var}(v_t)}.$$

The plim of the above expression is zero and, therefore, plim $\widehat{\beta}'_n = \beta^{\text{EMH}}$. We conclude that, like unit root tests on $s_{t+n} - f_{t,n}$, tests based on $\widehat{\beta}_n$ allow the researcher to differentiate between explosive behavior brought about by the process for fundamentals and by the presence of rational bubbles.

10

Hypothesis Testing with a Persistent Regressor. Fama's predictive regression (10) may contain a highly persistent, even explosive (under the alternative of bubbles), regressor which renders standard inference procedures invalid (see Campbell and Yogo, 2006; Cavanagh et al., 1995; Engsted and Nielsen, 2012; Jansson and Moreira, 2006; Phillips and Magdalinos, 2008). As shown by Stambaugh (1999), this inference problem may be further amplified if the innovations in Fama's regression and the innovations of the forward premium are correlated, i.e. the regressor is endogenous.

Consider the following bivariate system

$$y_{t+1} = \beta x_t + u_{1,t+1}, \tag{16}$$

$$x_{t+1} = \rho x_t + u_{2,t+1}, \tag{17}$$

where $y_{t+1} = s_{t+1} - f_{t,1}$, $x_t = f_{t,1} - s_t$, and $(u_{1,t+1}, u_{2,t+1})'$ is a martingale difference sequence with finite fourth moment and possibly non-zero correlation coefficient $\delta = \text{corr}(u_{1,t+1}, u_{2,t+1})$. Our objective is to test the null hypothesis of no bubbles or equivalently no predictability ($H_0 : \beta = 0$) against the alternative of an ongoing bubble ($H_1 : \beta > 0$) in the presence of uncertainty about the degree of regressor persistence ρ. It is typical in the predictive regression literature to model this uncertainty by adopting a local-to-unity setting where $\rho = 1 + c/T$. In this setting, widely employed test statistics, such as t-ratios, have distributions that depend on the unknown localizing coefficient c and the correlation coefficient δ (see Cavanagh et al., 1995). Although δ can be consistently estimated from the data, the nuisance parameter c cannot, which leads to a nonstandard and nonpivotal inference problem.

Several methods have been proposed in the literature to bypass this problem. By far, the most popular are Bonferroni-type methods (Campbell and Yogo, 2006; Cavanagh et al., 1995; Hjalmarsson, 2011; Torous et al., 2004). These methods involve, first, inversion of a unit root test to construct confidence intervals for c (Stock, 1991), and, second, inversion of the nonpivotal limit distribution of the test statistic corresponding to the null hypothesis of no bubble, $H_0 : \beta = 0$. Unfortunately, despite their popularity, Bonferroni-type methods have some undesirable properties. In a recent paper, Phillips (2014) shows that confidence intervals constructed in the first stage of the testing procedure are invalid at the limits of the domain of the localizing coefficient. The failure at the boundary has adverse implications for the performance of predictive regression tests. Specifically, for values of ρ distant from unity, the t-test of Cavanagh et al. (1995) is undersized, while the widely used two-sided Q-test of Campbell and Yogo (2006) is oversized, falsely rejecting the null of no predictability with probability approaching unity when in fact the null hypothesis is true.

A recently developed alternative to Bonferroni-type methods is the IVX method of Phillips and Magdalinos (2009), Phillips and Lee (2013) and Kostakis et al. (2015). The key idea behind the IVX method is to create an instrument, which relies solely on the regressor (hence the name IVX) but falls within the class of near-stationary processes, and use it to remove endogeneity and conduct valid statistical inference. This approach has several advantages that make it particularly attractive for our purposes: i) it allows stan-

dard chi-square testing for a variety of autoregressive processes for the regressor, from stationary to mildly explosive, ii) it has good finite-sample properties (Kostakis et al., 2015), iii) it is straightforward to implement because it is based on linear regression, and iv) it can be extended to long-horizon predictive regressions (Phillips and Lee, 2013). For all these reasons, we adopt the IVX method and a rolling-window framework to test for bubbles (for a description of the IVX testing procedure see Appendix B).[7]

Forward and Futures Prices Although our theoretical analysis has focused on the foreign exchange market, the proposed framework is general, and can be applied directly to any market for which forward or futures prices are available. It should be noted, however, that in practice the distinct characteristics of forward and futures contracts may cause their prices to diverge, and thus lead to different estimation results. On the one hand, forward contracts are unstandardized, they trade over the counter, and they settle at maturity. On the other hand, futures contracts are standardized with respect to their size and maturity date, and they trade in organized exchanges, in which a clearing house guarantees against credit risk, and ensures that all profits and losses are settled at a daily basis (*marking to market*). As a consequence, futures contracts usually have lower default risk and are more liquid.[8] Forward and futures contracts may also differ due to taxes, transaction costs, and margin requirements (Hull, 2015). Despite their distinct characteristics, the findings of the empirical literature suggest that, for financial assets, the prices of these two types of derivative contracts are not significantly different and therefore, for our purposes, they can be treated as equivalent (see, Chow et al., 2000, and the references therein).

---

[7]Alternatives to Bonferroni-type methods are also provided by, among others, Jansson and Moreira (2006) and Elliott et al. (2015). The first paper develops a conditional likelihood procedure based on a set of sufficient statistics. This testing procedure is theoretically appealing because it exhibits specific optimality properties in the local-to-unity model. However, simulation results in Kasparis et al. (2015) and Kostakis et al. (2015) show that, in practice, it exhibits substantial size distortions in cases of strong endogeneity. Moreover, Kasparis et al. (2015) find that there are severe difficulties with its numerical implementation. The second paper proposes a nearly optimal test that maximizes power over a range of values for the unknown localizing parameter c. An attractive feature of this test is that it switches to a standard t-test when the localizing coefficient estimate crosses a specific threshold and, thus, it avoids the problems of Bonferroni-type methods. Elliott et al. (2015) show that, for large sample sizes, the nearly optimal test displays good properties and almost uniformly dominates the Q-test of Campbell and Yogo (2006). In our context, it is imperative that the testing procedure also displays good small-sample properties since periodically collapsing bubbles, if they exist, may be short-lived. However, the small-sample performance of the Elliott et al. (2015) test has not been explored yet. By conducting a battery of simulation experiments, we find that, for small samples, the Elliott et al. (2015) test exhibits severe oversizing when there is strong endogeneity, the regressor is stationary, and there is no prior knowledge of any nuisance parameters. On the other hand, the test is undersized in the presence of an explosive regressor - a property which may well be important when testing for speculative bubbles. The simulation results are available upon request. We are grateful to Michael Jansson for kindly providing his Matlab code, and to Graham Elliott for making his code available online `http://econweb.ucsd.edu/~gelliott/RECENTPAPERS.htm`.

[8]An exception is the foreign exchange market. Forward foreign exchange contracts are more actively traded than futures contracts, and have comparable default risk since the majority of transactions involves major banks and securities dealers (see, for instance, Cornell and Reinganum, 1981).

3   Monte Carlo Simulations

In this section, we examine the performance of the recursive unit root test and the rolling Fama regression in detecting and dating bubble episodes by conducting a battery of Monte Carlo experiments. The design of the experiments is similar to that of Phillips et al. (2015a) and allows us to investigate the role of a number of factors in the performance of the tests. These factors include the number of bubble episodes, the location of the bubble in the sample, the duration of the bubble, the probability of the bubble collapsing, the periods to maturity, and the bubble type.

3.1   Processes with a Single Bubble Episode

For our first Monte Carlo exercise, we consider a scenario with a single bubble episode which starts at time $\tau_f = 31, 51$ or 71 and lasts for $d = 10, 15,$ or 20 periods. The sample size in this and the following exercises is set equal to $T = 100$ observations, the number of repetitions is 1,000, the nominal significance level is 5 percent, and the values for the parameters describing the economy are, unless otherwise stated, $\pi = \{0.5, 0.7\}$, $r = \{0.05, 0.1\}$, $n = 1$, $\sigma_\epsilon = 0.3$, and $\sigma_\theta = 1$.

Before proceeding to the main Monte Carlo simulation results, we demonstrate the applicability of the two methods on one of the realizations generated with $r = 0.05$, $\pi = 0.7$, $\tau_f = 71$ and $d = 15$. Figures 1 and 2 display test statistics (BSADF and IVX) and 95 percent critical value sequences for the Phillips et al. test and for the rolling Fama regression, respectively. Overall, the observed patterns are consistent with the theoretical analysis. Once the bubble starts the estimated test statistics gradually increase and eventually exceed their critical values rejecting the null hypothesis of no bubbles. When the bubble bursts both statistics fall below the critical bound almost instantaneously and remain there until the end of the sample.

–Figures 1 and 2–

Regarding the main simulation results, Table 2 reports the empirical power of the tests Rej., the percentage of times one, two and more than two bubble episodes are detected, the mean estimate of the start date of the bubble $\bar{\tau}_f$, and the mean estimated duration $\bar{d}$. Several interesting conclusions emerge from the results. Starting with the empirical power, we observe that, for both tests, it increases with the duration of the bubble d and the structural parameter r, and it decreases with the probability of the economy being in the bubble regime π. The positive relation between empirical power and bubble duration is intuitively obvious and does not require explanation. As far as r and π are concerned, these two factors determine the value of the AR coefficient in the bubble process (3) and, consequently, the degree of explosiveness of $s_{t+n} - f_{t,n}$ and $f_{t,n} - s_t$. When the value of r increases and the value of π decreases, the AR coefficient of

the bubble process takes larger values, and the two variables that enter the Fama regression equation become more explosive. This increase in the degree of explosiveness raises the power of the tests. Furthermore, it explains why, as π decreases and r increases, the tests detect the bubble episode earlier (see $\bar{\tau}_f$) and the mean estimated bubble duration ($\bar{d}$) becomes longer.

–Table 2–

Turning to the comparison of the results for the two methods, we see that the empirical power of the rolling IVX test is similar to or higher than the recursive unit root test. The difference between tests is particularly apparent (about 60 percentage points) when the bubble duration d is short, the value for π is large, and the value for r is small. In other words, when the degree of explosiveness of the Fama variables is low, and the bubble is short-lived. A possible explanation for the superiority of the IVX over the unit root test in terms of power is that, under the null hypothesis of no bubble, the $s_{t+n} - f_{t,n}$ process (9) is in fact stationary, rather than a unit root, which makes the unit root test conservative (see also the discussion of Phillips et al. (2015a, Section 5) about the application of unit root tests to price-to-fundamentals ratios).

With regards to date-stamping, the results in Table 2 (see $\bar{\tau}_f$) suggest that the Fama regression is again superior to the unit root test because it always detects earlier the emergence of the bubble. Both methods, however, perform remarkably well in detecting the collapse of the bubble, which is the reason the sum of $\bar{\tau}_f$ and $\bar{d}$ is the same as the date of collapse. The performance of the methods is also exceptional in terms of detecting the correct number of bubble episodes. Specifically, the maximum number of times that the methods falsely detect more than one bubble episodes never exceeds 6 percent of the total number of repetitions. Finally, it is worth noting that the location of the bubble does not affect the power of the tests, the time required to detect the bubble, $\bar{\tau}_f - \tau_f$, and the mean estimated duration.

3.2 Processes with Two Bubble Episodes

For the second Monte Carlo exercise, we consider a scenario with two bubble episodes, which start at times $\tau_{f,1} = 31$ and $\tau_{f,2} = 71$, and last for $d_1, d_2 = \{10, 15, 20\}$ periods. Not surprisingly, the results presented in Table 3 suggest that the empirical power in the presence of two bubbles is higher compared to that for one bubble and, like for the one-bubble data generating process (DGP), it depends positively on the duration of the bubble episodes d, and negatively on the probability of being in the bubble regime π. As before, the rolling Fama regression is equally or more powerful than the recursive unit root test, and it detects bubbles earlier. In the majority of cases, the Fama regression is also superior in detecting the correct number of bubble episodes. The largest differences between the two methods occur when one of the two bubbles is short-lived and the other

is not (i.e., d_1 or d_2 equals 10 periods, and $d_1 \neq d2$), $\pi = 0.7$, and $r = 0.05$. For these parameter values, the Fama regression detects two bubbles in more than 70 percent of the cases, while the recursive unit root test in less than or equal to 14 percent of the cases. A possible reason for these large differences is that, as explained before, the recursive unit root test may be conservative, frequently failing to detect short-lived bubbles when the probability of being in the bubble regime is high, i.e. $\pi = 0.7$. On the contrary, for $\pi = 0.5$, both methods perform well in terms of power and detecting the correct number of bubble episodes.[9]

–Table 3–

3.3 Different Periods to Maturity, n

In the previous two simulation exercises, we set the number of periods to maturity n equal to one. We now extend the analysis to $n = 2, 3$ and 4. The simulation design is identical to that used for the one-bubble simulation exercise with the exception that we now consider a single starting date $\tau_f = 31$, and a bubble duration of $d = 10$ periods.

–Table 4–

According to the results reported in Table 4, the number of periods to maturity has a substantial positive effect on the empirical power of the recursive unit root test for $\pi = 0.7$ but has no effect for $\pi = 0.5$. Specifically, when n increases from 2 to 4, the number of rejections almost doubles (from 22 to 41 percent) for $\pi = 0.7$ and $r = 0.05$, and it increases from 44 to 58 percent for $\pi = 0.7$ and $r = 0.1$. Regarding the Fama regression, the empirical power of the IVX test does not appear to depend on the periods to maturity n. However, it should be noted that as n increases the Fama regression behaves slightly worse in detecting the correct number of bubble episodes. This property is not shared by the recursive unit root test. Overall, the conclusion that emerges from these results is that, as n increases, the performance of the Phillips et al. (2015a) test can improve substantially, which makes this test more attractive for large n.

3.4 Periodically Collapsing Bubbles

In the last simulation exercise, we examine bubble episodes that do not occur at a specific part of the sample, as before, but instead emerge randomly. We also expand the

---

[9]We have run simulation experiments using DGPs with three bubble episodes. The results indicate that increasing the number of bubble episodes does not change the relative performance of the two methods.

analysis by considering the type of bubbles proposed by Evans (1991). This bubble DGP is given by

$$
B_{t+1} = \begin{cases} (1+r)B_t\eta_{t+1}, & \text{if } B_t \leq b, \\ \left[\lambda + \frac{1}{\pi}(1+r)\zeta_{t+1}\left(B_t - \frac{1}{(1+r)}\lambda\right)\right]\eta_{t+1}, & \text{if } B_t > b, \end{cases} \tag{18}
$$

where λ and b are positive constants with $\lambda < (1+r)b$, η_{t+1} is a lognormal variable scaled to have a mean of unity (that is, $\eta_{t+1} = \exp(\kappa_{t+1} - \sigma_\kappa^2/2)$ where $\kappa_{t+1} \sim \mathcal{N}(0, \sigma_\kappa^2)$), and ζ_{t+1} is a Bernoulli process that takes the value of one with probability π and the value of zero with probability $1 - \pi$. Like Blanchard's DGP, it is easy to verify that the DGP specified by (18) satisfies the condition for a rational bubble, i.e., $E_t[B_{t+1}] = (1+r)B_t$. However, contrary to Blanchard's DGP, bubbles of the Evans' type never fall to zero since $B_t > 0$ implies $B_s > 0$ for all $s > t$. Consider, for example, a bubble that starts below the threshold b. The bubble initially grows at a constant expected rate of $(1+r)$. Eventually the bubble exceeds b, and once this happens, its expected growth rate conditional on the bubble not collapsing increases to $(1+r)/\pi$. When the bubble collapses, it drops to a positive expected value of λ, and the cycle begins again.

The fact that the expected growth rate conditional on the bubble erupting is greater than the unconditional expected growth rate implies a gap between the future spot and forward exchange rates. By adopting the framework of Section 2, it is straightforward to show that the difference between the two exchange rates is given by

$$
\begin{aligned}
s_{t+n} - f_{t,n} &= (1+r)^n\left(\frac{1}{\pi}\prod_{i=1}^{n}\eta_{t+i} - 1\right)B_t + \left(1 - \frac{1}{\pi}\right)\lambda\sum_{i=1}^{n}\left(\frac{1+r}{\pi}\right)^{n-i}\prod_{j=i}^{n}\eta_{t+j} \\
&\quad + \sum_{i=1}^{n}\phi^{n-1}\theta_{t+i}.
\end{aligned} \tag{19}
$$

According to this expression, the $s_{t+n} - f_{t,n}$ process is explosive and does not depend on fundamentals. Thus, the application of the recursive econometric tests is valid for the bubble DGP of Evans (1991).

For the simulation exercise, we follow Evans (1991) and scale up bubbles generated from (18) by a factor of twenty. Furthermore, we set the parameters of the two periodically collapsing bubble processes and of the process for fundamentals equal to $\pi = \{0.3, 0.5, 0.7, 0.9\}$, $r = \{0.05, 0.1\}$, $n = \{1, 2, 3, 4\}$, $\lambda = 0.5$, $b = 1$, $\sigma_\epsilon = 0.3$, $\sigma_\theta = 1$, and $\sigma_\kappa = 0.05$. Finally, in order to make the analysis meaningful from a practical standpoint, we focus on bubble episodes that are not very short-lived, and discard series for which the duration of the longest bubble episode is less than or equal to five consecutive time periods. Table 5 presents the empirical power of the two econometric tests in this setting. Panel A reports results for the bubble DGP of Blanchard (1979), and Panel B for the bubble DGP of Evans (1991).

–Table 5–

16

Overall, the results are in line with those for the previous simulation exercises. Both tests display very high power (equal to or greater than 95 percent) when π is low. As π increases, their empirical power decreases. The decline becomes particularly apparent for the unit root test when the number of periods to maturity n and the structural parameter r take small values. For instance, for the DGP of Evans with $n = 1$ and $r = 0.05$, the power of the GSADF drops from 100 to 16 percent as π increases from 0.3 to 0.9, while the corresponding drop in power for the IVX test is from 100 to 86 percent.

The relationship between π and empirical power, however, is not always monotonic. Looking at the results for $r = 0.1$ and $\pi = \{0.7, 0.9\}$, we observe cases where an increase in π is associated with higher (although slightly) empirical power. To explain this finding, we note that raising π has two opposite effects in the presence of periodically collapsing bubbles. On the one hand, similarly to the previous simulation experiments, when π increases and the economy is in the bubble regime, the AR coefficient of the bubble process decreases, and $s_{t+n} - f_{t,n}$ and $f_{t,n} - s_t$ become less explosive. As an implication, an increase in π tends to lower the power of the tests. On the other hand, as π increases, bubble episodes become more frequent, which tends to raise power. Hence, what our results suggest is that the overall effect is ambiguous and depends on the specific values of the DGP parameters.

Finally, a comparison of the results for the two tests (IVX and GSADF) reveals that the rolling Fama regression method is again more (in some cases substantially) powerful than the recursive unit root. In accordance with the findings from the previous simulation experiment, the difference between methods is large when the number of periods to maturity n is small and the probability of the bubble regime π is high. As n increases the performance of the recursive unit root test generally improves, while the performance of the Fama regression remains the same or slightly worsens. As a consequence, for $n = 4$, the difference in performance is substantially smaller -in some cases, less than 5 percentage points.

4   Empirical Applications

In this section, we provide two empirical applications of the proposed bubble detection methods to US dollar exchange rates. The first concerns the interwar German hyperinflation, and the second the recent floating-rate period. We also illustrate how the proposed methods can be applied to stock market data, and, in particular, the S&P 500 index, using futures prices.

4.1   The German Hyperinflation

For the first empirical application, we use weekly German mark-US dollar spot and forward exchange rates spanning the period December 24th, 1921 to August 11th, 1923. The data is taken from Einzig (1937). As already mentioned, we restrict the analysis to the period before the middle of August 1923 to avoid potential biases related to 'peso' prob-

lems. The starting date, on the other hand, corresponds to the change of convention to a four-week forward rate. The use of weekly data allows us to analyze a much larger sample than those typically used in the German hyperinflation literature (see, e.g., Michael, Nobay, and Peel, 1994).

–Figures 3 and 4–

Figure 3 shows the evolution of the logarithm of the spot rate, s_t, over time. As it is evident from the figure and well documented in the literature, the mark depreciated rapidly against the dollar in the early 1920s. At the start of our sample period, December 24th, 1921, a dollar was trading for 190 German marks. By August 11th, 1923, the same amount of dollars could buy more than three million marks, which implies an increase by a factor of 17 thousand in less than two years.

Similarly to other hyperinflations, the large movements in the exchange rate were mainly driven by the rate of money supply growth. A rapid monetary expansion started in the summer of 1921, when the German authorities adopted a policy of selling the mark to buy foreign currencies in order to pay World War I reparations, and finished in November 1923 with the introduction of a new currency, the Rentenmark, which replaced the old currency at a conversion rate of one to a trillion. According to texts regarded as authoritative on the German hyperinflation (e.g., Bresciani-Turroni, 1937; Graham, 1930), the rate of growth of the money supply was the main but not the sole driving force of exchange rate movements. A second factor that appears to have played an important role in the foreign exchange market at that time is destabilizing speculation. In his analysis of foreign exchange trading, Bresciani-Turroni (1937, p. 101) underlines the role of self-fulfilling beliefs in the determination of the exchange rate: *"The dollar rate was increased rapidly from one day to the next and from one hour to the next by the expectation of a future rise"*. The question we address in this section is whether speculative bubbles arose in the foreign exchange market.

–Table 6–

We begin our empirical analysis by examining the integration properties of spot and forward exchange rates. Perhaps surprisingly, despite a large amount of research on the German hyperinflation, the properties of these series have not been examined so far.[10] The second and third columns of Table 6 report standard ADF and GSADF test statistics, respectively. The fifth column reports 95 percent critical values. The fact that the ADF statistics for the whole sample are greater than the 95 percent critical value indicates

---

[10]Funke et al. (1994) focus on Poland's hyperinflation period and run a regime switching unit root test. Nielsen (2008) examines the Yugoslavian hyperinflation using a vector autoregression model.

that the series have been explosive in the sample period. However, as shown in Section 2, explosive dynamics in asset prices does not constitute evidence neither in favor nor against bubbles since it may reflect bubbles or explosive market fundamentals, or both of these factors.

To shed light on the existence and duration of speculative bubbles, we investigate the integration properties of the difference between the future spot and forward rates, $s_{t+4} - f_{t,4}$, which is plotted in Figure 4. Looking at the results reported in the fourth column of Table 6, we observe that the GSADF statistic exceeds the 95 percent critical value but the ADF fails to do so. This implies that, although there are episodes of explosive dynamics, the $s_{t+4} - f_{t,4}$ series is not explosive throughout the sample period. The conclusion that emerges is that a speculative bubble process in the foreign exchange market exists, however, it is not ongoing for the entire sample -it collapses.

In order to identify the bubble period(s), we plot the BSADF statistics for $s_{t+4} - f_{t,4}$ and the corresponding 95 percent critical value sequence in Figure 5.[11] A comparison of the BSADF statistics with the critical value sequence suggests that self-fulfilling expectations drove the German mark-US dollar away from market fundamentals in two occasions: in July-August 1922, and in July 1923. The bubble episodes are short-lived, with the second episode lasting a single period. Given the short duration of exuberance, it is not surprising that the GSADF test rejects the null of a unit root in $s_{t+4} - f_{t,4}$, while the standard ADF does not.

–Figures 5 and 6–

We complement the unit root analysis by running rolling Fama regressions. Figure 6 displays the sequence of IVX statistics together with 95 percent critical values. The maximum IVX statistic is 4.069, which is greater than the 95 percent critical value, and thus the null hypothesis of no bubbles is again rejected by the data. Overall, the observed pattern of the IVX statistic is very similar to that of the BSADF statistic. Starting below the critical value sequence at the beginning of the sample, the statistic rises and becomes significant in July 1922, indicating a first bubble episode. It then falls back below the critical value and remains insignificant until June 1923. In June 1923, the statistic starts increasing sharply again and exceeds the critical value sequence for the second time. This second bubble episode continues until August 1923 –a month before the authorities suspended forward trading in the German mark due to its rapid depreciation.

Interestingly, the two periods identified as consistent with speculative bubbles coincide with those that Bresciani-Turroni (1937, p.102) gives as examples of destabilizing

---

[11]The minimum window size used for the recursive unit root test is 18 weeks, which is smaller than the window size of 25 weeks used for the rolling Fama regression. (A discussion about the selection of the (minimum) window size is provided in the Appendix.) As a consequence, the estimates for the IVX statistic start later than those for the BSADF statistic. Figures 5 and 6 plot the two test statistics for the same time period in order to allow a direct comparison of the results of the two methods.

speculation that pushed the mark away from fundamental conditions: "*For some time it was foreign speculation which provoked the great fluctuations of the exchanges...[ ] By August 1922 the dollar rate had risen beyond 1,900 marks; but a violent reaction set in and the dollar fell to little more than 1,200 marks...[ ] Towards the middle of August 1923 the dollar rate rose giddily to 1, 2, 3, and 5 million marks; later falling suddenly to 3 millions. These examples [] show that speculation did not exercise stabilizing influence on the exchanges, but had rather the opposite effect.*" Regarding the first period, Graham (1930, p.50) singles it out as representative of the *flight from the mark*: "*The greatest effect exerted by speculators in exchange is ordinarily supposed to have occurred after the middle of 1922...*". Moreover, this period coincides with the peak deviation of the mark from its purchasing power parity. The gap between the internal value of the currency and its exchange value was the largest among all European countries at the time (for historical evidence on the misalignment of the currency see Graham, 1930, p.119).

Our analysis does not employ a specific model of exchange rate determination and, therefore, no direct conclusions can be drawn about the time series properties of market fundamentals. Nonetheless, the results of the standard ADF test on s_t and $s_{t+4} - f_{t,4}$ taken together do provide indirect evidence. Bearing in mind that $s_{t+4} - f_{t,4}$ does not depend on market fundamentals, while s_t does, the fact that the standard ADF statistic is insignificant for $s_{t+4} - f_{t,4}$ but it is significant for s_t (see Table 6) suggests that the process for market fundamentals is explosive. This conclusion is consistent with the finding of Flood and Garber (1980) of an explosive money supply process during the German hyperinflation. It also supports Nielsen's (2008) view that hyperinflation data display explosive dynamics.

Before proceeding to the second empirical application, we would like to highlight two implications of the above results for standard statistical tests applied to the German hyperinflation. First, the fact that the process for market fundamentals is most likely explosive implies that statistical inference about bubbles based on unit root tests on spot prices alone is invalid. For this reason, either spot and forward rates should be used or an asset price determination model needs to be estimated. Second, because the identified bubble periods are short, standard statistical tests may fail to reject the null hypothesis of no bubbles, even if they are based on a correctly specified model, due to their low power. In this context, econometric models that allow for structural change, like the ones employed in this paper, seem more suitable.

4.2   The Recent Float

After the collapse of the Bretton Woods system, the world economy witnessed a dramatic change in the behavior of foreign exchange markets. Exchange rates among major currencies became, on the one hand, highly volatile in the short run and, on the other, highly persistent, moving over long periods of time in one direction (Engel and Hamilton, 1990). A prime example, that has attracted a considerable amount of attention in the literature, is the sustained appreciation of the US dollar against the British pound in the first half of the 1980s (see, e.g., De Grauwe and Rovira-Kaltwasser, 2012; Froot and

Thaler, 1990). During the period December 1980 to February 1985, the British pound price of the US dollar increased by more than 100 percent, from 0.418 to 0.924, which raised serious concerns about whether movements in the value of the dollar reflect only changes in macro fundamentals. Most notably, in June 1985, the then Federal Reserve Chairman, Paul Volcker, referred to an *overvalued* dollar and the possibility of a sharp decline (Feldstein, 1994). In line with Volcker's statement, the results of several studies from that period suggest that it is very difficult to reconcile the large dollar appreciation with the predictions of conventional economic models (based on variables such as inflation rates, interest rate differentials, and relative income) and thus other factors must also be taken into account, such as speculative bubbles (see, for instance, Evans, 1986; Sarno and Taylor, 2002).

In this section, we re-examine the presence of speculative bubbles in the British pound-US dollar foreign exchange market by applying the two methods proposed in Section 2 to monthly data on spot and forward rates for the period January 1979 to December 2013. The data is downloaded from the Bank of England website `http://www.bankofengland.co.uk/`. An advantage of using data for the recent float, in comparison to data for the interwar period, is the greater availability of forward rates, which allows us to obtain and compare results for different periods to maturity; namely one, three, and six months.

–Figures 7 and 8–

Figures 7 and 8 show the evolution of the logarithm of the spot rate and the difference between the logarithm of the future spot and the logarithm of the six-month forward over time. It is evident from Figure 7 that the long swings in the British Pound-US dollar rate is not a phenomenon unique to the 1980s. For instance, in the first half of 2000s, the dollar experienced another long (but not as extreme as that of the 1980s) period of depreciation, with the spot rate falling by more than one-third from June 2001 to October 2007, and then rebounding a year later. Thus, different periods of exuberance in the foreign exchange market may exist during the recent float. Turning to the time series plot of the difference between future spot and forward rates (Figure 8), we observe large and persistent deviations from the EMH value of zero, especially in the 1980s, early 1990s, and mid-to-late 2000s. However, the series does not exhibit trend-like behavior over very long time periods like the spot rate, which suggests a lower degree of persistence.

–Table 7–

The unit root tests results presented in Panel A of Table 7 provide statistical support to this conclusion. The ADF and GSADF statistics for the spot rate (second column) are much larger than those for the future spot rate minus the one-month forward rate (third

column) and, although this gap narrows as the number of periods to maturity increases, it remains substantial.[12] What is more important is that, according to the results, the null hypothesis of a unit root cannot be rejected in favor of the alternative of explosive dynamics at the 5 percent significance level for any of the series examined. The fact that the unit root tests cannot detect any period of exponential growth, even in the spot rate, suggests that there is no evidence in favor of rational speculative bubbles. This is also the conclusion that emerges from the results for rolling Fama regressions. Looking at Panel B of Table 7, we observe that, irrespective of the number of periods to maturity, the maximum IVX test statistic is below the 95 percent critical value. As an implication, the slope of the Fama regression never significantly exceeds the EMH value and, therefore, the no-bubbles hypothesis cannot be rejected.

In contrast to our findings, the studies of Meese (1986) and Evans (1986) provide evidence in support of speculative bubbles in the British pound-US dollar rate in the early 1980s. A possible explanation for the discrepancy between the former study and ours is related to the joint hypothesis problem. In particular, Meese (1986) employs the Hausman-type test of West (1987) which relies on a particular economic model and, consequently, may falsely reject the null hypothesis due to an erroneous assumption about fundamentals.[13] Regarding the latter study, Evans (1986) advocates a test that exploits the fact that the median of excess returns is different from zero in the presence of periodically collapsing bubbles. However, as Evans (1986) points out, a nonzero median may arise due to factors other than bubbles, such as nonsymmetric innovations of fundamentals. That is, the test is based on a necessary but not a sufficient condition for bubbles.

4.3 The US Equity Market

As a final application, we examine the US equity market. Our monthly dataset includes the S&P 500 price index, an observed market fundamental series (namely, S&P 500 dividends), and S&P 500 futures prices for horizons of three, six, and nine months.[14] The first two series are downloaded from Robert Shiller's website and the remaining from Bloomberg. The dataset starts in December 1982, which corresponds to the commencement of S&P 500 futures trading on the Chicago Mercantile Exchange in April of the same year, and ends in March 2015.

–Figures 9 & 10–

---

[12]Unit root tests results for one-, three-, and six-month forwards rates are not reported because they are virtually identical to those for the spot rate.

[13]Meese (1986, p. 367) too recognizes the possibility that the economic model adopted may not be well-suited for the dollar-pound rate.

[14]We follow Chernenko et al. (2004) and obtain regularly-spaced futures prices by linearly interpolating near-term contracts. We do not consider a twelve-month horizon to avoid issues related to low trading activity.

Figure 9 displays the logarithm of the real S&P 500 price index. As can be seen from the figure, the period under examination is characterized by a number of market rallies followed by market crashes (for a historical perspective, see Mishkin and White, 2002, and Shiller, 2015). The most remarkable episode is the great surge of the 1990s, the so-called *millennium boom*. In April 1994, the S&P 500 index stood (in nominal terms) at 447 units. In the subsequent years, stock prices increased at an unprecedented rate, which led Alan Greenspan to talk about *irrational exuberance* in the stock market, and by the early 2000s the index had more than tripled, to 1485 units. This rapid price increase was followed by a dramatic market crash. During the crash, the S&P 500 index lost in a few years almost 44 percent of its value, reaching a minimum of 837 units in February 2003. With the exception of the famous stock market run-up and collapse of the 1920s, no other historical episode comes close to resembling the behavior of the S&P 500 around 2000. Two other stock market episodes that, although lesser than the millennium boom, are worth noting are: *Black Monday* in October 1987 when, following a surge of stock prices, the S&P 500 index fell by about 20 percent, setting a new record one-day decline (see Carlson, 2006); and the 2007-2009 US bear market, during which the S&P 500 index lost around 50 percent of its value. As illustrated in Figure 10, which depicts the S&P 500 price-to-dividend ratio, and as argued by Shiller (2015) and many other authors, it is difficult to reconcile these stock price movements with the behavior of observed market fundamentals.

–Table 8–

To investigate the existence of speculative bubbles in the US stock market, we run unit root tests on the logarithm of the real S&P 500 price index, s_t^*, the logarithm of the price-to-dividend ratio, $s_t - d_t$, and the difference between the logarithm of the future spot price and the logarithm of the futures price, $s_{t+n} - f_{t,n}$ with $n = 3$, 6, and 9 months. We also run rolling Fama regressions. All results are reported in Table 8. By focusing on the GSADF test statistics (Panel A), we observe that the null hypothesis of non-explosive behavior can be rejected at the five percent significance level for all series but one, $s_{t+3} - f_{t,3}$. The finding that the S&P 500 price index and the S&P 500 price-to-dividend ratio display episodes of explosive dynamics is in accordance with the recent empirical literature on bubbles in the US stock market (see, e.g., Homm and Breitung, 2012; Phillips et al., 2015a). Our study complements this literature by showing that $s_{t+6} - f_{t,6}$ and $s_{t+9} - f_{t,9}$ also display explosive dynamics. By doing so, it provides novel evidence, based on a more agnostic approach about market fundamentals, in favor of speculative dynamics. Unlike the GSADF test, the results for the standard ADF test and rolling Fama regressions are not statistically significant (see Table 8). The failure of the standard ADF test to reject the null hypothesis is not surprising given, as already discussed, the high likelihood of Type II error in the presence of periodically collapsing bubbles. With regards to the results for the rolling Fama regressions, we note that the

power of the IVX test declines with the maturity horizon (see Section 3) which may explain why the test fails to reject the null of no-bubbles for maturity horizons of six and nine months, while the GSADF does so.

–Figures 11, 12 & 13–

For identifying the periods of exuberance, we plot the sequence of BSADF statistics together with the corresponding 95 percent critical value sequence for the logarithm of the real S&P 500 price index, s_t^*, the logarithm of the S&P 500 price-to-dividend ratio, $s_t - d_t$, and the difference between the logarithm of the future spot index and the logarithm of the futures price index, $s_{t+9} - f_{t,9}$, in Figures 11, 12, and 13, respectively. As can been seen from the two former figures, the BSADF test detects three episodes of exuberance in the real S&P 500 index and in the price-to-dividend ratio. These three episodes are also identified in the study of Phillips et al. (2015a), and correspond closely to Black Monday in October 1987, the millennium boom, and the US bear market of 2007-09. Turning to $s_{t+9} - f_{t,9}$, we observe that the BSADF test detects the latter two episodes but fails to detect speculative dynamics for Black Monday. By comparing the results for s_t^* and $s_t - d_t$ with those for $s_{t+9} - f_{t,9}$, we also observe that the period of exuberance corresponding to the millennium boom is shorter and takes place slightly later in the sample for the $s_{t+9} - f_{t,9}$ series.

In summary, the above analysis suggests that episodes of exuberance do exist in the US stock market. This conclusion holds when controlling for observed market fundamentals, and when futures prices are used to avoid specifying the asset pricing model, with the latter approach yielding somewhat more conservative results.

5   Conclusions

Econometric tests for bubble detection examine a joint hypothesis of no bubbles and a model of asset price determination. These tests fail to provide conclusive evidence about bubbles because rejection of the null may be due to the existence of bubbles or the misspecification of the asset price determination model, or to both of these reasons.

The present paper proposed two novel methods for bubble detection that use spot and forward (or futures) asset prices to alleviate this problem. The underlying idea is simple: when a periodically collapsing bubble is ongoing, both the forward price and the future spot price incorporate it but with different weights because of the probability of collapse. As a direct implication, the forward price becomes a biased predictor of the future spot price, and the difference between the two (as a linear function of the bubble) has an explosive root. On this basis, our first method employs the recursive unit root test of Phillips et al. (2015a) to test for explosive behavior in the difference between the future spot price and the forward price. The second method employs rolling Fama regressions to test the unbiasedness hypothesis. Both methods do not rely on a specific

model of asset price determination, are robust to an explosive root in the process for fundamentals, and are accompanied by date-stamping strategies.

As an empirical application, we analyzed the existence of speculative bubbles in the German mark-US dollar exchange rate during the interwar German hyperinflation, and the British pound-US dollar rate during the recent float. For the former period, we found two short-lived bubble episodes: one in July-August 1922 and another a year later in the summer of 1923. According to texts regarded as authoritative on the interwar hyperinflation, these episodes coincide with large and persistent speculative movements in the German foreign exchange market. Regarding the recent float, we found no evidence of speculative bubbles. To illustrate the applicability of the proposed methods to other markets, we also examined the behavior of the S&P 500 index from 1982 to 2015. Our findings indicate that, during this period, the S&P 500 index displayed at least two episodes of exuberance. These episodes correspond to the millennium boom and the bear market of 2007-09.

Appendix A: The Phillips et al. (2015) Testing Procedure

Consider the following Augmented Dickey-Fuller (ADF) regression equation

$$\Delta y_t = a_{r_1,r_2} + \gamma_{r1,r2} y_{t-1} + \sum_{i=1}^{k} \psi_{r_1,r_2}^i \Delta y_{t-i} + \epsilon_t, \tag{20}$$

where y_t denotes a time series process, $\epsilon_t \overset{iid}{\sim} N(0, \sigma_{r_1,r_2}^2)$, and r_1 and r_2 denote fractions of the total sample size that specify the starting and ending points of a subsample period. We are interested in testing the null hypothesis of a unit root, $H_0 : \gamma_{r_1,r_2} = 0$, against the alternative of explosive behavior in y_t, $H_1 : \gamma_{r_1,r_2} > 0$. Let

$$\text{ADF}_{r_1}^{r_2} = \widehat{\gamma}_{r1,r2} / \text{s.e.}(\widehat{\gamma}_{r1,r2})$$

denote the test statistic corresponding to this null hypothesis. Setting $r_1 = 0$ and $r_2 = 1$ yields the standard ADF test statistic, ADF_0^1. The limit distribution of the ADF_0^1 statistic is given by

$$\frac{\int_0^1 W dW}{\left(\int_0^1 W^2\right)^{1/2}},$$

where W is the standard brownian motion. The ADF_0^1 test entails comparing the ADF_0^1 statistic with the right-tailed critical value from its limit distribution. Although widely employed, this test has extremely low power in detecting episodes of explosive behavior when these episodes end with a large drop in prices, i.e. in the presence of boom-bust dynamics. As a matter of fact, nonlinear dynamics, such as those displayed by periodically collapsing bubbles, frequently lead to finding spurious stationarity even though the process under examination is inherently explosive (see Evans, 1991).

In order to deal with the effect of a price collapse on the performance of the test, Phillips et al. (2011) propose a a recursive procedure which is based on the estimation of the ADF regression on subsamples of the available data. In particular, the authors propose estimating (20) using a forward expanding sample with the end of the sample period r_2 increasing from r_0 (the minimum window size) to one (the entire sample period). Note that the beginning of the sample period is held constant at $r_1 = 0$. The test statistic, called sup ADF (SADF), is defined by

$$\text{SADF}(r_0) = \sup_{r_2 \in [r_0,1]} \text{ADF}_0^{r_2},$$

Under the null hypothesis, the limit distribution of the SADF statistic is given by

$$\sup_{r_2 \in [r_0,1]} \frac{\int_0^{r_2} W dW}{\int_0^{r_2} W^2}.$$

Similarly to the ADF_0^1 test, when the SADF statistic exceeds the right-tailed critical value from its limit distribution, the unit root hypothesis is rejected in favor of explosive behavior.

The SADF test performs well when there is a single boom-bust episode. Simulation experiments in Homm and Breitung (2012) indicate that the SADF outperforms alternative methods, such as the modified Bhargava, the modified Busetti-Taylor and the modified Kim, in terms of power. However, the test may perform poorly when there are more than one boom-bust episodes, and may also be inconsistent.

More recently, Phillips et al. (2015a) derive a new unit root test, the Generalized SADF (GSADF), that covers a larger number of subsamples than the SADF by allowing both the ending point, r_2, and the starting point, r_1, of the subsample to change. This extra flexibility results in substantial power gains in comparison to the SADF. Moreover, the test is consistent with multiple boom-bust episodes. The GSADF statistic is defined by

$$\text{GSADF}(r_0) = \sup_{r_2 \in [r_0,1], r_1 \in [0, r_2 - r_0]} \text{ADF}_{r_1}^{r_2}$$

Under the null hypothesis, the limit distribution of the GSADF statistic is

$$\sup_{r_2 \in [r_0,1], r_1 \in [0, r_2 - r_0]} \left\{ \frac{\frac{1/2}{r}{}_w [W(r_2)^2 - W(r_1)^2 - r_w] - \int_{r_1}^{r_2} W(r) dr [W(r_2) - W(r_1)]}{r_w^{1/2} \{ r_w \int_{r_1}^{r_2} W(r)^2 dr - [\int_{r_1}^{r_2} W(r) dr]^2 \}^{1/2}} \right\}.$$

where $r_w = r_2 - r_1$. Again, rejection of the unit root hypothesis in favor of explosive behavior requires that the test statistic exceeds the right-tailed critical value from its limit distribution.

Suppose that the GSADF test rejects the null hypothesis of a unit root. We are interested to identify the period(s) during which the series under examination displayed explosive dynamics. Phillips et al. recommend a dating strategy based on the Backward sup ADF statistic

$$\text{BSADF}_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \text{BADF}_{r_1}^{r_2}. \tag{21}$$

The authors define the origination date of the bubble as the first observation that the BSADF statistic exceeds its critical value

$$\widehat{r}_e = \inf_{r_2 \in [r_0,1]} \{ r_2 : \text{BSADF}_{r_2}(r_0) > scu_{r_2}^{\beta_T} \}$$

and the termination date as the first date after which the BSADF falls below its critical value

$$\widehat{r}_f = \inf_{r_2 \in [r_0,1]} \{ r_2 : \text{BSADF}_{r_2}(r_0) < scu_{r_2}^{\beta_T} \}$$

where $scu_{r_2}^{\beta_T}$ is the $100\beta_T\%$ critical value of the sup ADF based on $\lfloor r_2 T \rfloor$ observations and β_T is the chosen significance level.[15] The consistency of the above dating strategy in the presence of one or two bubbles is established in Phillips et al. (2012) Section 4.

The computation of the SADF, BSADF, and GSADF test statistics necessitates the selection of the minimum window size r_0. The minimum window size has to be large

---

[15]Note that the researcher may neglect very short periods of exuberance by setting a minimum duration period.

enough to allow initial estimation but it should not be too large in order to avoid missing short episodes of exuberance. Following Phillips et al. (2015a), we use the rule-of-thumb: $r_0 = 0.01 + 1.8/\sqrt{T}$.

The implementation of the unit root tests also requires the limit distributions of the BSADF and GSADF test statistics. These distributions are non-standard and depend on the minimum window size. Hence, critical values have to be obtained through Monte Carlo simulations. We obtain finite sample critical values for the GSADF and BSADF statistics by generating 2,000 random walk processes with $N(0,1)$ errors.

Appendix B: The IVX Testing Procedure

Consider the following bivariate system[16]

$$y_{t+1} = \beta x_t + u_{1,t+1}, \tag{22}$$
$$x_{t+1} = \rho x_t + u_{2,t+1}, \tag{23}$$

where the innovations $(u_{1,t+1}, u_{2,t+1})'$ follow a martingale difference sequence, and $\rho = 1 + c/T^\alpha$ for some $\alpha \geq 0$. In this setting, the AR coefficient for the regressor is not restricted to the vicinity of unity but is allowed to take a much wider range of values. Depending on the value of c and α, the regressor can belong to one of the following categories:

   i) integrated ($c = 0$ or $\alpha > 1$)

  ii) local-to-unity ($c \neq 0$ and $\alpha = 1$)

 iii) near stationary ($c < 0$ and $\alpha \in (0,1)$)

 iv) locally explosive ($c > 0$ and $\alpha = 1$)

  v) mildly explosive ($c > 0$ and $\alpha \in (0,1)$)

In the center of the IVX procedure is the creation of an instrument z_t which, although relies solely on the regressor, falls in the near stationary category iii. Specifically, given an artificial autoregressive scalar,[17]

$$\rho_z = 1 + c_z/T^\zeta, \quad \zeta \in (0,1), \quad c_z < 0, \tag{24}$$

the IVX instrument is initialized at zero and sequentially computed for the remaining periods according to

$$z_t = \rho_z z_{t-1} + \Delta x_t. \tag{25}$$

---

[16]For exposition purposes, we focus on the simple bivariate case. The model can easily be extended to include multiple regressors, and an intercept in the predictive regression. The model can also be extended to allow for autocorrelation in $u_{2,t+1}$, in which case Newey-West type estimators must be used for the computation of long-run covariance matrices. For a detailed description of the IVX procedure see Phillips and Magdalinos (2009), Kostakis et al. (2015), and Phillips and Lee (2013).

[17]Monte Carlo simulation results in Kostakis et al. (2015) suggest setting $c_z = -1$ and $\zeta = 0.95$.

It can be shown that the estimator

$$\hat{\beta}_{\text{IVX}} = \frac{\sum z_t y_{t+1}}{\sum z_t x_t},$$ (26)

has the following limit theory

$$T^{\frac{1+\zeta}{2}}(\hat{\beta}_{\text{IVX}} - \beta) \Longrightarrow \psi',$$

where ψ' is a mixed normal variable, and the IVX test statistic

$$\text{IVX} = \frac{\hat{\beta}_{\text{IVX}} - \beta}{\hat{\sigma}_{\text{IVX}}},$$

is standard normal (Kostakis et al., 2015; Phillips and Lee, 2013; Phillips and Magdalinos, 2009).[18]

The above limit theory is based on the assumption of no serial correlation in the error term of the predictive regression. For $n > 1$, the time to maturity exceeds the sampling interval, the data in regression (22) are overlapping, and consequently the error term u_1 exhibits serial correlation (Hodrick and Srivastava, 1986). Phillips and Lee (2013) propose an extension of the IVX procedure to long-horizon regressions which allows valid statistical inference. In particular, the authors show that testing the null hypothesis of no predictability is equivalent to testing the significance of the slope coefficient $\hat{\beta}_n$ in the empirical regression

$$y_{t+n} = \beta_n x_t^n + u_{1,t+n}, \quad x_t^n = \sum_{j=1}^{n} x_{t+j-1},$$ (27)

and propose the following long-horizon IVX estimator[19]

$$\hat{\beta}_{n,\text{IVX}} = \frac{\sum y_{t+n} z_t^n}{\sum x_t^n z_t^n},$$ (28)

whose asymptotic null distribution is again mixed gaussian regardless of the degree of persistence of the regressor.

In order to test for periodically collapsing bubbles, we use the IVX test inside a rolling-window regression framework. We select the size of the rolling window so that, on the one hand, the IVX test exhibits good size properties and, on the other, direct comparisons with the results obtained by the recursive unit root procedure can be made. With respect to the former criterion, we have conducted a battery of simulation experiments which indicate that the IVX test performs well for sample sizes as small as 25 observations. On this basis, we set the window size equal to the maximum of 25 observations and the size suggested by the rule of Phillips et al. (2015a) ($\lfloor r_0 \times T \rfloor$).

---

[18]We are grateful to Michael Stamatogiannis for sharing his Gauss code for implementing the IVX procedure.

[19]In empirical applications, data may only be available for $n > 1$ periods to maturity. In such cases, we use interpolation methods to approximate short-horizon variables.

In our rolling-window regression framework, the null hypothesis of no-bubbles is examined multiple times. As a consequence, the likelihood of false rejection (i.e. the likelihood of Type I error) may exceed the desired significance level. To address the issue of multiple hypotheses testing, we perform a Bonferroni correction, i.e. adjust the nominal significance level by the number of hypotheses. It is a well-known fact that the Bonferroni correction results in conservative tests. Nevertheless, the simulation experiments in Section 3 indicate that the IVX procedure is still as or more powerful than the recursive unit root test.[20] Finally, for dating episodes of exuberance, in the second stage of the rolling Fama regression procedure we compare the sequence of IVX statistics against standard normal critical values.

References

Abreu, D. and Brunnermeier, M. K. (2003), "Bubbles and Crashes," *Econometrica*, 71, 173–204.

Balke, N. S. and Wohar, M. E. (2009), "Market fundamentals versus Rational Bubbles in Stock Prices: A Bayesian Perspective," *Journal of Applied Econometrics*, 24, 35–75.

Bhargava, A. (1986), "On the Theory of Testing for Unit Roots in Observed Time Series," *Review of Economic Studies*, 53, 369–84.

Blanchard, O. J. (1979), "Speculative bubbles, crashes and rational expectations," *Economics Letters*, 3, 387–389.

Blanchard, O. J. and Watson, M. W. (1983), "Bubbles, rational expectations and financial markets," NBER Working Paper 945.

Bresciani-Turroni, C. (1937), *The Economics of Inflation*, London.

Brunnermeier, M. K. (2008), "Bubbles," *The New Palgrave Dictionary of Economics*, 2.

Burmeister, E. and Wall, K. D. (1982), "Kalman filtering estimation of unobserved rational expectations with an application to the German hyperinflation," *Journal of Econometrics*, 20, 255–284.

— (1987), "Unobserved Rational Expectations and the German Hyperinflation with Endogenous Money Supply," *International Economic Review*, 28, 15–32.

Busetti, F. and Taylor, A. M. R. (2004), "Tests of stationarity against a change in persistence," *Journal of Econometrics*, 123, 33–66.

Campbell, J. Y. and Yogo, M. (2006), "Efficient tests of stock return predictability," *Journal of Financial Economics*, 81, 27–60.

---

[20]An alternative approach is to employ a bootstrap technique. Such an approach may be more powerful but is too computationally intensive to permit Monte Carlo evaluations.

Carlson, M. A. (2006), "A brief history of the 1987 stock market crash with a discussion of the Federal Reserve response," Finance and Economics Discussion Series 2007-13, Board of Governors of the Federal Reserve System (U.S.).

Carriero, A. (2006), "Explaining US–UK Interest Rate Differentials: A Reassessment of the Uncovered Interest Rate Parity in a Bayesian Framework," *Oxford Bulletin of Economics and Statistics*, 68, 879–899.

Casella, A. (1989), "Testing for rational bubbles with exogenous or endogenous fundamentals : The German hyperinflation once more," *Journal of Monetary Economics*, 24, 109–122.

Cavanagh, C. L., Elliott, G., and Stock, J. H. (1995), "Inference in Models with Nearly Integrated Regressors," *Econometric Theory*, 11, 1131–1147.

Chernenko, S., Schwarz, K., and Wright, J. H. (2004), "The Information Content of Forward and Futures Prices: Market Expectations and the Price of Risk," *FRB International Finance discussion paper*, 1–27.

Chow, Y., McAleer, M., and Sequeira, J. (2000), "Pricing of Forward and Futures Contracts," *Journal of Economic Surveys*, 14, 215–253.

Christiano, L. J. (1987), "Cagan's Model of Hyperinflation under Rational Expectations," *International Economic Review*, 28, 33–49.

Cochrane, J. H. (1992), "Explaining the variance of price-dividend ratios," *Review of Financial Studies*, 5, 243–280.

Corbae, D., Lim, K.-G., and Ouliaris, S. (1992), "On cointegration and tests of forward market unbiasedness," *The Review of Economics and Statistics*, 728–732.

Cornell, B. and Reinganum, M. R. (1981), "Forward and Futures Prices: Evidence from the Foreign Exchange Markets," *The Journal of Finance*, 36, 1035–1045.

Craine, R. (1993), "Rational bubbles: A test," *Journal of Economic Dynamics and Control*, 17, 829–846.

De Grauwe, P. and Rovira-Kaltwasser, P. (2012), *The Exchange Rate in a Behavioral Finance Framework*, John Wiley & Sons, Inc., pp. 111–132.

Diba, B. T. and Grossman, H. I. (1988), "The Theory of Rational Bubbles in Stock Prices," *Economic Journal*, 98, 746–54.

Driffill, J. and Sola, M. (1998), "Intrinsic bubbles and regime-switching," *Journal of Monetary Economics*, 42, 357–373.

Dumas, B. (1992), "Dynamic Equilibrium and the Real Exchange Rate in Spatially Separated World," *Review of Financial Studies*, 5, 153–180.

Durlauf, S. N. and Hooker, M. A. (1994), "Misspecification Versus Bubbles in the Cagan Hyperinflation Model," in *Non-stationary time series analysis and cointegration*, ed. Hargreaves, C. P., Oxford: Oxford University Press.

Einzig, P. (1937), *The Theory of Forward Exchange*, London: Macmillan.

Elliott, G., Müller, U. K., and Watson, M. W. (2015), "Nearly Optimal Tests when a Nuisance Parameter is Present Under the Null Hypothesis," *Econometrica*, 83, 771–811.

Engel, C. (1996), "The Forward Discount Anomaly and the Risk Rremium: A survey of Recent Evidence," *Journal of Empirical finance*, 3, 123–192.

Engel, C. and Hamilton, J. D. (1990), "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?" *American Economic Review*, 80, 689–713.

Engel, C., Mark, N. C., and West, K. D. (2007), "Exchange rate models are not as bad as you think," Tech. rep., National Bureau of Economic Research.

Engel, C. and West, K. D. (2005), "Exchange Rates and Fundamentals," *Journal of Political Economy*, 113, 485–517.

Engsted, T. (1993), "Cointegration and Cagan's Model of Hyperinflation under Rational Expectations," *Journal of Money, Credit and Banking*, 25, 350–60.

— (1994), "The Classic European Hyperinflations Revisited: Testing the Cagan Model Using a Cointegrated VAR Approach," *Economica*, 61, 331–43.

— (2003), "Misspecification versus bubbles in hyperinflation data: comment," *Journal of International Money and Finance*, 22, 441–451.

Engsted, T. and Nielsen, B. (2012), "Testing for Rational Bubbles in a Coexplosive Vector Autoregression," *Econometrics Journal*, 15, 226–254.

Evans, G. W. (1986), "A Test for Speculative Bubbles in the Sterling-Dollar Exchange Rate: 1981-84," *American Economic Review*, 76, 621–36.

— (1991), "Pitfalls in Testing for Explosive Bubbles in Asset Prices," *American Economic Review*, 81, 922–30.

Fama, E. F. (1984), "Forward and spot exchange rates," *Journal of Monetary Economics*, 14, 319–338.

Feldstein, M. (1994), "American Economic Policy in the 1980s: A Personal View," in *American Economic Policy in the 1980s*, National Bureau of Economic Research, Inc, NBER Chapters, pp. 1–80.

Ferguson, N. (1995), "Keynes and the German inflation," *English Historical Review*, 368–391.

Flood, R. P. and Garber, P. M. (1980), "An Economic Theory of Monetary Reform," *Journal of Political Economy*, 88, 24–58.

— (1983), "Process consistency and monetary reform : Some further evidence," *Journal of Monetary Economics*, 12, 279–295.

Flood, R. P., Hodrick, R. J., and Paul, K. (1994), "An Evaluation on Recent Evidence on Stock Market Bubbles," in *Speculative Bubbles, Speculative Attacks, and Policy Switching*, eds. Flood, R. P. and Garber, P., Cambridge, MA: MIT Press, pp. 105–133.

Froot, K. A. and Obstfeld, M. (1992), "Intrinsic bubbles: The case of stock prices," Tech. rep., National Bureau of Economic Research.

Froot, K. A. and Thaler, R. H. (1990), "Anomalies: foreign exchange," *The Journal of Economic Perspectives*, 179–192.

Funke, M., Hall, S., and Sola, M. (1994), "Rational bubbles during Poland's hyperinflation: Implications and empirical evidence," *European Economic Review*, 38, 1257–1276.

Graham, F. D. (1930), *Exchange, prices, and production in Hyper-inflation: Germany, 1920-1923.*, Princeton, NJ: Princeton University Press.

Gürkaynak, R. S. (2008), "Econometric Tests Of Asset Price Bubbles: Taking Stock," *Journal of Economic Surveys*, 22, 166–186.

Hamilton, J. D. and Whiteman, C. H. (1985), "The observable implications of self-fulfilling expectations," *Journal of Monetary Economics*, 16, 353–373.

Hjalmarsson, E. (2011), "New methods for inference in long-horizon regressions," *Journal of Financial and Quantitative Analysis*, 46, 815–839.

Hodrick, R. J. and Srivastava, S. (1986), "The covariation of risk premiums and expected future spot exchange rates," *Journal of International Money and Finance*, 5, S5–S21.

Homm, U. and Breitung, J. (2012), "Testing for speculative bubbles in stock markets: a comparison of alternative methods," *Journal of Financial Econometrics*, 10, 198–231.

Hooker, M. A. (2000), "Misspecification versus bubbles in hyperinflation data: Monte Carlo and interwar European evidence," *Journal of International Money and Finance*, 19, 583–600.

Hull, J. (2015), *Options, Futures and Other Derivatives*, Pearson/Prentice Hall, 9th ed.

Jansson, M. and Moreira, M. J. (2006), "Optimal Inference in Regression Models with Nearly Integrated Regressors," *Econometrica*, 74, 681–714.

Kasparis, I., Andreou, E., and Phillips, P. C. B. (2015), "Nonparametric predictive regression," *Journal of Econometrics*, 185, 468–494.

Keynes, J. M. (–), "The Present State of the Foreign Exchanges," in *Collected Writings*, vol. xvii, p. 176.

Kim, J.-Y. (2000), "Detection of Change in Persistence of a Linear Time Series," *Journal of Econometrics*, 95, 97–116.

Kostakis, A., Magdalinos, T., and Stamatogiannis, M. P. (2015), "Robust econometric inference for stock return predictability," *Review of Financial Studies*, 28, 1506–1553.

LaHaye, L. (1985), "Inflation and Currency Reform," *Journal of Political Economy*, 93, 537–60.

Lansing, K. J. (2010), "Rational and Near-Rational Bubbles Without Drift," *The Economic Journal*, 120, 1149–1174.

LeRoy, S. F. (2004), "Rational Exuberance," *Journal of Economic literature*, 42, 783–804.

LeRoy, S. F. and Porter, R. D. (1981), "The Present-value Relation: Tests based on Implied Variance Bounds," *Econometrica*, 555–574.

Lettau, M., Ludvigson, S. C., and Wachter, J. A. (2008), "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?" *Review of Financial Studies*, 21, 1653–1687.

Mark, N. C. (1995), "Exchange rates and fundamentals: Evidence on long-horizon predictability," *The American Economic Review*, 201–218.

Maynard, A. (2003), "Testing for Forward-Rate Unbiasedness: On Regression in Levels and in Returns," *The Review of Economics and Statistics*, 85, 313–327.

Meese, R. A. (1986), "Testing for Bubbles in Exchange Markets: A Case of Sparkling Rates?" *The Journal of Political Economy*, 345–373.

Michael, P., Nobay, A. R., and Peel, D. A. (1994), "The German hyperinflation and the demand for money revisited," *International Economic Review*, 1–22.

Mishkin, F. S. and White, E. N. (2002), "US Stock Market Crashes and their Aftermath: Implications for Monetary Policy," Tech. rep., National bureau of economic research.

Nielsen, B. (2008), "On the Explosive Nature of Hyper-Inflation Data," Economics Discussion Papers 2008-9, Kiel Institute for the World Economy.

— (2010), "Analysis of coexplosive processes," *Econometric Theory*, 26, 882–915.

O'Connell, P. G. J. and Wei, S.-J. (2002), "The bigger they are, the harder they fall: Retail price differences across U.S. cities," *Journal of International Economics*, 56, 21–53.

Phillips, P. C. B. (2014), "On Confidence Intervals for Autoregressive Roots and Predictive Regression," *Econometrica*, 82, 1177–1195.

Phillips, P. C. B. and Lee, J. H. (2013), "Predictive regression under various degrees of persistence and robust long-horizon regression," *Journal of Econometrics*, 177, 250–264.

Phillips, P. C. B. and Magdalinos, T. (2008), "Limit Theory For Explosively Cointegrated Systems," *Econometric Theory*, 24, 865–887.

— (2009), "Econometric Inference in the Vicinity of Unity," CoFie Working Paper 7, Singapore Management University.

Phillips, P. C. B., Shi, S., and Yu, J. (2012), "Testing for Multiple Bubbles," Cowles Foundation Discussion Papers 1843, Cowles Foundation for Research in Economics, Yale University.

— (2014), "Specification Sensitivity in Right-Tailed Unit Root Testing for Explosive Behaviour," *Oxford Bulletin of Economics and Statistics*, 76, 315–333.

— (2015a), "Testing for Multiple Bubbles: Historical Episodes of Exuberance and Collapse in the S&P 500," *International Economic Review*, to appear.

— (2015b), "Testing for Multiple Bubbles: Limit Theory of Real Time Detector," *International Economic Review*, to appear.

Phillips, P. C. B., Wu, Y., and Yu, J. (2011), "Explosive Behavior in the 1990's Nasdaq: When did exuberance escalate asset values?" *International Economic Review*, 52, 201–226.

Rubin, H. (1950), "Consistency of Maximum Likelihood Estimates in the Explosive Case," in *Statistical Inference in Dynamic Economic Models*, ed. Koopmans, T., John Wiley & Sons, chap. XIV, pp. 356–364.

Sargent, T. J. (1977), "The Demand for Money During Hyperinflations under Rational Expectations: I," *International Economic Review*, 18, 59–82.

Sarno, L. and Taylor, M. P. (2002), *The Economics of Exchange Rates*, Cambridge England and New York: Cambridge University Press.

Scheinkman, J. A. and Xiong, W. (2003), "Overconfidence and Speculative Bubbles," *Journal of Political Economy*, 111, 1183–1220.

Shiller, R. J. (1981), "Do Stock Prices Move too much to be Justified by Subsequent Changes in Dividends," *American Economic Review*, 71, 421–436.

— (2015), *Irrational exuberance*, Princeton university press.

Stambaugh, R. F. (1999), "Predictive regressions," *Journal of Financial Economics*, 54, 375–421.

Stock, J. H. (1991), "Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series," *Journal of Monetary Economics*, 28, 435–459.

Taylor, M. P. (1987), "Risk premia and foreign exchange: a multiple time series approach to testing uncovered interest-rate parity," *Weltwirtschaftliches archiv*, 123, 579–591.

— (1991), "The Hyperinflation Model of Money Demand Revisited," *Journal of Money, Credit and Banking*, 23, 327–51.

Taylor, M. P. and Peel, D. A. (1998), "Periodically collapsing stock price bubbles: a robust test," *Economics Letters*, 61, 221–228.

Tirole, J. (1985), "Asset Bubbles and Overlapping Generations," *Econometrica: Journal of the Econometric Society*, 1071–1100.

Torous, W., Valkanov, R., and Yan, S. (2004), "On Predicting Stock Returns with Nearly Integrated Explanatory Variables," *The Journal of Business*, 77, 937–966.

van Norden, S. (1996), "Regime switching as a test for exchange rate bubbles," *Journal of Applied Econometrics*, 11, 219–251.

van Norden, S. and Schaller, H. (1993), "The Predictability of Stock Market Regime: Evidence from the Toronto Stock Exchange," *The Review of Economics and Statistics*, 505–510.

van Norden, S. and Vigfusson, R. (1998), "Avoiding the pitfalls: can regime-switching tests reliably detect bubbles?" *Studies in Nonlinear Dynamics & Econometrics*, 3, 1–24.

West, K. D. (1987), "A Specification Test for Speculative Bubbles," *The Quarterly Journal of Economics*, 102, 553–80.

— (1988), "Bubbles, fads and stock price volatility tests: a partial evaluation," *The Journal of Finance*, 43, 639–656.

— (1994), "Comments : Rational bubbles during Poland's hyperinflation: Implications and empirical evidence by M. Funke, S. Hall and M. Sola," *European Economic Review*, 38, 1282–1285.

Wu, Y. (1995), "Are there rational bubbles in foreign exchange markets? Evidence from an alternative test," *Journal of International Money and Finance*, 14, 27–46.

Tables and Figures

Panel A, Periods to Maturity $n = 1$						
		$r = 0.05$			$r = 0.10$	
	plim $\widehat{\beta}_n$	mean $(\widehat{\beta}_n)$	sd$(\widehat{\beta}_n)$	plim $\widehat{\beta}_n$	mean $(\widehat{\beta}_n)$	sd$(\widehat{\beta}_n)$
$\pi = 0.3$	49.00	49.00	6.45e-12	25.67	25.67	1.12e-12
$\pi = 0.5$	21.00	21.00	7.95e-07	11.00	11.00	1.37e-07
$\pi = 0.7$	9.00	9.00	1.91e-03	4.71	4.71	8.75e-04
$\pi = 0.9$	2.33	2.25	1.66e+00	1.22	1.21	3.77e-01
Panel B, Periods to Maturity $n = 4$						
		$r = 0.05$			$r = 0.10$	
	plim $\widehat{\beta}_n$	mean $(\widehat{\beta}_n)$	sd$(\widehat{\beta}_n)$	plim $\widehat{\beta}_n$	mean $(\widehat{\beta}_n)$	sd$(\widehat{\beta}_n)$
$\pi = 0.3$	690.69	690.69	2.13e-11	386.32	386.32	3.87e-12
$\pi = 0.5$	84.60	84.60	8.52e-07	47.32	47.32	1.47e-07
$\pi = 0.7$	17.85	17.85	1.58e-03	9.98	9.98	2.46e-04
$\pi = 0.9$	2.96	2.89	1.29e+00	1.65	1.65	2.40e-01

Table 1: The table reports the theoretical value of the slope coefficient (plim $\widehat{\beta}_n$) in the Fama regression in the presence of an ongoing bubble of the Blanchard (1979) type. The table also reports small sample properties (means and standard deviations, *sd*) of the least squares estimator obtained from Monte Carlo simulations. Details of the simulation exercise are provided in the main text.

Table 2: DGPs with a Single Bubble Episode

Panel A: $\pi = 0.5, r = 0.05$

τ_f	d	GSADF						IVX					
		Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
31	10	0.97	0.97	0.00	0.00	36.8	3.2	0.99	0.95	0.03	0.01	34.9	5.1
	15	1.00	1.00	0.00	0.00	37.0	8.0	1.00	0.97	0.02	0.01	35.0	10.0
	20	1.00	1.00	0.00	0.00	36.9	13.1	1.00	0.98	0.02	0.00	34.9	15.1
51	10	0.97	0.97	0.00	0.00	56.8	3.2	0.99	0.97	0.02	0.01	54.9	5.1
	15	1.00	1.00	0.00	0.00	57.1	7.9	1.00	0.97	0.02	0.01	55.0	10.0
	20	1.00	1.00	0.00	0.00	57.0	13.0	1.00	0.97	0.02	0.00	54.9	15.1
71	10	0.96	0.96	0.00	0.00	76.9	3.1	0.99	0.96	0.02	0.01	74.9	5.1
	15	1.00	1.00	0.00	0.00	77.1	7.9	1.00	0.97	0.03	0.01	75.0	10.0
	20	1.00	1.00	0.00	0.00	77.1	12.9	1.00	0.97	0.02	0.01	75.0	15.0

Continued on Next Page

Table 2 – Continued

		GSADF						IVX					
τ_f	d	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$

Panel B: $\pi = 0.7, r = 0.05$

		GSADF						IVX					
τ_f	d	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
31	10	0.06	0.06	0.00	0.00	39.4	0.6	0.66	0.61	0.04	0.01	37.0	3.0
	15	0.81	0.80	0.01	0.00	42.0	3.0	0.95	0.88	0.07	0.01	38.1	6.9
	20	0.98	0.97	0.01	0.00	42.9	7.1	1.00	0.91	0.07	0.01	38.6	11.4
51	10	0.06	0.06	0.00	0.00	59.6	0.4	0.68	0.63	0.05	0.01	57.1	2.9
	15	0.78	0.78	0.00	0.00	62.2	2.8	0.96	0.88	0.06	0.01	58.4	6.6
	20	0.97	0.97	0.01	0.00	63.1	6.9	1.00	0.93	0.07	0.01	58.6	11.4
71	10	0.05	0.05	0.00	0.00	79.6	0.4	0.66	0.61	0.04	0.01	77.2	2.8
	15	0.79	0.79	0.00	0.00	82.2	2.8	0.95	0.89	0.06	0.01	78.3	6.7
	20	0.97	0.97	0.01	0.00	83.0	7.0	0.99	0.91	0.07	0.02	78.5	11.5

Panel C: $\pi = 0.5, r = 0.1$

		GSADF						IVX					
τ_f	d	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
31	10	0.98	0.98	0.00	0.00	36.5	3.5	0.99	0.97	0.02	0.01	34.7	5.3
	15	1.00	1.00	0.00	0.00	36.6	8.4	1.00	0.98	0.02	0.00	34.8	10.2
	20	1.00	1.00	0.00	0.00	36.6	13.4	1.00	0.98	0.01	0.00	34.8	15.2
51	10	0.97	0.97	0.00	0.00	56.6	3.4	0.99	0.96	0.02	0.01	54.7	5.3
	15	1.00	1.00	0.00	0.00	56.7	8.3	1.00	0.98	0.01	0.00	54.8	10.2
	20	1.00	1.00	0.00	0.00	56.7	13.3	1.00	0.99	0.01	0.00	54.8	15.2
71	10	0.98	0.98	0.00	0.00	76.6	3.4	0.99	0.97	0.02	0.01	74.7	5.3
	15	1.00	1.00	0.00	0.00	76.7	8.3	1.00	0.98	0.01	0.00	74.8	10.2
	20	1.00	1.00	0.00	0.00	76.7	13.3	1.00	0.98	0.02	0.01	74.7	15.3

Panel D: $\pi = 0.7, r = 0.1$

		GSADF						IVX					
τ_f	d	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
31	10	0.20	0.20	0.00	0.00	39.2	0.8	0.78	0.73	0.04	0.00	36.9	3.1
	15	0.90	0.90	0.00	0.00	41.3	3.7	0.98	0.92	0.05	0.01	37.8	7.2
	20	0.99	0.99	0.00	0.00	41.7	8.3	1.00	0.95	0.04	0.01	37.9	12.1
51	10	0.18	0.18	0.00	0.00	59.3	0.7	0.75	0.71	0.04	0.01	57.0	3.0
	15	0.88	0.88	0.00	0.00	61.4	3.6	0.97	0.91	0.05	0.01	57.8	7.2
	20	0.99	0.99	0.00	0.00	61.9	8.1	1.00	0.95	0.05	0.01	58.0	12.0
71	10	0.19	0.19	0.00	0.00	79.4	0.6	0.76	0.71	0.05	0.00	77.0	3.0

Table 2 – Continued

		GSADF						IVX					
τ_f	d	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
	15	0.89	0.89	0.00	0.00	81.4	3.6	0.98	0.93	0.04	0.01	77.7	7.3
	20	0.99	0.99	0.00	0.00	81.9	8.1	1.00	0.94	0.05	0.01	77.9	12.1

Notes: The table reports results for the Phillips et al. (2015a) test on $s_{t+n} - f_{t,n}$ (GSADF), and for the rolling Fama (1984) regression (IVX) in the presence of a single bubble episode. The minimum window size for the Phillips et al. (2015a) test is 19 observations, and the window size for the rolling Fama regression is 25 observations. Rej. denotes the rejection rate, τ_f denotes the time the bubble starts, d denotes the duration of the bubble, 1,2 and 3 denote the percentage of times one, two and more than two bubble episodes are detected, respectively, $\bar{\tau}_f$ denotes the mean estimate of τ_f, and $\bar{d}$ denotes the mean estimated duration conditional on the correct number of bubble episodes detected. Further details for the simulation exercise are provided in the main text.

Table 3: DGPs with Two Bubble Episodes

Panel A: $\pi = 0.5$, $r = 0.05$, $\tau_{f,1} = 31$, $\tau_{f,2} = 71$

		GSADF								IVX							
d_1	d_2	Rej.	1	2	3	$\bar{\tau}_{f,1}$	$\bar{d}_1$	$\bar{\tau}_{f,2}$	$\bar{d}_2$	Rej.	1	2	3	$\bar{\tau}_{f,1}$	$\bar{d}_1$	$\bar{\tau}_{f,2}$	$\bar{d}_2$
10	10	1.00	0.06	0.94	0.00	36.7	3.3	76.9	3.1	1.00	0.01	0.98	0.01	34.9	5.1	74.9	5.1
	15	1.00	0.04	0.96	0.00	36.8	3.2	77.1	7.9	1.00	0.00	0.99	0.01	35.0	5.0	75.0	10.0
	20	1.00	0.04	0.96	0.00	36.8	3.2	77.1	12.9	1.00	0.01	0.98	0.01	34.9	5.1	75.0	15.0
15	10	1.00	0.03	0.97	0.00	36.9	8.1	76.9	3.1	1.00	0.01	0.99	0.01	35.0	10.0	74.9	5.1
	15	1.00	0.00	1.00	0.00	37.0	8.0	77.0	8.0	1.00	0.00	0.99	0.01	35.0	10.0	75.0	10.0
	20	1.00	0.00	1.00	0.00	36.9	8.1	77.1	12.9	1.00	0.00	0.99	0.01	34.9	10.1	75.0	15.0
20	10	1.00	0.04	0.96	0.00	36.9	13.1	76.9	3.1	1.00	0.01	0.99	0.01	34.9	15.0	75.5	4.5
	15	1.00	0.00	1.00	0.00	37.0	13.0	77.1	7.9	1.00	0.00	0.99	0.01	35.0	15.0	75.6	9.4
	20	1.00	0.00	1.00	0.00	36.9	13.1	77.1	12.9	1.00	0.00	0.99	0.01	35.0	15.0	75.6	14.4

Panel B: $\pi = 0.7$, $r = 0.05$

		GSADF								IVX							
d_1	d_2	Rej.	1	2	3	$\bar{\tau}_{f,1}$	$\bar{d}_1$	$\bar{\tau}_{f,2}$	$\bar{d}_2$	Rej.	1	2	3	$\bar{\tau}_{f,1}$	$\bar{d}_1$	$\bar{\tau}_{f,2}$	$\bar{d}_2$
10	10	0.10	0.09	0.01	0.00	39.7	0.3	79.6	0.4	0.88	0.26	0.57	0.06	37.6	2.4	77.4	2.6
	15	0.82	0.71	0.11	0.00	39.7	0.3	82.5	2.5	0.99	0.19	0.70	0.10	37.5	2.5	78.4	6.6
	20	0.98	0.84	0.14	0.00	39.7	0.3	83.3	6.7	1.00	0.19	0.72	0.09	37.5	2.5	78.5	11.5
15	10	0.80	0.71	0.09	0.00	42.1	2.8	79.8	0.2	0.98	0.20	0.71	0.07	38.2	6.6	77.5	2.5
	15	0.96	0.27	0.69	0.01	42.2	2.8	82.3	2.7	1.00	0.04	0.86	0.10	38.4	6.6	78.3	6.7
	20	0.99	0.18	0.80	0.00	42.2	2.8	83.1	6.9	1.00	0.03	0.86	0.11	38.5	6.5	78.6	11.4
20	10	0.98	0.85	0.12	0.00	42.7	7.1	79.7	0.3	1.00	0.19	0.73	0.07	38.6	11.1	77.4	2.6
	15	0.99	0.19	0.80	0.00	43.0	6.9	82.3	2.7	1.00	0.02	0.88	0.10	38.7	11.3	78.3	6.7
	20	1.00	0.05	0.95	0.01	42.9	7.1	83.1	6.9	1.00	0.01	0.89	0.10	38.6	11.4	78.6	11.4

Continued on Next Page

Table 3 – Continued

		GSADF								IVX							
d_1	d_2	Rej.	1	2	3	$\bar{\tau}_{f,1}$	$\bar{d}_1$	$\bar{\tau}_{f,2}$	$\bar{d}_2$	Rej.	1	2	3	$\bar{\tau}_{f,1}$	$\bar{d}_1$	$\bar{\tau}_{f,2}$	$\bar{d}_2$
Panel C: $\pi = 0.5, r = 0.1$																
10	10	1.00	0.04	0.96	0.00	36.5	3.5	76.6	3.4	1.00	0.01	0.98	0.01	34.8	5.2	74.8	5.2
10	15	1.00	0.02	0.98	0.00	36.5	3.5	76.7	8.3	1.00	0.01	0.98	0.01	34.8	5.2	74.8	10.2
10	20	1.00	0.02	0.98	0.00	36.5	3.5	76.7	13.3	1.00	0.01	0.98	0.01	34.8	5.2	74.8	15.2
15	10	1.00	0.02	0.98	0.00	36.6	8.4	76.6	3.4	1.00	0.01	0.99	0.01	34.8	10.2	74.7	5.3
15	15	1.00	0.00	1.00	0.00	36.6	8.4	76.7	8.3	1.00	0.00	0.99	0.01	34.8	10.2	74.8	10.2
15	20	1.00	0.00	1.00	0.00	36.6	8.4	76.7	13.3	1.00	0.00	0.99	0.01	34.8	10.2	74.8	15.2
20	10	1.00	0.02	0.98	0.00	36.6	13.4	76.6	3.4	1.00	0.00	0.99	0.01	34.8	15.2	75.4	4.6
20	15	1.00	0.00	1.00	0.00	36.6	13.4	76.7	8.3	1.00	0.00	0.99	0.01	34.7	15.3	75.5	9.5
20	20	1.00	0.00	1.00	0.01	36.6	13.4	76.7	13.3	1.00	0.00	1.00	0.00	34.8	15.2	75.5	14.5
Panel D: $\pi = 0.7, r = 0.1$																
10	10	0.37	0.28	0.08	0.00	39.4	0.5	79.5	0.5	0.94	0.20	0.67	0.06	37.1	2.8	77.1	2.9
10	15	0.91	0.62	0.29	0.00	39.5	0.5	81.6	3.4	0.99	0.14	0.78	0.08	37.2	2.8	77.9	7.1
10	20	0.99	0.67	0.32	0.00	39.5	0.5	82.0	8.0	1.00	0.15	0.78	0.07	37.2	2.8	78.1	11.9
15	10	0.91	0.65	0.26	0.00	41.5	3.5	79.5	0.5	0.99	0.14	0.78	0.07	37.8	7.1	77.2	2.8
15	15	0.99	0.17	0.82	0.00	41.3	3.7	81.5	3.5	1.00	0.03	0.90	0.08	37.9	7.1	77.9	7.1
15	20	1.00	0.08	0.91	0.01	41.4	3.6	82.0	8.0	1.00	0.01	0.91	0.08	37.8	7.2	77.9	12.1
20	10	0.99	0.71	0.28	0.00	41.7	8.2	79.6	0.4	1.00	0.15	0.80	0.06	38.0	11.9	77.3	2.7

Table 3 – Continued

		GSADF								IVX							
d_1	d_2	Rej.	1	2	3	$\bar{\tau}_{f,1}$	$\bar{d}_1$	$\bar{\tau}_{f,2}$	$\bar{d}_2$	Rej.	1	2	3	$\bar{\tau}_{f,1}$	$\bar{d}_1$	$\bar{\tau}_{f,2}$	$\bar{d}_2$
	15	1.00	0.10	0.89	0.01	41.9	8.1	81.6	3.4	1.00	0.01	0.91	0.07	37.9	12.0	77.9	7.1
	20	1.00	0.01	0.98	0.00	41.9	8.1	82.0	8.0	1.00	0.00	0.92	0.08	38.0	12.0	78.1	11.9

Notes: The table reports results for the Phillips et al. (2015a) test on $s_{t+n} - f_{t,n}$ (GSADF) and for the rolling Fama (1984) regressions (IVX) in the presence of two bubble episodes that start at times $\tau_{f,1} = 31$ and $\tau_{f,2} = 71$. The minimum window size for the Phillips et al. (2015a) test is 19 observations, and the window size for the rolling Fama regression is 25 observations. Rej. denotes the rejection rate, d_1 and d_2 denote the duration of the first and of the second bubble, respectively, 1,2 and 3 denote the percentage of times one, two and more than two bubble episodes are detected, respectively, $\bar{\tau}_{f,1}$ and $\bar{\tau}_{f,2}$ denote the mean estimate of $\tau_{f,1}$ and $\tau_{f,2}$, and $\bar{d}_1$ and $\bar{d}_2$ denote the mean estimated duration conditional on the correct number of bubble episodes detected. Further details for the simulation exercise are provided in the main text.

Table 4: Different Periods to Maturity, n

Panel A: $\pi = 0.5, r = 0.05, \tau_f = 31, d = 10$

| | | | GSADF | | | | | | IVX | | |
n	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
2	0.98	0.98	0.00	0.00	36.2	3.8	0.99	0.95	0.04	0.01	34.9	5.1
3	0.98	0.98	0.00	0.00	36.1	3.9	0.99	0.92	0.06	0.02	34.9	5.1
4	0.98	0.97	0.01	0.00	36.0	4.0	0.99	0.90	0.07	0.02	34.9	5.1

Panel B: $\pi = 0.7, r = 0.05$

| | | | GSADF | | | | | | IVX | | |
n	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
2	0.22	0.22	0.00	0.00	38.9	1.1	0.65	0.59	0.05	0.01	37.1	2.9
3	0.33	0.33	0.00	0.00	38.6	1.4	0.65	0.57	0.06	0.01	37.0	3.0
4	0.41	0.40	0.00	0.00	38.5	1.5	0.65	0.56	0.07	0.02	37.0	3.0

Panel C: $\pi = 0.5, r = 0.1$

| | | | GSADF | | | | | | IVX | | |
n	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
2	0.98	0.98	0.00	0.00	36.0	4.0	0.99	0.95	0.03	0.01	34.7	5.3
3	0.98	0.98	0.00	0.00	35.8	4.2	0.99	0.92	0.05	0.02	34.7	5.3
4	0.98	0.98	0.01	0.00	35.7	4.3	0.99	0.91	0.06	0.02	34.7	5.3

Panel D: $\pi = 0.7, r = 0.1$

| | | | GSADF | | | | | | IVX | | |
n	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$	Rej.	1	2	3	$\bar{\tau}_f$	$\bar{d}$
2	0.44	0.44	0.00	0.00	38.7	1.3	0.76	0.70	0.05	0.01	36.9	3.1
3	0.54	0.53	0.00	0.00	38.4	1.6	0.76	0.68	0.07	0.02	36.9	3.1
4	0.58	0.58	0.01	0.00	38.2	1.8	0.76	0.67	0.08	0.02	36.8	3.2

Notes: The table reports results for the Phillips et al. (2015a) test on $s_{t+n} - f_{t,n}$ (GSADF) and for the rolling Fama (1984) regression (IVX) in the presence of a single bubble episode that starts at time $\tau_f = 31$ and lasts for $d = 10$ periods. The minimum window size for the Phillips et al. (2015a) test is 19 observations, and the window size for the rolling Fama regression is 25 observations. Rej. denotes the rejection rate, 1,2 and 3 denote the percentage of times one, two and more than two bubble episodes are detected, respectively, n denotes the periods to maturity, $\bar{\tau}_f$ denotes the mean estimate of τ_f, and $\bar{d}$ denotes the mean estimated duration conditional on the correct number of bubble episodes detected. Further details for the simulation exercise are provided in the main text.

Table 5: Periodically Collapsing Bubbles

Blanchard's DGP

		GSADF				IVX			
r	n	$\pi = 0.3$	$\pi = 0.5$	$\pi = 0.7$	$\pi = 0.9$	$\pi = 0.3$	$\pi = 0.5$	$\pi = 0.7$	$\pi = 0.9$
0.05	1	0.95	0.49	0.20	0.10	0.98	0.78	0.51	0.46
	2	0.96	0.61	0.30	0.17	0.97	0.75	0.48	0.37
	3	0.96	0.66	0.33	0.21	0.97	0.74	0.48	0.37
	4	0.97	0.66	0.37	0.28	0.98	0.74	0.48	0.38
0.10	1	0.96	0.58	0.28	0.30	0.98	0.82	0.57	0.62
	2	0.97	0.66	0.40	0.42	0.97	0.77	0.57	0.60
	3	0.98	0.72	0.42	0.47	0.98	0.78	0.55	0.59
	4	0.97	0.70	0.44	0.51	0.96	0.77	0.53	0.59

Evans' DGP

		GSADF				IVX			
r	n	$\pi = 0.3$	$\pi = 0.5$	$\pi = 0.7$	$\pi = 0.9$	$\pi = 0.3$	$\pi = 0.5$	$\pi = 0.7$	$\pi = 0.9$
0.05	1	1.00	0.99	0.83	0.16	1.00	0.99	0.94	0.86
	2	1.00	0.99	0.88	0.49	1.00	0.99	0.92	0.81
	3	0.99	0.97	0.86	0.69	1.00	0.99	0.90	0.81
	4	0.99	0.95	0.80	0.73	1.00	0.99	0.88	0.76
0.10	1	1.00	0.97	0.60	0.46	1.00	0.98	0.86	0.82
	2	1.00	0.98	0.72	0.69	1.00	0.98	0.82	0.78
	3	1.00	0.96	0.75	0.75	1.00	0.98	0.83	0.78
	4	1.00	0.94	0.71	0.74	1.00	0.98	0.81	0.79

Notes: The table reports the empirical power of the Phillips et al. (2015a) test on $s_{t+n} - f_{t,n}$ (GSADF), and the empirical power of the rolling Fama (1984) regression (IVX) in the presence of periodically collapsing bubbles of the Blanchard (1979) type (Panel A) and of the Evans (1991) type (Panel B). The minimum window size for the Phillips et al. (2015a) test is 19 observations, and the window size for the rolling Fama regression is 25 observations. Further details for the simulation exercise are provided in the main text.

	Unit Root Tests			
	s_t	$f_{t,4}$	$s_{t+4} - f_{t,4}$	95% critical value
ADF	2.949	2.665	-0.244	-0.055
GSADF	5.504	4.641	2.150	1.941

Table 6: The table reports ADF and GSADF test statistics for the German mark-US dollar spot exchange rate, s_t, the four-week forward rate, $f_{t,4}$, and the difference between the future spot and forward rates, $s_{t+4} - f_{t,4}$. The finite-sample 95 percent critical value for the GSADF test is obtained using 2,000 simulations. The minimum window size for the GSADF test is 18 weeks.

Panel A: Unit Root Tests					
	s_t	$s_{t+1} - f_{t,1}$	$s_{t+3} - f_{t,3}$	$s_{t+6} - f_{t,6}$	Critical Value
ADF	-2.916	-13.388	-5.407	-4.040	-0.055
GSADF	1.793	-2.440	0.591	0.837	2.175

Panel B: Rolling Fama Regressions				
Periods to Maturity (n)	1 month	3 months	6 months	Critical Value
Maximum IVX statistic	3.310	2.789	1.765	3.64

Table 7: The table reports ADF and GSADF test statistics for the British pound-US dollar spot exchange rate, s_t, and for the difference between the future spot and forward rates $s_{t+n} - f_{t,n}$ with $n = 1, 3$ and 6. The table also reports results for rolling Fama regressions. The finite-sample 95 percent critical value for the GSADF test is obtained using 2,000 simulations. The minimum window size for the GSADF test and the window size for the rolling Fama regression are set equal to 41 observations.

Panel A: Unit Root Tests						
	s_t^*	$s_t - d_t$	$s_{t+3} - f_{t,3}$	$s_{t+6} - f_{t,6}$	$s_{t+9} - f_{t,9}$	Critical Value
ADF	-1.468	-1.971	-5.898	-4.561	-3.470	-0.055
GSADF	3.054	3.378	1.759	2.482	2.902	2.169

Panel B: Rolling Fama Regressions				
Periods to Maturity (n)	3 months	6 months	9 months	Critical Value
Maximum IVX statistic	1.448	2.462	1.874	3.62

Table 8: The table reports ADF and GSADF test statistics for the logarithm of the real S&P 500 index, s_t^*, the logarithm of the price-to-dividend ratio, $s_t - d_t$, and for the difference between the logarithm of the future spot price and the logarithm of the futures price, $s_{t+n} - f_{t,n}$ with $n = 3, 6$, and 9. The table also reports results for rolling Fama regressions. The finite-sample 95 percent critical value for the GSADF test is obtained using 2,000 simulations. The minimum window size for the GSADF test and the window size for the rolling Fama regression are set equal to 40 observations.
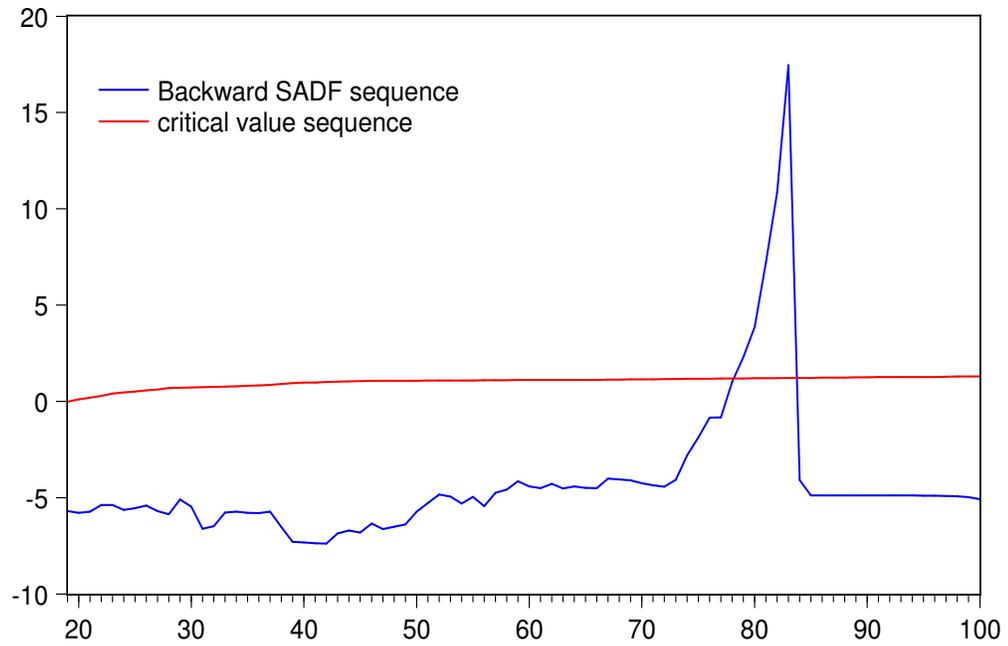
Figure 1: Simulation results for the Phillips et al. (2015a) test. The figure displays the sequence of BSADF statistics (blue line). The red line represents 95 percent critical values obtained from 2,000 Monte Carlo simulations. The minimum window size is 19 observations.
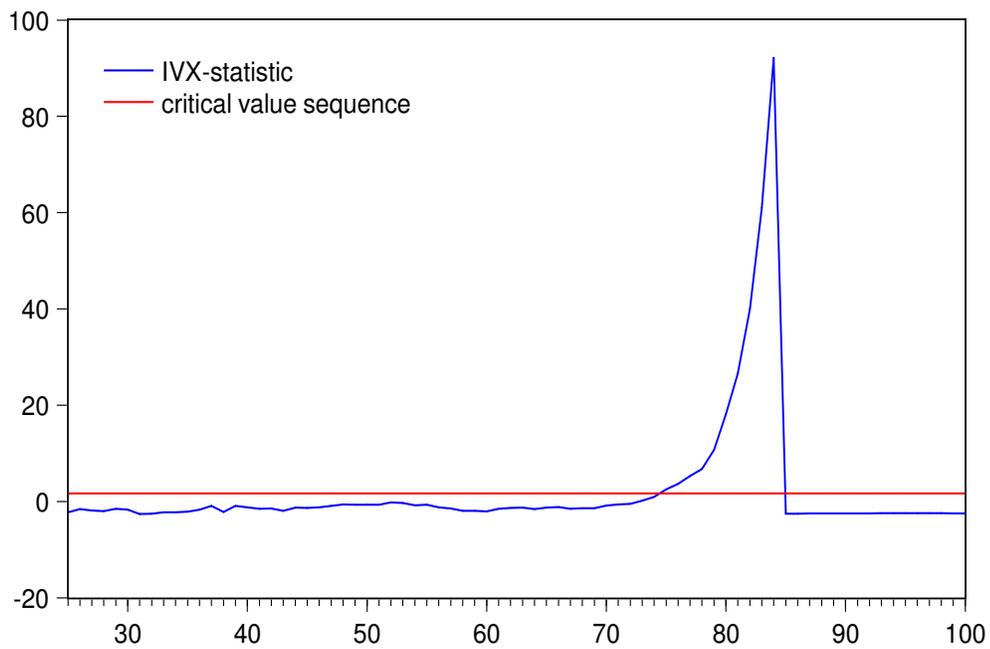
Figure 2: Simulation results for the rolling Fama (1984) regression. The figure displays the sequence of IVX statistics (blue line) corresponding to the null hypothesis that the slope coefficient in regression (10) is zero (no bubble) based on subsamples of 25 observations. The red line represents 95 percent critical values.
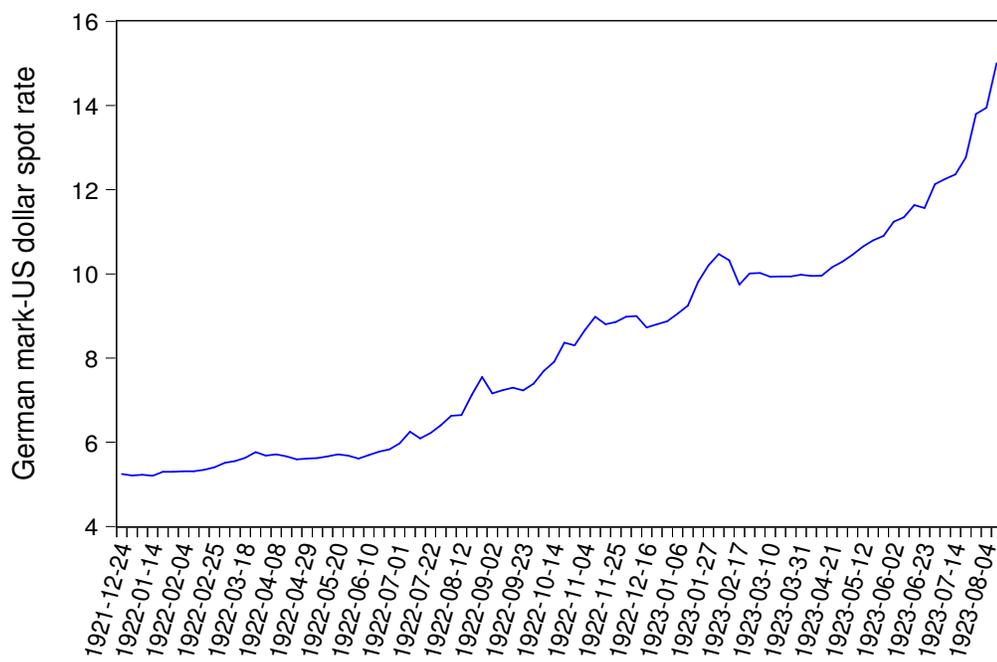
Figure 3: The figure displays the logarithm of the German mark-US dollar spot exchange rate from December 24th, 1921 to August 11th, 1923.
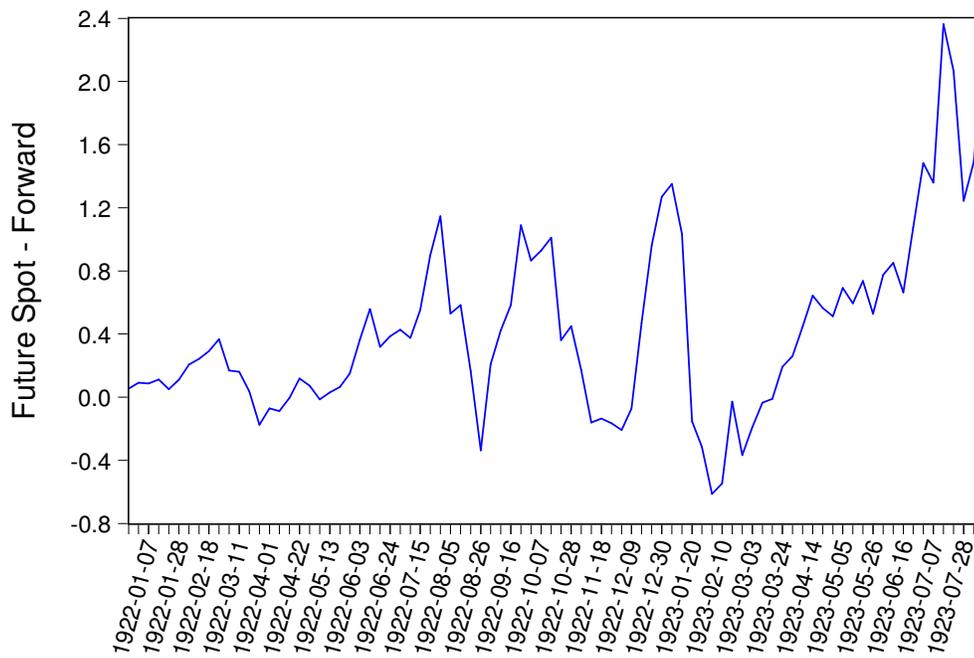
Figure 4: The figure displays the difference between the logarithm of the future German mark-US dollar spot rate and the logarithm of the forward rate, $s_{t+4} - f_{t,4}$, from December 24th, 1921 to August 11th, 1923.
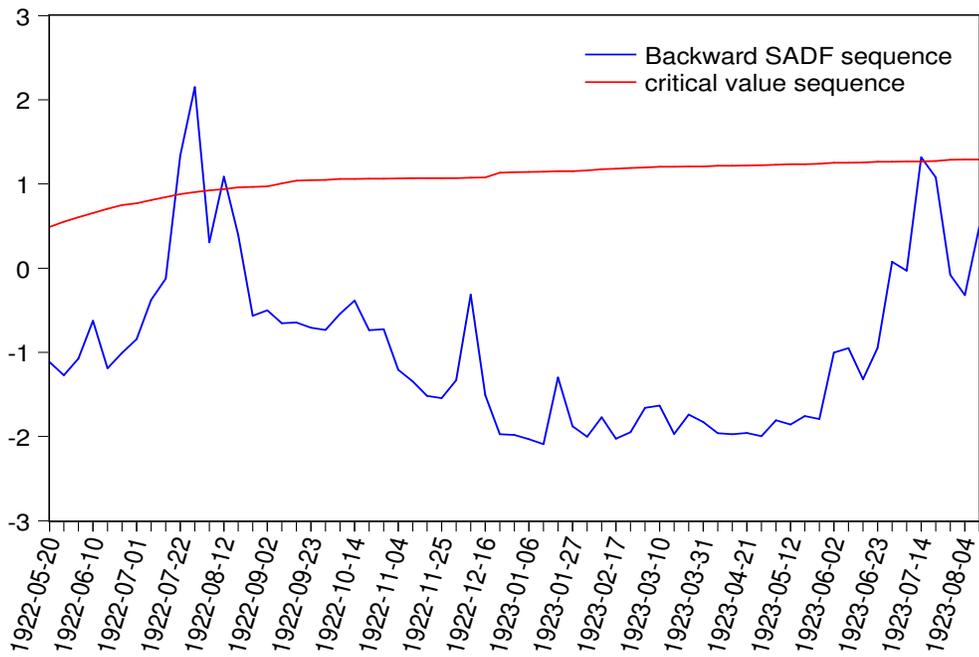
Figure 5: The figure displays the sequence of BSADF statistics for the future spot minus the forward rate, $s_{t+4} - f_{t,4}$, (blue line) together with the 95 percent critical value sequence (red line) obtained from 2,000 Monte Carlo simulations. The minimum window size is 18 weeks.
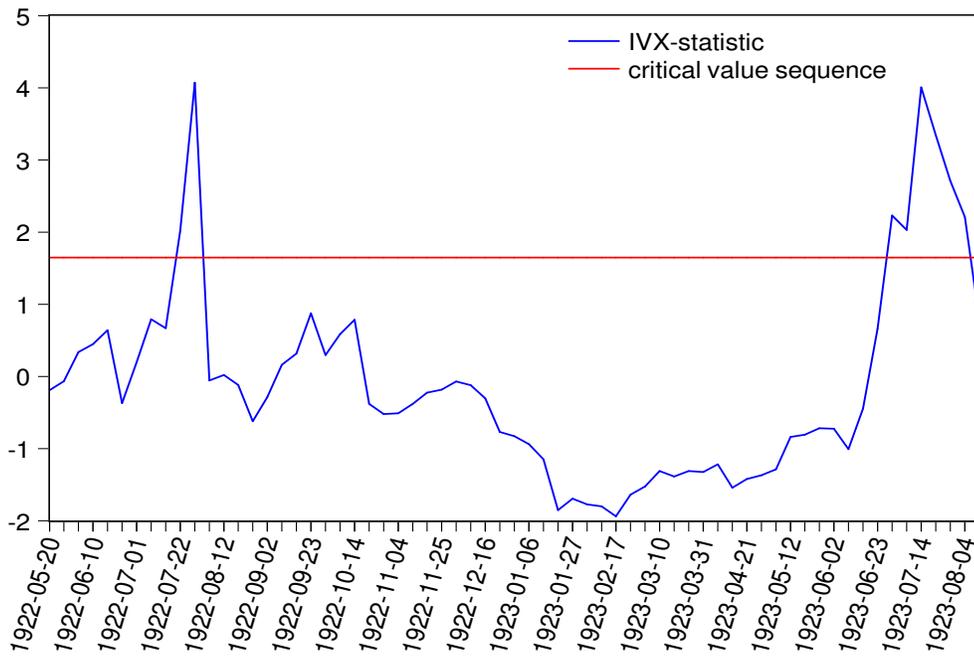
Figure 6: The figure displays the sequence of IVX statistics (blue line) corresponding to the null hypothesis that the slope coefficient in the rolling Fama regression (10) is zero (no bubble) based on subsamples of 25 observations. The red line represents 95 percent critical values.
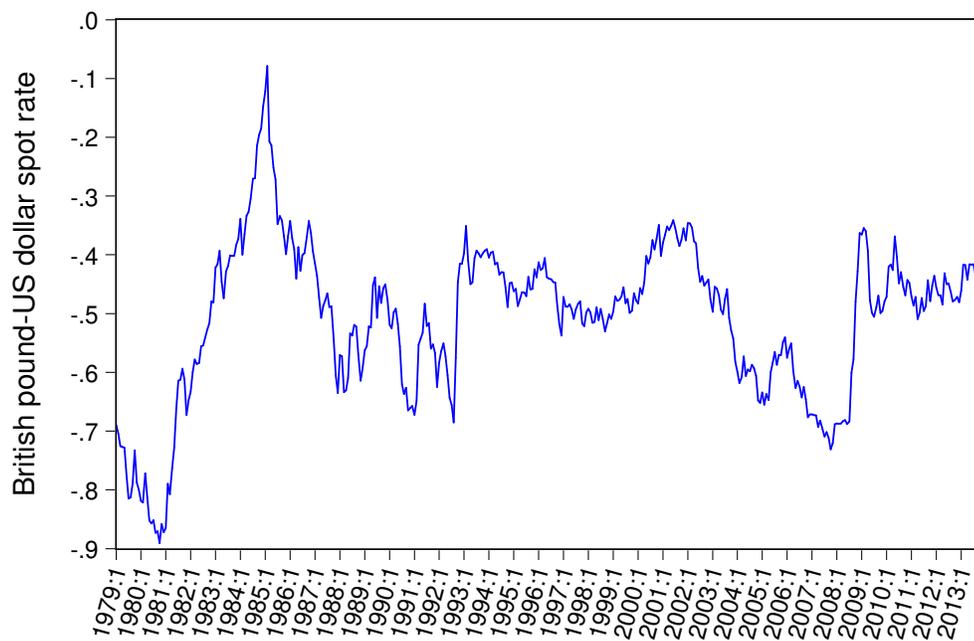
Figure 7: The figure displays the logarithm of the British pound-US dollar spot exchange rate from January 1979 to December 2013.
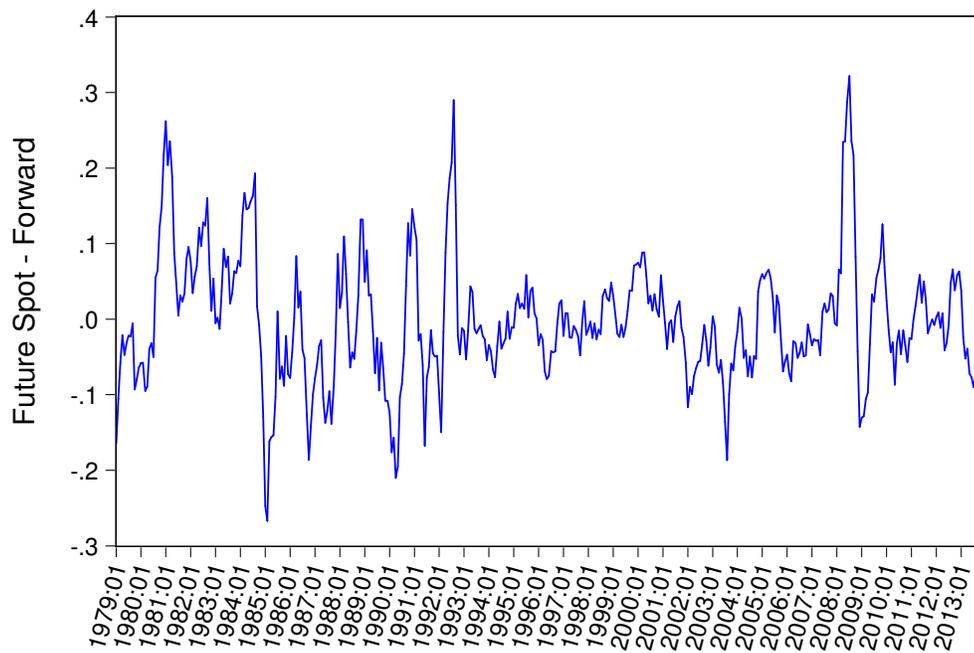
Figure 8: The figure displays the difference between the logarithm of the future British pound-US dollar spot rate and the logarithm of the forward rate, $s_{t+6} - f_{t,6}$, from January 1979 to December 2013.
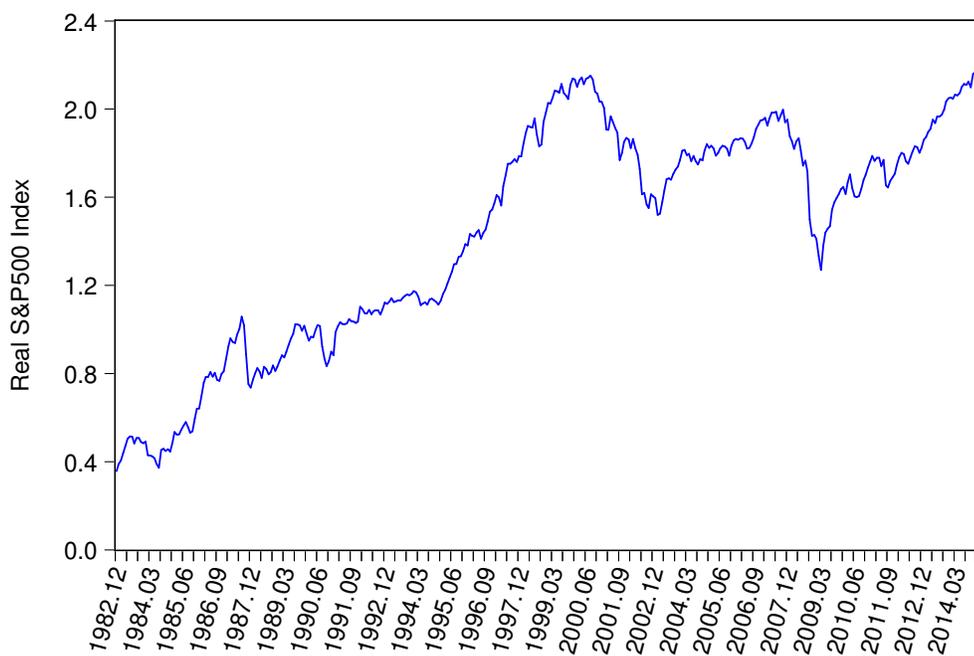
Figure 9: The figure displays the logarithm of the real S&P 500 index from December 1982 to March 2015.
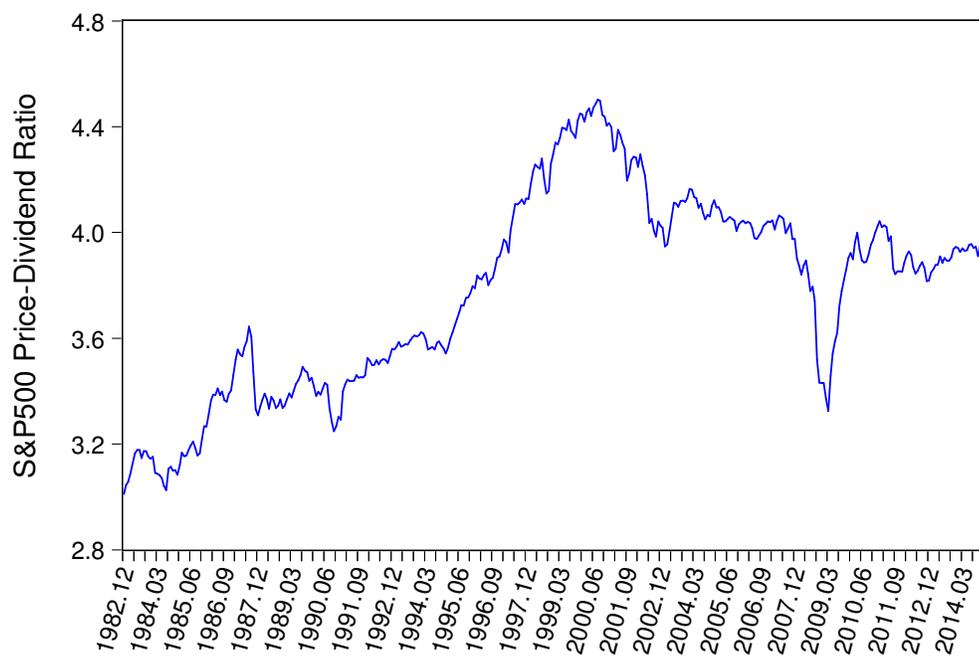
Figure 10: The figure displays the logarithm of the S&P 500 price-to-dividend ratio from December 1982 to March 2015.
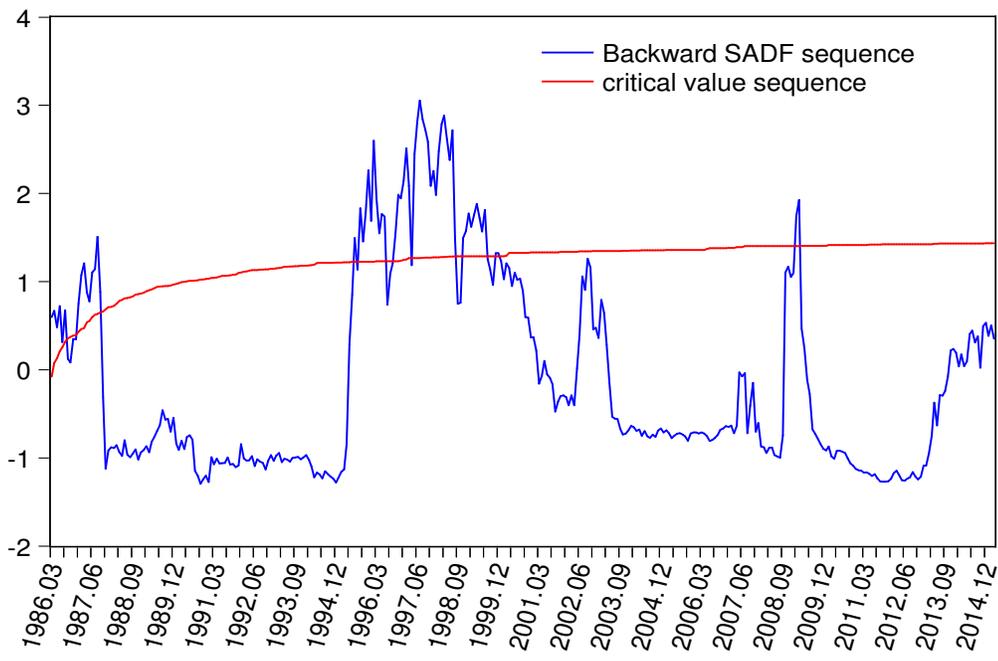
Figure 11: The figure displays the sequence of BSADF statistics for the real S&P 500 index (blue line) together with the 95 percent critical value sequence (red line) obtained from 2,000 Monte Carlo simulations. The minimum window size is 40 months.
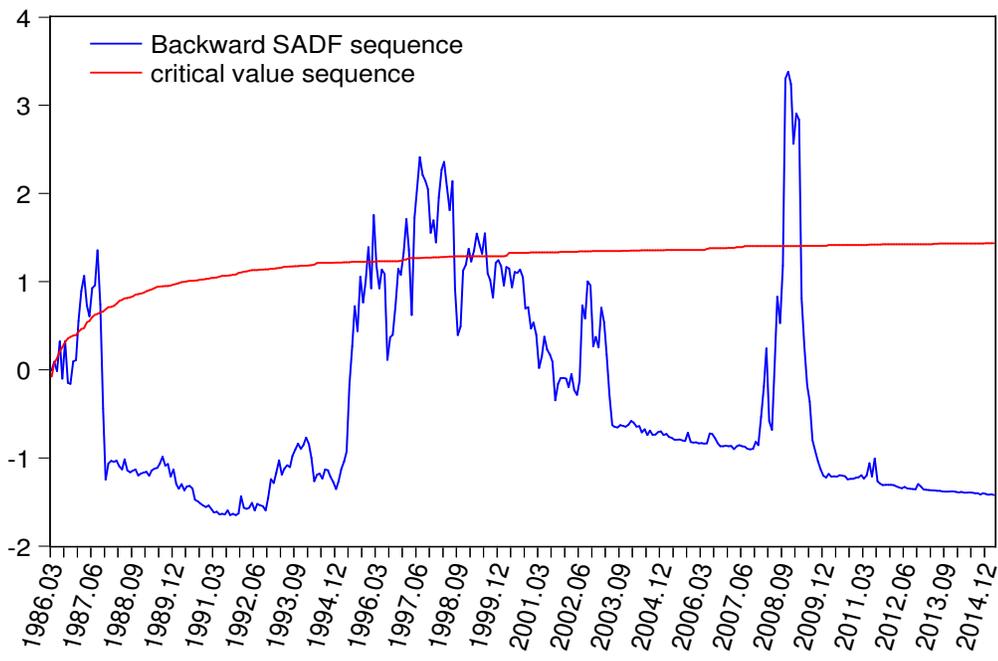
Figure 12: The figure displays the sequence of BSADF statistics for the S&P 500 price-to-dividend ratio (blue line) together with the 95 percent critical value sequence (red line) obtained from 2,000 Monte Carlo simulations. The minimum window size is 40 months.
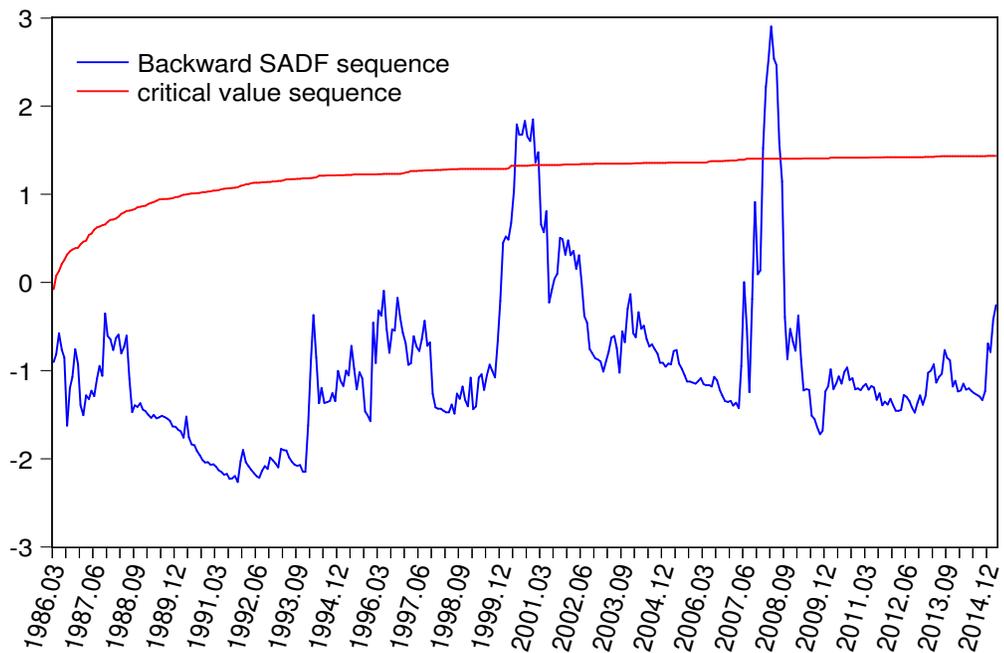
Figure 13: The figure displays the sequence of BSADF statistics for the S&P 500 future spot price minus the futures price, $s_{t+9} - f_{t,9}$, (blue line) together with the 95 percent critical value sequence (red line) obtained from 2,000 Monte Carlo simulations. The minimum window size is 40 months.