Detecting and Forecasting Large Deviations and Bubbles in a Near-Explosive Random Coefficient Model

Anurag Narayan Banerjee, Guillaume Chevillon, Marie Kratz

Detecting and Forecasting Large Deviations and Bubbles in a Near-Explosive Random Coefficient Model

Anurag Banerjee
Guillaume Chevillon
Marie Kratz

ESSEC
BUSINESS SCHOOL

Detecting and Forecasting Large Deviations and Bubbles in a Near-Explosive Random Coefficient Model.*

Anurag Banerjee
Durham Business School

Guillaume Chevillon
ESSEC Business School, CREAR

Marie Kratz
ESSEC Business School, CREAR

September 23, 2013

Abstract

This paper proposes a Near Explosive Random-Coefficient autoregressive model for asset pricing which accommodates both the fundamental asset value and the recurrent presence of autonomous deviations or bubbles. Such a process can be stationary with or without fat tails, unit-root nonstationary or exhibit temporary exponential growth. We develop the asymptotic theory to analyze ordinary least-squares (OLS) estimation. One important theoretical observation is that the estimator distribution in the random coefficient model is qualitatively different from its distribution in the equivalent fixed coefficient model. We conduct recursive and full-sample inference by inverting the asymptotic distribution of the OLS test statistic, a common procedure in the presence of localizing parameters. This methodology allows to detect the presence of bubbles and establish probability statements on their apparition and devolution. We apply our methods to the study of the dynamics of the Case-Shiller index of U.S. house prices. Focusing in particular on the change in the price level, we provide an early detection device for turning points of booms and bust of the housing market.

Keywords: Bubbles, Random Coefficient Autoregressive Model, Local asymptotics, Asset Prices.

JEL Codes: C22, C53, C58, G12.

1

1 Introduction and motivations

The aim of this paper is to propose a random-coefficient autoregressive model that accommodates the pricing of assets both when these follow *fundamentals* and in the presence of *bubbles*. The rationale behind our modeling choice comes from standard present value models (see e.g. Campbell and Shiller, 1987a,b) where the price P_t of a unique asset at time t (or possibly its logarithm) depends on the expected value of future associated cash flows, D_{t+1}, discounted using a time varying pricing kernel M_{t+1} as in $P_t = \mathsf{E}_t \left(M_{t+1} \left(P_{t+1} + D_{t+1} \right) \right)$. The price can be written $P_t = F_t + B_t$ where the so called fundamental price F_t is equal to the expected stream of discounted future cash flows and B_t denotes any process that satisfies $B_t = \mathsf{E}_t \left(M_{t+1} B_{t+1} \right)$. There exist solutions to this equation for which B_t exhibits exponential growth and can be labeled as a bubble, see inter alia Blanchard and Watson (1982), Abreu and Brunnermeier (2003), and Lee and Phillips (2011). Under the assumption that D_t is integrated of order 1, West (1987), and Diba and Grossman (1988) show that F_t is also integrated of the same order. Hence, unit root (or cointegration) tests have been used for testing that a function, say y_t, of asset prices does not exhibit a bubble. Different approaches have been proposed in a stream of papers by Peter Phillips, Jun Yu and several coauthors (see inter alia Phillips, Wu and Yu, 2011, and Phillips and Yu, 2009; respectively PWY and PY henceforth) where they perform recursive Dickey-Fuller tests.[1] To increase power, these authors adapt the critical values to the sample size, with the help of the distributions derived by Phillips and Magdalinos (2007, PM henceforth) under the alternative of a locally explosive root, for $t = 1, ..., T$:

$$y_t = \rho_0 y_{t-1} + \eta_t, \tag{1}$$
$$\rho_0 = \exp \left\{ \frac{\phi_0}{T^\alpha} \right\},$$

with $\phi_0 > 0$, $\alpha \in (0, 1)$, and where η_t is weakly dependent with mean zero. When η_t is independently and identically distributed (*i.i.d.*) with mean zero, we refer to model (1) as a Near Explosive Autoregressive process of order 1 (NEAR(1)) since $\rho_0 > 1$ but $\rho_0 \to 1$ as the sample size T increases. The NEAR(1) model requires $\alpha < 1$ for local explosiveness whereas $\alpha \geq 1$ in expression (1) implies y_t behaves as a near unit root process.

Unfortunately, the alternative (1) does not allow for the bursting of the bubble and the collapse of asset prices as pointed out by Diba and Grossman (1988) and Evans (1991). For this reason, several authors (such as Evans and PWY) have considered the possibility of regime switching or the deterministic collapse of bubbles. For instance, in the Phillips-Yu approach the estimation of the inception and termination of bubbles relies on the assumption that the process experiences deterministic breaks (or unmodelled regime shifts). To render the appearance and disappearance of bubbles stochastic, and avoid specifying their frequency, we generalize their approach to a Near

---

[1]The literature has also provided several other techniques to test for the presence of a bubble, see Gürkaynak (2008) for an overview.

Explosive Random Coefficient autoregressive process, a NERC(1) defined, for $t = 1, ..., T$ as:

$$y_t = \rho_t y_{t-1} + \eta_t, \tag{2}$$

where ρ_t is $i.i.d.$ such that its expectation $\mathsf{E}(\rho_t) = 1 + O\left(T^{-\alpha}\right)$ and its variance $\mathsf{V}(\rho_t) = O\left(T^{-\alpha}\right)$. We are specifically interested in the case where $\mathsf{E}(\rho_t)$ lies on the explosive side of unity. We parameterize:

$$\rho_t = \exp\left\{\frac{\phi + \lambda T^{\alpha/2} u_t}{T^\alpha}\right\},$$

with $(\phi, \lambda, \alpha) \in \mathbb{R} \times \mathbb{R}_+ \times (0, 1)$ and where u_t is $i.i.d$ with zero mean and unit variance. The model we consider is a local-asymptotic approximation to the random coefficient autoregressive model of Nicholls and Quinn (1982) and Granger and Swanson (1997); it nests the NEAR(1). Specifying that the autoregressive coefficient is stochastic, we can draw inference on the whole sample and there is no need to resort to rolling or recursive windows to test the presence of a bubble and estimate its magnitude; the absence of deterministic breaks avoids the usual trimming of observations at the beginning or end of the sample. To illustrate the idea, Figure 1 compares a random draw from the two processes (1) and (2) such that $\mathsf{E}(\rho_t) = \rho_0$ with common η_t. The figure illustrates the point that inception and collapse of bubbles are possible to model without resorting to deterministic breaks. In addition, we show in our empirical application that in the NERC(1), the emergence of the bubble relates to the value taken by the stochastic discount factor, so the model helps improving the structural interpretation of exuberant periods. Also, by a careful choice of y_t in our empirical application, we avoid the issue of negative bubbles pointed out by Diba and Grossman (1988).

The paper develops the asymptotic theory of the NERC(1) and derives the distribution of the ordinary least squares (OLS) estimator $\widehat{\rho}$ required in the Phillips-Yu methodology. Although the NERC(1) model parametrically nests the NEAR(1), its properties differ when $\mathsf{V}(\rho_t) \neq 0$. In particular, its asymptotics depend on the value of $c = \phi + \lambda^2$. When the NERC(1) process is weakly stationary ($c < 0$) the OLS estimate of ρ_t converges to a normal distribution, as under the NEAR(1) with $\phi_0 = c$, albeit with a larger variance. This is not surprising since random coefficient models usually exhibit larger variances than fixed coefficient models. More relevant and interesting, when the NERC(1) model is not weakly stationary ($c \geq 0$) the asymptotics is qualitatively different from the NEAR(1) in the sense that when $\lambda \to 0$, the asymptotic distribution of $\widehat{\rho}$ is not close to that of the NEAR(1) described by PM. When $\lambda \neq 0$, the NERC can generate processes that are stationary with fat tails or nonstationary with occasional explosive growth: bubbles in y_t (however defined) will eventually burst (as seen in Figure 1) and consecutive bubbles are also possible. Our model also provides an analytically tractable explanation for the simulation evidence of Evans (1991): he showed, although in a different setting, that tests for the presence of a bubble have low power when multiple bubbles are present.

Our choice of local asymptotic parameterization renders $\widehat{\rho}$ consistent (contrary to fixed parameter asymptotics, see Hwang and Basawa, 2005). Yet when $c \geq 0$, and as it is often the case under
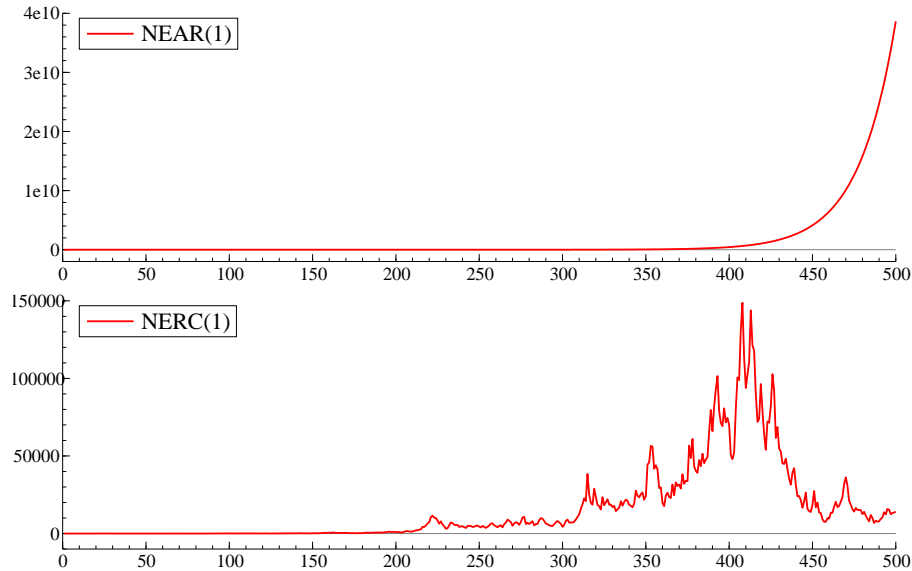
Figure 1: Panels (a) and (b): Simulated paths of the NEAR(1) and NERC(1) processes. The NEAR(1) process is defined as $y_t = \rho_0 y_{t-1} + \eta_t$ with $\rho_0 = \exp\left(\phi_0/T^\alpha\right)$ and the NERC(1) as $y_t = \rho_t y_{t-1} + \eta_t$ with $\rho_t = \exp\left(\left(\phi + \lambda T^{\alpha/2} u_t\right)/T^\alpha\right)$, η_t and u_t being independent standard Gaussian white noise. Parameter values are $\alpha = 1/2$ and $(\phi, \lambda) = (.5, .5)$ and $\phi_0 = .625$ so $\mathsf{E}\left(\rho_t\right) = \rho_0 = 1.028$.

local asymptotics, the estimator $\widehat{\rho}$ converges at a rate that does not allow consistent estimation of the parameters (ϕ, λ). We hence provide an inferential approach based on inverting the asymptotic distribution of $\widehat{\rho}$. The technique does not require the existence of consistent estimators with pivotal distributions. It was popularized by Stock (1991) and Andrews (1993) and various forms have widely been used in the near unit root and weak instrument literatures. In addition, the method can be performed in real-time since we need not resort to deterministic breaks. The distinctive asymptotic theory of NERC(1) is useful in that it allows to forecast the evolution (boom) and devolution (bust) of the bubble generation process. We evaluate our methodology empirically using the Case-Shiller index of U.S. house prices: we analyze both the logarithm of the price/rent ratio and the change in the price level. Analysis based on the latter series in particular helps providing an early detection device for turning points.

The structure of the paper is as follows. In section 2, we define the random-coefficient autoregressive process with local-asymptotic parameterization and derive its asymptotic properties. Section 3 presents the method of inference that we propose. Section 4 shows how the model can be used to forecast the probability of booms or busts. A Monte Carlo evaluation of the properties of the inferential methods is presented in section 5. We apply our methodology in section 6 to the inference regarding the dynamic properties of U.S. house prices. Proof are collected together with additional simulations in the appendix. Throughout the paper, $\lfloor \cdot \rfloor$ denotes the integer part.

2 The model and its properties

2.1 The Near-Explosive Random Coefficient autoregressive model.

The model we study in this paper belongs to the class of random-coefficient autoregressive (RCA) models as proposed and studied by Andel (1976), Nicholls and Quinn (1982), McCabe and Tremayne (1995) and Granger and Swanson (1997):

$$y_t = \rho_t y_{t-1} + \eta_t, \quad t = 1, \cdots, T; \tag{3}$$

where η_t is assumed to be identically and independently distributed with zero expectation, variance σ_η^2 and moment conditions specified in assumption 2 below; ρ_t is a nonnegative covariance stationary process that is independent of η_t. The RCA model (3) with

$$\mathsf{E}\left[\max\left\{\log|\eta_t|, 0\right\}\right] < \infty \quad \text{and} \quad \mathsf{E}\left[\max\left\{\log|\rho_t|, 0\right\}\right] < \infty \tag{4}$$

is known (see Aue, Horváth and Steinebach, 2006) to admit a strictly non-anticipatory stationary solution if and only if

$$\mathsf{E}\left[\log|\rho_t|\right] < 0, \tag{5}$$

and a covariance stationary solution if

$$\mathsf{E}\left[\rho_t^2\right] < 1. \tag{6}$$

Hence, the unit root hypothesis can take several forms: $\mathsf{E}[\rho_t] = 1$, or $\mathsf{E}\left[\rho_t^2\right] = 1$, see Granger and Swanson (1997) for a discussion.[2] When $\mathsf{E}\left[\rho_t^2\right] > 1$, Hwang and Basawa (2005) denote this model an Explosive Random Coefficient Autoregressive model (ERCA) and study processes such that both $\mathsf{E}\left[\rho_t^2\right] \geq 1$ and $\mathsf{E}[\log|\rho_t|] < 0$ (which are strictly stationary but do not possess finite second moments).[3]

Here we follow Aue (2008) and deviate from the existing literature on RCA models *à la* Granger-Swanson in the sense that we assume that both the expectation and variance of $(\rho_t - 1)$ are very close to zero: we model the moments using extensions to standard local-asymptotic frameworks so that as $T \to \infty$ $(\mathsf{E}[\rho_T], \mathsf{V}[\rho_T]) \to (1, 0)$. This framework builds on Bobkoski (1983), Chan and Wei (1987), Phillips (1987) and the more recent work of Giraitis and Phillips (2006) and PM.

The data generating process (DGP) we consider is formally defined as a triangular array since the distribution of y_t, for $t \leq T$, is allowed to depend on the actual sample size T: we parameterize

---

[2]Several Lagrange-Multiplier tests of the unit root hypothesis have been proposed in this framework, see Leybourne, McCabe and Tremayne (1996), Hwang and Basawa (2005), Distaso (2008) and Aue and Horváth (2011).

[3]Also, expression (3) implies that y_t exhibits conditional heteroskedasticity: assume $\rho_t \sim iid\left(\rho, \sigma_\rho^2\right)$ then

$$\mathsf{E}\left[y_t|y_{t-1}\right] = \rho y_{t-1}, \quad \mathsf{Var}\left[y_t|y_{t-1}\right] = \sigma_\rho^2 y_{t-1}^2 + \sigma_\eta^2$$

see inter alia Tsay (1987), Yoon (2002), and Hwang and Basawa (2005). These authors, as well as others have also proposed functional forms that differ from (3) and that belong to the classes of double-autoregressive or bilinear processes.

the distribution of ρ_t to ensure that its realizations take the form of local deviation from a unit root, with an interest on deviations on the explosive side, hence the terminology *Near Explosive Random Coefficient* autoregressive model (NERC).

Throughout the paper, we make the following assumptions, where $\mathbb{R}_{+,*}$ denotes the set of strictly positive real scalars.

Assumption 1

$$\rho_t = \exp\left\{\frac{\phi + \lambda T^{\alpha/2} u_t}{T^\alpha}\right\} \qquad with \quad u_t \sim i.i.d.\,(0,1),$$

where $(\phi, \lambda, \alpha) \in \mathbb{R} \times \mathbb{R}_{+,*} \times (0,1)$, and where u_t and η_t are mutually independent.

Assumption 2 $y_0 = o_p\left(T^{\alpha/2}\right)$ and

$$\mathsf{E}\,|\eta_t|^\nu < \infty \quad for \quad \nu \geq \frac{2}{\alpha};$$

$$\mathsf{E}\,|u_t|^\omega < \infty \quad for \quad \omega \geq \frac{2}{\alpha}.$$

Assumption 1 implies that the parameters ϕ and λ^2 play similar roles in determining the magnitude of $\mathsf{E}\,[\rho_t] = \exp\left\{\left(\phi + \frac{1}{2}\lambda^2\right)/T^\alpha\right\}$. Also $\mathsf{V}\,[\rho_t] = \exp\left\{\left(2\phi + \lambda^2\right)/T^\alpha\right\}\left(\exp\left\{\lambda^2/T^\alpha\right\} - 1\right) = \frac{\lambda^2}{T^\alpha} + O\left(T^{-2\alpha}\right)$, so ρ_t admits the following stochastic expansion:

$$\rho_t = 1 + \frac{\phi + \frac{1}{2}\lambda^2}{T^\alpha} + \frac{\lambda}{T^{\alpha/2}} u_t + \frac{\lambda^2}{2T^\alpha}\left(u_t^2 - 1\right) + O_p\left(T^{-2\alpha}\right). \tag{7}$$

Assumption 2 ensures that the assumption (4) from Aue *et al.* (2006) is satisfied. It also implies that a strong approximation is possible, see Csörgő and Horváth (1993) and PM, according to which we can construct an expanded probability space with standard Brownian motions W, B such that, as $T \to \infty$,

$$\begin{cases} \sup_{s \in [0, T^{1-\alpha}]} \left|T^{-\alpha/2} \sum_{t=1}^{\lfloor sT^\alpha \rfloor} u_t - W_s\right| = o_{a.s.}\,(1)\,; \\ \sup_{r \in [0, T^{1-\alpha}]} \left|T^{-\alpha/2} \sigma_\eta^{-1} \sum_{t=1}^{\lfloor rT^\alpha \rfloor} \eta_t - B_r\right| = o_{a.s.}\,(1)\,. \end{cases} \tag{8}$$

In order to map the values of (ϕ, λ) corresponding to different properties of y_t, we define the following subsets of $\mathbb{R} \times \mathbb{R}_+$:

$$\mathcal{S}_w = \left\{(\phi, \lambda) \in \mathbb{R} \times \mathbb{R}_+,\ \phi + \lambda^2 < 0\right\};$$

$$\mathcal{S}_s = \left\{(\phi, \lambda) \in \mathbb{R} \times \mathbb{R}_+,\ \phi < 0\right\}.$$

The conditions (5) and (6) for strict and weak stationarity correspond respectively to $(\phi, \lambda) \in \mathcal{S}_s$ and $(\phi, \lambda) \in \mathcal{S}_w$. We also define the subset $\mathcal{S}_{s\setminus w} = \mathcal{S}_s \setminus \mathcal{S}_w$ of processes that are strictly stationary yet non weakly so. Using the results of Kesten (1973) and Goldie (1991) applied by Lux and Sornette (2002) to periodically collapsing bubble, the distribution of y_t for $(\phi, \lambda) \in \mathcal{S}_{s\setminus w}$ with $\lambda > 0$ can be shown to be characterized by a power law, in the sense that there exist $\tau > 0$
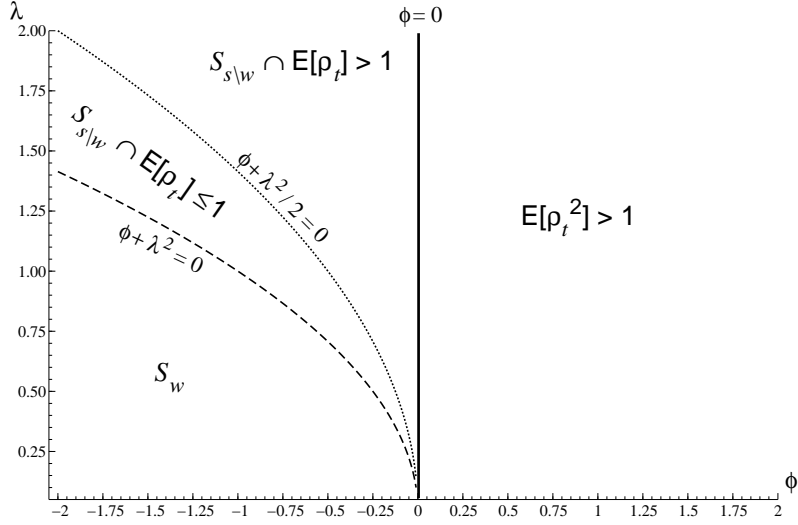
Figure 2: Values of $(\phi, \lambda) \in \mathbb{R} \times \mathbb{R}_+$ belonging to the subsets $\mathcal{S}_w$ and $\mathcal{S}_{s \setminus w}$ which correspond respectively to y_t being weakly stationary and strictly yet non weakly stationary. The figure also reports whether $\mathsf{E}[\rho_t] > 1$ and $\mathsf{E}[\rho_t^2] > 1$.

such that $\Pr\left(|y_t| > a\right) \sim \tau a^{-\sqrt{-2\phi}/\lambda}$ as $a \to \infty$. Hence moments[4] of y_t exist up to the order $\sqrt{-2\phi}/\lambda - 1 \leq \sqrt{2} - 1$. Hence $(\phi, \lambda) \in \mathcal{S}_{s \setminus w}$ implies that the process is not characterized by temporary explosive behavior (as when $\mathsf{E}[\rho_t^2] > 1$) but instead by *large deviations* caused by the fat tailed nature of the stationary distribution. Yet, fat tails can generate processes which appear to exhibit temporary bubbles (see the appendix, section F).

Notice the condition $\mathsf{E}[\rho_t] < 1 \Leftrightarrow \phi + \frac{1}{2}\lambda^2 < 0$ differs from those defining $\mathcal{S}_w$ and $\mathcal{S}_s$ as

$$\mathsf{E}\left[\log|\rho_t|\right] \leq \mathsf{E}\left[\rho_t\right] \leq \mathsf{E}\left[\rho_t^2\right],$$

where both equalities hold if and only if $\lambda = 0$, i.e. in the moderately explosive processes of PM and PY. The difference here is that $\rho_t \in [0, \infty)$: the autoregressive coefficient is allowed over time to enter the mean reversion region $(0, 1)$, to be close to unity and to lie on the explosive side $(1, \infty)$. We show in Figure 2 which values of (ϕ, λ) belong to the various subsets.

The model we propose deviates non-trivially from that of Aue (2008, Aue henceforth) in that we allow for a greater role played by the stochastic variation in ρ_t. In his setting $\mathsf{E}[\rho_t] - 1 = O\left(T^{-\alpha}\right)$ with $\alpha \in (1/2, 1)$, and $\mathsf{V}[\rho_t] = o\left(T^{-1}\right)$ which implies that $\mathsf{V}[\rho_t]$ lies in a tighter neighborhood of unity and so does not asymptotically impact[5] the tail distributions or explosiveness of y_t. In his framework, the asymptotic distributions of the least-squares estimator of the AR(1) regression parameter coincide with PM. Our assumptions extend Aue (2008) to the situation where $\mathsf{V}[\rho_t]$ lies further away from zero[6] and we show that this affects significantly the asymptotic distributions.

---

[4]See theorem 1 in Lux and Sornette (2002) where the moment conditions are satisfied under our assumption 2.

[5]For Aue, conditions $\mathsf{E}\left[\rho_t^2\right] < 1$ and $\mathsf{E}\left[\rho_t\right] < 1$ are asymptotically equivalent so $\mathcal{S}_{s \setminus w} = \varnothing$.

[6]We rule out the assumption of fixed (non-local) parameterization, $\alpha = 0$.

Accordingly with PY, we restrict $\alpha < 1$ to ensure that ρ_t is sufficiently away from unity for y_t to exhibit properties distinctively different from those of a random walk (in a sense that will become clear).

An empirical analysis of the ERCA model with $\mathsf{E}[\rho_t] > 1$ and non-local parameters ($\alpha = 0$) was made by Charemza and Deadman (1995) in the context of periodically collapsing bubbles (see also, Aue and Horváth, 2011, and Wang and Gosh, 2009). We show here that, following the recent work by P. C. B. Phillips and his coauthors, the introduction of a local-asymptotic framework yields benefits. We present simulated paths of the NERC process in section F of the appendix.

2.2 Asymptotic distribution

The first step of our analysis is to provide a Functional Central Limit Theorem (FCLT) for the NERC model. For this we define, for $(\phi, \lambda) \in \mathbb{R} \times \mathbb{R}_+$ and $r \in \mathbb{R}_+$, the stochastic integral of a geometric Brownian motion as the diffusion:

$$K_{\phi,\lambda}(r) = \int_0^r \exp\left\{(r - s)\phi + \lambda(W_r - W_s)\right\} dB_s. \tag{9}$$

The FCLT follows.

Proposition 1 *Let the process y_t be defined for $t \geq 0$ by (3) under assumptions 1 and 2. Then, for $r \in [0, T^{1-\alpha}]$ and as $T \to \infty$,*

$$T^{-\alpha/2} y_{\lfloor rT^\alpha \rfloor} \Rightarrow \sigma_\eta K_{\phi,\lambda}(r).$$

Corollary 2 *Proposition 1 also holds when $\alpha = 1$, so $T^{-1/2} y_{\lfloor rT \rfloor} \Rightarrow \sigma_\eta K_{\phi,\lambda}(r)$ for $r \in [0, 1]$.*

Throughout the paper, asymptotic behaviors depend on the sign of $\log \mathsf{E}[\rho_t^2] = 2T^{-\alpha}(\phi + \lambda^2)$ so we define

$$c = \phi + \lambda^2, \tag{10}$$

which extends the role played by ϕ in PM. Proposition 1 shows that several cases arise depending on whether the distribution of $K_{\phi,\lambda}(r)$ remains bounded. Indeed, $K_{\phi,\lambda}(r) \sim \mathsf{N}\left(0, \int_0^r e^{2cs} ds\right)$ and it reduces when $\lambda = 0$ to the Ornstein-Uhlenbeck diffusion considered in PM. Since $\mathsf{E}[y_T] = 1 + o(1)$, the magnitude of y_T is similar to that which PM obtain when $c \leq 0$: denoting by sd the standard deviation,

$$\mathsf{sd}(y_T) = \begin{cases} O\left(T^{\frac{\alpha}{2}}\right), & \text{if} \quad c < 0; \\ O\left(T^{\frac{1}{2}}\right), & \text{if} \quad c = 0; \\ O\left(T^{\alpha/2} e^{cT^{1-\alpha}}\right), & \text{if} \quad c > 0, \end{cases} \tag{11}$$

where c differs from ϕ when $\lambda \neq 0$. The latter expression shows that when $c > 0$, the process exhibits explosiveness in its second moment as pointed out by Hwang and Basawa (2005). Clearly Corollary 2 and expression (11) together imply that, as is the case for the NEAR(1) model, explosive patterns may only arise if $\alpha < 1$; this is the case we consider in the paper.

3 Inference

This section delineates a methodology for drawing inference on the parameters (ϕ, λ) of the NERC(1). In the spirit of PY and PWY, we first derive the distribution of the ordinary least-squares (OLS) $\widehat{\rho}$, which is the estimator originally proposed by Quinn and Nicholls (1980). Since alternative quasi-maximum likelihood estimators are known to present consistency issues in the ERCA (i.e. when $\alpha = 0$, see Berkes et al., 2009), we then propose to draw inference on (ϕ, λ) using solely the distribution of $\widehat{\rho}$.

3.1 Least Squares Estimator

We now consider the distribution of the OLS estimator $\widehat{\rho}$ in the regression of y_t on y_{t-1}. The expansion (7) implies that, as $T \to \infty$,

$$y_t = \left(\mathsf{E}\left(\rho_t\right) + \lambda T^{-\alpha/2} u_t + O_p\left(T^{-\alpha}\right) \right) y_{t-1} + \eta_t.$$

Hence, letting $S_{yyu} = \sum_{t=1}^T y_{t-1}^2 u_t$, $S_{y\eta} = \sum_{t=1}^T y_{t-1} \eta_t$ and $S_{yy} = \sum_{t=1}^T y_{t-1}^2$, the OLS estimator satisfies

$$\widehat{\rho} - \mathsf{E}\left(\rho_t\right) = \lambda T^{-\alpha/2} \frac{S_{yyu}}{S_{yy}} + \frac{S_{y\eta}}{S_{yy}} + O_p\left(T^{-3\alpha/2}\right), \tag{12}$$

and its asymptotic distribution is driven by the the sum with higher magnitude between $T^{-\alpha/2} S_{yyu}$ and $S_{y\eta}$. For this analysis, we introduce the following random variables:

$$V_{T^{1-\alpha}} = \int_0^{T^{1-\alpha}} e^{2(\phi r + \lambda W_r)} dW_r, \quad Z_{T^{1-\alpha}} = \int_0^{T^{1-\alpha}} e^{2(\phi r + \lambda W_r)} dr,$$

$$X_{T^{1-\alpha}} = \int_0^{T^{1-\alpha}} e^{-(\phi r + \lambda W_r)} dB_r \quad , Y_{T^{1-\alpha}} = \int_0^{T^{1-\alpha}} e^{\phi r + \lambda W_r} dB_r.$$

Let a tilde denote the centered random variable scaled by its standard deviation (e.g. $\tilde{V}_{T^{1-\alpha}} = (\mathsf{V}\left[V_{T^{1-\alpha}}\right])^{-1/2} (V_{T^{1-\alpha}} - \mathsf{E}\left[V_{T^{1-\alpha}}\right])$). We show in the appendix that, as $T \to \infty$, $\tilde{V}_{T^{1-\alpha}}$, $\tilde{X}_{T^{1-\alpha}}$, $\tilde{Y}_{T^{1-\alpha}}$ and $\tilde{Z}_{T^{1-\alpha}}$ converge weakly to random variables V, X, Y, and Z: the variables X, Y and V are standard normal and Z is a random variable with zero expectation and unit variance.[7] In addition Z does not correlate with V. We can now provide the weak convergence of the sample moments:

Lemma 3 *Let the process y_t be defined for $t \geq 0$ by (3) under assumptions 1 and 2, then as $T \to \infty$:*
· if $c < 0$,

$$T^{-(1+\alpha)} S_{yy} \xrightarrow{P} \frac{\sigma_\eta^2}{2c}, \quad T^{-\frac{1+\alpha}{2}} S_{y\eta} \xrightarrow{L} \mathsf{N}\left(0, \frac{\sigma_\eta^2}{2c}\right), \quad T^{-\frac{1+2\alpha}{2}} S_{yyu} \xrightarrow{L} \mathsf{N}\left(0, \frac{12\sigma_\eta^2}{c^2}\right);$$

---

[7]Matsumoto and Yor (2005), theorem 7.4, show how the distribution of Z can be expressed (for some values of the parameters) in terms of transforms of Brownian motions involving a Gamma variable.

$\cdot$ *if $c \geq 0$ and for $x \in \{u, \eta\}$, there exist (μ^x, ϕ_T^x) functions of (ϕ, λ) such that*

$$T^{-\alpha/2} \phi_T^u S_{yy} \Rightarrow \frac{\sigma_\eta^2}{\mu^u \sqrt{c + 2\lambda^2}} X^2 Z, \quad \phi_T^\eta S_{y\eta} \Rightarrow \frac{\sigma_\eta^2}{\mu^\eta} XY, \quad \phi_T^u S_{yyu} \Rightarrow \frac{\sigma_\eta^2}{\mu^u} X^2 V,$$

with $\phi_T^u / \phi_T^\eta = o\left(e^{-2\lambda^3 T^{1-\alpha}}\right)$.

The lemma implies the following: (i) when $c < 0$, i.e. when the process is weakly stationary, then both S_{yyu} and $S_{y\eta}$ impact the asymptotic distributions; but (ii) when $\lambda \neq 0$ and $c \geq 0$ S_{yyu} dominates. This setting differs markedly from that of Aue where the variance of ρ_t is of lower magnitude so $S_{y\eta}$ is the dominant term in the expansion (12). It also differs from the fixed-asymptotics framework of Hwang and Basawa (2005) where the ratio S_{yyu}/S_{yy} is not premultiplied by $T^{-\alpha/2}$ and hence diverges: the OLS estimator is inconsistent there. This is not the case here as the following theorem shows.

Theorem 4 *Let the process y_t be defined for $t \geq 0$ by (3) under assumptions 1 and 2, with $\lambda \neq 0$. Letting $c = \phi + \lambda^2$, the OLS estimator $\widehat{\rho}$ in the regression of y_t on y_{t-1} then satisfies as $T \to \infty$:*
if $c < 0$, $T^{\frac{1+\alpha}{2}} (\widehat{\rho} - \mathsf{E}[\rho_t]) \Rightarrow \mathsf{N}\left(0, 3\lambda^2 - 2c\right),$
if $c \geq 0$, $T^\alpha (\widehat{\rho} - \mathsf{E}[\rho_t]) \Rightarrow \lambda \sqrt{c + 2\lambda^2}\, \dfrac{V}{Z}.$

This theorem presents several key differences from the existing literatures on near unit roots and random coefficients when $c \geq 0$. When $c < 0$, the asymptotic distribution of the OLS estimator $\widehat{\rho} - \mathsf{E}[\rho_t]$ is comparable to the results of PM and Aue that $T^{\frac{1+\alpha}{2}} (\widehat{\rho} - \rho) \Rightarrow \mathsf{N}\left(0, -2\phi + \lambda^2\right)$: the presence of the stochastic root does not affect the asymptotic normality of $\widehat{\rho}$ or the rate of convergence; the only difference is that the asymptotic variance is increased by λ^2.

By contrast, when $c \geq 0$ the results are new. Here the OLS estimator converges more slowly than under the constant parameter AR(1) : it does not achieve the $O_p\left(T^{-1}\right)$ of unit root processes or the exponential rate of PM where $(2\phi)^{-1} T^\alpha e^{\phi T^{1-\alpha}} (\widehat{\rho} - \rho)$ tends to a standard Cauchy variable. Convergence can be arbitrarily slow here if α is close to zero: the limit $\alpha \to 0$ corresponds to the fixed-asymptotics of Hwang and Basawa (2005) where the estimator is shown to be inconsistent. Also, the limiting distribution is expressed, as in PM or Aue, as the ratio of two uncorrelated random variables. Yet, Z is not standard normal (although it has zero expectation and unit variance). This implies that V/Z does not define a Cauchy variable contrary to the limiting distribution in PM.

Theorem 4 shows that $\widehat{\rho}$ allows to estimate $\phi + \lambda^2/2$ consistently when $c < 0$ since

$$T^\alpha (\widehat{\rho} - 1) = \phi + \frac{1}{2}\lambda^2 + O_p\left(T^{-\frac{1-\alpha}{2}}\right).$$

This is not the case for $c \geq 0$ as the convergence of $\widehat{\rho}$ is then too slow.

The theorem also shows that under the NERC model, the unit root problem does not exist when $c \geq 0$ since the asymptotic distribution does not show the usual knife-edge problem as c

10

tends to zero from above (see Berkes et al., 2009, for a discussion). This may pose difficulties as the following corollary shows.

Corollary 5 *Under the assumptions and conditions of Theorem 4, define the test statistic $\tau_{0,T}$ for the null* $\mathsf{H}_0 : (\phi, \lambda) = (\phi_0, \lambda_0)$ *as*

$$\tau_{0,T} = \begin{cases} T^{\frac{1+\alpha}{2}} \left(\widehat{\rho} - \mathsf{E}_{\mathsf{H}_0} \left(\rho_t \right) \right), & \text{if } c_0 < 0; \\ T^{\alpha} \left(\widehat{\rho} - \mathsf{E}_{\mathsf{H}_0} \left(\rho_t \right) \right), & \text{if } c_0 \geq 0. \end{cases}$$

where $c_0 = \phi_0 + \lambda_0^2$. Then under $\mathsf{H}_1 : (\phi, \lambda) = (\phi_1, \lambda_1) \neq (\phi_0, \lambda_0)$ *and, as $T \to \infty$,*

$$\tau_{0,T} \underset{\mathsf{H}_1}{=} \begin{cases} O_p \left(T^{\frac{1-\alpha}{2}} \right), & \text{if } c_0 < 0; \\ O_p \left(1 \right), & \text{if } c_0 \geq 0. \end{cases}$$

The corollary shows that the test based on the OLS estimator is asymptotically powerful when the null implies that the process is weakly stationary. Yet the test statistic does not diverge asymptotically (so the test has low asymptotic power) when the null implies that y_t is not weakly stationary. This holds irrespective of the alternative hypothesis. The corollary sheds light on the reason why the simulations of Evans (1991) and Charemza and Deadman (1995) show that the Dickey-Fuller test has low power in the presence of periodically collapsing bubbles.[8]

In addition Theorem 4 shows that the distribution of $\tau_{0,T}$, as defined in the corollary, does not depend on the nuisance parameter σ_η^2. This is a key feature. Indeed, the OLS estimator is less efficient than the quasi-maximum likelihood estimators (QMLE) of the expectation and variance of ρ_t that have been proposed e.g. by Quinn and Nicholls (1981). Yet the QMLE does not exist in closed form and requires a consistent estimator of σ_η^2. Berkes et al. (2009) show that in the fixed-parameter explosive model (i.e. $\alpha = 0$ and $c > 0$) the QMLE of σ_η^2 is inconsistent. The estimator suggested by Schick (1996) could be used in combination with other estimators, see e.g. Aue and Horvath, (2012) or Hwang and Basawara, (2005). However, the properties of the estimator of Schick are still unknown for the ERCA model. For this reason, we propose in the next subsection an alternative methodology that is feasible although it relies on an underidentified parameterization.

3.2   Inference using Grid Testing

The DGP we consider uses a local-asymptotic parameterization and Theorem 4 shows that consistent estimation of the localizing parameters (ϕ, λ) may be unfeasible when $c \geq 0$.[9] To conduct inference, we resort hence to the technique which is now standard under local asymptotics and consists in inverting a test statistic. There exists a significant literature where such an approach is

---

[8]This is not the only such case in the literature, indeed the locally best invariant Lagrange-Multiplier test of Leybourne *et al.* (1996) was also shown not to be consistent under the unit root hypothesis against explosive alternatives (see Nagakura, 2009).

[9]Hence, we do not consider the nonlinear Kalman or particle filters.

used for inference in the near-unit root framework (originating in Stock, 1991).[10] Instead of focusing on point estimation, inference consists, here, in constructing asymptotically valid confidence sets for the parameters of interest.

The technique relies on introducing a scalar function $\tau_{\theta,T}$ of $y_1, ..., y_T$ (a test statistic) that satisfies

$$\tau_{\theta,T} \Rightarrow \mathcal{D}_\theta, \tag{13}$$

where $\theta = (\phi, \lambda)' \in \Theta$ denote the parameters of interest, here $\theta = (\phi, \lambda)$ and $\mathcal{D}_\theta$ denotes a distribution that depends on θ. Under the null $\mathsf{H}_0 : \theta = \theta_0$, Stock (1991) constructs asymptotic $(1-\omega)\%$ confidence sets as $\Theta_\omega \subset \Theta$ consisting of the values θ_0 which are not rejected at the $\omega\%$ significance level by the limit of $\tau_{\theta_0,T}$ as $T \to \infty$. The finite sample corrections of Andrews (1993) and Hansen (1999) have been shown by Mikusheva (2007, see also 2012) to be uniformly valid. In this setting, the least rejected parameter θ^* may constitute a biased estimator of θ but median-unbiased estimation is feasible under the weak convergence assumption, provided that the quantile function is monotonic (Stock, 1991, Andrews, 1993). When τ is a Generalized Method of Moments (GMM) statistic, θ^* can be seen as the continuously-updated estimator (see Stock, Wright and Yogo, 2002) and it inherits its properties.

Here we conduct inference under the null

$$\mathsf{H}_0 : (\phi, \lambda) = (\phi_0, \lambda_0).$$

Since $y_t - \mathsf{E}_{\mathsf{H}_0} [\rho_T] y_{t-1} = (\rho_t - \mathsf{E}_{\mathsf{H}_0} [\rho_T]) y_{t-1} + \eta_t$, we use the moment condition:

$$\mathsf{Cov}(y_t - \mathsf{E}_{\mathsf{H}_0} [\rho_T] y_{t-1}, y_{t-1}) \underset{H_0}{=} 0.$$

The test we choose for simplicity follows the pseudo Dickey-Fuller autoregression:

$$y_t - \mathsf{E}_{\mathsf{H}_0} [\rho_T] y_{t-1} = (\widehat{\rho} - \mathsf{E}_{\mathsf{H}_0} [\rho_T]) y_{t-1} + \eta_t, \tag{14}$$

and we set $\tau_{\theta,T}$ to be the OLS estimator $\widehat{\rho} - \mathsf{E}_{\mathsf{H}_0} [\rho_T]$ scaled by the asymptotic rate given in Theorem 4. Confidence sets are obtained by grid search over all possible values of (ϕ, λ) and critical values are obtained by simulation. The variance σ_η^2 constitutes a scaling parameter that does not affect the asymptotic distribution of $\widehat{\rho} - \mathsf{E}_{\mathsf{H}_0} [\rho_T]$ so we may fix it to unity. Also, α is not identified using the method: it constitutes only a scaling parameter since it does not enter the asymptotic distributions in Theorem 4[11]. In the following, we consider testing against either a one-sided alternative (rejection in the upper tail) or a two-sided alternative. The least-rejected parameter values in the two-sided test correspond to an under-identified Method of Moment estimator. The one-sided test can only be used to construct confidence sets.

---

[10]This technique is also common in the context of weak instruments where there exists no fully robust estimation method, but robust tests can be constructed (see Dufour, 1997, and Staiger and Stock, 1997). For papers that discuss the mechanics of the inversion of robust tests to form confidence sets, see Andrews and Stock (2005) and references therein.

[11]Since α is a scaling parameter it will be fixed as $\alpha = 1/2$ in the empirical applications.

Following Phillips (2012) we recognize that as $|\phi| \to \infty$ or $\lambda \to \infty$, the asymptotic distribution of the estimator becomes diffuse so the confidence sets may become empty when the true data generating process does not present local parameters. Also, Corollary 5 shows that although we obtain valid asymptotic confidence sets under the null, the asymptotic power is low and the proposed confidence sets may be too wide: we assess their coverage probabilities by simulation in section 5.

4 Forecasting

An attractive feature of the model we propose, is that it provides a distributional assumption about ρ_t contrary to models where ρ_t breaks deterministically. As a consequence, we can answer questions on the probability that a bubble forms, bursts, continues and so on. There exist several ways to define a bubble and to characterize its timing and magnitude (see e.g. White and Granger, 2011), but our purpose here is not to provide an extensive characterization. Many definitions of a bubble imply that the growth rate is greater than some preassigned growth rate over a finite horizon, say $k > 0$ periods. This implies that we should be concerned with events such as $\{y_{t+k}/y_t \geq \gamma\}$ for some $\gamma > 0$. We define the probability of this event as

$$\mathcal{P}_{t,k}^{\gamma} \equiv \mathsf{P}\left(\frac{y_{t+k}}{y_t} \geq \gamma\right). \tag{15}$$

An example of a question of interest may for instance concern the probability $\mathcal{P}_{t,k}^{y_t/y_{t-k}}$ that over the horizon $k > 0$, the process grows at least as fast as has been observed over the last k periods.

We define $f_c(r) = (e^{cr} - 1)/c$ for $c \neq 0$ and $f_0(r) = r$. The following proposition shows how the questions above frame into a simple analytic expression using our model.

Proposition 6 *Under the assumptions and conditions of Theorem 4, then*
(i) for $(r,s) \in \left(0, T^{1-\alpha}\right]^2$ and, as $T \to \infty$,

$$\frac{y_{\lfloor T^{\alpha}(r+s)\rfloor}}{y_{\lfloor T^{\alpha}r\rfloor}} \Rightarrow \exp\left\{\phi s + \lambda\left(W_{r+s} - W_r\right)\right\} + \sqrt{\frac{f_{2c}(s)}{f_{2c}(r)}}\, C,$$

where C is a standard Cauchy variable, and
(ii) if $c \geq 0$ with $\lambda \neq 0$, then as $(T, r) \to (\infty, \infty)$ such that $s/r \to 0$,

$$\mathcal{P}_{\lfloor T^{\alpha}r\rfloor, \lfloor T^{\alpha}s\rfloor}^{\gamma} \to \Phi\left(\frac{\phi s - \log \gamma}{\lambda\sqrt{s}}\right),$$

where Φ denotes the standard normal cumulative distribution function.

The proposition shows that the distribution of the ratio $\frac{y_{\lfloor T^{\alpha}(r+s)\rfloor}}{y_{\lfloor T^{\alpha}r\rfloor}}$ is driven by the sum of two random variables which are not independent of each other: (i) the sum of the increment

$W_{r+s} - W_r$ which is independent on W_{r-s}, $s \geq 0$ and (ii) a Cauchy variable C. When $r \to \infty$ such that $s/r \to 0$ the scaling factor $f_{2c}(s)/f_{2c}(r)$ tends to zero if $c \geq 0$ so the impact of the Cauchy C vanishes: then probabilities such as $\mathcal{P}^\gamma_{\lfloor T^\alpha r \rfloor, \lfloor T^\alpha s \rfloor}$ are given by the normal distribution and can be factored when they relate to non-overlapping events.

For instance the definition of a technical bubble in White and Granger (2011), requires that there exist $t_1 < t_2 < t_3$ such that the bubble builds up between t_1 and t_2, i.e. $y_{t+1}/y_t > 1$ for $t_1 \leq t < t_2$, and it collapses between t_2 and $t_3 : y_{t+1}/y_t < 1$ for $t_2 \leq t < t_3$. The proposition shows that when $c \geq 0$, this bubble condition can therefore be written asymptotically as the product of individual probabilities:[12]

$$\mathcal{P}^1_{t_1,1} ... \mathcal{P}^1_{t_2-1,1} \left(1 - \mathcal{P}^1_{t_2,1}\right) ... \left(1 - \mathcal{P}^1_{t_3-1,1}\right).$$

In addition, proposition $6(ii)$ shows that probabilities such as $\mathcal{P}^\gamma_{t,k}$ asymptotically tend to very simple expressions, where the functional central limit theorem yields normality. A finite sample approximation based on the asymptotic distributions obtains, for $k << t$:

$$\mathcal{P}^\gamma_{t,k} \approx \Phi\left(\frac{\phi k T^{-\alpha} - \log \gamma}{\lambda \sqrt{k T^{-\alpha}}}\right).$$

The expression above relies crucially on $\lambda \neq 0$; when this is not the case and under the local-asymptotic approximation, $\mathcal{P}^\gamma_{\lfloor T^\alpha r \rfloor, \lfloor T^\alpha s \rfloor} \to 1_{\{\phi s - \log \gamma \geq 0\}}$, with $1_{\{\cdot\}}$ the indicator function.

Proposition 6 shows how to compute asymptotic probability statements parametrically when the DGP is known. When the parameters (ϕ, λ) are unknown, it is possible to use the inference method we suggested in section 3.2 in order to obtain confidence intervals for the probability forecasts $\mathcal{P}^\gamma_{t,k}$. As before, let Θ_ω denote an asymptotic $(1-\omega)\%$ confidence set for (ϕ, λ). Then if all $(\phi, \lambda) \in \Theta_\omega$ satisfy $\phi + \lambda^2 \geq 0$, we can compute the infimum and supremum of $\mathcal{P}^\gamma_{t,k}$ by considering all parameter combinations in the confidence set. We show an example of this methodology in the empirical section.

5 Monte Carlo

5.1 Finite Sample Confidence Sets

We now provide a short Monte Carlo evaluation of the finite sample probability coverage of confidence sets. Asymptotic distributions are obtained via simulation, using samples of $T = 10,000$ observations. All Monte Carlo distributions are obtained using 10,000 replications. We set $\alpha = 1/2$ since it is only a scaling parameter that does not affect the asymptotic distribution.

---

[12]White and Granger also impose the condition $(y_{t_2} - y_{t_1})/(t_2 - t_1) \leq |(y_{t_3} - y_{t_2})/(t_3 - t_2)|$ so the collapse is sharper than the build-up, this event has probability:

$$1 - \mathcal{P}^{1 - \left(1 - y_{t_1}/y_{t_2}\right)\frac{t_3 - t_2}{t_2 - t_1}}_{t_2, t_3 - t_2}.$$

under the conditions that $y_{t_3} \leq y_{t_2}$ and $y_{t_2} \neq 0$.

	Probability coverage for a nominal 95% Interval																				
	(a) **One-sided test,** $T = 3,000$							(b) **One-sided test,** $T = 300$							(c) **Two-sided test,** $T = 300$						
$\lambda =$	0	0.1	0.2	0.5	1	1.5	2	0	0.1	0.2	0.5	1	1.5	2	0	0.1	0.2	0.5	1	1.5	2
ϕ	**Asymptotic Distribution**																				
-0.2	0.97	0.97	0.97	0.93	0.90	0.89	0.89	0.98	0.98	0.98	0.85	0.74	0.69	0.70	0.92	0.92	0.92	0.68	0.72	0.91	0.94
0	0.95	0.90	0.92	0.94	0.93	0.92	0.91	0.95	0.77	0.83	0.87	0.87	0.83	0.78	0.98	0.64	0.64	0.63	0.82	0.90	0.93
0.01	1.00	0.91	0.93	0.94	0.93	0.93	0.93	1.00	0.78	0.84	0.87	0.87	0.83	0.80	1.00	0.63	0.64	0.63	0.82	0.91	0.93
0.05	1.00	0.95	0.95	0.94	0.94	0.94	0.93	1.00	0.86	0.90	0.91	0.91	0.88	0.81	1.00	0.55	0.59	0.64	0.83	0.91	0.93
0.1	1.00	0.95	0.95	0.95	0.94	0.93	0.94	1.00	0.91	0.94	0.93	0.91	0.90	0.86	1.00	0.60	0.68	0.71	0.85	0.90	0.93
0.2	1.00	0.95	0.95	0.95	0.94	0.94	0.94	1.00	0.95	0.95	0.94	0.93	0.92	0.90	1.00	0.83	0.86	0.84	0.87	0.91	0.92
0.3	1.00	0.94	0.95	0.95	0.95	0.95	0.95	1.00	0.95	0.95	0.95	0.94	0.93	0.92	1.00	0.93	0.94	0.92	0.89	0.91	0.92
ϕ	**Gaussian Wild Bootstrap**																				
-0.2	1.00	1.00	1.00	0.99	0.99	0.99	0.99	1.00	1.00	1.00	1.00	0.98	0.98	0.98	1.00	1.00	1.00	1.00	0.99	0.02	0.02
0	1.00	0.99	0.97	0.97	0.98	0.99	0.99	1.00	1.00	1.00	0.99	0.97	0.97	0.98	1.00	1.00	1.00	1.00	1.00	0.03	0.02
0.01	1.00	0.98	0.97	0.96	0.98	0.98	0.99	1.00	1.00	1.00	0.99	0.97	0.97	0.98	1.00	1.00	1.00	1.00	1.00	0.03	0.02
0.05	1.00	0.98	0.97	0.95	0.97	0.99	0.99	1.00	1.00	1.00	0.99	0.97	0.97	0.98	1.00	1.00	1.00	1.00	1.00	0.03	0.02
0.1	1.00	0.99	0.96	0.93	0.97	0.98	0.99	1.00	1.00	1.00	0.99	0.96	0.97	0.98	1.00	1.00	1.00	1.00	1.00	0.03	0.02
0.2	1.00	1.00	0.96	0.89	0.96	0.97	0.98	1.00	1.00	1.00	0.99	0.95	0.96	0.97	1.00	1.00	1.00	1.00	0.61	0.04	0.03
0.3	1.00	1.00	0.96	0.85	0.94	0.97	0.98	1.00	1.00	1.00	0.99	0.94	0.96	0.97	1.00	1.00	1.00	1.00	0.25	0.04	0.03

Table 1: Simulated finite sample probability coverage of confidence intervals constructed at a probability of 0.95 using the asymptotic distribution or that obtained using the Gaussian wild bootstrap. The parameter $\alpha = 1/2$. The simulated sample size is $T = 3,000$ in columns labeled (a) and $T = 300$ for columns labeled (b) and (c). The tests are one-sided for (a) and (b) and two-sided for (c). The number of Monte Carlo replications is $10,000$ and so is the sample size used in computing the asymptotic distribution.

The method of asymptotic inference introduced by Stock (1991) was modified in Hansen (1999) who recommended the use of a so called *grid bootstrap*. Such bootstrap aims at replacing the use of the asymptotic distribution (13) by the finite-sample bootstrap distribution whose critical values can be obtained by resampling from the empirical distribution of the error $v_t \equiv y_t - \mathsf{E}_{\mathsf{H}_0}[\rho_T] y_{t-1}$ (which is observed under H_0). Indeed $v_t = (\rho_t - \mathsf{E}_{\mathsf{H}_0}[\rho_T]) y_{t-1} + \eta_t$ exhibits conditional heteroskedasticity but is serially uncorrelated under the null. Hence, we also consider bootstrap techniques immune to heteroskedasticity, such as the wild bootstrap.[13]

We first analyze one-sided (upper tailed) tests. Columns labeled (a) and (b) in Table 1 report the simulated finite sample (respectively $T = 3,000$ and $T = 300$ observations) coverage probability of 95% confidence sets constructed using the asymptotic and bootstrap distributions. The tables show that the coverage is reasonable under the asymptotic distribution when $\phi > 0$ and $\lambda > 0$ although it is slightly lower than the nominal 95%. For $\phi < 0$ and the larger values of $\lambda > 0$, the probability coverage is even lower (we only report one value $\phi = -0.2$ as it does not seem to play an influential role). By contrast, coverage is slightly too wide when $c < 0$. Finally, when $\lambda = 0$, the test has low power and the coverage rate is inappropriately large, both in small and medium-sized samples. Notice as shown in our theoretical results that when $\lambda \to 0$, the distribution exhibits a nonlinearity: coverage is low for $\lambda > 0$ but large when $\lambda = 0$. The lower part of the table reports the coverage probabilities using the wild bootstrap. These lead to wider coverage and low discriminatory power. Notice the exception of the case $\lambda^2 \approx \phi$ (here for $\lambda = .5$ and ϕ between $.2$ and $.3$) in columns labeled (a) in Table 1; this corresponds to a discontinuity in the distribution of $X_{T^{1-\alpha}}$ defined in section 3 (see also expression (B.6) in the appendix).

The columns labeled (c) in Table 1 report the corresponding small sample $(T = 300)$ probability coverage using a two-sided test. Coverage rates using the asymptotic distributions are lower and the wild bootstrap is inadequate.

5.2   Power

We assess the power of the inference technique to reject the null of a constant autoregressive coefficient ρ_t under the alternative that it is random with same expectation. Figures 3, 4 and 5 report the rejection probabilities at the asymptotic nominal size of 10% of the null

$$\mathsf{H}_0 : (\phi_0, 0) \quad \text{vs} \quad \mathsf{H}_1 : (\phi, \lambda) \text{ with } \phi + \lambda^2/2 = \phi_0,$$

so the alternative preserves $\mathsf{E}[\rho_t]$. Figures 3 and 5 consider upper-tailed and bilateral tests for $\alpha = 1/2$. To consider larger values of the parameters (relative to the sample size), we also report one-sided tests for $\alpha = 1/4$ in Figure 4. The figures report the power for $T = 3,000$ and 300 observations. Left- and right-hand side columns show the same rejection probabilities but where

---

[13]Following Davidson and Flachaire (2008), we used the wild bootstrap with standard normal or Bernoulli distributed corrective factors but only report the former. Given the strong dependence in y_t when $c \geq 0$, we also considered the Maximum Entropy bootstrap (see Vinod, 2006) but we do not report it as it is computationally much slower yet does not improve upon the wild bootstrap here.

we parameterize the parameter space as (ϕ, λ) (left) or $\left(\phi + \lambda^2, \phi + \lambda^2/2\right)$ (right). For readability of the figures, we should stress for the reader that the axes have been rotated between one-sided and two-sided tests so great care must be taken when comparing the figures.

Starting with Figure 3, rejections probabilities are always larger than 0.5 and increase with λ. When ϕ is positive, and λ close to zero the power is at its minimum. This corresponds to nonstationary processes under either hypothesis. Correspondingly, the right-hand side columns show that when $\phi + \lambda^2$ is positive but close to zero and when $\phi + \lambda^2/2$ is large, then the power is itself low: the variance of the random coefficient is low but its expectation large. Comparing the right-hand side panels, we notice that the power does not increase with the sample size. This is in line with the results of Corollary 5, and confirms the analysis by Evans (1991) that stochastic bubbles, being non-permanent by nature, can be difficult to detect even when their magnitudes and the sample size are large.

Turning to Figure 4, where a lower α implies a higher magnitude of the parameters relative to the sample size, we see that the small sample $(T = 300)$ power is not affected, but that rejection probabilities drop to zero for large values of λ. We interpret this observation in light of Phillips (2012) who argues that inverting test statistics can lead to zero power when the distribution becomes diffuse, as is the case here when $\lambda \to \infty$, i.e. the first and second moments of ρ_t become large (upper right-hand side panel).

Finally, the bilateral rejection probabilities presented in Figure 5 show that this test has power of at least 50% when $\phi + \lambda^2 > 0$. Yet the simulation shows that test is unable to discriminate between a constant or a random coefficient when the DGP belongs to $\mathcal{S}_{s \setminus w}$ ($\phi + \lambda^2 < 0$ under the alternative), but is tested nonstationary under the null ($\phi_0 = \phi + \lambda^2/2 \geq 0$). In other words, the bilateral test does not reject the null of an explosive AR(1) under the alternative that it is a stationary NERC with power law distribution. Interestingly, the one-sided test does.

6   Empirical Application to Housing Prices

6.1   The Data

We now show how the model and results above can be used for the detection of bubbles in asset prices and their prediction. We follow the examples of PWY and PY and consider U.S. housing prices. Standard models relate P_t, the price at t, to the cash flow (the rent) D_{t+1} it generates between t and $t + 1$ so the ex-post realized return is r_{t+1} defined as

$$1 + r_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t}.$$

When D_t is integrated of order 1, PY show that when r_t varies, there may exist subperiods where the fundamental price is explosive as opposed to a martingale. In their application, they test the null that P_t/D_t follows a random walk against a near explosive alternative (NEAR(1)). Cointegration properties imply that P_t/D_t is expected to be stationary in the absence of bubble. The upper tailed unit root tests by PY hence consists in testing a null at the boundary of the set
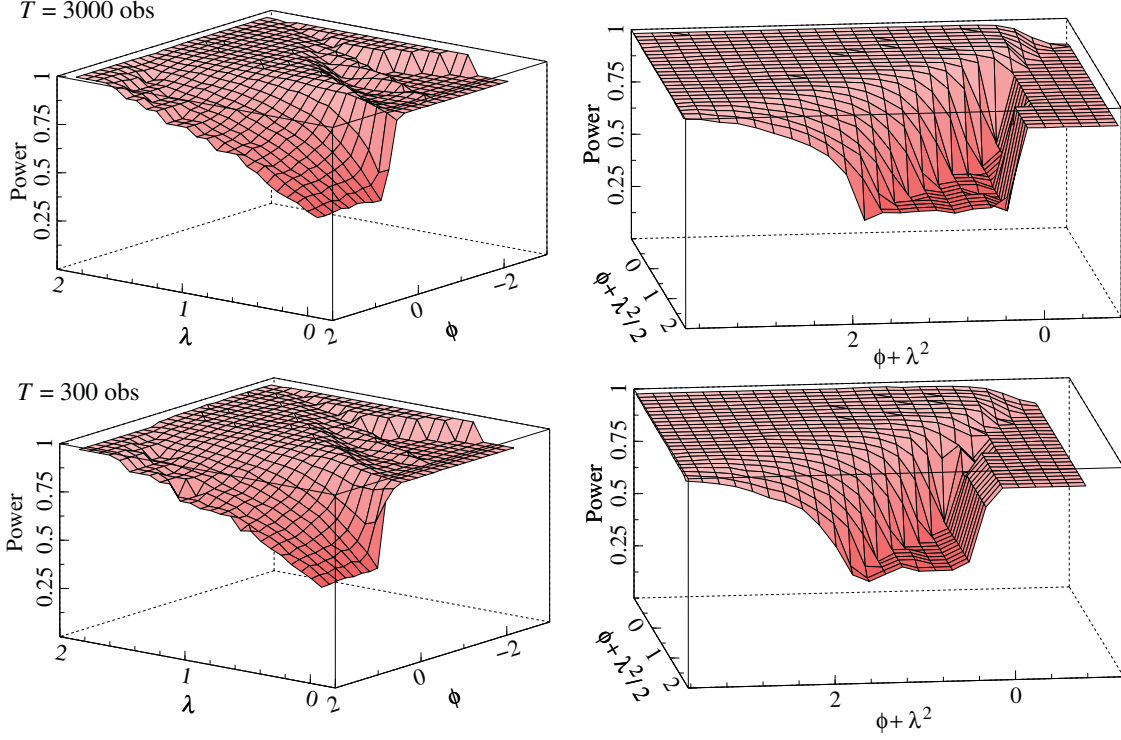
Figure 3: Upper-tail rejection probabilities at the asymptotic nominal size of 10% corresponding to the null $\mathsf{H}_0 : (\phi_0, 0)$ under the alternative $\mathsf{H}_1 : (\phi, \lambda)$ with $\phi + \lambda^2/2 = \phi_0$. We set the parameter $\alpha = .50$. Panels on the left-hand side reproduce those on the right-hand side, but with different axes where $\log \mathsf{E}(\rho_t) = \phi + \lambda^2/2$ and $\log \mathsf{E}(\rho_t^2) = \phi + \lambda^2$.
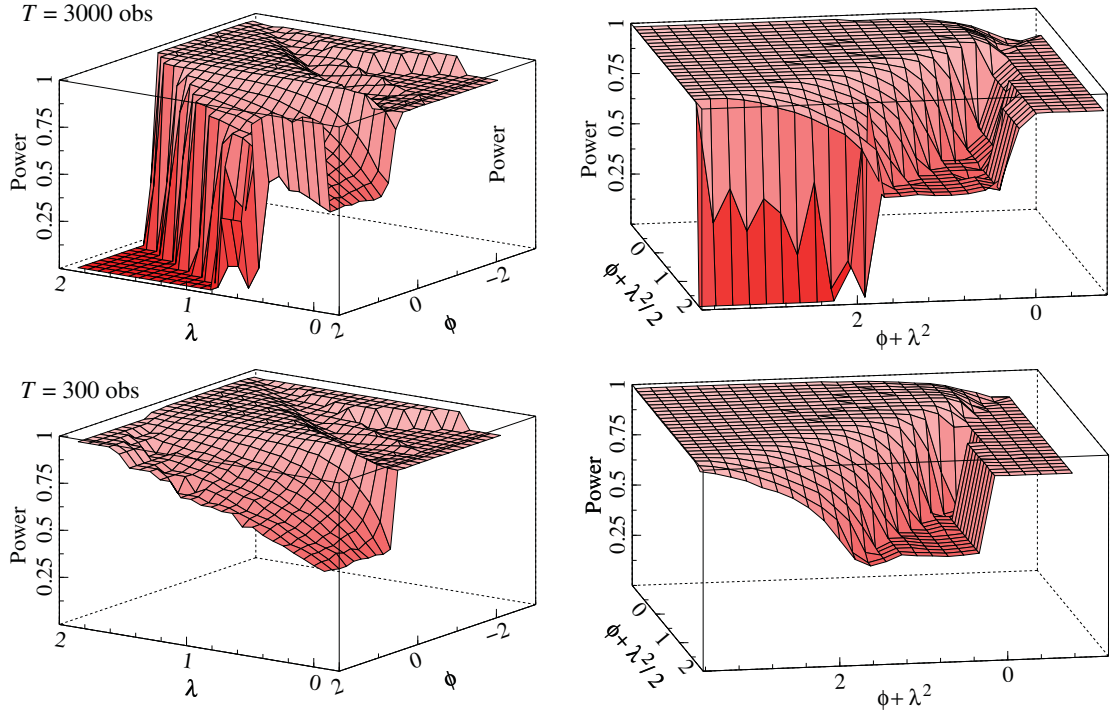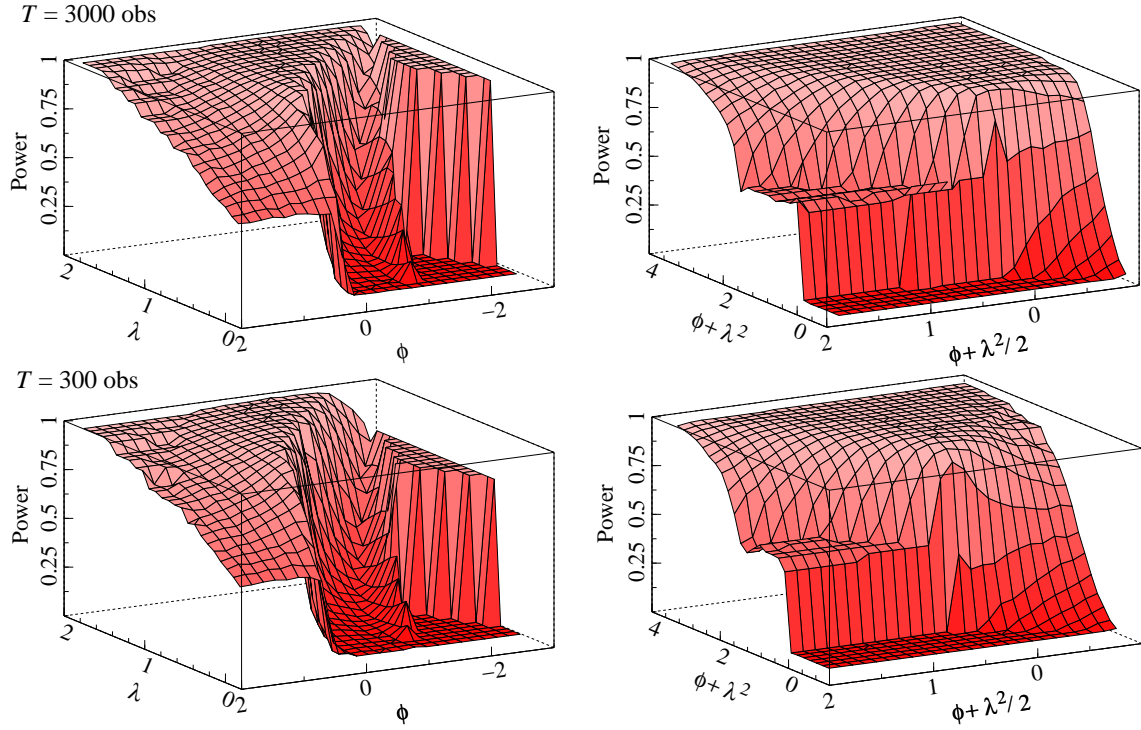
18

Figure 4: Upper-tail rejection probabilities at the nominal size of 10% corresponding to the null $\mathsf{H}_0 : (\phi_0, 0)$ under the alternative $\mathsf{H}_1 : (\phi, \lambda)$ with $\phi + \lambda^2/2 = \phi_0$. We set $\alpha = .25$. Panels on the left-hand side reproduce those on the right-hand side, but with different axes where $\log \mathsf{E}(\rho_t) = \phi + \lambda^2/2$ and $\log \mathsf{E}(\rho_t^2) = \phi + \lambda^2$.

Figure 5: Bilateral rejection probabilities at the nominal size of 10% corresponding to the null $H_0 : (\phi_0, 0)$ under the alternative $H_1 : (\phi, \lambda)$ with $\phi + \lambda^2/2 = \phi_0$. We set $\alpha = .5$. Panels on the left-hand side reproduce those on the right-hand side, but with different axes where $\log \mathsf{E}\,(\rho_t) = \phi + \lambda^2/2$ and $\log \mathsf{E}\,(\rho_t^2) = \phi + \lambda^2$.

of weakly stationary roots. In the NERC(1) model, the null can imply more general dynamics and we explore this below.

In addition, we recognize that the NERC may in principle generate downward negative bubbles which may not be consistent with applying it to P_t/D_t. Hence we also consider the standard present-value model of Campbell and Shiller (1987):

$$P_t = \mathsf{E}_t \left[\frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}} \right], \tag{16}$$

where $\mathsf{E}_t [\cdot]$ denotes the expectation conditional on information available at time t and R_{t+1} is the stochastic discount factor. We show in the appendix that under the standard assumption that D_t follows a random walk (driven by the $i.i.d.$ shock ζ_t) and the simplifying assumption that the ex-post return is constant and equal to R, the present-value relation (16) then admits the solution (with minimal number of state variables, see McCallum, 1983):

$$\Delta P_t = (1 + R + \delta (R_t - R)) \Delta P_{t-1} - \zeta_t, \tag{17}$$

for any $\delta \in [0,1]$.[14] Assuming R_t iid and uncorrelated with ζ_t, the dynamics of ΔP_t can be represented using the NERC(1) model if the expectation and variance of R_t are low. Expressions (7) and (17) exemplify in particular the parametric similarities.[15] Here, a large value of R_t – and hence a preference for the present in valuing assets – may generate explosiveness in ΔP_t.

In the following, we apply our methodology both to P_t/D_t and ΔP_t where P_t is the seasonally adjusted monthly Case-Shiller housing market price index maintained by Standard and Poor's (288 observations from 1987:1 to 2010:12).[16] For D_t, we follow PY and use the quarterly rental data imputed using the method of Davis, Lehnert, and Martin (2008) and linearly interpolated to a monthly frequency.[17] The series is presented in Figure 6: the price exhibits sustained growth over the 1987-2005 period followed by a sharp collapse. The figure shows that P_t/D_t and ΔP_t both exhibit patterns similar to those that arise under the NERC model.[18]

6.2 Inference

We first conduct inference using the whole sample at our disposal. To construct confidence sets, we perform grid searches using 3,000 uniform draws of the parameters $\phi \in [-1,1]$ and $\lambda \in [0,1]$, setting $\alpha = 1/2$.

---

[14]We are grateful to Kevin Lansing for suggesting the use of ΔP_t to us.

[15]The purpose of the assumption of constant ex-post returns is to maintain the assumption that ρ_t is $i.i.d.$ in a simple way. Extensions to the NERC(1) would allow more realistic models.

[16]PY use the Composite-10 index, but we use the Composite-30 instead as it provides longer series, but the results do not significantly differ.

[17]All the analysis is based on the series in deviation from their first observations. The non-negligible impact of nonzero origins on inference on nonstationary data is well known and warrants further study in the NERC model.

[18]RCA models such as (3) have also been used in the literature for the price level or log price of an asset, see e.g. Leybourne, McCabe and Mills (1996), Gonzalo and Lee (1998).
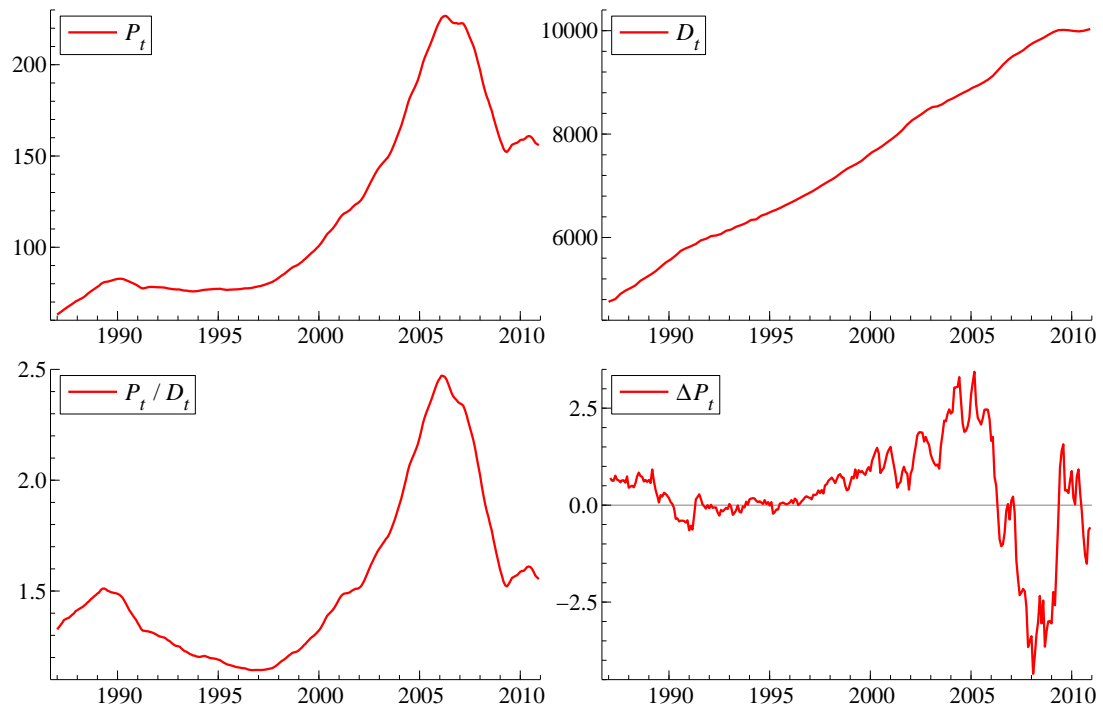
Figure 6: The seasonally adjusted monthly Case-Shiller Housing Composite-30 price index for the United-States (P_t) and rental price (D_t), together with the first-order difference ΔP_t and ratio P_t/D_t.

Figures 7 and 8 report on the top row the confidence sets obtained using both upper-tail (in red disks) and bilateral (in blue squares) tests. We adjust the significance levels to visualize the impact of right- and left-tail rejection.[19] We report in particular conference sets with low probability coverage to show the set of least rejected parameter values. The bottom rows record the loci of median estimates. The bottom left panel presents (in red) the parameters which imply a test statistic whose distance to the median is less than five times the smallest observed distance; it also records (in black) the locus of parameters corresponding to the 30 (i.e. 1% of parameter values considered) smallest distances to the medians. These distances are reported in increasing order in the bottom right panel. Figures 7 and 8 present inference based on ΔP_t and P_t/D_t respectively.

Inference based on ΔP_t and P_t/D_t imply different dynamics so we comment on them in turn. Figure 7 (ΔP_t) shows that upper-tail tests based on the asymptotic distribution lead to confidence sets that are predominantly within the $\mathcal{S}_{s\backslash w}$ region that correspond to fat-tailed DGPs. Yet, some of the parameter values close to the limit $\mathsf{E}\left[\rho_t^2\right] = 1$ cannot be rejected. This is the region for which our simulations showed that the technique has little power. Hence we also report the confidence sets based on the two-sided tests: these reject parameter values in $\mathcal{S}_w$ and even yield confidence sets for which not only $\mathsf{E}\left[\rho_t^2\right] > 1$ but even $\mathsf{E}\left[\rho_t\right] \geq 1$. This is also reflected by the locus of median estimates. We see on the bottom left panel that two parameter combinations yield similar distances from the estimate to the median of the distribution. For clarity, we report in Table 2 the minimum median-distance estimates, $(\phi, \lambda)^+ = (.17, .20)$, which fall in the explosive region.

Figure 8 reports inference on P_t/D_t based on the asymptotic distribution. The unilateral confidence sets are now predominantly close to the frontier between $\mathcal{S}_w$ and $\mathcal{S}_{s\backslash w}$. Again, the low power against weakly stationary alternative leads us to consider bilateral tests: these together with the locus of median estimates indicates that P_t/D_t might be strictly stationary with fat tails, the minimum median distance estimate is $(\phi, \lambda) = (-.25, .59)$ such that $c = \phi + \lambda^2 = 0.10$.

6.3   Bubble Detection

We now turn to real time detection and prediction using the NERC model. Since it is not possible to extract the latent $i.i.d$ process u_t, we apply grid testing to the recursive methodology suggested by PY. We specifically ask when the assumption $\mathsf{E}\left[\rho_t\right] = 1$ can be rejected using an upper tailed test. Here the null is composite in the sense that we test a set of parameter combinations such that $\phi + \lambda^2/2 = 0$. Specifically, Figure 9, panels (a) and (c), reports the minimum p-values obtained for 5,000 random draws of $\phi \in [0, 1/2)$ under the maintained null $\mathsf{H}_0 : \phi + \lambda^2/2 = 0$. We use the asymptotic distribution and report minimum p-values as indexed according to the end-of-sample date on the horizontal axis. The tests are one-sided so rejection of the composite null does not preclude that the series is strictly stationary but with very fat tails and explosive finite sample second moments.

---

[19]For sets with probabilities above 50%, and only for those, the confidence levels are chosen so that the bilateral confidence set is part of the unilateral set; hence the reader should therefore consider that the one-sided confidence set comprises both red disk and blue square parameter combinations.
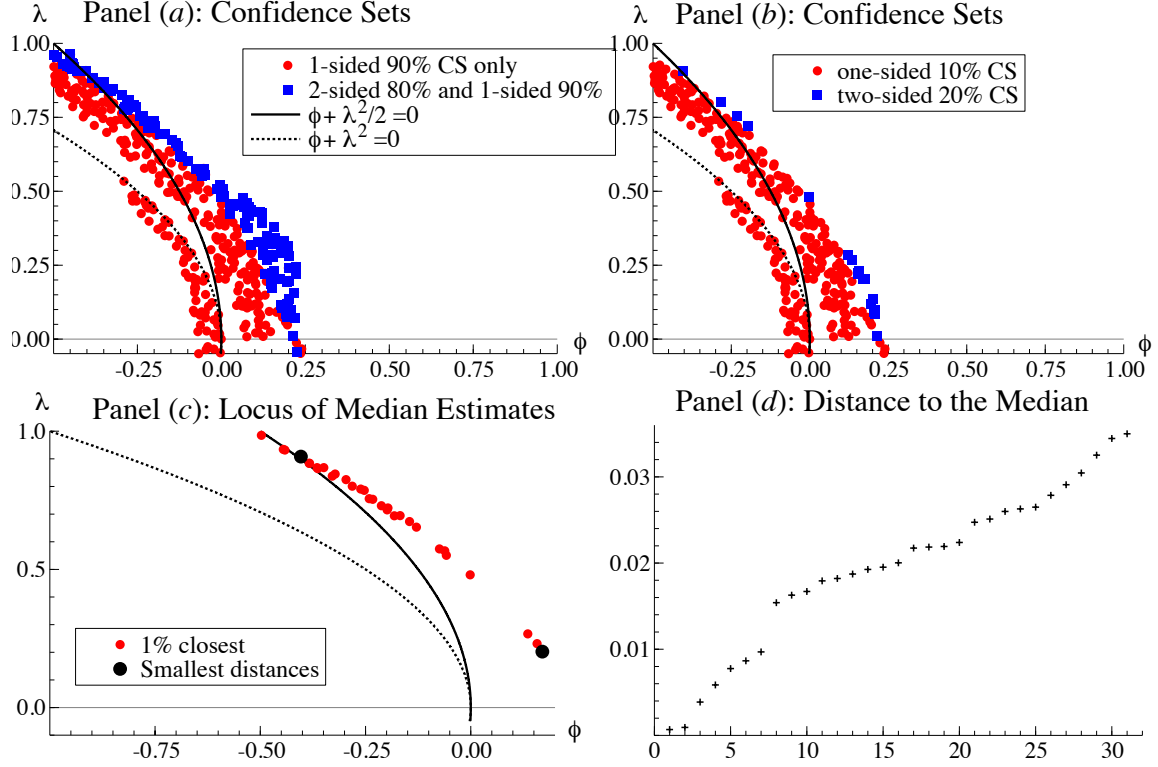
Figure 7: The figure reports inferential results on the NERC applied to ΔP_t. The top row records the confidence sets computed as parameter combinations which are not rejected using the asymptotic distribution ot the OLS estimator of $\widehat{\rho}$. Panel (a) : the dots define the 90% 1-sided (red disks) and 80% 2-sided (blue squares) confidence sets. Panel (b) : the dots define the 10% 1-sided (red disks) and 20% 2-sided (blue squares) confidence sets. The bottom row refers to the locus of median estimates (parameters implying $\widehat{\rho}$ is closest to the median of its asymptotic distribution). Panel (c) : the dots represent the set comprising the 1% parameters for which distance to the median is smallest (in red) as well as those whose distance is less than five times the smallest observed distance (in black). Panel (d) : the panel reports the 1% smallest observed distances in increasing order.
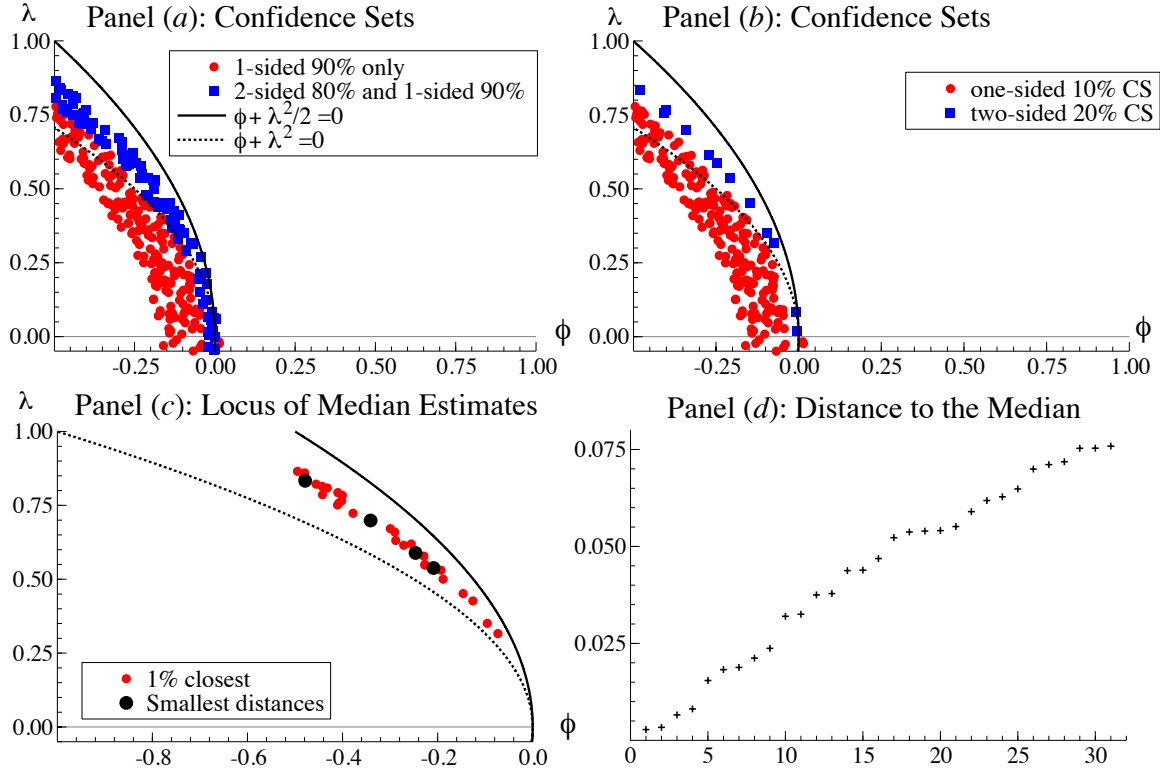
Figure 8: The figure reports inferential results on the NERC applied to P_t/D_t. The top row records the confidence sets computed as parameter combinations which are not rejected using the asymptotic distribution ot the OLS estimator of $\widehat{\rho}$. Panel (a) : the dots define the 90% 1-sided (red disks) and 80% 2-sided (blue squares) confidence sets. Panel (b) : the dots define the 10% 1-sided (red disks) and 20% 2-sided (blue squares) confidence sets. The bottom row refers to the locus of median estimates (parameters implying $\widehat{\rho}$ is closest to the median of its asymptotic distribution). Panel (c) : the dots represent the set comprising the 1% parameters for which distance to the median is smallest (in red) as well as those whose distance is less than five times the smallest observed distance (in black). Panel (d) : the panel reports the 1% smallest observed distances in increasing order.

	Least Rejected	Median Estimate	Univariate Confidence Interval at Probability			Test $(\phi,\lambda)=(0,0)$
	$(\phi,\lambda)^*$	$(\phi,\lambda)^+$	0.90	0.80	0.10	p-value
Upper-Tailed Tests						
$\Delta P_t, \widehat{\rho}=0.972$	$(.05,.01)$ $E\left[\rho_t^*\right]=1.00$ $p\text{-value}:1.00$	$(.17,.20)$ $E\left[\rho_t^+\right]=1.01$ $p\text{-value}:0.50$	$\lambda:[0,.99]$ $\phi:[-.50,.24]$	$\lambda:[0,.99]$ $\phi:[-.50,.24]$	$\lambda:[0,.93]$ $\phi:[-.50,.24]$	.95
$P_t/D_t, \widehat{\rho}=1.00$	$(.51,.02)$ $E\left[\rho_t^*\right]=1.03$ $p\text{-value}:1.00$	$(-.25,.59)$ $E\left[\rho_t^+\right]=1.00$ $p\text{-value}:0.51$	$\lambda:[0,.87]$ $\phi:[-.50,.01]$	$\lambda:[0,.84]$ $\phi:[-.50,.01]$	$\lambda:[0,.78]$ $\phi:[-.50,.01]$	.32
Two-Tailed Tests						
$\Delta P_t, \widehat{\rho}=0.972$	$(.17,.20)$ $E\left[\rho_t^*\right]=1.01$ $p\text{-value}:1.00$	$(.17,.20)$ $E\left[\rho_t^+\right]=1.01$ $p\text{-value}:1.00$	$\lambda:[0,.99]$ $\phi:[-.50,.22]$	$\lambda:[0.01,.99]$ $\phi:[-.50,.22]$	$\lambda:[0.01,.99]$ $\phi:[-.50,.22]$	.11
$P_t/D_t, \widehat{\rho}=1.00$	$(-.25,.59)$ $E\left[\rho_t^*\right]=1.00$ $p\text{-value}:.99$	$(-.25,.59)$ $E\left[\rho_t^+\right]=1.00$ $p\text{-value}:0.99$	$\lambda:[0,.87]$ $\phi:[-.50,.002]$	$\lambda:[0,.84]$ $\phi:[-.50,0]$	$\lambda:[0.01,.83]$ $\phi:[-.50,-.005]$	.64

Table 2: The table reports statistics regarding inference on the dynamics of ΔP_t and P_t/D_t using upper-tailed and two-tailed grid testing. The first two columns report the least rejected and median estimates using the asymptotic distributions of the test statistic. The next three columns present univariate confidence intervals obtained by projection of the bivariate confidence sets on either one-dimensional parameter space. The last column records the p-value associated with the test of a pure random walk.

| | Timing of the main bubble | |
Method	Inception	Burst
ΔP_t	April 2000	March 2006
D_t/P_t	May 2002	Nov 2007
PY	Feb 2002	Dec 2007

Table 3: The table report the dates of inception and burst of the main upward bubble as detected by the NERC(1) model for ΔP_t or D_t/P_t, and as reported by Phillips and Yu (2011), denoted PY.

Despite the large number of parameter combinations which are tested, we do not proceed to a Bonferroni correction and report, on panels (b) and (d) of the figure, the periods where the null is rejected at the 5% and 0.5% size respectively. The minimum p-value in the case of P_t/D_t is chosen much smaller as we follow PY who state that for consistent detection the nominal size must tend to zero with the sample size.[20] We follow PY in interpreting rejection as evidence of a bubble. The one-sided test for ΔP_t first detects a possible bubble at the 5% significance level in April 2000 and a turning point in March 2006. It also detects a downward bubble for the period October 2007 – November 2008 (with an exception of June 2008). Panel (d) reports the inception and termination for P_t/D_t (in May 2002 and November 2007). In their work, PY detect a bubble in house prices that starts in February 2002 and ends in December 2007. Our results based on P_t/D_t are comparable: we detect a slightly shorter bubble at the 0.5% level (they use 1%). Together with PY we find here a bubble that bursts slightly after the first evidence of the emerging subprime crisis (which they date to start in August 2007). Table 3 summarizes the dates of the main upward bubble detected using ΔP_t and P_t/D_t together with those reported by PY.

Interestingly, the results based in ΔP_t provide new evidence on the bubble: analysis based on this series provides evidence in changes in growth rates and hence may constitute an early detection device for turning points. The minimum p-value is much more volatile than for P_t/D_t and we observe that it first drops early, as soon as April 2000 (although it only settles at low values in March 2002). Hence there might have been tentative bubbles at play before one properly settled. In addition, inference based on ΔP_t detects the end of the bubble in March 2006, while the growth in prices was still positive: this appears to be the turning point in the bubble, before it properly burst, and we detect a negative bubble over the subprime crisis in late 2007 and throughout 2008.[21]

6.4 Forecasting

We finally apply our methodology to forecasting the probability of growth of the variable of interest. We recursively compute the minimum $\mathcal{P}_{t,k}^{y_t/y_{t-k}}$, i.e. minimum probabilities that growth over a

---

[20] Although our choices of significance levels seem *ad hoc* here, this should not concern us too much as they are based on the observation of a sharp drop in the minimum p-value. We leave considerations on the appropriate choice of significance level to further research.

[21] To the exception of June 2008: on the 11th, the Securities and Exchange Commission unveiled its comprehensive reform of credit ratings.
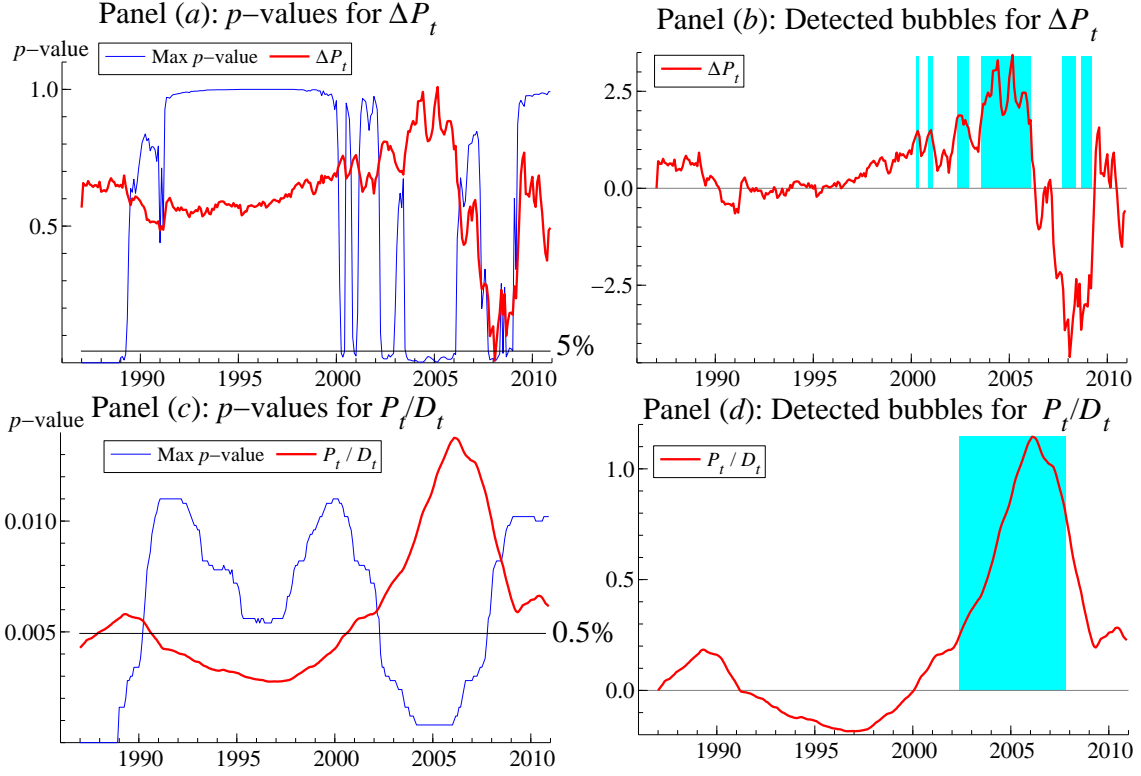
Figure 9: The figure reports the ouput from detection of explosive behavior in ΔP_t and P_t/D_t. The maximum p-values are computed as the maximum obtained over the asymptotic distributions corresponding to 5,000 uniform draws of the parameters (ϕ, λ) such that $\phi + \lambda^2/2 = 0$, i.e. $\mathsf{E}[\rho_t] = 1$. The test statistics are the scaled recursive OLS estimator estimated over the sample until the dates on the horizontal axis, with a minimum of 24 observations. Tests are one-sided. Panel (a) corresponds to inference based on ΔP_t. The p-values are reported on the left axis, together with a scaled series of ΔP_t. Panel (b) : The shaded areas refer to periods where the composite null $\mathsf{E}[\rho_t] = 1$ is rejected at the 5% significance level. Panels (c) and (d) report the equivalent of panels (a) and (b) for the series of P_t/D_t, the significance level is 0.5%.

forecast horizon of $k = 1, 6$, or 12 months will be as high as that observed over the latest k periods. The minimum probabilities are obtained over the set of parameters (ϕ, λ) such that $c = \phi + \lambda^2 \geq 0$. We only consider parameter sets which constitute the nominal 50% or 90% asymptotic confidence set.[22] The confidence sets for these minimum probabilities are reported in figures 10 and 11.

Minimum probability forecasts relating to ΔP_t are volatile and only nonzero at the beginning and end of the sample; they do not appear easily interpretable and somewhat inaccurate. By contrast, those relating to P_t/D_t are interesting: Figure 11 records three periods where minimum probabilities are strictly positive in the 50% confidence set. The first period (1991-early 1992) corresponds to decreasing or stable P_t/D_t ratio. The second starts in January 2000 and its duration depends on the horizon considered. The bottom panel in the figure shows that this corresponds to the period where the growth in P_t/D_t started accelerating, before subsiding until the third period (March 2004 - February 2006) where growth picked up. The latter period only appears when considering the narrower confidence set (50%) and ends too late at longer horizons.

7 Conclusion

The paper proposes a local asymptotic model that builds on random coefficient autoregressive processes and shows how this NERC model can be applied to the modelling of asset prices.

We show that the process generated by a NERC converges towards the stochastic integral of a geometric Brownian motion. We also derive the asymptotic distributions of OLS estimators of the first-order autocorrelation coefficient. We then provide a technique of inference on the parameters of the process based on inverting a test statistic.

As with some existing models for bubbles, the presence of a random coefficient introduces flexibility in the modelling of multiple bubbles. Here, bubbles may – or not – appear, and by avoiding regime switching, we do not imply that they periodically do. Instead, their emergence depends on the values taken by a latent process that relates to the stochastic discount factor. The generalization we propose presents benefits that are similar to the univariate locally explosive AR(1) with breaks of Phillips, Wu and Yu (2011), while allowing for full-sample inference. Also the flexible model we propose allows the so-called bubbles either to reflect nonstationary behavior or be caused by large deviations within a strictly stationary model.

Under the NERC DGP, it is also possible to provide density forecasts and hence to establish statements on the probability of bubbles. Finally, we apply our methodology to U.S. data on house prices and show how our method compares to the existing literature. In particular, we show how a simple model for the change in prices can help anticipating turning points in bubble dynamics.

Possible extensions of the model we propose comprise considering multivariate models where a unique latent process may be causing bubbles that spill over into different markets (as in Phillips and Yu, 2011). This might require relaxing the assumption that u_t is $i.i.d.$ In turn, it would

---

[22]Minimum probabilities based on the 10% to 50% confidence sets are almost identical, and so are those in the 80%-99% range.
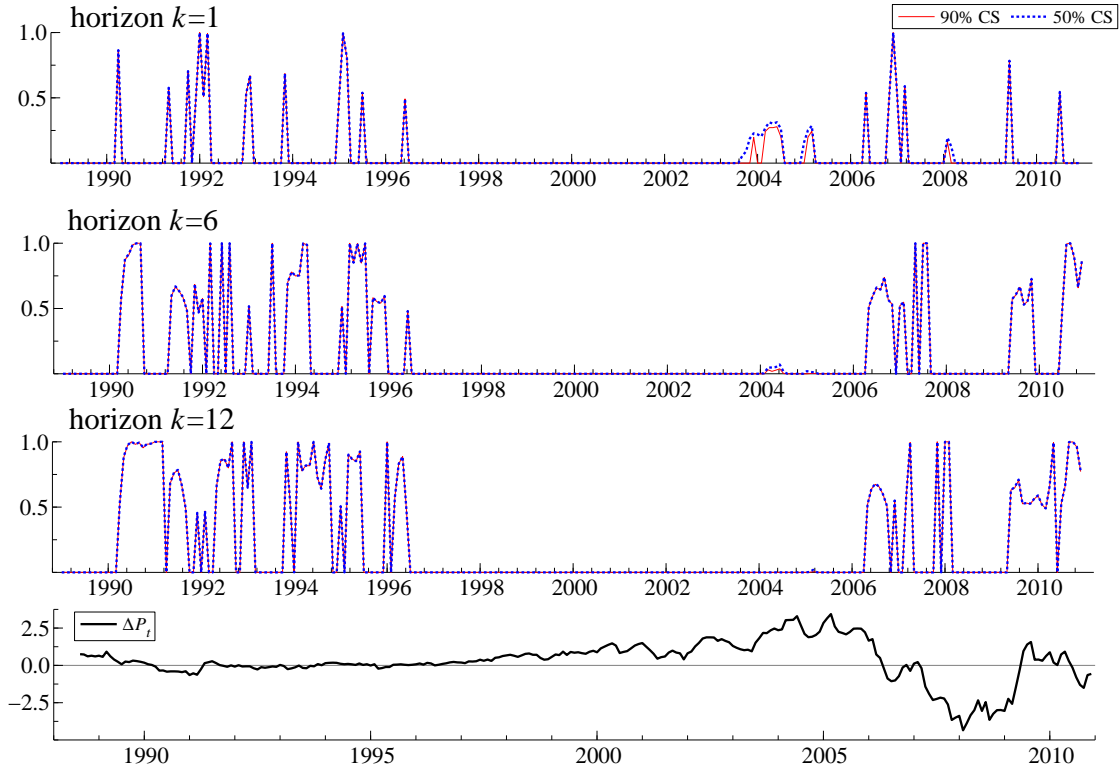
Figure 10: The figure reports the estimated minimum probabilities $\mathsf{P}\left(y_{t+k}/y_t \geq y_t/y_{t-k}\right)$ computed under the asymptotic distribution under the null. The minimum is computed over 5,000 draws of parameters (ϕ, λ) such that $c = \phi + \lambda^2 \geq 0$ and (ϕ, λ) belong to the set of parameters in the nominal 90% (in blue) or 50% (in red) confidence set (under the asymptotic distribution). All panels correspond to the case where the series of interest, y_t, is ΔP_t. Each panel reports a different horizon: $k = 1$ (top panel), 6 (middle) or 12 months (bottom).

be possible to filter out an estimate of the latent process u_t or the stochastic discount factor. Alternatively, our model allows simply to postulate and test a candidate variable for u_t.
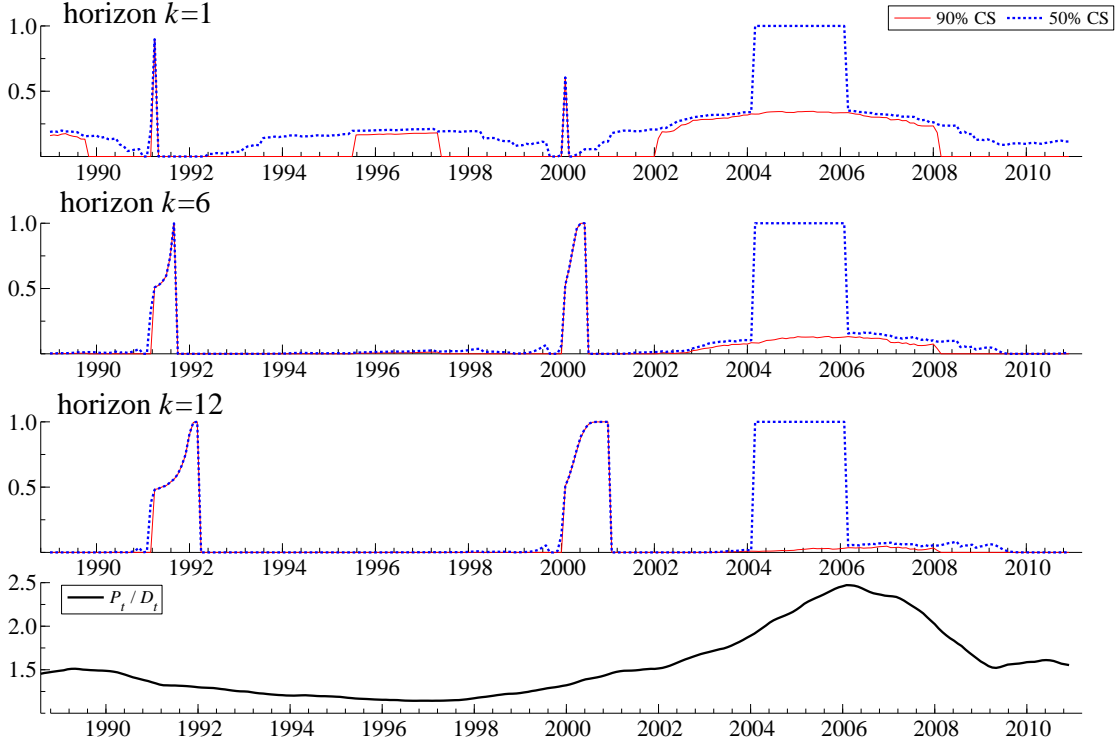
Figure 11: The figure reports the estimated minimum probabilities $\mathsf{P}\left(y_{t+k}/y_t \geq y_t/y_{t-k}\right)$ computed under the asymptotic distribution under the null. The minimum is computed over 5,000 draws of parameters (ϕ, λ) such that $c = \phi + \lambda^2 \geq 0$ and (ϕ, λ) belong to the set of parameters in the nominal 90% (in blue) or 50% (in red) confidence set (under the asymptotic distribution). All panels correspond to the case where the series of interest, y_t, is P_t/D_t. Each panel reports a different horizon: $k = 1$ (top panel), 6 (middle) or 12 months (bottom).

Appendix

A   Proof of proposition 1

We have, given y_0, and setting $\prod_{j=0}^{-1} \rho_j \equiv 1$

$$
\begin{aligned}
y_t &= \left(\prod_{j=0}^{t-1} \rho_{t-j}\right) y_0 + \sum_{i=0}^{t-1} \left(\prod_{j=0}^{i-1} \rho_{t-j}\right) \eta_{t-i} \\
&= \left(\prod_{i=1}^{t} \rho_i\right) y_0 + \sum_{i=1}^{t} \left(\prod_{j=i+1}^{t} \rho_j\right) \eta_i \\
&= \exp\left\{\frac{tT^{-\alpha/2}\phi + \lambda S_t}{T^{\alpha/2}}\right\} y_0 + \sum_{i=1}^{t} \exp\left\{\frac{(t-i)T^{-\alpha/2}\phi + \lambda(S_t - S_i)}{T^{\alpha/2}}\right\} \eta_i.
\end{aligned}
$$

We evaluate the increment $y_t - y_0$ using the blocking method of Phillips and Magdalinos (2004, 2007). Letting, for $t = 1$ to T, $t = \lfloor jT^\alpha \rfloor + k$ ($\lfloor x \rfloor$ denoting the integer part of x) for $j = 0, \cdots, \lfloor T^{1-\alpha} \rfloor - 1$, and $k = 1, \cdots, \lfloor T^\alpha \rfloor$, and letting $k = \lfloor pT^\alpha \rfloor$ for some $p \in [0,1]$, we can write

$$\frac{1}{T^{\alpha/2}} (y_{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor} - y_0)$$

$$= \frac{1}{T^{\alpha/2}} \left(\exp \left\{ \frac{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor}{T^\alpha} \phi + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor}}{T^{\alpha/2}} \right\} - 1 \right) y_0$$

$$+ \sigma_\eta \sum_{i=1}^{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor} \exp \left\{ \frac{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor - i}{T^\alpha} \phi + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor} - S_i}{T^{\alpha/2}} \right\} \frac{\eta_i}{\sqrt{\sigma_\eta^2 T^\alpha}}$$

$$= \frac{1}{T^{\alpha/2}} \left(\exp \left\{ \frac{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor}{T^\alpha} \phi + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor}}{T^{\alpha/2}} \right\} - 1 \right) y_0$$

$$+ \sigma_\eta \int_0^{j+p} \exp \left\{ \frac{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor - \lfloor sT^\alpha \rfloor}{T^\alpha} \phi + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor} - S_{\lfloor sT^\alpha \rfloor}}{T^{\alpha/2}} \right\} dB_{T^\alpha}(s),$$

using Proposition A1 in Phillips and Magdalinos (2004) in the last equality, where

$$B_{T^\alpha}(s) \equiv \frac{1}{\sigma_\eta T^{\alpha/2}} \sum_{i=1}^{\lfloor sT^\alpha \rfloor} \eta_i. \tag{A.1}$$

When applying the Functional Central Limit Theorem (FCLT) to the process $\tilde{S}_T$ defined by $\tilde{S}_T(s) \equiv T^{-\alpha/2} S_{\lfloor sT^\alpha \rfloor}$ ($0 \le s \le 1$), we obtain that $\tilde{S}_T$ converges in distribution, as $T \to \infty$, to a Brownian motion (BM) on $[0,1]$ that we denote by W.

The FLCT also implies that the process B_{T^α} defined in (A.1) converges in distribution, as $T \to \infty$, to a BM on $[0,1]$, say B, which, by assumption on the sequences (u_i) and (η_j), is independent of W.

Then we can deduce, using e.g. Theorem 8.3.1 in Liptser and Shiryaev (1989), that

$$\int_0^{j+p} \exp \left\{ \phi \frac{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor - \lfloor sT^\alpha \rfloor}{T^\alpha} + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + \lfloor pT^\alpha \rfloor} - S_{\lfloor sT^\alpha \rfloor}}{T^{\alpha/2}} \right\} dB_{T^\alpha}(s)$$

converges, as $T \to \infty$, to

$$\int_0^r \exp \left\{ \phi(r - s) + \lambda (W_r - W_s) \right\} dB_s, \quad \text{with} \quad r = j + p.$$

Corollary 2 follows since the proof above also holds when $\alpha = 1$. $\qquad\square$

B Proof of Theorem 4

We have, as $T \to \infty$,

$$y_t = \left(\mathsf{E}(\rho_t) + \lambda T^{-\alpha/2} u_t + O_p \left(T^{-\alpha} \right) \right) y_{t-1} + \eta_t.$$

Then the OLS estimator given by $\widehat{\rho} = \dfrac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2}$ satisfies

$$\widehat{\rho} - \mathsf{E}(\rho_t) = \lambda T^{-\alpha/2} \frac{\sum_t y_{t-1}^2 u_t}{\sum_t y_{t-1}^2} + \frac{\sum_t y_{t-1} \eta_t}{\sum_t y_{t-1}^2} + O_p \left(T^{-3\alpha/2} \right). \tag{B.2}$$

This shows that the asymptotic distribution of the estimator is driven by the term with higher magnitude between $T^{-\alpha/2} \sum_t y_{t-1}^2 u_t$ and $\sum_t y_{t-1}\eta_t$.

So we need to study the three sums appearing in the expression of the OLS estimator. Throughout we assume $y_0 = 0$ without loss of generality as our assumption that $y_0 = o_p\left(T^{\alpha/2}\right)$ implies it is neglible compare to $y_{\lfloor rT^\alpha \rfloor}$ for $0 < r < T^{1-\alpha}$.

Recall that $c = \phi + \lambda^2$. We will consider different cases depending on the sign of c.

B.1 Case $c < 0$

Proposition 1 gives $T^{-\alpha/2} y_{\lfloor rT^\alpha \rfloor} \Rightarrow K_{\phi,\lambda}(r) \sim \mathsf{N}\left(0, \dfrac{e^{2cr}-1}{2c}\sigma_\eta^2\right)$;

for $c < 0$, we can write $K_{\phi,\lambda}^*(r) = e^{cr}K_{\phi,\lambda}^*(0) + K_{\phi,\lambda}(r)$. Hence $K_{\phi,\lambda}^*(r) \sim \mathsf{N}\left(0, -\frac{\sigma_\eta^2}{2c}\right)$ and is stationary. We can deduce, via the Law of Large Numbers (LLN), that

$$\frac{1}{T^{1+\alpha}}\sum_{t=1}^{T} y_t^2 \Rightarrow \mathsf{E}[K_{\phi,\lambda}^*(r)^2] = \frac{-\sigma_\eta^2}{2c}$$

and that

$$T^{-\frac{1+\alpha}{2}}\sum_{t=1}^{T} y_{t-1}\eta_t \Rightarrow \mathsf{N}\left(0, -\frac{\sigma_\eta^4}{2c}\right)$$

The result concerning $\sum_{t=1}^{T} y_{t-1}^2 u_t$ similarly follows. Indeed, define the martingale difference sequence $\xi_t = T^{-\frac{1+2\alpha}{2}} y_{t-1}^2 u_t$ which admits conditional variance satisfying

$$\sum_{t=1}^{T} \mathsf{E}_{t-1}\left(\xi_t^2\right) = \frac{1}{T^{1+2\alpha}}\sum_{t=1}^{T} y_{t-1}^4 \Rightarrow \frac{3\sigma_\eta^4}{4c^2},$$

using the consistency of the empirical estimator of the kurtosis. A martingale analogue of the Lindeberg condition (see e.g. Pollard, 1984) ensures then that

$$T^{-\frac{1+2\alpha}{2}}\sum_{t=1}^{T} y_{t-1}^2 u_t \Rightarrow \mathsf{N}\left(0, \frac{3\sigma_\eta^4}{4c^2}\right).$$

B.2 Case $c \geq 0$

The proof follows the main schemata given in Phillips & Magdalinos (2004); hence we keep their notation to help the reader, and set $T = n$,

$$\kappa_n = n^\alpha \left\lfloor n^{1-\alpha} \right\rfloor \quad \text{and} \quad q = n^{1-\alpha} - \left\lfloor n^{1-\alpha} \right\rfloor,$$

Sample variance We first consider the sample variance of y_t and show the following.

Lemma 7 *Define*

$$
\phi_{n^{1-\alpha}} = \begin{cases} \frac{e^{2cn^{1-\alpha}}-1}{2c}, & \text{if } c \neq 0; \\ n^{1-\alpha}, & \text{if } c = 0; \end{cases}
$$

$$
\psi_{n^{1-\alpha}} = \begin{cases} \frac{1}{2\sqrt{(c+2\lambda^2)(c+\lambda^2)}} e^{2(c+\lambda^2)n^{1-\alpha}}, & \text{if } c \neq 0; \\ \frac{e^{2\lambda^2 n^{1-\alpha}}}{2\sqrt{2}\lambda^2}, & \text{if } c = 0; \end{cases}
$$

$$
\text{and} \quad \varphi_{n^{1-\alpha}} = \begin{cases} \frac{1}{2(\phi-\lambda^2)}, & \text{if } \lambda^2 - \phi < 0; \\ \sqrt{n^{1-\alpha}}, & \text{if } \lambda^2 - \phi = 0; \\ \frac{e^{(\lambda^2-\phi)n^{1-\alpha}}}{\sqrt{2(\lambda^2-\phi)}}, & \text{if } \lambda^2 - \phi > 0. \end{cases}
$$

Then, as $n \to \infty$,

$$
\sigma_\eta^{-2} n^{-2\alpha} \psi_{n^{1-\alpha}}^{-1} \varphi_{n^{1-\alpha}}^{-2} \sum_{t=1}^{n} y_t^2 \Rightarrow X^2 Z,
$$

where the random variables X and Z are defined, respectively, by

$$
\sigma_\eta^{-1} \frac{\varphi_{\lfloor n^{1-\alpha} \rfloor}^{-1}}{n^{\alpha/2}} \sum_{i=1}^{\lfloor \kappa_n \rfloor} \exp\left(-\frac{\phi}{n^\alpha} i - \frac{\lambda}{n^{\alpha/2}} U_i \right) \eta_i \Rightarrow X \sim \mathsf{N}(0,1).
$$

and

$$
\psi_{\lfloor n^{1-\alpha} \rfloor}^{-1} \left(\int_0^{\lfloor n^{1-\alpha} \rfloor} e^{2\phi s + 2\lambda W_s} ds \right) \Rightarrow Z,
$$

with mean 0 and unit variance.

Proof of Lemma 7.

We write

$$
\frac{1}{n^{2\alpha}} \sum_{t=1}^{n} y_t^2 = \frac{1}{n^{2\alpha}} U_{1n} + \frac{1}{n^{2\alpha}} U_{2n} + O_p\left(\frac{1}{n^\alpha} \right), \tag{B.3}
$$

with

$$
U_{1n} \equiv \sum_{j=0}^{\lfloor n^{1-\alpha} \rfloor - 1} \sum_{k=1}^{\lfloor n^\alpha \rfloor} y_{\lfloor n^\alpha j \rfloor + k}^2,
$$

$$
\text{and} \quad U_{2n} \equiv \sum_{t=\lfloor \kappa_n \rfloor}^{n} y_t^2.
$$

Note that the index of the last summation term in the definition of U_{1n}, given by $\lfloor \kappa_n - n^\alpha \rfloor + \lfloor n^\alpha \rfloor$, is bounded by $\lfloor \kappa_n \rfloor - 1 \leq \lfloor \kappa_n - n^\alpha \rfloor + \lfloor n^\alpha \rfloor \leq \lfloor \kappa_n \rfloor$.

The study of U_{1n} leads to the following result.
As $n \to \infty$,

$$
\sigma_\eta^{-2} n^{-2\alpha} \psi_{\lfloor n^{1-\alpha} \rfloor}^{-1} \varphi_{\lfloor n^{1-\alpha} \rfloor}^{-2} U_{1n} \Rightarrow X^2 Z, \tag{B.4}
$$

34

where the random variables X and Z are defined in Lemma 7.

Proof of (B.4).

Notice that

$$
\begin{aligned}
y_k &= \sum_{i=0}^{k-1} \exp\left(\frac{\phi}{n^\alpha} i + \frac{\lambda}{n^{\alpha/2}} \sum_{j=k-i+1}^{t} u_j \right) \eta_{k-i} \\
&= \sum_{i=1}^{k} \exp\left(\frac{\phi}{n^\alpha}(k-i) + \frac{\lambda}{n^{\alpha/2}}(U_k - U_i) \right) \eta_i \qquad \text{with} \quad U_i \equiv \sum_{j=1}^{i} u_j \\
&= \exp\left(\frac{\phi}{n^\alpha}k + \frac{\lambda}{n^{\alpha/2}}U_k \right) \sum_{i=1}^{k} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i,
\end{aligned}
$$

so

$$
\begin{aligned}
\sum_{k=1}^{t} y_k^2 &= \sum_{k=1}^{t} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k \right) \left[\sum_{i=1}^{k} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \right]^2 \\
&= \sum_{k=1}^{t} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k \right) \\
&\qquad \times \left[\sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i - \sum_{i=k+1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \right]^2 \\
&= \left(\sum_{k=1}^{t} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k \right) \right) \left[\sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \right]^2 + R_t
\end{aligned}
$$

where

$$
\begin{aligned}
R_t &= \sum_{k=1}^{t} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k \right) \left[\sum_{i=1}^{k} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \right]^2 \\
&\quad - \left(\sum_{k=1}^{t} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k \right) \right) \left[\sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \right]^2 .
\end{aligned}
$$

Therefore we obtain

$$
U_{1n} = \left(\sum_{k=1}^{\lfloor \kappa_n \rfloor} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k \right) \right) \left[\sum_{i=1}^{\lfloor \kappa_n \rfloor} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \right]^2 + R_{\lfloor \kappa_n \rfloor}. \tag{B.5}
$$

Let us study the three elements given in this last equation (B.5).

We start by showing that $n^{-2\alpha} \psi_{\lfloor n^{1-\alpha} \rfloor}^{-1} \varphi_{\lfloor n^{1-\alpha} \rfloor}^{-2} R_{\lfloor \kappa_n \rfloor}$ is negligible w.r.t. $\psi_{\lfloor n^{1-\alpha} \rfloor}^{-1} \varphi_{\lfloor n^{1-\alpha} \rfloor}^{-2} n^{-2\alpha} (U_{1n} + U_{2n})$.

Indeed, writing $R_t = R_{1t} - 2R_{2t}$ with

$$R_{1t} = \sum_{k=1}^{t} \left[\sum_{i=k+1}^{t} \exp\left(\frac{\phi}{n^\alpha}(k-i) - \frac{\lambda}{n^{\alpha/2}}(U_k - U_i) \right) \eta_i \right]^2$$

$$R_{2t} = \left(\sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \right) \times \sum_{k=1}^{t} \sum_{i=k+1}^{t} \exp\left(\frac{\phi}{n^\alpha}(2k-i) + \frac{\lambda}{n^{\alpha/2}}(2U_k - U_i) \right) \eta_i$$

$$\equiv \left(\sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \right) \times \overline{R}_{2t},$$

then, on one hand,

$$n^{-2\alpha}\,\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\varphi_{\lfloor n^{1-\alpha}\rfloor}^{-2}R_{1\lfloor \kappa_n\rfloor} = \psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\int_0^{\lfloor n^{1-\alpha}\rfloor} \left(\sigma_\eta \varphi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\int_r^{\lfloor n^{1-\alpha}\rfloor} e^{\phi(r-s)+\lambda(W_r-W_s)}dB_s \right)^2 dr + o_p(1)$$

uniformly, and since we have

$$\mathsf{E}\left(\int_r^{\lfloor n^{1-\alpha}\rfloor} e^{\phi(r-s)-\lambda(W_r-W_s)}dB_s \right)^2 = \int_r^{\lfloor n^{1-\alpha}\rfloor} e^{2(\phi+\lambda^2)(r-s)}ds$$

$$= \frac{1 - e^{-2(\phi+\lambda^2)(\lfloor n^{1-\alpha}\rfloor - r)}}{2(\phi+\lambda^2)} = O(1)$$

uniformly in $r \leq n^{1-\alpha}$, we obtain

$$n^{-2\alpha}\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\varphi_{\lfloor n^{1-\alpha}\rfloor}^{-2}R_{1\lfloor \kappa_n\rfloor} = O_p(1) \times O\left(\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\int_0^{\lfloor n^{1-\alpha}\rfloor} dr \right) = O_p\left(\lfloor n^{1-\alpha}\rfloor \psi_{\lfloor n^{1-\alpha}\rfloor}^{-1} \right).$$

On the other hand, we have $n^{-\alpha}\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\overline{R}_{2\lfloor \kappa_n\rfloor} = O_p\left(\lfloor n^{1-\alpha}\rfloor \psi_{\lfloor n^{1-\alpha}\rfloor}^{-1} \right)$, so it comes

$$n^{-2\alpha}\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\varphi_{\lfloor n^{1-\alpha}\rfloor}^{-2}R_{2\lfloor \kappa_n\rfloor} = O_p\left(n^{1-3/2\alpha}\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\varphi_{\lfloor n^{1-\alpha}\rfloor}^{-1} \right),$$

hence the result concerning $R_{\lfloor \kappa_n\rfloor}$.

Now let us look at the second element on the RHS of equation (B.5). We can write it as

$$\frac{1}{n^{\alpha/2}} \sum_{i=1}^{\lfloor \kappa_n\rfloor} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i = \sigma_\eta \int_0^{\lfloor n^{1-\alpha}\rfloor} e^{-\phi n^\alpha s - \lambda W_{n^\alpha s}}dB_{n^\alpha}(s) + o_p(1),$$

B_{n^α} being defined in (A.1).

When $\lambda^2 < \phi$, it admits limit $\sigma_\eta \int_0^\infty e^{-\phi s - \lambda W_s}dB(s)$.

When $\lambda^2 \geq \phi$, the stochastic integral is not defined, but since $\int_0^{\lfloor n^{1-\alpha}\rfloor} e^{-(\phi s + \lambda W_s)}dB_s$ is normally distributed, it will be enough to scale it by its standard deviation, using that

$$\mathsf{V}\left[\int_0^{\lfloor n^{1-\alpha}\rfloor} e^{-(\phi s + \lambda W_s)}dB_s \right] = \begin{cases} \frac{e^{2(\lambda^2-\phi)\lfloor n^{1-\alpha}\rfloor}-1}{2(\lambda^2-\phi)}, & \text{if } \lambda^2 > \phi; \\ \lfloor n^{1-\alpha}\rfloor, & \text{if } \lambda^2 = \phi; \\ \frac{1-e^{-2(\phi-\lambda^2)\lfloor n^{1-\alpha}\rfloor}}{2(\phi-\lambda^2)}, & \text{if } \lambda^2 < \phi. \end{cases} \tag{B.6}$$

Hence we obtain that

$$\sigma_\eta^{-1}\varphi_{\lfloor n^{1-\alpha}\rfloor}^{-1}n^{-\alpha/2}\sum_{i=1}^{\lfloor \kappa_n\rfloor}\exp\left(-\frac{\phi}{n^\alpha}i-\frac{\lambda}{n^{\alpha/2}}U_i\right)\eta_i \Rightarrow X \sim \mathsf{N}\,(0,1)\,.$$

Finally, let us look at the first element on the RHS of (B.5). We have

$$\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}n^{-\alpha}\sum_{k=1}^{\lfloor \kappa_n\rfloor}\exp\left(\frac{2\phi}{n^\alpha}k+\frac{2\lambda}{n^{\alpha/2}}U_k\right)=\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\left(\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2\phi s+2\lambda W_s}ds\,-\,\phi_{\lfloor n^{1-\alpha}\rfloor}\right)+o_p\,(1)\,,$$

$\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\phi_{\lfloor n^{1-\alpha}\rfloor}$ tending to 0 as $n\to\infty$.

Note that the expectation of $\displaystyle\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2(\phi s+\lambda W_s)}ds$ is given by

$$\phi_{\lfloor n^{1-\alpha}\rfloor}\equiv\mathsf{E}\left[\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2\phi s+2\lambda W_s}ds\right]=\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2cs}ds=\begin{cases}\frac{e^{2c\lfloor n^{1-\alpha}\rfloor}-1}{2c},&\text{if }c\neq 0;\\[2mm]\lfloor n^{1-\alpha}\rfloor,&\text{if }c=0;\end{cases}$$

and that the rate $\psi_{\lfloor n^{1-\alpha}\rfloor}$ comes from the second moment of $\displaystyle\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2(\phi s+\lambda W_s)}ds.$
Indeed, straightforward computations lead, for $c\geq 0$, to

$$\mathsf{E}\left[\left(\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2(\phi s+\lambda W_s)}ds\right)^2\right]=\mathsf{E}\left[\left(\int_0^{\lfloor n^{1-\alpha}\rfloor}\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2\phi(s+r)+2\lambda(W_s+W_r)}dsdr\right)\right]$$

$$=\int_0^{\lfloor n^{1-\alpha}\rfloor}\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2\phi(s+r)+2\lambda^2(s+2\min(r,s)+r)}dsdr$$

$$=\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2cr}\int_0^r e^{2(c+2\lambda^2)s}dsdr+\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2(c+2\lambda^2)r}\int_r^{\lfloor n^{1-\alpha}\rfloor}e^{2cs+2(c+2\lambda^2)r}dsdr$$

$$=\begin{cases}\frac{e^{4(c+\lambda^2)\lfloor n^{1-\alpha}\rfloor}}{4(c+2\lambda^2)(c+\lambda^2)}-\frac{e^{2cn^{1-\alpha}}}{2c(c+2\lambda^2)}+\frac{1}{4c(c+\lambda^2)}=\frac{e^{4(c+\lambda^2)\lfloor n^{1-\alpha}\rfloor}}{4(c+2\lambda^2)(c+\lambda^2)}+O\left(e^{2cn^{1-\alpha}}\right),&\text{if }c\neq 0;\\[3mm]\frac{e^{4\lambda^2\lfloor n^{1-\alpha}\rfloor}}{8\lambda^4}+O\left(n^{1-\alpha}\right),&\text{if }c=0.\end{cases}$$

Now, Matsumoto and Yor (2005), Theorem 7.4, implies that $\left(\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2(\phi s+\lambda W_s)}ds\right)^2$ divided by its expectation converges weakly to a random variable as $n\to\infty$. The continuous mapping theorem implies that the square root thereof also admits a weak limit. Hence there exists Z with unit variance and zero expectation such that

$$\psi_{\lfloor n^{1-\alpha}\rfloor}^{-1}\left(\int_0^{\lfloor n^{1-\alpha}\rfloor}e^{2(\phi s+\lambda W_s)}ds-\phi_{\lfloor n^{1-\alpha}\rfloor}\right)\;\Rightarrow\;Z.$$

The result of (B.4) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Let us now consider the second term U_{2n} in equation (B.3). We have

$$\frac{1}{n^{2\alpha}}U_{2n}=\frac{1}{n^{2\alpha}}\sum_{j=0}^{n-[\kappa_n]}y_{j+[\kappa_n]}^2=\int_0^q\left(\frac{1}{n^{\alpha/2}}y_{[\kappa_n]+[n^\alpha p]}\right)^2dp+O_p\left(n^{-2\alpha}\right),$$

where, for all $j = 0, ..., \lfloor n^{1-\alpha} \rfloor - 1$, as $n \to \infty$,

$$n^{-\alpha/2} y_{[n^a j] + [n^\alpha p]} \Rightarrow \sigma_\eta \int_0^{j+p} e^{\phi(j+p-s) + \lambda(W_{j+p} - W_s)} dB_s$$

$$= \sigma_\eta e^{\phi(j+p) + \lambda W_{j+p}} \int_0^{j+p} e^{-\phi s - \lambda W_s} dB_s.$$

Then it comes

$$\frac{1}{n^{2\alpha}} U_{2n} = \int_0^q e^{2\phi\left(\lfloor n^{1-\alpha} \rfloor + p\right) + 2\lambda W_{\lfloor n^{1-\alpha} \rfloor + p}} \left(\sigma_\eta \int_0^{\lfloor n^{1-\alpha} \rfloor + p} e^{-\phi s - \lambda W_s} dB_s \right)^2 dp + o_p(1)$$

$$= \left(\sigma_\eta \int_0^{\lfloor n^{1-\alpha} \rfloor + q} e^{-\phi s - \lambda W_s} dB_s \right)^2 \int_0^q e^{2\left(\phi\left(\lfloor n^{1-\alpha} \rfloor + s\right) + \lambda W_{\lfloor n^{1-\alpha} \rfloor + s}\right)} ds + o_p(1)$$

$$= \left(\sigma_\eta \int_0^{\lfloor n^{1-\alpha} \rfloor + q} e^{-\phi s - \lambda W_s} dB_s \right)^2$$

$$\times \left(\int_0^{\lfloor n^{1-\alpha} \rfloor + q} e^{2(\phi s + \lambda W_s)} ds - \int_0^{\lfloor n^{1-\alpha} \rfloor} e^{2(\phi s + \lambda W_s)} ds \right) + o_p(1),$$

hence

$$U_{2n} = \left(\sigma_\eta \int_0^{n^{1-\alpha}} e^{-\phi s - \lambda W_s} dB_s \right)^2 \int_0^{n^{1-\alpha}} e^{2(\phi s + \lambda W_s)} ds \tag{B.7}$$

$$- \frac{1}{n^{2\alpha}} \left(\frac{\int_0^{n^{1-\alpha}} e^{-\phi s - \lambda W_s} dB_s}{\int_0^{\lfloor n^{1-\alpha} \rfloor} e^{-\phi s - \lambda W_s} dB_s} \right)^2 U_{1n} + o_p(1). \tag{B.8}$$

Combining (B.3), (B.4), (B.7), and the asymptotic equivalence $\psi_{n^{1-\alpha}}^{-1} \psi_{\lfloor n^{1-\alpha} \rfloor} = 1 + o(1)$ allows to conclude to Lemma 7.

Sample covariances Now consider the covariance terms.

Lemma 8 *We have, as $n \to \infty$,*

$$\sigma_\eta^{-2} n^{-\alpha} \varphi_{n^{1-\alpha}}^{-1} \phi_{n^{1-\alpha}}^{-1} \sum_{t=1}^n y_{t-1} \eta_t \Rightarrow XY,$$

and

$$\sigma_\eta^{-2} \varphi_{n^{1-\alpha}}^{-2} \chi_{n^{1-\alpha}}^{-1} n^{-3\alpha/2} \sum_{t=1}^n y_{t-1}^2 u_t \Rightarrow X^2 V,$$

where $X \sim \mathsf{N}(0,1)$, $Y \sim \mathsf{N}(0,1)$, $V \sim \mathsf{N}(0,1)$,
and with $(\varphi_{n^{1-\alpha}}, \phi_{n^{1-\alpha}})$ defined in lemma 7, and

$$\chi_{n^{1-\alpha}} = \frac{e^{2(c+\lambda^2)n^{1-\alpha}}}{2\sqrt{c+\lambda^2}}.$$

Proof of Lemma 8.

Note that

$$\mathsf{V}\left(\int_0^{n^{1-\alpha}} e^{\phi s + \lambda W_s} dB_s\right) = \mathsf{E}\left[\left(\int_0^{n^{1-\alpha}} e^{\phi s + \lambda W_s} dB_s\right)^2\right]$$

$$= \mathsf{E}\left[\int_0^{n^{1-\alpha}} e^{2\phi s + 2\lambda W_s} ds\right] = \phi_{n^{1-\alpha}},$$

using in this last equality the computation of $\phi_{n^{1-\alpha}}$ made in the proof of Lemma 7.
So we have

$$\phi_{n^{1-\alpha}}^{-1} \int_0^{n^{1-\alpha}} e^{\phi s + \lambda W_s} dB_s \ \Rightarrow \ Y \sim \mathsf{N}\left(0,1\right).$$

Hence we can write

$$\frac{\varphi_{n^{1-\alpha}}^{-1} \phi_{n^{1-\alpha}}^{-1}}{n^\alpha} \sum_{t=1}^{n} y_{t-1}\eta_t = \left(\sigma_\eta \varphi_{n^{1-\alpha}}^{-1} \int_0^{n^{1-\alpha}} e^{-(\phi s + \lambda W_s)} dB_s\right)\left(\sigma_\eta \phi_{n^{1-\alpha}}^{-1} \int_0^{n^{1-\alpha}} e^{\phi r + \lambda W_r} dB_r\right) + I_n,$$

where I_n can be shown to be negligible, referring to Phillips and Magdalinos (2004). Then

$$\sigma_\eta^{-2} \varphi_{n^{1-\alpha}}^{-1} \phi_{n^{1-\alpha}}^{-1} n^{-\alpha} \sum_{t=1}^{n} y_{t-1}\eta_t \Rightarrow XY.$$

Now let us consider $\sum_{t=1}^{n} y_{t-1}^2 u_t$. It can be expressed as

$$\sum_{t=0}^{n-1} y_t^2 u_{t+1} = \sum_{t=1}^{n-1} \exp\left(\frac{2\phi}{n^\alpha}t + \frac{2\lambda}{n^{\alpha/2}}U_t\right)\left[\sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i\right)\eta_i\right]^2 (U_{t+1} - U_t)$$

$$= \left(\sum_{k=1}^{n-1} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k\right)(U_{k+1} - U_k)\right)\left[\sum_{i=1}^{n} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i\right)\eta_i\right]^2 + R_t^*, \quad \text{(B.9)}$$

where

$$R_t^* = \sum_{t=0}^{n-1} \exp\left(\frac{2\phi}{n^\alpha}t + \frac{2\lambda}{n^{\alpha/2}}U_t\right)\left[\sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i\right)\eta_i\right]^2 (U_{t+1} - U_t)$$

$$- \left(\sum_{k=1}^{n-1} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k\right)(U_{k+1} - U_k)\right)\left[\sum_{i=1}^{n} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i\right)\eta_i\right]^2$$

$$\equiv R_{1t}^* - 2R_{2t}^*,$$

with

$$R_{1t}^* = \sum_{k=1}^{t} \left[\sum_{i=k+1}^{t} \exp\left(\frac{\phi}{n^\alpha}(k-i) - \frac{\lambda}{n^{\alpha/2}}(U_k - U_i) \right) \eta_i \right]^2 (U_{k+1} - U_k)$$

$$R_{2t}^* = \sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i$$

$$\times \sum_{k=1}^{t} \sum_{i=k+1}^{t} \exp\left(\frac{\phi}{n^\alpha}(2k-i) + \frac{\lambda}{n^{\alpha/2}}(2U_k - U_i) \right) \eta_i (U_{k+1} - U_k)$$

$$= \sum_{i=1}^{t} \exp\left(-\frac{\phi}{n^\alpha}i - \frac{\lambda}{n^{\alpha/2}}U_i \right) \eta_i \times \overline{R}_{2t}^*.$$

The proof follows then the same line as for R_t (in the proof of (B.4)).

Finally, let us look at the summation (B.9).

Notice that

$$\mathsf{E}\left(\int_0^{n^{1-\alpha}} e^{4(\phi r + \lambda W_r)} dr \right) = \frac{e^{4(c+\lambda^2)n^{1-\alpha}}}{4(c+\lambda^2)} - \frac{1}{4(c+\lambda^2)} = \chi_{n^{1-\alpha}}^2 + O(1).$$

Again, we will use a Lindberg Condition, this time regarding

$$\zeta_{k+1} \equiv n^{-\alpha/2} \chi_{n^{1-\alpha}}^{-1} \exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k \right) (U_{k+1} - U_k),$$

which admits conditional variance such that

$$\sum_{k=1}^{n-1} \mathsf{E}_k \left[\zeta_{k+1}^2 \right] = n^{-\alpha} \chi_{n^{1-\alpha}}^{-2} \sum_{k=1}^{n-1} \exp\left(\frac{4\phi}{n^\alpha}k + \frac{4\lambda}{n^{\alpha/2}}U_k \right)$$

$$= \chi_{n^{1-\alpha}}^{-2} \int_0^{n^{1-\alpha}} e^{4(\phi r + \lambda W_r)} dr + o_p(1)$$

$$= O_p(1).$$

It follows that

$$\sum_{k=1}^{n-1} \frac{\exp\left(\frac{2\phi}{n^\alpha}k + \frac{2\lambda}{n^{\alpha/2}}U_k \right)}{\chi_{n^{1-\alpha}}} \frac{U_{k+1} - U_k}{n^{\alpha/2}} = \chi_{n^{1-\alpha}}^{-1} \int_0^{n^{1-\alpha}} e^{2(\phi r + \lambda W_r)} dW_r + o_p(1)$$

so

$$n^{-3\alpha/2} \varphi_{n^{1-\alpha}}^{-2} \chi_{n^{1-\alpha}}^{-1} \sum_{t=1}^{n} y_{t-1}^2 u_t = \left(\varphi_{n^{1-\alpha}}^{-1} \sigma_\eta \int_0^{n^{1-\alpha}} e^{-(\phi s + \lambda W_s)} dB_s \right)^2$$

$$\times \chi_{n^{1-\alpha}}^{-1} \int_0^{n^{1-\alpha}} e^{2(\phi r + \lambda W_r)} dW_r + o_p(1),$$

and

$$\sigma_\eta^{-2} \varphi_{n^{1-\alpha}}^{-2} \chi_{n^{1-\alpha}}^{-1} n^{-3\alpha/2} \sum_{t=1}^{n} y_{t-1}^2 u_t \Rightarrow X^2 V,$$

where V is defined as

$$\chi_{n^{1-\alpha}}^{-1} \int_0^{n^{1-\alpha}} e^{2(\phi r + \lambda W_r)} dW_r \; \Rightarrow \; V \sim \mathsf{N}(0,1).$$

B.3 Conclusion

We can then summarize in the following table the results obtained above, considering the three cases, $c < 0$, $c = 0$ and $c > 0$ respectively, and introducing the notation

$$S_{yy} = \sum_{t=1}^{T} y_t^2, \quad S_{y\eta} = \sum_{t=1}^{T} y_{t-1}\eta_t, \quad \text{and} \quad S_{yyu} = \sum_{t=1}^{T} y_{t-1}^2 u_t.$$

Let the process (y_t) be defined as in (3)-(1) for $t \geq 0$, with $y_0 = 0$.
As $T \to \infty$ and for $x \in \{yy, y\eta, yyu\}$,

$$\sigma_\eta^{-2} \, \mu^x \, \phi_T^x \, S_x \Rightarrow U_x,$$

where (μ^x, ϕ_T^x, U_x) are defined as follows (we assume $(\phi, \lambda) \neq (0,0)$).

	ϕ_T^{yy}	$\phi_T^{y\eta}$	ϕ_T^{yyu}
$c < 0$	$T^{-(1+\alpha)}$	$T^{-\frac{1+\alpha}{2}}$	$T^{-\frac{1+2\alpha}{2}}$
$c = 0$	$T^{-2\alpha}e^{-6\lambda^2 T^{1-\alpha}}$	$T^{-1}e^{-2\lambda^2 T^{1-\alpha}}$	$T^{-3\alpha/2}e^{-6\lambda^2 T^{1-\alpha}}$
$c > 0$			
$\lambda^2 < \phi$	$T^{-2\alpha}e^{-2(c+\lambda^2)T^{1-\alpha}}$	$T^{-\alpha}e^{-2cT^{1-\alpha}}$	$T^{-3\alpha/2}e^{-2(c+\lambda^2)T^{1-\alpha}}$
$\lambda^2 = \phi$	$T^{-(1+\alpha)}e^{-2(c+\lambda^2)T^{1-\alpha}}$	$T^{-\frac{1+\alpha}{2}}e^{-2cT^{1-\alpha}}$	$T^{-(1+\frac{\alpha}{2})}e^{-2(c+\lambda^2)T^{1-\alpha}}$
$\lambda^2 > \phi$	$T^{-2\alpha}e^{-6\lambda T^{1-\alpha}}$	$T^{-\alpha}e^{-(c+2\lambda^2)T^{1-\alpha}}$	$T^{-3\alpha/2}e^{-6\lambda^2 T^{1-\alpha}}$

with

	μ^{yy}	$\mu^{y\eta}$	μ^{yyu}
$c < 0$	$-2c$	$\sqrt{-2c}$	$-2c/\sqrt{3}$
$c = 0$	$8\sqrt{2}\,\lambda^4$	2λ	$8\lambda^3$
$c > 0$			
$\lambda^2 < \phi$	$8(c-2\lambda^2)^2\,\sqrt{(c+2\lambda^2)(c+\lambda^2)}$	$4c\left(c-2\lambda^2\right)$	$8(c-2\lambda^2)^2\,\sqrt{c+\lambda^2}$
$\lambda^2 = \phi$	$2\sqrt{(c+2\lambda^2)(c+\lambda^2)}$	$2c$	$2\sqrt{c+\lambda^2}$
$\lambda^2 > \phi$	$4(2\lambda^2-c)\sqrt{(c+2\lambda^2)(c+\lambda^2)}$	$2c\sqrt{2(2\lambda^2-c)}$	$4(2\lambda^2-c)\sqrt{c+\lambda^2}$

and

	U_{yy}	$U_{y\eta}$	U_{yyu}
$c < 0$	1	$N(0,1)$	$N(0,1)$
$c \geq 0$	$X^2 Z$	XY	$X^2 V$

where $X \sim N(0,1)$, $Y \sim N(0,1)$, $V \sim N(0,1)$, $X \perp V$, and
Z such that $E(Z) = 0$, $V(Z) = 1$ and $Cov(V, Z) = 0$.
Theorem 4 can be directly deduced from the results of this table.

Indeed, in the case $c < 0$, we can write, after noticing that $\sum_{t=1}^{T} y_{t-1}^2 u_t$ is asymptotically uncorrelated

with $\sum_{t=1}^{T} y_{t-1}\eta_t$, that

$$
\begin{aligned}
T^{\frac{1+\alpha}{2}}\left(\widehat{\rho}-E\left(\rho_t\right)\right) &= \lambda \frac{T^{-\frac{1+2\alpha}{2}}\sum_t y_{t-1}^2 u_t}{T^{-1-\alpha}\sum_t y_{t-1}^2} + \frac{T^{-\frac{1+\alpha}{2}}\sum_t y_{t-1}\eta_t}{T^{-1-\alpha}\sum_t y_{t-1}^2} \\
&= \lambda \frac{\mu_T^{yy}}{\mu_T^{yyu}}\frac{\sigma_\eta^{-2}}{\sigma_\eta^{-2}}\frac{\mu_T^{yyu}}{\mu_T^{yy}}\frac{T^{-\frac{1+2\alpha}{2}}\sum_t y_{t-1}^2 u_t}{T^{-1-\alpha}\sum_t y_{t-1}^2} + \frac{\mu_T^{yy}}{\mu_T^{y\eta}}\frac{\sigma_\eta^{-2}}{\sigma_\eta^{-2}}\frac{\mu_T^{y\eta}}{\mu_T^{yy}}\frac{T^{-\frac{1+\alpha}{2}}\sum_t y_{t-1}\eta_t}{T^{-1-\alpha}\sum_t y_{t-1}^2} \\
&\Rightarrow \; \mathsf{N}\left(0, 3\lambda^2 - 2c\right).
\end{aligned}
$$

Assume now that $c \geq 0$. We can write

$$
T^{\alpha/2}\frac{\phi_T^{yyu}}{\phi_T^{yy}}\left(\widehat{\rho}-E\left(\rho_t\right)\right) \;\Rightarrow\; \lambda\frac{\mu^{yy}}{\mu^{yyu}}\frac{U_{yyu}}{U_{yy}},
$$

where the various ratios are calculated using the previous table and provide the same results for all cases when $c \geq 0$, namely

$$
T^{\alpha/2}\;\phi_n^{yyu}/\phi_n^{yy} = T^\alpha \qquad \text{and} \qquad \lambda^{-1}\mu^{yyu}/\mu^{yy} = \frac{1}{\lambda\sqrt{c+2\lambda^2}},
$$

hence the result. $\qquad\qquad\square$

C   Proof of Corollary 5

Recall that under the null, the statistic is defined by

$$
\tau_{0,T} = \begin{cases} T^{\frac{1+\alpha}{2}}\left(\widehat{\rho}-E_{\mathsf{H}_0}\left(\rho_t\right)\right), & \text{if } \phi_0 + \lambda_0^2 < 0; \\ T^\alpha\left(\widehat{\rho}-E_{\mathsf{H}_0}\left(\rho_t\right)\right), & \text{if } \phi_0 + \lambda_0^2 \geq 0. \end{cases}
$$

Let us write

$$
\widehat{\rho}-E_{\mathsf{H}_0}\left[\rho_t\right] = \left(\widehat{\rho}-E_{\mathsf{H}_1}\left[\rho_t\right]\right) + \left(E_{\mathsf{H}_1}\left[\rho_t\right]-E_{\mathsf{H}_0}\left[\rho_t\right]\right).
$$

and consider the two elements of the sum in turn.
The null and alternative hypotheses are local to each other:

$$
E_{\mathsf{H}_1}\left[\rho_t\right]-E_{\mathsf{H}_0}\left[\rho_t\right] = \frac{\phi_1 - \phi_0 + \frac{1}{2}\left(\lambda_1^2 - \lambda_0^2\right)}{T^\alpha} + o\left(T^{-\alpha}\right),
$$

hence $T^{\frac{1+\alpha}{2}}\left(E_{\mathsf{H}_1}\left[\rho_t\right]-E_{\mathsf{H}_0}\left[\rho_t\right]\right)$ diverges but $T^\alpha\left(E_{\mathsf{H}_1}\left[\rho_t\right]-E_{\mathsf{H}_0}\left[\rho_t\right]\right)$ does not.
Also, under the alternative, $T^{\frac{1+\alpha}{2}}\left(\widehat{\rho}-E_{\mathsf{H}_1}\left[\rho_t\right]\right)$ diverges only if $\phi_1 + \lambda_1^2 \geq 0$ but $T^\alpha\left(\widehat{\rho}-E_{\mathsf{H}_1}\left[\rho_t\right]\right)$ does not diverge.
Finally, if both $T^{\frac{1+\alpha}{2}}\left(\widehat{\rho}-E_{\mathsf{H}_1}\left[\rho_t\right]\right)$ and $T^{\frac{1+\alpha}{2}}\left(E_{\mathsf{H}_1}\left[\rho_t\right]-E_{\mathsf{H}_0}\left[\rho_t\right]\right)$ diverge, their sum is $O_p\left(T^{\frac{1-\alpha}{2}}\right)$ so they do not cancel each other.
To conclude, $\tau_{0,T}$ diverges under H_1 only if $\phi_0 + \lambda_0^2 < 0$, irrespective of $\left(\phi_1, \lambda_1\right)$.

D   Proof of proposition 6

Consider the projection

$$y_{t+k} = \exp\left\{\frac{k\phi + \lambda T^{\alpha/2}\sum_{j=1}^{k}u_{t+j}}{T^{\alpha}}\right\}y_t + \sum_{i=1}^{k}\exp\left\{\frac{(k-i)\phi + \lambda T^{\alpha/2}\sum_{j=i+1}^{k}u_{t+j}}{T^{\alpha}}\right\}\eta_{t+i}.$$

Let $(r,s)\in\left(0,T^{1-\alpha}\right)$, with $s>0$, then

$$\frac{y_{[T^{\alpha}(r+s)]}}{y_{[T^{\alpha}r]}} = \exp\left\{\frac{[T^{\alpha}s]\phi + \lambda T^{\alpha/2}\sum_{j=[T^{\alpha}r]+1}^{[T^{\alpha}(r+s)]}u_j}{T^{\alpha}}\right\}$$

$$+\ \frac{1}{y_{[T^{\alpha}r]}}\sum_{i=[T^{\alpha}r]+1}^{[T^{\alpha}(r+s)]}\exp\left\{\frac{([T^{\alpha}(r+s)]-[T^{\alpha}r]-i)\phi + \lambda T^{\alpha/2}\sum_{j=i+1}^{[T^{\alpha}(r+s)]}u_j}{T^{\alpha}}\right\}\eta_{[T^{\alpha}r]+i},$$

where Proposition 1 implies that

$$\exp\left\{\frac{[T^{\alpha}s]\phi + \lambda T^{\alpha/2}\sum_{j=[T^{\alpha}r]+1}^{[T^{\alpha}(r+s)]}u_j}{T^{\alpha}}\right\} \Rightarrow \exp\left\{s\phi + \lambda\left(W_{r+s}-W_r\right)\right\},$$

and

$$T^{-\alpha/2}\sum_{i=[T^{\alpha}r]+1}^{[T^{\alpha}(r+s)]}\exp\left\{\frac{([T^{\alpha}(r+s)]-[T^{\alpha}r]-i)\phi + \lambda T^{\alpha/2}\sum_{j=i+1}^{[T^{\alpha}(r+s)]}u_j}{T^{\alpha}}\right\}\eta_{[T^{\alpha}r]+i}$$

$$\Rightarrow K_{\phi,\lambda}\left(r+s\right) - e^{\phi s + \lambda(W_{r+s}-W_r)}K_{\phi,\lambda}\left(r\right)$$

$$= \int_{r}^{r+s}\exp\left\{\phi(r+s-u) + \lambda\left(W_{r+s}-W_u\right)\right\}dB_s,$$

hence $K_{\phi,\lambda}\left(r+s\right) - e^{\phi s + \lambda(W_{r+s}-W_r)}K_{\phi,\lambda}\left(r\right)$ is independent of $K_{\phi,\lambda}\left(r\right)$ and

$$K_{\phi,\lambda}\left(r+s\right) - e^{\phi s + \lambda(W_{r+s}-W_r)}K_{\phi,\lambda}\left(r\right) \sim \mathsf{N}\left(0,f_{2c}\left(s\right)\right).$$

It follows that we can define a Cauchy variable C such that

$$\frac{y_{[T^{\alpha}(r+s)]}}{y_{[T^{\alpha}r]}} \Rightarrow e^{\phi s + \lambda(W_{r+s}-W_r)} + \sqrt{\frac{f_{2c}\left(s\right)}{f_{2c}\left(r\right)}}\,C, \tag{D.10}$$

which constitutes the first half of the proposition.

Now let us turn to the proof of the second part of the proposition. If $c\geq 0$ then $s/r\to 0$ implies $\frac{f_{2c}(s)}{f_{2c}(r)}\to 0$. Hence the impact of the Cauchy variable in expression (D.10) vanishes. This is not the case of $e^{\phi s + \lambda(W_{r+s}-W_r)}$ unless s itself tends to zero, which we do not consider.

Let $t = \lfloor T^\alpha r \rfloor$, $k = \lfloor T^\alpha s \rfloor$ then, as $T \to \infty$,

$$\mathsf{P}\left(\frac{y_{t+k}}{y_t} \geq \gamma\right) \to \mathsf{P}\left(e^{\phi s + \lambda(W_{r+s} - W_r)} \geq e^{\log \gamma}\right)$$

$$= \mathsf{P}\left(\frac{W_{r+s} - W_r}{\sqrt{s}} \geq \frac{\log \gamma - \phi s}{\lambda \sqrt{s}}\right),$$

where $\frac{W_{r+s} - W_r}{\sqrt{s}} \sim N(0, 1)$ so

$$\mathsf{P}\left(\frac{y_{t+k}}{y_t} \geq \gamma\right) \to 1 - \Phi\left(\frac{\log \gamma - \phi s}{\lambda \sqrt{s}}\right) = \Phi\left(\frac{\phi s - \log \gamma}{\lambda \sqrt{s}}\right),$$

i.e.

$$\mathsf{P}\left(\frac{y_{t+k}}{y_t} \geq \gamma\right) - \Phi\left(\frac{\phi k T^{-\alpha} - \log \gamma}{\lambda \sqrt{k T^{-\alpha}}}\right) \xrightarrow[T \to \infty]{} 0.$$

E   Present Value Model

Consider the standard definition of an ex-post asset return

$$r_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} - 1,$$

(see e.g. Campbell, Lo and McKinlay, 1996, expression (7.1.1)) and assume r_{t+1} constant and equal R. Then

$$P_t = \frac{P_{t+1} + D_{t+1}}{1 + R},$$

which is compatible with

$$\Delta P_t = (1 + (1 - \delta) R + \delta R_t) \Delta P_{t-1} - \zeta_t,$$

where R_t is iid and $\mathsf{E}\left[(1 + R_t)^{-1}\right] = (1 + R)^{-1}$.

Indeed, the expression

$$\Delta P_t = (1 + (1 - \delta) R + \delta R_t) \Delta P_{t-1} - \zeta_t \tag{E.11}$$

implies that

$$P_{t+1} + D_{t+1} = P_t + (1 + (1 - \delta) R + \delta R_{t+1}) \Delta P_t - \zeta_{t+1} + D_t + \zeta_{t+1}$$

$$\frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}} = \frac{P_t + (1 + (1 - \delta) R + \delta R_{t+1}) \Delta P_t}{1 + R_{t+1}} + \frac{D_t}{1 + R_{t+1}}$$

$$= \frac{P_t + (1 + (1 - \delta) R) \Delta P_t}{1 + R_{t+1}} + \delta \frac{R_{t+1}}{1 + R_{t+1}} \Delta P_t + \frac{D_t}{1 + R_{t+1}}$$

Now, if

$$P_t = \mathsf{E}_t\left[\frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}}\right],$$

44

then taking conditional expectations on either side gives

$$\mathsf{E}_t\left[\frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}}\right] = \frac{P_t + (1 + (1 - \delta)\,R)\,\Delta P_t}{1 + R}$$

$$+ \frac{D_t}{1 + R} + \delta\mathsf{E}_t\left[\frac{1 + R_{t+1}}{1 + R_{t+1}} - \frac{1}{1 + R_{t+1}}\right]\Delta P_t$$

$$= \frac{P_t + (1 + (1 - \delta)\,R)\,\Delta P_t}{1 + R} + \frac{D_t}{1 + R} + \delta\left(1 - \frac{1}{1 + R}\right)\Delta P_t$$

$$= \frac{P_t + (1 + R)\,\Delta P_t}{1 + R} + \frac{D_t}{1 + R}$$

$$= \frac{P_t + D_t}{1 + R} + P_t - P_{t-1}.$$

Now let r_t such that $P_{t-1} = \dfrac{P_t + D_t}{1 + r_t}$ then

$$\mathsf{E}_t\left[\frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}}\right] = \frac{P_t + D_t}{1 + R} + P_t - \frac{P_t + D_t}{1 + r_t},$$

which shows that if r_t is constant and equal to R, (E.11) implies that the present value model holds:

$$P_t = \mathsf{E}_t\left[\frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}}\right].$$

F Simulated NERC paths

In order to show the sort of dynamics the model generates, Figure 12 records simulations of the process over samples of $T = 1000$ observations using two sets of draws of (u_t, η_t). Exuberant periods become clearly more pronounced and explosive as ϕ increases or α decreases. For $\alpha = 1$, the processes exhibit near-unit roots as in Phillips (1987) and no type of what could be called a "bubble" seems to appear visually; we disregard this situation in the paper. As α decreases, some bubbles appear. Some local explosive pattern appears and disappears alternatively. Although, by visual inspection, some draws seem to exhibit volatility clustering (random draw 1, left column), this is generically not an observed pattern (see random draw 2).
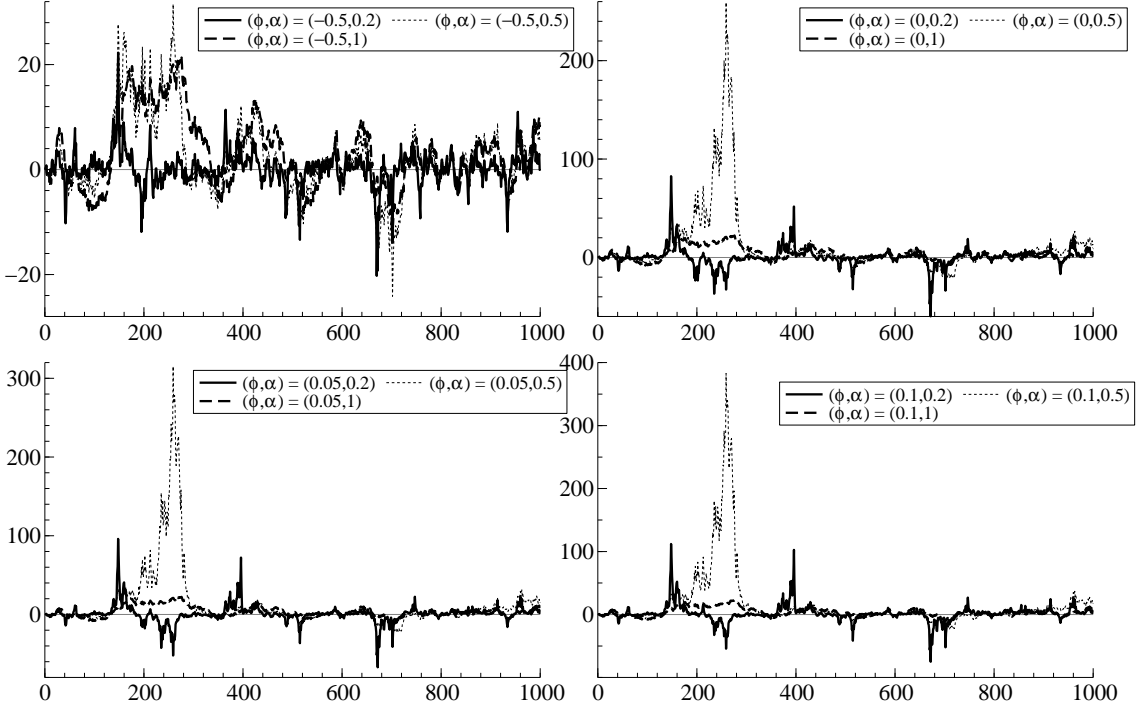
RANDOM DRAW 1
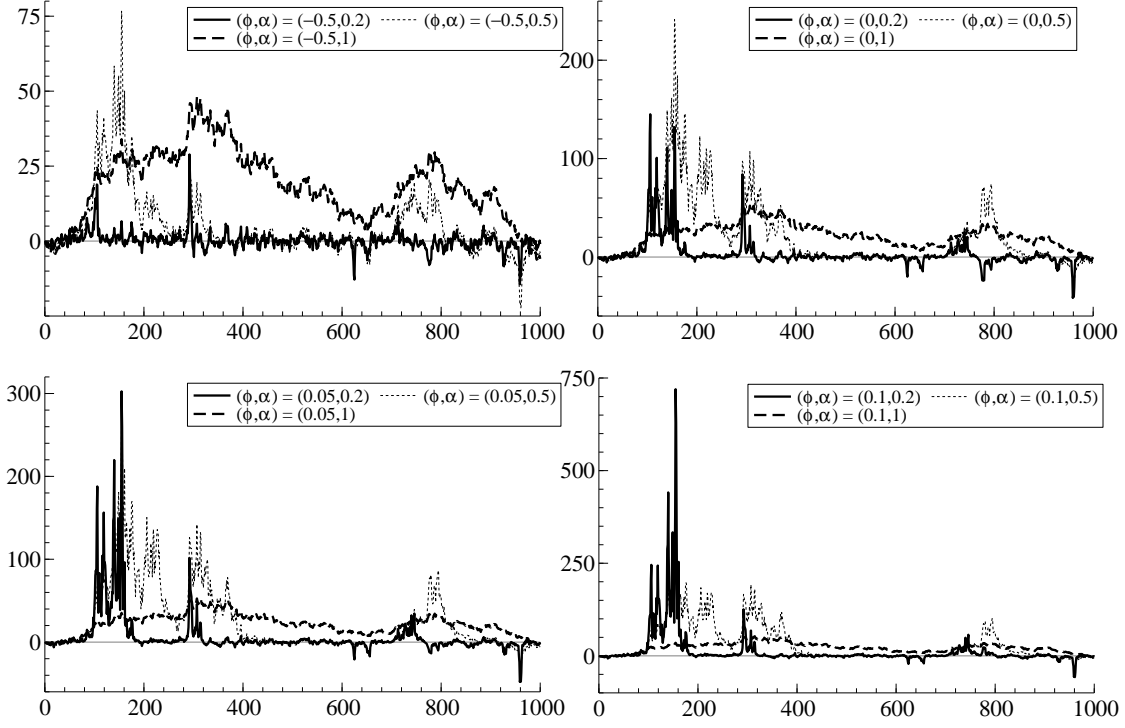
RANDOM DRAW 2

Figure 12: Simulated realizations from the model of autoregressive conditional exuberance for different parameter values.

Informations

Alison Bougi
+33 (0)1 34 43 33 58
bougi@essec.fr
www.essec.fr
research.center@essec.fr

ESSEC
BUSINESS SCHOOL