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

Anurag Banerjee
Durham Business School

Guillaume Chevillon*
ESSEC Business School

Marie Kratz
ESSEC Business School

April 25, 2013

Work in progress.

Abstract
This paper proposes a Near Explosive Random-Coefficient Autoregressive model that can accommodate the pricing of assets under standard present-value relations, both according to fundamentals and in the presence of bubbles. The distribution of the random coefficient is parameterized in a local-asymptotics framework as a moderate deviation from a stochastic unit root. An application to inference regarding the dynamics of U.S. house prices shows the pertinence of the model.

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

*Corresponding author: chevillon@essec.edu. The authors would like to Kajal Lahiri, Kevin Lansing, Huamin Peng, Barbara Rossi and Shu-Ping Shi as well as the participants in the 2011 SNDE Symposium, the 2012 SMU-ESSC Symposium on Financial Econometrics, and seminar participants at SUNY, Albany.
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) for 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 $P_t = E_t (M_{t+1} (P_{t+1} + D_{t+1}))$. 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 $B_t = E_t (M_{t+1}B_{t+1})$. There exist solutions to this condition 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 or 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 $y_t$ (a function 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.\footnote{The literature has also provided several other techniques to test for the presence of a bubble, see G"urkaynak (2008) for an overview.}

To increase power, these authors adapt the critical values to the sample size, with the help the distributions derived by Phillips and Magdalinos (2007, PM henceforth) under the alternative of a locally explosive root

\begin{equation}
\begin{align*}
y_t &= \rho_0 y_{t-1} + \eta_t, \quad t = 1, ..., T; \\
\rho_0 &= \exp \left\{ \frac{\phi_0}{T^\alpha} \right\}, \text{ with } \phi_0 > 0 \text{ and } \alpha \in (0, 1),
\end{align*}
\end{equation}

where $\eta_t$ is weakly dependent with mean zero. When $\eta_t$ 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, a NEAR(1), since $\rho_0 > 1$ but $\rho_0 \to 1$ as the sample size $T$ increases.

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 introduced 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. To avoid predetermining the appearance or disappearance of bubbles or their frequency, we generalize their approach to a Near Explosive Random Coefficient autoregressive process, a NERC(1) defined as:

\begin{equation}
\begin{align*}
y_t &= \rho_t y_{t-1} + \eta_t, \quad t = 1, \cdots, T;
\end{align*}
\end{equation}

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

\[
\rho_t = \exp \left\{ \phi + \lambda T^{\alpha/2} u_t \right\}
\]

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; also the absence of deterministic breaks avoids 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 $\mathbb{E}(\rho_t) = \rho_0$ and common $\eta_t$. 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 the appendix and 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 $y_t$, we avoid the issue of negative bubbles pointed by Diba and Grossman (1998).

Although the NERC(1) model parametrically nests the NEAR(1), its properties differ when $V(\rho_t) \neq 0$. In particular, its asymptotics depends on the value of $c = \phi + \lambda^2$. When the NERC(1) process is weakly stationary ($c < 0$) the OLS estimate of $\rho_t$ converges to Normal distribution, as under the NEAR(1) 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 non weakly stationary ($c \geq 0$) the asymptotic is qualitatively different from the NEAR(1) in the sense that when $\lambda \to 0$, the asymptotic distribution is not close to the distribution corresponding to NEAR(1) model as 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 multiple 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.

Based on the asymptotic theory developed in this paper, we provide an inferential approach for the NERC(1) by inverting the asymptotic distribution the least-squares estimator $\hat{\rho}$. The method was popularized by Stock (1991) and Andrews (1993) and has been used widely in the near unit root and weak instrument literatures. This method of estimation can be performed in real-time since we need not resort to deterministic breaks. The distinctive asymptotic theory of NERC(1) is useful we allows to forecast the evolution (boom) and devolution (bust) of the bubble generation process. Thus given a history of observations, we can forecast in real time the probability of a bust or a boom. We evaluate our methodology empirically using the Case-Shiller index of US house prices.

The structure of the paper is as follows. In section 2, we define the random-coefficient au-
toregressive 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 model properties and inferential methods is shown in section 5. We apply our methodology in section 6 to the inference on the dynamic properties of U.S. house prices.

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):

\[ y_t = \rho_t y_{t-1} + \eta_t, \quad t = 1, \cdots, T; \]

where \( \eta_t \) is assumed to be identically and independently distributed with zero expectation, variance \( \sigma_\eta^2 \) and moment conditions that we specify below; \( \rho_t \) is a nonnegative covariance stationary process that is independent of \( \eta_t \). The RCA model (3) with

\[ \mathbb{E} [\max \{ \log |\eta_t|, 0\}] < \infty \quad \text{and} \quad \mathbb{E} [\max \{ \log |\rho_t|, 0\}] < \infty \]
is known (see Aue, Horváth and Steinebach, 2006) to admit a strictly non-anticipatory stationary solution if and only if
\[ E[\log|\rho_t|] < 0 \] (5)
and a covariance stationary solution if
\[ E[\rho_t^2] < 1. \] (6)

Hence, the unit root hypothesis can take several forms: \( E[\rho_t] = 1 \), or \( E[\rho_t^2] = 1 \), see Granger and Swanson (1997) for a discussion.\(^2\) When \( E[\rho_t^2] > 1 \), Hwang and Basawa (2005) denote this model an Explosive Random Coefficient Autoregressive model (ERCA) and study processes such that \( E[\rho_t^2] \geq 1 \) and \( E[\log|\rho_t|] < 0 \) (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 Model à la Granger-Swanson in the sense that we assume that both the expectation and variance of \( \rho_t \) are very close to unity: we model the moments using extensions to standard local-asymptotic frameworks so that as \( T \to \infty \) (\( E[\rho_T], V[\rho_T] \to 1,0 \)). This framework builds on Bobkoski (1983), Chan and Wei (1987), Phillips (1987) and the more recent work of Giraitis and Phillips (2006) and Phillips and Magdalinos. It constitutes an extension to PWY, and PY). The process we consider are formally defined as triangular arrays as the distribution of \( y_t \), for \( t \leq T \), is allowed to depend on the actual sample size \( T \) : we parameterize the distribution of \( \rho_t \) 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:

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

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

**Assumption 2** \( y_0 = o_p(T^{\alpha/2}) \) and
\[ E|\eta_t|^\nu < \infty \quad \text{for} \quad \nu \geq \frac{2}{\alpha} \]
\[ E|u_t|^\omega < \infty \quad \text{for} \quad \omega \geq \frac{2}{\alpha} \]

\(^2\)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), Distasio (2008) and Aue and Horváth (2011).\(^3\) Also, expression (3) implies that \( y_t \) exhibits conditional heteroskedasticity: assume \( \rho_t \sim \text{iid} (\rho, \sigma_\rho^2) \) then
\[ E[y_t|y_{t-1}] = \rho y_{t-1}, \quad \text{Var}[y_t|y_{t-1}] = \sigma_\rho^2 y_{t-1}^2 + \sigma_\eta^2 \]
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.
Assumption 2 ensures that the assumption (4) in Aue et al. (2006) is satisfied. The assumption also implies that a strong approximation is possible, see Cs"{o}rg"{o} and Horv"{a}th (1993) and PM according to which we can construct an expanded probability space with standard Brownian motions $W$, $B$ such that as $T \to \infty$,

$$
\begin{align*}
\sup_{t \in [0,T_1 - \alpha]} & T^{-\alpha/2} \sum_{i=1}^{\lfloor t^\alpha \rfloor} y_i - W_t = o_{a.s.} \tag{1}, \\
\sup_{t \in [0,T_1 - \alpha]} & T^{-\alpha/2} \sigma^{-1} \sum_{i=1}^{\lfloor t^\alpha \rfloor} \eta_i - B_t = o_{a.s.} \tag{1}
\end{align*}
$$

(7)

where $[\cdot]$ denotes the integer part.

Under assumptions 1 and 2, the expectation of $\rho_t$ is $E[\rho_t] = \exp \left\{ \left( \phi + \frac{1}{2} \lambda^2 \right) / T^\alpha \right\}$ and its variance satisfies $V[\rho_t] = \exp \left\{ \frac{\lambda^2}{T^\alpha} \right\} \left( \exp \left\{ \frac{\lambda^2}{T^\alpha} \right\} - 1 \right) = \frac{\lambda^2}{T^\alpha} + O \left( T^{-2\alpha} \right)$. So $\rho_t$ admits the following stochastic expansion:

$$
\rho_t = E[\rho_t] + \frac{\lambda}{T^{\alpha/2}} y_t + \frac{\lambda^2}{2T^\alpha} (y_t^2 - 1) + O_p \left( T^{-2\alpha} \right) \tag{8}
$$

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

$$
\begin{align*}
S_w &= \{(\phi, \lambda) \in \mathbb{R} \times \mathbb{R}_+, \phi + \lambda^2 < 0 \} \\
S_s &= \{(\phi, \lambda) \in \mathbb{R} \times \mathbb{R}_+, \phi < 0 \}
\end{align*}
$$

The conditions (5) and (6) for strict and covariance stationarity correspond respectively to $(\phi, \lambda) \in S_s$ and $(\phi, \lambda) \in S_w$. We also define the subset $S_s \setminus w = S_s \setminus S_w$ 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 $(\phi, \lambda) \in S_s \setminus w$ with $\lambda > 0$ can be shown to be characterized by a power law, in the sense that there exist $\tau > 0$ such that $Pr \left( |y_t| > a \right) \sim \tau a^{-\sqrt{-2\phi}/\lambda}$ as $a \to \infty$. Hence moments of $y_t$ exist up to the order $\sqrt{-2\phi}/\lambda - 1 \leq \sqrt{2} - 1$. Hence $(\phi, \lambda) \in S_s \setminus w$ implies that the process is not characterized by temporary explosive behavior (as when $E[\rho_t^2] > 1$) 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 8.1.).

Notice the condition $E[\rho_t] < 1 \Leftrightarrow \phi + \frac{1}{2} \lambda^2 < 0$ differs from those defining $S_w$ and $S_s$ as

$$
E[\log |\rho_t|] \leq E[\rho_t] \leq E[\rho_t^2]
$$

where the 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 present in figure 2 which values of $(\phi, \lambda)$ 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 $\rho_t$. In his setting $E[\rho_t] - 1 = \ldots$
Figure 2: Values of $(\phi, \lambda) \in \mathbb{R} \times \mathbb{R}_+$ belonging to the subsets $S_w$ and $S_{s\setminus w}$ which correspond respectively to $y_t$ being weakly stationary and strictly yet non weakly stationary. The figure also reports whether $E[\rho_t] > 1$ and $E[\rho_t^2] > 1$.

$O(T^{-\alpha})$ with $\alpha \in (1/2, 1)$, and $V[\rho_t] = o(T^{-1})$ which implies that $\rho_t - E[\rho_t]$ lies in a tighter neighborhood of unity and so does not impact the 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 $\rho_t$ lies further away from unity and we show that this affects significantly the asymptotic distributions. Accordingly with PY, we restrict $\alpha < 1$ to ensure that $\rho_t$ 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 NERC with $E[\rho_t] > 1$ and non-local parameters 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 in the appendix simulated paths of the NERC process.

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 diffusion:

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

The FCLT hence follows.

---

5For him conditions $E[\rho_t^2] < 1$ and $E[\rho_t] < 1$ are asymptotically equivalent.

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

7
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 $T \to \infty$,

$$T^{-\alpha/2} y_{[r_T^\alpha]} \Rightarrow \sigma_n K_{\phi, \lambda}(r).$$

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

$$c = \phi + \lambda^2$$

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

$$y_T = \begin{cases} 
O_p \left( T^{\frac{\alpha}{2}} \right), & \text{if } c < 0; \\
O_p \left( T^{\frac{1}{2}} \right), & \text{if } c = 0;
\end{cases}$$

(11)

where $c$ differs from $\phi$ when $\lambda > 0$. When $c > 0$, the process satisfies

$$y_T = \begin{cases} 
O_p \left( T^{\alpha/2} e^{\lambda^2 T^{1-\alpha}} \right), & \text{if } \lambda^2 < \phi; \\
O_p \left( T^{\frac{1}{2}} e^{2\lambda T^{1-\alpha}} \right), & \text{if } \lambda^2 = \phi; \\
O_p \left( T^{\alpha/2} e^{2\lambda^2 T^{1-\alpha}} \right), & \text{if } \lambda^2 > \phi.
\end{cases}$$

(12)

The latter expression shows that when $c > 0$, the process exhibits explosiveness

### 3 Inference

In this section, we show how theorem 3 can be used to conduct inference on model parameters.

#### 3.1 Least Squares Estimator

We now consider the distribution of the ordinary least-squares (OLS) estimator $\hat{\rho}$ in the regression of $y_t$ on $y_{t-1}$. The expansion (8) implies that, as $T \to \infty$,

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

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

$$\hat{\rho} - E \left( \rho_t \right) = \lambda T^{-\alpha/2} S_{yyu} S_{yy} + O_p \left( T^{-3\alpha/2} \right),$$

(13)
and its asymptotic distribution is driven by the the sum with higher magnitude between $T^{-\alpha/2}S_{yuu}$ and $S_{y\eta}$. For this analysis, we introduce the following random variables:

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

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

Letting a tilde denote the random variable scaled by its standard deviation (e.g. $\tilde{V}_{T^{1-\alpha}} = \sqrt{\text{Var}[V_{T^{1-\alpha}}]}^{-1/2} V_{T^{1-\alpha}}$) we show in the appendix that they jointly tend as $T \to \infty$ to

$$\left(\tilde{V}_{T^{1-\alpha}}, \tilde{X}_{T^{1-\alpha}}^{(-)}, \tilde{X}_{T^{1-\alpha}}^{(+)}, \tilde{Z}_{T^{1-\alpha}}\right) \Rightarrow \left(V, X^{(-)}, X^{(+)}, Z\right),$$

where $X^{(-)}$, $X^{(+)}$ and $V$ are standard normal variables and $Z$ has zero expectation and unit variance.\footnote{Matsumoto and Yor (2005) 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.} In addition $Z$ does not correlate with $V$. We can now provide our result on the distribution of the sample moments:

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

if $c < 0$,

$$T^{-1}(1+\alpha)S_{y\eta} \overset{p}{\to} \frac{\sigma^2}{\lambda c},$$

$$T^{-1}S_{y\eta} \overset{d}{\to} \mathcal{N} \left(0, \frac{\sigma^2}{\lambda c^2}\right),$$

$$T^{-1+2\alpha} S_{yuu} \overset{d}{\to} \mathcal{N} \left(0, \frac{12\sigma^2}{\lambda c^2}\right);$$

if $c \geq 0$ and for $x \in \{u, \eta\}$, there exist $(\mu^x, \phi^x_T)$ functions of $(\phi, \lambda)$ such that

$$T^{-\alpha/2} \phi^y_T S_{y\eta} \Rightarrow \frac{\sigma^2}{\mu^y \sqrt{\lambda + 2\lambda}} X^2 Z,$$

$$\phi^u_T S_{y\eta} \Rightarrow \frac{\sigma^2}{\mu^u} X^2 Y,$$

$$\phi^u_T S_{yuu} \Rightarrow \frac{\sigma^2}{\mu^u} X^2 V,$$

with $\phi^y_T / \phi^u_T = o \left(e^{-2\lambda^2 T^{1\alpha}}\right)$: joint convergence of the three moments also holds.

The lemma shows that when $c < 0$, so the process is weakly stationary, then both $S_{yuu}$ and $S_{y\eta}$ impact the asymptotic distributions, but that when $\lambda \neq 0$ and $c \geq 0$ $S_{yuu}$ dominates. This setting differs markedly from that of Aue (2008) where the variance of $\rho_t$ is of lower magnitude so $S_{y\eta}$ is the dominant term in the expansion (13). By contrast, in the fixed-asymptotics framework of Hwang and Basawa (2005), the ratio $S_{yuu}/S_{yy}$ is not premultiplied by $T^{-\alpha/2}$ and diverges, yielding an inconsistent OLS estimator. This is not the case here as the following theorem shows.

**Theorem 3** Let the process $y_t$ be defined as in (3) under assumptions 1 and 2 for $t \geq 0$. When $\lambda > 0$, the OLS estimator $\hat{\rho}$ in the regression of $y_t$ on $y_{t-1}$ satisfies the following properties as $T \to \infty$:
if $c < 0$, $T^{\frac{1+c}{c}} (\hat{\rho} - \mathbb{E} [\rho_t]) \Rightarrow \mathcal{N} (0, 3\lambda^2 - 2c)$
if $c \geq 0$, $T^\alpha (\hat{\rho} - \mathbb{E} [\rho_t]) \Rightarrow \lambda \sqrt{c + 2\lambda^2 V}$

where $c = \phi + \lambda^2$.

The theorem shows several key differences from the existing literatures on near unit roots and random coefficients when $c \geq 0$. Indeed, when the process is weakly stationary, $c < 0$, the asymptotic distribution of the OLS estimator $\hat{\rho} - \mathbb{E} [\rho_t]$ is comparable to the results of PM and Aue that $T^{\frac{1+c}{c}} (\hat{\rho} - \rho) \Rightarrow \mathcal{N} (0, -2\phi)$: the presence of the stochastic root does not affect the asymptotic normality of $\hat{\rho}$ or the rate of convergence; the only difference is that the asymptotic variance is larger.

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 $O_p (T^{-1})$ of unit root processes or the exponential rate of PM where $(2\phi)^{-1} T^\alpha e^{\phi T^{-1-\alpha}} (\hat{\rho} - \rho)$ tends to a standard Cauchy variable. Convergence can be arbitrarily slow here if $\alpha$ 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.

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

**Corollary 4** Under the assumptions and conditions of theorem 3, define the test statistic $\tau_{0,T}$ for the null $H_0 : (\phi, \lambda) = (\phi_0, \lambda_0)$ as

$$\tau_{0,T} = \begin{cases} 
T^{\frac{1+c}{c}} (\hat{\rho} - \mathbb{E}_{H_0} (\rho_t)) , & \text{if } \phi_0 + \lambda^2_0 < 0; \\
T^\alpha (\hat{\rho} - \mathbb{E}_{H_0} (\rho_t)) , & \text{if } \phi_0 + \lambda^2_0 \geq 0.
\end{cases}$$

Then under $H_1 : (\phi, \lambda) = (\phi_1, \lambda_1) \neq (\phi_0, \lambda_0)$ and as $T \to \infty$

$$\tau_{0,T} = \begin{cases} 
O_p \left( T^{\frac{1+c}{c}} \right) , & \text{if } \phi_0 + \lambda^2_0 < 0; \\
O_p (1) , & \text{if } \phi_0 + \lambda^2_0 \geq 0.
\end{cases}$$

The corollary shows that the test based on the OLS estimator is asymptotically powerful for weakly stationary nulls, yet it has low power not for a null of non weak stationarity even when the process $y_t$ exhibits explosiveness (if $\phi_1 > 0$ as expression (11) shows). This sheds light on why the simulations of Evans (1991) and Charemza and Deadman (1995) found that the Dickey-Fuller test has low power in the presence of periodically collapsing bubbles. These authors used

---

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)
various test statistics but we notice that $\tau_{0,T}$, as defined in the corollary, is scale invariant, whereas those proposed in the literature (e.g. Aue and Horvath (2012) or Hwang and Basawara, 2005) require estimating $\sigma^2$ in a first step; estimators thereof where suggested by Schick (1996) under the assumption of covariance stationarity but the properties of these when $y_t$ is explosive are not established. For this reason, we do not consider these estimators in the following.

### 3.2 Inference using Grid Testing

The DGP we consider uses a local-asymptotic parameterization and it is well known that localizing parameters may not be consistently estimable using standard techniques (see Phillips, 1987). Yet, these are the parameters of interest, and the key assumption we want to test is whether $(\phi, \lambda) \neq (0,0)$, i.e. whether the market price differs from its fundamental.

To conduct inference, we resort to the standard technique that consists in inverting a test statistic. There exists now a significant literature where such an approach is used for inference in the near-unit root framework (originated from Stock, 1991). The technique relies on introducing a scalar function $\tau_{\theta,T}$ (a test statistic) of $Y_T = (y_1, ..., y_T)'$ that satisfies

$$\tau_{\theta,T}(Y_T) \Rightarrow \tau_{\theta}(Y)$$

where $\theta = (\phi, \lambda)' \in \Theta$ denote the parameters of interest, here $\theta = (\phi, \lambda)$. Under the null $H_0 : \theta = \theta_0$, Stock (1991) constructs asymptotic $(1-\omega)$% confidence sets as $\Theta^* \subset \Theta$ consisting of the values $\theta^*$ which are not rejected by $\tau_{\theta^*}(Y)$ at size $\omega$. 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 $\theta^*$ may constitute a biased estimator of $\theta$ but median-unbiased estimation is feasible under the weak convergence assumption, provided that the quantile function is monotonic (Stock, 1991, Andrews, 1993). When $\tau$ is a GMM statistic, $\theta^*$ 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

$$H_0 : (\phi, \lambda) = (\phi_0, \lambda_0).$$

Since $y_t - E_{H_0}[\rho_T] y_{t-1} = (\rho_t - E_{H_0}[\rho_T]) y_{t-1} + \eta_t$, we use the moment condition:

$$\text{Cov} \left( y_t - E_{H_0}[\rho_T] y_{t-1}, y_{t-1} \right) = 0_{H_0}$$

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

$$y_t - E_{H_0}[\rho_T] y_{t-1} = \beta y_{t-1} + \eta_t$$

9We hence rule out nonlinear extensions to the Kalman filter and the particle filter.

10This 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.
and we set $\tau_{\theta, T}$ to be the OLS estimator $\hat{\beta} = \hat{\rho} - \mathbb{E}_{H_0}[\rho_T]$ scaled by the asymptotic rate given in theorem 3. Confidence sets are obtained by grid search over all possible values of $(\phi, \lambda)$ and critical values are obtained by simulation. The variance $\sigma^2_\eta$ constitutes a scaling parameter that does not affect the asymptotic distribution of $\hat{\beta}$ so we may fix it to unity. Also, $\alpha$ is not identified using the method: it constitutes only a scaling parameter since it does not enter the asymptotic distributions in theorem 3\textsuperscript{11}. In the following, we consider testing against either a one-sided alternative (rejection in the right 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.

Following Phillips (2012) we recognize that as $|\phi| \rightarrow \infty$ or $\lambda \rightarrow \infty$, 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 4 shows that although we obtain valid asymptotic confidence sets under the null, the power is low and the proposed confidence sets may be too wide. For this reason, we assess their coverage probabilities by simulation in the next subsection.

### 4 Forecasting

An attractive feature of the model we propose, is that it provides a distributional assumption about $\rho_t$ contrary to models where $\rho_t$ 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 what a bubble and to characterize its timing and magnitudes, see e.g. White and Granger (2011) but our purpose here is not to provide an extensive characterization. Many definitions of a bubble imply some positive growth rate that is sustained over a finite horizon, say $k > 0$ periods. This implies that we are concerned with events such as \{\(y_{t+k}/y_t \geq \gamma\)\} for some $\gamma > 0$. We define the probability of this event as

\[
\mathbb{P}^\gamma_{t,k} \equiv \mathbb{P}\left(\frac{y_{t+k}}{y_t} \geq \gamma\right).
\]

An example of a question of interest may for instance concern the probability that over the horizon $k > 0$, the process grows at least as fast as has been observed over the last $k$ periods. We may express this probability as $\mathbb{P}^{y_{t+k}/y_t}_{t,k}$ (where we implicit assume to condition on $y_t/y_{t-k}$). It is also straightforward to extend (16) to the probability of joint events (such as growth $y_{t+k}/y_t \geq \gamma$ followed by decay $y_{t+k+1}/y_{t+k} < 1$).

The following proposition shows how the questions above frame into a simple analytic expression using our model.

**Proposition 5** Under the assumptions and conditions of theorem 3, then for $(r, s) \in [0, T^{1-\alpha}]^2$ and $c \geq 0$, as $T \rightarrow \infty$

\textsuperscript{11}Since $\alpha$ is a scaling parameter it will be fixed as $\alpha = 1/2$ in the empirical applications.
\[
\frac{y_{t}|T_{n}^{\alpha}(r+s)}{y_{t}|T_{n}^{r}} \Rightarrow \exp \left\{ \phi s + \lambda (W_{r+s} - W_{r}) \right\}, \text{ and}
\]
\[
P_{\gamma}^{T_{n}^{r}|T_{n}^{s}} \to \Phi \left( \frac{\phi s - \log \gamma}{\lambda \sqrt{s}} \right),
\]
where \( \Phi \) denotes the standard normal cumulative distribution function.

The proposition shows that probabilities such as \( P_{\gamma} t,k \) asymptotically tend to very simple expressions, where the functional central limit theorem yields normality. It follows that an approximation based on the asymptotic distributions is obtained as
\[
P_{\gamma} t,k \approx \Phi \left( \frac{\phi k T^{-\alpha} - \log \gamma}{\lambda k^{1/2}} \right).
\]
The expressions above rely crucially on \( \lambda \neq 0 \); when this is not the case and under the local-asymptotic approximation, \( P_{\gamma} T_{n}^{r}|T_{n}^{s} \to 1 \{ \phi s - \log \gamma \geq 0 \} \), with \( 1 \{ \cdot \} \) the indicator function.

Combining proposition 5 with the method for inference that we proposed previously, we can compute confidence intervals for \( P_{\gamma} t,k \) relying on \((\phi,\lambda) \in \Theta^{*}\), where \( \Theta^{*} \) is the set comprising the non-rejected values at specified significance level (provided \( \Theta^{*} \cap \{ (\phi,\lambda), \phi + \lambda^{2} < 0 \} \) is empty).

In particular, if we let \( \gamma \) vary with the sample size, and define \( \kappa \) such that
\[
k^{-1} T^{\alpha} \log \gamma \to \kappa,
\]
i.e. \( \gamma = \left( e^{\kappa T^{-\alpha}} \right)^{k} \), then \( P_{\gamma}^{T_{n}^{r}|T_{n}^{s}} \to \Phi \left( \frac{\phi \kappa - \kappa \lambda}{\sqrt{s}} \right) \). Hence, letting \( (\tilde{\phi}_{\kappa}, \tilde{\lambda}_{\kappa}) = \arg \min_{(\phi,\lambda) \in \Theta^{*}} \frac{\phi \kappa - \kappa \lambda}{\lambda} \), then
\[
\lim_{T \to \infty} P_{\gamma}^{T_{n}^{r}|T_{n}^{s}} \geq \Phi \left( \frac{\tilde{\phi}_{\kappa} - \kappa \lambda}{\tilde{\lambda}_{\kappa} \sqrt{s}} \right). \]
The parameter \( (\tilde{\phi}_{\kappa} - \kappa) / \tilde{\lambda}_{\kappa} \) yields a bound that is independent of the horizon.

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 intervals. 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.

The method of asymptotic inference introduced by Stock (1991) is modified in Hansen (1999) who recommends the use of a so called grid bootstrap. Such bootstrap aims at replacing the use of the asymptotic distribution (14) by the finite-sample bootstrap distribution whose critical values can be obtained by repetitive sampling from the empirical distribution of the errors \( v_{t} = y_{t} - E_{H_{0}} [ \rho T ] y_{t-1} \) (which are observed under \( H_{0} \)). Noticing that
\[
v_{t} = \left[ \frac{\lambda}{T^{\alpha/2}} u_{t} + \frac{\lambda^{2}}{2T^{\alpha}} (u_{t}^{2} - 1) + O_{p} (T^{-2\alpha}) \right] y_{t-1} + \eta_{t},
\]
is heteroskedastic but asymptotically serially uncorrelated (as \( T \to \infty \)) it appears possibly sufficient to use a bootstrapping technique that is immune to heteroskedasticity, such as the wild bootstrap as we consider below.\footnote{Following Davidson and Flachaire (2008), we used the wild bootstrap with standard normal or Bernoulli dis-}
We first consider one-sided (upper tailed) tests. Tables 1 and 2 report the simulated finite sample (respectively $T = 3,000$ and $T = 300$ observations) coverage probability of 95% confidence intervals constructed using the asymptotic and bootstrap distributions. The tables show that the coverage is reasonable under the asymptotic distribution, yet narrow – corresponding to a conservative test – when $\phi > 0$ and $\lambda > 0$. For $\lambda > 0, \phi < 0$, the test is even more conservative and coverage is 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 $\lambda \leq 0$, the test has low power and the coverage rate is inappropriately large, both in small and medium-sized samples. The lower panels in the table report the coverage probabilities using the bootstrapped distributions. These lead to wider coverage and low discriminatory power. Notice the exception of the case $\lambda^2 \approx \phi$ (here for $\lambda = .5$ and $\phi$ between .2 and .3) in table 1; this correspond to a discontinuity of the asymptotic distributions as expression (12) shows.

Table 3 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 $\rho_t$ under the alternative that it is random with same expectation. Figures 5, 4 and 3 report, for a given value of $E[\rho_t] = \rho$, the rejection probabilities of the null $H_0 : (\phi_0, \lambda_0) = (T^a \log \rho, 0)$ at the asymptotic nominal size of 10% under the alternative $(\phi, \lambda)$ which preserves $E[\rho_t] = \rho$ (so $\phi + \lambda^2/2 = \phi_0$). Figures 5 and 3 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 report the same rejection probabilities but where we parameterize the parameter space as $(\phi, \lambda)$ (left) or $(\phi + \lambda^2, \phi + \lambda^2/2)$ (right). For readability of the figures, we should like to 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 5, rejections probabilities are always larger than 0.5 and increase with $\lambda$. When $\phi$ is positive, and $\lambda$ close to zero the power is at its minimum. This corresponds to nonstationary process. Correspondingly, the right-hand side columns show that the power is low when $\phi + \lambda^2$ is positive but low and $\phi + \lambda^2/2$ is large, i.e. when the variance of the random coefficient is low. 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 4. This confirms the analysis by Evans (1991) that stochastic bubbles, being non-permanent by nature, can be difficult to detect even when their magnitude is large.
<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymptotic Distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.2</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.93</td>
<td>0.90</td>
<td>0.89</td>
<td>0.89</td>
</tr>
<tr>
<td>0</td>
<td>0.95</td>
<td>0.92</td>
<td>0.94</td>
<td>0.93</td>
<td>0.92</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>1.00</td>
<td>0.91</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>1.00</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.94</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>0.1</td>
<td>1.00</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.93</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td>0.2</td>
<td>1.00</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>0.3</td>
<td>1.00</td>
<td>0.94</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>$\phi$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian Wild Bootstrap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
<td>0.99</td>
<td>0.97</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>0.01</td>
<td>1.00</td>
<td>0.98</td>
<td>0.97</td>
<td>0.96</td>
<td>0.98</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td>0.05</td>
<td>1.00</td>
<td>0.98</td>
<td>0.97</td>
<td>0.95</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>0.1</td>
<td>1.00</td>
<td>0.99</td>
<td>0.96</td>
<td>0.93</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td>0.2</td>
<td>1.00</td>
<td>1.00</td>
<td>0.96</td>
<td>0.89</td>
<td>0.96</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>0.3</td>
<td>1.00</td>
<td>1.00</td>
<td>0.96</td>
<td>0.85</td>
<td>0.94</td>
<td>0.97</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Table 1: Simulated Finite Sample Probability Coverage of one-sided confidence intervals at an asymptotic nominal probability of 0.95, together without quantiles obtained using the Gaussian wild bootstrap. The simulated sample size is $T = 3,000$ with $\alpha = 1/2$. The number of Monte Carlo replications is 10,000 and so is the sample size used in computing the asymptotic distribution.
<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi )</td>
<td>Asymptotic Distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.2</td>
<td>0.98</td>
<td>0.98</td>
<td>0.85</td>
<td>0.74</td>
<td>0.69</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.95</td>
<td>0.77</td>
<td>0.83</td>
<td>0.87</td>
<td>0.83</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>1.00</td>
<td>0.78</td>
<td>0.84</td>
<td>0.87</td>
<td>0.83</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>1.00</td>
<td>0.86</td>
<td>0.90</td>
<td>0.91</td>
<td>0.88</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>1.00</td>
<td>0.91</td>
<td>0.94</td>
<td>0.93</td>
<td>0.91</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>1.00</td>
<td>0.95</td>
<td>0.95</td>
<td>0.93</td>
<td>0.92</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>1.00</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.93</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>( \phi )</td>
<td>Gaussian Wild Bootstrap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>0</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.97</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>0.01</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.97</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>0.05</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.97</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>0.1</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.96</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>0.2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.95</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>0.3</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.94</td>
<td>0.96</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table 2: Simulated Finite Sample Probability Coverage of one-sided confidence intervals at an asymptotic nominal probability of 0.95, together without quantiles obtained using the Gaussian wild bootstrap. The simulated sample size is \( T = 300 \) with \( \alpha = 1/2 \). The number of MonteCarlo replications is 10,000 and so is the sample size used in computing the asymptotic distribution.
<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.5</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.2</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.68</td>
<td>0.72</td>
<td>0.91</td>
<td>0.94</td>
</tr>
<tr>
<td>0</td>
<td>0.98</td>
<td>0.64</td>
<td>0.64</td>
<td>0.63</td>
<td>0.82</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>0.01</td>
<td>1.00</td>
<td>0.63</td>
<td>0.64</td>
<td>0.63</td>
<td>0.82</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>0.05</td>
<td>1.00</td>
<td>0.55</td>
<td>0.59</td>
<td>0.64</td>
<td>0.83</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>0.1</td>
<td>1.00</td>
<td>0.60</td>
<td>0.68</td>
<td>0.71</td>
<td>0.85</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>0.2</td>
<td>1.00</td>
<td>0.83</td>
<td>0.86</td>
<td>0.84</td>
<td>0.87</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>0.3</td>
<td>1.00</td>
<td>0.93</td>
<td>0.94</td>
<td>0.92</td>
<td>0.89</td>
<td>0.91</td>
<td>0.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\phi$</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian Wild Bootstrap</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>-0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Simulated Finite Sample Probability Coverage of two-sided confidence intervals at an asymptotic nominal probability of 0.95, together without quantiles obtained using the Gaussian wild bootstrap. The simulated sample size is $T = 300$ with $\alpha = 1/2$. The number of MonteCarlo replications is 10,000 and so is the sample size used in computing the asymptotic distribution.
Figure 3: Upper-tail rejection probabilities at the asymptotic nominal size of 5% 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 the parameter $\alpha = .50$. Panels on the left-hand side reproduce those on the right-hand side, but with different axes where $\log \mathbb{E}(\rho_t) = \phi + \lambda^2 / 2$ and $\log \mathbb{E}(\rho_t^2) = \phi + \lambda^2$.

Turning to figure 4, where a lower $\alpha$ 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 $\lambda$. 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 $\rho_t$ become large (upper right-hand side panel).

Finally, the bilateral rejection probabilities presented in figure 3 show that this test has power of at least 50% when $\phi + \lambda^2 > 0$ but that it is unable to discriminate between a constant and a random coefficient when the process is strictly stationary but not weakly so ($\phi + \lambda^2 < 0$) under the alternative, but nonstationary under the null ($\phi + \lambda^2 / 2 \geq 0$): the bilateral test does not reject the null of an explosive AR(1) under the alternative that it is a stationary RCA with power law distribution. Interestingly, the one-sided test does.
Figure 4: Upper-tail rejection probabilities at the nominal size of 5% corresponding to the null hypothesis $H_0 : (\phi_0, 0)$ under the alternative $H_1 : (\phi, \lambda)$ 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 $\log E(\rho_t) = \phi + \lambda^2 / 2$ and $\log E(\rho_t^2) = \phi + \lambda^2$. 
Figure 5: Bilateral rejection probabilities at the nominal size of 5% 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 \( \log E(\rho_t) = \phi + \lambda^2/2 \) and \( \log E(\hat{\rho}_t^2) = \phi + \lambda^2 \).
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 housing prices in the U.S. Standard models relate the price at time $t$ ($P_t$) 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

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

PY show that when $r_t$ varies, there may exist subperiods where the ratio of price over rent $P_t/D_t$ exhibits explosiveness as opposed to martingale behavior (assuming $\lambda = 0$ in their models). Hence, in their application, they test whether they can reject the null that $P_t/D_t$ follows a random walk against an explosive alternative. We extend their work here to the NERC model.

In addition, we recognize that the NERC may generate downward 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),

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

(17)

where $\mathbb{E}_t[\cdot]$ 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 and the simplifying assumption that the ex-post return is constant and equal to $R$, the present-value relation (17) then admits the solution (with minimal number of state variables, see McCallum, 1983)

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

(18)

where $\Delta D_t = \zeta_t$ is i.i.d and $\delta \in [0,1]$. Hence a large value of $R_t$ may generate explosiveness in $\Delta P_t$. Assuming $R_t$ iid and uncorrelated with $\zeta_t$, the dynamics of $\Delta P_t$ can be represented using the NERC model considered previously.

In the following, we apply our methodology both to $P_t/D_t$ and $\Delta P_t$ 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). 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. The series is presented 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 $\Delta P_t$ both exhibit patterns similar to those that arise under the NERC model.
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 $\Delta P_t$ and ratio $P_t/D_t$. 

22
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 \( \alpha = 1/2 \).

Figures 7 and 8 report on the top row the confidence sets obtained in the absence of both upper-tail (in red) and bilateral (in blue) rejection. We set the significance levels to see the impact of right- and left-tail rejection. We report in particular confidence 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 the 1% 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 \( \Delta P_t \) and \( P_t/D_t \) respectively (a supplementary appendix records the corresponding sets based on the distributions obtained using the wild bootstrap).

Inference based on \( \Delta P_t \) and \( P_t/D_t \) lead to different dynamics so we comment on them in turn. Figure 7 shows that upper-tail tests based on the asymptotic distribution lead to confidence sets that are predominantly with then \( S_{\lambda\wedge w} \) region that correspond to strictly yet non-weakly stationary processes. Yet, some of the parameter values close to the limit \( E[\rho_t^2] = 1 \) 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 test: these reject parameter values in \( S_w \) and even yield confidence sets for which not only \( E[\rho_t^2] > 1 \) but even \( E[\rho_t] \geq 1 \). 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 4 the minimum median-distance estimates, \((\phi, \lambda) = (0.17, 0.20)\), which fall in the explosive region.

Now, figure 8 reports inference on \( P_t/D_t \) based on the asymptotic distribution. The unilateral confidence sets are now predominantly close to the border between \( S_w \) and \( S_{\lambda\wedge w} \). 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 \((\phi, \lambda) = (-0.25, 0.59)\) such that \( \phi + \lambda^2 = 0.10 \).

\footnote{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 and blue parameter combinations.}
Figure 7: The figure reports inferential results on the NERC applied to $\Delta P_t$. The top row records the confidence sets computed as parameter combinations which are not rejected using the asymptotic distribution of the OLS estimator of $\hat{\rho}$. Panel (a): the dots define the 90% 1-sided and 80% 2-sided confidence sets. Panel (b): the dots define the 10% 1-sided and 20% 2-sided confidence sets. The bottom row refers to the locus of median estimates (parameters implying $\hat{\rho}$ 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 of the OLS estimator of $\hat{\rho}$. Panel (a) : the dots define the 90% 1-sided and 80% 2-sided confidence sets. Panel (b) : the dots define the 10% 1-sided and 20% 2-sided confidence sets. The bottom row refers to the locus of median estimates (parameters implying $\hat{\rho}$ 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.
<table>
<thead>
<tr>
<th>Least Rejected Median Estimate</th>
<th>Univariate Confidence Interval</th>
<th>Test ((\phi, \lambda) = (0, 0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\phi, \lambda)^*)</td>
<td>((\phi, \lambda)^+)</td>
<td>0.90 0.80 0.10 (pvalue)</td>
</tr>
</tbody>
</table>

**One-Sided Test**

<table>
<thead>
<tr>
<th>(\Delta P_t, \hat{\rho} = 0.972)</th>
<th>(E[\rho_t^*] = 1.00)</th>
<th>(E[\rho_t^{*+}] = 1.01)</th>
<th>(\lambda : [0, .99])</th>
<th>(\lambda : [0, .99])</th>
<th>(\lambda : [0, .93])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(pvalue : 1.00)</td>
<td>(pvalue : 0.50)</td>
<td>(\phi : [-.50, .24])</td>
<td>(\phi : [-.50, .24])</td>
<td>(\phi : [-.50, .24])</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(P_t/D_t, \hat{\rho} = 1.00)</th>
<th>(E[\rho_t^*] = 1.03)</th>
<th>(E[\rho_t^{*+}] = 1.00)</th>
<th>(\lambda : [0, .87])</th>
<th>(\lambda : [0, .84])</th>
<th>(\lambda : [0, .78])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(pvalue : 1.00)</td>
<td>(pvalue : 0.51)</td>
<td>(\phi : [-.50, .01])</td>
<td>(\phi : [-.50, .01])</td>
<td>(\phi : [-.50, .01])</td>
<td></td>
</tr>
</tbody>
</table>

**Two-Sided Test**

<table>
<thead>
<tr>
<th>(\Delta P_t, \hat{\rho} = 0.972)</th>
<th>(E[\rho_t^*] = 1.01)</th>
<th>(E[\rho_t^{*+}] = 1.01)</th>
<th>(\lambda : [0, .99])</th>
<th>(\lambda : [0, .99])</th>
<th>(\lambda : [0, .99])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(pvalue : 1.00)</td>
<td>(pvalue : 1.00)</td>
<td>(\phi : [-.50, .22])</td>
<td>(\phi : [-.50, .22])</td>
<td>(\phi : [-.50, .22])</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(P_t/D_t, \hat{\rho} = 1.00)</th>
<th>(E[\rho_t^*] = 1.00)</th>
<th>(E[\rho_t^{*+}] = 1.00)</th>
<th>(\lambda : [0, .87])</th>
<th>(\lambda : [0, .84])</th>
<th>(\lambda : [0, .83])</th>
</tr>
</thead>
<tbody>
<tr>
<td>(pvalue : .99)</td>
<td>(pvalue : 0.99)</td>
<td>(\phi : [-.50, .002])</td>
<td>(\phi : [-.50, .0])</td>
<td>(\phi : [-.50, -.005])</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: The table reports statistics regarding inference on the dynamics of \(\Delta P_t\).
6.3 Bubble Detection

We now turn to real time detection and prediction using the NERC model. Since no technique is available to extract the latent \( i.i.d \) process \( u_t \), we apply grid testing to the recursive methodology addressed by PY. We specifically ask when the assumption \( E[\rho_t] = 1 \) can be rejected using an upper tailed test. Here the null is composite in the sense that we test all parameter combinations such that \( \phi + \lambda^2/2 = 0 \). Figure 9, panels (a) and (c), reports the minimum \( p \)-values obtained for 5,000 random combinations of the parameters \( (\phi, \lambda) \) under the null \( H_0 : \phi + \lambda^2/2 = 0 \). The distribution is asymptotic and the minimum \( p \)-value are computed recursively and indexed according to the end-of-sample date on the horizontal axis. The tests are one-sided. 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.\(^\text{17}\) We follow PY in interpreting rejection as evidence of a bubble. The one-sided test for \( \Delta P_t \) first detects a possible bubble at the 5% 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 \( P_t/D_t \) from May 2002 to November 2007. In their work, PY find 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).

Interestingly, the results based in \( \Delta P_t \) 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 in 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 \( \Delta P_t \) 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.\(^\text{18}\)

6.4 Forecasting

We finally apply our methodology the forecasting growth. We recursively compute the minimum \( p_{y_t,y_{t-k}}^{y_{t+k},y_{t-k}} \), i.e. minimum probabilities that growth over a forecast horizon of \( k = 1, 6, \) or 12 months

\(^{17}\)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.

\(^{18}\)Tto the exception of June 2008: on the 11th, the Securities and Exchange Commission unveiled its comprehensive reform of credit ratings.
The figure reports the output from detection of explosive behavior in $\Delta P_t$ 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 $(\phi, \lambda)$ such that $\phi + \lambda^2 = 0$, i.e. $E[\rho_t] = 1$. 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 $\Delta P_t$. The $p$-values are reported on the left axis, together with a scaled series of $\Delta P_t$. Panel (b): The shaded areas refer to periods where the composite null $E[\rho_t] = 1$ 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%.
will be as high as that observed over the latest \( k \) periods. The minimum probabilities are obtained over the set of parameters \((\phi, \lambda)\) such that \( c = \phi + \lambda^2 \geq 0 \). We only consider parameter sets which constitute the nominal 50% or 90% asymptotic confidence set.\textsuperscript{19} These minimum probabilities are reported in figures 10 and 11.

Minimum probability forecasts relating to \( \Delta P_t \) 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 positive in the 50% confidence set. The first period (1991-early 1992) correspond 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.

\section*{7 Conclusion}

The paper proposes a local asymptotic model that builds on random coefficient autoregressive processes. 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 appearance relates to the values taken by a latent process that relates to the stochastic discount factor. The generalization we propose presents similar inferential benefits to the univariate locally explosive AR(1) with break, while allowing for full-sample inference.

The paper aims also to provide an empirical application to validate the model of local-asymptotic NERC and show its applicability. On a theoretical side, it seems important to relax the assumption that \( u_t \) is i.i.d. since the latter is unlikely to hold in practice. Some persistence in the stochastic discount factor is indeed expected. Multivariate extensions would allow to consider spillovers between different sectors of the economy.

\section*{References}


\textsuperscript{19}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 \( P(\frac{y_{t+k}}{y_t} \geq \frac{y_t}{y_{t-k}}) \) computed under the asymptotic distribution under the null. The minimum is computed over 5,000 draws of parameters \((\phi, \lambda)\) such that \( c = \phi + \lambda^2 \geq 0 \) and \((\phi, \lambda)\) 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 \( \Delta P_t \). Each panel reports a different horizon: \( k = 1 \) (top panel), 6 (middle) or 12 months (bottom).
Figure 11: The figure reports the estimated minimum probabilities $P(y_{t+k}/y_t \geq y_t/y_{t-k})$ computed under the asymptotic distribution under the null. The minimum is computed over 5,000 draws of parameters $(\phi, \lambda)$ such that $c = \phi + \lambda^2 \geq 0$ and $(\phi, \lambda)$ 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).


8 Appendix

8.1 Simulated NERC paths.

In order to show the sort of dynamics the model generates, figure 8.1 records simulations of the process over samples of $T = 1000$ observations using two sets of draws of $(u_t, \eta_t)$. Exuberant periods become clearly more pronounced and explosive as $\phi$ increases or $\alpha$ 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 $\alpha$ 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).

Proof of Proposition 1

The proof of Proposition 1 is based on the following lemma.

**Lemma 6** Let the process $y_t$ be defined for $t \geq 0$ by (3) under assumptions 1 and 2. For $r \in \left[0, T^{1-\alpha}\right]$ and as $T \to \infty$, it holds

$$T^{-\alpha/2}y_{\lfloor rT^{\alpha}\rfloor} \Rightarrow K_{\phi, \lambda}(r) \equiv \int_{0}^{r} \exp \{ (r-s)\phi + \lambda (W_r - W_s) \} dB_s$$

where $W, B$ are two independent standard Brownian motion such that, for $(s,r) \in [0,1]^2$,

$$T^{-1/2} \left( \sum_{t=1}^{\lfloor sT \rfloor} u_t, \sigma_\eta^{-1} \sum_{t=1}^{\lfloor rT \rfloor} \eta_t \right) \Rightarrow (W_s, B_r),$$

$\lfloor \cdot \rfloor$ denoting the integer part.

**Proof.** Proof of Lemma 6
Figure 12: Simulated realizations from the model of autoregressive conditional exuberance for different parameter values.
We have, given \( y_0 \), and setting \( \prod_{j=0}^{t-1} \rho_j := 1 \)
\[
y_t = \left( \prod_{j=0}^{t-1} \rho_{t-j} \right) y_0 + \sum_{i=0}^{t-1} \left( \prod_{j=0}^{i-1} \rho_{t-j} \right) \eta_{t-i}
\]
\[
= \left( \prod_{i=1}^{t} \rho_i \right) y_0 + \sum_{i=1}^{t} \left( \prod_{j=i+1}^{t} \rho_j \right) \eta_i
\]
\[
= \exp \left\{ \frac{tT^{-\alpha/2} \phi + \lambda S_t}{T^{\alpha/2}} \right\} y_0 + \sum_{i=1}^{t} \exp \left\{ \frac{(t-i)T^{-\alpha/2} \phi + \lambda (S_i - S_t)}{T^{\alpha/2}} \right\} \eta_i
\]

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 \( j = 0, \ldots, \lfloor T^{1-\alpha} \rfloor - 1 \), and \( k = 1, \ldots, \lfloor T^\alpha \rfloor \), and letting \( k = \lfloor pT^\alpha \rfloor \) for some \( p \in [0, 1] \), we can write
\[
\frac{1}{T^{\alpha/2}} (|y|_{jT^\alpha} + |pT^\alpha| - y_0) = \frac{1}{T^{\alpha/2}} \left( \exp \left\{ \frac{|jT^\alpha| + |pT^\alpha|}{T^{\alpha/2}} \phi + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + |pT^\alpha|} - S_t}{T^{\alpha/2}} \right\} - 1 \right) y_0 +
\]
\[
\sigma_\eta \sum_{i=1}^{\lfloor jT^\alpha \rfloor + |pT^\alpha|} \exp \left\{ \frac{|jT^\alpha| + |pT^\alpha| + i - \lfloor jT^\alpha \rfloor}{T^{\alpha/2}} \phi + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + |pT^\alpha|} + pT^\alpha} - S_{\lfloor jT^\alpha \rfloor + |pT^\alpha|} \right\} \frac{\eta_i}{\sigma_\eta T^{\alpha/2}}
\]
\[
= \frac{1}{T^{\alpha/2}} \left( \exp \left\{ \frac{|jT^\alpha| + |pT^\alpha|}{T^{\alpha/2}} \phi + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + |pT^\alpha|} - S_t}{T^{\alpha/2}} \right\} - 1 \right) y_0 +
\]
\[
\sigma_\eta \int_0^{j+p} \exp \left\{ \frac{|jT^\alpha| + |pT^\alpha| - |sT^\alpha|}{T^{\alpha/2}} \phi + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + |pT^\alpha|} - S_{\lfloor sT^\alpha \rfloor}}{T^{\alpha/2}} \right\} dB_{T^\alpha}(s)
\]

using Proposition A1 in Phillips and Magdalinos (2004) in the last equality, where
\[
B_{T^\alpha}(s) := \frac{1}{\sigma_\eta T^{\alpha/2}} \sum_{i=1}^{\lfloor sT^\alpha \rfloor} \eta_i
\]

When applying the Functional Central Limit Theorem (FCLT) to the process \( \tilde{S}_T \) defined by \( \tilde{S}_T(s) := \frac{S_{\lfloor sT^\alpha \rfloor}}{\sqrt{T^\alpha}} \) \( 0 \leq s \leq 1 \), 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 \( B_{T^\alpha} \) defined in (19) converges in distribution, as \( T \to \infty \), to a BM on \([0, 1]\), say \( B \), which, by assumption on the sequences \( (u_t) \) and \( (\eta_j) \), is independent of \( W \).

Then we can deduce, using e.g. theorem 8.3.1 in Liptser and Shiryaev (1989), that
\[
\int_0^{j+p} \exp \left\{ \phi \frac{|jT^\alpha| + |pT^\alpha| - |sT^\alpha|}{T^{\alpha/2}} + \lambda \frac{S_{\lfloor jT^\alpha \rfloor + |pT^\alpha|} - S_{\lfloor sT^\alpha \rfloor}}{T^{\alpha/2}} \right\} dB_{T^\alpha}(s)
\]
converges, as \( T \to \infty \), to
\[
\int_0^r \exp \{ (r-s)\phi + \lambda (W_r - W_s) \} dB_s, \quad \text{with} \quad r = j + p.
\]

This last integral can be written as
\[
e^{r\phi + \lambda W_r} \int_0^r X_s dY_s \quad \text{where} \quad X_s = e^{-\lambda W_s} \quad \text{and} \quad dY_s = e^{-s\phi} dB_s.
\]

37
The covariation process of two independent BM being identically 0 (see e.g. Klebaner (2005), theorem 4.19), the stochastic integration by parts reduces to the usual integration by parts formula and provides
\[
e^{r\phi + \lambda W_r} \int_0^r X_s dY_s = \int_0^r e^{r(s-s)}\phi dB_s - e^{r\phi + \lambda W_r} \int_0^r Y_s dX_s = J_\phi(r) - e^{r\phi + \lambda W_r} \int_0^r Y_s dX_s \quad (20)
\]
where \( J_\phi(r) = \int_0^r e^{(r-s)}\phi dB_s \) is the Ornstein-Uhlenbeck process that corresponds to the limit obtained in Phillips and Magdalinos (2004).

Since \( X_s \) satisfies the stochastic differential equation (SDE) \( dX_s = \lambda X_s ds - \lambda X_s dW_s \), the second term on the right-hand side (RHS) of (20) can be written as
\[
e^{r\phi + \lambda W_r} \left( \frac{\lambda}{2} \int_0^r X_s Y_s ds - \lambda \int_0^r X_s Y_s dW_s \right) \quad (21)
\]
with \( Y_s = \int_0^s e^{-u\phi} dB_u = e^{-s\phi} B_s + \phi \int_0^s e^{-u\phi} B_u du \). The results of the Lemma then follow. □

The process \( K_{\phi,\lambda}(r) \) defined in Lemma 6 is distributed as
\[
K_{\phi,\lambda}(r) \sim N \left( 0, \sigma^2 \int_0^r e^{2cs} ds \right), \quad \text{for } r \in [0, T^{1-\alpha}]
\]
with
\[
c = \phi + \lambda^2 \quad (22)
\]
and where the integral \( \int_0^r e^{2cs} ds = \frac{e^{2cr} - 1}{2c} I_{c \neq 0} + r I_{c=0} \). This shows that the parameter \( c \) plays a major role for the asymptotics of the process: only if \( c < 0 \) does \( K_{\phi,\lambda}(r) \) admit a stationary version; it is then given by
\[
K_{\phi,\lambda}^*(r) = e^{cr} K_{\phi,\lambda}^*(0) + K_{\phi,\lambda}(r)
\]
where \( K_{\phi,\lambda}^*(0) \sim N \left( 0, \sigma^2 / (-2c) \right) \) is independent of \( K_{\phi,\lambda}(r) \). □

8.2 Proof of Theorem 3

- We have, as \( T \to \infty \),
\[
y_t = \left( E \left( \rho_t \right) + \lambda T^{-\alpha/2} u_t + O_p \left( T^{-\alpha} \right) \right) y_{t-1} + \eta_t
\]
Then the OLS estimator given by \( \hat{\rho} = \frac{\sum_{t=1}^T y_{t-1} y_t}{\sum_{t=1}^T y_{t-1}^2} \) satisfies
\[
\hat{\rho} - E \left( \rho_t \right) = \lambda T^{-\alpha/2} \frac{\sum_{t=1}^T y_{t-1}^2 u_t}{\sum_{t=1}^T y_{t-1}^2} + \frac{\sum_{t=1}^T y_{t-1} \eta_t}{\sum_{t=1}^T y_{t-1}^2} + O_p \left( T^{-3\alpha/2} \right). \quad (23)
\]
This shows that the asymptotic distribution of the estimator is driven by the term with higher magnitude between \( T^{-\alpha/2} \sum_{t=1}^T y_{t-1}^2 u_t \) and \( \sum_{t=1}^T y_{t-1} \eta_t \).
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 \( y_0 = o_p(T^{\alpha/2}) \) implies it is negligible compare to \( y_{[T^{\alpha}]} \) for \( 0 < r < T^{1-\alpha} \).

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

- **Case** \( c < 0 \)

  Proposition 1 gives \( T^{-\alpha/2}y_{[T^{\alpha}]} \Rightarrow K_{\phi,\lambda}(r) \sim N\left(0, \frac{e^{2cr} - 1}{2c} \sigma_n^2 \right) \); for \( c < 0 \), we can write \( K^*_{\phi,\lambda}(r) = e^{cr}K^*_{\phi,\lambda}(0) + K_{\phi,\lambda}(r) \). So

  \[
  K^*_{\phi,\lambda}(r) \sim N \left(0, \frac{-\sigma_n^2}{2c} \right)
  \]

  and is stationary.

  We can deduce that, via the LLN (Law of Large Numbers),

  \[
  \frac{1}{T^{1+\alpha}} \sum_{t=1}^{T} y_t^2 \Rightarrow E[K^*_{\phi,\lambda}(r)] = -\frac{\sigma_n^2}{2c}
  \]

  and that

  \[
  T^{-\frac{1+2\alpha}{\alpha}} \sum_{t=1}^{T} y_{t-1} \eta_t \Rightarrow N\left(0, \frac{-\sigma_n^2}{2c} \right)
  \]

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

  \[
  \sum_{t=1}^{T} E_{t-1} \left(\xi_t^2 \right) = \frac{1}{T^{1+2\alpha}} \sum_{t=1}^{T} y_{t-1}^4 \Rightarrow 3\sigma_n^4 \frac{\alpha}{4c^2}
  \]

  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

  \[
  T^{-\frac{1+2\alpha}{\alpha}} \sum_{t=1}^{T} y_{t-1}^2 u_t \Rightarrow N\left(0, \frac{3\sigma_n^4}{4c^2} \right)
  \]

- **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 \( T = n \),

  \[
  \kappa_n = n^{\alpha} \left[n^{1-\alpha} \right] \quad \text{and} \quad q = n^{1-\alpha} - \left[n^{1-\alpha} \right]
  \]

  (i) We first consider the sample variance of \( y_t \) and show the following.
Lemma 7 Define

\[ \psi_{n^{1-\alpha}} = \begin{cases} 
\frac{1}{2\sqrt{(c+2\lambda^2)(c+\lambda^2)}} & \text{if } c \neq 0 \\
2c^{2}\lambda_{n^{1-\alpha}} & \text{if } c = 0 
\end{cases} \]

and \[ \varphi_{n^{1-\alpha}} = \begin{cases} 
\frac{1}{2(\phi-\lambda^2)} & \text{if } \lambda^2 - \phi < 0 \\
\sqrt{n^{1-\alpha}} & \text{if } \lambda^2 - \phi = 0 \\
\frac{\lambda_{n^{1-\alpha}}}{\sqrt{2(\lambda^2-\phi)}} & \text{if } \lambda^2 - \phi > 0 
\end{cases} \]

Then, as \( n \to \infty \),

\[ n^{-2\alpha}\psi_{n^{1-\alpha}}^{-1}\varphi_{n^{1-\alpha}}^{-2} \sum_{t=1}^{n} y_t^2 \Rightarrow X^2Z \]

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

\[ \frac{\varphi_{[n^{1-\alpha}]}}{n^{\alpha/2}} \sum_{i=1}^{[\kappa_n]} \exp \left( -\frac{\phi}{n^{\alpha}} i - \frac{\lambda}{n^{\alpha/2}} U_i \right) \eta_i \Rightarrow X \sim N(0, \sigma_n^2) \]

and

\[ \psi_{[n^{1-\alpha}]^{-1}} \int_{0}^{[n^{1-\alpha}]} e^{2\phi s+2\lambda W_s} ds \Rightarrow Z \text{ mean } 0 \text{ and unit variance} \]

Proof of Lemma 7.

We write

\[ \frac{1}{n^{2\alpha}} \sum_{t=1}^{n} y_t^2 = \frac{1}{n^{2\alpha}} U_{1n} + \frac{1}{n^{2\alpha}} U_{2n} + O_p \left( \frac{1}{n^{\alpha}} \right) \tag{24} \]

with

\[ U_{1n} := \sum_{j=0}^{[n^{1-\alpha}] - 1} \sum_{k=1}^{n^{\alpha} - |n^{\alpha} + k|} y_{[n^{\alpha} + k]}^2 \]

and

\[ U_{2n} := \sum_{t=\left[\kappa_n\right]}^{n} y_t^2 \]

Note that the index of the last summation term in the definition of \( U_{1n} \), given by \( \left| \kappa_n - n^{\alpha} \right| + \left| n^{\alpha} \right| \), is bounded by \( \left| \kappa_n \right| - 1 \leq \left| \kappa_n - n^{\alpha} \right| + \left| n^{\alpha} \right| \leq \left| \kappa_n \right| \).

The study of \( U_{1n} \) leads to the following result.

As \( n \to \infty \),

\[ n^{-2\alpha}\psi_{[n^{1-\alpha}]}^{-1}\varphi_{[n^{1-\alpha}]}^{-2} U_{1n} \Rightarrow X^2Z \tag{25} \]

where the random variables \( X \) and \( Z \) are defined in Lemma 7.
Proof of (25).

Notice that

\[ y_k = \sum_{i=0}^{k-1} \exp \left( \frac{\phi}{n^\alpha} k + \frac{\lambda}{n^{\alpha/2}} \sum_{j=k-i+1}^{t} u_j \right) \eta_{k-i} \]

\[ = \sum_{i=1}^{k} \exp \left( \frac{\phi}{n^\alpha} (k-i) + \frac{\lambda}{n^{\alpha/2}} (U_k - U_i) \right) \eta_i \quad \text{with} \quad U_i := \sum_{j=1}^{i} u_j \]

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

so

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

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

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

where

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

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

Therefore we obtain

\[ U_{1n} = \left( \sum_{k=1}^{\lfloor n^\alpha \rfloor} \exp \left( \frac{2\phi}{n^\alpha} k + \frac{2\lambda}{n^{\alpha/2}} U_k \right) \right) \left[ \sum_{i=1}^{\lfloor n^\alpha \rfloor} \exp \left( -\frac{\phi}{n^\alpha} i - \frac{\lambda}{n^{\alpha/2}} U_i \right) \eta_i \right] + R_{\lfloor n^\alpha \rfloor} \quad (26) \]

Let us study the three elements given in this last equation (26).

We start by showing that \( n^{-2\alpha} \psi^{-1}_{n^{1-\alpha}} \varphi^{-2}_{n^{1-\alpha}} R_{\lfloor n^\alpha \rfloor} \) is negligible w.r.t. \( \psi^{-1}_{n^{1-\alpha}} \varphi^{-2}_{n^{1-\alpha}} n^{-2\alpha} (U_{1n} + U_{2n}) \).

Indeed, writing \( R_t = R_{1t} - 2R_{2t} \) with

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

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

\[ := \left( \sum_{i=1}^{t} \exp \left( -\frac{\phi}{n^\alpha} i - \frac{\lambda}{n^{\alpha/2}} U_i \right) \eta_i \right) \times \mathcal{R}_{2t} \]
Hence we obtain that
\[ n^{-2\alpha} \psi^{-1}_{[n^{1-\alpha}]} \varphi^{-2}_{[n^{1-\alpha}]} R_1[\kappa_\alpha] = \psi^{-1}_{[n^{1-\alpha}]} \int_0^{[n^{1-\alpha}]} \left( \varphi^{-1}_{[n^{1-\alpha}]} \int_r^{[n^{1-\alpha}]} e^{\phi(r-s) + \lambda(W_r - W_s)} dB_s \right)^2 dr + o_p(1) \]
uniformly, and since we have
\[ E \left( \int_r^{[n^{1-\alpha}]} e^{\phi(r-s) - \lambda(W_r - W_s)} dB_s \right)^2 = \int_r^{[n^{1-\alpha}]} e^{2(\phi + \lambda^2)(r-s)} ds \]
uniformly in \( r \leq n^{1-\alpha} \), we obtain
\[ n^{-2\alpha} \psi^{-1}_{[n^{1-\alpha}]} \varphi^{-2}_{[n^{1-\alpha}]} R_1[\kappa_\alpha] = O_p(1) \times O \left( \psi^{-1}_{[n^{1-\alpha}]} \int_0^{[n^{1-\alpha}]} dr \right) = O_p \left( [n^{1-\alpha}] \psi^{-1}_{[n^{1-\alpha}]} \right) \]
On the other hand, we have \( n^{-\alpha} \psi^{-1}_{[n^{1-\alpha}]} \varphi^{-2}_{[n^{1-\alpha}]} R_2[\kappa_\alpha] = O_p \left( [n^{1-\alpha}] \psi^{-1}_{[n^{1-\alpha}]} \right) \), so it comes
\[ n^{-2\alpha} \psi^{-1}_{[n^{1-\alpha}]} \varphi^{-2}_{[n^{1-\alpha}]} R_2[\kappa_\alpha] = O_p \left( n^{1-3/2\alpha} \psi^{-1}_{[n^{1-\alpha}]} \varphi^{-1}_{[n^{1-\alpha}]} \right) \]
hence the result concerning \( R[\kappa_\alpha] \).

Now let us look at the second element on the RHS of equation (26). We can write it as
\[ \frac{1}{n^{\alpha/2}} \sum_{i=1}^{[\kappa_\alpha]} \exp \left( -\frac{\phi}{n^\alpha i} - \frac{\lambda}{n^{\alpha/2} U_i} \right) \eta_i = \sigma_\eta \int_0^{[n^{1-\alpha}]} e^{-\phi n^\alpha s - \lambda W_n^{\alpha s}} dB_n^{\alpha s} (s) + o_p(1) \]
\( B_n \) being defined in (19).
When \( \lambda^2 < \phi \), it admits limit \( \sigma_\eta \int_0^\infty e^{-\phi s - \lambda W_n^{\alpha s}} dB(s) \).
When \( \lambda^2 \geq \phi \), the stochastic integral is not defined, but since \( \int_0^{[n^{1-\alpha}]} e^{-(\phi s + \lambda W_n^{\alpha s})} dB_n^{\alpha s} \) is normally distributed, it will be enough to scale it by its standard deviation, using that
\[ \text{Var} \left( \int_0^{[n^{1-\alpha}]} e^{-(\phi s + \lambda W_n^{\alpha s})} dB_s \right) = \left\{ \begin{array}{ll} \frac{e^{2(\phi + \lambda^2)(n^{1-\alpha})} - 1}{n^{1-\alpha}} & \text{if } \lambda^2 > \phi \\ \frac{2(\phi + \lambda^2)(n^{1-\alpha})}{2(\phi + \lambda^2)} & \text{if } \lambda^2 = \phi \end{array} \right. \]
Hence we obtain that
\[ \frac{\varphi^{-1}_{[n^{1-\alpha}]}}{n^{\alpha/2}} \sum_{i=1}^{[\kappa_\alpha]} \exp \left( -\frac{\phi}{n^\alpha i} - \frac{\lambda}{n^{\alpha/2} U_i} \right) \eta_i \Rightarrow X \sim N(0, \sigma_\eta^2) \]
Finally, let us look at the first element on the RHS of (26). We have
\[ \psi^{-1}_{[n^{1-\alpha}]} \eta^{-2}_{[n^{1-\alpha}]} \sum_{k=1}^{[\kappa_\alpha]} \exp \left( \frac{2\phi}{n^\alpha k} + \frac{2\lambda}{n^{\alpha/2} U_k} \right) = \psi^{-1}_{[n^{1-\alpha}]} \int_0^{[n^{1-\alpha}]} e^{2\phi + 2\lambda W_n^{\alpha s}} ds + o_p(1) \]
Note that the rate $\psi_{[n^{1-\alpha}]}$ comes from the second moment of $\int_0^{[n^{1-\alpha}]} e^{2(\phi s + \lambda W_s)} ds$. Indeed, straightforward computations lead, for $c \geq 0$, to

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

$$= \int_0^{[n^{1-\alpha}]} \int_0^{[n^{1-\alpha}]} e^{2\phi(s+r) + 2\lambda^2(s+2 \min(r,s)+r)} ds dr$$

$$= \int_0^{[n^{1-\alpha}]} e^{2\phi r} e^{2(c+2\lambda^2)rs} ds dr + \int_0^{[n^{1-\alpha}]} \int_r^{[n^{1-\alpha}]} e^{2(c+2\lambda^2)r} ds dr$$

$$= \left\{ \begin{array}{ll}
\frac{e^{4(c+\lambda^2)}[n^{1-\alpha}^2]}{4(c+2\lambda^2)(c+\lambda^2)} - \frac{e^{2cn^{1-\alpha}}}{2(c+2\lambda^2)} + \frac{1}{4(c+\lambda^2)} = \frac{e^{4(c+\lambda^2)}[n^{1-\alpha}^2]}{4(c+2\lambda^2)(c+\lambda^2)} + O(e^{2cn^{1-\alpha}}) & \text{if } c \neq 0 \\
\frac{e^{4\alpha^2}}{8\lambda^2} + O(n^{1-\alpha}) & \text{if } c = 0
\end{array} \right.$$ 

Now, since $\psi_{[n^{1-\alpha}]}^{-1} n^{1-\alpha}$ is bounded, using Matsumoto and Yor (2005), theorem 7.4, provides there exists $Z$ such that

$$\psi_{[n^{1-\alpha}]}^{-1} \int_0^{[n^{1-\alpha}]} e^{2(\phi s + \lambda W_s)} ds \Rightarrow Z$$

where $Z$ is a random variable with mean zero expectation and unit variance.

Hence the result of (25). \qed

Let us now consider the second term $U_{2n}$ in equation (24). We have

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

where, for all $j = 0, \ldots, [n^{1-\alpha}] - 1$, as $n \to \infty$,

$$n^{-\alpha/2} y_j [y_j + [n^{\alpha}]] \Rightarrow \int_0^{j+p} e^{\phi(j+p-s)+\lambda(W_j+p-W_s)} dB_s = e^{\phi(j+p)+\lambda W_{j+p}} \int_0^{j+p} e^{-\phi s - \lambda W_s} dB_s$$

Then it comes

$$\frac{1}{n^{2\alpha}} U_{2n} = \int_0^q e^{2\phi([n^{1-\alpha}] + \lambda W_s)} \left( \int_0^{[n^{1-\alpha}] + \lambda W_s} e^{-\phi s - \lambda W_s} dB_s \right)^2 dp + o_p (1)$$

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

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

hence

$$U_{2n} = \left( \int_0^{[n^{1-\alpha}]} e^{-\phi s - \lambda W_s} dB_s \right)^2 \int_0^{[n^{1-\alpha}]} e^{2(\phi s + \lambda W_s)} ds - \frac{1}{n^{2\alpha}} \left( \int_0^{[n^{1-\alpha}]} e^{-\phi s - \lambda W_s} dB_s \right)^2 U_{1n} + o_p (1)$$

43
Combining (24), (25), (27), and the asymptotic equivalence \( \psi_{n^{-\alpha}}^{-1} \psi_{n^{-\alpha}} = 1 + o(1) \) allows to conclude to Lemma 7.

(ii) Now consider the covariance terms.

**Lemma 8** We have, as \( n \to \infty \),

\[
n^{-\alpha} \varphi_{n^{-\alpha}}^{-1} \phi_{n^{-\alpha}}^{-1} \sum_{t=1}^{n} y_{t-1} \eta_{t} \Rightarrow XY
\]

and

\[
n^{-3\alpha/2} \varphi_{n^{-\alpha}}^{-2} \lambda_{n^{-\alpha}}^{-1} \sum_{t=1}^{n} y_{t-1}^{2} \eta_{t} \Rightarrow X^{2}V
\]

where \( X \sim N(0, \sigma_{\eta}^{2}) \), \( Y \sim N(0, \sigma_{\eta}^{2}) \), \( V \sim N(0, 1) \),

and with \( \varphi_{n^{-\alpha}} : \text{defined in Lemma 7} \)

\[
\chi_{n^{-\alpha}} = \frac{e^{2(\alpha+\lambda^{2})n^{-\alpha}}}{2\sqrt{c+\lambda^{2}}}
\]

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

**Proof of Lemma 8.**

Note that

\[
\text{var} \left( \int_{0}^{n^{-\alpha}} e^{\phi s + \lambda W_{s}} dB_{s} \right) = \mathbb{E} \left[ \left( \int_{0}^{n^{-\alpha}} e^{\phi s + \lambda W_{s}} dB_{s} \right)^{2} \right] = \mathbb{E} \left[ \int_{0}^{n^{-\alpha}} e^{2\phi s + 2\lambda W_{s}} ds \right]
\]

\[
= \int_{0}^{n^{-\alpha}} e^{2cs} ds = \begin{cases} 
\frac{e^{2cn^{-\alpha}-1}}{2c} & \text{if } c \neq 0 \\
n^{-\alpha} & \text{if } c = 0 
\end{cases}
\]

i.e. \( \text{var} \left( \int_{0}^{n^{-\alpha}} e^{\phi s + \lambda W_{s}} dB_{s} \right) = \phi_{n^{-\alpha}} \)

so we have

\[
\phi_{n^{-\alpha}}^{-1} \sigma_{\eta} \int_{0}^{n^{-\alpha}} e^{\phi s + \lambda W_{s}} dB_{s} \Rightarrow Y \sim N(0, \sigma_{\eta}^{2})
\]

Hence we can write

\[
\frac{\varphi_{n^{-\alpha}}^{-1} \phi_{n^{-\alpha}}^{-1} \sum_{t=1}^{n} y_{t-1} \eta_{t}}{n^{\alpha}} = \left( \varphi_{n^{-\alpha}}^{-1} \int_{0}^{n^{-\alpha}} e^{-\phi s + \lambda W_{s}} dB_{s} \right) \phi_{n^{-\alpha}}^{-1} \sigma_{\eta} \left( \int_{0}^{n^{-\alpha}} e^{\phi r + \lambda W_{r}} dB_{r} \right) + I_{n}
\]

44
where $I_n$ can be shown to be negligible, referring to Phillips and Magdalinos (2004). Then

$$n^{-\alpha} \phi_{n^1-\alpha}^{-1} \sum_{t=1}^{n} y_{t-1} \eta_t \Rightarrow XY$$

Now let us consider $\sum_{t=1}^{n} y_t^2 u_{t+1}$. It can be expressed as

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

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

where

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

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

$$:= R_{1t}^* - 2R_{2t}^*$$ with

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

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

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

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

The proof follows then the same line as for $R_t$ (in the proof of (25)).

Finally, let us look at the summation (28).

Notice that

$$\mathbb{E} \left( \int_{0}^{n^{1-a}} e^{4(\phi r + \lambda W_r)} dr \right) = \frac{e^{4(c+\lambda^2)n^{1-a}}}{4(c + \lambda^2)} - \frac{1}{4(c + \lambda^2)} = \chi_n^{2} + O(1)$$

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

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

45
which admits conditional variance such that
\[
\sum_{k=1}^{n-1} \frac{n^{-\alpha} \chi_n^{-2} \sum_{k=1}^{n-1} \exp \left( \frac{4\lambda}{n^{\alpha/2}} U_k \right) \chi_n^{1-\alpha} \int_0^{n^{1-\alpha}} e^{4(\phi r + \lambda W_r)} dr + o_p(1) \\
= O_p(1)
\]

It follows that
\[
\sum_{k=1}^{n-1} \frac{\exp \left( \frac{2\phi}{n^{\alpha/2}} U_k \right) \chi_n^{1-\alpha} \int_0^{n^{1-\alpha}} e^{2(\phi r + \lambda W_r)} dW_r + o_p(1)
\]

so
\[
n^{-3\alpha/2} \varphi_n^{1-\alpha} \chi_n^{1-\alpha} \sum_{t=1}^{n} y_{t-1}^2 u_t = \left( \varphi_n^{1-\alpha} \sigma_n \int_0^{n^{1-\alpha}} e^{-(\phi s + \lambda W_s)} dB_s \right)^2 \chi_n^{1-\alpha} \int_0^{n^{1-\alpha}} e^{2(\phi r + \lambda W_r)} dW_r + o_p(1)
\]
and
\[
n^{-3\alpha/2} \varphi_n^{1-\alpha} \chi_n^{1-\alpha} \sum_{t=1}^{n} y_{t-1}^2 u_t \Rightarrow X^2 V
\]

where $V$ is defined as
\[
\chi_n^{1-\alpha} \int_0^{n^{1-\alpha}} e^{2(\phi r + \lambda W_r)} dW_r \Rightarrow V \sim N(0, 1).
\]

- 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

\[
S_{y} = \sum_{t=1}^{T} y_t^2, \quad S_{yd} = \sum_{t=1}^{T} y_{t-1} \eta_t \quad \text{and} \quad S_{yyu} = \sum_{t=1}^{T} y_{t-1}^2 u_t
\]

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 $x \in \{yy, y\eta, yyu\}$,

\[
\sigma_n^{-2} \mu^x \phi_T^x S_x \Rightarrow U_x
\]

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

<table>
<thead>
<tr>
<th>$\phi_T^{yy}$</th>
<th>$\phi_T^{y\eta}$</th>
<th>$\phi_T^{yyu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c &lt; 0$</td>
<td>$T^{-1(1+\alpha)}$</td>
<td>$T^{-\frac{1+\alpha}{2}}$</td>
</tr>
<tr>
<td>$c = 0$</td>
<td>$T^{-2\alpha} e^{-6\lambda^2 T^{1-\alpha}}$</td>
<td>$T^{-1} e^{-2\lambda^2 T^{1-\alpha}}$</td>
</tr>
<tr>
<td>$c &gt; 0$</td>
<td>$T^{-2\alpha} e^{-2(\phi+c\lambda^2) T^{1-\alpha}}$</td>
<td>$T^{-\alpha} e^{-2\phi T^{1-\alpha}}$</td>
</tr>
<tr>
<td>$\lambda^2 &lt; \phi$</td>
<td>$T^{-2\alpha} e^{-2(\phi+c\lambda^2) T^{1-\alpha}}$</td>
<td>$T^{-\alpha} e^{-2\phi T^{1-\alpha}}$</td>
</tr>
<tr>
<td>$\lambda^2 = \phi$</td>
<td>$T^{-2\alpha} e^{-2(\phi+c\lambda^2) T^{1-\alpha}}$</td>
<td>$T^{-\alpha} e^{-2\phi T^{1-\alpha}}$</td>
</tr>
<tr>
<td>$\lambda^2 &gt; \phi$</td>
<td>$T^{-2\alpha} e^{-6\lambda T^{1-\alpha}}$</td>
<td>$T^{-\alpha} e^{-(\phi+2\lambda^2) T^{1-\alpha}}$</td>
</tr>
</tbody>
</table>
\[
\begin{array}{|c|c|c|c|}
\hline
& \mu_{yy} & \mu_{y\eta} & \mu_{yyu} \\
\hline
c < 0 & -2c & \sqrt{-2c} & -2c/\sqrt{3} \\
c = 0 & 8\sqrt{2} \lambda^4 & 2\lambda & 8\lambda^3 \\
c > 0 & \lambda^2 < \phi & 8(c - 2\lambda^2) \sqrt{(c + 2\lambda^2)(c + \lambda^2)} & 4c (c - 2\lambda^2) & 8(c - 2\lambda^2)^2 \sqrt{c + \lambda^2} \\
\lambda^2 = \phi & 2\sqrt{(c + 2\lambda^2)(c + \lambda^2)} & 2c & 2\sqrt{c + \lambda^2} \\
\lambda^2 > \phi & 4(2\lambda^2 - c) \sqrt{(c + 2\lambda^2)(c + \lambda^2)} & 2c\sqrt{2(2\lambda^2 - c)} & 4(2\lambda^2 - c) \sqrt{c + \lambda^2} \\
\hline
\end{array}
\]

with \(X \sim N(0, 1), V \sim N(0, 1), X \perp V\), and 
\(Z\) such that \(E(Z) = 0, \text{Var}(Z) = 1\) and \(\text{Cov}(V, Z) = 0\).

- Theorem 3 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

\[
T^{1/2} \left( \hat{\rho} - E(\rho_t) \right) = \lambda \frac{\mu_{yy} T^{-1/2} \sum_{t=1}^{T} y_{t-1}^2 u_t}{\mu_{y\eta} T^{-1/2} \sum_{t=1}^{T} y_{t-1} \eta_t} + \frac{\mu_{yyu} T^{-1/2} \sum_{t=1}^{T} y_{t-1} u_t}{\mu_{y\eta} T^{-1/2} \sum_{t=1}^{T} y_{t-1} \eta_t}
\]

\[
= \lambda \frac{\mu_{yy} T^{1/2} \sum_{t=1}^{T} y_{t-1}^2 u_t}{\mu_{y\eta} T^{1/2} \sum_{t=1}^{T} y_{t-1} \eta_t} + \frac{\mu_{yyu} T^{1/2} \sum_{t=1}^{T} y_{t-1} u_t}{\mu_{y\eta} T^{1/2} \sum_{t=1}^{T} y_{t-1} \eta_t}
\]

\[
\Rightarrow N \left( 0, 3\lambda^2 - 2c \right)
\]

Suppose now that \(c \geq 0\). We can write

\[
T^{\alpha/2} \frac{\phi_{yyu}^{T}}{\phi_{y}^{T}} (\hat{\rho} - E(\rho_t)) \Rightarrow \lambda \frac{\mu_{yy}}{\mu_{yyu}} \frac{U_{yyu}}{U_{yy}}
\]

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

\[
T^{\alpha/2} \frac{\phi_{yyu}^{T}}{\phi_{y}^{T}} = T^\alpha \quad \text{and} \quad \lambda^{-1} \frac{\mu_{yyu}}{\mu_{yy}} = \frac{1}{\lambda \sqrt{c + 2\lambda^2}}
\]

hence the result. \(\square\)
8.3 Proof of corollary 4

Under the null, the statistic is

$$\tau_{0,T} = \begin{cases} \frac{T^{1+\alpha}}{2} (\hat{\rho} - E_{H_0}(\rho_t)), \quad \text{if } \phi_0 + \lambda_0^2 < 0; \\ T^\alpha (\hat{\rho} - E_{H_0}(\rho_t)), \quad \text{if } \phi_0 + \lambda_0^2 \geq 0. \end{cases}$$

where

$$\hat{\rho} - E_{H_0}[\rho_t] = (\hat{\rho} - E_{H_1}[\rho_t]) + (E_{H_1}[\rho_t] - E_{H_0}[\rho_t]).$$

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

$$E_{H_1}[\rho_t] - E_{H_0}[\rho_t] = \frac{\phi_1 - \phi_0}{T^\alpha} + o(T^{-\alpha})$$

hence $T^{1+\alpha} (E_{H_1}[\rho_t] - E_{H_0}[\rho_t])$ diverges but $T^\alpha (E_{H_1}[\rho_t] - E_{H_0}[\rho_t])$ does not. Also, under the alternative, $T^{1+\alpha} (\hat{\rho} - E_{H_1}[\rho_t])$ diverges only if $\phi_1 + \lambda_1^2 \geq 0$ but $T^\alpha (\hat{\rho} - E_{H_1}[\rho_t])$ does not diverge. Finally, if both $T^{1+\alpha} (\hat{\rho} - E_{H_1}[\rho_t])$ and $T^{1+\alpha} (E_{H_1}[\rho_t] - E_{H_0}[\rho_t])$ diverge, their sum is $O_p(T^{1+\alpha})$ so they do not cancel each other. To summarize, $\tau_{0,T}$ diverges under $H_1$ only if $\phi_0 + \lambda_0^2 < 0$, irrespective of $(\phi_1, \lambda_1)$.

8.4 Proof of proposition 5

Consider the projection:

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

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

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

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

where proposition 1 implies that:

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

$$T^{-\alpha/2} \sum_{i=[T^\alpha r]+1}^{[T^\alpha(r+s)]} e^{\left\{ \frac{[T^\alpha r] s - [T^\alpha r] - i + \lambda T^\alpha/2}{T^\alpha} \sum_{j=i+1}^{[T^\alpha(r+s)]} u_j \right\}} \eta_{[T^\alpha r]+i} \Rightarrow \int_{rT^{1-\alpha}}^{(r+s)T^{1-\alpha}} e^{\phi(s-u) + \lambda(W_{r+s} - W_u)} dB_{r+u}$$
hence, since $T^\alpha/2 y^{-1}_{[T^\alpha r]} \to 0$ for $c \geq 0$, it follows that
\[
\frac{y_{[T^\alpha (r+s)]}}{y_{[T^\alpha r]}} \Rightarrow e^{\phi s + \lambda (W_{r+s} - W_r)}.
\]
which constitutes the first half of the proposition.

Now we prove expression for the conditional probability.

Let $t = \lceil T^\alpha r \rceil$, $k = \lceil T^\alpha s \rceil$ then as $T \to \infty$
\[
P \left( \frac{y_{t+k}}{y_t} \geq \gamma \right) \to P \left( e^{\phi s + \lambda (W_{r+s} - W_r)} \geq e^{\log \gamma} \right)
\]
\[
= P \left( \frac{W_{r+s} - W_r}{\sqrt{s}} \geq \frac{\log \gamma - \phi s}{\lambda \sqrt{s}} \right)
\]
where \( \frac{W_{r+s} - W_r}{\sqrt{s}} \sim N(0, 1) \) so
\[
P \left( \frac{y_{t+k}}{y_t} \geq \gamma \right) \to 1 - \Phi \left( \frac{\phi k T^{-\alpha} - \log \gamma}{\lambda \sqrt{kT^{-\alpha}}} \right) = \Phi \left( \frac{\phi s - \log \gamma}{\lambda \sqrt{s}} \right)
\]
We can also rewrite
\[
P \left( \frac{y_{t+k}}{y_t} \geq \gamma \right) - \Phi \left( \frac{\phi k T^{-\alpha} - \log \gamma}{\lambda \sqrt{kT^{-\alpha}}} \right)
\]
\[
\to 0
\]
Notice that $T^\alpha/2 y^{-1}_{[T^\alpha r]} \to 0$ also holds for $c \geq 0$ when conditioning on $y_{[T^\alpha r]}/y_{[T^\alpha (r-s)]} = \gamma$ since $y_{[T^\alpha r]} = \gamma y_{[T^\alpha (r-s)]}$; so the proof for
\[
P \left( \frac{y_{t+k}}{y_t} \geq \gamma \bigg| \frac{y_{t+k}}{y_{t-k}} \right)
\]
follows similarly.

### 8.5 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 $R^c$ then
\[
P_t = \frac{P_{t+1} + D_{t+1}}{1 + R^c}
\]
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 $\mathbb{E}[(1 + R_t)^{-1}] = (1 + R)^{-1}$.

Proof:
\[
\Delta P_t = (1 + (1 - \delta) R + \delta R_t) \Delta P_{t-1} - \zeta_t
\]
implies that

\[
\begin{align*}
P_{t+1} + D_{t+1} &= P_t + (1 + (1 - \delta) R + \delta R_{t+1}) \Delta P_t - \zeta_{t+1} + D_t + \zeta_{t+1} \\
\frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}} &= \frac{P_t + (1 + (1 - \delta) R + \delta R_{t+1}) \Delta P_t}{1 + R_{t+1}} + \frac{D_t}{1 + R_{t+1}} \\
&= \frac{P_t + (1 + (1 - \delta) R) \Delta P_t}{1 + R_{t+1}} + \frac{D_t}{1 + R_{t+1}} \\
&= \frac{P_t}{1 + R} + \Delta P_t
\end{align*}
\]

Now, if

\[
P_t = E_t \frac{P_{t+1} + D_{t+1}}{1 + R_{t+1}}
\]

then

\[
\begin{align*}
P_t &= \frac{P_t + (1 + (1 - \delta) R) \Delta P_t}{1 + R} + \frac{D_t}{1 + R} + \delta E_t \left[ \frac{1 + R_{t+1}}{1 + R_{t+1}} - \frac{1}{1 + R_{t+1}} \right] \Delta P_t \\
&= \frac{P_t}{1 + R} + \Delta P_t
\end{align*}
\]

i.e.

\[
P_{t-1} = \frac{P_t + D_t}{1 + R}
\]

qed.