Journal of Banking and Finance 161 (2024) 107113

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Reconciling interest rates evidence with theory: Rejecting unit roots when the HD(1) is a competing alternative

Alessandro Palandri

Dipartimento di Statistica, Informatica, Applicazioni (DiSIA) G. Parenti, Università degli Studi di Firenze, Italy

ARTICLE INFO

JEL classification:
G12
C22
C12

Keywords:
Interest rates
Bond yields
Hyperbolic reversion
Unit root test
Critical values

ABSTRACT

The paper introduces the HD(1), a Markovian process of order one with reversion rates that are faster the farther the process is from equilibrium. The aHD(1) approximation is introduced to allow for an estimation-calibration procedure based on available ARMA routines. Critical values of unit root tests with aHD(1) alternative are tabulated for the signed likelihood-ratio statistic. Revisiting the *non-stationarity of interest rates* stylized fact, the aHD(1) is found to be preferred to ARMA, SETAR and RCA and the resulting tests to reject the unit root hypothesis for all rates and yields considered.

1. Introduction

Since Rose (1988), which identifies interest rates as being non-stationary I(1) processes, Campbell and Shiller (1991) found evidence of cointegration among spreads at different maturities, and Hall et al. (1992) established that a common non-stationary factor underlies the evolution of various (cointegrating) yields to maturity. However, the non-stationarity of interest rates, which consolidated into an empirical stylized fact,[1] stands in contrast with the observation that returns on other assets (e.g. equities) are stationary. A similar discrepancy arises in economic and financial models where interest rates determine equilibria jointly with stationary variables (e.g. consumption growth rate). Clearly, knowledge of the stationarity of the interest rate, or lack thereof, is crucial for the subsequent modeling of its dynamics. This includes whether a stationary VAR (as in Sarno et al., 2007) or a VECM (as in Sarno et al., 2006) is employed and the validity of the ensuing findings.

While certain studies have imposed *tout court* stationarity on interest rates (e.g. Bekaert and Hodrick, 2001 and Shapiro and Watson, 1998), others have explored alternative tests that discriminate between stationarity and non-stationarity based on how quickly processes or their variances diverge such as the KPSS test by Kwiatkowski et al. (1992) and the random coefficient tests by Distaso (2008), Nagakura (2009) and Horváth and Trapani (2019). Nevertheless, the outcomes of these testing procedures do not alter the fact that, when taken to the data, existing models do not exhibit mean-reversion. Further efforts to reconcile theory and observations have investigated whether mean-reverting specifications of the alternative, beyond standard ARMA models, may provide more accurate descriptions of interest rate dynamics and, as a result, lead to the rejection of the I(1) hypothesis. In particular, motivated by considerations such as transaction costs, market frictions and liquidity constraints, Balke and Fomby (1997), Enders and Granger (1998), and Caner and Hansen (2001) investigate threshold error correction specifications of the alternative. Anderson (1997) adds the smooth error correction model of Granger and Teräsvirta (1993) to the analysis while Gray (1996) and Bansal and Zhou (2002) introduce regime switching models of the interest rate.

Overall, the prevailing literature aligns in asserting that, if interest rates are indeed stationary, identifying an accurate specification of their reversion to equilibrium is essential to reconcile the dissonance between the theoretical I(0) and the apparent experimental I(1) behavior. Consequently, the pursuit is directed towards reversion mechanisms to the long-run equilibrium other than the geometric reversion found in ARMA models. A readily available alternative to ARMA is provided by station-

E-mail address: alessandro.palandri@unifi.it.

[1] Corroborative evidence supporting interest rates non-stationarity is also found in the macroeconomics literature, for example, in MacDonald and Murphy (1989), Rapach and Weber (2004), among others.

https://doi.org/10.1016/j.jbankfin.2024.107113
Received 19 October 2022; Accepted 13 February 2024
Available online 23 February 2024

ary ARFIMA processes,[2] which allow for a slower hyperbolic reversion and long-memory,[3] Hosking (1981). Efforts to extend reversion or *error-correction* beyond linear[4] have resulted in the threshold autoregressive (TAR) and the self-exciting TAR (SETAR) models. While TAR processes are globally stationary, they alternate between unit root dynamics and reversion to the equilibrium, depending on the value attained by the process with respect to the threshold. However, from a financial modeling perspective, it must be acknowledged that TAR models depict a scenario where equilibrium forces (profitability, arbitrage, speculative behavior, etc.) remain dormant in the central regime but are fully activated once the system crosses the threshold.

This paper presents the *hyperbolic decay* HD(1), a Markovian process of order one characterized by the geometric decay of autocorrelations (no long-memory) and *hyperbolic* reversion to equilibrium. In other words, it exhibits a non-linear *error-correction* devoid of thresholds and the ensuing abrupt changes in dynamics. More specifically, the reversion speed of the HD(1) is negligible near the equilibrium (almost indistinguishable from a unit root process), accelerates as the distance increases, and reaches its maximum speed (similar to that of a serially uncorrelated process) at infinity. Notably, an HD(1) at infinity in time t will take on a finite value in $t+1$, determined by the sum of the deterministic component and the contemporaneous innovation. This unique feature distinguishes the HD(1) from the autoregressive components of ARMA, ARFIMA, SETAR and random coefficient autoregressive (RCA) specifications, where the process, once at infinity, remains at infinity. A first-order approximation of the HD(1), denoted as aHD(1), is introduced to allow for parameter estimation via any econometric and statistical software package. The aHD(1), nesting the AR(1), readily accommodates the inclusion of AR and MA terms, thereby encompassing both stationary and non-stationary ARMA processes. This flexibility extends to the generalization of Dickey and Fuller (1979, 1981) ADF tests, aiming to detect unit roots in an environment where the model under the alternative hypothesis is a stationary aHD(1) characterized by location-dependent speed of reversion. Critical values for both non-trending and trending processes are tabulated from 10^8 Monte Carlo simulations.

The empirical study tests for the presence of unit roots in selected Treasury Bill rates and long term Government Bond yields under ARMA, SETAR, RCA and aHD(1) specifications of the stationary alternative. The results show that among the nine rates and yields considered, ARMA provides the best data description (as measured by the Bayes Information Criterion) in only one instance (10-year Bond of Germany), SETAR is optimal in one instance (10-year Bond of Australia), and aHD(1) outperforms in seven instances. Overall, the I(1) hypothesis is rejected for six of the nine series considered: by aHD(1) in five cases and in SETAR in one case. Furthermore, the aHD(1) modeling of the spreads between the apparent I(1) yields and appropriate I(0) benchmark yields, allows to reject the unit root hypothesis for the three remaining bonds.

The paper is organized as follows. Section 2 presents the HD(1) process along with its properties, accompanied by a simulation study focusing on reversion to equilibrium. Section 3 introduces both no-trend and trend versions of the aHD(1), proposes a two-step estimation-calibration procedure of model parameters, derives the limiting distributions of the signed likelihood-ratio statistics for unit root testing and tabulates the corresponding non-standard critical values. The empirical study, investigating the stationarity of selected Treasury Bill rates and long term Government Bond yields, is presented in Section 4. Section 5 concludes.

2. The HD(1) process

The HD(1) is an iterated random process with parameters $(\alpha,\kappa) \in (0,+\infty)$, $\nu \in (0,1]$ and additive i.i.d. innovations $\varepsilon_{t+1} \sim (0,\sigma_\varepsilon^2)$:

$$y_{t+1} = g(y_t) + \varepsilon_{t+1} \tag{1}$$
$$g(y_t) = S_t \cdot \left[\kappa^{-1/\alpha} + \nu^{-1/\alpha} \cdot |y_t|^{-1/\alpha}\right]^{-\alpha}$$

where $S_t = \{-1,+1\}$ is the sign of y_t. The following properties are established:

1. The HD(1) is ergodic. Isolating the effects of the innovations ε_{t+1}, for the resulting recursive deterministic process $y_{t+1} = g(y_t)$ it holds that $|y_{t+1}|^{-1/\alpha} = \kappa^{-1/\alpha} + \nu^{-1/\alpha}|y_t|^{-1/\alpha}$. Given that the autoregressive coefficient $\nu^{-1/\alpha}$ is no less than unity, it follows that $|y_{t+1}|^{-1/\alpha}$ is explosive ($|y_\infty|^{-1/\alpha} \to \infty$), which implies that $|y_\infty| \to 0$. Hence, zero is the unique fixed point.[5] Since $g'(|y_t|) < \nu \leq 1$, the domain of attraction of g covers the entire Euclidean space, thereby meeting the absorbing requirements outlined in assumptions (A1)-(A5) of Theorem 3.3.2 of Chan and Tong (2001) for ergodicity and stationarity.

   Furthermore, as $\sup_{y\neq y'} \frac{\|g(y)-g(y')\|_p}{|y-y'|} < \nu$, it follows that for $\nu \in (0,1]$, the HD(1) meets the criteria outlined in Theorem 2.1 in Berkes et al. (2014), which extends Komlós-Major-Tusnády results to dependent sequences. The proof is provided in Appendix A.
2. Unlike autoregressive processes, for which $y_t = \pm\infty$ is inevitably followed by $y_{t+1} = \pm\infty$, the HD(1) re-enters the set of finite values at $y_{t+1} = \pm\kappa$ within a single period. The re-entry values $\pm\kappa$ bear a resemblance to the technical analysis concepts of *support* and *resistance*. In the context of the stochastic equation (1), the value $-\kappa$ can be viewed as the support, with a high probability of y_{t+1} rebounding, while κ serves as a resistance, with a high probability of y_{t+1} turning downward.[6] Notably, as $\kappa \to 0$, the HD(1) converges towards a white noise process while as $\kappa \to \infty$ it converges to an AR(1) with coefficient ν.
3. The HD(1) exhibits hyperbolic rates of reversion to equilibrium. Consider h-iterations of g:

$$g^h(y_t) = \left[\frac{\nu^{-h/\alpha}-1}{\nu^{-1/\alpha}-1} \cdot \left(\frac{|y_t|}{\kappa}\right)^{1/\alpha} + \nu^{-h/\alpha}\right]^{-\alpha} y_t \tag{2}$$

   When $\nu = 1$ and $|y_t| = \kappa$, equation (2) simplifies to $g^h(y_t) = (h+1)^{-\alpha}\, y_t$, which characterizes the α-hyperbolic reversion of the process from κ to zero. For instance, it behaves hyperbolically with $\alpha = 1$, quadratically hyperbolic with $\alpha = 2$, etc. The presence of

[2] Defining fractional differencing as $(1-L)^d = \sum_{i=0}^{\infty}\binom{d}{i}(-L)^i$, where $\binom{d}{i}$ is the binomial coefficient and L the lag-operator, stationarity requires a fractional integration parameter $|d| < 0.5$, see Granger and Joyeux (1980).

[3] Long-memory is defined from the summability of the absolute value of the autocorrelations ρ_j: $\lim_{n\to\infty}\sum_{j=-n}^{n}|\rho_j| \to \infty$. Estimation of the differencing parameter d, which captures the long-memory structure, requires long time series of observations due to non-negligible autocorrelations even at large lags. For an exact maximum-likelihood approach see Sowell (1992) while for some semi-parametric approaches see Geweke and Porter-Hudak (1983) and Robinson (1995), among others. For a survey of long-memory models see Baillie (1996).

[4] Examples where larger rates of reversion are associated to larger deviations from equilibrium include purchasing power parity (Sercu et al., 1995, Michael et al., 1997 and Taylor, 2001), output growth (Pesaran and Potter, 1997), arbitrage (Anderson, 1997), term structure of interest rates (Enders and Granger, 1998) and prices (Lo and Zivot, 2001), among others.

[5] Interestingly, when $\nu \in (1,+\infty)$, the fixed point $y_* = 0$ becomes repulsive and two fixed points emerge: $y_* = -(1-\nu^{-1/\alpha})^\alpha \cdot \kappa$ and $y_* = (1-\nu^{-1/\alpha})^\alpha \cdot \kappa$. These new fixed points act as attractors within the ranges $(-\infty,0)$ and $(0,+\infty)$, respectively. Therefore, when the variance of the innovations ε_{t+1} is non-negligible compared to $2(1-\nu^{-1/\alpha})^\alpha\kappa$, the process will transition randomly between the two regions of attraction, giving rise to y_{t+1} trajectories resembling those of a regime switching model. However, this specific feature of the HD(1) will not be discussed further in this context.

[6] In the case of the HD(1) and AR(1) processes shown in Fig. 1, when y_t reaches the support $-\kappa$, the probability $\mathbb{P}(y_{t+1} > -\kappa)$ is nearly indistinguishable from unity $(1-8.9\cdot 10^{-16})$ for the HD(1) while it stands at 0.5526 for the AR(1). The same probabilities apply to $\mathbb{P}(y_{t+1} < \kappa)$ when y_t reaches the resistance κ.

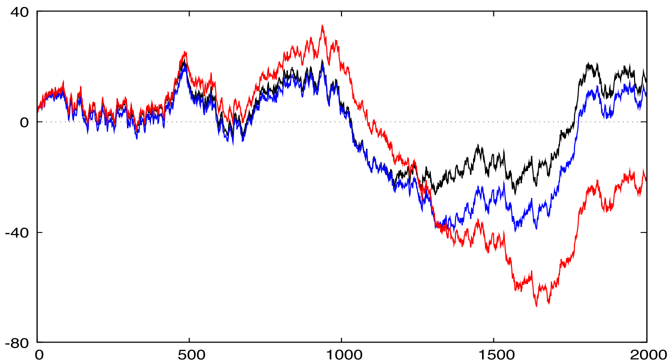


Fig. 1. Time-series trajectory of the HD(1) process with parameter values of $(\alpha = 0.25, \kappa = 50, \nu = 1)$ in black, AR(1) with parameter 0.997357 to match the lag-1 autocorrelation of the HD(1) in blue and Random Walk in red.

$|y_t|$ within the square brackets of (2) results in a faster rate of reversion to equilibrium for larger initial values $|y_t|$, in contrast to standard autoregressive models. Fig. 2 shows the hyperbolic reversion to equilibrium for an HD(1) process with parameters $(\alpha = 0.5, \kappa = 4, \nu = 1)$ alongside the geometric reversion of an AR(1) with identical initial decay.

4. The rate at which the HD(1) reverts to equilibrium is influenced by all model parameters. By examining the sign of the derivatives, as detailed in Appendix B, for $g^h(y_t)$ in (2) with respect to α, κ and ν:
   i. $\partial|g^h(y_t)|/\partial\alpha < 0$ indicates that larger values of α correspond to smaller values of $|y_{t+h}|$, which represent positions closer to equilibrium, as illustrated in Fig. 2b. Thus, a direct relationship exists between the magnitude of α and the rate of reversion (e.g. quadratic hyperbolic with $\alpha = 2$ reverts faster than hyperbolic with $\alpha = 1$).
   ii. $\partial|g^h(y_t)|/\partial\kappa > 0$ implies an indirect relationship between κ and the rate of reversion, as shown by the comparisons of Figs. 2c and 2b. It is noteworthy that $\lim_{\kappa\to\infty} g(y_t) = \nu \cdot y_t$ represents the deterministic component of an AR(1) with coefficient ν, whereas $\lim_{\kappa\to 0} g(y_t) = 0$ is the deterministic component of an i.i.d. process.
   iii. $\partial|g^h(y_t)|/\partial\nu > 0$ signifies that smaller ν values lead to faster rates of reversion to equilibrium, as conveyed by the comparison of Figs. 2c and 2d. Notably when $\nu < 1$, the rate of reversion exceeds that of an α-hyperbolic process.
5. The derivatives of $|g^h(y_t)|$ w.r.t. the model parameters, where $|g^h(y_t)|$ measures the distance of the process from equilibrium, allow to establish that for the HD(1) lag-1 autocorrelation ρ, the following relationships hold: $\partial\rho/\partial\alpha < 0$, $\partial\rho/\partial\kappa > 0$ and $\partial\rho/\partial\nu > 0$.
6. Scaling (change of variance) of the HD(1) process is achieved by the simultaneous scaling of the innovations variance (same as for ARMA processes) and the parameter κ. Letting $c > 0$ and $\tilde{y}_t = c \cdot y_t$ for every t, equation (1) may be rewritten as:

$$\tilde{y}_{t+1} = S_t \cdot \left[(c\kappa)^{-1/\alpha} + \nu^{-1/\alpha} \cdot |\tilde{y}_t|^{-1/\alpha}\right]^{-\alpha} + c \cdot \varepsilon_{t+1}$$

7. The impact on HD(1) dynamics of increasing (decreasing) the innovation variance σ_ε^2 is analogous to decreasing (increasing) the parameter κ while keeping the variance constant. Without loss of generality, assume $c > 1$ and consider the following three processes:

$$y_{t+1} = S_t \cdot \left[\kappa^{-1/\alpha} + \nu^{-1/\alpha} \cdot |y_t|^{-1/\alpha}\right]^{-\alpha} + \varepsilon_{t+1} \tag{3}$$

$$y_{t+1} = S_t \cdot \left[\kappa^{-1/\alpha} + \nu^{-1/\alpha} \cdot |y_t|^{-1/\alpha}\right]^{-\alpha} + c \cdot \varepsilon_{t+1} \tag{4}$$

$$y_{t+1} = S_t \cdot \left[(c\kappa)^{-1/\alpha} + \nu^{-1/\alpha} \cdot |y_t|^{-1/\alpha}\right]^{-\alpha} + c \cdot \varepsilon_{t+1} \tag{5}$$

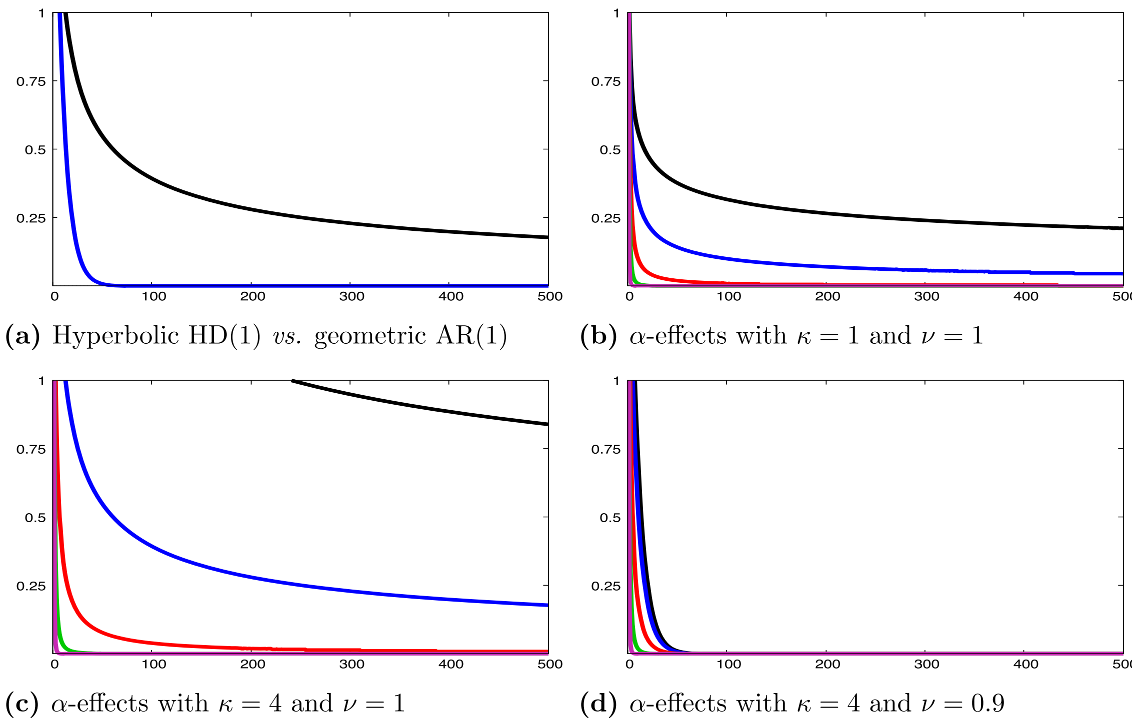
Given that, apart from the variance, (3) and (5) exhibit identical dynamic properties, including autocorrelations and relative speeds of reversion, the distinctions in the characteristics of (3) and (4) are analogous to those between (4) and (5). Since (4) can be viewed as a version of (5) with a smaller κ coefficient, it establishes an inverse relationship between κ and the variance of innovations σ_ε^2 concerning process dynamics.
8. The lag-1 autocorrelation ρ of the HD(1) exhibits an inverse relationship with the innovation variance σ_ε^2. The inverse relationship, discussed in 7., between κ and σ_ε^2, along with $\partial\rho/\partial\kappa > 0$, derived in 5., implies that changes in σ_ε^2 affect ρ in the direction of $\partial\rho/\partial\sigma_\varepsilon^2 < 0$. In Fig. 3, the relationship $\partial\rho/\partial\sigma_\varepsilon^2 < 0$ is illustrated by the comparison of the left and right panels. In the left panels, where σ_ε^2 is smaller, ρ is larger, leading to slower convergence to equilibrium for the AR(1) processes with parameter ρ.
9. The HD(1) may be interpreted as an AR(1) with location-dependent autoregressive coefficient:

$$y_{t+1} = \nu \cdot \left[1 + \left(\frac{\nu}{\kappa}\right)^{1/\alpha} \cdot |y_t|^{1/\alpha}\right]^{-\alpha} \cdot y_t + \varepsilon_{t+1}$$

The maximum ν of the autoregressive coefficient is attained when $y_t = 0$, while the minimum of 0 occurs as $y_t \to \pm\infty$. Consequently, the correction is negligible, resembling a random walk, when the process is near the equilibrium. As the magnitude of the deviation increases, the correction steadily intensifies, reaching full correction, akin to white noise behavior, as the disequilibrium diverges to $\pm\infty$. This property is illustrated in Fig. 1, which compares the time-series trajectories, driven by the same innovations, of an HD(1) with parameters $(\alpha = 0.25, \kappa = 50, \nu = 1)$, an AR(1) with matching lag-1 autocorrelation of 0.997357, and a Random Walk. It is evident that the HD(1) exhibits faster reversion than the AR(1) when the processes are far from equilibrium (range of realizations 1250-1750). Conversely, the reversion of the HD(1) is slower than that of the AR(1) when the processes are near the equilibrium (range of realizations 1750-2000).
10. The HD(1) has 1-step-ahead forecasts $\mathbb{E}_t[y_{t+1}] = S_t\left[\kappa^{-1/\alpha} + \nu^{-1/\alpha}|y_t|^{-1/\alpha}\right]^{-\alpha}$ with location-dependent speed of reversion to equilibrium. When $|y_t| > (\rho^{-1/\alpha} - \nu^{-1/\alpha})^{\alpha}\kappa$, where ρ is the lag-1 autocorrelation of the process, the HD(1) exhibits a faster reversion compared to an AR(1) with parameter ρ. In Fig. 3, as indicated by the derived threshold, it is observed that when $|y_t|$ is greater (less) than 12 when $\rho = 0.9992$ or 17.5 when $\rho = 0.9963$, HD(1) forecasts revert faster (slower) than those of the AR(1).
11. Calculating h-step ahead forecasts of the HD(1), when $h > 1$, entails computing h-dimensional integrals, which require knowledge of the distribution of the innovations. In the case of non-linear models like the HD(1), a common approach is to resample the residuals $\{\hat{\varepsilon}_\tau\}_{\tau=0}^{t}$ to generate S trajectories $\{y_{t+j}^s\}_{j=2}^{h}$ for $s = 1, \dots, S$, and then estimate $\mathbb{E}_t[y_{t+j}]$ by averaging y_{t+j}^s over s.
12. To a first-order approximation, the behavior of the h-period forecasts is described by $g^h(y_t)$. Iterating $y_{t+h} = g(y_{t+h-1}) + \varepsilon_{t+h}$ backward until time t, results in the expression $y_{t+h} = \varepsilon_{t+h} + g(\varepsilon_{t+h-1} + g(\varepsilon_{t+h-2} + g(\cdots + g(y_t))\cdots)$. A first-order Taylor expansion around $\varepsilon_{t+1} = 0, \dots, \varepsilon_{t+h-1} = 0$ gives:

$$y_{t+h} \approx g^h(y_t) + \varepsilon_{t+h} + \sum_{i=1}^{h-1}\left[\prod_{j=1}^{i} g'(g^{h-j}(y_t))\right] \cdot \varepsilon_{t+h-i}$$

and thus, $\mathbb{E}_t[y_{t+h}] \approx g^h(y_t)$. Building upon property 3., and observing that $g^h(y_t) = S_t \cdot g^h(|y_t|)$, it follows that as h increases, the forecasts converge toward equilibrium, albeit at a decreasing rate.
13. The forecasts revert to equilibrium faster as the variance of the innovations increases, all else equal. From 4ii., which establishes that $\partial|g^h(y_t)|/\partial\kappa > 0$, and 7., which outlines the inverse relationship between κ and σ_ε^2, it follows that larger (smaller) values of σ_ε^2 lead to forecasts converging faster (slower) to equilibrium. The comparison between the left and right panels of Fig. 3 exemplifies this relationship: larger values of σ_ε^2 lead to faster reverting forecasts, whereas smaller values of σ_ε^2 lead to slower reversion.



(a) Hyperbolic HD(1) *vs.* geometric AR(1)

(b) α-effects with $\kappa = 1$ and $\nu = 1$

(c) α-effects with $\kappa = 4$ and $\nu = 1$

(d) α-effects with $\kappa = 4$ and $\nu = 0.9$

Fig. 2. Panel 2a: HD(1) with parameters ($\alpha = 0.5, \kappa = 4, \nu = 1$) in black, and AR(1) with identical initial reversion in blue. Panels 2b-2d depict HD(1) reversions from $y_0 = 2$ for the triplet of parameters (α, κ, ν) with $\alpha = 0.25$ in black, $\alpha = 0.5$ in blue, $\alpha = 1$ in red, $\alpha = 2$ in green and $\alpha = 4$ in magenta. Panel 2b: parameters set to ($\alpha, \kappa = 1, \nu = 1$). Panel 2c: parameters set to ($\alpha, \kappa = 4, \nu = 1$). Panel 2d: parameters set to ($\alpha, \kappa = 4, \nu = 0.9$).

14. The HD(1) model in equation (1) can be extended to converge to the time-varying equilibrium $\mu_t \in \mathbb{R}$ by redefining $S_t = \{-1, +1\}$ as the sign of $(y_t - \mu_t)$ and specifying dynamics:

$$y_{t+1} = \mu_{t+1} + S_t \cdot \left[\kappa^{-1/\alpha} + \nu^{-1/\alpha} \cdot |y_t - \mu_t|^{-1/\alpha}\right]^{-\alpha} + \varepsilon_{t+1} \tag{6}$$

15. Oscillatory reversion, analogous to that of an AR(1) with negative autoregressive coefficient, can be introduced defining $S_t = \{-1, +1\}$ as the sign of $(\mu_t - y_t)$.
16. Consistent estimates of the HD(1) parameters, along with standard asymptotic inference, can be obtained via quasi maximum likelihood (see Gourieroux et al., 1984, among others) by specifying a Gaussian likelihood, regardless of the true distribution of the innovations ε_{t+1}.
17. Empirically, identification issues may arise for the HD(1) parameters (α, κ, ν) when the data generating process exhibits simple autocorrelation structures. This challenge is substantiated by the presence of countless combinations of (α, κ, ν) that result in y_{t+1} dynamics differing only at the second-order level, as demonstrated in Appendix C. To address this, the econometrician may choose to constrain either α or ν. In particular, selecting α is akin to defining a function that links the endogenous variable to its lag. This parallels the selection of a link function in the study of binary data, a decision made prior to estimation. Alternatively, the econometrician can set $\nu = 1$ to exclusively examine processes characterized by precisely α-hyperbolic rates of reversion to equilibrium (property 3). With $\nu = 1$, it is also possible to derive the aHD(1) approximation presented in Section 3, enabling parameter estimation using readily available ARMA routines.

3. The aHD(1) and unit root testing

The aHD(1) is a power-autoregressive specification that provides a first-order approximation to the HD(1) and its varying speed of reversion to equilibrium. Besides reconnecting the salient features of the HD(1) to the familiar autoregressive structure, the aHD(1) presents a substantial operational advantage as it allows to model and test time-series properties using ARMA routines that are readily available in every econometric and statistical software package.

Considering HD(1) processes in equation (6) with α-hyperbolic ($\nu = 1$) rates of reversion, we observe that as $\kappa^{-1/\alpha} \to 0$, the HD(1) simplifies to a random-walk. A first-order Taylor expansion around the unit-root case $\kappa^{-1/\alpha} = 0$ results in:

$$\left(y_{t+1} - \mu_{t+1}\right) = \left(y_t - \mu_t\right) + \pi \left(y_t - \mu_t\right) |y_t - \mu_t|^{1/\alpha} + \varepsilon_{t+1} \tag{7}$$

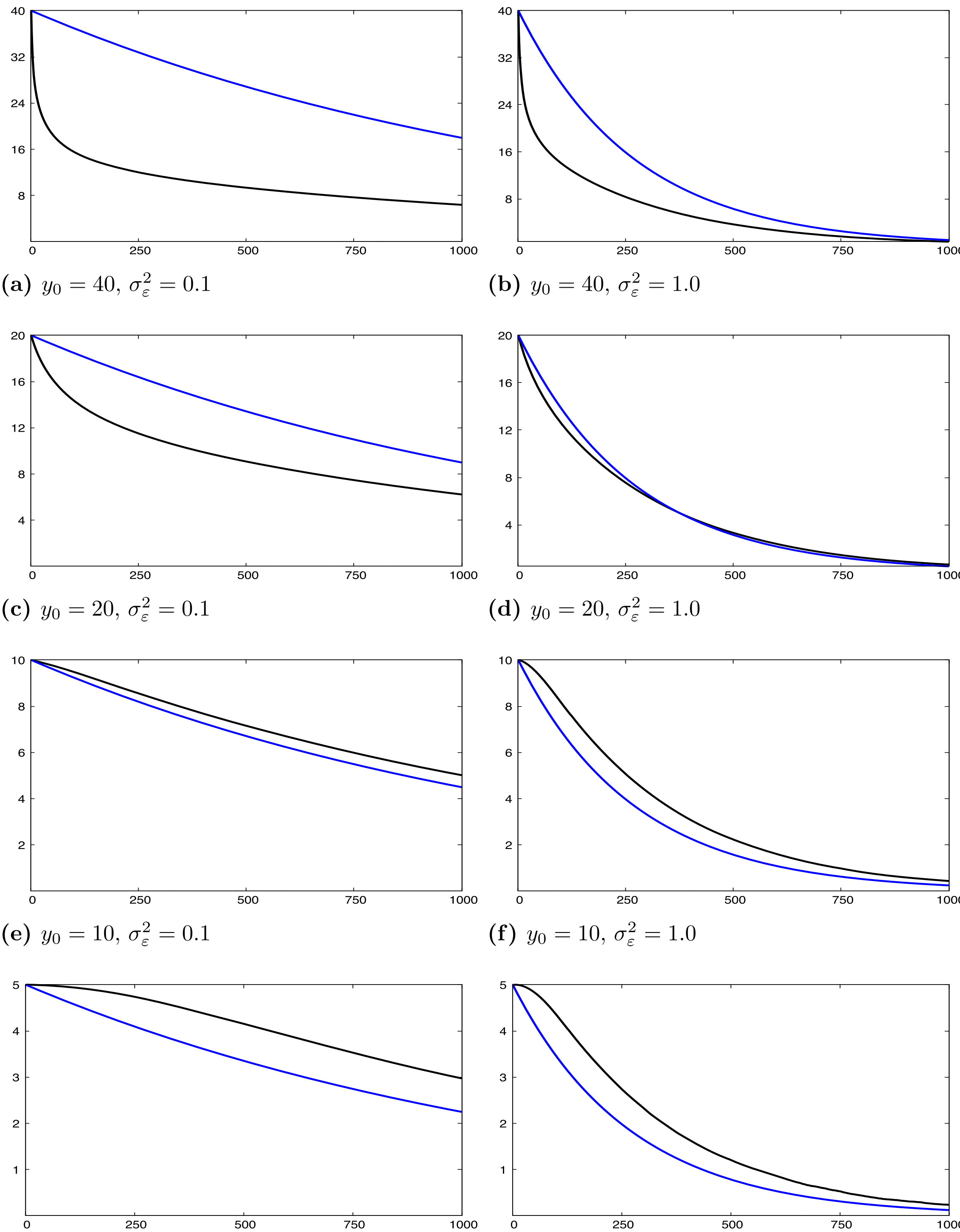
where $\pi = -\alpha\kappa^{-1/\alpha} < 0$. In contrast to the error-correction representation of an AR(1), which is linear in $(y_t - \mu_t)$, the aHD(1) exhibits a reversion which is faster (slower) when the process is distant from (close to) equilibrium. Specifically, the percentage offset of the disequilibrium $(\Delta y_{t+1} - \Delta\mu_{t+1})/(y_t - \mu_t)$ is determined by $\pi \left|y_t - \mu_t\right|^{1/\alpha}$. At equilibrium, no correction occurs, akin to a random walk. However, as the deviation's magnitude increases, the correction intensifies, eventually achieving complete offset resembling white noise behavior when $\left|y_t - \mu_t\right|^{1/\alpha} = -\pi^{-1}$. Notably, unlike the HD(1) model, the aHD(1) can overshoot the correction to equilibrium when $\left|y_t - \mu_t\right|^{1/\alpha} > -\pi^{-1}$. Conditions ensuring the stationarity of the aHD(1) in the event of overshoots and resulting oscillatory behavior are detailed in Appendix D.

3.1. Non-trending aHD(1)

The augmented aHD(1) that nests a non-trending unit root process is obtained by setting $\mu_{t+1} = \delta$ for every t in equation (7) and by adding stationary ARMA(p,q) components:

$$\Delta y_{t+1} = \pi \left(y_t - \delta\right) |y_t - \delta|^{1/\alpha} + \sum_{i=1}^{p} \phi_i \Delta y_{t+1-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j} + \varepsilon_{t+1} \tag{8}$$

which reduces to the usual augmented unit root specification for $\pi = 0$. Furthermore, for $\alpha \to +\infty$ equation (8) converges to a standard ARMA(p,q) specification of the process under the alternative hypothesis.



(a) $y_0 = 40$, $\sigma_\varepsilon^2 = 0.1$ (b) $y_0 = 40$, $\sigma_\varepsilon^2 = 1.0$

(c) $y_0 = 20$, $\sigma_\varepsilon^2 = 0.1$ (d) $y_0 = 20$, $\sigma_\varepsilon^2 = 1.0$

(e) $y_0 = 10$, $\sigma_\varepsilon^2 = 0.1$ (f) $y_0 = 10$, $\sigma_\varepsilon^2 = 1.0$

(g) $y_0 = 5$, $\sigma_\varepsilon^2 = 0.1$ (h) $y_0 = 5$, $\sigma_\varepsilon^2 = 1.0$

Fig. 3. In black: time-series trajectories of the h-period forecasts generated by the HD(1) model with parameter values ($\alpha = 0.25, \kappa = 50, \nu = 1$) and two different levels σ_ε^2: 0.1 (left panels) and 1.0 (right panels). In blue: forecast trajectories of the AR(1) models with matching lag-1 autocorrelation coefficients ρ: 0.9992 (left panels) and 0.9963 (right panels). The forecasts, arranged from top to bottom, are conditioned on the initial y_0 values of 40, 20, 10 and 5.

3.2. Trending aHD(1)

The augmented aHD(1) that nests a trending unit root process is obtained by introducing a deterministic trend $(\delta+\gamma t)$ in the equilibrium and by adding stationary ARMA(p,q) components:

$$\Delta y_{t+1} = \gamma + \pi \left(y_t - \delta - \gamma t\right) |y_t - \delta - \gamma t|^{1/\alpha} + \sum_{i=1}^{p} \phi_i \Delta y_{t+1-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j} + \varepsilon_{t+1} \quad (9)$$

which reduces to a unit root process with drift for $\pi = 0$ and converges to a standard trend-stationary ARMA(p,q) for $\alpha \to +\infty$.

3.3. Estimation

In the aHD(1)models defined in of equations (8) and (9), unit root testing strictly follows the standard procedure of detecting positive unit roots only, owing to the prevalent positive autocorrelations observed in the time-series of interest. Given that the parameters in (8) and (9) cannot be estimated jointly by readily available ARMA routines due to residual non-linearities, the estimation process will adopt a two-step procedure akin to the Engle and Granger (1987) approach to cointegration. The long-run equilibrium is estimated in the first step and the resulting residuals are used in the second step to estimate the error-correction specification. Some calibration of the α parameter completes the proposed approach.

3.3.1. First-step

The first-step estimates the parameters of the long-run equilibrium (whether without or with trend) and generates the residuals $\hat{u}_t$:

$$\text{No Trend}: \; y_t = \delta + u_t \quad \to \quad \hat{u}_t = y_t - \hat{\delta} \quad (10)$$

$$\text{Trend}: \; y_t = \delta + \gamma t + u_t \to \quad \hat{u}_t = y_t - \hat{\delta} - \hat{\gamma} t \quad (11)$$

3.3.2. Second-step (conditional on α)

Conditional on some α value, construct $x_t(\alpha) \equiv \hat{u}_t |\hat{u}_t|^{1/\alpha}$ and estimate the reversion parameter π and the coefficients ϕ_i and θ_j using standard ARMA routines with exogenous regressors $x_t(\alpha)$:

$$\begin{aligned} \text{No Trend}: \quad \Delta y_{t+1} &= \pi x_t(\alpha) + \textstyle\sum_{i=1}^{p} \phi_i \Delta y_{t+1-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t+!-j} + \varepsilon_{t+1} \\ \text{Trend}: \quad (\Delta y_{t+1} - \hat{\gamma}) &= \pi x_t(\alpha) + \textstyle\sum_{i=1}^{p} \phi_i \Delta y_{t+1-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j} + \varepsilon_{t+1} \end{aligned}$$

3.3.3. Poor man's α calibration

The two-step estimation procedure allows to estimate all aHD(1) model parameters, conditional on α. While the joint estimation of α in the second step is also possible, it is not feasible using readily available ARMA routines. Therefore, the proposed poor man's procedure focuses on a predetermined set of α values including hyperbolic ($\alpha = 1$), quadratic-hyperbolic ($\alpha = 2$), quartic-hyperbolic ($\alpha = 4$) and their reciprocals ($\alpha = 0.25$ and $\alpha = 0.5$). The two-step estimation procedure is then carried out for each $\alpha = \{0.25, 0.5, 1, 2, 4\}$ followed by the selection of the α that maximizes the value of the log-likelihood.

Obviously, a calibrated α will result in larger sums of squared residuals and lower log-likelihood values compared to those attainable through the estimation of the α parameter. Therefore, in unit root testing, rejecting the null hypothesis in favor of the alternative when α is calibrated, implies rejecting the alternative with estimated α. On the other hand, failure to reject the null when α is calibrated does not imply it won't be rejected if using the estimated α. In conclusion, adopting the poor man's approach to the α parameter leads to unit root tests with lower statistical *power* than those based on estimated α.

3.4. Asymptotics of I(1) testing

3.4.1. No-trend

Under the null, the data generating process is $\Delta y_{t+1} = \varepsilon_{t+1}$, with $\varepsilon_{t+1} \sim N(0, \sigma_\varepsilon^2)$. In the first step, the parameter δ of equation (10) is estimated by $\hat{\delta} = T^{-1}\sum_{t=1}^{T} y_t$ for which $T^{-1/2}\hat{\delta} \to \sigma_\varepsilon \int_0^1 W(s)ds$, where $W(s)$ is a standard Brownian motion defined on $s \in [0,1]$. For the resulting residuals $\hat{u}_t = y_t - \hat{\delta}_t$, it follows that $T^{-1/2}\hat{u}_t \to \sigma_\varepsilon \left[W(r) - \int_0^1 W(s)ds\right]$. Restricting the attention to the specification without ARMA components and conditioning on α, the least-squares estimator $\hat{\pi}$ of the reversion parameter π is:

$$\hat{\pi} = \pi + \frac{\sum_{t=1}^{T-1} \hat{u}_t |\hat{u}_t|^{1/\alpha} \Delta y_{t+1}}{\sum_{t=1}^{T-1} \hat{u}_t^{2+2/\alpha}}$$

Under the alternative $(\pi < 0)$ the process is stationary and $\hat{\pi}$ exhibits a standard rate of convergence $T^{1/2}$ to π and $T^{1/2}(\hat{\pi} - \pi)$ converges in distribution to a Gaussian random variable. On the other hand, under the null $(\pi = 0)$ the process is non-stationary with $T^{-1-1/2\alpha}\sum_{t=1}^{T-1} \hat{u}_t|\hat{u}_t|^{1/\alpha}\Delta y_{t+1} \to \sigma_\varepsilon^{2+1/\alpha} Z_1$ and $T^{-2-1/\alpha}\sum_{t=1}^{T-1}\hat{u}_t^{2+2/\alpha} \to \sigma_\varepsilon^{2+2/\alpha} Z_2$, where:

$$Z_1 = \int_0^1 \left(W(r) - \int_0^1 W(s)ds\right) \cdot \left|W(r) - \int_0^1 W(s)ds\right|^{1/\alpha} dW(r)$$

$$Z_2 = \int_0^1 \left(W(r) - \int_0^1 W(s)ds\right)^{2+2/\alpha} dr$$

Since $T^{-1-1/2\alpha}\hat{\pi} \to \sigma_\varepsilon^{-1/\alpha} Z_1 Z_2^{-1}$, it follows that $\hat{\pi}$ converges to zero at the rate $T^{1+1/2\alpha}$, which is faster than the rate T of the error-correction in ARMA and SETAR models. The t-statistic $t_{\pi=0}$ of the unit root test converges to the non-standard distribution $Z_1 Z_2^{-1/2}$. The signed likelihood-ratio test, henceforth LR, which compares the *unconstrained* Gaussian log-likelihood l_u of the aHD(1) of equation (8) to the *constrained* log-likelihood l_c of the I(1) processes resulting from $\pi = 0$ has the same asymptotic distribution:

$$LR_{\pi=0} \equiv \text{sign}(\hat{\pi}) \cdot \sqrt{2\left(l_u - l_c\right)} \to Z_1 Z_2^{-1/2}$$

3.4.2. Trend

The data generating process under the null, $\Delta y_{t+1} = \delta + \varepsilon_{t+1}$ with $\varepsilon_{t+1} \sim N(0, \sigma_\varepsilon^2)$, may be rewritten as $y_t = y_0 + \delta t + \eta_t$, where $\eta_t = \sum_{i=1}^{t} \varepsilon_t$. In the first step, the parameters δ and γ of $y_t = \delta + \gamma t + u_t$, in equation (10), are jointly estimated:

$$\begin{bmatrix} \hat{\delta} - y_0 \\ \hat{\gamma} - \delta \end{bmatrix} = \begin{bmatrix} T & \sum_{t=1}^{T} t \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} \eta_t \\ \sum_{t=1}^{T} t \cdot \eta_t \end{bmatrix}$$

It follows that $T^{-1/2}(\hat{\delta} - y_0) \to 4\sigma_\varepsilon Z_3 - 6\sigma_\varepsilon Z_4$ and $T^{1/2}(\hat{\gamma} - \delta) \to -6\sigma_\varepsilon Z_3 + 12\sigma_\varepsilon Z_4$, with $Z_3 = \int_0^1 W(s)ds$ and $Z_4 = \int_0^1 sW(s)ds$. The regression residuals may be rewritten as $\hat{u}_t = \eta_t - (\widetilde{\delta} - y_0) - (\hat{\gamma} - \delta)t$, from which it follows that $T^{-1/2}\hat{u}_t \to \sigma_\varepsilon \psi(r)$, where $\psi(r) = W(r) - 4Z_3 + 6Z_4 + r(6Z_3 - 12Z_4)$. Restricting the attention to the specification without ARMA components and conditioning on α, the least-squares estimator $\hat{\pi}$ is given by:

$$\hat{\pi} = \pi + \frac{\sum_{t=1}^{T-1} \hat{u}_t |\hat{u}_t|^{1/\alpha} \Delta\hat{u}_{t+1}}{\sum_{t=1}^{T-1} \hat{u}_t^{2+2/\alpha}}$$

where $\Delta\hat{u}_{t+1} = \Delta y_{t+1} - \hat{\gamma}$. Under the alternative $(\pi < 0)$ the process is trend-stationary and $\hat{\pi}$ exhibits a standard rate of convergence $T^{1/2}$ to π and $T^{1/2}(\hat{\pi} - \pi)$ converges in distribution to a Gaussian random variable. On the other hand, under the null $\pi = 0$ the process

Table 1
NO TREND critical values of the signed likelihood-ratio test in 3.4.1 with null $H_0 : \pi = 0$ in $\Delta y_{t+1} = \pi x_t(\alpha) + \sum_{i=1}^{p} \phi_i \Delta y_{t+1-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t+1-j} + \epsilon_{t+1}$, where $x_t(\alpha) \equiv \hat{u}_t|\hat{u}_t|^{1/\alpha}$ is calculated from the first-step regression residuals $\hat{u}_t = y_t - \hat{\delta}$. RAW column reports critical values estimated over 10^8 Monte Carlo simulations. FITTED column reports fitted values $\hat{c}$ from the regression function $c = a + bT^{-1}$ where c are the RAW critical values and T are the sample sizes. Asymptotic values ($T = \infty$) are estimated by $\hat{a}$.

$\alpha = 0.25$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.2316	−2.9348	−2.6905	−3.2187	−2.9243	−2.6821
100	−3.2870	−3.0027	−2.7667	−3.3154	−3.0260	−2.7854
250	−3.3710	−3.0859	−2.8469	−3.3735	−3.0870	−2.8474
500	−3.4108	−3.1213	−2.8789	−3.3928	−3.1074	−2.8681
∞	−3.4122	−3.1277	−2.8887	−3.4122	−3.1277	−2.8887
$\alpha = 0.50$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.3360	−3.0498	−2.8129	−3.3325	−3.0477	−2.8115
100	−3.3915	−3.1059	−2.8661	−3.3995	−3.1106	−2.8692
250	−3.4403	−3.1487	−2.9040	−3.4396	−3.1483	−2.9038
500	−3.4568	−3.1632	−2.9169	−3.4530	−3.1609	−2.9153
∞	−3.4664	−3.1735	−2.9269	−3.4664	−3.1735	−2.9269
$\alpha = 1.00$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.4214	−3.1300	−2.8854	−3.4210	−3.1299	−2.8854
100	−3.4504	−3.1562	−2.9084	−3.4514	−3.1564	−2.9085
250	−3.4698	−3.1726	−2.9225	−3.4697	−3.1724	−2.9224
500	−3.4762	−3.1776	−2.9269	−3.4758	−3.1777	−2.9270
∞	−3.4818	−3.1830	−2.9316	−3.4818	−3.1830	−2.9316
$\alpha = 2.00$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.4563	−3.1564	−2.9044	−3.4563	−3.1566	−2.9045
100	−3.4651	−3.1643	−2.9102	−3.4651	−3.1638	−2.9100
250	−3.4703	−3.1682	−2.9133	−3.4703	−3.1682	−2.9133
500	−3.4721	−3.1693	−2.9143	−3.4721	−3.1696	−2.9144
∞	−3.4738	−3.1710	−2.9155	−3.4738	−3.1710	−2.9155
$\alpha = 4.00$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.4612	−3.1563	−2.8998	−3.4612	−3.1566	−2.8999
100	−3.4600	−3.1553	−2.8978	−3.4599	−3.1546	−2.8976
250	−3.4589	−3.1539	−2.8962	−3.4591	−3.1535	−2.8962
500	−3.4589	−3.1534	−2.8957	−3.4588	−3.1531	−2.8958
∞	−3.4585	−3.1527	−2.8953	−3.4585	−3.1527	−2.8953

is non-stationary with $T^{-1-1/2\alpha} \sum_{t=1}^{T-1} \hat{u}_t|\hat{u}_t|^{1/\alpha} \Delta\hat{u}_{t+1} \to \sigma_\epsilon^{2+1/\alpha} Z_5$ and $T^{-2-1/\alpha} \sum_{t=1}^{T-1} \hat{u}_t^{2+2/\alpha} \to \sigma_\epsilon^{2+2/\alpha} Z_6$, where:

$$Z_5 = \int_0^1 \psi(r)|\psi(r)|^{1/\alpha} d\psi(r)$$

$$Z_6 = \int_0^1 \psi(r)^{2+2/\alpha} dr$$

Since $T^{-1-1/2\alpha}\hat{\pi} \to \sigma_\epsilon^{-1/\alpha} Z_5 Z_6^{-1}$, $\hat{\pi}$ converges to zero at the rate $T^{1+1/2\alpha}$, which is faster than the rate T of the error-correction in ARMA and SETAR models. The t-statistic $t_{\pi=0}$ of the unit root test converges to the non-standard distribution $Z_5 Z_6^{-1/2}$. The signed LR test, which compares the *unconstrained* Gaussian log-likelihood l_u of the aHD(1) of equation (9) to the *constrained* log-likelihood l_c of the I(1) processes resulting from $\pi = 0$ has the same asymptotic distribution:

$$LR_{\pi=0} \equiv \text{sign}(\hat{\pi}) \cdot \sqrt{2\left(l_u - l_c\right)} \to Z_5 Z_6^{-1/2}$$

3.4.3. *Critical values*

Critical values of the non-standard distributions are calculated via Monte Carlo simulations for the NO TREND and TREND cases and reported in Tables 1 and 2, respectively. The data generating process in the NO TREND case is $\Delta y_{t+1} = \epsilon_{t+1}$ while in the TREND case is $\Delta y_{t+1} = \gamma + \epsilon_{t+1}$. The aHD(1) with NO TREND and TREND are estimated following the two-step estimation procedure of Section 3.3 for $\alpha = \{0.25, 0.5, 1, 2, 4\}$. Critical values of LR (RAW) are estimated from 10^8 simulations for sample sizes of $T = \{50, 100, 250, 500\}$. Following standard procedure, the estimated critical values c are regressed on $a + bT^{-1}$ to produce fitted values (FITTED) and to estimate asymptotic ($T = \infty$) critical values $\hat{a}$. For a detailed discussion and analysis of the powers of T^{-1} see MacKinnon (2010).

4. Empirical study

Using data from FRED, of the Federal Reserve Bank of St. Louis, the empirical analysis studies the time-series properties of interest rates in various countries. Short term rates (as in Rose, 1988) are the United State's 3-month and 1-year Treasury Bills from 1954:01:08 to 2020:01:31 (3448 weekly obs.) and from 1959:07:17 to 2001:08:03 (2195 weekly obs.), respectively. Long term yields (as in Rapach and Weber, 2004) are the 10-year Government Bonds of the United States from 1960:01 to 2019:12 (720 monthly obs.), Australia from 1969:07 to 2019:12 (606 monthly obs.), Canada from 1960:01 to 2019:12 (720 monthly obs.), Germany from 1960:01 to 2019:12 (720 monthly obs.), France from 1960:01 to 2019:12 (720 monthly obs.), Great Britain from

Table 2
TREND critical values of the signed likelihood-ratio test in 3.4.2 with null $H_0: \pi = 0$ in $\Delta y_{t+1} - \hat{\gamma} = \pi x_t(\alpha) + \sum_{i=1}^{p} \phi_i \Delta y_{t+1-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j} + \varepsilon_{t+1}$, where $x_t(\alpha) \equiv \hat{u}_t |\hat{u}_t|^{1/\alpha}$ is calculated from the first-step regression residuals $\hat{u}_t = y_t - \hat{\delta} - \hat{\gamma} t$. RAW column reports critical values estimated over 10^8 Monte Carlo simulations. FITTED column reports fitted values $\hat{c}$ from the regression function $c = a + bT^{-1}$ where c are the RAW critical values and T are the sample sizes. Asymptotic values ($T = \infty$) are estimated by $\hat{a}$.

$\alpha = 0.25$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.6751	−3.3895	−3.1486	−3.6629	−3.3783	−3.1385
100	−3.8061	−3.5203	−3.2794	−3.8324	−3.5446	−3.3015
250	−3.9297	−3.6413	−3.3970	−3.9342	−3.6444	−3.3993
500	−3.9867	−3.6940	−3.4463	−3.9681	−3.6777	−3.4319
∞	−4.0020	−3.7110	−3.4645	−4.0020	−3.7110	−3.4645
$\alpha = 0.50$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.7780	−3.5155	−3.2950	−3.7771	−3.5163	−3.2968
100	−3.9604	−3.6882	−3.4573	−3.9632	−3.6868	−3.4536
250	−4.0774	−3.7912	−3.5493	−4.0748	−3.7892	−3.5476
500	−4.1113	−3.8207	−3.5754	−4.1121	−3.8233	−3.5790
∞	−4.1493	−3.8574	−3.6103	−4.1493	−3.8574	−3.6103
$\alpha = 1.00$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.8243	−3.5752	−3.3660	−3.8301	−3.5824	−3.3734
100	−4.0356	−3.7676	−3.5387	−4.0235	−3.7522	−3.5227
250	−4.1439	−3.8573	−3.6146	−4.1396	−3.8541	−3.6123
500	−4.1677	−3.8767	−3.6313	−4.1783	−3.8881	−3.6422
∞	−4.2169	−3.9220	−3.6721	−4.2169	−3.9220	−3.6721
$\alpha = 2.00$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.8237	−3.5778	−3.3734	−3.8319	−3.5876	−3.3830
100	−4.0478	−3.7813	−3.5522	−4.0308	−3.7602	−3.5313
250	−4.1553	−3.8672	−3.6229	−4.1501	−3.8638	−3.6203
500	−4.1758	−3.8835	−3.6360	−4.1899	−3.8983	−3.6499
∞	−4.2297	−3.9328	−3.6796	−4.2297	−3.9328	−3.6797
$\alpha = 4.00$	RAW			FITTED		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−3.8083	−3.5638	−3.3620	−3.8175	−3.5656	−3.3726
100	−4.0393	−3.7729	−3.5438	−4.0199	−3.7655	−3.5206
250	−4.1470	−3.8575	−3.6119	−4.1414	−3.8854	−3.6094
500	−4.1661	−3.8725	−3.6238	−4.1438	−3.9254	−3.6390
∞	−4.2223	−3.9654	−3.6686	−4.2223	−3.9654	−3.6686

1960:01 to 2019:12 (720 monthly obs.) and Japan from 1989:01 to 2019:12 (372 monthly obs.). The time span of all rates and yields is that provided by FRED with the sole exception of the 1-year Treasury Bill whose series has been trimmed to 2001:08:03, the last date prior to the US Treasury discontinuing issue until recently. Time-series of T-Bill rates and Bond yields are plotted in Fig. 4.

While, in principle, ARFIMA models should be added to the set of competing specifications and possibly used for testing,[7] interest rates' stylized facts include the rejection of fractional integration (long-memory) in favor of unit roots. Although here no formal testing is carried out for ARFIMA parameterizations, no autocorrelogram displays the hyperbolic telltale sign of stationary ARFIMA processes and long-memory. Instead, correlograms in Fig. 5 display the characteristic geometric decay of persistent processes. Geweke and Porter-Hudak (1983) estimates of the fractional integration parameter range from 0.8689 to 1.1080, well above the stationary threshold value of 0.5. Estimates performed on first-differences, in Table 3, produce fractional integration parameters that are not statistically different from zero for each of the 10-year Bonds (at 5%) and US Treasury Bills (at 1% but not at 2.5%), suggesting the presence of a unit root in the levels.

The benchmark SETAR specification closely follows that of Bec et al. (2004) with two regimes instead of three. In fact, the time series trajectories of the interest rates in Fig. 4 and preliminary estimates do not support a third regime. Hence, the SETAR dynamics are given by:

$$\Delta y_{t+1} = \begin{cases} \sum_{i=1}^{r} \phi_{1,i} \Delta y_{t+1-i} + \sum_{j=1}^{s} \theta_{1,i} \Delta\varepsilon_{t+1-i} + \varepsilon_{t+1} & \text{if } y_t \in A \\ \delta + \pi y_t + \sum_{i=1}^{q} \phi_{2,i} \Delta y_{t+1-i} + \sum_{j=1}^{v} \theta_{2,i} \Delta\varepsilon_{t+1-i} + \varepsilon_{t+1} & \text{otherwise} \end{cases} \tag{12}$$

where the set A is defined by the threshold parameter c either as $A = \{y | y \le c\}$ or $A = \{y | y \ge c\}$, depending on the data features. Since threshold parameters are not identified under the null, following the work of Tong (1990) and Hansen (1996), unit root testing is performed using the infimum, over possible values of the thresholding parameter c, of the signed likelihood-ratio test:

$$LR = \inf_{i \in \Gamma} \operatorname{sign}(\hat{\pi}) \cdot \sqrt{2\left(l_u^{(i)} - l_c\right)} \tag{13}$$

[7] For unit root testing with fractional integration alternative see Robinson (1994), Phillips (1999), Tanaka (1999), Dolado et al. (2002, 2008), Lobato and Velasco (2007), Cho et al. (2015) and Chang and Perron (2017), among others.

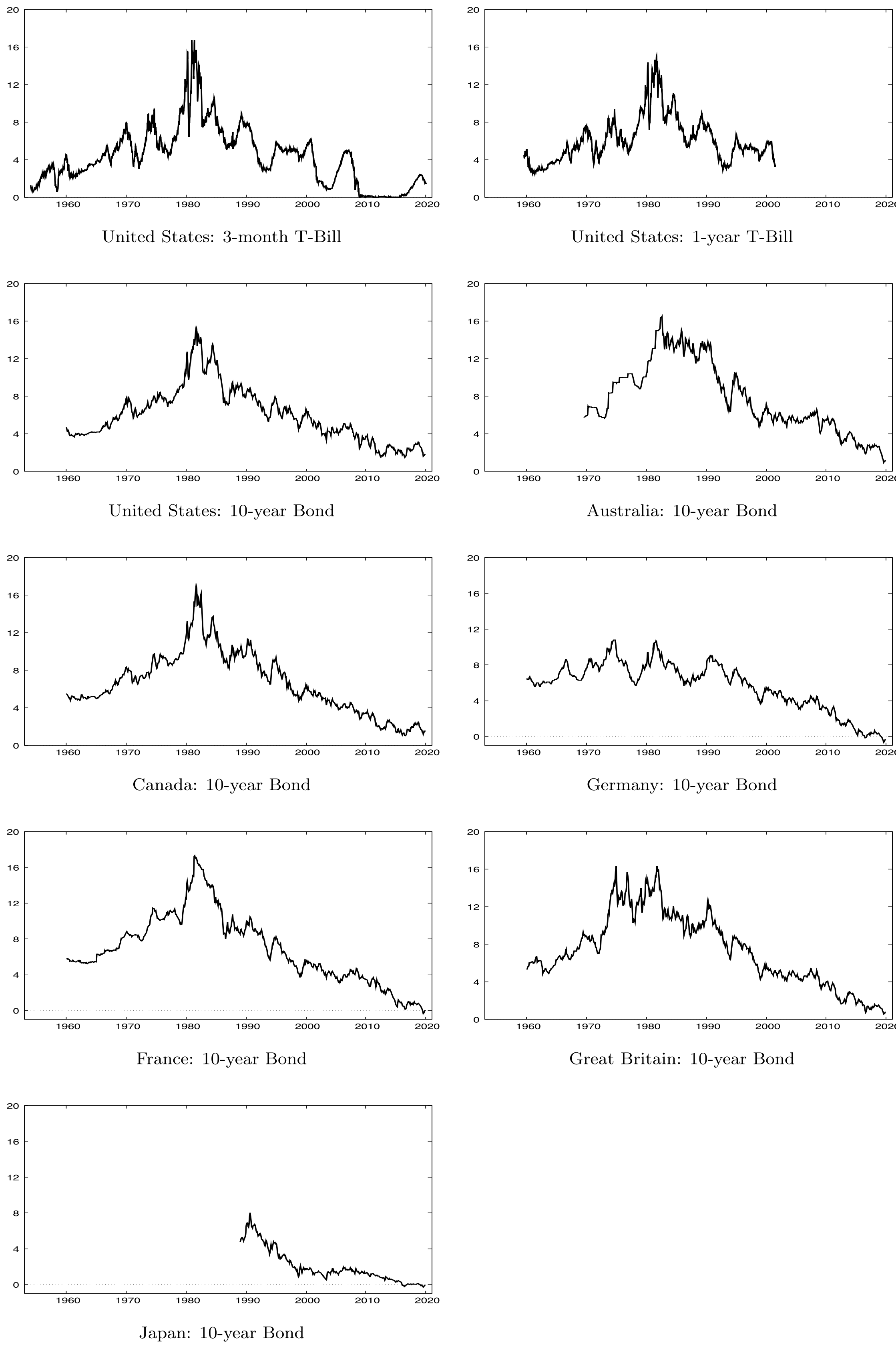


Fig. 4. Time-series of T-Bill rates and Bond yields.

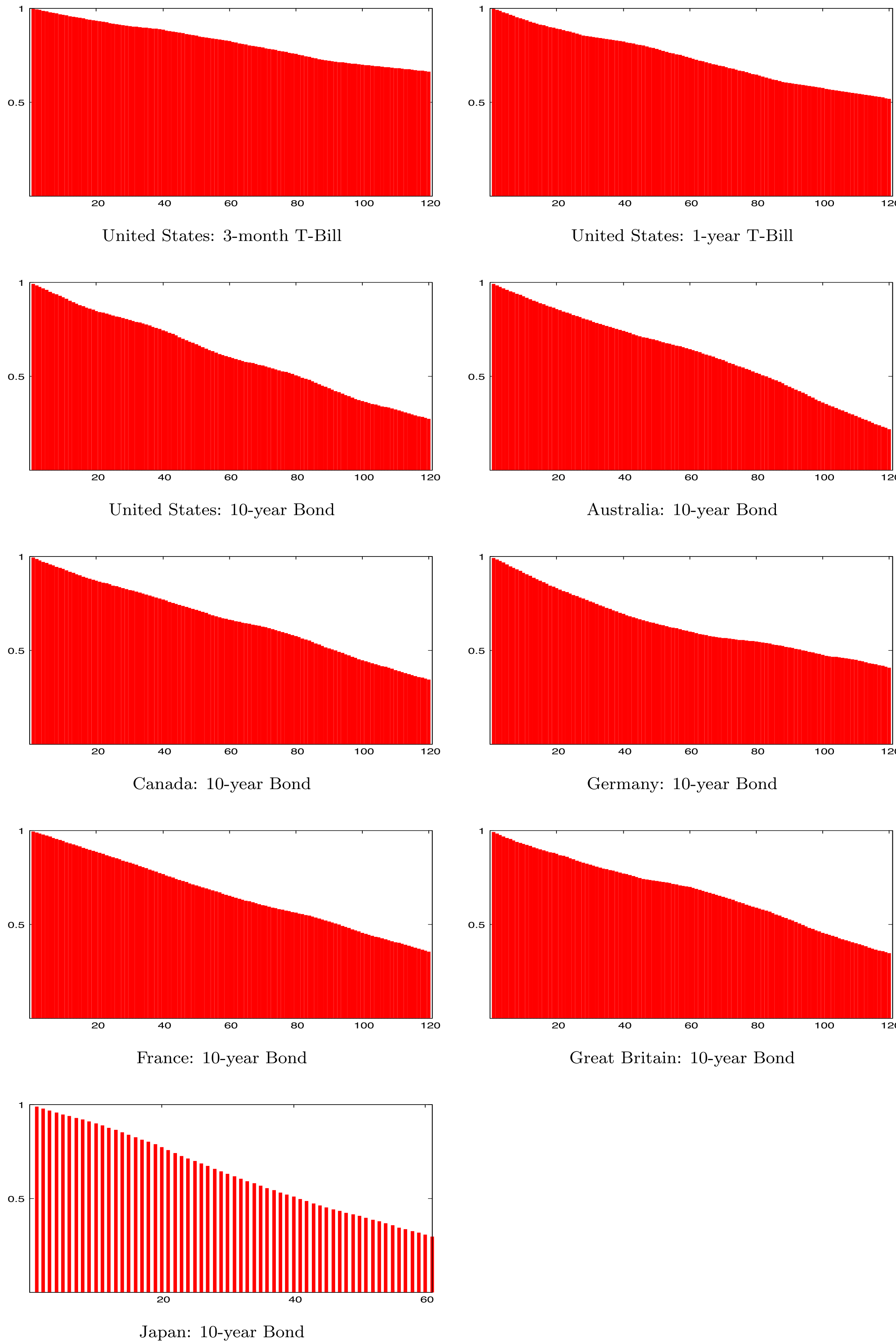


Fig. 5. Empirical Autocorrelations. All estimates are statistically significant at 1% (confidence interval not plotted).

Table 3
Geweke and Porter-Hudak (1983) estimates of the fractional differencing parameter and associated test of significance for the level of rates and yields (d) and the first-difference of rates and yields ($d-1$). Estimation lag-orders are 120 for Treasury Bills and 50 for Government Bonds. ^{a}, ^{b}, and ^{c} indicate significance at 1%, 2.5% and 5%, respectively.

	d	s.e.	z	$d-1$	s.e.	z
US: 3-month T-Bill	0.8774	0.0562	15.6115^a	−0.1279	0.0558	−2.2898^b
US: 1-year T-Bill	0.8689	0.0555	15.6457^a	−0.1229	0.0534	−2.3013^b
US: 10-year Bond	1.1016	0.1074	10.2535^a	0.1361	0.1012	1.3459
Australia: 10-year Bond	1.1055	0.1141	9.6869^a	−0.1072	0.1037	−1.0338
Canada: 10-year Bond	1.0197	0.0912	11.1756^a	0.0651	0.0844	0.7715
Germany: 10-year Bond	0.9872	0.0797	12.3796^a	0.0244	0.0938	0.2601
France: 10-year Bond	1.0712	0.0987	10.8569^a	0.1433	0.0965	1.4850
GB: 10-year Bond	0.9232	0.1129	8.1781^a	−0.0826	0.1261	−0.6548
Japan: 10-year Bond	1.1080	0.0916	12.0982^a	−0.0663	0.0896	−0.7405

Table 4
Critical values of the inf signed likelihood-ratio test in equation (13) with null $H_0 : \pi = 0$ in the two-regime SETAR parameterization of equations (12). Raw column reports critical values estimated over 10^8 Monte Carlo simulations. Fitted column reports fitted values $\hat{c}$ from the regression function $c = a + bT^{-1}$ where c are the Raw critical values and T are the sample sizes. Asymptotic values ($T = \infty$) are estimated by $\hat{a}$.

	Raw			Fitted		
T	1.0%	2.5%	5.0%	1.0%	2.5%	5.0%
50	−4.13	−3.86	−3.64	−4.12	−3.85	−3.63
100	−4.19	−3.92	−3.70	−4.21	−3.94	−3.72
250	−4.27	−4.00	−3.78	−4.26	−3.99	−3.77
500	−4.29	−4.02	−3.80	−4.28	−4.01	−3.79
∞	−4.30	−4.03	−3.81	−4.30	−4.03	−3.81

where l_c is the log-likelihood of the I(1) process under the null and $l_u^{(i)}$ is the log-likelihood value, of the estimated SETAR, obtained from the i-th point of the threshold parameter grid set Γ. Corresponding critical values, reported in Table 4, are estimated via 10^6 Monte Carlo simulations for sample sizes $T = \{50, 100, 250, 500, \infty\}$.

An additional class worth considering is that of Random Coefficient Autoregressive (RCA) models, which encompasses the Stochastic Unit Root (STUR) processes studied, among others, by Granger and Swanson (1997), Leybourne et al. (1996), and McCabe and Tremayne (1995):

$$\Delta y_{t+1} = \delta + \pi_{t+1} y_t + \sigma_\epsilon \epsilon_{t+1}$$

$$\ln(1 + \pi_{t+1}) = \alpha + \phi_1 \ln(1 + \pi_t) + \sigma_\eta \eta_{t+1}$$

with $\epsilon_{t+1} \sim (0,1)$ independent of $\eta_{t+1} \sim N(0,1)$. In contrast to the proposed aHD(1) model, the error-correction of the RCA is linear in y_t and its random coefficient π_{t+1} induces time-varying speeds of adjustment that are independent of the distance from equilibrium. While several tests have been developed to assess the random walk hypothesis against the STUR alternative (see Nagakura, 2009, Distaso, 2008, Leybourne et al., 1996, and McCabe and Tremayne, 1995), and others test for stationarity based on the rate of divergence of the process trajectory (see Busetti and Harvey, 2010, and Trapani, 2021), there is a void when it comes to testing the STUR null hypothesis against the stationary RCA alternative. To circumvent this testing challenge, a Bayesian inferential approach (see Lubrano, 1995 and Hoek et al., 1995 among others) based on 10^6 Markov chain Monte Carlo (MCMC) iterations[8] is adopted by constructing the posterior from uninformative uniform priors and assuming that $\epsilon_{t+1} \sim N(0,1)$. The main advantage of working in the Bayesian framework is that it readily delivers *credible intervals* from the sampling of the unknown parameters. This, in turn, enables reporting of the p-values for the RCA non-stationarity in mean (Ms) and variance (Vs) in Tables 5–10. Specifically, Ms and Vs display the average values of $\delta + 0.5\sigma_\eta^2/(1-\phi_1)$ and $\delta + \sigma_\eta^2/(1-\phi_1)$, respectively, along with their significance: $\mathbb{P}(\delta + 0.5\sigma_\eta^2/(1-\phi_1) \geq 0)$ and $\mathbb{P}(\delta + \sigma_\eta^2/(1-\phi_1) \geq 0)$ falling below the established significance levels. Similarly, π is the average value of $\delta + 0.5\sigma_\eta^2/(1+\phi_1)$, determining the unconditional expectation of π_{t+1}, with its significance calculated from $\mathbb{P}(\delta + 0.5\sigma_\eta^2/(1+\phi_1) \geq 0)$ also falling below the significance level. Notice that, as shown in Appendix E, $\mathbb{E}(\pi_{t+1}) < 1$ implies neither mean- nor variance-stationarity.

Since initial data analysis ruled out stochastic trends in the time series considered (compatibly with the trajectories plotted in Fig. 4, only No Trend parameterizations of the alternatives and null are presented and discussed. The Bayes Information Criterion (Bic) is adopted to perform specific-to-general model selection and attain the best parameterization[9] for each of the unconstrained ARMA, SETAR, RCA[10] and aHD(1) specifications. Estimation and unit-root testing output for every yield and specification considered is presented in Tables 5–7.

To begin, consider ARMA, SETAR, RCA and aHD(1) testing results as if each one of them is the sole parameterization of the alternative. Within the framework of ARMA and RCA modeling, non-stationarity cannot be rejected for any of the nine yields. Furthermore, Bic favors ARMA over RCA in every instance. SETAR modeling of the alternative would allow to reject a unit-root in the US 3-month Treasury Bill and the 10-year bonds of Australia and Great Britain. However, since Bic selects ARMA over SETAR for the yield of Great Britain, ARMA and SETAR modeling reject the unit-root hypothesis only for the US 3-month Treasury Bill and Australian 10-year Bond. Instead, focusing on the aHD(1) specification of the alternative, allows to reject the presence of a unit-root for the US 3-month and 1-year Treasury Bills and the 10-year bonds of the US, Canada and Great Britain. Finally, considering all four parameterizations and following the consolidated good practice of performing unit root tests on the unconstrained specification that best fits the data, Tables 5–7 show that ARMA is the best specification for the German yield (Table 6), SETAR for the Australian yield (Table 6) while the aHD(1) provides the best data description for the remaining seven series. Correspondingly, the unit-root hypothesis is rejected

[8] While several techniques are available for updating π_{t+1} in $(t+1)$ based on the prediction made in t, in this context, estimation was carried out using the Kalman filter. Despite the non-linear nature of RCA models, closed-form solutions can be derived for all predictions and updates in the filter, except for the updated (*a posteriori*) estimate covariance. Nevertheless, due to the Gaussian nature of the innovations η_{t+1}, this update can be effectively approximated by interpolating the boundary values of zero and the predicted (*a priori*) estimate covariance with the parameter ϕ_1.

[9] The best parameterization is the one that strikes the best balance, as measured by the Bic, between goodness of fit and parameters' parsimony. Alternative criteria defining what is meant by best specification are the elimination of insignificant coefficients (general-to-specific) and Box-Jenkins (minimal parameterization yielding serially uncorrelated residuals).

[10] In the case of RCA, estimated by MCMC, Lik represents the maximum value of the log-likelihood attained across the 10^6 iterations. Importantly, the Lik values of RCA consistently align with the corresponding, albeit not reported, Lik values AR(1) models with constant π.

Table 5

Bond yields: estimation and unit-root testing results of No Trend specifications. Every sub-panel reports the estimation output under the unconstrained alternatives: ARMA, two-regimes SETAR, RCA and aHD(1). Autoregressive coefficients are ϕ_i, moving average coefficients are θ_j, and the variance of the innovations to the RCA latent process is σ_η^2. The lowest Bic of the four competing parameterizations is reported in bold, Lik is the log-likelihood, and cLik the constrained log-likelihood ($\delta = 0$, $\pi = 0$). LR is the value of the signed likelihood-ratio statistic for unit root testing in ARMA, SETAR and aHD(1), while Ms and Vs are the values of the mean- and variance-stationarity statistics for unit root testing in RCA. Significance levels are marked by [a], [b], and [c], indicating 1%, 2.5% and 5% significance, respectively.

	δ	π	α	ϕ_1	ϕ_2/σ_η^2	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0098	−0.0022		0.2638		−2135.577	1080.007	1077.368	−2.297
	(0.0052)	(0.0010)		(0.0164)					
SETAR(< 7.19)			0.1653						
				(0.0292)					
SETAR(> 7.19)	0.1240	−0.0140		0.3133		−2150.408	1095.568	1085.467	−4.495[a]
	(0.0318)	(0.0032)		(0.0198)					
RCA	0.0106	−0.0106[a]	−0.0287	0.2744	0.0532	−2127.912	1080.246	0.008	0.044
	(0.0061)	(0.0062)	(0.0214)	(0.0181)	(0.0427)				
aHD(1)	4.3271	$-1.6\cdot10^{-6}$	0.25	0.2835		**−2187.259**	1105.848	1077.368	−7.547[a]
	(0.0527)	$(2.4\cdot10^{-7})$		(0.0167)					

US: 3-month T-Bill from 1954:08 to 2020:01 (3448 weekly obs.)

	δ	π	α	ϕ_1	ϕ_2/σ_η^2	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0201	−0.0033		0.3347		−1694.509	858.795	856.153	−2.299
	(0.0095)	(0.0014)		(0.0201)					
SETAR(< 7.93)				0.2947					
				(0.0321)					
SETAR(> 7.93)	0.1454	−0.0150		0.3638		−1688.972	863.721	857.318	−3.579
	(0.0443)	(0.0043)		(0.0258)					
RCA	0.0309	−0.0099[a]	−0.0150	0.3682	0.0231	−1682.092	856.433	0.003	0.020
	(0.0143)	(0.0058)	(0.0145)	(0.0425)	(0.0287)				
aHD(1)	6.1343	$-4.0\cdot10^{-6}$	0.25	0.3197	0.0634	**−1709.492**	870.134	855.843	−5.346[a]
	(0.0515)	$(7.5\cdot10^{-7})$		(0.0212)	(0.0213)				

US: 1-year T-Bill from 1959:07 to 2001:08 (2195 weekly obs.)

	δ	π	α	ϕ_1	ϕ_2/σ_η^2	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0278	−0.0053		0.3889	−0.0997	109.153	−41.418	−42.220	−1.266
	(0.0312)	(0.0033)		(0.0244)	(0.0254)				
SETAR(< 8.70)				0.3125	−0.1613				
				(0.0511)	(0.0510)				
SETAR(> 8.70)	0.2956	−0.0279		0.4212	−0.2415	131.101	−42.523	−44.858	−2.161
	(0.1490)	(0.0133)		(0.0522)	(0.0528)				
RCA	0.0392	−0.0312[a]	−0.0737	0.3202	0.1369	146.456	−60.072	0.027	0.127
	(0.0334)	(0.0199)	(0.0562)	(0.0038)	(0.1123)				
aHD(1)	6.0491	$-8.1\cdot10^{-6}$	0.25		0.4437	**96.420**	−38.341	−45.818	−3.867[a]
	(0.1072)	$(9.0\cdot10^{-7})$			(0.0220)				

US: 10-year Bond from 1960:01 to 2019:12 (720 monthly obs.)

for the US 3-month and 1-year Treasury Bills and the 10-year bonds of the US, Canada and Great Britain (all with aHD(1) being the best specification under the alternative) and the 10-year Australian bond (with SETAR being the best specification under the alternative). From the aHD(1) modeling perspective, the selective rejection of the null hypothesis across different cases dismisses the notion that the aHD(1) is merely capturing the common *tent shape* observed in Fig. 4.

For the cases in which the unit root could not be rejected, it is worth considering the possibility that yields are stationary but the aHD(1) parameterization of the alternative, despite being Bic preferred to ARMA and SETAR, might still not adequately describe the data. While a possible course of action would involve higher order HD parameterizations, generalizing the HD(1) is not straightforward. Unlike the AR(1), which has a unique generalization to AR(p), there are multiple candidate parameterizations HD(p) that reduce to the proposed HD(1) for p = 1. Instead, consider decomposing the yield b_y into a benchmark yield b_x plus the spread $s_{y,x} = (b_y - b_x)$ so that $b_y = b_x + s_{y,x}$. Consequently, if the benchmark b_x, is stationary, the yield b_y must also be stationary if the spread $s_{y,x}$ is stationary. The rationale behind such decoupling is that the resulting b_y would be the sum of two (possibly correlated) stationary HD(1) processes. In essence, the sum of HD(1) processes extends the HD(1) much like the sum of two uncorrelated AR(1) processes, which yields an ARMA(2,1), captures a moving-average component without requiring an explicit specification of what an MA component entails.

Spreads of the 10-year bonds, calculated with respect to that of the United States, are plotted in Fig. 6 and depict the disappearance of the tent shape pattern in the yields. The results in Tables 8 and 9 show that ARMA modeling is Bic preferred to RCA in every instance. Moreover, while RCA rejects non-stationarity of the mean, but not the variance, only for Australia-US (at 1%), ARMA rejects unit-roots only for Canada-US (at 5%) and GB-US (at 2.5%). Since the Canadian, GB and US yields test I(1) when the alternative is ARMA (Tables 5–7), the fact that the resulting spreads are found to be I(0) may only be interpreted as evidence of cointegration. Similarly, SETAR modeling of the alternative does not allow to reject a unit-root in the spreads except for Canada-US (at 1%) and France-US (at 2.5%). Given that the individual yields test I(1)) when the alternative is SETAR (Tables 5–7), the interpretation of I(0) spreads is, once more, cointegration. On the other hand, with aHD(1) alternative, presence of a unit-root is *rejected* for all spreads with the exception of Australia-US and France-US. Considering that the US yield tests I(0) when the alternative is aHD(1) (Table 5), it follows that all the yields, whose spread with the US yield is I(0), must also be I(0). Furthermore, when the benchmark yield b_x for the 10-year Bonds of Australia and France is replaced with the more financially and eco-

Table 6

Bond yields: estimation and unit-root testing results of NO TREND specifications. Every sub-panel reports the estimation output under the unconstrained alternatives: ARMA, two-regimes SETAR, RCA and aHD(1). Autoregressive coefficients are ϕ_i, moving average coefficients are θ_j, and the variance of the innovations to the RCA latent process is σ^2_η. The lowest BIC of the four competing parameterizations is reported in bold, LIK is the log-likelihood, and CLIK the constrained log-likelihood ($\delta = 0$, $\pi = 0$). LR is the value of the signed likelihood-ratio statistic for unit root testing in ARMA, SETAR and aHD(1), while Ms and Vs are the values of the mean- and variance-stationarity statistics for unit root testing in RCA. Significance levels are marked by [a], [b], and [c], indicating 1%, 2.5% and 5% significance, respectively.

	δ	π	α	ϕ_1	ϕ_2/σ^2_η	BIC	LIK	CLIK/Ms	LR/Vs
ARMA	0.0101	−0.0021		0.1685		297.269	−139.024	−139.346	−0.802
	(0.0293)	(0.0034)		(0.0403)					
SETAR(< 12.30)				0.1936					
				(0.0528)					
SETAR(> 12.30)	2.3343	−0.1704		0.2009		**284.521**	−126.244	−139.075	−5.066[a]
	(0.4608)	(0.0335)		(0.0606)					
RCA	0.0095	−0.0162	−0.0834	0.1790	0.1628	303.514	−138.947	0.015	0.113
	(0.0350)	(0.0138)	(0.0724)	(0.0560)	(0.1445)				
aHD(1)	7.8416	$-5.6\cdot10^{-6}$	0.25	0.1796		290.933	−135.856	−139.346	−2.642
	(0.1501)	($2.2\cdot10^{-6}$)		(0.0402)					

Australia: 10-year Bond from 1969:07 to 2019:12 (606 monthly obs.)

	δ	π	α	ϕ_1	$\phi_2/\theta_1/\sigma^2_\eta$	BIC	LIK	CLIK/Ms	LR/Vs
ARMA	0.0228	−0.0042			0.3940	84.916	−32.589	−33.233	−1.135
	(0.0363)	(0.0034)			(0.0210)				
SETAR(< 6.49)				0.1438	−0.0959				
				(0.0859)	(0.0835)				
SETAR(> 6.49)	0.1114	−0.0123		0.4082	−0.1531	106.293	−30.119	−32.104	−1.992
	(0.0607)	(0.0063)		(0.0409)	(0.0414)				
RCA	0.0268	−0.0249[a]	−0.0585	0.3290	0.1098	109.066	−41.377	0.023	0.104
	(0.312)	(0.0161)	(0.0430)	(0.0380)	(0.0858)				
aHD(1)	6.6507	$-5.4\cdot10^{-6}$	0.25		0.4091	**68.435**	−24.348	−33.233	−4.215[a]
	(0.1222)	($4.9\cdot10^{-7}$)			(0.0221)				

Canada: 10-year Bond from 1960:01 to 2019:12 (720 monthly obs.)

	δ	π	α	ϕ_1	$\phi_2/\theta_1/\sigma^2_\eta$	BIC	LIK	CLIK/Ms	LR/Vs
ARMA	−0.0044	−0.0009			0.4105	**−440.286**	230.012	229.447	−1.063
	(0.0256)	(0.0036)			(0.0259)				
SETAR(< 7.15)				0.3207					
				(0.0509)					
SETAR(> 7.15)	0.3723	−0.0451		0.4203		−425.524	229.210	223.281	−3.444
	(0.1119)	(0.0133)		(0.0473)					
RCA	−0.0052	−0.0175[c]	−0.0387	0.3779	0.0757	−420.643	223.477	0.022	0.082
	(0.0207)	(0.0148)	(0.0341)	(0.0374)	(0.0680)				
HD(1)	5.6924	$-2.9\cdot10^{-6}$	0.25		0.4119	−439.420	229.579	229.447	−0.514
	(0.0989)	($7.1\cdot10^{-6}$)			(0.0259)				

Germany: 10-year Bond from 1960:01 to 2019:12 (720 monthly obs.)

nomically proximate GB (Fig. 7), the results in Table 10 show that the Australia-GB and France-GB spreads are I(0) when the alternative is either ARMA or aHD(1), but they become I(1) when the alternative is either RCA or SETAR. Nevertheless, the conclusions that may be drawn for ARMA and aHD(1) are substantially different. On the one hand, ARMA modeling implies that the yields of Australia and France cointegrate with that of GB. On the other, aHD(1) modeling implies that the yields of Australia and France are stationary, in line with the suggested extension via the sum of two HD(1) components. Notably, with BIC consistently selecting the aHD(1) as the best specification, the unit-root hypothesis is rejected for every bond yield considered.

In a forecasting framework, the in-sample estimation of model parameters involves a trade-off between larger samples, which reduce estimation errors, and smaller samples (comprising more recent observations), which better capture structural breaks in the parameters. On the other hand, distinguishing between unit root and persistent stationary processes involves studying their long-run behavior, such as mean reversion or lack thereof. This, in turn, requires working with large samples of data to draw accurate conclusions. Nevertheless, to provide additional validation to the in-sample findings, one-step ahead forecasts are generated from each model for the yields and spreads that tested stationary. The out-of-sample rates are the US 3-month and 1-year[11] Treasury Bills from 2020:02:07 to 2023:07:14 (180 obs.) and from 2019:11:29 to 2023:07:14 (189 obs.), respectively. Additionally, the 10-year Government Bonds of the US, Canada and GB along with the spreads of Australia-GB, Germany-US, France-GB and Japan-US are considered from 2020:01:01 to 2023:05:01 (41 obs.). Table 11, which presents the forecast mean-square-errors (FMSE), demonstrates that the aHD(1) provides the most accurate forecasts in six instances, ARMA in two and SETAR in one. Diebold-Mariano tests of superior forecasting ability against competing alternatives are insignificant and are therefore not reported. Nevertheless, the null that the aHD(1) doesn't outperform other specifications out-of-sample is rejected at the 1% significance level (5% when excluding RCA from the comparison).

11 Although the 1-year Treasury Bill data resumes in 2008:05:30, it covers a period during which the US economy was in recession until 2010, followed by five years with the yield consistently close to zero. As a result, to avoid abnormal behaviors for which none of the models was trained in-sample, the out-of-sample period has been set to match that of the 3-month yield.

Table 7
Bond yields: estimation and unit-root testing results of No Trend specifications. Every sub-panel reports the estimation output under the unconstrained alternatives: ARMA, two-regimes SETAR, RCA and aHD(1). Autoregressive coefficients are ϕ_i, moving average coefficients are θ_j, and the variance of the innovations to the RCA latent process is σ_η^2. The lowest Bic of the four competing parameterizations is reported in bold, Lik is the log-likelihood, and cLik the constrained log-likelihood ($\delta = 0$, $\pi = 0$). LR is the value of the signed likelihood-ratio statistic for unit root testing in ARMA, SETAR and aHD(1), while Ms and Vs are the values of the mean- and variance-stationarity statistics for unit root testing in RCA. Significance levels are marked by [a], [b], and [c], indicating 1%, 2.5% and 5% significance, respectively.

	δ	π	α	ϕ_1	ϕ_2/σ_η^2	Bic	Lik	cLik/Ms	LR/Vs
ARMA	−0.0020	−0.0006		0.2359		−94.273	57.005	56.701	−0.780
	(0.0173)	(0.0022)		(0.0364)					
SETAR(< 10.71)				0.3048					
				(0.0469)					
SETAR(> 10.71)	0.4554	−0.0052		0.1384		−87.274	60.085	59.352	−1.211
	(0.1448)	(0.0108)		(0.0571)					
RCA	−0.0053	−0.0120	−0.0477	0.2332	0.0944	−87.142	56.727	0.014	0.075
	(0.0211)	(0.0114)	(0.0451)	(0.0375)	(0.0903)				
aHD(1)	6.8565	$-9.6{\cdot}10^{-7}$	0.25	0.2408		**−96.770**	58.254	56.701	−1.762
	(0.1425)	$(5.4{\cdot}10^{-7})$		(0.0363)					

France: 10-year Bond from 1960:01 to 2019:12 (720 monthly obs.)

	δ	π	α	ϕ_1	$\phi_2/\theta_1/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0250	−0.0042			0.3705	346.401	−163.332	−163.940	−1.103
	(0.0508)	(0.0044)			(0.0260)				
SETAR(< 11.74)				0.3473	−0.1024				
				(0.0523)	(0.0500)				
SETAR(> 11.74)	1.3666	−0.1041		0.4369	−0.1963	346.648	−150.297	−159.138	-4.205^{b}
	(0.3432)	(0.0255)		(0.0509)	(0.0546)				
RCA	0.0301	-0.0255^{a}	−0.0573	0.3377	0.1073	368.870	−171.279	0.024	0.104
	(0.0361)	(0.0167)	(0.0424)	(0.0371)	(0.0843)				
aHD(1)	7.3498	$-8.3{\cdot}10^{-6}$	0.25		0.3806	**332.113**	−156.188	−163.940	-3.94^{a}
	(0.1402)	$(1.2{\cdot}10^{-6})$			(0.0274)				

GB: 10-year Bond from 1960:01 to 2019:12 (720 monthly obs.)

	δ	π	α	ϕ_1	$\phi_2/\theta_1/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0030	−0.0077			0.3939	−315.991	166.874	165.353	−1.744
	(0.0215)	(0.0050)			(0.0417)				
SETAR(< 3.55)				0.2128	−0.1130				
				(0.0721)	(0.0717)				
SETAR(> 3.55)	0.0637	−0.0159		0.5909	−0.2652	−308.241	174.837	173.671	−1.527
	(0.0990)	(0.0180)		(0.0709)	(0.0728)				
RCA	0.0030	-0.0432^{s}	−0.0931	0.3614	0.1737	−303.421	163.543	0.041	0.175
	(0.0154)	(0.0312)	(0.0807)	(0.0543)	(0.1612)				
aHD(1)	2.0521	$-3.5{\cdot}10^{-5}$	0.25		0.3977	**−318.244**	168.000	165.353	−2.301
	(0.0978)	$(8.2{\cdot}10^{-6})$			(0.0415)				

Japan: 10-year Bond from 1989:01 to 2019:12 (372 monthly obs.)

5. Conclusions

The paper presents the HD(1) process of order one, incorporating hyperbolic reversion to equilibrium while explicitly excluding long memory. Other properties include ergodicity, stationarity, and the nesting of the AR(1). Importantly, the reversion speed is inversely related to the proximity of the process to equilibrium, exhibiting accelerated reversions as the process moves farther from equilibrium and negligible rates in its immediate vicinity.

The augmented aHD(1), introduced for straightforward estimation and testing, results in an ARMA with power error-correction, providing a good approximation to the HD(1) speed of reversion. Through a two-step estimation-calibration procedure based on ARMA routines, all model parameters, with the exception of the calibrated α, can be estimated using a standard econometric and statistical software package. Testing for unit roots is also straightforward, employing the signed LR statistic of Dickey and Fuller (1981) and the critical values tabulated for the stationary aHD(1) alternative.

The empirical study, involving Treasury Bill rates and Government Bond yields, shows that the aHD(1) provides a better description of the data (measured by the Bic) than competing ARMA, SETAR and RCA specifications. Testing, performed on the best specification, results in the rejection of the unit root hypothesis by SETAR, in one instance, and the aHD(1), in five instances. Testing the spreads of the apparent non-stationary yields against a stationary reference bond, further allows to reject the presence of a unit root in the 10-year bond of Germany (vs. US), Japan (vs. US), Australia (vs. GB) and France (vs. GB). In contrast, when considering only ARMA, SETAR and RCA as parameterizations of the alternative, the I(1) hypothesis can only be rejected for the US 3-month T-Bill and the Australian 10-year Bond. Therefore, by introducing error-corrections that are weak near equilibrium but grow stronger as the system deviates from it, the aHD(1) appears capable of reconciling empirical evidence with theoretical expectations.

Generalizations to HD(p), VAR models with hyperbolic rates of reversion, and other extensions for modeling the term structure of interest rates are left to future research.

CRediT authorship contribution statement

Alessandro Palandri: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Table 8
Bond spreads: estimation and unit-root testing results of No Trend specifications. Every sub-panel reports the estimation output under the unconstrained alternatives: ARMA, two-regimes SETAR, RCA and aHD(1). Autoregressive coefficients are ϕ_i, moving average coefficients are θ_j, and the variance of the innovations to the RCA latent process is σ_η^2. The lowest Bic of the four competing parameterizations is reported in bold, Lik is the log-likelihood, and cLik the constrained log-likelihood ($\delta = 0$, $\pi = 0$). LR is the value of the signed likelihood-ratio statistic for unit root testing in ARMA, SETAR and aHD(1), while Ms and Vs are the values of the mean- and variance-stationarity statistics for unit root testing in RCA. Significance levels are marked by [a], [b], and [c], indicating 1%, 2.5% and 5% significance, respectively.

	δ	π	α	ϕ_1	$\phi_2/\theta_1/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0395	−0.0252		0.1112	−0.1286	418.043	−196.208	−200.078	−2.782
	(0.0194)	(0.0091)		(0.0403)	(0.0404)				
SETAR(> 0.19)					−0.0853				
					(0.0430)				
SETAR(< 0.19)	−0.0577	−0.2088			−0.3819	421.052	−194.509	−198.814	−2.934
	(0.0475)	(0.0744)			(0.1103)				
RCA	0.0409	−0.0283[a]	−0.0526	0.0559	0.0533	422.269	−204.729	−0.025[a]	0.003
	(0.0201)	(0.0103)	(0.0228)	(0.1187)	(0.0399)				
aHD(1)	1.5098	−0.0088	1.00	0.1128	−0.1275	**416.847**	−195.610	−200.078	−2.989[c]
	(0.0624)	(0.0029)		(0.0403)	(0.0404)				

Australia−US: 10-year Bond spread from 1969:07 to 2019:12 (606 monthly obs.)

	δ	π	α	ϕ_1	$\phi_2/\theta_1/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0135	−0.0255		0.1660	−0.1204	−570.943	298.630	294.008	−3.040[c]
	(0.0078)	(0.0084)		(0.0370)	(0.0372)				
SETAR(< 0.44)				0.0502					
				(0.0739)					
SETAR(> 0.44)	0.1007	−0.0990		0.2013		−570.477	301.687	288.720	−5.093[a]
	(0.0231)	(0.0198)		(0.0423)					
RCA	0.0167	−0.0785[a]	−0.2463	0.1986	0.4331	−567.001	296.656	0.024	0.295
	(0.0086)	(0.0236)	(0.0889)	(0.0432)	(0.1755)				
aHD(1)	0.6016	−0.0363	0.50		0.2097	**−576.572**	298.155	290.726	−3.855[a]
	(0.0269)	(0.0097)			(0.0310)				

Canada−US: 10-year Bond spread from 1960:01 to 2019:12 (720 monthly obs.)

	δ	π	α	ϕ_1/θ_1	$\phi_2/\theta_2/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	−0.0089	−0.0097		0.2698	−0.1758	−49.897	38.107	36.150	−1.978
	(0.0110)	(0.0043)		(0.0214)	(0.0261)				
SETAR(> −1.23)				0.2222	−0.2027				
				(0.0542)	(0.0540)				
SETAR(< −1.23)	−0.1369	−0.0532		0.2621	−0.2027	−26.454	36.255	31.208	−3.177
	(0.0466)	(0.0168)		(0.0491)	(0.0497)				
RCA	−0.0087	−0.0340[a]	−0.0944	0.2270	0.1657	−5.231	15.771	0. 013	0.119
	(0.0106)	(0.0118)	(0.0406)	(0.0434)	(0.0802)				
aHD(1)	−0.3568	$-1.1{\cdot}10^{-4}$	0.25	0.2909	−0.1589	**−60.785**	43.551	36.150	−3.847[a]
	(0.0687)	$(1.9{\cdot}10^{-5})$		(0.0234)	(0.0290)				

Germany−US: 10-year Bond spread from 1960:01 to 2019:12 (720 monthly obs.)

Data availability

Data will be made available on request.

Appendix A. Wiener approximation of the HD(1)

Berkes et al. (2014) extend the KMT approximation of Komlós et al. (1975, 1976) to the class of dependent sequences $y_{t+1} = G(\dots, \varepsilon_t, \varepsilon_{t+1}, \varepsilon_{t+2}, \dots)$, where ε_{t+1} are i.i.d. random variables and $G : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}$ is a measurable function, by studying the rate at which the dependence condition $\Theta_{t+1,p}$ goes to zero as $(t+1)$ diverges. In addition, they demonstrate the fulfillment of the general conditions in Theorem 2.1 for iterated random processes of the type $y_{t+1} = G(y_t, \varepsilon_{t+1})$ where ε_{t+1} are i.i.d. and G is a measurable function, provided that:

$$\sup_{y \neq y'} \frac{\|g(y) - g(y')\|_p}{|y - y'|} < 1 \tag{14}$$

Consider the HD(1) representation with additive innovations $y_{t+1} = g(y_t) + \varepsilon_{t+1}$ with $\nu \in (0,1]$, and let $\mathcal{A} = \{y, y' \text{ s.t. } |y| > \lambda \text{ and } |y'| > \lambda\}$ for some $\lambda > 0$ s.t. $\mathbb{P}(\mathcal{A}) > 0$. The derivative $\partial g(y)/\partial y = \nu^{-1/\alpha}\left[\kappa^{-1/\alpha}|y|^{1/\alpha} + \nu^{-1/\alpha}\right]^{-\alpha-1}$ is such that $\sup_{y\in\mathcal{A}^c} g'(y) \le 1$ and $\sup_{y\in\mathcal{A}} g'(y) = \theta$ for some $\theta < 1$. Consequently, as $g(y)$ is Lipschitz with $|g(y) - g(y')|^p \le \left(\theta^p \mathbb{1}_{\mathcal{A}} + \mathbb{1}_{\mathcal{A}^c}\right) \cdot |y - y'|^p$, it follows that $\|y - y'\|_p^p \le (\theta^p \mathbb{P}(\mathcal{A}) + \mathbb{P}(\mathcal{A}^c))\,|y - y'|^p$, which can be rearranged to obtain condition (14).

Appendix B. Reversion: the role of α, κ and ν

B.1. The α parameter

Consider the absolute value of the deterministic component of the HD(1) in equation (1) and define $x = -\alpha^{-1}$ and $B_{t+h-1} = \max\{\kappa, \nu|y_{t+h-1}|\}$. Notice that B_{t+h-1} is independent of α for all $h \ge 1$. Specifically, for $h = 1$, $B_t = \max\{\kappa, \nu|y_t|\}$, where $|y_t|$ is a given value. As for $h > 1$, since $|y_{t+h-1}| < \kappa$ for every $h > 1$, it follows that $B_{t+h-1} = \kappa$. Subsequently, $|y_{t+h}|$ can be expressed as $|y_{t+h}| = B_{t+h-1}\left(1 + \theta_{t+h-1}^x\right)^{x^{-1}}$, with $\theta_{t+h-1} \in (0,1)$ for all $h \ge 1$. In particular, for $h > 1$, $\theta_{t+h-1} = \nu|y_{t+h-1}|/\kappa$, while for $h = 1$, $\theta_{t+h-1} = \min\{\nu|y_{t+h-1}|/\kappa, \kappa/\nu|y_{t+h-1}|\}$. Taking logs and differentiating w.r.t. x yields:

$$\frac{\partial}{\partial x}\ln|y_{t+h}| = \eta_{t+h-1} + \beta_{t+h-1}\frac{\partial}{\partial x}\ln\theta_{t+h-1}$$

Table 9
Bond spreads: estimation and unit-root testing results of No Trend specifications. Every sub-panel reports the estimation output under the unconstrained alternatives: ARMA, two-regimes SETAR, RCA and aHD(1). Autoregressive coefficients are ϕ_i, moving average coefficients are θ_j, and the variance of the innovations to the RCA latent process is σ_η^2. The lowest Bic of the four competing parameterizations is reported in bold, Lik is the log-likelihood, and cLik the constrained log-likelihood ($\delta = 0$, $\pi = 0$). LR is the value of the signed likelihood-ratio statistic for unit root testing in ARMA, SETAR and aHD(1), while Ms and Vs are the values of the mean- and variance-stationarity statistics for unit root testing in RCA. Significance levels are marked by [a], [b], and [c], indicating 1%, 2.5% and 5% significance, respectively.

	δ	π	α	ϕ_1/θ_1	$\phi_2/\theta_2/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0106	−0.0181		0.1916	−0.1412	181.336	−77.509	−80.368	−2.391
	(0.0187)	(0.0099)		(0.0244)	(0.0270)				
SETAR(< 1.86)				0.1363	−0.2083				
				(0.0505)	(0.0505)				
SETAR(> 1.86)	−0.0238	−0.0258		0.2966	−0.1755	181.591	−67.768	−76.485	−4.175[b]
	(0.0848)	(0.0304)		(0.0528)	(0.0537)				
RCA	0.0135	−0.0385[a]	−0.0944	0.1922	0.1474	207.406	−90.547	−0.005	0.085
	(0.0133)	(0.0129)	(0.0494)	(0.0507)	(0.0955)				
aHD(1)	0.8074	−0.0031	0.50	0.1940	−0.1372	**179.749**	−76.716	−80.368	−2.703
	(0.0526)	(0.0013)		(0.0246)	(0.0276)				

France−US: 10-year Bond spread from 1960:01 to 2019:12 (720 monthly obs.)

	δ	π	α	ϕ_1	$\phi_2/\theta_1/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0.0375	−0.0305			0.3595	434.880	−207.571	−212.869	−3.255[b]
	(0.0263)	(0.0067)			(0.0266)				
SETAR(< 2.72)				0.3038	−0.1799				
				(0.0527)	(0. 0492)				
SETAR(> 2.72)	0.2249	−0.0611		0.4295	−0.2356	438.957	−196.451	−201.600	−3.209
	(0.0981)	(0.0212)		(0.0501)	(0. 0547)				
RCA	0.0223	−0.0741[a]	−0.1362	0.3284	0.2229	469.174	−221.431	0.029	0.194
	(0.0183)	(0.0218)	(0.0564)	(0.0430)	(0.1107)				
aHD(1)	1.3006	$-4.3{\cdot}10^{-5}$	0.25		0.3665	**424.767**	−202.515	−212.869	−4.551[a]
	(0.0653)	$(4.4{\cdot}10^{-6})$			(0.0279)				

GB−US: 10-year Bond spread from 1960:01 to 2019:12 (720 monthly obs.)

	δ	π	α	ϕ_1/θ_1	θ_2/σ_η^2	Bic	Lik	cLik/Ms	LR/Vs
ARMA	−0.0809	−0.0338		0.1937	−0.1393	−85.380	54.528	50.876	−2.703
	(0.0378)	(0.0133)		(0.0529)	(0.0503)				
SETAR(>-3.32)				0.1849					
				(0.0594)					
SETAR(<-3.32)	−0.7179	−0.1923		0.1141		−75.128	49.402	44.956	−2.982
	(0.2780)	(0.0711)		(0.1011)					
RCA	−0.1031	−0.0639[a]	−0.1571	0.1753	0.2418	−74.949	49.307	−0.010	0.137
	(0.0412)	(0.0252)	(0.0969)	(0.0774)	(0.1912)				
aHD(1)	−2.5859	−0.0380	1.00		0.2454	**−88.295**	53.026	46.876	−3.507[a]
	(0.0474)	(0.0109)			(0.0529)				

Japan−US: 10-year Bond spread from 1989:01 to 2019:12 (372 monthly obs.)

where $\beta_{t+h-1} = \theta^x_{t+h-1}/(1+\theta^x_{t+h-1}) \in (0,1)$ and $\eta_{t+h-1} = x^{-2}\left[\beta_{t+h-1}\ln\theta^x_{t+h-1} - \ln(1+\theta^x_{t+h-1})\right] < 0$. Since, for $h=1$, the derivative $\partial \ln\theta_t/\partial x$ equals zero, it follows that $\partial \ln|y_{t+1}|/\partial x = \eta_t$. For all other $h>1$, however, $\partial\ln|y_{t+h}|/\partial x = \eta_{t+h-1} + \beta_{t+h-1}\partial\ln|y_{t+h-1}|/\partial x$. Consequently, it can be concluded that $\partial\ln|y_{t+h}|/\partial x < 0$, and thus, $\partial|y_{t+h}|/\partial\alpha < 0$ for all $h \geq 1$.

B.2. *The κ and ν parameters*

Expressing the deterministic component of the HD(1) in (1) in the autoregressive form $|y_{t+h}|^{-1/\alpha} = \kappa^{-1/\alpha} + \nu^{-1/\alpha}|y_{t+h-1}|^{-1/\alpha}$ and differentiating w.r.t. κ yields:

$$\frac{\partial}{\partial\kappa}|y_{t+h}|^{-1/\alpha} = -\frac{1}{\alpha}\kappa^{-1-1/\alpha} + \nu^{-1/\alpha}\frac{\partial}{\partial\kappa}|y_{t+h-1}|^{-1/\alpha}$$

Since, for $h=1$, $\partial|y_t|^{-1/\alpha}/\partial\kappa$ equals zero, it follows that $\partial|y_{t+1}|^{-1/\alpha}/\partial\kappa < 0$. Consequently, iterating for all other $h>1$, gives $\partial|y_{t+h}|^{-1/\alpha}/\partial\kappa < 0$. Therefore, $\partial|y_{t+h}|/\partial\kappa > 0$ for all $h \geq 1$. By following analogous steps, with appropriate adjustments made where necessary, it can be shown that also $\partial|y_{t+h}|/\partial\nu > 0$ for all $h \geq 1$.

Appendix C. First-order equivalence

To streamline the notation for HD(1), consider the following simplifications: $y = |y_{t+1}|$, $x = |y_t|$, $b = \kappa^{-1/\alpha}$, $c = \nu^{-1/\alpha} > 1$. Next rewrite the deterministic component $g(y_t)$ of the process as $y = (b + cx^{-1/\alpha})^{-\alpha}$. Define $d = \mathbb{1}[x \leq 1]$. When $d = 1$, the first-order Taylor expansion of y around $x = 1$ is $y = (b+c)^{-\alpha} + (b+c)^{-\alpha-1}c(x-1) + \mathcal{O}[(x-1)^2]$. On the other hand, when $d = 0$, the first-order expansion around $x^{-1} = 1$ is: $y = (b+c)^{-\alpha} - (b+c)^{-\alpha-1}c(x^{-1}-1) + \mathcal{O}[(x^{-1}-1)^2]$, whereas when $d = 0$ the first-order expansion around $x^{-1} = 1$. Combining the two cases gives:

$$y = (b+c)^{-\alpha} + (b+c)^{-\alpha-1}c\left(x^d - x^{d-1}\right) + \mathcal{O}\left[\left(x^d - x^{d-1}\right)^2\right]$$

Consider now the process $\tilde{y}$ with parameters $\tilde{\alpha} \neq \alpha$, $\tilde{b}$ and $\tilde{c}$. For the difference $\tilde{y} - y$ to be of order $\mathcal{O}\left[\left(x^d - x^{d-1}\right)^2\right]$, it is necessary that the constants and the coefficients of x and x^{-1} cancel each other in pairs. Specifically, $(\tilde{b}+\tilde{c})^{-\tilde{\alpha}} = (b+c)^{-\alpha}$ and $(\tilde{b}+\tilde{c})^{-\tilde{\alpha}-1}\tilde{c} = (b+c)^{-\alpha-1}c$. Solving for $\tilde{b}$ and $\tilde{c}$ yields $\tilde{b} = (b+c)^{\alpha/\tilde{\alpha}-1}b$ and $\tilde{c} = (b+c)^{\alpha/\tilde{\alpha}-1}c$. Therefore, there are infinitely many combinations of parameters (α, κ, ν) that result in processes being first-order equivalent. Recall that, due to the station-

Table 10

Bond spreads: estimation and unit-root testing results of No Trend specifications. Every sub-panel reports the estimation output under the unconstrained alternatives: ARMA, two-regimes SETAR, RCA and aHD(1). Autoregressive coefficients are ϕ_i, moving average coefficients are θ_j, and the variance of the innovations to the RCA latent process is σ_η^2. The lowest Bic of the four competing parameterizations is reported in bold, Lik is the log-likelihood, and cLik the constrained log-likelihood ($\delta = 0$, $\pi = 0$). LR is the value of the signed likelihood-ratio statistic for unit root testing in ARMA, SETAR and aHD(1), while Ms and Vs are the values of the mean- and variance-stationarity statistics for unit root testing in RCA. Significance levels are marked by ^{a}, ^{b}, and ^{c}, indicating 1%, 2.5% and 5% significance, respectively.

	δ	π	α	ϕ_1/θ_1	$\phi_2/\theta_2/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	0. 0150	−0.0296			0.2275	613.851	−297.315	−302.201	-3.126^b
	(0.0202)	(0.0065)			(0.0288)				
SETAR(>-2.00)				0.0629	−0.1226				
				(0.0496)	(0.0480)				
SETAR(<-2.00)	−0.5022	−0.1549		0.4439	−0.1960	606.274	−280.713	−287.720	−3.744
	(0.1588)	(0.0434)		(0.0663)	(0.0744)				
RCA	0.0183	−0.0530	−0.1291	0.2186	0.2071	625.769	−300.074	0.001	0.131
	(0.0189)	(0.0214)	(0.0871)	(0.0590)	(0.1701)				
aHD(1)	0.3119	$-5.8\cdot10^{-5}$	0.25	0.1802	−0.1056	**604.503**	−289.438	−298.034	-4.146^a
	(0.0847)	$(8.9\cdot10^{-6})$		(0.0310)	(0.0329)				

Australia−GB: 10-year Bond spread from 1969:07 to 2019:12 (606 monthly obs.)

	δ	π	α	ϕ_1	$\phi_1/\theta_1/\sigma_\eta^2$	Bic	Lik	cLik/Ms	LR/Vs
ARMA	−0.0293	−0.0557			0.3538	327.644	−153.953	−164.044	-4.492^a
	(0.0158)	(0.0073)			(0.0260)				
SETAR(>-1.28)				0.1988	−0.1521				
				(0.0455)	(0.0454)				
SETAR(<-1.28)	−0.1752	−0.1057		0.4797	−0.2210	339.095	−146.520	−153.452	−3.723
	(0.0785)	(0.0336)		(0.0600)	(0.0625)				
RCA	−0.0279	-0.0899^a	−0.1349	0.3150	0.1829	361.669	−167.679	−0.003	0.128
	(0.0145)	(0.0210)	(0.0495)	(0.0535)	(0.0929)				
aHD(1)	−0.4932	−0.0189	1.00		0.3514	**324.840**	−152.551	−164.044	-4.794^a
	(0.0456)	(0.0040)			(0.0375)				

France−GB: 10-year Bond spread from 1960:01 to 2019:12 (720 monthly obs.)

Table 11

Forecast mean-squared-errors (FMSE) of the ARMA, SETAR, RCA and aHD(1) specifications. The lowest FMSE of the four competing parameterizations is reported in bold.

	T	ARMA	SETAR	RCA	aHD(1)		T	ARMA	SETAR	RCA	aHD(1)
US3M:	180	0.007219	0.007871	0.007766	**0.007096**	AU10-GB10:	41	0.042648	0.045977	0.042598	**0.042565**
US01:	189	0.011166	**0.011150**	0.011840	0.011431	DE10-US10:	41	0.019454	0.020656	0.020546	**0.018932**
US10:	41	0.045032	0.046273	0.047417	**0.043775**	FR10-GB10:	41	**0.023062**	0.023528	0.023248	0.023728
CA10:	41	0.041578	0.044548	0.042901	**0.041076**	JP10-US10:	41	0.054547	0.055247	0.054510	**0.053015**
GB10:	41	**0.075002**	0.076693	0.079778	0.075575						

arity of the process, the difference $\mathcal{O}\left[\left(x^d - x^{d-1}\right)^2\right]$ between $\tilde{y}_{t+1}$ and y_{t+1} diminishes in magnitude from period $t+2$ onward.

Appendix D. Stationarity of the oscillatory aHD(1)

Although the aHD(1) process has a root strictly less than unity for any negative π, a necessary condition for its stationarity is that the root is strictly greater than minus one. Since the aHD(1) is a power-autoregression, a necessary condition for stationarity is that innovations are bounded, as outlined in Chan and Tong (2001). Without loss of generality, consider the process in equation (7) with $\mu_t = 0$ for all t and bounded support for the innovations $\varepsilon_{t+1} \in (-a, a)$ for every t:

$$y_{t+1} = y_t + \pi y_t |y_t|^{1/\alpha} + \varepsilon_{t+1}$$

The necessary and sufficient condition for the stationarity of y_{t+1} is the existence of a global attractor $|\partial y_{t+1}/\partial y_t| < 1$, corresponding to $-1 < 1 + \frac{\pi(\alpha+1)}{\alpha}|y_t|^{1/\alpha} < 1$. Given $\pi < 0$, the right-hand-side of the inequality is always satisfied. As for the left-hand-side, it implies that $|y_t| < \left[-2\alpha/\pi(\alpha+1)\right]^\alpha$. Thus, the lower and upper bounds of the process are, respectively, $y_L = -\left[-2\alpha/\pi(\alpha+1)\right]^\alpha$ and $y_H = \left[-2\alpha/\pi(\alpha+1)\right]^\alpha$. Since the upper bound is attained at $y_H = \left(1 - \alpha\theta|y_L|^{1/\alpha}\right) y_L + a$ and since $y_L = -y_H$, it follows that $a = 2y_H + \pi y_H^{1+1/\alpha}$. Finally, substituting the expression for y_H in that of a gives the bound on the innovations:

$$a = \frac{2}{\alpha+1}\left(\frac{2\alpha}{\alpha+1}\right)^\alpha |\pi|^{-\alpha}$$

Therefore, when π has a small modulus, the magnitude of the innovations bound a - proportional to $|\pi|^{-\alpha}$ - is large. Consequently, both the lower an upper bounds of the process exhibit ample magnitudes.

Appendix E. Stationarity of RCA

The infinite moving average representation of the RCA model is:

$$y_{t+1} = \delta\left[1 + \sum_{j=0}^{\infty} S_j\right] + \sigma_\epsilon\left[\epsilon_{t+1} + \sum_{j=0}^{\infty} S_j \epsilon_{t-j}\right] + \lim_{\kappa\to\infty} S_\kappa y_{t-\kappa}$$

where $\ln S_j \equiv \sum_{n=0}^{j} \ln(1 + \pi_{t+1-n})$ with $\mathbb{E}[\ln S_j] = (j+1)\alpha(1-\phi_1)^{-1}$ and $\mathbb{V}[\ln S_j] = \sigma_\eta^2(j+1)(1+\phi_1)(1-\phi_1)^{-1}(1-\phi_1^2)^{-1} - 2\sigma_\eta^2\phi_1(1-\phi_1^{j+1})(1-\phi_1)^{-2}(1-\phi_1^2)^{-1}$. The term $S_\kappa y_{t-\kappa}$ becomes negligible as $\kappa \to \infty$, provided that $\lim_{\kappa\to\infty} \kappa^{-1}\sum_{n=0}^{\kappa} \ln(1 + \pi_{t+1-n}) < 0$ holds, or equivalently, when the expectation $\mathbb{E}\left[\ln(1+\pi_{t+1})\right]$ is negative. Consequently, under the constraint $\alpha < 0$, the RCA representation simplifies to:

$$y_{t+1} = \delta\left[1 + \sum_{j=0}^{\infty} S_j\right] + \sigma_\epsilon\left[\epsilon_{t+1} + \sum_{j=0}^{\infty} S_j \epsilon_{t-j}\right]$$

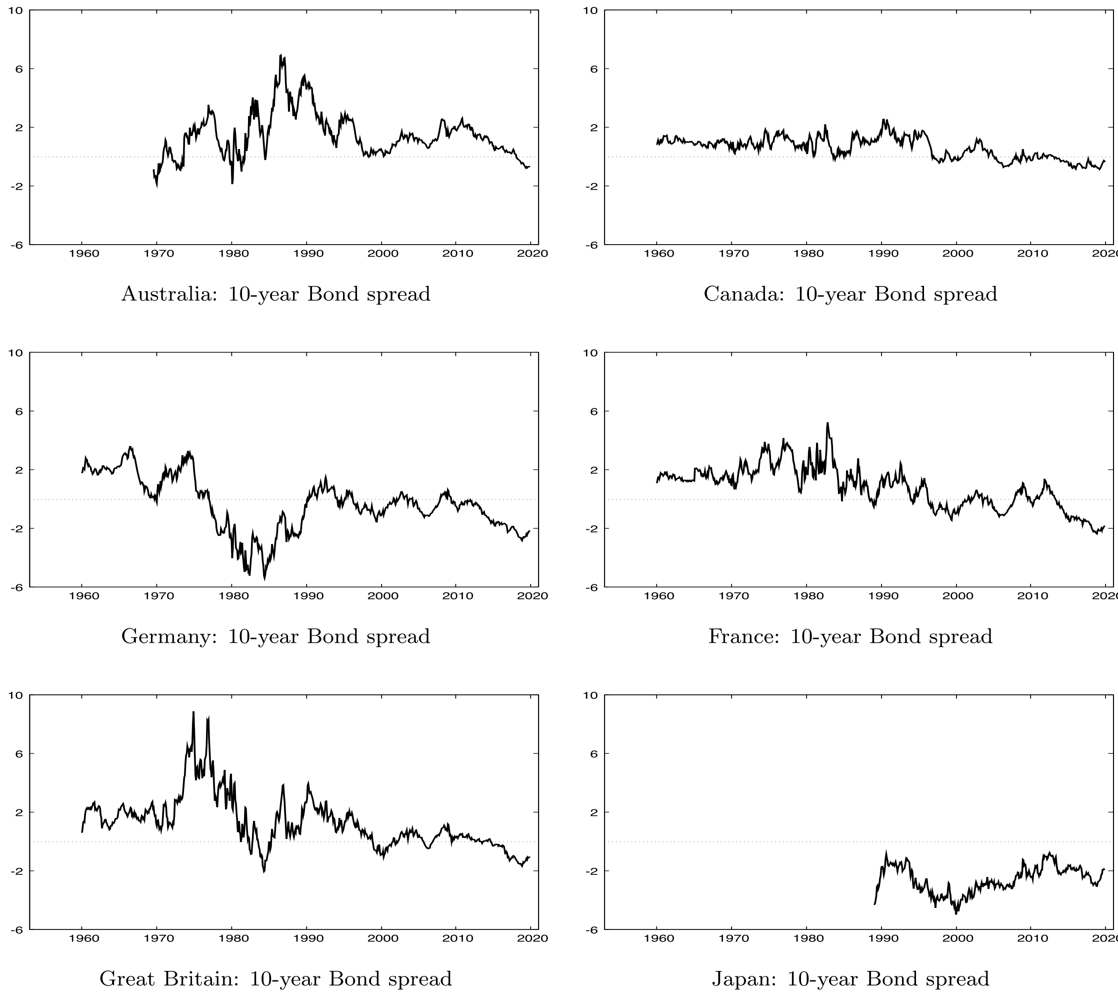


Fig. 6. Time-series of Bond spreads (US 10-year Bond benchmark).

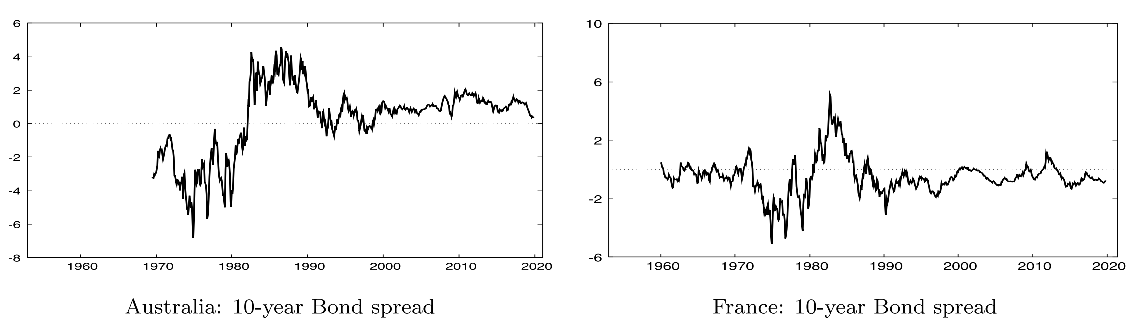


Fig. 7. Time-series of selected Bond spreads (GB 10-year Bond benchmark).

Mean stationarity implies that $\mathbb{E}\left[y_{t+1}\right] = \delta\left[1 + \sum_{j=0}^{\infty}\mathbb{E}(S_j)\right]$ should be finite, a criterion met when the series is absolutely convergent: $\lim_{j\to\infty}\mathbb{E}(S_j)/\mathbb{E}(S_{j-1}) < 1$. Under of the assumption of Gaussian innovations η_{t+1}, the condition becomes:

$$\lim_{j\to\infty}\exp\left\{\mathbb{E}[\ln S_j] + \frac{1}{2}\mathbb{V}[\ln S_j] - \mathbb{E}[\ln S_{j-1}] - \frac{1}{2}\mathbb{V}[\ln S_{j-1}]\right\} < 1$$

$$\lim_{j\to\infty}\frac{\alpha}{1-\phi_1} + \frac{\sigma_\eta^2}{2(1-\phi_1^2)}\cdot\left[\frac{1+\phi_1}{1-\phi_1} - \frac{2\phi_1^{j+1}}{1-\phi_1}\right] < 0$$

$$\alpha + \frac{\sigma_\eta^2}{2(1-\phi_1)} < 0$$

Exploiting the fact that the ϵ_{t+1} are serially uncorrelated, the requirement for variance stationarity is that $\sum_{j=0}^{\infty}\mathbb{E}(S_j^2)$ remains finite. Recognizing the relationship $\mathbb{E}(S_j^2) = \mathbb{E}(\exp(2\ln S_j))$, the series is absolutely convergent if:

$$\lim_{j\to\infty}\exp\left\{2\mathbb{E}[\ln S_j] + 2\mathbb{V}[\ln S_j] - 2\mathbb{E}[\ln S_{j-1}] - 2\mathbb{V}[\ln S_{j-1}]\right\} < 1$$

$$\lim_{j\to\infty}\frac{2\alpha}{1-\phi_1} + \frac{2\sigma_\eta^2}{1-\phi_1^2}\cdot\left[\frac{1+\phi_1}{1-\phi_1} - \frac{2\phi_1^{j+1}}{1-\phi_1}\right] < 0$$

$$\alpha + \frac{\sigma_\eta^2}{1-\phi_1} < 0$$

References

Anderson, H.M., 1997. Transaction costs and nonlinear adjustment toward equilibrium in the US Treasury Bill market. Oxf. Bull. Econ. Stat. 59, 465–484.

Baillie, R., 1996. Long memory processes and fractional integration in econometrics. J. Econom. 73, 5–59.

Balke, N.S., Fomby, T.B., 1997. Threshold cointegration. Int. Econ. Rev. 38, 627–645.

Bansal, R., Zhou, H., 2002. Term structure of interest rates with regime shifts. J. Finance 57, 1997–2043.

Bec, F., Ben Salem, M., Carrasco, M., 2004. Tests for unit-root versus threshold specification with an application to the purchasing power parity relationship. J. Bus. Econ. Stat. 22, 382–395.

Bekaert, G., Hodrick, R.J., 2001. Expectations hypotheses tests. J. Finance 56, 1357–1394.

Berkes, I., Liu, W., Wu, W.B., 2014. Komlós-Major-Tusnády approximation under dependence. Ann. Probab. 42, 794–817.

Busetti, F., Harvey, A., 2010. Test of strict stationarity based on quantile indicators. J. Time Ser. Anal. 31, 435–450.

Campbell, J.Y., Shiller, R.J., 1991. Yield spreads and interest rate movements: a bird's eye view. Rev. Econ. Stud. 58, 495–514.

Caner, M., Hansen, B.E., 2001. Threshold autoregression with a unit root. Econometrica 69, 1555–1596.

Chan, K.-S., Tong, H., 2001. Chaos: a Statistical Perspective. Springer-Verlag, New York.

Chang, S.Y., Perron, P., 2017. Fractional unit root tests allowing for a structural change in trend under both the null and alternative hypotheses. Econometrics 5, 1–26.

Cho, C.-K., Amsler, C., Schmidt, P., 2015. A test of the null of integer integration against the alternative of fractional integration. J. Econom. 187, 217–237.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc. 72, 427–431.

Dickey, D.A., Fuller, W.A., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica 49, 1057–1072.

Distaso, W., 2008. Testing for unit root processes in random coefficient autoregressive models. J. Econom. 142, 581–609.

Dolado, J.J., Gonzalo, J., Mayoral, L., 2002. A fractional Dickey-Fuller test for unit roots. Econometrica 70, 1963–2006.

Dolado, J.J., Gonzalo, J., Mayoral, L., 2008. Wald tests of i(1) against i(d) alternatives: some new properties and an extension to processes with trending components. Stud. Nonlinear Dyn. Econom. 12, 1–35.

Enders, W., Granger, C.W.J., 1998. Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates. J. Bus. Econ. Stat. 16, 304–311.

Engle, R., Granger, C.W.J., 1987. Co-integration and error correction: representation, estimation and testing. Econometrica 55, 251–276.

Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time series models. J. Time Ser. Anal. 4, 221–238.

Gourieroux, C., Monfort, A., Trognon, A., 1984. Pseudo maximum likelihood methods: theory. Econometrica 52, 681–700.

Granger, C.W.J., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. J. Time Ser. Anal. 1, 15–30.

Granger, C.W.J., Swanson, N.R., 1997. An introduction to stochastic unit-root processes. J. Econom. 80, 35–62.

Granger, C.W.J., Teräsvirta, T., 1993. Modeling Economic Nonlinear Relationships. Oxford University Press, Oxford.

Gray, S.F., 1996. Modeling the conditional distribution of interest rates as a regime-switching process. J. Financ. Econ. 42, 27–62.

Hall, A.D., Anderson, H.M., Granger, C.W.J., 1992. A cointegration analysis of Treasury Bill yields. Rev. Econ. Stat. 74, 116–126.

Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 414–430.

Hoek, H., Lucas, A., van Dijk, H.K., 1995. Classical and Bayesian aspects of robust unit root inference. J. Econom. 69, 27–59.

Horváth, L., Trapani, L., 2019. Testing for randomness in a random coefficient autoregression. J. Econom. 209, 338–352.

Hosking, J.R.M., 1981. Fractional differencing. Biometrika 68, 165–176.

Komlós, J., Major, P., Tuesnády, G., 1975. An approximation of partial sums of independent RV's and the sample DF. I. Z. Wahrscheinlichkeitstheor. Verw. Geb. 32, 111–131.

Komlós, J., Major, P., Tuesnády, G., 1976. An approximation of partial sums of independent RV's and the sample DF. II. Z. Wahrscheinlichkeitstheor. Verw. Geb. 34, 33–58.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesys of stationarity against the alternative of a unit root: how sure are we that economic times series have a unit root? J. Econom. 54, 159–178.

Leybourne, S.J., McCabe, B.P., Tremayne, A.R., 1996. Can economic time series be differenced to stationarity? J. Bus. Econ. Stat. 14, 435–446.

Lo, M.C., Zivot, E., 2001. Threshold cointegration and nonlinear adjustment to the law of one price. Macroecon. Dyn. 5, 533–576.

Lobato, I.N., Velasco, C., 2007. Efficient Wald tests for fractional unit roots. Econometrica 75, 575–589.

Lubrano, M., 1995. Testing for unit roots in a Bayesian framework. J. Econom. 69, 81–109.

MacDonald, R., Murphy, P., 1989. Testing for the long-run relationship between nominal interest rates and inflation using cointegration techniques. Appl. Econ. 21, 439–447.

MacKinnon, J.G., 2010. Critical values for cointegration tests. Working Paper 1227. Economics Department, Queen's University.

McCabe, B.P., Tremayne, A.R., 1995. Testing a time series for difference stationarity. Ann. Stat. 23, 1015–1028.

Michael, P., Nobay, R.A., Peel, D.A., 1997. Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation. J. Polit. Econ. 105, 862–879.

Nagakura, D., 2009. Asymptotic theory for explosive random coefficient autoregressive models and inconsistency of a unit root test against a stochastic unit root process. Stat. Probab. Lett. 79, 2476–2483.

Pesaran, M.H., Potter, S.M., 1997. A floor and ceiling model of US output. J. Econ. Dyn. Control 21, 661–695.

Phillips, P.C.B., 1999. Discrete Fourier transforms of fractional processes. Working Paper 1243. Cowles Foundation.

Rapach, D.E., Weber, C.E., 2004. Are real interest rates really nonstationary? New evidence from tests with good size and power. J. Macroecon. 26, 409–430.

Robinson, P.M., 1994. Efficient tests of nonstationary hypotheses. J. Am. Stat. Assoc. 89, 1420–1437.

Robinson, P.M., 1995. Log-periodogram regression of time series with long range dependence. Ann. Stat. 23, 1048–1072.

Rose, A.K., 1988. Is the real interest rate stable? J. Finance 43, 1095–1112.

Sarno, L., Taylor, M.P., Valente, G., 2006. The role of asymmetries and regime shifts in the term structure of interest rates. J. Bus. 79, 1193–1224.

Sarno, L., Thornton, D.L., Valente, G., 2007. The empirical failure of the expectation hypothesis of the term structure of bond yields. J. Financ. Quant. Anal. 42, 81–100.

Sercu, P., Uppal, R., van Hulle, C., 1995. The exchange rate in the presence of transaction costs: implications for tests of purchasing power parity. J. Finance 50, 1309–1319.

Shapiro, M., Watson, M., 1998. Sources of business cycle fluctuations. In: Fischer, S. (Ed.), NBER Macroeconomics Annual. MIT Press.

Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. J. Econom. 53, 165–188.

Tanaka, K., 1999. The nonstationary fractional unit root. Econom. Theory 15, 549–582.

Taylor, A., 2001. Potential pitfalls for the ppp puzzle? Sampling and specification biases in mean reversion tests of the LOOP. Econometrica 69, 473–498.

Tong, H., 1990. Nonlinear Time Series: a Dynamical System Approach. Oxford University Press, Oxford.

Trapani, L., 2021. Testing for strict stationarity in a random coefficient autoregressive model. Econom. Rev. 40, 220–256.