On the Small Sample Properties of Dickey-Fuller and Maximum Likelihood Unit Root Tests on Discrete-Sampled Short-Term Interest Rates[*]

Paulo M.M. Rodrigues
Faculty of Economics
University of Algarve
Portugal
E-mail: prodrig@ualg.pt

Antonio Rubia
Department of Financial Economics
University of Alicante
Spain
E-mail: antonio.rubia@ua.es

February 2004

Abstract

Testing for unit roots in short-term interest rates plays a key role in the empirical modelling of these series. It is widely assumed that the volatility of interest rates follows some time-varying function which is dependent of the level of the series. This may cause distortions in the performance of conventional tests for unit root nonstationarity since these are typically derived under the assumption of homoskedasticity. Given the relative unfamiliarity on the issue, we conducted an extensive Monte Carlo investigation in order to assess the performance of the DF unit root tests, and examined the effects on the limiting distributions of test procedures (t- and likelihood ratio tests) based on maximum likelihood estimation of models for short-term rates with a linear drift.

Key Words: Unit root, interest rates, CKLS model.

JEL Classification: C12, C15, C52, E43.

1 Introduction

There is a vast literature that deals with modelling risk free short-term interest rates dynamics since this is a major economic variable. Models that describe interest rate dynamics normally assume a continuous time-generating process. Stochastic differential equations (SDE) are commonly used in this context to characterise the term structure and to price security assets. A key underlying assumption behind the usual empirical applications of such models on interest rates time series concerns the mean-reversion property. Nonstationarity is usually evaluated based on estimated parameters from the SDE solution and/or within the unit root testing framework. In choosing the former case, evidence against the null of integration is taken by performing significance tests on the relevant parameters from the SDE solution, while the tests proposed by Dickey and Fuller [DF] (1979, 1981) continue to be the leading methods for the latter case.

Interest rates display time-varying conditional volatility patterns which, unlike most financial time series, are not regarded as a measurable function of squared innovations, but are largely considered as some function of the lagged level of the series itself (recall that the former is the basic structure under GARCH-type models). Thus, fitting GARCH-type models to these time series often cause parameter estimates that exhibit integration and even explosive patterns as a result of model misspecification. Although there are a number of studies that deal with asymptotic distributions and properties of unit root tests in the presence of conditionally heteroskedastic errors (see Li, Ling and McAleer, 2002, for an interesting overview), little attention has been given to the driving process of volatility following a highly nonlinear, level-dependent functional form, characteristic of interest rate time series. In fact, available research has generally centred on either the assumption of unconditional heteroskedasticity or stationary GARCH type errors, and has avoided alternative conditional heteroskedasticity structures which are relevant in practice.

Given the importance of testing for nonstationarity when modelling short-term interest rate dynamics, the aim of this paper is twofold. First, we investigate the effects of level-dependent volatility on the DF tests (t-test and likelihood ratio test) and a suitable nonparametric alternative. Second, we assess the behaviour of test procedures based on maximum likelihood estimates of parameters taken from the approximate solution of the SDE. Specifically, we consider the class of one-factor diffusion models with a linear drift and a nonlinear volatility function and study both the limiting distribution of the t-test on the autoregressive parameter and the likelihood ratio test on the drift parameters.

The t-test is extensively applied in empirical papers in this context. It should be noted that most applied papers tend to refrain from making explicit reference to the limiting distribution used when eliciting the significance of a statistic. This ambiguity is problematic because standard theory (as implicitly used) does not hold under the hypothesis of a unit root and hence statistical inference may be misleading. Furthermore, since there is no formal proof that supports the presence of a unique asymptotic distribution for the t-test, this paper also looks to shed some light on this issue. Given that this type of level dependent volatility imposes analytical limitations and renders the derivation of limit distributions unfeasible, we look to Monte Carlo experimentation as a valuable means for assessing the effects of such patterns on the performance of the test procedures considered.

Non-stationarity in interest rates seems to be a controversial point in practice. There are strong theoretical reasons supporting stationarity of short-term rates, but empirical

evidence mostly shows weak or no reversion at all. A prudent explanation for this evidence could be based on the fact that misspecified linear models tend to find spurious evidence of nonstationary behaviour, as suggested by, *inter alia*, Aït-Sahalia (1996). However, it has been shown that methods that estimate nonlinear drifts in diffusions, such as the nonparametric methods proposed by Aït-Sahalia (1996), tend to find nonlinearities as an artifact of the procedure; *c.f.* Chapman and Pearson (2000). In addition, it is often found that simple linear models cannot be rejected for more flexible, nonlinear models (see, *inter alia*, Hong and Li, 2002; and references in Chapman and Pearson, 2001).

Further, one-factor diffusions are potentially too parsimonious for correctly adjusting the complex dynamics of interest rates; *c.f.* Hong and Li (2002)[1]. However, single-factor models with a linear drift are widely used in practice because of their tractability and ability to fit yield curve of securities reasonably well (see Rogers and Stummer, 2000, and references therein). This is because most of the observed variability of the bond prices can be explained through a single factor[2]. Given the importance of these types of models in the empirical modelling of short-term interest rate time series, the scope of this paper looks to assess the effects of nonlinear patterns in volatility when testing against a unit root, to better comprehend the modelling and practical fitting of these series.

This paper is organised as follows. Section 2 discusses a general diffusion model for interest rate dynamics as well as its econometric extensions. Section 3 presents the Monte Carlo simulations with Section 4 reporting the main results. Finally, Section 5 summarises the empirical findings.

2   Short-term Interest Rates Modelling

For notational purposes, recall that short-term interest rates, say $(r_t)_{t>0}$, are often regarded as state variables following a one-factor time-homogeneous diffusion in continuous time and represented as,

$$dr_t = \mu\left(r_t; \theta\right) dt + \sigma\left(r_t; \theta\right) dW_t \qquad (1)$$

where W_t is a Wiener (Brownian) process and both the drift, $\mu\left(\cdot\right)$, and the volatility function, $\sigma\left(\cdot\right)$, depend solely on the state variable, r_t, and a set of parameters θ. Continuous-time diffusion processes have been extensively used in economics and finance. In the case of research based on interest rates, these processes provide a basic tool for pricing the term structure in both the arbitrage-free and the equilibrium framework.

We can observe specific differences of one-factor SDEs when applied to short-term rate dynamics in the linear or nonlinear nature of $\mu\left(\cdot\right)$ and in the way in which the volatility function $\sigma\left(\cdot\right)$ is related to the level of the process. Particular cases of (1) are nonlinear diffusions with linear drift and time-varying volatility. A more general model of this type was proposed by Chan, Karolyi, Longstaff and Sanders [CKLS] (1992), and expressed as,

$$dr_t = \left(\alpha_0 + \alpha_1 r_t\right) dt + \sigma r_t^{\gamma} dW_t \qquad (2)$$

---

[1] Hong and Li (2002) compare the application of different models on US interest-rate data. They find that all diffusions with a single factor are rejected, regardless of the linear or non-linear characteristics of the drift function. They also establish that the nonlinear drift model of Aït-Sahalia (1996), earlier applied on the same data, does not significantly improve the goodness of fit over the linear drift alternative.

[2] As reported in several studies, roughly 90% of variation in US treasury rates can be traced to changes in levels of interest rates. In one-factor models, this factor is identified with short-term interest rates.

where $\gamma \geq 0$ represents the elasticity of volatility against the level of the series. When this parameter is non-zero, the conditional volatility of the model increases according to the level of the interest rate.

Remark 2.1: The CKLS diffusion in (2) combines a large number of well-known models from the literature; such as, the Orstein-Uhlenbek process used in Vasicek (1977), the model proposed by Brennan and Schwartz (1979), and the square-root model of Cox, Ingerson and Ross [CIR] (1985), among others. Because of the generality inherited in the CKLS model, it will later serve as benchmark in our analysis.

Despite the characteristics of continuous-time diffusions, the parameters involved are always estimated in the discrete-time domain based on simulations of the SDE solution. Explicit solutions are only available for very few models (*e.g.*, the Ornstein-Uhlenbek and the CIR model) and as such, parameters are generally estimated from approximations based on numerical techniques. The Euler scheme is the preferred method because of its simplicity and computational convenience. In fact, it is often applied in many papers even when exact solutions are available.

Thus, given that model (2) is in general not explicitly solvable, its Euler discretisation results in,

$$\Delta_m r_t = \alpha_0 m + \alpha_1 m r_{t-1} + \sigma \sqrt{m} r_{t-1}^\gamma \eta_t \qquad (3)$$

where $\eta_t \sim iidN(0,1)$ and the changes in the variable are assumed to be measured over short enough sampling intervals (m) to avoid large discretisation bias[3]. Broze, Saillet and Zaköian (1995) discussed several statistical properties of discrete-time approximations of the CKLS model. In particular, they show that (3) is ergodic and second-order stationary if and only if $|1 + \alpha_1 m| < 1$ and $\gamma \in [0,1)$. When $\gamma = 1$, second-order stationarity is still possible, provided $(1 + \alpha_1 m)^2 < 1 - \sigma^2$; and ergodicity is guaranteed if $E(\ln|1 + \alpha_1 m + \sigma \eta_t|) < 0$. It should be noted that the properties of the discrete model are not exactly the same as those in the continuous time model $(m \to 0)$. For instance, strict stationarity depends on the parameters of the drift and volatility functions in the continuous model, so the process may exhibit so-called 'volatility-induced' stationarity in continuous time even for $\alpha_1 \geq 0$. Model (3) is typically estimated by (quasi) maximum likelihood (ML) based on the normality assumption, or by the general method of moments (GMM) under suitable conditions, although more sophisticated methods that do not rely on the Euler approximation have since been developed (see, *inter alia*, Aït-Sahalia, 2002 and references therein, and Kloeden and Platen, 1995).

Model (2) and its discrete-time approximation (3) were generalised in several ways to fit more complex dynamics in volatility, in the conditional mean, or both. A natural extension in the econometric framework lies in allowing the scale parameter, σ^2, to be time-varying following its own dynamics, resulting in what has been termed mixed models for interest rates. A number of processes have considered stochastic volatility patterns (Ball and Torous, 1995) and GARCH models that include dependence of unexpected shocks and cluster effects in volatility. Models under the latter approach are largely based on

---

[3] This approach is based on the heuristic argument that the approximation converges to the continuous-time model as $m \to 0$. Data in empirical analysis are mainly sampled on a daily and weekly basis but also on a monthly basis, even though the latter can imply large biases. Given the trade-off between discretisation bias in low-frequency samples and data problems related to market microstructure in high-frequency observations, weekly data is often considered as the best alternative.

GARCH(1,1)-type equations, such as,

$$\sigma_{t,m}^2 = \omega_0 + \omega_1 \xi_{t-1,m} + \omega_2 \sigma_{t-1,m}^2; \quad \omega_0 > 0; \ \omega_1, \omega_2 \geq 0 \tag{4}$$

$$\varepsilon_{t,m} = \Delta_m r_t - \alpha_0 m - \alpha_1 m r_{t-1,m};$$

where several alternatives for $\xi_{t-1,m}$ are common. Brenner, Harjes and Kroner (1996) specify,

$$\xi_{t-1,m} = \varepsilon_{t-1,m}^2 \tag{5}$$

which makes the scale dependent on both shocks and level through lagged values of the disturbance term. Note that the basic assumptions of GARCH processes do not apply to the driving errors $\varepsilon_{t,m}$, thus the resulting process is no longer an ordinary GARCH(1,1) model. Koedijk, Nissen, Schotman and Wolff (1997); Anderson and Lund (1997); and Bali (2003), among others, overcome this inconvenience by modelling 'standardised' innovations rather than the error term itself, that is,

$$\xi_{t-1,m} = \varepsilon_{t-1,m}^2 / r_{t-2,m}^{2\gamma} \tag{6}$$

so that the scale dynamics is unaffected by the level. There are several methodological advantages supporting formulation (6) over (5). It allows for a parsimonious and neat distinction between the clustering effect (related to the GARCH dynamics of the scale) and the level effect (related to the elasticity parameter). Moreover, GARCH parameters in the scale maintain their standard meaning and accept the common restrictions to ensure stationarity and existence of higher moments, unobserved in (5). Finally, the drift function may be augmented in order to include nonlinear effects owing to nonlinear or polynomial structures; see Aït-Sahalia (1996). The resulting models may be used together with the former volatility models originating heavily parametrized yet powerful specifications (see Bali, 2003; for recent estimation of such models). However, it should be noted that precise estimation under such specifications is difficult because of the strong degree of nonlinearity, hence ML procedures are likely to find problems in reaching a global optimum.

3 Testing for Unit Roots: A Monte Carlo Investigation

Little or nothing is formally known about the statistical properties concerning stationarity or volatility persistence of mixed models. We provide some insights on the behaviour of interest rate dynamics as a function of mixed models through Monte Carlo analysis in this section. We assess the behaviour of different statistical procedures when testing the null hypothesis of a unit root in time series that exhibit some form of level-dependent conditional heteroskedasticity. In a general setting, the basic framework for this analysis is based on a discrete-time model (subscript m defining the sample frequency is hereafter omitted for simplicity of notation) related to the CKLS specification as,

$$\Delta r_t = \mu + \alpha r_{t-1} + \varepsilon_t; \ \varepsilon_t = \sqrt{h_t}\eta_t; \quad \eta_t | I_{t-1} \sim Niid(0,1); \ t = 1, ..., T \tag{7}$$

where I_t denotes the information set up to time t and the sequence $\{h_t\}$ denotes the I_{t-1}−measurable conditional variance of the noise term such that $E\left(\varepsilon_t^2 | I_{t-1}\right) = h_t$. As

motivated in the previous section, we shall focus on two different sources driving h_t in order to conduct experimentation: a pure level-dependent process given by,

$$h_t = \sigma^2 r_{t-1}^{2\gamma}; \quad \sigma > 0, \ \gamma \geq 0 \tag{8}$$

and a mixed model with a scale parameter following a GARCH(1,1) process as in (6), i.e[4],

$$h_t = \sigma_t^2 r_{t-1}^{2\gamma}; \quad \gamma \geq 0 \tag{9}$$
$$\sigma_t^2 = \omega_0 + \omega_1 \left(\sigma_{t-1}\eta_{t-1}\right)^2 + \omega_2 \sigma_{t-1}^2; \ \omega_i > 0.$$

Remark 3.1: In considering that $(\mu, \alpha, \gamma) = \mathbf{0}$ and $\sigma_t^2 = \sigma^2$ in the above models, we obtain the homoskedastic driftless random walk considered under the DF tests null hypothesis. The series r_t is conditionally heteroskedastic when $\gamma > 0$, so that testing for unit roots by means of procedures derived under the iid assumption may result in power and size distortions.

Remark 3.2: The disturbance term related to (8) is not unconditionally second-order stationary when r_t is I(1). This occurs because the conditional variance of ε_t is defined as a scaled power of a lagged integrated process. However, since the dynamics of r_t are maintained, the unconditional expectation is not well defined. Note that this complicates the theoretical derivation of suitable limiting distributions, since second-order stationarity is normally required in order to apply the functional central limit theorem to show convergence to a functional of a Brownian motion. Although the DF test can be shown to convergence to a ratio of functionals of Lévy processes when innovations have infinite variance (*c.f.* Chan and Tran, 1989; Chan, 1990; Phillips, 1990; Ahn, Fotopoulos and He, 1990), this is generally based on strong assumptions regarding the error distribution. The reader is referred to Hansen (1992,1995) for a general analysis of parameter estimation under nonstationary variances.

Remark 3.3: Recall that r_t is not second-order stationary if $\gamma > 1$ regardless of the value of α. As such, the (unconditional) variance of this series may not be finite even when the true process has no unit roots. Moreover, the statistical properties of the model when the scale parameter σ_t^2 is time-varying are unknown.

Remark 3.4: Park (2002) has investigated several statistical properties of time series built as a function of an integrated process. These series are relevant because they can generate volatility clustering patterns, in other words, nonstationary nonlinear heteroskedasticity, that behave similar to nonstationary GARCH structures. Note that (8) cannot reproduce cluster effects in the conditional variance unless r_t is integrated. However, specification (9) works to provide such flexibility.

In obtaining the t-test for $H_0 : \alpha = 0$ and the likelihood ratio test for $H_0 : (\mu, \alpha) = \mathbf{0}$ (i.e., the standard DF unit root test statistics) we investigate the presence of an autoregressive unit root in r_t. The t-ratio test is also widely used for testing against nonstationarity after estimating the discrete version (or any econometric extension) of the CKLS model.

---

[4]Although both GARCH-type models are econometric extensions of CKLS and there are no theoretical reasons to prefer one model over the other, the approach based on (6) seems to have more popularity in empirical applications owing to the reasons commented in the text. Furthermore, Anderson and Lund (1997, p.350) reported that both alternatives seemed to be indistinguishable when fitting volatility dynamics to their series.

The next section will consider a univariate time series $\{r_t\}$ generated according to a conditionally Gaussian random walk without drift. In specific, we simulate random paths of model (7) with $(\mu, \alpha) = \mathbf{0}$ and innovations driven by either (8) or (9). Note that a random walk with a non-zero drift displays an upward/downward deterministic trend in the level of the series which is unrealistic in terms of accommodating interest rate dynamics[5]. Therefore, the additional constant term, μ, in the estimated models is primarily intended to fit the value to which the series reverts to under stationarity, rather than fitting the possible drift of a random walk process. We do not consider short-run dynamics in the DGP for the sake of simplicity, although the analysis including these patterns is straightforward.

In order to evaluate the performance of the test procedures considered, three Monte Carlo experiments were performed, which we denote as Experiments I, II and III.

3.1   Experiment I

In the first set of experiments, we explore the small sample size properties of the DF tests (t-test and LR test) when applied on time series with level dependent heteroskedasticity. The DGP in this analysis is defined as a random walk with a level-dependent volatility specification that considers (a) constant scale, σ^2, and (b) scale dynamics, σ_t following GARCH(1,1) dynamics. Since GARCH parameters cause some analytical complexities in the latter case, we focus on values raised by existing empirical work.

Recall that the basic DF testing procedure is based on estimating an auxiliary first-order autoregressive model such as,

$$\Delta r_t = \mu + \alpha r_{t-1} + u_t \tag{10}$$

with $u_t \sim iid\,(0, \sigma_u^2)$, $\sigma_u < \infty$. The relevant statistics (a t-test for the significance of α, and an F-test for the joint significance of (μ, α)) when testing against the null of integration are obtained from the ordinary least-square (OLS) estimates of (10). Hence, both estimation and inference stages of this procedure are performed under the assumption of an iid error term with finite, constant variance. As these assumptions are unfulfilled in the present context, this experiment looks to analyse the effect of such departures on the size and power performance of the DF tests.

Remark 3.5: Ling and Li (1997b) derived the distribution of the test based on the least-square estimates when the errors followed a covariance-stationary GARCH(1,1) process (this is the case considered in (b) when γ is set to zero). They show that the limiting distribution is identical to the one tabulated by Dickey and Fuller (1979). Thus, the DF test procedure remains valid under the considered heteroskedasticity since the distribution is asymptotically invariant. On the other hand, the limit distribution of the DF test when GARCH errors are not stationary (and hence divergence of the unconditional variance occurs) is as yet unknown.

Remark 3.6: Kim and Schmidt (1993) show through Monte Carlo simulation that integrated GARCH (IGARCH) errors drive the DF tests to moderate size departures even in large samples.

---

[5]The analysis based on a DGP which includes deterministic trends as well as a unit root, though of econometric interest, is beyond the scope of this paper, since it does not seem plausible for the present context.

Along with the DF tests, we also investigate the performance of a nonparametric test recently introduced by Breitung (2002), which we denote henceforth as NP. The NP test rejects the null of integration for small values of the variance ratio statistic,

$$NP = \frac{T^{-2} \sum_{t=1}^{T} \hat{U}_t^2}{\sum_{t=1}^{T} \hat{u}_t^2} \tag{11}$$

where $\hat{U}_t^2 = \sum_{i=1}^{t} \hat{u}_i^2$, and $\hat{u}_i = r_i - \bar{r}$ are the residuals from an OLS regression after adjusting for a non-zero mean. Critical values for this test at the confident levels of 10%, 5% and 1% for a sample of 500 observations are, respectively, 0.01473, 0.01046 and 0.00536 (see Breitung, 2002; for further details).

This procedure served firstly to provide robustness against misspecified short-run dynamics and nuisance parameters in the mean equation. Interestingly, preliminary analysis showed that this test is also robust against ARCH errors even under volatility integration or explosive patterns, a property not generally satisfied by DF tests. To illustrate the properties of the DF and NP tests, the empirical sizes of these tests against both IGARCH errors and ARCH errors (regarded as a function of the values of the ARCH coefficient, ω_1) are plotted in Figure 3.1. Here it is clear that the DF tests are fairly sensitive to ARCH effects, whereas the nonparametric test only undergoes moderate size distortions. Thus, as this test may feature better properties than the DF tests under level and mixed models it deserves closer investigation.

[**Insert Figure 3.1: IGARCH Effects**]

Our attention is on the small-sample properties of these tests thus, the analysis at this stage was based on 50,000 replications of samples of 500 observations. This, sample length is of empirical relevance for data sampled on a weekly or monthly basis (roughly corresponding to a decade of weekly observations)[6]. To minimise the effects of initial values on the findings, 900 observations were generated, with the first 400 having been removed from the simulated paths. The empirical size and the power of the tests are evaluated at the 1% and 5% nominal levels. To assess the power of the tests against stationary alternatives we concentrated on the near-integrated region, $\alpha = \{-0.01, -0.05\}$. In both cases the drift parameter was considered as $\mu = r_0 (1 + \alpha)$ so that the long-run expectation could take the same value. The value $r_0 = 0.15$ was arbitrarily chosen and assigned to initialise the simulated paths in all experiments. Finally, since the constant scale parameter, σ, has been empirically estimated to be approximately 1e-02, we restrict our interest to this range of values. All the runs of the Monte Carlo analysis were drawn with $\sigma^2 = 0.0003$.

In experiment (b), we also consider an unconditional scale parameter equal to $\sigma^2 = 0.0003$, where it follows that $\sigma^2 = E(\sigma_t^2) = \omega_0 (1 - \omega_1 - \omega_2)^{-1}$. With empirical evidence having provided quite different results when estimating GARCH parameters (see Experiment III below for further details). We began experimentation by taking a scale parameter that followed stationary GARCH dynamics given by i) a high degree of persistence together with low short-run effects ($\omega_1 = 0.05, \omega_2 = 0.90$), namely GARCH-L process, and ii) high persistence together with a relatively high ARCH effect ($\omega_1 = 0.20$,

---

[6]Simulations were also carried out on samples of length T=1000, with little differences with respect to the results observed from T=500. We avoid the presentation of the results regarding T=1000 in order to save space, but these are available from the authors upon request.

$\omega_2 = 0.70$), namely GARCH-H process, for comparative purposes. Note that empirical application of the CKLS models have found similar values (see, for instance, Koedijk *et al.*, 1997).

3.2 Experiment II

The second Monte Carlo analysis focuses on the *t*-test on the autoregressive parameter estimate and on the LR test obtained from ML estimation of the discrete approximation of the CKLS model. The GMM estimation proposed in CKLS is a natural alternative to the ML method, though orthogonality conditions require finite moments and hence could be inadequate if the time series considered is not weakly stationary. Here, we consider a DGP rendered by a pure random walk with driving errors following level-dependent volatility with constant scale. The analysis is based on the ML estimation of the following specification, *i.e.*,

$$\Delta r_t = \mu + \alpha r_{t-1} + \varepsilon_t; \ \varepsilon_t | I_{t-1} \sim N\left(0, \sigma^2 r_{t-1}^{2\gamma}\right). \tag{12}$$

The set of parameters to be estimated is $\theta = (\mu, \alpha, \gamma, \sigma)'$ and where the parameter space is constrained to avoid negative values of γ.

The unrestricted log-likelihood function, $\mathcal{L}(\theta_U)$, is then constructed by considering the prediction error decomposition,

$$\mathcal{L}(\theta_U) = \sum_{t=1}^{T} \ln f\left(\Delta r_t | I_{t-1}\right) \tag{13}$$

where $f(\cdot)$ corresponds to the Normal distribution density. The ML estimator becomes the optimum of the above function and the nonstationarity test drives from the left-sided *t*-test statistic on $\hat{\alpha}$, that is,

$$t_\alpha = \frac{\hat{\alpha}}{\sqrt{\hat{\omega}_\alpha}} \tag{14}$$

where $\hat{\omega}_\alpha$ is the corresponding diagonal element of the covariance matrix. Since the assumption of ε_t as normally distributed is unrealistic, the covariance matrix is typically estimated using a robust covariance estimator (referred to as the White method) from the QML procedure.

In addition, the null of a random walk with zero-drift may also be proved using a LR test after computing both the restricted and unrestricted log-likelihood functions. In this case, the appropriate statistic is,

$$LR = -2\left[\mathcal{L}\left(\hat{\theta}_R\right) - \mathcal{L}\left(\hat{\theta}_U\right)\right] \tag{15}$$

with $\mathcal{L}(\theta_R)$ denoting the log-likelihood function valued at the restricted estimated parameter vector.

Both the *t*- and the LR statistics are computed from ML estimation of model (8) based on 50,000 replications of samples of length T=500. We take γ as a nuisance parameter and carry out the simulations for values of this parameter in the range $[0, 1.5]$. Note that this is the usual range of values found in empirical work. As indicated earlier, the scale parameter is set as $\sigma^2 = 0.0003$. Since optimization of the likelihood functions in the Monte Carlo framework is computationally quite demanding, the optimisation routine sets off by taking the true value as an initial guess[7].

---

[7] Optimisation is carry out with the CML library of GAUSS 3.2.

Remark 3.7: Recall that the ML estimation method is more efficient than OLS when stationary ARMA-type models with conditional heteroskedastic innovations are involved. This property could extend to nonstationary time series, in which case unit root tests, based on ML estimates, can provide more powerful results. Hence, we assess the possible gains in power over OLS-type tests when the true process has no unit roots with $\alpha = \{-0.01, -0.05\}$ and $\mu = r_0 (1 + \alpha)$.

Remark 3.8: Note that the distribution of the t-statistic is scale-invariant in standard theory and in the unit root framework under errors with long-run finite variance. However, this property can be lost when errors bear infinite variance. We also analyse the sensitivity of the asymptotic distribution of these tests against different values of the scale parameter, σ^2. Taking the parameter of γ in the CIR model as a benchmark $(\gamma = 1/2)$, we consider an increasing sequence σ_j^2 of values for the scale parameter in which case $\sigma_j^2 = j\sigma^2$, $j = \{0.10, 1, 10, 100\}$ and $\sigma^2 = 0.0003$. The simulation takes 50,000 paths of length T=500. We also consider the effect of changing the value of σ on the OLS t-statistic from Experiment Ia)

3.3 Experiment III

Finally, the last experiment is concerned with the t-test and the LR test under the ML framework. In this case, the focus is on the econometric extension of the CKLS model which allows for the scale parameter to follow time-varying dynamics. The data generating process is a pure random walk with conditional variance following (9), and the model to be estimated is defined as,

$$\Delta r_t = \mu + \alpha r_{t-1} + \varepsilon_t; \ \varepsilon_t | I_{t-1} \sim N\left(0, \sigma_t^2 r_{t-1}^{2\gamma}\right) \tag{16}$$
$$\sigma_t^2 = \omega_0 + \omega_1 \left(\varepsilon_{t-1}/r_{t-2}^{\gamma}\right)^2 + \omega_2 \sigma_{t-1}^2$$

where the set of parameters to be estimated through ML is now given by $\theta = (\theta_1, \theta_2)'$, $\theta_1 = (\mu, \alpha)'$ and $\theta_2 = (\gamma, \omega_0, \omega_1, \omega_2)'$.

Remark 3.9: The distribution of the ML t-test with errors following a GARCH process (*i.e.*, by taking $\gamma = 0$ in the former equation) was analytically derived under restrictive assumptions in Seo (1999). Under the fourth-moment condition, the limiting distribution converges to a mixture of DF and standard normal distributions. The mixture coefficient is a nuisance term that depends on the values of the GARCH parameters and the fourth moment of the standardised errors. Ling, Li and McAleer (2001) propose several transformations of the t-statistic in order to obtain a test with a limiting distribution free of nuisance parameters.

Remark 3.10: The distribution of the likelihood ratio test is derived by Boswijk (2001) when innovations follow a near-integrated GARCH process. It is shown that there is little improvement over the OLS tests when volatility displays strong persistence in the presence of low short-term effects (high values of ω_2 together with low values of ω_1 in our framework), stylised features exhibited by most financial data.

GARCH modelling requires parameter restrictions to ensure nonnegative values ($\omega_0 > 0; \ \omega_1, \omega_2 \geq 0$) and frequently parametric space constraints through the restriction $\omega_1 + \omega_2 < 1$ to ensure second-order stationarity. It is worth noting that it is still possible to define meaningful stationary volatility paths under (5) even when using values beyond the unit boundary; recall that the common restrictions to ensure stationarity under GARCH

errors are meaningless for this model. However, under GARCH(1,1) dynamics this is no longer the case since the scale and hence the whole conditional variance, explodes fairly quickly. As it is unfeasible to report an exhaustive analysis of all possible combinations of parameters, we will focus on values found in the empirical literature. Empirical research presents different values for parameter estimates when considering series from different countries, different sample frequencies and different time periods. Still, we can observe some common features.

Firstly, the level effect measured through the elasticity coefficient is always found to be strongly reduced when mixed-level models are used. This is typically observed when $0 < \gamma < 1$ and often when the parameter estimate is not significantly different from that of the CIR model ($\gamma = 1/2$). Secondly, a number of empirical papers show that the scale parameter follows a stationary process, mainly characterised by low values for ω_1 and high values for ω_2 inside the stationary region (see, for instance, Koedijk *et al.*, 1997). Some authors have found parameter estimates to exceed far beyond the stationarity bound, when using daily and weekly data (see, for instance, Bali, 2003). As mentioned earlier, such an occurrence is extremely unappealing and non-intuitive, and suggestive that volatility follows an explosive process that ultimately diverges with time[8]. A possible explanation for this may lie in the fact that parameter stability is always assumed when modelling interest rates, thus yielding misspecification when samples span over several years (often decades) of observations. As the GARCH estimation under the QML procedure is sensitive to the correct specification of the model, results may be biased.

Therefore, since nonstationary values cannot be considered reliable, we consider values falling inside the stationarity region. We initially allow for a high degree of persistence in the scale dynamics through $\omega_1 = 0.05$ and $\omega_2 = 0.90$ (GARCH-L). As in Experiment I, we also treat the case of high persistence together with relatively high ARCH effects, assigned by $\omega_1 = 0.20$ and $\omega_2 = 0.70$ (GARCH-H). Given the great computational effort inherent in this experiment, we concentrate on a fixed parameter for elasticity, namely that obtained in the CIR model, *i.e.* $\gamma = 1/2$. The unconditional expectation of the scale is again set to $E\left(\sigma_t^2\right) = 0.0003$ and the analysis based on 25,000 simulations of paths of length $T = 500$.

4 Simulation Results

Whereas interest rates are positive time series, the discretised CKLS model and their econometric extensions are likely to take negative values under the unit root hypothesis[9]. Furthermore, the process with level-dependent volatility has an absorving barrier in zero that is attainable under the null. The implication is that the asymptotic distribution of the unit root test is degenerated because the underlying DGP collapses in zero, since the random walk has the property of almost surely visiting any value of the real line in asymptotic samples. Of course, the aim when testing for a unit root in real applications

---

[8]The values reported in Bali (2003) were inadequate to simulate interest rate paths in a model with a linear drift, given that the simulated series either diverged quickly or fell into the nonpositive region owing to the explosive volatility.

[9]The parametric restrictions ensuring non-negativity are not fulfilled under the unit root process. However, note that even if the parametric restrictions are preserved, the approximated solution of a SDE intended for a positive process may take negative values (see Schurz, 1996 for a discussion of this issue and an application of the CKLS model).

is tacitely that of checking out if the observed process behaves like a non-degenarate, positive random walk in period analyzed, which only is possible in finite series[10].

We therefore focus on finite-sampled bounded random walks on $(0, \infty)$ as DGPs for our experiment. There is growing interest in the literature towards bounded random walks (see Nicolau 2002; and the discussion therein for a recent application of these procedures). Simulations are easily obtained by discarding paths which do not fall into this range. The resulting series provide economical meaning *a-priori*, and the intuition behind our results is close to that tied to boostrap procedures in finite samples.

Finally, note that all simulations (and estimations) follow a discrete-time framework, that considers $m = 1$. We proceed in this way because the analysis of the bias induced by the SDE discretisation on the statistical properties of the tests involved is not the primary aim of this paper. Instead, a context isolated of such an effect is regarded so that we can analyse the ideal case and thus provide a reference benchmark. Of course, the power of the t-ratio in empirical applications related to Experiment II, in which we observe that the application of the Euler's scheme can lead to large estimate bias, is expected to be no greater than that commented below. The analysis of the specific effect related to the discretisation bias is an interesting topic and is currently being investigated by the authors.

4.1   Experiment I

Our findings relating to empirical size from Experiment Ia) for the nominal levels of 1% and 5% are summarised in Table 4.1. Small deviations from the nominal sizes arise as a function of γ in the case of the parametric tests, while the nonparametric procedure appears quite robust against this form of heteroskedasticity. As the value of γ increases, however, the distribution tends rapidly to the DF distribution. The underlying distribution seems to be therefore the DF distributions.

[**Insert Table 4.1**: **DF Size**]

The effect on power after adjusting for size can be summarised as follows (see Figures 4.1 and 4.2). Firstly, the parametric tests suffer from distortions depending on the value of the eslasticiy parameter; the case of maximum power for the parametric test is observed for $\gamma = 0$,(corresponding to the context where the model is correctly specified). The LR test is widely known to be less powerful than the t-test due to its two-sided nature, and it is also shown here to be more sensitive to level effects than the t-ratio. Both tests are indeed sensitive to γ, particularly as the root reaches towards unity, though interestingly, the impact on the DF test weakens as the significance level decreases. Hence, researchers who tend to avoid large significant levels when applying the DF test due to the characteristic low power of this procedure should note in this another reason for doing so when testing the properties of interest rate time series. The power decreases in the case of the parametric DF tests, and we show that these procedures have low ability to identify the mean-reversion behaviour of the series for several values of γ in a near-integration context.

---

[10]In other words, testing the null of a unit root seems to make sense for (arbitrarily long yet) finite series, where the random walk is able to show a non-degenarate, positive behaviour, as that exhibited in observed series.

[**Insert Figures 4.1 and 4.2**: DF Power]

On the other hand, the nonparametric test proposed by Breitung (2002) certainly displays robustness against this type of volatility. This test is generally not more powerful than the DF t-test when the root is relatively far from unity, it may indeed display better performance in the near-integrated region given it is less sensitive than the DF test. Therefore, this test may be the better option when one suspects of the presence of this type of volatility, than the Dickey-Fuller test. Experiment Ib) highlights the results discussed earlier. The effect on the unit root tests when allowing a time-varying scale with stationary GARCH dynamics does not seem significantly different from that under constant scale, and the empirical percentiles (in the lower tail) are very similar. Given the strong similarity, we only present results for the empirical size on the 5% nominal size (see Table 4.2) for the sake of brevity and space, but remaining results are available from the authors upon request.

[**Insert Table 4.2: Size Mixed Model**]

The power results are essentially the same as those obtained with a pure level model for all the tests here considered. Although there are slight differences for small values of γ, the behaviour is roughly the same for high values of the same parameter. Figure 4.3 shows the effect on the DF t-test in the near-integrated region $\alpha = -0.01$ and nominal size of 5%, exhibiting predominance of the level effect under the hypothesis of no mean-reverting behaviour.

[**Insert Figure 4.3: Power Mixed Effects**]

4.2 Experiment II

In Tables 4.3 and 4.4 we report the simulated critical values for the test statistics from ML estimations. It is clear that the distribution of the t-test and the LR test gravitate around the Dickey-Fuller distributions. Moreover, as γ increases the t-statistic is progressively attracted to the same limiting distribution as the standard DF test, although some slight deviations from these distributions can be observed when γ takes values under $1/2$. The empirical distribution of these tests seems to be (at least) influenced in small samples by the value of the elasticity parameter γ, although remarkably the DF distribution of the t-ratio indicates good critical values for the empirical range considered (see Figure 4.4).

[**Insert Tables 4.3 and 4.4 and Figure 4.4: ML Simulations**]

In terms of the scale parameter effect on the shape of the limit distribution, it appears that the tails of the t-test are affected when this parameter takes relatively high values, so the scale-invariance of the ordinary t-stastistic is indeed lost when considering these types of stochastic processes (see Table 4.5). The differences seem to be of little importance when the scale parameter is small or moves into the relevant range for empirical applications. However, it becomes evident that high parameter values shift the distribution to the left, so the probability of taking parameter estimates of α greater than zero increases, given the value[11] of γ. Interestingly, this phenomenon is also present

---

[11]We repited the simulations for several values of γ with no qualitatively differences from results reported in the main text.

when the DGP is bounded, though the magnitude of the shift is found to be smaller and concentrated in the upper tail. Of course, the scale-invariance property of the OLS t-statistc in Experiment I is also lost, and the limiting distributions depend on the value of the scale. Again, the higher the value of that parameter, the higher the departure.

[Insert Table 4.5: Scale Invariance]

With respect to the power properties of the ML tests, our analysis shows that parameter γ still conditions the ability of these tests to reject the null. Table 4.6 presents the power results of the ML when using the empirical critical values from the above simulations. It can be observed that the power of the ML tests is nearly halved when the elasticity parameter is different from zero and the root is near-integrated $\alpha = -0.01$. The power of the t-test is higher than that of the LR statistic and in general ML based tests exhibit higher power than those based on the OLS method from Experiment Ia). Interestingly, the power of the ML and OLS based procedures tends to be equal for values $\gamma > 1$. It is seen that the critical values offered by the DF distributions are precise for most of the values of γ in the range analyzed, although it should be remarked that empirical sizes are slightly higher than the respective nominal sizes when γ takes low values.

[Insert Tables 4.6: Power ML]

4.3   Experiment III

The general results stated in the previous section can be extended to Experiment III, in which case, the scale parameter follows a stationary yet highly-persistent GARCH dynamics. The search for the optimum is considerably more complex than in the level case as a result of tolerating a high degree variance nonlinearity (despite the offset of the optimisation routine being the true solution), though convergence is reached in all cases.

The GARCH parameters governing the scale dynamics are clearly nuisance parameters and therefore make the limiting distribution differ from the one with constant scale (see Figure 4.8). For example, when $\gamma = 1/2$, the percentiles (1%, 5%, 10%) and [90%, 95%, 99%] of the empirical distributions are (-3.33, -2.77, -2.50) and [0.04,0.40,0.90] respectively, when GARCH dynamics includes low short-run effects. However, the corresponding percentiles are (-3.33,-2.76,-2.44) and [0.25, 0.58, 1] when higher short-run dynamics is allowed for. Recall that the computed critical values are (-3.44, -2.89, -2.60) and [-0.11, 0.23, 0.61], respectively, when only level-dependence drives the conditional variance, so the GARCH dynamics push the distributions to the left. Note that this result is in the same line that the main results evidenced by Seo (1999). The limiting distribution of the LR statistic is also seen to depend on the particular GARCH dynamics [12].

[Insert Figure 4.8]

Thus, it is important to note that the limiting distribution of the t-test obtained in the estimation of generalisations of the CKLS model for a particular value of γ is far from normal and, in fact, is much closer to the DF distribution. Nonetheless, the normal distribution is (implicitly) used as the limit distribution when conducting inference in

---

[12]We avoid the presentation of these statistics for the sake of space, although these are available upon request.

such models, (though, no explicit reference to the distribution of the test statistic used is made). Hence, conclusions regarding mean-reversion in this non-formal manner are dubious and potentially misleading (see, Bali, 2003 among others, for an example of this ambiguity[13]). It should be noted that the normal distribution would only make sense if the true model were intended as a unit root process with a non-zero drift, which is not a realistic premise for interest rates. Yet, even in such a case, it has been shown in the conventional unit root literature that the limit distribution could differ from the Gaussian distribution in samples of moderate size and small drift; see Hylleberg and Mizon (1989) and Haldrup and Hylleberg (1995).

5   Conclusion

In this paper we provide several insights into the behaviour of the more popular unit root test procedures, commonly applied in the empirical modelling of risk-free, short-term interest rates. Unit root dynamics do not seem to be a proper model for interest rates, though testing against this type of behaviour is common in literature related to the topic, either as part of a previous descriptive analysis or as part of the diagnostic analysis on the estimates from the parametric model involved.

As a first approach, we assessed the performance of the well-known DF tests as well as a recently introduced nonparametric procedure proposed by Breitung (2002, 2003) when random innovations are level-dependent. The effect of GARCH dynamics and other nonlinear structures, like regime switching models, have been analysed in the literature related to unit root tests, however to level-dependent conditional variances little attention has been given. This type of process is better suited in the framework of interest rates and, hence, raises an interesting issue of empirical relevance. Our analysis represents an empirical extension of the basic homoskedastic framework of unit root testing procedures towards this form of heteroskedasticity. We observe that the statistical properties of the DF procedures are sensitive to a particular value of γ. This form of heteroskedasticity leads to high losses of power in samples of moderate size, particularly when a high significance level is used in a near-integrated context.

On the other hand, the nonparametric procedure analysed reveals to be quite robust against this class of heteroskedasticity and yields better results in contexts where the performance of parametric tests fails. Given that the strong persistence of the series is capable of generating a quasi-integrated context, the exact functional form of conditional volatility is unknown and can only be estimated. Procedures that provide robustness against conditional volatility and even against the presence of a second moment may constitute a better alternative than parametric tests. Approaches such as the recently proposed procedure by Luger (2003) could in this context prove useful, and clearly deserve

---

[13]The procedure used by Bali (2003) is as follows: firstly, the DF t-test is performed under the null of a random walk without drift, based on an auxiliary regression with a constant. Evidence against the unit root is found using a 10% significance level. Then, the CKLS model (among others) is estimated according to the discrete approximation of the continuous process, and subsequently parameter estimates and their corresponding asymptotic t-statistics presented. It is found that "*The maximum likelihood estimates of the linear drift model imply mean reversion of the spot rate since α_1* [the autoregressive parameter] *is found to be negative and highly significant*", pp. 211. Note that the t-statistic for the level model (see Table 4, p.213) takes the value -1.48 with $\hat{\gamma} = 1.57$, as a result this parameter is not very significant at any of the conventional confident levels when the true empirical distribution is used.

further investigation.

Next, given the importance of modelling and testing for the interest rate literature, we analysed the shape of the limiting distributions of both the t-test and the LR test obtained from ML estimation of discretely observed diffusion processes. We considered a general nonlinear CKLS diffusion model and an econometric extension that allows the scale parameter to evolve according to stationary GARCH dynamics. The limiting distributions of the relevant statistics are thampered by nuisance parameters and depend (at least) on the particular value of the elasticity parameter, particularly when γ takes low values. Furthermore, the t-statistic is not scale-invariant, as is generally the case when there is a well-defined long-term variance. The power of tests based on ML estimation shows to be sensitive to the particular value of γ, although not to the same extent as observed in OLS procedures. It was also found that the limiting distribution is also sensitive to a particular value of the driving parameters when the scale follows time-varying dynamics. It is likewise the case in the standard context in which the variance does not depend on the level, since the distribution switchs to the left.

In conclusion, the t-test for estimates taken from the discretely observed CKLS model or its econometric extensions, require finding evidence against the unit root hypothesis based on distributions free of nuisance parameters; alternatively the limiting distribution for each particular realization can be obtained by means of bootstrapping procedures. Nevertheless, it should be noted that the DF distributions provide precise critical values for most of the relevant values of γ and σ which are found in practice. When the scale is allowed to follow time-varying dynamics, the critical values from the DF distribution represent only approximated values.

The findings of this paper emphasise not only the need for closer analytical investigation of proposed procedures but further introduction of new parametric tests based on ML estimation with better properties than those presently available. Also this study importantly touched on the particularities connected to the difficulty in deriving asymptotic distributions. The highly nonlinear structure displayed by the random innovations are not treated in the standard theory of random walk processes. The interest rate series exhibit conditional heteroskedasticity as well as infinite-variance errors under the null, introducing yet more intricacies to our investigation. Although econometric literature has already dealt with finite-variance errors in the unit root framework under the assumption of errors drawn from a α-stable distribution, the general cases here discussed remain a challenging problem.

References

[1] Ahn. S.K.; Fotopoulos, S. B.; He L.; (1990). Unit Root Tests with Infinite Variance Errors. Econometric Reviews, 20(4), pp. 461-483.

[2] Aït-Sahalia, Y.; (1996). Testing Continuous Time Models of the Spot Interest Rate. Review of Financial Studies, 9, pp. 385-426.

[3] Aït-Sahalia, Y.; (2002). Maximum Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approach, Econometrica 70, pp.223-262.

[4] Anderson, T.G.; Lund, J.; (1997). Estimating Constinous-time Stochastic Volatility Models of the Short-term Interest Rates, Journal of Econometrics, 77, pp. 343-377.

[5] Bali, T.G.; (2003). Modeling the Stochastic Behavior of Short-Term Interest Rates: Pricing Implications for Discount Bonds. Journal of Banking & Finance, 27, pp. 201-228.

[6] Ball, C.A.; Torous, W.N.; (1999). The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence. Journal of Finances, 54(6), pp. 2339-2359.

[7] Breitung, J.; (2002). Nonparametric Tests for Unit Roots and Cointegration. Journal of Econometrics, 108, pp. 343-363.

[8] Boswijk, H.P.; (2001). Testing for a Unit Root with Near-Integrated Volatility. Timberger Institute Discussion Paper, TI 2001-077/4.

[9] Brenner, R.J.; Harjes, R.H.; Kroner, K.F.; (1996). Another Look at Models of the Short-Term Interest Rate. Journal of Finance and Quantitative Analysis, 31(1), pp. 85-107.

[10] Broze, L.; Scaillet, O.; Zaköian, J.M.; (1995). Testing for Continuous-Time Models of the Short-Term Interest Rate. Journal of Empirical Finance, 2, pp. 199-223.

[11] Chan, K.S.; (1990). Inference for Near-Integrated time series with Infinite Variance. Journal of the American Statistical Association, 85, pp. 1069-1074.

[12] Chan, K.C.; Karolyi, G.A.; Longstaff, L.A.; Sanders, S.A.; (1992). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. Journal of Finance, 47, pp. 1209-1227.

[13] Chan, K.S.; Tran, L.T.; (1989). On the First-order Autoregressive Processes with Infinite Variance. Econometric Theory, 5, pp. 354-362.

[14] Chapman, D.; Pearson, N.; (2000). Is the Short Rate Actually Non-linear? Journal of Finance, 55, pp. 355-388.

[15] Chapman, D.; Pearson, N.; (2001). Recent Advances in Estimating Term-Structure Models, Financial Analyst Journal, July/August, pp. 77-95.

[16] Cox, J.; Ingersoll, J.; Ross, S.; (1985). A Theory of the Term Structure of Interest Rates. Econometrica, 53, pp. 385-407.

[17] Dickey, D.A.; Fuller, W.A.; (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association, 74, pp. 427-431.

[18] Dickey, D.A.; Fuller, W.A.; (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. Econometrica, 49, pp. 1057-1079.

[19] Haldrup, N.; Hylleberg, S.; (1995). A Note on the Least Squares Estimator of a Random Walk with Drift,- Some Analytical Evidence. Economics Letters, 48, pp. 221-228.

[20] Hansen, B.A.; (1992). Heteroskedastic Cointegration. Journal of Econometrics, 54, pp. 139-158.

[21] Hansen, B.A.; (1995). Regression with Non-stationary Volatility. Econometrica, 63, pp. 1113-1123.

[22] Hylleberg, S.; Mizon, G. E.; (1989). A Note on the Distribution of the Least Squares Estimator of a Random Walk with Drift. Economics Letters, 29, pp. 225-230.

[23] Hong, Y.; Li, H.; (2002). Nonparametric Specification Testing for Continous-Time Models with Application to Spot Interest Rates, unpublished manuscript, Cornell University.

[24] Kim, K.; Schmidt, P.; (1997). Unit Root Tests with Conditional Heteroskedasticity. Journal of Econometrics, 59, pp. 287-300.

[25] Koedijk, K.G.; Nissen, F.G.J.; Schotman, P.C.; Wolff, C.; (1997). The Dynamics of Short-Term Interest Rate Volatility Reconsidered. European Finance Review, 1, pp. 105-130.

[26] Kloeden, P.E.; Platen, E.; (1995). Numerical Solution of Stochastic Differential Equations. Springer.

[27] Li, W.K.; Ling, S.; McAleer, M.; (2002). Recent Theoretical Results for Time Series Models with GARCH Errors. Journal of Economic Surveys, 16(3), pp.245-269.

[28] Ling, S.; Li, W.K.; (1997b). Estimating and Testing for Unit Root Processes with GARCH(1,1) Errors. Technical Report, Department of Statistics, The University of Hong Kong.

[29] Li, W.K.; Ling, S.; McAleer, M.; (2001). Estimating and Testing for Unit Root Processes with GARCH(1,1) Errors: Theory and Monte Carlo Evidence. Technical Report, Department of Statistics, The University of Hong Kong.

[30] Litterman, R.; Scheinkman, J.; (1991). Common Factors Affecting Bond Returns. Journal of Fixed Income, 3(3), pp. 54-61.

[31] Luger, R.; (2003). Exact Non-Parametric Tests for a Random Walk with Unknown drift under Conditional Heteroskedasticity. Journal of Econometrics, 115, pp. 259-276.

[32] Nicolau, J.; (2002). Stationary Processes that look like Random Walks -The Bonded Random Walk in Discrete and Continuous Time. Econometric Theory, 18, pp. 99-118.

[33] Pantula, S.G.; (1988). Estimation of Autoregressive Models with ARCH Errors, Sankhya B, 50, pp.119-138.

[34] Park, J.Y.; (2002). Nonstationary Nonlinear Heteroskedasticity. Journal of Econometrics, 110, pp. 383-415.

[35] Phillips, P.C.B.; (1990). Time Series Regression with a Unit Root and Infinite Variance Errors. Econometric Theory, 6, pp. 44-62.

[36] Rogers, L.C.G.; Stummer, W.; (2000). Consistent Fitting of One-Factor Models to Interest Rate Data. Insurance: Mathematics and Finance, 27, pp. 45-63.

[37] Schurz, H.; (1996). Numerical Regularization for SDEs: Construction of Nonnegative Solutions. Dyn. Syst. Appl. 5 (1), pp. 323-352.

[38] Seo, B.; (1999). Distribution Theory for Unit Root Tests with Conditional Heteroskedasticity. Jounal of Econometrics, 91, pp. 113-144.

[39] Vasicek, O.; (1977); An Equilibrium Characterization of the Term Structure, Journal of Financal Economics, 5, pp. 177-188.

Tables

Table 4.1: Empirical size of unit root tests against stationarity at 1% and 5% nominal level. The DGP is $\Delta r_t = \sigma r_{t-1}^{\gamma} \eta_t, \quad \eta_t | I_{t-1} \sim iidN(0,1)$ and the test regression used is $\Delta r_t = \mu + \alpha r_{t-1} + u_t$, where u_t is regarded as an iid process. Empirical size is computed using 50,000 replications of samples of 500 observations for the i) Dickey-Fuller t-test (DF-T) on $H_0 : \alpha = 0$, ii) Dickey-Fuller likelihood ratio (DF-LR) on $H_0 : (\mu, \alpha) = \mathbf{0}$, and iii) Non-parametric test of Beitrung (2002) (NP) on $H_0 : \alpha = 0$.

	1%				5%		
γ	DF-T	DF-LR	NP	γ	DF-T	DF-LR	NP
0.00	0.013	0.012	0.011	0.00	0.061	0.056	0.066
0.05	0.012	0.010	0.011	0.05	0.059	0.052	0.067
0.10	0.012	0.010	0.011	0.10	0.057	0.051	0.067
0.25	0.011	0.009	0.011	0.25	0.054	0.046	0.066
0.40	0.013	0.011	0.012	0.40	0.055	0.047	0.064
0.50	0.016	0.017	0.010	0.50	0.062	0.059	0.059
0.55	0.018	0.021	0.010	0.55	0.065	0.064	0.055
0.60	0.022	0.027	0.009	0.60	0.070	0.075	0.052
0.65	0.025	0.031	0.009	0.65	0.071	0.078	0.050
0.70	0.025	0.035	0.009	0.70	0.075	0.084	0.052
0.75	0.021	0.030	0.009	0.75	0.069	0.079	0.053
0.80	0.019	0.027	0.009	0.80	0.066	0.075	0.053
0.95	0.016	0.022	0.009	0.95	0.061	0.068	0.053
1.00	0.015	0.020	0.009	1.00	0.060	0.066	0.053
1.10	0.014	0.018	0.009	1.10	0.057	0.063	0.053
1.25	0.013	0.016	0.009	1.25	0.054	0.059	0.053
1.50	0.011	0.013	0.009	1.50	0.052	0.054	0.052
1.75	0.011	0.011	0.085	1.75	0.050	0.052	0.051
2.00	0.010	0.011	0.089	2.00	0.049	0.050	0.051

Table 4.2: Empirical size of unit root tests against stationarity at a 5% nominal level. The DGP is $\Delta r_t = \sigma_t r_{t-1}^\gamma \eta_t, \quad \eta_t | I_{t-1} \sim N(0,1)$, with $\sigma_t^2 = \omega_0 + \omega_1 \left(\eta_{t-1}\sigma_{t-1}\right)^2 + \omega_2 \sigma_{t-1}^2, \sigma_t^2 = 0.0003 = \omega_0 (1 - \omega_1 - \omega_2)^{-1}$ and either $(\omega_1, \omega_2) = (0.05, 0.90), (GARCH - L)$, or $(\omega_1, \omega_2) = (0.20, 0.70), (GARCH - H)$. The estimated model is $\Delta r_t = \mu + \alpha r_{t-1} + u_t$, where u_t is regarded as an iid process. Empirical size is computed using 50,000 replications of samples of 500 observations for the i) Dickey-Fuller t-test (DF-T) on $H_0 : \alpha = 0$, ii) Dickey-Fuller likelihood ratio (DF-LR) on $H_0 : (\mu, \alpha) = \mathbf{0}$, and iii) Non-parametric test of Beitrung (2002) (NP) on $H_0 : \alpha = 0$.

	GARHC-L				GARCH-H		
γ	DF-T	DF-LR	NP	γ	DF-T	DF-LR	NP
0.00	0.063	0.059	0.066	0.00	0.063	0.059	0.066
0.05	0.061	0.055	0.066	0.05	0.068	0.063	0.065
0.10	0.063	0.055	0.065	0.10	0.067	0.061	0.066
0.25	0.056	0.048	0.065	0.25	0.064	0.054	0.064
0.40	0.059	0.051	0.063	0.40	0.064	0.056	0.062
0.50	0.063	0.061	0.058	0.50	0.070	0.064	0.059
0.55	0.068	0.068	0.054	0.55	0.075	0.072	0.054
0.60	0.072	0.075	0.053	0.60	0.077	0.080	0.051
0.65	0.075	0.083	0.052	0.65	0.079	0.087	0.050
0.70	0.074	0.083	0.051	0.70	0.080	0.088	0.051
0.75	0.072	0.081	0.052	0.75	0.076	0.085	0.049
0.80	0.070	0.077	0.052	0.80	0.074	0.082	0.051
0.95	0.063	0.070	0.052	0.95	0.069	0.074	0.052
1.00	0.062	0.068	0.052	1.00	0.067	0.072	0.052
1.10	0.060	0.065	0.052	1.10	0.065	0.069	0.052
1.25	0.058	0.061	0.053	1.25	0.054	0.057	0.053
1.50	0.055	0.057	0.053	1.50	0.062	0.066	0.053
1.75	0.053	0.055	0.052	1.75	0.059	0.059	0.052
2.00	0.052	0.053	0.051	2.00	0.059	0.058	0.052

Table 4.3: Empirical percentiles for the t-test of the null hypothesis $H_0 : \alpha = 0$ in the model $\Delta r_t = \mu + \alpha r_{t-1} + \sigma r_{t-1}^{\gamma} \eta_t$; $\eta_t | I_{t-1} \sim iidN(0,1)$, where parameters are estimated based on the ML method. The DGP takes $(\mu, \alpha) = \mathbf{0}$. DF-T represents the percentiles of the t-test DF distribution (T=500). Percentiles are computed based on 50,000 replications.

	Probability of a smaller value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
DF-T	-3.44	-3.13	-2.87	-2.57	-0.43	-0.07	0.24	0.63
γ								
0.00	-3.52	-3.20	-2.94	-2.65	-0.60	-0.25	0.06	0.42
0.10	-3.48	-3.17	-2.91	-2.62	-0.55	-0.20	0.13	0.51
0.25	-3.48	-3.17	-2.91	-2.66	-0.56	-0.19	0.14	0.54
0.40	-3.44	-3.14	-2.89	-2.60	-0.50	-0.12	0.23	0.63
0.50	-3.45	-3.13	-2.87	-2.54	-0.50	-0.12	0.22	0.64
0.65	-3.44	-3.13	-2.86	-2.56	-0.41	-0.04	0.29	0.68
0.75	-3.44	-3.12	-2.87	-2.57	-0.46	-0.08	0.25	0.66
0.90	-3.43	-3.12	-2.86	-2.57	-0.44	-0.07	0.27	0.68
0.95	-3.45	-3.13	-2.87	-2.57	-0.42	-0.04	0.29	0.68
1.00	-3.44	-3.13	-2.87	-2.57	-0.41	-0.04	0.30	0.68
1.25	-3.43	-3.12	-2.85	-2.56	-0.41	-0.03	0.29	0.68
1.50	-3.40	-3.10	-2.84	-2.54	-0.40	-0.03	0.29	0.67

Table 4.4: Empirical percentiles of the LR test of the null hypothesis $H_0 : (\mu, \alpha) = \mathbf{0}$ in the model $\Delta r_t = \mu + \alpha r_{t-1} + \sigma r_{t-1}^{\gamma} \eta_t$; $\eta_t | I_{t-1} \sim iidN(0,1)$, where parameters are estimated by ML in both the restricted and unrestricted models. The DGP takes $(\mu, \alpha) = \mathbf{0}$. DF-LR represents the percentiles of the LR DF distribution (T=500). Percentiles are computed based on 50,000 replications.

	Probability of a smaller value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
DF-LR	*0.60*	*0.78*	*1.02*	*1.34*	*7.58*	*9.22*	*10.82*	*12.94*
γ								
0.00	0.64	0.87	1.13	1.50	7.90	9.52	11.09	13.14
0.10	0.65	0.89	1.15	1.52	7.78	9.42	10.96	13.04
0.25	0.58	0.86	1.17	1.59	7.80	9.48	11.03	13.09
0.40	0.48	0.74	1.06	1.50	7.85	9.42	10.91	12.90
0.50	0.42	0.67	0.97	1.40	7.65	9.24	10.79	12.83
0.65	0.41	0.67	0.94	1.33	7.52	9.11	10.67	12.65
0.75	0.44	0.69	0.95	1.32	7.48	9.06	10.57	12.67
0.90	0.46	0.69	0.95	1.31	7.53	9.12	10.66	12.57
0.95	0.47	0.70	0.95	1.31	7.55	9.15	10.69	12.72
1.00	0.47	0.70	0.95	1.31	7.55	9.16	10.69	12.71
1.25	0.51	0.72	0.96	1.31	7.53	9.16	10.69	12.78
1.50	0.54	0.74	0.98	1.31	7.48	9.10	10.62	12.70

Table 4.5 : Empirical percentiles for the t-test of the null hypothesis $H_0 : \alpha = 0$ in the model $\Delta r_t = \mu + \alpha r_{t-1} + \sigma r_{t-1}^{\gamma} \eta_t$; $\eta_t | I_{t-1} \sim iidN(0,1)$, where parameters are estimated by ML. The DGP takes $\gamma = 1/2$, $(\mu, \alpha) = \mathbf{0}$ and the values of the parameter scale, σ^2, presented in the first column. DF-T denote the percentile for the t and LR test obtained from the respective DF distributions (T=500). Percentiles are computed based on 50,000 replications.

	Probability of a smaller value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
DF-T	-3.43	-3.12	-2.86	-2.57	-0.44	-0.07	0.23	0.60
σ^2								
0.00003	-3.42	-3.11	-2.84	-2.54	-0.42	-0.06	0.26	0.64
0.0003	-3.45	-3.13	-2.87	-2.54	-0.50	-0.12	0.22	0.64
0.003	-3.45	-3.13	-2.87	-2.54	-0.40	-0.05	0.30	0.74
0.03	-3.42	-3.11	-2.86	-2.56	-0.43	-0.03	0.35	0.77

|	Probability of a smaller value							
	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99
DF-LR	0.60	0.78	1.02	1.34	7.58	9.22	10.82	12.94
σ^2								
0.00003	0.57	0.77	1.00	1.32	7.44	9.05	10.63	12.66
0.0003	0.42	0.67	0.97	1.40	7.65	9.24	10.79	12.83
0.003	0.38	0.65	1.02	1.57	8.26	9.85	11.44	13.58
0.03	0.38	0.65	1.02	1.60	8.32	9.95	11.50	13.55

Table 4.6: Power of ML t-test tests at 1% and 5% nominal levels. The estimated model is $\Delta r_t = \mu + \alpha r_{t-1} + \sigma r_{t-1}^{\gamma}\eta_t, \ \eta_t|I_{t-1} \sim iidN(0,1)$, where the DGP considers mean-reverting paths with $\alpha = \{-0.01, -0.05\}$. Power is computed based on 50,000 replications of samples of 500 observations for the i) t-test on $H_0 : \alpha = 0$, and ii) LR test on $H_0 : (\mu, \alpha) = \mathbf{0}$. The value of μ is set as $\mu_0(1 + \alpha)$.

| | $\alpha = -0.01$ | | | | $\alpha = -0.05$ | | | |
| | 1% | | 5% | | 1% | | 5% | |
γ	t-test	LR	t-test	LR	t-test	LR	t-test	LR
0.00	0.053	0.039	0.210	0.159	0.813	0.740	0.987	0.970
0.10	0.049	0.034	0.189	0.138	0.743	0.650	0.971	0.942
0.25	0.030	0.022	0.138	0.105	0.727	0.645	0.965	0.935
0.40	0.028	0.020	0.120	0.092	0.740	0.660	0.964	0.933
0.50	0.026	0.020	0.122	0.098	0.728	0.661	0.965	0.940
0.65	0.026	0.020	0.119	0.099	0.728	0.672	0.963	0.942
0.75	0.026	0.020	0.118	0.100	0.726	0.666	0.963	0.942
0.90	0.026	0.020	0.120	0.095	0.730	0.670	0.963	0.936
0.95	0.024	0.019	0.118	0.093	0.718	0.652	0.964	0.934
1.00	0.025	0.019	0.117	0.090	0.723	0.652	0.962	0.933
1.25	0.024	0.017	0.120	0.087	0.726	0.637	0.965	0.930
1.50	0.024	0.017	0.118	0.085	0.740	0.638	0.967	0.930

Figures

Figure 3.1:

Empirical sizes of unit root tests when errors follow IGARCH and explosive ARCH patterns.
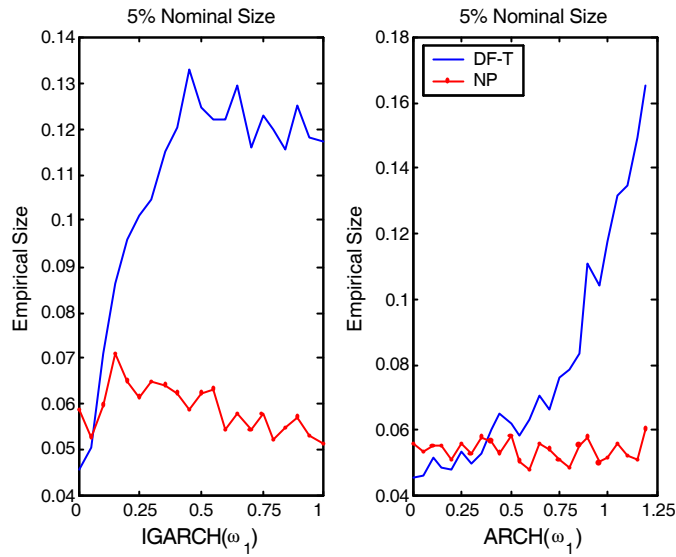
Empirical size of the DF t-test (DF-T) and Breitung's nonparametric test (NP) when errors follow an IGARCH process, $h_t = \omega_0 + \omega_1 \varepsilon_{t-1}^2 + (1 - \omega_1) h_{t-1}, (0 \leq \omega_1 \leq 1)$; and ARCH patterns ($h_t = \omega_0 + \omega_1 \varepsilon_{t-1}^2$), $(0 \leq \omega_1 \leq 1.20)$
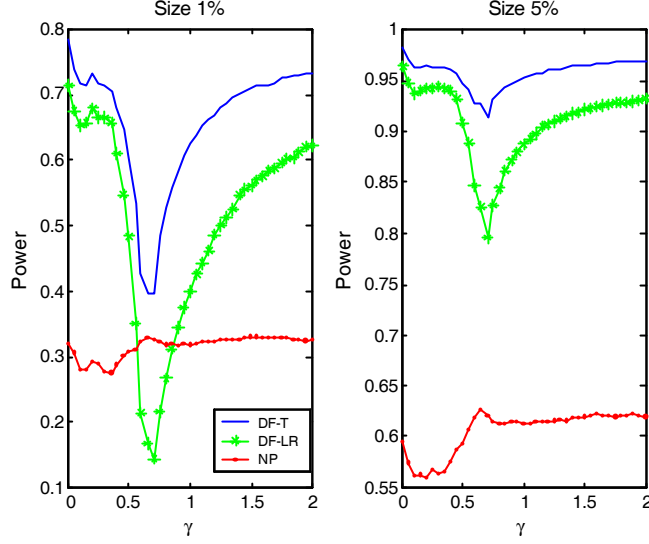
Figure 4.1:
Power of unit root tests under different values of $\gamma, \alpha = -0.01$.
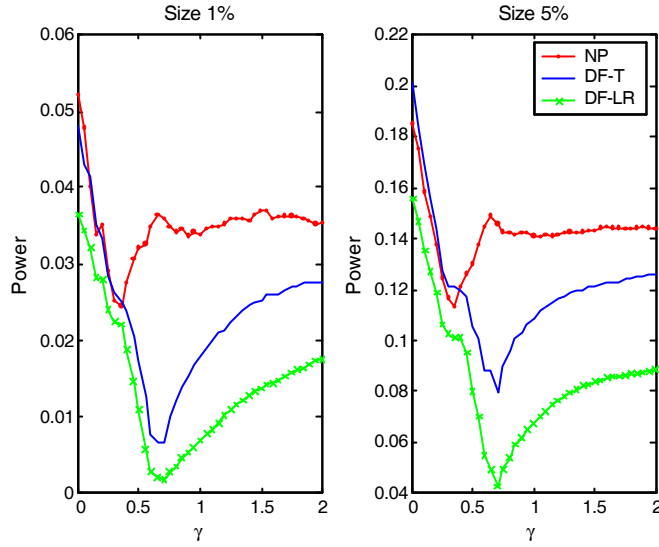
Power of unit root tests against stationarity at nominal sizes of 1% and 5%. The DGP is given by $\Delta r_t = \mu + \alpha r_{t-1} + \varepsilon_t, \varepsilon_t | I_{t-1} \sim N\left(0, \sigma^2 r_{t-1}^{2\gamma}\right)$ with $(\mu, \alpha) = (0.0015, -0.01)$. Only positive realization of the DGP are considered in order to compute size and power. The estimated model is $\Delta r_t = \mu + \beta r_{t-1} + u_t$ where the noise is regarded as an iid process. Power (corrected by empirical size) is computed for the i)Dickey-Fuller t-test (DF-T), $H_0 : \alpha = 0$; ii) Dickey-Fuller likelihood-ratio test (DF-LR), $H_0 : (\mu, \alpha) = 0$; and iii) Breitung's nonparametric tests (NP), $H_0 : \alpha = 0$.

Figure 4.2:
Power of unit root tests under different values of $\gamma, \alpha = -0.05$.

Power of unit root tests against stationarity at nominal sizes of 1% and 5%. The DGP is given by $\Delta r_t = \mu + \alpha r_{t-1} + \varepsilon_t, \varepsilon_t | I_{t-1} \sim N\left(0, \sigma^2 r_{t-1}^{2\gamma}\right)$ with $(\mu, \alpha) = (0.0075, -0.05)$. The estimated model is $\Delta r_t = \mu + \beta r_{t-1} + u_t$, where the noise is regarded as an iid process. Power (corrected by empirical size) is computed for the i) Dickey-Fuller t-test (DF-T), $H_0 : \alpha = 0$; ii) Dickey-Fuller likelihood-ratio test (DF-LR), $H_0 : (\mu, \alpha) = 0$; and iii) Breitung's nonparametric tests (NP), $H_0 : \alpha = 0$.

Figure 4.3:
Power of unit root tests under Mixed errors, $\alpha = -0.01$
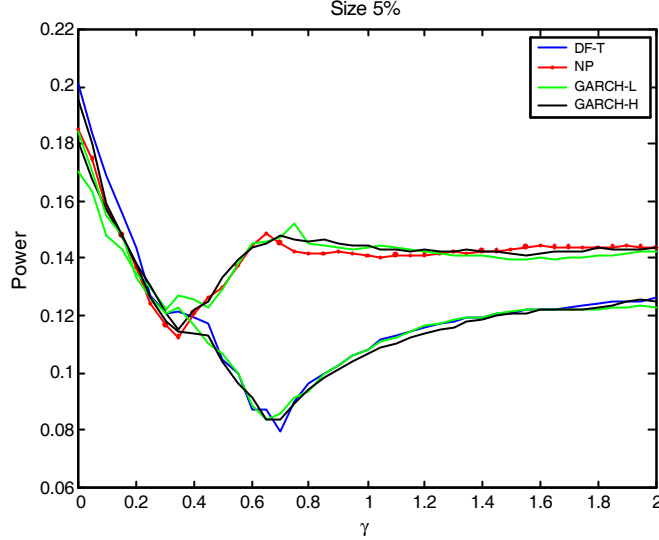
Power of unit roots tests against stationarity at 5% nominal size. The DGP is given by $\Delta r_t = \mu + \alpha r_{t-1} + r_{t-1}^{\gamma} \xi_t$, $\xi_t = \sigma_t \eta_t$ $\eta_t | I_{t-1} \sim iidN(0,1)$ where $\sigma_t^2 = \bar{\omega}_0 + \omega_1 \xi_{t-1}^2 + \omega_2 \sigma_t^2$. The DGP in all cases assumes $(\mu, \alpha) = (0.0015, -0.01)$ and $\gamma \in [0,2]$. The conditional variance dynamcis is given by:

i) Level model: $\boldsymbol{\omega} = (\sigma_0^2, 0, 0)$, $\sigma_0^2 = 0.003$.

ii) mixed model with strong persistence and low ARCH effect (GARCH-L): $\boldsymbol{\omega} = (\bar{\omega}_0, 0.05, 0.80)$

iii) mixed model with strong persistence and high ARCH effect (GARCH-H): $\omega = (\bar{\omega}_0, 0.20, 0.70)$ where $\bar{\omega}_0 = \sigma_0^2 (1 - \omega_1 - \omega_2)$.

The estimated model is $\Delta r_t = \mu + \beta r_{t-1} + u_t$, where the noise is regarded as an iid process. Power (corrected by empirical size) is computed for the i) Dickey-Fuller t-test (DF-T), $H_0 : \alpha = 0$; ii) Dickey-Fuller likelihood-ratio test (DF-LR), $H_0 : (\mu, \alpha) = 0$; and iii) Breitung's nonparametric tests (NP), $H_0 : \alpha = 0$.

Figure 4.4:
Empirical Cumulative Density Function of the *t*-stastic
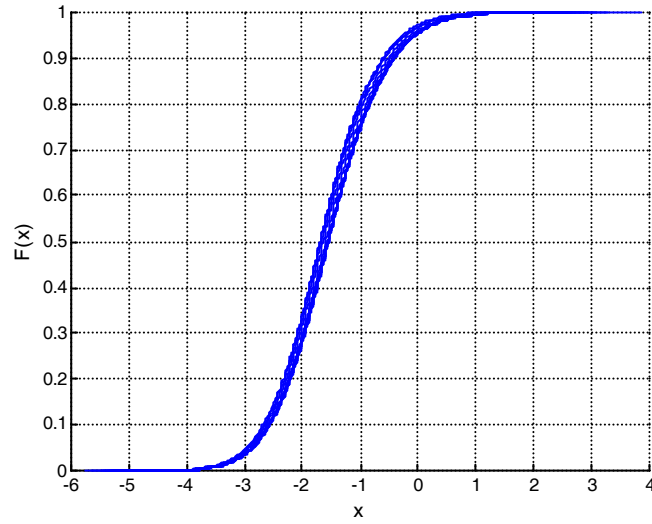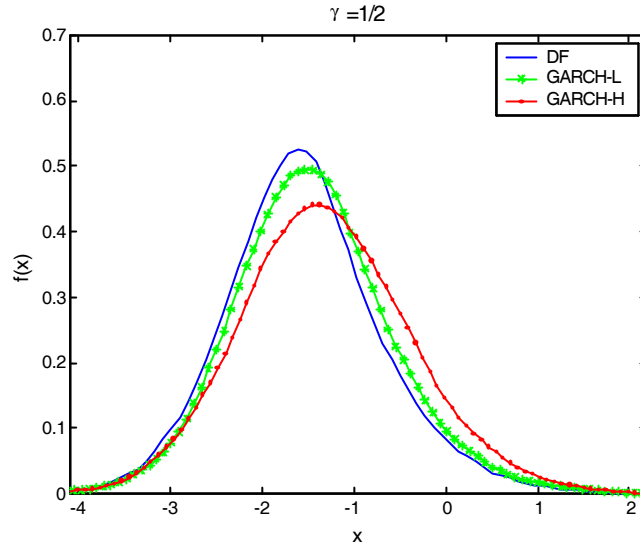for values of γ reported in Table 3.1

Kernel Densities of t-statistics under level and Mixed errors

Kernel densities of the empirical distributions of the t-stastistic with and without a time-varying scale when $\gamma = 1/2$. The conditional variance dynamcis is given by:

i) level model (Level): $\boldsymbol{\omega} = \left(\sigma_0^2, 0, 0\right), \sigma_0^2 = 0.003$;

ii) mixed model with strong persistence and low ARCH coefficient (GARCH-H): $\boldsymbol{\omega} = \left(\bar{\omega}_0, 0.05, 0.80\right)$;

iii) mixed model with high strong persistence and high ARCH coefficient (GARCH-L): $\boldsymbol{\omega} = \left(\bar{\omega}_0, 0.20, 0.70\right)$.