HYBRID STOCHASTIC LOCAL UNIT ROOTS

By

Offer Lieberman and Peter C. B. Phillips

November 2017

COWLES FOUNDATION DISCUSSION PAPER NO. 2113

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Hybrid Stochastic Local Unit Roots

Offer Lieberman* and Peter C. B. Phillips†

November 28, 2017

Abstract

Two approaches have dominated formulations designed to capture small departures from unit root autoregressions. The first involves deterministic departures that include local-to-unity (LUR) and mildly (or moderately) integrated (MI) specifications where departures shrink to zero as the sample size $n \to \infty$. The second approach allows for stochastic departures from unity, leading to stochastic unit root (STUR) specifications. This paper introduces a hybrid local stochastic unit root (LSTUR) specification that has both LUR and STUR components and allows for endogeneity in the time varying coefficient that introduces structural elements to the autoregression. This hybrid model generates trajectories that, upon normalization, have non-linear diffusion limit processes that link closely to models that have been studied in mathematical finance, particularly with respect to option pricing. It is shown that some LSTUR parameterizations have a mean and variance which are the same as a random walk process but with a kurtosis exceeding 3, a feature which is consistent with much financial data. We develop limit theory and asymptotic expansions for the process and document how inference in LUR and STUR autoregressions is affected asymptotically by ignoring one or the other component in the more general hybrid generating mechanism. In particular, we show how confidence belts constructed from the LUR model are affected by the presence of a STUR component in the generating mechanism. The import of these findings for empirical research are explored in an application to the spreads on US investment grade corporate debt.

Key words and phrases: Autoregression; Nonlinear diffusion; Stochastic unit root; Time-varying coefficient.

JEL Classification: C22

1 Introduction

For over four decades various devices have been employed to study and to model the progressive deterioration of Gaussian asymptotics in the simple first order autoregression (AR(1)) as the autoregressive coefficient (β) approaches unity from below. Edgeworth and saddlepoint approximations (Phillips, 1977, 1978) showed clearly with analytic formulae the extent of the error in the stationary asymptotics as $\beta \to 1$ and numerical computations (Evans and Savin, 1981) revealed that the unit root (UR)

*Bar-Ilan University. Support from Israel Science Foundation grant No. 1182-17 and from the Sapir Center in Tel Aviv University are gratefully acknowledged. Correspondence to: Department of Economics and Research Institute for Econometrics (RIE), Bar-Ilan University, Ramat Gan 52900, Israel. E-mail: offer.lieberman@gmail.com

†Yale University, University of Auckland, Southampton University, and Singapore Management University. Research support from the Kelly Foundation, University of Auckland is gratefully acknowledged. Email: peter.phillips@yale.edu

limit distribution typically provides better approximations than stationary limit theory in the immediate neighborhood of unity. The use of local-to-unit root (LUR) autoregressions provided a direct approach to modeling processes with a root near unity. In independent work using different methods and assumptions, Chan and Wei (1987) and Phillips (1987) explored LUR models of the form

$$Y_t = \beta_n Y_{t-1} + \varepsilon_t,\ \beta_n = e^{c/n} \sim 1 + \frac{c}{n};\ t = 1, ..., n, \tag{1}$$

where c is constant and β_n is nearly nonstationary in the sense that c/n is necessarily small as the sample size $n \to \infty$.

Under quite general conditions on ε_t and the initial condition Y_0, the asymptotic distribution of the least squares estimator of β_n takes the form of a ratio of quadratic functionals of a linear diffusion process that depends on the localizing coefficient c in (1) and nonparametric quantities that depend on the one-sided and two-sided long run variances of ε_t. These results provided a natural path to the analysis of power functions (Phillips, 1987) and power envelopes for UR tests (Elliott *et. al.*, 1995; Elliott and Stock, 1996), as well as the construction of confidence intervals (Stock, 1991) and prediction intervals (Campbell and Yogo, 2006; Phillips, 2014) in models where persistence in the regressors is relevant in practical work.

The array mechanism of (1) has also proved useful in developing methods of uniform inference. Giraitis and Phillips (2006) established uniform asymptotic theory for the OLS estimator of β_n in models like (1) but where β_n is more distant from unity so that $(1 - \beta_n)\, n \to \infty$. These models allow values of stationary β_n that include neighborhoods of unity beyond the immediate $O\left(n^{-1}\right)$ vicinity of unity, such as when $\beta_n = 1 - L_n/n$, where $L_n \to \infty$ is slowly varying at infinity. These cases were explored in detail by Phillips and Magdalinos (2007a, 2007b) by using moderate deviations from unity of the form

$$\beta_n = 1 + \frac{c}{k_n}, \quad \text{with } c \text{ constant and } \frac{1}{k_n} + \frac{k_n}{n} \to 0. \tag{2}$$

Models with such roots are considered mildly integrated (MI) as β_n lies outside the LUR region as $n \to \infty$. Phillips and Magdalinos (2007a) and developed central limit theory for the near-stationary case ($c < 0$) and, somewhat surprisingly, for the near-explosive case ($c > 0$), finding a Cauchy limit theory in the latter case that matched the known Cauchy limit that applies in the pure explosive case under Gaussian errors (White, 1958; Anderson, 1959). In a significant advance, Mikusheva (2007, 2012) demonstrated that careful approaches to confidence interval (CI) construction with appropriate centering were capable of producing uniform inferences about the true in a wide interval that includes stationary, MI, LUR, and UR specifications.

A different approach was considered by Lieberman and Phillips (2014, 2017a, 2017b), who considered localized stochastic departures from unity via the stochastic unit root (STUR) model

$$\begin{aligned} Y_1 &= \varepsilon_1, \\ Y_t &= \mu + \exp\left(\frac{a' u_t}{\sqrt{n}}\right) Y_{t-1} + \varepsilon_t,\ t = 2, ..., n, \end{aligned} \tag{3}$$

where μ can be zero or otherwise and in which departures from unity are driven by a possibly endogenous $K \times 1$ vector of explanatory variables u_t. In their formulation, Lieberman and Phillips (2017b) allowed $\{u_t, \varepsilon_t\}$ to follow a general linear process satisfying mild summability and moment conditions. This stochastic formulation of departures from unity has proved useful in empirical applications that include dual stocks pricing (Lieberman, 2012), Exchange Traded Fund pricing (Lieberman and Phillips, 2014) and call option pricing (Lieberman and Phillips, 2017a). This line of stochastic departure from a UR follows in the tradition of earlier contributions by Leybourne, McCabe and Mills (1996), Leybourne, McCabe and Tremayne (1996), Granger and Swanson (1997), McCabe and Smith, (1998), and Yoon (2006).

The present paper investigates a hybrid model that combines both LUR and STUR specifications in a localized stochastic unit root (LSTUR) model of the following form

$$\begin{aligned} Y_1 &= \varepsilon_1, \\ Y_t &= \beta_{nt} Y_{t-1} + \varepsilon_t, \ t = 2, ..., n, \end{aligned} \tag{4}$$

where

$$\beta_{nt} = \exp\left(\frac{c}{n} + \frac{a' u_t}{\sqrt{n}}\right).$$

In this model the autoregressive coefficient is a stochastic time varying parameter that fluctuates in the vicinity of unity according to the properties of u_t, the value of the localizing constant c, and the size of the sample n. The time series $w_t = (u_t', \varepsilon_t)'$ is assumed to be generated according to a linear process framework that allows for both contemporaneous and serial cross dependence, thereby allowing the random coefficient β_{nt} to be endogenous.

The paper establishes limit theory for the normalized form of the output process Y_t in (4) and for nonlinear least squares (NLLS) estimation of the components, a and c, of β_{nt}. It turns out that the limiting output process of (4) is a nonlinear diffusion process that satisfies a nonlinear stochastic differential equation corresponding to a structural model of option pricing that has been considered in the continuous time mathematical finance literature (Föllmer and Schweizer, 1993). So the model may be considered a discrete time version of such a system. Working directly with this nonlinear continuous time system, Tao et. al. (2017) developed an estimation procedure for the structural parameters of the stochastic differential equation using a realized variance approach and established asymptotic properties of these estimates under infill asymptotics. The model considered in the present paper therefore links to the continuous time finance literature and to ongoing work on continuous time econometrics.

A primary goal of the current paper is to examine the properties of this hybrid model and, in doing so, study the implied empirical features of the model in comparison with the discrete time random walk (RW), LUR and STUR models. In particular, we show that certain LSTUR parametrizations are consistent with a mean and variance which are equal to those of a RW process but with a kurtosis coefficient which is greater than 3 - a feature which is arguably consistent with much financial data. The analysis helps to document how inference in LUR and STUR autoregressions is affected by the presence of the other component in the time varying autoregressive coefficient β_{nt} in the generating mechanism. In particular, we show how asymptotic confidence belts constructed using the LUR model (Stock, 1991) are affected by the omission of a random coefficient STUR component. The implications for empirical work of such misspecification of random departures from unity by deterministic from unity models are explored in an empirical application.

The plan for the rest of the paper is as follows. Notation, assumptions and limit theory for $n^{-1/2}Y_t$ are given in Section 2. Asymptotic theory for parameter estimation follows in Section 3. Some further results including asymptotic expansions are given in Section 4. Robustness of the misspecified STUR-based NLLS and IV estimators of a and the covariance parameters are established in Section 5. A simulation study to the effects of an omitted STUR component on the confidence belts given by Stock

(1991) for c and β in the LUR model is provided in Section 6. An empirical application supporting the analytical findings and simulations follows in Section 7. Section 8 concludes. All proofs are placed in the appendix.

2 Preliminary Limit theory for the LSTUR Model

We start with the following assumption that will be used in the sequel detailing the generating mechanism for w_t.

Assumption 1. *The vector w_t is a linear process satisfying*

$$w_t = G(L)\eta_t = \sum_{j=0}^{\infty} G_j \eta_{t-j},\ \sum_{j=1}^{\infty} j \|G_j\| < \infty,\ G(1) \text{ has full rank } K+1, \tag{5}$$

η_t is iid, zero mean with $E(\eta_t\eta_t') = \Sigma_\eta > 0$ and $\max E|\eta_{i0}|^p < \infty$, for some $p > 4$.

Under Assumption 1, w_t is zero mean, strictly stationary and ergodic, with partial sums satisfying the invariance principle

$$n^{-1/2}\sum_{t=1}^{\lfloor n\cdot \rfloor} w_t \Rightarrow B(\cdot) \equiv \text{BM}\left(\Sigma^{\ell r}\right),\ \Sigma^{\ell r} = \begin{pmatrix} \Sigma_u^{\ell r} & \Sigma_{u\varepsilon}^{\ell r} \\ \Sigma_{u\varepsilon}^{\ell r \prime} & \left(\sigma_\varepsilon^{\ell r}\right)^2 \end{pmatrix}, \tag{6}$$

where $\lfloor \cdot \rfloor$ is the floor function and $B = (B_u, B_\varepsilon)'$ is a vector Brownian motion. The matrix $\Sigma^{\ell r} = G(1)\Sigma_\eta G(1)' > 0$ is the long run covariance matrix of w_t, with $K \times K$ submatrix $\Sigma_u^{\ell r} > 0$, scalar $\left(\sigma_\varepsilon^{\ell r}\right)^2 > 0$ and $K \times 1$ vector $\Sigma_{u\varepsilon}^{\ell r}$. In component form, we write (5) as

$$\begin{aligned} w_t &= \begin{pmatrix} u_t \\ \varepsilon_t \end{pmatrix} = \begin{pmatrix} G_{11}(L) & G_{12}(L) \\ G_{21}(L) & G_{22}(L) \end{pmatrix} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} = \begin{pmatrix} G_1(L) \\ G_2(L) \end{pmatrix} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=0}^{\infty} G_{1,j}\eta_{t-j} \\ \sum_{j=0}^{\infty} G_{2,j}\eta_{t-j} \end{pmatrix} \end{aligned} \tag{7}$$

where η_{1t} is $K \times 1$, η_{2t} is scalar, $G_{1,j}$ is $K \times (K+1)$ and $G_{2,j}$ is $1 \times (K+1)$.

We denote the contemporaneous covariance matrix of w_t by $\Sigma > 0$, with corresponding components $\Sigma_{u\varepsilon} = E(u_t u_t') > 0$, $\Sigma_{u\varepsilon} = E(u_t \varepsilon_t)$ and $\sigma_\varepsilon^2 = E\left(\varepsilon_t^2\right) > 0$. The one-sided long run covariance matrices are similarly denoted by $\Lambda = \sum_{h=1}^{\infty} E(w_0 w_h')$ and $\Delta = \sum_{h=0}^{\infty} E(w_0 w_h') = \Lambda + \Sigma$, with corresponding component submatrices $\Lambda_{u\varepsilon} = \sum_{h=1}^{\infty} E(u_0 \varepsilon_h)$, $\lambda_{\varepsilon\varepsilon} = \sum_{h=1}^{\infty} E(\varepsilon_0 \varepsilon_h)$, $\Delta_{u\varepsilon} = \sum_{h=0}^{\infty} E(u_0 \varepsilon_h)$, $\Delta_{\varepsilon\varepsilon} = \sum_{h=0}^{\infty} E(\varepsilon_0 \varepsilon_h)$.

We use H and L to denote the zero-one duplication and elimination matrices for which

$$vec(A) = Hvech(A) \text{ and } vech(A) = Lvec(A), \tag{8}$$

where A is a symmetric matrix of order $K+1$. Under Assumption 1, centred partial sums of $\eta_t\eta_t'$ satisfy the invariance principle

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor nr \rfloor} vech\left(\eta_t\eta_t' - \Sigma_\eta\right) \Rightarrow \xi(r), \tag{9}$$

where $\xi(r)$ is vector Brownian motion with covariance matrix

$$\Sigma_{\eta\otimes\eta} = E\left(L\left((\eta_t \otimes \eta_t) - E(\eta_t \otimes \eta_t)\right)\left((\eta'_t \otimes \eta'_t) - E(\eta'_t \otimes \eta'_t)\right)L'\right).$$

Furthermore, for any $l \neq 0$ we denote by $\zeta(r)$ the vector Brownian motion with covariance matrix

$$\mathbb{E}\left((\eta_t\eta'_t \otimes \eta_{t-l}\eta'_{t-l})\right) = E(\eta_t\eta'_t) \otimes E(\eta_{t-l}\eta'_{t-l}) = \Sigma_\eta \otimes \Sigma_\eta.$$

Finally, the matrix of third order moments of η_t is denoted

$$M_3 = E\left((\eta_t \otimes \eta_t)\eta'_t\right). \tag{10}$$

The limit process of the scaled time series Y_t is given in the following Lemma.

Lemma 1 *For the model (4), under Assumption 1,*

$$\frac{Y_{t=\lfloor nr \rfloor}}{\sqrt{n}} \Rightarrow G_{a,c}(r) := e^{rc+a'B_u(r)}\left(\int_0^r e^{-pc-a'B_u(p)}dB_\varepsilon(p) - a'\Delta_{u\varepsilon}\int_0^r e^{-pc-a'B_u(p)}dp\right). \tag{11}$$

Lemma 1 extends the limit theory for the special case where there is no LUR component ($c = 0$) and the case where there is no STUR component ($a = 0$). The latter case leads to the familiar limit

$$\frac{Y_t}{\sqrt{n}} \Rightarrow \sigma_\varepsilon^{\ell r}\int_0^r e^{(r-s)c}dW(s) =: \sigma_\varepsilon^{lr}J_c(r) = G_{0,c}(r) =: G_c(r), \text{ say}$$

where $W(r)$ is standard BM and $J_c(r)$ is a linear diffusion (Phillips, 1987).

3 Parameter Estimation

Let $\hat{a}_n$ and $\hat{c}_n$ denote the NLLS of a and c. Explicit formulae for these estimates are not available but first order conditions are given in (65) of the Appendix. This section presents the limit theory for these estimates in various cases. We use the following sample covariance limit theory.

Lemma 2 *For the model (4), under Assumption 1*

$$\frac{1}{n}\sum_{t=2}^n \varepsilon_t Y_{t-1} \Rightarrow \int_0^1 G_{a,c}(r)\,dB_\varepsilon(r) + \Lambda'_{u\varepsilon}a\int_0^1 G_{a,c}(r)\,dr + \lambda_{\varepsilon\varepsilon}. \tag{12}$$

The limit in (12) reduces to the standard result $\int_0^1 G_{0,c}(r)\,dB_\varepsilon(r) + \lambda_{\varepsilon\varepsilon}$ when $a = 0$.

We start with the case where a is known, which enables us to relate results to earlier literature on the LUR model in a convenient way. This simplification is relaxed below.

Theorem 3 *For the model (4), under Assumption 1 and when a is known,*

$$\hat{c}_n - c \Rightarrow \left(\int_0^1 G_{a,c}^2(r)\,dr\right)^{-1}\left(\int_0^1 G_{a,c}(r)\,dB_\varepsilon(r) + \Delta'_{u\varepsilon}a\int_0^1 G_{a,c}(r)\,dr + \lambda_{\varepsilon\varepsilon}\right). \tag{13}$$

When $a=0$ the result in (13) reduces to the standard limit theory for the least squares estimate $\hat{c}_n$ of the localizing coefficient c in a LUR model, viz.,

$$\hat{c}_n - c \Rightarrow \left(\int_0^1 G_c^2(r)\,dr\right)^{-1}\left(\int_0^1 G_c(r)\,dB_\varepsilon(r) + \lambda_{\varepsilon\varepsilon}\right). \tag{14}$$

The presence of the stochastic UR component alters the usual limit theory (14) by (i) modifying the limiting output process to $G_{a,c}(r)$ in which the effects of the random autoregressive coefficient figure, and (ii) introducing the additional bias term, $\Delta'_{u\varepsilon} a \int_0^1 G_{a,c}(r)\,dr$ to the limit distribution.

Next consider the case in which a is unknown.

Theorem 4 *For the model (4) under Assumption 1 with $\Sigma_{u\varepsilon} \neq 0$*

$$(\hat{a}_n - a) \Rightarrow \left(\int_0^1 G_{a,c}^2(r)\,dr\right)^{-1}\left(\int_0^1 G_{a,c}(r)\,dr\right)\Sigma_u^{-1}\Sigma_{u\varepsilon}, \tag{15}$$

and

$$\begin{aligned}(\hat{c}_n - c) \Rightarrow & \left(\int_0^1 G_{a,c}^2(r)\,dr\right)^{-1}\left(\int_0^1 G_{a,c}(r)\,dB_\varepsilon(r) + \Lambda'_{u\varepsilon} a \int_0^1 G_{a,c}(r)\,dr + \lambda_{\varepsilon\varepsilon}\right) \\ & -\Sigma'_{u\varepsilon}\Sigma_u^{-1}\frac{\int_0^1 G_{a,c}^2(r)\,dB_u(r) + 2\left(\Lambda'_{uu} a \int_0^1 G_{a,c}^2(r)\,dr + \Lambda_{u\varepsilon}\int_0^1 G_{a,c}(r)\,dr\right)}{\left(\int_0^1 G_{a,c}^2(r)\,dr\right)^2}\int_0^1 G_{a,c}(r)\,dr.\end{aligned}$$

When $\Sigma_{u\varepsilon} = 0$,

$$\begin{aligned}\sqrt{n}(\hat{a}_n - a) \Rightarrow & \frac{\Sigma_u^{-1}}{\int_0^1 G_{a,c}^2(r)\,dr}\Bigg(\sum_{j=0}^{\infty}(G_{2,j}\otimes G_{1,j})\,H\int_0^1 G_{a,c}(r)\,d\xi(r) \\ & + \sum_{j=1}^{\infty}(G_{2,j}\otimes G_{1,j})\,M_3\left(\left(\sum_{i=0}^{j-1}G_{1,i}\right)' a\int_0^1 G_{a,c}(r)\,dr + \left(\sum_{i=0}^{j-1}G_{2,i}\right)'\right) \\ & + \sum_{j\neq k}(G_{2,k}\otimes G_{1,j})\int_0^1 G_{a,c}(r)\,d\zeta(r) + E\left(\varepsilon_t u_t u_t' \hat{a}_n\right)\int_0^1 G_{a,c}(r)\,dr\Bigg)\end{aligned}$$

and $(\hat{c}_n - c) \Rightarrow \left(\int_0^1 G_{a,c}^2(r)\,dr\right)^{-1}\left(\int_0^1 G_{a,c}(r)\,dB_\varepsilon(r) + \Lambda'_{u\varepsilon} a \int_0^1 G_{a,c}(r)\,dr + \lambda_{\varepsilon\varepsilon}\right)$.

The distribution of $\hat{a}_n$ depends on the localizing coefficient c through $G_{a,c}(r)$. The estimator is consistent when $\Sigma_{u\varepsilon} = 0$. When $\Sigma_{u\varepsilon} \neq 0$, the parameter a may be estimated consistently using instrumental variables (Lieberman and Phillips, 2017b) or by infill asymptotics via a two-stage process involving realized variance when high frequency data is available (Tao et al., 2017)). Unlike $\hat{a}_n$, $\hat{c}_n$ is inconsistent irrespective of whether $\Sigma_{u\varepsilon} = 0$ and this accords with known results for simpler models without STUR effects (Phillips, 1987). However, the localizing coefficient c may be estimated consistently under certain conditions when the data support joint large span and infill asymptotics, as shown in Tao et al. (2017).

The next result concerns the OLS estimator of the autoregressive coefficient β_{nt}. Its asymptotic distribution and that of the t-statistic for testing the hypothesis of a unit root are used later in the paper to construct confidence intervals for the autoregressive parameter.

Theorem 5 *The ols estimator of β_{nt} in the model (4) and under Assumption 1 satisfies*

$$n\left(\hat{\beta}_{nt}-1\right) \Rightarrow c+\frac{a'\Sigma_u a}{2}$$
$$+\frac{a'\left(\int_0^1 G_{a,c}^2(r)\,dB_u(r)+2\left(\Lambda_{uu}'a\int_0^1 G_{a,c}^2(r)\,dr+\frac{3}{2}\Lambda_{u\varepsilon}\int_0^1 G_{a,c}(r)\,dr\right)\right)+\int_0^1 G_{a,c}(r)\,dB_\varepsilon(r)+\lambda_{\varepsilon\varepsilon}}{\int_0^1 G_{a,c}^2(r)\,dr}.$$

When w_t is a martingale difference, the one-sided long run covariances are zero and the limit result reduces to

$$n\left(\hat{\beta}_{nt}-1\right) \Rightarrow c+\frac{a\prime\Sigma_u a}{2}+\frac{a'\int_0^1 G_{a,c}^2(r)\,dB_u(r)+\int_0^1 G_{a,c}(r)\,dB_\varepsilon(r)}{\int_0^1 G_{a,c}(r)^2\,dr}. \tag{16}$$

4 Empirical Implications and Further Results

This section explores the relationships among the RW, LUR and LSTUR models in more detail in the univariate case ($K=1$) with $\Sigma_{u\varepsilon}=0$ and for iid (u_t', ε_t). This special case highlights the distinguishing features of these models and some key elements in their relationships that are important for empirical work. The output limit process (11) in this case has the simpler form

$$G_{a,c}(r)=e^{rc+aB_u(r)}\int_0^r e^{-pc-a'B_u(p)}dB_\varepsilon(p), \tag{17}$$

which satisfies the generating differential equation

$$dG_{a,c}(r)=aG_{a,c}(r)\,dB_u(r)+dB_\varepsilon(r)+\left(c+\frac{b}{2}\right)G_{a,c}(r)\,dr, \tag{18}$$

where $b=(a\sigma_u)^2$. The covariance kernel and moments of the output process $G_{a,c}(r)$ are given in the following result.

Lemma 6 *For the model (4), under the assumptions that $K=1$, $\Sigma_{u\varepsilon}=0$, and u_t and ε_t are iid,*

$$E\left(G_{a,c}(r)\right)=0,$$

$$Cov\left(G_{a,c}(r),G_{a,c}(s)\right)=\sigma_\varepsilon^2 e^{\left(c+\frac{b}{2}\right)(r\vee s-r\wedge s)}\frac{e^{2(c+b)r\wedge s}-1}{2(c+b)}=:\gamma_{G_{a,c}}(r,s), \tag{19}$$

and

$$E\left(G_{a,c}^4(r)\right)=\frac{3\sigma_\varepsilon^4 e^{4(c+2b)r}}{c+b}\left(\frac{1-e^{-2(c+3b)r}}{2(c+3b)}-\frac{1-e^{-4(c+2b)r}}{4(c+2b)}\right). \tag{20}$$

An immediate consequence of Lemma 6 is that

$$Var\left(G_{a,c}(r)\right)=E\left(G_{a,c}^2(r)\right)=\sigma_\varepsilon^2\frac{e^{2(c+b)r}-1}{2(c+b)}. \tag{21}$$

The function $\left(e^{zr}-1\right)/z$ is monotonically increasing and equals r at $z=0$. It follows from (21) that an LSTUR process with $c=-b$ has a limit process with variance $\sigma_{\varepsilon}^{2}r$, which is the variance of a Brownian motion. However, the process $G_{a,c}\left(r\right)$ is non-Gaussian in this case and has covariance kernel $\gamma_{G_{a,c=-b}}\left(r,s\right)=\sigma_{\varepsilon}^{2}e^{-\frac{b}{2}\left(r\vee s-r\wedge s\right)}r\wedge s\neq r\wedge s$. Thus, the particular case where $c+b=0$ provides an interesting example of a non-Gaussian LSTUR limit process whose first two moments match those of Brownian motion. For $c+b<0$ the variance of the LSTUR limit is less than that of Brownian motion and for $c+b>0$ the variance is larger and increasing with the value of $c+b$. In particular, given c, the variance of the process increases with b (equivalently, with either $|a|$ or σ_{u}). Alternatively, given b, the variance of the process increases with c. A small b expansion of (21) yields

$$Var\left(G_{a,c}\left(r\right)\right)=\sigma_{\varepsilon}^{2}\left(\frac{e^{2cr}-1}{2c}+\frac{\left(1+e^{2cr}\left(2cr-1\right)\right)}{2c^{2}}b+O\left(b^{2}\right)\right),$$

showing that the lead term of the variance is the variance of the linear diffusion LUR process, as expected, coupled with a second linear term in b.

Even though the special case $c+b=0$ matches the first two moments of the LSTUR limit process with a Brownian motion, the kurtosis of the processes differ. In particular, using Lemma 6, we have

$$\lim_{b+c\rightarrow0}E\left(G_{a,c}^{4}\left(r\right)\right)=\frac{3\sigma_{\varepsilon}^{4}\left(e^{-4cr}+4cr-1\right)}{8c^{2}}=3\sigma_{\varepsilon}^{4}\left(r^{2}+O\left(c\right)\right),\ \text{and}\lim_{b+c\rightarrow0}E\left(G_{a,c}^{2}\left(r\right)\right)=\sigma_{\varepsilon}^{2}r, \tag{22}$$

so that in this case the kurtosis of the process, $\left\{3\sigma_{\varepsilon}^{4}\left(r^{2}+O\left(c\right)\right)\right\}/\left(\sigma_{\varepsilon}^{2}r\right)^{2}=3+O\left(c\right)$, matches that of Brownian motion when $c\rightarrow0$ because the variances are the same when $c+b=0$. However, kurtosis exceeds 3 in the case $c+b=0$ and $c<0$ and kurtosis increases as c becomes more negative when $c+b=0$. The case $c+b=0$ and $c>0$ is excluded because $b=\left(a\sigma_{u}\right)^{2}\geq0$.

An instantaneous kurtosis measure for the process increments $dG_{a,c}(r)$ at r may be defined as

$$\kappa_{b,c}\left(r\right)=\frac{E\left(E\left[\left(dG_{a,c}\left(r\right)\right)^{4}|\mathcal{F}_{r}\right]\right)}{\left\{E\left(E\left[\left(dG_{a,c}\left(r\right)\right)^{2}|\mathcal{F}_{r}\right]\right)\right\}^{2}},$$

which has the following explicit form for the diffusion process (18)

$$\kappa_{b,c}\left(r\right)=3+\frac{3b^{2}\left[E\left(G_{a,c}^{4}\left(r\right)\right)-\left(E\left(G_{a,c}^{2}\left(r\right)\right)\right)^{2}\right]}{b^{2}\left(E\left(G_{a,c}^{2}\left(r\right)\right)\right)^{2}+\sigma_{\varepsilon}^{4}+2b\sigma_{\varepsilon}^{2}E\left(G_{a,c}^{2}\left(r\right)\right)}+o_{p}\left(1\right), \tag{23}$$

as shown in Lemma 12 of the Appendix. The second term on the right side of (23) shows the excess kurtosis in the process increments arising from the non-Gaussianity of $G_{a,c}(r)$. As $b\rightarrow0$ we have $\kappa_{b,c}\left(r\right)\rightarrow3$, as expected since in that case $G_{a,c}\left(r\right)\rightarrow G_{c}\left(r\right)=\int_{0}^{r}e^{-(r-p)c}dB_{\varepsilon}\left(p\right)=\sigma_{\varepsilon}J_{c}\left(r\right)$, which is a linear Gaussian diffusion. But when $c\rightarrow0$, $G_{a,c}\left(r\right)\rightarrow G_{a}\left(r\right)=e^{aB_{u}(r)}\int_{0}^{r}e^{-a'B_{u}(p)}dB_{\varepsilon}\left(p\right)$ which is still non-Gaussian and $\kappa_{b,0}\left(r\right)>3$. A large b expansion of (23) shows that $\kappa_{b,c}\left(r\right)\sim\frac{9}{6}e^{4br}$, with kurtosis increasing exponentially with $b=a^{2}\sigma_{u}^{2}$, measuring the impact of non-Gaussianity in the process $G_{a,c}(r)$ as either a^{2} or σ_{u}^{2} rise, which originates in the nonlinear dependence of $G_{a,c}(r)$ on $aB_{u}\left(r\right)$.

These results are summarized in the following remark.

For the model (4) with $K = 1$, $\Sigma_{u\varepsilon} = 0$, and iid (u_t, ε_t), the instantaneous kurtosis measure of the increment process $dG_{a,c}(r)$ is

$$\kappa_{b,c}(r) = 3 + \frac{3b^2 Var\left(G_{a,c}^2(r)\right)}{b^2\left(E\left(G_{a,c}^2(r)\right)\right)^2 + \sigma_\varepsilon^4 + 2b\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right)} + o_p(1),$$

and the kurtosis of the process $G_{a,c}(r)$ itself satisfies

$$\lim_{c+b\to 0} \frac{E\left(G_{a,c}^4(r)\right)}{\left(E\left(G_{a,c}^2(r)\right)\right)^2} = \frac{3\left(e^{-4cr} + 4cr - 1\right)}{8(cr)^2},$$

which rises as $c \to -\infty$ and has minimum of 3 at $c = 0$.

Financial data are well known to resemble trajectories generated by a RW but with the important exception that the kurtosis coefficient of asset returns exceeds 3, typically by a large margin. This stylized feature of financial times series matches the corresponding characteristic of the LSTUR limit process $G_{a,c}(r)$, which has random wandering behavior similar to a Gaussian RW but with kurtosis of its increments in excess of Gaussian increments. These features give the LSTUR process a desirable property for empirical work.

In spite of their common features, the limit processes corresponding to RW, LUR, and LSTUR time series are very different, including the special parameter configuration $c + b = 0$ in LSTUR. In particular, when $K = 1$, $\Sigma_{u\varepsilon} = 0$, and (u_t, ε_t) are iid, the limit process $G_{a,c}(r)$ satisfies the stochastic differential equation (18). Non-Gaussianity in the process $G_{a,c}(r)$ is then governed by the magnitude of the coefficient $b = a^2\sigma_u^2$. The following result sheds light on the composition of the process $G_{a,c}(r)$ when the parameter b is small.

Lemma 7 *For the model (4) when $K = 1$, $\Sigma_{u\varepsilon} = 0$, and u_t and ε_t are iid,*

$$G_{a,c}(r) = G_c(r) + V_{c,a}(r) + O_p(b), \tag{24}$$

where $G_c(r) = \int_0^r e^{(r-p)c} dB_\varepsilon(p)$ is a Gaussian process, $V_{c,a}(r) = a\int_0^r e^{(r-p)c}\left(B_u(r) - B_u(p)\right) dB_\varepsilon(p)$ is a mixed Gaussian process, and $G_c(r)$ and $V_{c,a}(r)$ are uncorrelated. To first order in b

$$Var\left(G_{a,c}(r)\right) = \sigma_\varepsilon^2\left(\frac{e^{2cr} - 1}{2c} + \frac{b\left(e^{2cr}(2cr - 1) + 1\right)}{(2c)^2}\right) + O(b). \tag{25}$$

According to (24) and (25) the STUR component effect is small when $b = a^2\sigma_u^2$ is small, in which case the limit process $G_{a,c}(r)$ is approximately mixed Gaussian, with variance that exceeds the variance of the LUR process component, viz.,

$$\sigma_\varepsilon^2\left(\frac{e^{2cr} - 1}{2c} + \frac{b\left(e^{2cr}(2cr - 1) + 1\right)}{(2c)^2}\right) \geq \sigma_\varepsilon^2\left(\frac{e^{2cr} - 1}{2c}\right).$$

In the special configuration $c + b = 0$ when b is small, c is also small and then the LSTUR process is approximately Brownian motion with variance $\sigma_\varepsilon^2 r$.

5 Robustness to Misspecification

This section explores the robustness of STUR-based NLLS and IV parameter estimation to misspecification that arises from an LSTUR generating mechanism. Let $\left(\tilde{a}_n, \tilde{\sigma}^2_{\varepsilon,n}\right)$ be the STUR-based NLLS estimates of $\left(a, \sigma^2_\varepsilon\right)$, so that

$$\tilde{a}_n = \arg\min_a \sum_t \left(Y_t - e^{a' u_t/\sqrt{n}} Y_{t-1}\right)^2, \; \tilde{\sigma}^2_{\varepsilon,n} = \frac{1}{n}\sum_t \left(Y_t - e^{\tilde{a}'_n u_t/\sqrt{n}} Y_{t-1}\right)^2.$$

When $\Sigma_{u\varepsilon} = 0$, (u_t, ε_t) is iid, and the generating mechanism is LSTUR, $\tilde{a}_n$ and $\tilde{\sigma}^2_{\varepsilon,n}$ are still consistent for a and σ^2_ε, as shown below.

Lemma 8 *For the model (4) when $\Sigma_{u\varepsilon} = 0$ and (u_t, ε_t) is iid,*

(i) $\sqrt{n}\left(\tilde{a}_n - a\right) \Rightarrow \frac{1}{\int_0^1 G_{a,c}(r)dr}\Sigma_u^{-1}\left\{\left\{E\left(\varepsilon_t u_t u_t'\right)\right\} a \int_0^1 G_{a,c}(r)\, dr + \int_0^1 G_{a,c}(r)\, dB_{u\varepsilon}(r)\right\}$,

(ii) $\tilde{\sigma}^2_{\varepsilon,n} \rightarrow_p \sigma^2_\varepsilon$,

(iii) If $\tilde{Y}^p_t = \left(1 + \frac{\tilde{a}'_n u_t}{\sqrt{n}} + \frac{(\tilde{a}'_n u_t)^2}{2n}\right) Y_{t-1}$ *and* $\hat{Y}^p_t = \left(1 + \frac{\tilde{a}'_n u_t}{\sqrt{n}} + \frac{1}{n}\left(c + \frac{(\tilde{a}'_n u_t)^2}{2}\right)\right) Y_{t-1}$ *are in-sample predictors based on STUR and LSTUR specifications, then*

$$\frac{1}{\sqrt{n}}\sum_t \left(Y_t - \tilde{Y}^p_t\right) \Rightarrow B_\varepsilon(1) + c\int_0^1 G_{a,c}(r)\, dr, \quad \frac{1}{\sqrt{n}}\sum_t \left(Y_t - \hat{Y}^p_t\right) \Rightarrow B_\varepsilon(1),$$

and $\sum_t \left(\tilde{Y}^p_t - \hat{Y}^p_t\right)^2 \Rightarrow c^2 \int_0^1 G^2_{a,c}(r)\, dr$.

Parts (i) and (ii) of Lemma 8 are obtained in the same way as Theorems 2 and 3 of Lieberman and Phillips (2017a). The only difference in the limit distribution in (i) compared to the case where STUR is the correct specification the limit process is now $G_{a,c}(r)$ rather than $G_a(r)$. An implication of this result is that the n^{-1}-normalized sum of squared errors of (the misspecified) STUR and LSTUR will be identical asymptotically and therefore, for large enough n, AIC and BIC should always favor STUR over LSTUR, even when LSTUR is the true DGP. This finding corresponds with the known result that information criteria such as BIC are typically blind to local departures of the LUR variety (Phillips and Ploberger, 2003; Leeb and Pötscher, 2005).

In part (iii) of the Lemma, $\tilde{Y}^p_t$ and $\hat{Y}^p_t$ are the STUR- and LSTUR-based predictors of Y_t. The latter is infeasible as c is unknown but may be replaced by an inconsistent estimate or by imposing a special restriction such as $c = -b$, which is discussed in Section 4. In this case, the $n^{-1/2}$-normalized error sums differ by the term $c\int_0^1 G_{a,c}(r)\, dr$ and the sum of squared discrepancies between the two predictors converges to $c^2 \int_0^1 G_{a,c}(r)^2\, dr$ so that the value of the localizing coefficient c affects these differentials directly as well as through the correct limit process $G_{a,c}(r)$ corresponding to LSTUR rather than $G_a(r)$

In the case $\Sigma_{u\varepsilon} \neq 0$, even the correctly specified LSTUR-based NLLS estimator is inconsistent. Fortunately, for the LSTUR model the misspecified STUR-based IV estimators (Lieberman and Phillips, 2017b) of a and the covariance parameters are still consistent. Let $\tilde{a}^{IV}_n$, $\tilde{\gamma}^{IV}_{\varepsilon,n}(j)$ and $\tilde{\gamma}^{IV}_{u,\varepsilon,n}(j)$ be the STUR-based IV estimators of a, $\gamma_\varepsilon(j) = Cov\left(\varepsilon_t, \varepsilon_{t-j}\right)$ and $\gamma_{u,\varepsilon}(j) = Cov\left(u_t, \varepsilon_{t-j}\right)$ for $(j = 0, 1, 2, \ldots)$. That is, $\tilde{a}^{IV}_n$ solves the K-moment conditions

$$\sum_{t=2}^n \left(Y_t - \beta_{nt}\left(\tilde{a}^{IV}_n\right) Y_{t-1}\right) Z_t = 0, \tag{26}$$

where Z_t is a vector of instruments which satisfy Assumption 3 of Lieberman and Phillips (2017b),

$$\tilde{\gamma}_{\varepsilon,n}^{IV}(j) = \frac{1}{n}\sum_{t=j+2}^{n} e_t^{IV} e_{t-j}^{IV},\ \tilde{\gamma}_{u,\varepsilon,n}^{IV}(j) = \frac{1}{n}\sum_{t=j+2}^{n} u_t e_{t-j}^{IV},$$

and

$$e_t^{IV} = Y_t - e^{\tilde{a}_n^{IV\prime} u_t/\sqrt{n}} Y_{t-1},\ t = 2, ..., n. \tag{27}$$

In the misspecified model case, the STUR-IV estimators are still consistent. In particular, for the model (4) and under Assumptions 2-3 of Lieberman and Phillips (2017b), we have $\tilde{a}_n^{IV} - a = O_p\left(n^{-1/2}\right)$, $\tilde{\gamma}_{\varepsilon,n}^{IV}(j) - \gamma_{\varepsilon}(j) = O_p\left(n^{-1/2}\right)$ and $\tilde{\gamma}_{u,\varepsilon,n}^{IV}(j) - \gamma_{u,\varepsilon}(j) = O_p\left(n^{-1/2}\right)$ for all fixed and finite j. The proof follows the arguments given in Theorems 3 and 4 of Lieberman and Phillips (2017b) and is omitted. These results are employed in the empirical section below.

6 The Effects of Misspecification on CI Construction

Stock (1991) constructed confidence belts for the localizing coefficient c in the LUR model from which confidence intervals (CIs) valid within a vicinity of unity for the autoregressive coefficient β could be deduced from unit root tests. Application of this methodology to the Nelson Plosser (1982) data produced very wide confidence bands. Hansen (1999) showed how the accuracy of these simulation-based CIs deteriorated as the stationary regrion was approached. He suggested a grid bootstrap procedure for the construction of the CIs which helped to improve coverage accuracy of the bands. Phillips (2014) provided an asymptotic analysis that explained the deterioration of the CIs as the generating process moves deeper into the LUR region and ultimately the stationary region, reinforcing the work of Hansen (1999) and Mikusheva (2007) on the role of correctly centred statistics in the development of uniformly valid confidence bands.

This work was all conducted using LUR formulations of departures from unity. The present section addresses the issue of how confidence band accuracy is affected by an underspecification of an LSTUR process as an LUR process. To this end, we consider the limit distribution given in (16). The t-ratio for the UR hypothesis is given by

$$t_{\beta} = \frac{n\left(\hat{\beta}_n - 1\right)}{\left(\hat{\sigma}_{\varepsilon}^2/n^{-2}\sum_t Y_{t-1}^2\right)^{1/2}} \Rightarrow \frac{1}{\sigma_{\varepsilon}}\left(\int_0^1 G_{a,c}(r)^2\, dr\right)^{1/2}\left(c + \frac{a\prime\Sigma_u a}{2} + \frac{a'\int_0^1 G_{a,c}^2(r)\, dB_u(r) + \int_0^1 G_{a,c}(r)\, dB_{\varepsilon}(r)}{\int_0^1 G_{a,c}(r)^2\, dr}\right), \tag{28}$$

where $\hat{\sigma}_{\varepsilon}^2$ is a consistent estimator of σ_{ε}^2, such as the IV estimator $\tilde{\gamma}_{\varepsilon,n}^{IV}(0)$, discussed in Section 5. When $a = 0$, the result (28) reduces to equation (2) of Phillips (2014), or equation (5) of Stock (1991)[1]. To get an idea of how the confidence belts of Stock (1991) would be affected by the omission of a stochastic component, we simulated the right side of (28) with parameter settings $\sigma_{\varepsilon}^2 = 1$, $\Sigma_{u\varepsilon} = 0$, $a = (0, 1, 2, 3, 4)$, $\sigma_u^2 = (0.1, 1)$ and $c = (1, 0, -1, -5, -10, ..., -35)$. As $b = (a\sigma_u)^2$, the setting includes parameter combinations under which $-35 < c + b < 17$. Table 1 was constructed with 5000 replications and 400 integral points and includes the 5th, 10th, 50th, 90th and 95th percentiles of the simulated asymptotic distribution, as well as the width of the 80%- and 90%-CIs in each case.

The most striking feature of the results is that the CIs become wider as the value of $c + b$ increases.

[1] To be precise, Stock (1991, equation (5)) used a demeaned ADF t-statistic in constructing the confidence belts.

In other words, *given* a c-value, the effect of misspecification becomes more pronounced as the value of a and/or σ_u^2 increases. This is expected, as a very negative value of $(c+b)$, for instance, is consistent with a dominant LUR - relative to STUR - component. Each value of c gives a point on each confidence line in Stock' (1991) confidence belts, from which the permissible values of the test statistic can be implied, given a confidence level, and vice versa. Therefore, wider CIs for the test statistic for larger a and/or σ_u^2 values translate to wider CIs for c and for β, implying that Stock's (1991) conclusion that the CIs for β are typically wide applies with greater force in the presence of a STUR component in the process. In effect, the CIs grow wider as the STUR signal becomes more dominant. For instance, suppose that the observed value $t_{\hat{\beta}} = -2.0$. Reading from Table 1, the value $c = 0$ is not in the 90% CI if $a = 0, 1, 2$ and $\sigma_u^2 = 0.1$, but it is inside the 90% CI if $a = 3, 4$ and $\sigma_u^2 = 0.1$. Put differently, when $b = [0, 0.4]$, $c = 0$ is not in the 90% CI, given a $t_{\hat{\beta}}$-value of -2.0, but for larger b-values, the value of $c = 0$ is within the 90% CI.

The above discussion pertains to a given c-value. In practice, as shown in the next section, a fitted LSTUR model may lead to a substantially narrower CI for c, compared with the CI for c that would be obtained from an LUR model. The results shown in this simulation are simply illustrative of the implications of having a generating mechanism that involves random as well as deterministic departures from unity. Comprehensive tabulation is a multidimensional task, involving a constellation of conceivable parameter values, and the limit theory is non-pivotal so that practical work would require consistent estimates of many unknown parameters and an approach that led to uniformly valid (over LUR and STUR departures from unity as well as stationary departures) confidence intervals. Such a program is beyond the scope of the present paper.

7 An Empirical Application

Lieberman and Phillips (2017b) estimated a STUR model in which the dependent variable is the log spread between an index of U.S. dollar denominated investment grade rated corporate debt publicly issued in the U.S. domestic market and the spot Treasury curve. The variable u_t was taken to be the demeaned $100 \log(SP_{US,t}/SP_{US,t-1})$, where $SP_{US,t}$ is the opening rate of the SPDR S&P 500 ETF Trust. The sample correlation between u_t and ΔY was -0.52, supporting Kwan's (1996) report of a negative correlation between stock returns and bond spread changes. In this case the NLLS estimator is inconsistent. The IV estimator, which is consistent, was estimated with 1454 daily observations over the period January 5, 2010, through to December 30, 2015, giving a value $\hat{a}_n^{IV} = -0.245$. In addition, the misspecified STUR-based IV estimators of the covariance parameters are consistent as discussed in Section 5. Using these results we calculated the t-statistic (28) with error variance estimated by $\tilde{\gamma}_{\varepsilon,n}^{IV}(0)$ obtaining a value of $t_{\hat{\beta}} = -0.659$. The 5th, 10th, 50th, 90th and 95th percentiles of the asymptotic distribution were simulated[2] using (28), with parameters replaced by their IV-consistent estimates. The 90% CI for c is given by the intersection of the horizontal line $t_{\hat{\beta}} = -0.659$ and the 5th and 95th percentiles lines, in the $\left(c, t_{\hat{\beta}}\right)$ plane, as shown in Figure 1, yielding the CI lower and upper limits $c_L^a = -0.64$ and $c_U^a = 0.53$. The intersection points of $t_{\hat{\beta}} = -0.659$ with the percentiles are summarized in Table 2, from which we deduce that the median unbiased estimate of c in the LSTUR model is $\hat{c}_{med} = -0.21$. The procedure was repeated for the LUR model, where the asymptotic distribution is

[2] A MATHEMATICA program was written to evaluate the percentiles using 400 integration points, 5000 replications and a grid of 0.1 over the c-values.

given by (28), with $a = 0$, $\sigma_u^2 = 0$, $\Sigma_{u\varepsilon} = 0$. The results are shown in in Figure 2 and Table 2. For this model we obtain the 90% CI limits $c_L = -4.05$ and $c_U = 3.27$, and a mean unbiased estimator for c equal to -0.35.

Figures 1-2 as well as Table 2 reveal that the 90% CI for c, which is LSTUR-based, is much narrower and is in fact fully within the 90% LUR-based CI. Thus, at least in this case, LSTUR attenuates the estimated impact of c on the time varying autoregressive coefficient β_{nt}. The induced 90% CI for β_{nt} which is LUR-based is approximately $[1 - 4.05/n, 1 + 3.27/n]$, whereas the variation of u_t needs to be accounted for in the construction of an LSTUR-based 90% CI for β_{nt}. Conditional on u_t and on the values of the nuisance parameters, the LSTUR-based 90% CI for β_{nt} is $\left[e^{-0.64/n + au_t/\sqrt{n}}, e^{0.53/n + au_t/\sqrt{n}}\right]$, so that the width of the interval is approximately $1.17/n$, compared with a width of $7.31/n$ for the LUR-based CI. The means of the CI bounds, taken with respect to u_t and assuming that w_t is multivariate normal, are $Ee^{c_L^a + au_t/\sqrt{n}} = e^{\left(c_L^a + b/2\right)/n}$ and $Ee^{c_U^a + au_t/\sqrt{n}} = e^{\left(c_U^a + b/2\right)/n}$. Plugging in the IV estimates, $\hat{a}_n^{IV} = -0.245$ and $\hat{\sigma}_u^2 = n^{-1}\sum u_t^2 = 0.983$[3] into these formulae, the estimated means of the bounds are $1 - 0.61/n$ and $1 + 0.56/n$, which are much smaller in absolute values than the respective LUR-based bounds. Furthermore, Given the model parameters, and assuming that w_t is multivariate normal,

$$\Pr\left(e^{\frac{c}{n} + \frac{au_t}{\sqrt{n}}} < L\right) = \alpha \;\text{ iff } L = e^{\frac{c}{n} + \frac{c_\alpha \sqrt{b}}{\sqrt{n}}},$$

where c_α is the α'th percentile of the standard normal distribution. Thus, given the model parameters and the distribution of w_t, the induced 90% CI for β_{nt} is

$$\left[e^{-\frac{0.64}{n} - \frac{0.4}{\sqrt{n}}}, e^{\frac{0.53}{n} + \frac{0.4}{\sqrt{n}}}\right].$$

So, the width of the CI is approximately $0.8/\sqrt{n} + 1.17/n$. Compared with the LUR-based induced CI for β_{nt}, the LSTUR-based induced CI has a term which is $O\left(n^{-1/2}\right)$, to account for the additional variability in β_{nt} which is due to u_t. On the other hand, the $O\left(n^{-1}\right)$ term in the CI which is due to c and b in LSTUR and due to c only in LUR, is much smaller in absolute value in the LSTUR-based bounds than in LUR. These findings are illustrated in Figure 3.

We remark that an 'exact' analytical CI which accounts for the variability in the estimates of a and the covariance parameters is analytically intractable, because these estimates influence both the percentiles of $t_{\hat{\beta}}$ (and, hence, the values c_L^a and c_U^a) as well as the summand $\hat{a}_n^{IV} u_t/\sqrt{n}$. Nevertheless, qualitatively, the message from the empirical application is that the reported CI for c can be substantially wrong and, in reality, much wider when an LSTUR process is misspecified as a LUR model. On the other hand, unconditionally, the induced CI for β_{nt} is wider when a STUR component is present as is expected from the additional random variability that is embodied in the LSTUR representation of the time variation in the autoregressive coefficient.

8 Discussion

It is widely acknowledged that with much economic and financial data the unit root hypothesis may only hold approximately or in some sense on average over a given sample. A more general modeling perspective that offers greater flexibility is that the generating mechanism may involve temporary de-

[3] The variable u_t is demeaned and its standard variance estimator is consistent as it does not depend on a.

partures from unity at any sample point that can move the process in stationary or explosive directions. Recognition of this type of functional coefficient flexibility and its relevance for applied work has led to the literature on LUR, functional LUR (Bykovskaya and Phillips, 2017a, 2017b), and STUR models, which seek to capture certain non-random and random departures from an autoregressive unit root process. The hybrid model introduced in this paper incorporates two streams of this literature as special cases and the limit theory generalizes results already known for the LUR and STUR models. As expected, ignoring one or other of these component departures introduces inferential bias. Both simulations and empirics reveal how the construction of uniform confidence intervals for autoregressive coefficients using a LUR model formulation are affected by misspecification in which the random departures of the LSTUR mechanism are neglected. Of particular relevance in applications is the fact that an LSTUR process, may have the same mean and variance as a Gaussian random walk but with kurtosis that is well in excess of 3, a feature that is consonant with the heavy tails of much observed financial return data.

References

Anderson, T. W. (1959). "On asymptotic distributions of estimates of parameters of stochastic difference equations," *The Annals of Mathematical Statistics* 30(3), 676–687.

Bykhovskaya, A. and P. C. B. Phillips "Point Optimal Testing with Roots that are Functionally Local to Unity", Cowles Foundation Discussion Paper #3007, Yale University.

Bykhovskaya, A. and P. C. B. Phillips "Boundary Limit Theory for Functional Local to Unity Regression", Cowles Foundation Discussion Paper #3008, Yale University.

Campbell, J.Y. and M. Yogo (2006). "Efficient tests of stock return predictability," *Journal of Financial Economics* 81(1), 27–60.

Chan, N. H. and C. Z. Wei (1987). "Asymptotic inference for nearly nonstationary AR(1) processes," *Annals of Statistics* 15, 1050–1063.

Evans, G. B. A. and N. E. Savin (1981). "Testing for a unit root: 1" *Econometrica* 49(3), 753–779.

Elliott, G., T. J. Rothenberg and J. H. Stock (1996). "Efficient tests of an autoregressive unit root," *Econometrica,* Vol. 64, 813 - 836.

Föllmer, H. and M. Schweizer (1993). "A microeconomic approach to diffusion models for stock prices," *Mathematical Finance* 3(1), 1–23.

Giraitis, L. and P. C. B. Phillips (2006) "Uniform limit theory for stationary autoregression," *Journal of Time Series Analysis* 27, 51–60.

Granger, C. W. and N. R. Swanson (1997). "An introduction to stochastic unit-root processes," *Journal of Econometrics 80(1), 35–62.*

Ibragimov, R. and P. C. B. Phillips (2008). "Regression asymptotics using martingale convergence methods," *Econometric Theory* 24, 888–947.

Kwan, S. H. (1996). "Firm-specific information and the correlation between individual stocks and bonds," *Journal of Financial Economics,* 40, 63–80.

Leeb H. and B. Pötscher (2005). "Model selection and inference: facts and fiction," *Econometric Theory*, 21, pp. 21-59.

Leybourne, S. J., McCabe, B. P., and T. C. Mills (1996). "Randomized unit root processes for modelling and forecasting financial time series: theory and applications," *Journal of Forecasting* 15(3), 253-270.

Leybourne, S. J., McCabe, B. P. and A. R. Tremayne (1996). "Can economic time series be differenced to stationarity?," *Journal of Business and Economic Statistics* 14(4), 435-446.

Lieberman, O. (2012). "A similarity-based approach to time-varying coefficient non-stationary autoregression," *Journal of Time Series Analysis* 33, 484–502.

Lieberman, O. and P. C. B. Phillips (2014). "Norming rates and limit theory for some time-varying coefficient autoregressions," *Journal of Time Series Analysis* 35, 592–623.

Lieberman, O. and P. C. B. Phillips (2017a). "A multivariate stochastic unit root model with an application to derivative pricing," *Journal of Econometrics,* 196, 99-110.

Lieberman, O. and P. C. B. Phillips (2017b). "IV and GMM Inference in Endogenous Stochastic Unit Root Models". *Econometric Theory,* doi.org/10.1017/S0266466617000330.

McCabe, B. P. and R. J. Smith (1998). "The power of some tests for difference stationarity under local heteroscedastic integration," *Journal of the American Statistical Association* 93(442), 751-761.

Mikusheva, A. (2007). "Uniform inference in autoregressive models," *Econometrica* 75(5), 1411–1452.

Mikusheva, A. (2012). "One-dimensional inference in autoregressive models with the potential presence of a unit root," *Econometrica* 80(1), 173–212.

Phillips, P. C. B. (1987). "Towards a unified asymptotic theory for autoregression," *Biometrika* 74, 535–547.

Phillips, P. C. B. (1977). "Approximations to Some Finite Sample Distributions Associated with a First Order Stochastic Difference Equation," *Econometrica* 45:2, 463–485.

Phillips, P. C. B. (1978). "Edgeworth and saddlepoint approximations in a first order non-circular autoregression," *Biometrika* 65:1, 91–98.

Phillips, P. C. B. (2014). "On confidence intervals for autoregressive roots and predictive regression," *Econometrica* 82(3), 1177-1195.

Phillips, P. C. B. and T. Magdalinos (2007a). "Limit theory for moderate deviations from a unit root," *Journal of Econometrics* 136, 115-130.

Phillips, P. C. B., T. Magdalinos (2007b), "Limit Theory for Moderate Deviations from a Unit Root Under Weak Dependence," in G. D. A. Phillips and E. Tzavalis (Eds.) *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis.* Cambridge: Cambridge University Press, pp.123-162.

Ploberger, W. and P. C. B. Phillips (2003). "Empirical limits for time series econometric models" *Econometrica,* 71, 627-673.

Stock, J. H. (1991). "Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series," *Journal of Monetary Economics* 28(3), 435–459.

Tao, Y., P. C. B. Phillips and J. Yu (2017). "Random coefficient continuous systems: testing for extreme sample path behavior", Working paper, Singapore Management University.

White, J. S. (1958). "The limiting distribution of the serial correlation coefficient in the explosive case," *The Annals of Mathematical Statistics* 29(4), 1188-1197.

Yoon, G. (2006). "A note on some properties of STUR processes," *Oxford Bulletin of Economics and Statistics* 68(2), 253-260.

Table 1. Percentiles and Confidence Intervals for $t_{\hat{\beta}}$, $\sigma_u^2 = 0.1$

c	a	5	10	50	90	95	80% CI	90% CI
1	0	-1.724	-1.361	0.3	2.466	3.144	3.827	4.868
1	1	-1.73	-1.355	0.273	2.588	3.423	3.943	5.152
1	2	-1.716	-1.367	0.151	2.935	4.199	4.302	5.915
1	3	-1.807	-1.4917	-0.1127	3.988	6.432	5.479	8.239
1	4	-1.974	-1.641	-0.467	4.921	9.538	6.562	11.512
0	0	-1.913	-1.62	-0.508	0.9445	1.311	2.564	3.224
0	1	-1.919	-1.607	-0.534	0.907	1.374	2.514	3.293
0	2	-1.976	-1.641	-0.562	1.068	1.753	2.71	3.729
0	3	-2.03	-1.712	-0.736	1.25	2.264	2.962	4.294
0	4	-2.054	-1.817	-0.94	1.882	3.437	3.698	5.492
-1	0	-2.145	-1.852	-0.877	0.125	0.432	1.977	2.576
-1	1	-2.136	-1.812	-0.887	0.101	0.472	1.914	2.608
-1	2	-2.195	-1.889	-0.955	0.171	0.592	2.059	2.786
-1	3	-2.195	-1.905	-1.021	0.21	0.862	2.115	3.057
-1	4	-2.255	-1.986	-1.187	0.32	1.277	2.306	3.532
-5	0	-2.746	-2.472	-1.642	-0.996	-0.81	1.476	1.936
-5	1	-2.755	-2.501	-1.656	-0.996	-0.802	1.505	1.953
-5	2	-2.757	-2.496	-1.664	-1.01	-0.819	1.486	1.938
-5	3	-2.788	-2.52	-1.7	-1.023	-0.804	1.497	1.983
-5	4	-2.875	-2.556	-1.755	-1.102	-0.838	1.454	2.037
-10	0	-3.283	-3.047	-2.257	-1.628	-1.468	1.419	1.815
-10	1	-3.274	-3.026	-2.269	-1.63	-1.463	1.396	1.811
-10	2	-3.297	-3.053	-2.285	-1.642	-1.475	1.411	1.821
-10	3	-3.308	-3.064	-2.272	-1.624	-1.462	1.44	1.846
-10	4	-3.35	-3.116	-2.299	-1.685	-1.496	1.431	1.854

Note: The entries in the table are the percentiles- and confidence interval width (last two columns) of the limit distribution of the statistic $t_{\hat{\beta}}$, based on 5000 replications and 400 integral points, with $\sigma_\varepsilon^2 = 1$, $\Sigma_{u\varepsilon} = 0$.

Table 1 (continued). Percentiles and Confidence Intervals for $t_{\hat{\beta}}$, $\sigma_u^2 = 0.1$

c	a	5	10	50	90	95	80% CI	90% CI
-15	0	−3.744	−3.514	−2.738	−2.093	−1.917	1.421	1.827
-15	1	−3.756	−3.51	−2.756	−2.099	−1.941	1.411	1.815
-15	2	−3.74	−3.522	−2.75	−2.082	−1.931	1.44	1.81
-15	3	−3.754	−3.488	−2.763	−2.103	−1.918	1.385	1.835
-15	4	−3.792	−3.547	−2.762	−2.122	−1.958	1.424	1.834
-20	0	−4.11	−3.885	−3.104	−2.462	−2.3	1.424	1.81
-20	1	−4.108	−3.909	−3.146	−2.488	−2.332	1.421	1.776
-20	2	−4.15	−3.923	−3.136	−2.485	−2.315	1.438	1.835
-20	3	−4.144	−3.896	−3.143	−2.474	−2.31	1.422	1.834
-20	4	−4.192	−3.938	−3.159	−2.492	−2.331	1.446	1.861
-25	0	−4.469	−4.236	−3.48	−2.785	−2.627	1.451	1.842
-25	1	−4.46	−4.24	−3.452	−2.79	−2.645	1.449	1.815
-25	2	−4.473	−4.225	−3.456	−2.784	−2.602	1.441	1.872
-25	3	−4.491	−4.218	−3.469	−2.793	−2.628	1.425	1.863
-25	4	−4.507	−4.247	−3.469	−2.818	−2.641	1.428	1.866
-30	0	−4.74	−4.512	−3.75	−3.073	−2.895	1.44	1.845
-30	1	−4.756	−4.54	−3.757	−3.087	−2.935	1.453	1.821
-30	2	−4.744	−4.52	−3.768	−3.097	−2.926	1.423	1.818
-30	3	−4.759	−4.524	−3.768	−3.102	−2.926	1.421	1.833
-30	4	−4.78	−4.554	−3.777	−3.093	−2.909	1.46	1.871
-35	0	−5.038	−4.794	−4.012	−3.325	−3.131	1.469	1.906
-35	1	−5.013	−4.802	−4.034	−3.357	−3.171	1.446	1.842
-35	2	−5.077	−4.812	−4.037	−3.357	−3.193	1.455	1.884
-35	3	−5.044	−4.811	−4.036	−3.357	−3.177	1.454	1.868
-35	4	−5.021	−4.804	−4.039	−3.337	−3.165	1.466	1.857

Note: The entries in the table are the percentiles- and confidence interval width (last two columns) of the limit distribution of the statistic $t_{\hat{\beta}}$, based on 5000 replications and 400 integral points, with $\sigma_\varepsilon^2 = 1$, $\Sigma_{u\varepsilon} = 0$.

Table 1 (continued). Percentiles and Confidence Intervals for $t_{\hat{\beta}}$, $\sigma_u^2 = 1$

c	a	5	10	50	90	95	80% CI	90% CI
1	0	−1.685	−1.329	0.316	2.427	3.058	3.756	4.743
1	1	−1.789	−1.469	−0.1421	4.009	6.849	5.478	8.638
1	2	−8.473	−4.854	−1.68	7.822	22.755	12.676	31.228
1	3	−110.893	−46.739	−4.215	9.609	54.083	56.348	164.976
1	4	−1398.42	−423.387	−13.242	13.956	142.768	437.343	1541.18
0	0	−2.0288	−1.652	−0.5106	0.9	1.286	2.552	3.315
0	1	−2.01	−1.701	−0.746	1.333	2.599	3.034	4.609
0	2	−6.253	−3.799	−1.816	2.35	7.526	6.148	13.779
0	3	−68.45	−29.095	−3.747	3.687	25.273	32.782	93.723
0	4	−736.13	−233.141	−11.1434	3.057	63.583	236.199	799.713
-1	0	−2.114	−1.823	−0.865	0.102	0.414	1.925	2.528
-1	1	−2.159	−1.866	−1.014	0.217	0.891	2.083	3.05
-1	2	−4.517	−3.164	−1.887	0.458	3.22	3.62	7.737
-1	3	−42.348	−19.889	−3.537	0.222	7.327	20.111	49.676
-1	4	−397.271	−142.957	−8.921	−0.499	20.427	142.459	417.698
-5	0	−2.731	−2.473	−1.646	−1.019	−0.832	1.454	1.899
-5	1	−2.767	−2.494	−1.716	−1.04	−0.828	1.454	1.939
-5	2	−3.067	−2.808	−2.103	−1.345	−0.964	1.462	2.1
-5	3	−9.954	−6.511	−3.1	−1.879	−0.91	4.632	9.043
-5	4	−63.291	−28.508	−5.263	−2.579	−0.089	25.928	63.202
-10	0	−3.316	−3.068	−2.279	−1.622	−1.47	1.446	1.846
-10	1	−3.317	−3.071	−2.305	−1.65	−1.463	1.422	1.853
-10	2	−3.45	−3.191	−2.437	−1.794	−1.601	1.398	1.849
-10	3	−4.579	−3.973	−3.018	−2.189	−1.898	1.784	2.681
-10	4	−16.424	−10.077	−4.282	−2.789	−2.23	7.287	14.194

Note: The entries in the table are the percentiles- and confidence interval width (last two columns) of the limit distribution of the statistic $t_{\hat{\beta}}$, based on 5000 replications and 400 integral points, with $\sigma_\varepsilon^2 = 1$, $\Sigma_{u\varepsilon} = 0$.

Table 1 (continued). Percentiles and Confidence Intervals for $t_{\hat{\beta}}$, $\sigma_u^2 = 1$

c	a	5	10	50	90	95	80% CI	90% CI
-15	0	-3.7256	-3.476	-2.728	-2.085	-1.93	1.391	1.791
-15	1	-3.721	-3.501	-2.743	-2.093	-1.923	1.408	1.798
-15	2	-3.879	-3.593	-2.828	-2.181	-2.01	1.412	1.869
-15	3	-4.229	-3.968	-3.177	-2.418	-2.143	1.55	2.086
-15	4	-7.522	-5.798	-3.999	-2.866	-2.476	2.932	5.047
-20	0	-4.127	-3.879	-3.112	-2.48	-2.313	1.4	1.815
-20	1	-4.148	-3.908	-3.129	-2.4891	-2.323	1.419	1.825
-20	2	-4.213	-3.975	-3.185	-2.528	-2.359	1.447	1.853
-20	3	-4.471	-4.22	-3.433	-2.69	-2.468	1.53	2.003
-20	4	-5.761	-5.154	-3.958	-3.002	-2.699	2.152	3.063
-25	0	-4.451	-4.235	-3.477	-2.796	-2.629	1.439	1.821
-25	1	-4.504	-4.26	-3.473	-2.808	-2.622	1.452	1.883
-25	2	-4.526	-4.292	-3.502	-2.827	-2.647	1.465	1.879
-25	3	-4.675	-4.449	-3.686	-2.911	-2.714	1.539	1.961
-25	4	-5.43	-5.084	-4.104	-3.19	-2.901	1.894	2.53
-30	0	-4.796	-4.538	-3.768	-3.069	-2.91	1.468	1.886
-30	1	-4.828	-4.555	-3.773	-3.085	-2.914	1.471	1.914
-30	2	-4.815	-4.552	-3.791	-3.109	-2.916	1.443	1.899
-30	3	-4.977	-4.732	-3.948	-3.194	-2.973	1.538	2.003
-30	4	-5.419	-5.148	-4.212	-3.343	-3.073	1.804	2.346
-35	0	-5.043	-4.819	-4.037	-3.354	-3.184	1.465	1.859
-35	1	-5.064	-4.809	-4.029	-3.336	-3.16	1.473	1.903
-35	2	-5.1	-4.879	-4.083	-3.359	-3.165	1.52	1.935
-35	3	-5.249	-5.014	-4.177	-3.418	-3.218	1.596	2.031
-35	4	-5.605	-5.305	-4.406	-3.562	-3.317	1.743	2.288

Note: The entries in the table are the percentiles- and confidence interval width (last two columns) of the limit distribution of the statistic $t_{\hat{\beta}}$, based on 5000 replications and 400 integral points, with $\sigma_\varepsilon^2 = 1$, $\Sigma_{u\varepsilon} = 0$.

Table 2. Intersections of $t_{\hat{\beta}}$ with confidence lines.

Model	Percentiles 5th	10th	50th	90th	95th
LUR	3.268	2.367	-0.347	-3.260	-4.047
LSTUR	0.532	0.442	-0.213	-0.546	-0.640

Note: The figures in the Table are the intersections of the line $t_{\hat{\beta}} = -0.659$ with the confidence lines for the LUR and LSTUR models.

9 Proofs

9.1 Proofs of Lemmas and Supplementary Results

Proof of Lemma 2. From (7), $\frac{1}{n}\sum_{t=2}^{n}\varepsilon_t Y_{t-1} = \frac{1}{n}\sum_{t=2}^{n}\sum_{j=0}^{\infty} G_{2,j}\eta_{t-j}Y_{t-1}$. As Y_{t-1} is uncorrelated with η_t,

$$\frac{1}{n}G_{2,0}\sum_{t=2}^{n}\eta_t Y_{t-1} \Rightarrow G_{2,0}\int_0^1 G_{a,c}(r)\,dB_\eta(r)\,. \tag{29}$$

Next decompose the contemporaneous sample covariance as

$$\begin{aligned}\frac{1}{n}G_{2,1}\sum_t \eta_{t-1}Y_{t-1} &= \frac{1}{n}G_{2,1}\sum_t \eta_{t-1}\left(\exp\left(\frac{c}{n}+\frac{a'u_{t-1}}{\sqrt{n}}\right)Y_{t-2}+\varepsilon_{t-1}\right)\\ &= \frac{1}{n}G_{2,1}\sum_t \eta_{t-1}\left\{\left(1+\frac{a'u_{t-1}}{\sqrt{n}}+o_p\left(n^{-1/2}\right)\right)Y_{t-2}+\varepsilon_{t-1}\right\},\end{aligned} \tag{30}$$

where $\frac{1}{n}\sum_t \eta_{t-1}Y_{t-2} \Rightarrow \int_0^1 G_{a,c}(r)\,dB_\eta(r)$ from (29) and

$$\begin{aligned}&\frac{1}{n^{3/2}}\sum_t \eta_{t-1}a'u_{t-1}Y_{t-2} = \frac{1}{n^{3/2}}\sum_{t=2}^{\infty}\eta_{t-1}\left(\sum_{j=0}^{\infty}G_{1,j}\eta_{t-1-j}\right)' aY_{t-2}\\ &= \frac{1}{n^{3/2}}\sum_{t=2}^{\infty}\eta_{t-1}\eta_{t-1}'G_{1,0}'aY_{t-2} + \frac{1}{n^{3/2}}\sum_{t=2}^{\infty}\eta_{t-1}\sum_{j=1}^{\infty}\eta_{t-1-j}'G_{1,j}'a\left(\exp\left(\frac{c}{n}+\frac{a'u_{t-2}}{\sqrt{n}}\right)Y_{t-3}+\varepsilon_{t-2}\right)\\ &= \Sigma_\eta G_{1,0}'a\int_0^1 G_{a,c}(r)\,dr + \frac{1}{n^{3/2}}\sum_{t=2}^{\infty}\eta_{t-1}\sum_{j=1}^{\infty}\eta_{t-1-j}'G_{1,j}'a\varepsilon_{t-2} + o_p(1)\,.\end{aligned} \tag{31}$$

Further,

$$\begin{aligned}&\frac{1}{n^{3/2}}\sum_{t=2}^{\infty}\eta_{t-1}\sum_{j=1}^{\infty}\eta_{t-1-j}'G_{1,j}'a\varepsilon_{t-2} = \frac{1}{n^{3/2}}\sum_{t=2}^{\infty}\eta_{t-1}\sum_{j=1}^{\infty}\eta_{t-1-j}'G_{1,j}'a\sum_{k=0}^{\infty}G_{2,k}\eta_{t-2-k}\\ &= \frac{1}{n^{3/2}}\sum_{t=2}^{\infty}\eta_{t-1}\sum_{j=1}^{\infty}\sum_{k=0}^{\infty}G_{2,k}\eta_{t-2-k}\eta_{t-1-j}'G_{1,j}'a = O_p\left(n^{-1}\right),\end{aligned}$$

so that

$$\frac{1}{n^{3/2}}\sum_t \eta_{t-1}a'u_{t-1}Y_{t-2} \Rightarrow \Sigma_\eta G'_{1,0}a\int_0^1 G_{a,c}(r)\,dr. \tag{32}$$

The last non-vanishing term in (30) involves

$$\frac{1}{n}\sum_{t=2}^{n}\eta_{t-1}\varepsilon_{t-1} = \frac{1}{n}\sum_{t=2}^{n}\eta_{t-1}\sum_{k=0}^{\infty}G_{2,k}\eta_{t-1-k} \to_p \Sigma_\eta G'_{2,0}, \tag{33}$$

and so

$$\frac{1}{n}G_{2,1}\sum_t \eta_{t-1}Y_{t-1} \Rightarrow G_{2,1}\left(\int_0^1 G_{a,c}(r)\,dB_\eta(r) + \Sigma_\eta G'_{1,0}a\int_0^1 G_{a,c}(r)\,dr + \Sigma_\eta G'_{2,0}\right). \tag{34}$$

Continuing,

$$\begin{aligned}
Y_{t-1} &= \exp\left(\frac{c}{n} + \frac{a'u_{t-1}}{\sqrt{n}}\right)Y_{t-2} + \varepsilon_{t-1} \\
&= \exp\left(\frac{c}{n} + \frac{a'u_{t-1}}{\sqrt{n}}\right)\left(\exp\left(\frac{c}{n} + \frac{a'u_{t-2}}{\sqrt{n}}\right)Y_{t-3} + \varepsilon_{t-2}\right) + \varepsilon_{t-1} \\
&= \exp\left(\frac{2c}{n} + \frac{a'(u_{t-1}+u_{t-2})}{\sqrt{n}}\right)Y_{t-3} + \exp\left(\frac{c}{n} + \frac{a'u_{t-1}}{\sqrt{n}}\right)\varepsilon_{t-2} + \varepsilon_{t-1}. \qquad (35)
\end{aligned}$$

We therefore have

$$\frac{1}{n}G_{2,2}\sum_t \eta_{t-2}Y_{t-1} = \frac{1}{n}G_{2,2}\sum_t \eta_{t-2}\left(\exp\left(\frac{2c}{n} + \frac{a'(u_{t-1}+u_{t-2})}{\sqrt{n}}\right)Y_{t-3} + \exp\left(\frac{c}{n} + \frac{a'u_{t-1}}{\sqrt{n}}\right)\varepsilon_{t-2} + \varepsilon_{t-1}\right) \tag{36}$$

$$= \frac{1}{n}G_{2,2}\sum_t \eta_{t-2}\left(1 + \frac{a'(u_{t-1}+u_{t-2})}{\sqrt{n}}\right)Y_{t-3} \tag{37}$$

$$+ \frac{1}{n}G_{2,2}\sum_t \eta_{t-2}\exp\left(\frac{c}{n} + \frac{a'u_{t-1}}{\sqrt{n}}\right)\varepsilon_{t-2} \tag{38}$$

$$+ \frac{1}{n}G_{2,2}\sum_t \eta_{t-2}\varepsilon_{t-1} + o_p(1). \tag{39}$$

To deal with (37), we write

$$\frac{1}{n}\sum_t \eta_{t-2}\left(1 + \frac{a'(u_{t-1}+u_{t-2})}{\sqrt{n}}\right)Y_{t-3} = \int_0^1 G_{a,c}(r)\,dB_\eta(r) + \frac{1}{n^{3/2}}\sum_t \eta_{t-2}a'(u_{t-1}+u_{t-2})Y_{t-3} + o_p(1), \tag{40}$$

and using (32) gives

$$\frac{1}{n^{3/2}}\sum_t \eta_{t-2}a'u_{t-2}Y_{t-3} \Rightarrow \Sigma_\eta G'_{1,0}a\int_0^1 G_{a,c}(r)\,dr, \tag{41}$$

so that

$$\frac{1}{n^{3/2}}\sum_t \eta_{t-2}a'u_{t-1}Y_{t-3} = \frac{1}{n^{3/2}}\sum_t \eta_{t-2}\left(\sum_{j=0}^{\infty} G_{1,j}\eta_{t-1-j}\right)' aY_{t-3} = \frac{1}{n^{3/2}}\sum_t \eta_{t-2}\eta'_{t-1}G'_{1,0}aY_{t-3}$$

$$+\frac{1}{n^{3/2}}\sum_t \eta_{t-2}\eta'_{t-2}G'_{1,1}aY_{t-3} + \frac{1}{n^{3/2}}\sum_t \eta_{t-2}\sum_{j=3}^{\infty}\eta'_{t-j}G'_{1,j-1}aY_{t-3}. \tag{42}$$

Two types of terms occur in (42): one with equal lags of η_t and the other with non-equal lags. Since $Y_{t-3} = \exp\left(\frac{c}{n} + \frac{a'u_{t-3}}{\sqrt{n}}\right)Y_{t-4} + \varepsilon_{t-3}$, the first term is

$$\frac{1}{n^{3/2}}\sum_t \eta_{t-2}\eta'_{t-1}G'_{1,0}aY_{t-3} = \frac{1}{n^{3/2}}\sum_t \eta_{t-2}\eta'_{t-1}G'_{1,0}a\varepsilon_{t-3} + o_p(1)$$

$$= \frac{1}{n^{3/2}}\sum_t \eta_{t-2}\eta'_{t-1}G'_{1,0}a\left(G_{2,0}\eta_{t-3} + \cdots\right) + o_p(1) = o_p(1). \tag{43}$$

The second term in (42) gives

$$\frac{1}{n^{3/2}}\sum_t \eta_{t-2}\eta'_{t-2}G'_{1,1}aY_{t-3} \Rightarrow \Sigma_\eta G'_{1,1}a\int_0^1 G_{a,c}(r)\,dr, \tag{44}$$

and the third term in (42) is

$$\frac{1}{n^{3/2}}\sum_t \eta_{t-2}\sum_{j=3}^{\infty}\eta'_{t-j}G'_{1,j-1}aY_{t-3} = \frac{1}{n^{3/2}}\sum_t \eta_{t-2}\sum_{j=3}^{\infty}\eta'_{t-j}G'_{1,j-1}a\varepsilon_{t-3} + o_p(1) = o_p(1). \tag{45}$$

It follows from (42)-(45) that

$$\frac{1}{n^{3/2}}\sum_t \eta_{t-2}a'u_{t-1}Y_{t-3} \Rightarrow \Sigma_\eta G'_{1,1}a\int_0^1 G_{a,c}(r)\,dr. \tag{46}$$

Then, from (40), (41) and (46) we deduce that

$$\frac{1}{n}\sum_t \eta_{t-2}\left(1 + \frac{a'\left(u_{t-1}+u_{t-2}\right)}{\sqrt{n}}\right)Y_{t-3} \Rightarrow \int_0^1 G_{a,c}(r)\,dB_\eta(r) + \Sigma_\eta \sum_{j=0}^{1} G'_{1,j}a\int_0^1 G_{a,c}(r)\,dr. \tag{47}$$

Using (33), we have

$$\frac{1}{n}\sum_t \eta_{t-2}\exp\left(\frac{c}{n} + \frac{a'u_{t-1}}{\sqrt{n}}\right)\varepsilon_{t-2} \Rightarrow \Sigma_\eta G'_{2,0}, \tag{48}$$

and

$$\frac{1}{n}\sum_t \eta_{t-2}\varepsilon_{t-1} = \frac{1}{n}\sum_t \eta_{t-2}\sum_{j=0}^{\infty}\eta'_{t-1-j}G'_{2,j}\varepsilon_{t-1} \Rightarrow \Sigma_\eta G'_{2,1}. \tag{49}$$

Combining results from (35), (47), (48), and (49) gives

$$\frac{1}{n}G_{2,2}\sum_t \eta_{t-2}Y_{t-1} \Rightarrow G_{2,2}\left(\int_0^1 G_{a,c}(r)\,dB_\eta(r) + \Sigma_\eta \sum_{j=0}^{1} G'_{1,j}a\int_0^1 G_{a,c}(r)\,dr + \Sigma_\eta \sum_{j=0}^{1} G'_{2,j}\right) \quad (50)$$

In view of (29), (34) and (50), we have

$$\begin{aligned}\frac{1}{n}\sum_{t=2}^{n}\varepsilon_t Y_{t-1} &= \frac{1}{n}\sum_{t=2}^{n}\sum_{j=0}^{\infty} G_{2,j}\eta_{t-j}Y_{t-1} \Rightarrow \sum_{m=0}^{\infty} G_{2,m}\int_0^1 G_{a,c}(r)\,dB_\eta(r)\\ &+ \sum_{m=1}^{\infty} G_{2,m}\left(\Sigma_\eta \sum_{j=0}^{m-1} G'_{1,m}a\int_0^1 G_{a,c}(r)\,dr + \Sigma_\eta \sum_{j=0}^{m-1} G'_{2,j}\right).\end{aligned}$$

Now, $\sum_{m=0}^{\infty} G_{2,m} = G_2(1)$ and B_η has variance matrix Σ_η, so that $G_2(1)\,B_\eta(r) = B_\varepsilon(r)$ has variance matrix $G_2(1)\,\Sigma_\eta G_2(1)'$ and

$$\left(\sum_{m=0}^{\infty} G_{2,m}\right)\int_0^1 G_{a,c}(r)\,dB_\eta(r) = \int_0^1 G_{a,c}(r)\,dB_\varepsilon(r).$$

Further, recalling that $\Lambda_{u\varepsilon} = \sum_{h=1}^{\infty} E(u_0\varepsilon_h)$, we obtain

$$\begin{aligned}&\Lambda'_{u\varepsilon}a = E\left\{\left(G_{1,0}\eta_t + G_{1,1}\eta_{t-1} + \cdots\right)\left[\left(G_{2,0}\eta_{t+1} + G_{2,1}\eta_t + \cdots\right) + \left(G_{2,0}\eta_{t+2} + G_{2,1}\eta_{t+1} + \cdots\right)\right]\right\}' a\\ &= \left\{G_{1,0}\Sigma_\eta\left(G_{2,1} + G_{2,2} + \cdots\right)' + G_{1,1}\Sigma_\eta\left(G_{2,2} + G_{2,3} + \cdots\right)' + \cdots\right\}' a\\ &= \left\{G_{1,0}\Sigma_\eta G'_{2,1} + \left(G_{1,0} + G_{1,1}\right)\Sigma_\eta G'_{2,2}\cdots\right\}' a\\ &= \left\{G_{2,1}\Sigma_\eta G'_{1,0} + G_{2,2}\Sigma_\eta\left(G_{1,0} + G_{1,1}\right)'\cdots\right\} a\\ &= \sum_{j=1}^{\infty} G_{2,j}\Sigma_\eta\left(\sum_{k=0}^{j-1} G'_{1,k}\right)' a. \end{aligned} \quad (51)$$

Similarly, $\lambda_{\varepsilon\varepsilon} = E\sum_{h=1}^{\infty}\varepsilon_t\varepsilon_{t+h}$ is

$$\begin{aligned}& E\left\{\left(G_{2,0}\eta_t + G_{2,1}\eta_{t-1} + \cdots\right)\left[\left(G_{2,0}\eta_{t+1} + G_{2,1}\eta_t + \cdots\right) + \left(G_{2,0}\eta_{t+2} + G_{2,1}\eta_{t+1} + \cdots\right) + \cdots\right]'\right\}\\ =\ & G_{2,0}\Sigma_\eta\left(G_{2,1} + G_{2,2} + \cdots\right)' + G_{2,1}\Sigma_\eta\left(G_{2,2} + G_{2,3} + \cdots\right)' + \cdots\\ =\ & G_{2,0}\Sigma_\eta {G_{2,1}}' + \left(G_{2,0} + G_{2,1}\right)\Sigma_\eta G'_{2,2} + \cdots\\ =\ & G_{2,1}\Sigma_\eta G'_{2,0} + G_{2,2}\Sigma_\eta\left(G_{2,0} + G_{2,1}\right)' + \cdots = \sum_{j=1}^{\infty} G_{2,j}\Sigma_\eta\left(\sum_{k=0}^{j-1} G'_{2,k}\right). \end{aligned} \quad (52)$$

This implies the result given in (12) and the Lemma is established. ■

Lemma 9 *For the model (4), under Assumption 1,*

$$\frac{1}{n^2}\sum_t u_t u_t' Y_{t-1}^2 \Rightarrow \Sigma_u \int_0^1 G_{a,c}^2(r)\,dr.$$

Proof of Lemma 9. The result follows immediately from

$$\frac{1}{n^2}\sum_t u_t u_t' Y_{t-1}^2 = \Sigma_u \frac{1}{n}\sum_t \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 + \frac{1}{\sqrt{n}}\sum_t \left(\frac{u_t u_t' - \Sigma_u}{\sqrt{n}}\right)\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 \Rightarrow \Sigma_u \int_0^1 G_{a,c}^2(r)\,dr.$$

■

Lemma 10 *For the model (4), under Assumption 1,*

$$\frac{1}{n^{3/2}}\sum_t u_t Y_{t-1}^2 \Rightarrow \int_0^1 G_{a,c}^2(r)\,dB_u(r) + 2\left(\Lambda_{uu}' a \int_0^1 G_{a,c}^2(r)\,dr + \Lambda_{u\varepsilon}\int_0^1 G_{a,c}(r)\,dr\right).$$

Proof of Lemma 10. We have

$$\frac{1}{n^{3/2}}\sum_t u_t Y_{t-1}^2 = \frac{1}{n^{3/2}}\sum_{t=2}^n \sum_{j=0}^{\infty} G_{1,j}\eta_{t-j} Y_{t-1}^2.$$

For $j=0$,

$$\frac{1}{n^{3/2}}\sum_{t=2}^n G_{1,0}\eta_t Y_{t-1}^2 \Rightarrow G_{1,0}\int_0^1 G_{a,c}^2(r)\,dB_\eta(r).$$

For $j=1$,

$$\begin{aligned}
&\frac{1}{n^{3/2}}\sum_{t=2}^n G_{1,1}\eta_{t-1} Y_{t-1}^2 = \frac{1}{n^{3/2}}\sum_{t=2}^n G_{1,1}\eta_{t-1}\left(\exp\left(\frac{c}{n} + \frac{a' u_{t-1}}{\sqrt{n}}\right) Y_{t-2} + \varepsilon_{t-1}\right)^2 \\
= \;& G_{1,1}\int_0^1 G_{a,c}^2(r)\,dB_\eta(r) + \frac{1}{n^{3/2}}\sum_{t=2}^n G_{1,1}\eta_{t-1}\frac{2a' u_{t-1}}{\sqrt{n}} Y_{t-2}^2 + \frac{2}{n^{3/2}}\sum_{t=2}^n G_{1,1}\eta_{t-1} Y_{t-2}\varepsilon_{t-1} + o_p(1) \\
\Rightarrow \;& G_{1,1}\int_0^1 G_{a,c}^2(r)\,dB_\eta(r) + 2G_{1,1}\Sigma_\eta G_{1,0}' a\int_0^1 G_{a,c}^2(r)\,dr + 2G_{1,1}\Sigma_\eta G_{2,0}\int_0^1 G_{a,c}(r)\,dr.
\end{aligned}$$

For $j=2$,

$$\begin{aligned}
&\frac{1}{n^{3/2}}\sum_{t=2}^n G_{1,2}\eta_{t-2} Y_{t-1}^2 \\
= \;& \frac{1}{n^{3/2}}\sum_{t=2}^n G_{1,2}\eta_{t-2}\left(\exp\left(\frac{2c}{n} + \frac{a'(u_{t-1}+u_{t-2})}{\sqrt{n}}\right) Y_{t-3} + \exp\left(\frac{c}{n} + \frac{a' u_{t-1}}{\sqrt{n}}\right)\varepsilon_{t-2} + \varepsilon_{t-1}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&= G_{1,2}\int_0^1 G_{a,c}^2(r)\,dB_\eta(r) + \frac{1}{n^{3/2}}\sum_{t=2}^n G_{1,2}\eta_{t-2}\frac{2a'(u_{t-1}+u_{t-2})}{\sqrt{n}}Y_{t-3}^2 \\
&\quad + \frac{2}{n^{3/2}}\sum_{t=2}^n G_{1,2}\eta_{t-2}Y_{t-3}(\varepsilon_{t-2}+\varepsilon_{t-1}) + o_p(1) \\
&\Rightarrow G_{1,2}\int_0^1 G_{a,c}^2(r)\,dB_\eta(r) + 2G_{1,2}\Sigma_\eta(G_{1,0}+G_{1,1})'a\int_0^1 G_{a,c}^2(r)\,dr + 2G_{1,2}\Sigma_\eta(G_{2,0}+G_{2,1})'\int_0^1 G_{a,c}(r)\,dr.
\end{aligned}
$$

Continuing this scheme and using summability, we deduce that

$$
\frac{1}{n^{3/2}}\sum_t u_t Y_{t-1}^2 \Rightarrow \sum_{j=0}^\infty G_{1,j}\int_0^1 G_{a,c}^2(r)\,dB_\eta(r) + 2\sum_{j=1}^\infty G_{1,j}\Sigma_\eta\left(\sum_{k=0}^{j-1}G_{1,k}'a\int_0^1 G_{a,c}^2(r)\,dr + \sum_{k=0}^{j-1}G_{2,k}'\int_0^1 G_{a,c}(r)\,dr\right).
$$

and the proof of the Lemma is completed by using (51) and (52). ■

Lemma 11 *For the model (4), under Assumption 1,*

$$
\begin{aligned}
\sum_{t=2}^n u_t\varepsilon_t Y_{t-1} &= n^{3/2}\Sigma_{u\varepsilon}\int_0^1 G_{a,c}(r)\,dr + n\Bigg\{\sum_{j=0}^\infty (G_{2,j}\otimes G_{1,j})H\int_0^1 G_{a,c}(r)\,d\xi(r) \\
&+ \sum_{j=1}^\infty (G_{2,j}\otimes G_{1,j})M_3\left(\left(\sum_{i=0}^{j-1}G_{1,i}\right)'a\int_0^1 G_{a,c}(r)\,dr + \left(\sum_{i=0}^{j-1}G_{2,i}\right)'\right) \\
&+ \sum_{j\neq k}(G_{2,k}\otimes G_{1,j})\int_0^1 G_{a,c}(r)\,d\zeta(r)\Bigg\} + o_p(n).
\end{aligned}
$$

Proof of Lemma 11. The proof is similar to that of Lemma 8 of Lieberman and Phillips (2017b) and is omitted.

Lemma 12 *For the model (18) where $K = 1$, $\Sigma_{u\varepsilon} = 0$, u_t and ε_t are iid, and with the filtration $\mathcal{F}_r = \sigma\{(B_u(s), B_\varepsilon(s)), 0 \le s \le r\}$ the instantaneous kurtosis measure is*

$$
\kappa_{b,c}(r) = \frac{E\left(E\left[(dG_{a,c}(r))^4|\mathcal{F}_r\right]\right)}{\left\{E\left(E\left[(dG_{a,c}(r))^2|\mathcal{F}_r\right]\right)\right\}^2} = 3 + \frac{3b^2\left[E\left(G_{a,c}^4(r)\right) - \left(E\left(G_{a,c}^2(r)\right)\right)^2\right]}{b^2\left(E\left(G_{a,c}^2(r)\right)\right)^2 + \sigma_\varepsilon^4 + 2b\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right)} + o_p(1)
$$

Proof of Lemma 12. The process increments $dG_{a,c}(r)$ at r satisfy (18)

$$
dG_{a,c}(r) = aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r) + \left(c + \frac{b}{2}\right)G_{a,c}(r)\,dr, \tag{53}
$$

where $b = a^2\sigma_u^2$. Then

$$
E\left[(dG_{a,c}(r))^4|\mathcal{F}_r\right] = E\left[aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r) + \left(c+\frac{b}{2}\right)G_{a,c}(r)\,dr|\mathcal{F}_r\right]^4
$$

$$= E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^4 |\mathcal{F}_r\right] + 4\left(c+\frac{b}{2}\right) E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^3 G_{a,c}(r)\,|\mathcal{F}_r\right] dr$$

$$+ 6\left(c+\frac{b}{2}\right)^2 E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^2 G_{a,c}(r)^2\,|\mathcal{F}_r\right](dr)^2$$

$$+ 4\left(c+\frac{b}{2}\right)^3 E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right) G_{a,c}(r)^3\,|\mathcal{F}_r\right](dr)^3 + \left(c+\frac{b}{2}\right)^4 E\left[G_{a,c}(r)^4\,|\mathcal{F}_r\right](dr)^4$$

$$= \left[3b^2 G_{a,c}(r)^4 + 6b\sigma_\varepsilon^2 G_{a,c}(r)^2 + 3\sigma_\varepsilon^4\right](dr)^2 + 6\left(c+\frac{b}{2}\right)^2\left[bG_{a,c}(r)^4 + G_{a,c}(r)^2\sigma_\varepsilon^2\right](dr)^3 + O_p\left((dr)^4\right) \tag{54}$$

since $E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^3 G_{a,c}(r)\,|\mathcal{F}_r\right] dr = 0$,

$$\begin{aligned} E\left\{E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^4 |\mathcal{F}_r\right]\right\} &= 3a^4\sigma_u^4 E\left(G_{a,c}^4(r)\right) + 6a^2\sigma_u^2\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right) + 3\sigma_\varepsilon^4 \\ &= 3b^2 E\left(G_{a,c}^4(r)\right) + 6b\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right) + 3\sigma_\varepsilon^4, \end{aligned} \tag{55}$$

and

$$E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^2 G_{a,c}(r)^2\,|\mathcal{F}_r\right] = \left[a^2\sigma_u^2 G_{a,c}^4(r) + \sigma_\varepsilon^2 G_{a,c}(r)^2\right] dr.$$

Similarly

$$\begin{aligned} E\left[\left(dG_{a,c}(r)\right)^2 |\mathcal{F}_r\right] &= E\left[aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r) + \left(c+\frac{b}{2}\right)G_{a,c}(r)\,dr|\mathcal{F}_r\right]^2 \\ &= E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^2 |\mathcal{F}_r\right] + \left(c+\frac{b}{2}\right)^2 E\left[G_{a,c}(r)^2\,|\mathcal{F}_r\right](dr)^2 \\ &= E\left[\left(a^2\sigma_u^2 G_{a,c}(r)^2 + \sigma_\varepsilon^2\right)|\mathcal{F}_r\right] dr + \left(c+\frac{b}{2}\right)^2 E\left[G_{a,c}(r)^2\,|\mathcal{F}_r\right](dr)^2 \\ &= \left[bG_{a,c}(r)^2 + \sigma_\varepsilon^2\right] dr + \left(c+\frac{b}{2}\right)^2 G_{a,c}(r)^2 (dr)^2. \end{aligned} \tag{56}$$

Using (54) - (56) gives

$$\begin{aligned} \kappa_{b,c}(r) &= \frac{E\left(E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^4 |\mathcal{F}_r\right]\right) + o_p\left((dr)^2\right)}{\left\{E\left(E\left[\left(aG_{a,c}(r)\,dB_u(r) + dB_\varepsilon(r)\right)^2 |\mathcal{F}_r\right]\right) + o_p\left((dr)^2\right)\right\}^2} \\ &= \frac{3b^2 E\left(G_{a,c}^4(r)\right) + 6b\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right) + 3\sigma_\varepsilon^4}{\left(bE\left(G_{a,c}^2(r)\right) + \sigma_\varepsilon^2\right)^2} + o_p(1) \end{aligned}$$

$$= \frac{3b^2 E\left(G_{a,c}^4(r)\right) + 6b\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right) + 3\sigma_\varepsilon^4}{b^2\left(E\left(G_{a,c}^2(r)\right)\right)^2 + \sigma_\varepsilon^4 + 2b\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right)} + o_p(1)$$

$$= 3 + \frac{3b^2\left[E\left(G_{a,c}^4(r)\right) - \left(E\left(G_{a,c}^2(r)\right)\right)^2\right]}{b^2\left(E\left(G_{a,c}^2(r)\right)\right)^2 + \sigma_\varepsilon^4 + 2b\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right)} + o_p(1),$$

as stated. For large b, note that

$$\begin{aligned} E\left(G_{a,c}^4(r)\right) &= \frac{3\sigma_\varepsilon^4 e^{4(c+2b)r}}{c+b}\left(\frac{1}{2(c+3b)} - \frac{1}{4(c+2b)}\right) = \frac{3\sigma_\varepsilon^4 e^{4(c+2b)r}}{4(c+b)}\frac{2(c+2b)-(c+3b)}{(c+3b)(c+2b)} \\ &= \frac{3\sigma_\varepsilon^4 e^{4(c+2b)r}}{4(c+b)}\frac{c+b}{(c+3b)(c+2b)} = \frac{3\sigma_\varepsilon^4 e^{4(c+2b)r}}{4(c+3b)(c+2b)}, \end{aligned}$$

and $E\left(G_{a,c}^2(r)\right) = \sigma_\varepsilon^2 \frac{e^{2(c+b)r}-1}{2(c+b)} \sim \sigma_\varepsilon^2 \frac{e^{2(c+b)r}}{2(c+b)}$. Hence, as $b \to \infty$

$$\begin{aligned} \kappa_{b,c}(r) &= \frac{3b^2 E\left(G_{a,c}^4(r)\right) + 6b\sigma_\varepsilon^2 E\left(G_{a,c}^2(r)\right) + 3\sigma_\varepsilon^4}{\left(bE\left(G_{a,c}^2(r)\right) + \sigma_\varepsilon^2\right)^2} \sim \frac{3b^2 E\left(G_{a,c}^4(r)\right)}{b^2 E\left(G_{a,c}^2(r)\right)^2} \sim \frac{3\frac{3\sigma_\varepsilon^4 e^{4(c+2b)r}}{4(c+3b)(c+2b)}}{\left(\sigma_\varepsilon^2\frac{e^{2(c+b)r}}{2(c+b)}\right)^2} \\ &= 9\frac{e^{4br}(c+b)^2}{(c+3b)(c+2b)} \sim \frac{9}{6}e^{4br}, \end{aligned}$$

and kurtosis of the process increments $dG_{a,c}(r)$ grows exponentially with b irrespective of the fixed value of c.

9.2 Proofs of the Main Results

Proof of Lemma 1. By repeated substitution, we obtain

$$Y_t = \sum_{j=1}^{t} \exp\left(\frac{(t-j)c}{n} + \frac{a'\sum_{i=j+1}^{t} u_i}{\sqrt{n}}\right)\varepsilon_j, \ t \geq 2. \tag{57}$$

Therefore, setting $t = \lfloor nr \rfloor$,

$$\begin{aligned} \frac{Y_{t=\lfloor nr\rfloor}}{\sqrt{n}} &= e^{rc+a'B_u(r)+o_p(1)}\sum_{j=1}^{\lfloor nr\rfloor}\exp\left(-\frac{jc}{n} - \frac{a'\sum_{i=1}^{j}u_i}{\sqrt{n}}\right)\frac{\varepsilon_j}{\sqrt{n}} \\ &= e^{rc+a'B_u(r)}\sum_{j=1}^{\lfloor nr\rfloor}\exp\left(-\frac{(j-1)c}{n} - \frac{a'\sum_{i=1}^{j-1}u_i}{\sqrt{n}}\right)\left\{1 - \frac{a'u_j}{\sqrt{n}} + O_p\left(n^{-1}\right)\right\}\frac{\varepsilon_j}{\sqrt{n}} + o_p(1) \\ &= e^{rc+a'B_u(r)}\left\{\sum_{j=1}^{\lfloor nr\rfloor} e^{-\frac{(j-1)c}{n} - \frac{a'\sum_{i=1}^{j-1}u_i}{\sqrt{n}}}\frac{\varepsilon_j}{\sqrt{n}} - \sum_{j=1}^{\lfloor nr\rfloor} e^{-\frac{(j-1)c}{n} - a'B_u\left(\frac{j-1}{n}\right)}\frac{a'u_j\varepsilon_j}{n}\right\} + o_p(1). \end{aligned} \tag{58}$$

Let $e^{-\frac{(j-1)c}{n}-\frac{a'\sum_{i=1}^{j-1}u_j}{\sqrt{n}}} =: f\left(-\frac{(j-1)c}{n} - a'\frac{1}{\sqrt{n}}\sum_{i=1}^{j-1}u_j\right)$. Then,

$$\frac{\partial}{\partial X}f\left(-\frac{(j-1)c}{n} - a'X\right)_{X=\frac{1}{\sqrt{n}}\sum_{i=1}^{j-1}u_i} = -af\left(-\frac{(j-1)c}{n} - a'\frac{1}{\sqrt{n}}\sum_{i=1}^{j-1}u_i\right)$$

and by Ibragimov and Phillips (2008; equation (4.9)) we obtain the following sample covariance limit

$$\sum_{j=1}^{\lfloor nr\rfloor}\exp\left(-\frac{(j-1)c}{n} - \frac{a'\sum_{i=1}^{j-1}u_i}{\sqrt{n}}\right)\frac{\varepsilon_j}{\sqrt{n}} \Rightarrow -a'\Lambda_{u\varepsilon}\int_0^r e^{-pc-a'B_u(p)}dp + \int_0^r e^{-pc-a'B_u(p)}dB_\varepsilon(p). \quad (59)$$

Furthermore,

$$\sum_{j=1}^{\lfloor nr\rfloor}\exp\left(-\frac{(j-1)c}{n} - a'B_u\left(\frac{j-1}{n}\right)\right)\frac{a'u_j\varepsilon_j}{n} = a'\Sigma_{u\varepsilon}\int_0^r e^{-pc-a'B_u(p)}dp + o_p(1). \quad (60)$$

It follows from (58), (59) and (60) that

$$\begin{aligned}\frac{Y_t}{\sqrt{n}} &\Rightarrow e^{rc+a'B_u(r)}\left\{-a'\Lambda_{u\varepsilon}\int_0^r e^{-pc-a'B_u(p)}dp + \int_0^r e^{-pc-a'B_u(p)}dB_\varepsilon(p) - a'\Sigma_{u\varepsilon}\int_0^r e^{-pc-aB_u(p)}dp\right\}\\ &= e^{rc+a'B_u(r)}\left\{\int_0^r e^{-pc-a'B_u(p)}dB_\varepsilon(p) - a'\Delta_{u\varepsilon}\int_0^r e^{-pc-aB_u(p)}dp\right\},\end{aligned}$$

which is the stated result. ■

Proof of Theorem 3. The NLLS $\hat{c}_n$ of c, given known a, is defined as the solution to the equation

$$\sum_t (Y_t - \beta_{nt}(\hat{c}_n, a)Y_{t-1})\dot{\beta}_{nt}(\hat{c}_n, a)Y_{t-1} = 0, \quad (61)$$

where $\dot{\beta}_{nt}(c,a) = \frac{\partial\beta_{nt}(c,a)}{\partial c} = \frac{1}{n}\beta_{nt}(c,a)$. The solution to (61) is equivalent to the solution of

$$\sum_t Y_t\beta_{nt}(\hat{c}_n, a)Y_{t-1} = \sum_t \beta_{nt}^2(\hat{c}_n, a)Y_{t-1}^2.$$

or

$$\sum_t(\beta_{nt}(c,a)Y_{t-1} + \varepsilon_t)\beta_{nt}(\hat{c}_n, a)Y_{t-1} = \sum_t\beta_{nt}(2\hat{c}_n, 2a)Y_{t-1}^2.$$

Rearranging the last equation, we seek a solution to

$$\sum_t e^{2a'u_t/\sqrt{n}}\left[e^{2\hat{c}/n} - e^{(c+\hat{c}_n)/n}\right]Y_{t-1}^2 = \sum_t e^{a'u_t/\sqrt{n}}e^{\hat{c}_n/n}\varepsilon_tY_{t-1}. \quad (62)$$

Expanding the left side of (62) we get

$$\begin{aligned}
&\sum_t e^{2a'u_t/\sqrt{n}} \left[\sum_{j=1}^{\infty} \frac{(2\hat{c}_n)^j - (c+\hat{c}_n)^j}{n^j j!} \right] Y_{t-1}^2 = \sum_t e^{2a'u_t/\sqrt{n}} \left[\sum_{j=1}^{\infty} \frac{(2\hat{c}_n)^j - (2\hat{c}_n + [c-\hat{c}_n])^j}{n^j j!} \right] Y_{t-1}^2 \\
= \; & \sum_t e^{2a'u_t/\sqrt{n}} \left[\sum_{j=1}^{\infty} \frac{(2\hat{c}_n)^j - \sum_{k=0}^{j} \binom{j}{k} (2\hat{c}_n)^k [c-\hat{c}_n]^{j-k}}{n^j j!} \right] Y_{t-1}^2 \\
= \; & \sum_t e^{2a'u_t/\sqrt{n}} \left[\sum_{j=1}^{\infty} \frac{- \sum_{k=0}^{j-1} \binom{j}{k} (2\hat{c}_n)^k [c-\hat{c}_n]^{j-k}}{n^j j!} \right] Y_{t-1}^2 \\
= \; & \sum_t e^{2a'u_t/\sqrt{n}} \left[(\hat{c}_n - c) \sum_{j=1}^{\infty} \frac{\sum_{k=0}^{j-1} \binom{j}{k} (2\hat{c}_n)^k [c-\hat{c}_n]^{j-1-k}}{n^j j!} \right] Y_{t-1}^2 .
\end{aligned}$$

At the true value of c, the objective function $Q_n^a(c) = n^{-1}(Y_t - \beta_{nt}(c,a) Y_{t-1})^2$ converges in probability to σ_ε^2 and therefore, the only term in the square brackets which contributes asymptotically is the first order term, $(\hat{c}_n - c)/n$. The leading term on the left side of (62) is therefore

$$\sum_t \left(1 + \frac{2a'u_t}{\sqrt{n}} + \frac{2a'\Sigma_u a}{n} + \frac{2a'(u_t u_t' - \Sigma_u) a}{n} + o_p\left(\frac{1}{n}\right) \right) \left[\frac{\hat{c}_n - c}{n} + O_p\left(n^{-2}\right) \right] Y_{t-1}^2 .$$

Upon scaling by $1/n$, we have the following asymptotic form

$$\begin{aligned}
& (\hat{c}_n - c) \left\{ \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 + \frac{2}{n} \sum_t \frac{a'u_t}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 + \frac{2}{n} \sum_t a' \left(\frac{u_t u_t' - \Sigma_u}{\sqrt{n}} \right) a \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 \right. \\
& \left. + \frac{2a'\Sigma_u a}{n^2} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 + o_p\left(\frac{1}{n}\right) \right\} \\
= \; & (\hat{c}_n - c) \left\{ \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 + O_p\left(\frac{1}{n}\right) \right\} \sim_a (\hat{c}_n - c) \int_0^1 G_{a,c}^2(r)\, dr. \qquad (63)
\end{aligned}$$

Scaling by $1/n$, the dominant term on the right side of (62) is

$$\frac{1}{n} \sum_t \varepsilon_t Y_{t-1} + \frac{1}{n} \sum_t a' u_t \varepsilon_t \frac{Y_{t-1}}{\sqrt{n}},$$

and

$$\frac{1}{n^{3/2}} \sum_t a' u_t \varepsilon_t Y_{t-1} = \frac{a'\Sigma_{u\varepsilon}}{n} \sum_t \frac{Y_{t-1}}{\sqrt{n}} + \frac{1}{\sqrt{n}} \sum_t a' \left(\frac{u_t \varepsilon_t - \Sigma_{u\varepsilon}}{\sqrt{n}} \right) \frac{Y_{t-1}}{\sqrt{n}} \Rightarrow a'\Sigma_{u\varepsilon} \int_0^1 G_{a,c}(r)\, dr.$$

Using Lemma 2, we obtain

$$\frac{1}{n}\sum_t \varepsilon_t Y_{t-1} + \frac{1}{n^{3/2}}\sum_t a' u_t \varepsilon_t Y_{t-1} \Rightarrow \left(\int_0^1 G_{a,c}(r)\, dB_\varepsilon(r) + a'\Lambda_{u\varepsilon}\int_0^1 G_{a,c}(r)\, dr + \lambda_{\varepsilon\varepsilon}\right) + a'\Sigma_{u\varepsilon}\int_0^1 G_{a,c}(r)\, dr$$
$$= \int_0^1 G_{a,c}(r)\, dB_\varepsilon(r) + a'\Delta_{u\varepsilon}\int_0^1 G_{a,c}(r)\, dr + \lambda_{\varepsilon\varepsilon}. \tag{64}$$

The result of the theorem then follows from (63) and (64). ■

Proof of Theorem 4. Write the autoregressive coefficient as $\beta_{nt} = e^{\frac{c}{n}+\frac{a'u_t}{\sqrt{n}}} =: e^{x_t'\gamma}$, with $x_t' = \left(\frac{1}{n}, \frac{u_t'}{\sqrt{n}}\right)$ and $\gamma' = (c, a')$. The NLLS $\hat{\gamma}_n$ of γ is defined as the solution to the equation

$$\sum_t \left(Y_t - \beta_{nt}(\hat{\gamma}_n) Y_{t-1}\right) \dot{\beta}_{nt}(\hat{\gamma}_n) Y_{t-1} = 0. \tag{65}$$

Since the derivative vector $\dot{\beta}_{nt}(\gamma) = x_t \beta_{nt}$, we need to solve the system

$$\sum_t Y_t x_t \beta_{nt}(\hat{\gamma}_n) Y_{t-1} = \sum_t x_t \beta_{nt}^2(\hat{\gamma}_n) Y_{t-1}^2. \tag{66}$$

Using the fact that $\beta_{nt}(\hat{\gamma}_n)\beta_{nt}(\gamma) = \beta_{nt}(\hat{\gamma}_n + \gamma)$, (66) is equivalent to

$$\sum_t x_t \beta_{nt}(\hat{\gamma}_n)\left(\beta_{nt}(\gamma) Y_{t-1} + \varepsilon_t\right) Y_{t-1} = \sum_t x_t \beta_{nt}^2(\hat{\gamma}_n) Y_{t-1}^2,$$

or

$$\sum_t x_t \left[\beta_{nt}(2\hat{\gamma}_n) - \beta_{nt}(\hat{\gamma}_n + \gamma)\right] Y_{t-1}^2 = \sum_t x_t \varepsilon_t \beta_{nt}(\hat{\gamma}_n) Y_{t-1}. \tag{67}$$

Expanding (67) we obtain

$$\begin{aligned}
&\sum_t x_t \Bigg[1 + \frac{2\hat{c}_n}{n} + \frac{2\hat{a}_n' u_t}{\sqrt{n}} + \frac{1}{2}\left(2\left(\frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}}\right)\right)^2 + O_p\left(n^{-3/2}\right) \\
&- \left(1 + \frac{(\hat{c}_n + c)}{n} + \frac{(\hat{a}_n + a)' u_t}{\sqrt{n}} + \frac{1}{2}\left(\frac{\hat{c}_n + c}{n} + \frac{(\hat{a}_n + a)' u_t}{\sqrt{n}}\right)^2\right) + O_p\left(n^{-3/2}\right)\Bigg] Y_{t-1}^2 \\
= \;& \sum_t x_t \varepsilon_t \left(1 + \frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}} + \frac{1}{2}\left(\frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}}\right)^2 + O_p\left(n^{-3/2}\right)\right) Y_{t-1},
\end{aligned}$$

which becomes

$$\sum_t x_t \left\{\frac{(\hat{c}_n - c)}{n} + \frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + \frac{1}{2}\left(2\left(\frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}}\right)\right)^2 - \frac{1}{2}\left(\frac{\hat{c}_n + c}{n} + \frac{(\hat{a}_n + a)' u_t}{\sqrt{n}}\right)^2 + O_p\left(n^{-3/2}\right)\right\} Y_{t-1}^2$$
$$= \sum_t x_t \varepsilon_t \left(1 + \frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}} + \frac{1}{2}\left(\frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}}\right)^2 + O_p\left(n^{-3/2}\right)\right) Y_{t-1}. \tag{68}$$

Using $x_t' = \left(\frac{1}{n}, \frac{u_t'}{\sqrt{n}}\right)$ we have the system

$$\sum_t \begin{pmatrix} \frac{1}{n} \\ \frac{u_t}{\sqrt{n}} \end{pmatrix} \left\{ \frac{(\hat{c}_n - c)}{n} + \frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + \frac{1}{2}\left(2\left(\frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}}\right)\right)^2 - \frac{1}{2}\left(\frac{\hat{c}_n + c}{n} + \frac{(\hat{a}_n + a)' u_t}{\sqrt{n}}\right)^2 + O_p\left(n^{-3/2}\right) \right\} Y_{t-1}^2$$

$$= \sum_t \begin{pmatrix} \frac{1}{n} \\ \frac{u_t}{\sqrt{n}} \end{pmatrix} \varepsilon_t \left(1 + \frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}} + \frac{1}{2}\left(\frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}}\right)^2 + O_p\left(n^{-3/2}\right)\right) Y_{t-1}. \qquad (69)$$

The leading term of the upper element of the left side of (69) is

$$\begin{aligned}
& \frac{1}{n}\sum_t \left(\frac{(\hat{c}_n - c)}{n} + \frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + 2\frac{\hat{a}_n' u_t u_t' \hat{a}_n'}{n} - \frac{1}{2}\frac{(\hat{a}_n + a)' u_t u_t' (\hat{a}_n + a)}{n} \right) Y_{t-1}^2 \\
= \;& \frac{1}{n}\sum_t \left(\frac{(\hat{c}_n - c)}{n} + \frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + 2\frac{\hat{a}_n' (u_t u_t' - \Sigma_u) \hat{a}_n'}{n} + 2\frac{\hat{a}_n' \Sigma_u \hat{a}_n'}{n} \right) Y_{t-1}^2 \\
& -\frac{1}{2n}\sum_t \left[\frac{(\hat{a}_n + a)' (u_t u_t' - \Sigma_u) (\hat{a}_n + a)}{n} + \frac{(\hat{a}_n + a)' \Sigma_u (\hat{a}_n + a)}{n} \right] Y_{t-1}^2 \\
= \;& (\hat{c}_n - c)\frac{1}{n}\sum_t \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 + (\hat{a}_n - a)' \sum_t \frac{u_t}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 + \frac{2}{\sqrt{n}}\hat{a}_n' \sum_t \left(\frac{u_t u_t' - \Sigma_u}{\sqrt{n}}\right)\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 \hat{a}_n \\
& +2\hat{a}_n' \Sigma_u \hat{a}_n \frac{1}{n}\sum_t \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 - \frac{(\hat{a}_n + a)' \Sigma_u (\hat{a}_n + a)}{2n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 - \frac{1}{2\sqrt{n}}\sum_t \left[(\hat{a}_n + a)' \left(\frac{u_t u_t' - \Sigma_u}{\sqrt{n}}\right)(\hat{a}_n + a)\right] (
\end{aligned}$$

$$= (\hat{c}_n - c)\frac{1}{n}\sum_t \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 + (\hat{a}_n - a)' \sum_t \frac{u_t}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 + O_p\left(\frac{1}{\sqrt{n}}\right). \qquad (70)$$

Now,

$$(\hat{c}_n - c)\frac{1}{n}\sum_t \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 \sim_a (\hat{c}_n - c)\int_0^1 G_{a,c}^2(r)\, dr, \qquad (71)$$

and by Lemma 10,

$$(\hat{a}_n - a)' \sum_t \frac{u_t}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2 \sim_a (\hat{a}_n - a)' \left\{ \int_0^1 G_{a,c}^2(r)\, dB_u(r) + 2\left(\Lambda_{uu}' a \int_0^1 G_{a,c}^2(r)\, dr + \Lambda_{u\varepsilon}\int_0^1 G_{a,c}(r)\, dr\right)\right\}. \qquad (72)$$

By Lemma 2, the asymptotic form of the leading term of the top element of the right side of (69) is

$$\frac{1}{n}\sum_t \varepsilon_t Y_{t-1} \sim_a \left(\int_0^1 G_{a,c}(r)\, dB_\varepsilon(r) + \Lambda_{u\varepsilon}' a \int_0^1 G_a(r)\, dr + \lambda_{\varepsilon\varepsilon}\right).$$

Using (70), (71) and (72), the solution to the first equation has the asymptotic form

$$(\hat{c}_n - c) \sim_a \frac{\int_0^1 G_{a,c}(r)\, dB_\varepsilon(r) + \Lambda'_{u\varepsilon} a \int_0^1 G_a(r)\, dr + \lambda_{\varepsilon\varepsilon}}{\int_0^1 G_{a,c}^2(r)\, dr} - \frac{(\hat{a}_n - a)' \left\{ \int_0^1 G_{a,c}^2(r)\, dB_u(r) + 2\left(\Lambda'_{uu} a \int_0^1 G_{a,c}^2(r)\, dr + \Lambda_{u\varepsilon} \int_0^1 G_{a,c}(r)\, dr \right) \right\}}{\int_0^1 G_{a,c}^2(r)\, dr}. \tag{73}$$

Next continue with the lower element of the system (69). The leading term of the left side of the lower element of (69) is

$$\begin{aligned} & \sum_t \frac{u_t}{\sqrt{n}} \frac{(\hat{c}_n - c)}{n} Y_{t-1}^2 + \sum_t \frac{u_t}{\sqrt{n}} \frac{u_t' (\hat{a}_n - a)}{\sqrt{n}} Y_{t-1}^2 \\ = \;& (\hat{c}_n - c) \sum_t \frac{u_t}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 + \sum_t u_t u_t' (\hat{a}_n - a) \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 \end{aligned} \tag{74}$$

and by Lemma 9,

$$\sum_t u_t u_t' (\hat{a}_n - a) \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 = n \Sigma_u (\hat{a}_n - a) \int_0^1 G_{a,c}^2(r)\, dr + o_p(n(\hat{a}_n - a)). \tag{75}$$

Taking the lower element of the right side of (69) and scaling by $1/n$ we have

$$\begin{aligned} & \frac{1}{n} \sum_t \frac{u_t}{\sqrt{n}} \varepsilon_t \left(1 + \frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}} + \frac{1}{2} \left(\frac{\hat{c}_n}{n} + \frac{\hat{a}_n' u_t}{\sqrt{n}} \right)^2 + O_p\left(n^{-3/2}\right) \right) Y_{t-1} \\ = \;& \frac{1}{n} \sum_t u_t \varepsilon_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + \frac{1}{\sqrt{n}} \sum_t \frac{u_t \varepsilon_t}{\sqrt{n}} \frac{u_t' \hat{a}_n}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + \frac{1}{2n} \sum_t u_t \varepsilon_t \left(\frac{u_t' \hat{a}_n}{\sqrt{n}} \right)^2 \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + O_p\left(n^{-3/2}\right) \\ = \;& \Sigma_{u\varepsilon} \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + \frac{1}{\sqrt{n}} \sum_t \left(\frac{u_t \varepsilon_t - \Sigma_{u\varepsilon}}{\sqrt{n}} \right) \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + \frac{E(\varepsilon_t u_t u_t')}{\sqrt{n}} \hat{a}_n \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) \\ & + \frac{1}{\sqrt{n}} \sum_t \frac{[\varepsilon_t u_t u_t' - E(\varepsilon_t u_t u_t')] \hat{a}_n'}{n} \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + \frac{E\left(\varepsilon_t u_t (u_t' \hat{a}_n)^2 \right)}{2n} \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + O_p\left(n^{-3/2}\right). \end{aligned} \tag{76}$$

Thus, the leading term of the lower elements of the right side of (69) is

$$\Sigma_{u\varepsilon} \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) \sim_a \Sigma_{u\varepsilon} \int_0^1 G_{a,c}(r)\, dr. \tag{77}$$

and the remaining terms in (76) are all no larger than $O_p\left(n^{-1/2}\right)$. Rescaling (74) by $1/n$ we have

$$\frac{1}{n} (\hat{c}_n - c) \sum_t \frac{u_t}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 + \frac{1}{n} \sum_t u_t u_t' (\hat{a}_n - a) \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 \sim_a \Sigma_u (\hat{a}_n - a) \int_0^1 G_{a,c}(r)^2\, dr. \tag{78}$$

Combining (78) with (77) we obtain the following asymptotics for $\hat{a}_n$ in the case where $\Sigma_{u\varepsilon} \neq 0$

$$\hat{a}_n - a \sim_a \Sigma_u^{-1} \Sigma_{u\varepsilon} \frac{\int_0^1 G_{a,c}(r)\, dr}{\int_0^1 G_{a,c}(r)^2\, dr}, \tag{79}$$

Using (79) in (73) we find that

$$\begin{aligned}(\hat{c}_n - c) &\sim_a \frac{\int_0^1 G_{a,c}(r)\, dB_\varepsilon(r) + \Lambda'_{u\varepsilon} a \int_0^1 G_a(r)\, dr + \lambda_{\varepsilon\varepsilon}}{\int_0^1 G_{a,c}^2(r)\, dr} \\ &- \int_0^1 G_{a,c}(r)\, dr \frac{\Sigma_{\varepsilon u} \Sigma_u^{-1} \left\{ \int_0^1 dB_u(r)\, G_{a,c}^2(r) + 2\left(\Lambda'_{uu} a \int_0^1 G_{a,c}^2(r)\, dr + \Lambda_{u\varepsilon} \int_0^1 G_{a,c}(r)\, dr \right) \right\}}{\left(\int_0^1 G_{a,c}^2(r)\, dr \right)^2},\end{aligned} \tag{80}$$

which, together with (79), gives the first part of the theorem.

When $\Sigma_{u\varepsilon} = 0$, we have

$$(\hat{c}_n - c) \sim_a \frac{\int_0^1 G_{a,c}(r)\, dB_\varepsilon(r) + \Lambda'_{u\varepsilon} a \int_0^1 G_a(r)\, dr + \lambda_{\varepsilon\varepsilon}}{\int_0^1 G_{a,c}^2(r)\, dr}, \tag{81}$$

in place of (80). Also, in place of (76) after scaling by $\sqrt{n}$, we have

$$\begin{aligned} &= \sum_t \left(\frac{u_t \varepsilon_t}{\sqrt{n}} \right) \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + E\left(\varepsilon_t u_t u_t'\right) \hat{a}_n \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + \frac{1}{\sqrt{n}} \sum_t \frac{\left[\varepsilon_t u_t u_t' - E\left(\varepsilon_t u_t u_t'\right)\right] \hat{a}_n'}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}} \right) \\ &\quad + \frac{E\left(\varepsilon_t u_t \left(u_t' \hat{a}_n\right)^2\right)}{2\sqrt{n}} \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + O_p\left(\frac{1}{n} \right) \\ &= \sum_t \left(\frac{u_t \varepsilon_t}{\sqrt{n}} \right) \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + E\left(\varepsilon_t u_t u_t'\right) \hat{a}_n \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + O_p\left(\frac{1}{\sqrt{n}} \right). \end{aligned} \tag{82}$$

Correspondingly rescaling (78) by $\sqrt{n}$ we have

$$\frac{1}{\sqrt{n}} (\hat{c}_n - c) \sum_t \frac{u_t}{\sqrt{n}} \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 + \frac{1}{n} \sum_t u_t u_t' \left[\sqrt{n} (\hat{a}_n - a) \right] \left(\frac{Y_{t-1}}{\sqrt{n}} \right)^2 \sim_a \Sigma_u \sqrt{n} (\hat{a}_n - a) \int_0^1 G_{a,c}(r)^2\, dr. \tag{83}$$

It now follows from (83) and (82) that

$$\sum_t \left(\frac{u_t \varepsilon_t}{\sqrt{n}} \right) \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + E\left(\varepsilon_t u_t u_t'\right) \hat{a}_n \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + O_p\left(\frac{1}{\sqrt{n}} \right) \sim_a \Sigma_u \sqrt{n} (\hat{a}_n - a) \int_0^1 G_{a,c}(r)^2\, dr. \tag{84}$$

Using Lemma 15 we deduce from (84) that when $\Sigma_{u\varepsilon} = 0$

$$\sqrt{n} (\hat{a}_n - a) \sim_a \left(\int_0^1 G_{a,c}(r)^2\, dr \right)^{-1} \Sigma_u^{-1} \left\{ \sum_t \left(\frac{u_t \varepsilon_t}{\sqrt{n}} \right) \left(\frac{Y_{t-1}}{\sqrt{n}} \right) + E\left(\varepsilon_t u_t u_t'\right) \hat{a}_n \frac{1}{n} \sum_t \left(\frac{Y_{t-1}}{\sqrt{n}} \right) \right\}$$

$$\sim_a \Sigma_u^{-1}\left\{\sum_{j=0}^{\infty}(G_{2,j}\otimes G_{1,j})\,H\int_0^1 G_{a,c}(r)\,d\xi(r)+\sum_{j=1}^{\infty}(G_{2,j}\otimes G_{1,j})\,M_3\left(\sum_{i=0}^{j-1}G_{1,i}'a\int_0^1 G_{a,c}(r)\,dr+\sum_{i=0}^{j-1}G_{2,i}'\right)\right.$$

$$\sim_a \Sigma_u^{-1}\left\{\sum_{j=0}^{\infty}(G_{2,j}\otimes G_{1,j})\,H\int_0^1 G_{a,c}(r)\,d\xi(r)+\sum_{j=1}^{\infty}(G_{2,j}\otimes G_{1,j})\,M_3\left(\sum_{i=0}^{j-1}G_{1,i}'a\int_0^1 G_{a,c}(r)\,dr+\sum_{i=0}^{j-1}G_{2,i}'\right)\right.$$
$$\left.+\sum_{j\neq k}(G_{2,k}\otimes G_{1,j})\int_0^1 G_{a,c}(r)\,d\zeta(r)+E\left(\varepsilon_t u_t u_t' a\right)\int_0^1 G_{a,c}(r)\,dr\right\}\Big/\left(\int_0^1 G_{a,c}(r)^2\,dr\right). \tag{85}$$

which establishes the final part of the theorem, in conjunction with (81). ■

Proof of Theorem 5: The ols estimator of β_{nt} in (4) satisfies

$$\hat{\beta}_{nt}=\frac{\sum_{t=2}^n Y_tY_{t-1}}{\sum_{t=2}^n Y_{t-1}^2}=\frac{\sum_{t=2}^n \beta_{nt}Y_{t-1}^2}{\sum_{t=2}^n Y_{t-1}^2}+\frac{\sum_{t=2}^n \varepsilon_tY_{t-1}}{\sum_{t=2}^n Y_{t-1}^2}.$$

The first term above yields

$$\begin{aligned}\frac{\sum_{t=2}^n \beta_{nt}Y_{t-1}^2}{\sum_{t=2}^n Y_{t-1}^2} &= \frac{\sum_{t=2}^n\left(1+\frac{a'u_t}{\sqrt{n}}+\frac{c}{n}+\frac{1}{2}\left(\frac{a'u_t}{\sqrt{n}}\right)^2+o_p\left(n^{-1}\right)\right)Y_{t-1}^2}{\sum_{t=2}^n Y_{t-1}^2}\\ &= 1+\frac{\sum_{t=2}^n\left(\frac{a'u_t}{\sqrt{n}}+\frac{c}{n}+\frac{1}{2}\left(\frac{a'u_t}{\sqrt{n}}\right)^2+o_p\left(n^{-1}\right)\right)Y_{t-1}^2}{\sum_{t=2}^n Y_{t-1}^2}.\end{aligned}$$

Thus,

$$n\left(\hat{\beta}_{nt}-1\right)=\frac{n^{-1}\sum_{t=2}^n\left(\frac{a'u_t}{\sqrt{n}}+\frac{c}{n}+\frac{1}{2}\left(\frac{a'u_t}{\sqrt{n}}\right)^2+o_p\left(n^{-1}\right)\right)Y_{t-1}^2}{n^{-2}\sum_{t=2}^n Y_{t-1}^2}+\frac{n^{-1}\sum_{t=2}^n \varepsilon_tY_{t-1}}{n^{-2}\sum_{t=2}^n Y_{t-1}^2}$$

By Lemmas 9 and 10, the first term satisfies

$$\begin{aligned}&\frac{n^{-1}\sum_{t=2}^n\left(\frac{a'u_t}{\sqrt{n}}+\frac{c}{n}+\frac{1}{2}\left(\frac{a'u_t}{\sqrt{n}}\right)^2+o_p\left(n^{-1}\right)\right)Y_{t-1}^2}{n^{-2}\sum_{t=2}^n Y_{t-1}^2}\\ \Rightarrow\ &\frac{a'\left(\int_0^1 G_{a,c}^2(r)\,dB_u(r)+2\left(\Lambda_{uu}'a\int_0^1 G_{a,c}^2(r)\,dr+\Lambda_{u\varepsilon}\int_0^1 G_{a,c}(r)\,dr\right)\right)}{\int_0^1 G_{a,c}^2(r)\,dr}\\ &+\frac{\left(c+\frac{a'\Sigma_u a}{2}\right)\int_0^1 G_{a,c}^2(r)\,dr}{\int_0^1 G_{a,c}^2(r)\,dr},\end{aligned}$$

and by Lemma 2, the second term yields

$$\frac{n^{-1}\sum_{t=2}^{n}\varepsilon_t Y_{t-1}}{n^{-2}\sum_{t=2}^{n}Y_{t-1}^2} \Rightarrow \frac{\int_0^1 G_{a,c}(r)\,dB_\varepsilon(r) + \Lambda_{u\varepsilon}' a\int_0^1 G_{a,c}(r)\,dr + \lambda_{\varepsilon\varepsilon}.}{\int_0^1 G_{a,c}(r)^2\,dr}.$$

Hence,

$$\begin{aligned} n\left(\hat{\beta}_{nt}-1\right) \Rightarrow c + \frac{a'\Sigma_u a}{2} + \frac{a'\left(\int_0^1 G_{a,c}^2(r)\,dB_u(r) + 2\left(\Lambda_{uu}' a\int_0^1 G_{a,c}^2(r)\,dr + \Lambda_{u\varepsilon}\int_0^1 G_{a,c}(r)\,dr\right)\right)}{\int_0^1 G_{a,c}^2(r)\,dr} \\ + \frac{\int_0^1 G_{a,c}(r)\,dB_\varepsilon(r) + \Lambda_{u\varepsilon}' a\int_0^1 G_{a,c}(r)\,dr + \lambda_{\varepsilon\varepsilon}}{\int_0^1 G_{a,c}(r)^2\,dr}, \end{aligned}$$

which simplifies to the stated result. ■

Proof of Lemma 6. As $B_u(p)$ is independent of $dB_\varepsilon(p)$, the expected value of $G_{a,c}(r)$ is zero. The covariance of the process is given by

$$\begin{aligned} & Cov\left(G_{a,c}(r), G_{a,c}(s)\right) \\ = \; & E\left(\int_0^r\int_0^s e^{(r+s-p-q)c+a(B_u(r)+B_u(s)-B_u(p)-B_u(q))}dB_\varepsilon(p)\,dB_\varepsilon(q)\right) \\ = \; & \sigma_\varepsilon^2\int_0^{r\wedge s} e^{(r+s-2p)c}E\left(e^{a(B_u(r\vee s)-B_u(r\wedge s)+2(B_u(r\wedge s)-B_u(p)))}\right)dp \\ = \; & \sigma_\varepsilon^2\int_0^{r\wedge s} e^{(r\vee s+r\wedge s-2p)c}e^{\frac{a^2\sigma_u^2(r\vee s-r\wedge s)}{2}+2a^2\sigma_u^2(r\wedge s-p)}dp \\ = \; & \sigma_\varepsilon^2\int_0^{r\wedge s} e^{(r\vee s-r\wedge s+2(r\wedge s-p))c}e^{\frac{b(r\vee s-r\wedge s)}{2}+2b(r\wedge s-p)}dp \\ = \; & \sigma_\varepsilon^2 e^{\left(c+\frac{b}{2}\right)(r\vee s-r\wedge s)}\frac{e^{2(c+b)r\wedge s}-1}{2(c+b)}, \end{aligned}$$

as stated. The fourth order moment is given by

$$\begin{aligned} E\left(G_{a,c}^4(r)\right) &= \int_0^r\cdots\int_0^r e^{\sum_{i=1}^4(r-p_i)c}E\left(e^{\sum_{i=1}^4 B_u(r)-B_u(p_i)}\right)E\prod_{i=1}^4 dB_\varepsilon(p_i) \\ &= 6\sigma_\varepsilon^4\int_0^r\int_0^q e^{2(2r-p-q)c}E\left(e^{4a(B_u(r)-B_u(q))+2a(B_u(q)-B_u(p))}\right)dpdq \\ &= 6\sigma_\varepsilon^4\int_0^r\int_0^q e^{2(2r-p-q)c+8a^2\sigma_u^2(r-q)+2a^2\sigma_u^2(q-p)}dpdq \\ &= \frac{3\sigma_\varepsilon^4}{c+b}\left[\frac{1-e^{4(c+2b)r}}{4(c+2b)} - e^{2(c+b)r}\frac{1-e^{2(c+3b)r}}{2(c+3b)}\right] \end{aligned}$$

which gives the stated result. ■

Proof of Lemma 7: Expansion of the limit process in this case yields

$$G_{a,c}(r) = e^{rc+a'B_u(r)} \int_0^r e^{-pc-a'B_u(p)} dB_\varepsilon(p) = \int_0^r e^{(r-p)c} \left\{1 + a\left(B_u(r) - B_u(p)\right) + O_p\left(a^2\right)\right\} dB_\varepsilon(p)$$
$$= \int_0^r e^{(r-p)c} dB_\varepsilon(p) + a \int_0^r e^{(r-p)c} \left(B_u(r) - B_u(p)\right) dB_\varepsilon(p) + O_p\left(a^2\right) =: G_c(r) + V_{c,a}(r) + O_p\left(a^2\right). \tag{86}$$

Here $G_c(r)$ is the limit process of the LUR process and has finite dimensional distribution $\mathcal{N}\left(0, \sigma_\varepsilon^2\left(e^{2cr} - 1\right)/2c\right)$ (Phillips, 1987)). The process $V_{c,a}(r)$ has mean $E\left(V_{c,a}(r)\right) = 0$, variance

$$Var\left(V_{c,a}(r)\right) = a^2\sigma_\varepsilon^2\sigma_u^2 \int_0^r e^{2(r-p)c}(r-p)\,dp = \frac{\sigma_\varepsilon^2 b}{4c^2}\left(1 + e^{2cr}(2cr-1)\right), \tag{87}$$

and fourth moment

$$E\left(V_{c,a}^4(r)\right) = a^4 \int_0^r \cdots \int_0^r e^{\sum_{i=1}^4 (r-p_i)c} E\prod_{i=1}^4 \left(\left(B_u(r) - B_u(p_i)\right)\right) E\prod_{i=1}^4 dB_\varepsilon(p_i)$$
$$= \quad 6a^4\sigma_u^4\sigma_\varepsilon^4 \int_0^r \int_0^q e^{(4r-2(p+q))c}(r-p)(r-q)\,dp\,dq = 3a^4\sigma_u^4\sigma_\varepsilon^4 \frac{\left(1 + e^{2cr}(2cr-1)\right)^2}{16c^4}.$$

It follows that $E\left(V_{c,a}^4(r)\right) = 3\left(Var\left(V_{c,a}(r)\right)\right)^2$ and $V_{c,a}$ has kurtosis 3. We deduce that $V_{c,a}(r)$ is a mixed normal ($\mathcal{MN}$) process with finite dimensional distribution

$$V_{c,a}(r) \sim_d \mathcal{MN}\left(0, \frac{\sigma_\varepsilon^2 b}{4c^2}\left(1 + e^{2cr}(2cr-1)\right)\right).$$

Finally,

$$E\left(G_c(r)V_{c,a}(r)\right) = a\sigma_\varepsilon^2 \int_0^r e^{2(r-p)c} E\left(B_u(r) - B_u(q)\right) dp = 0,$$

and so

$$Var\left(G_{a,c}(r)\right) = Var\left(G_c(r)\right) + Var\left(V_{c,a}(r)\right) + O\left(b^2\right) = \sigma_\varepsilon^2 \frac{e^{2cr}-1}{2c} + \frac{\sigma_\varepsilon^2 b}{4c^2}\left(1 + e^{2cr}(2cr-1)\right) + O\left(b^2\right), \tag{88}$$

giving (25). The moment expansion (88) is valid based on the stochastic expansion (86) because all moments of the component Gaussian processes $\left(B_u(r), B_\varepsilon(r)\right)$ are finite and bounded. ■

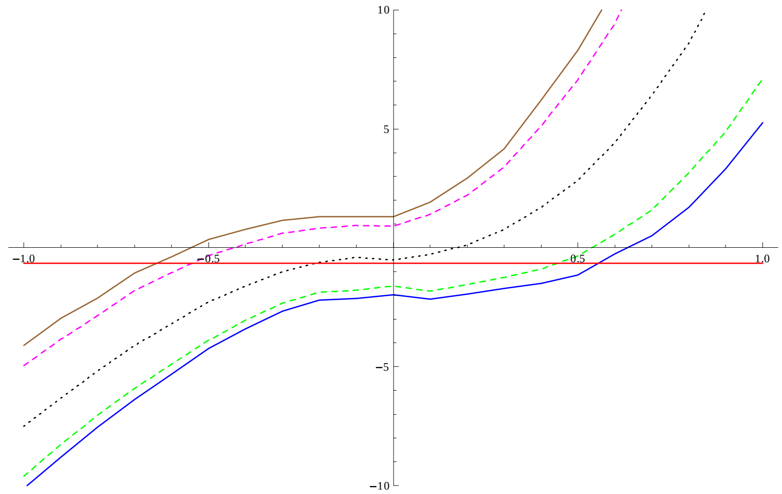

Figure 1: Asymptotic confidence belts for $t_{\hat{\beta}}$ - the LSTUR case. Y-axis - $t_{\hat{\beta}}$ values, X-axis - c values, solid blue - 5th percentile belt, dashed green - 10th percentile belt, dotted black - median belt, dashed magenta, 90th - percentile belt, solid brown - 95th - percentile belt, horizontal red line - the sample's $t_{\hat{\beta}}$, $a = -0.245$, $\rho = -0.150$, $\sigma_u^2 = 0.983$, $\sigma_\varepsilon^2 = 7 \times 10^{-5}$, $\hat{t}_\beta = -0.659$. Calculated with a grid step of 0.1, 400 integral points and 5000 replications.

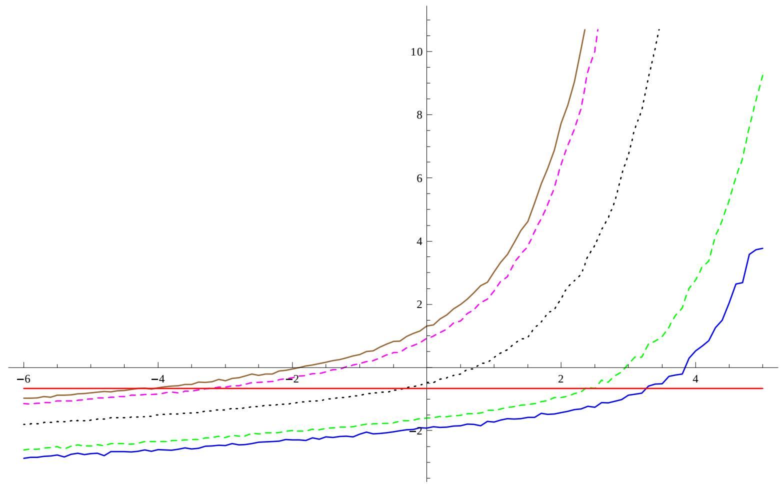

Figure 2: Asymptotic confidence belts for $t_{\hat{\beta}}$ - the LUR case. Y-axis - $t_{\hat{\beta}}$ values, X-axis - c values, solid blue - 5th percentile belt, dashed green - 10th percentile belt, dotted black - median belt, dashed magenta, 90th - percentile belt, solid brown - 95th - percentile belt, horizontal red line - the sample's $t_{\hat{\beta}}$, $a = 0$, $\sigma_{\varepsilon}^2 = 7 \times 10^{-5}$, $\hat{t}_{\beta} = -0.659$. Calculated with a grid step of 0.1, 400 integral points and 5000 replications.

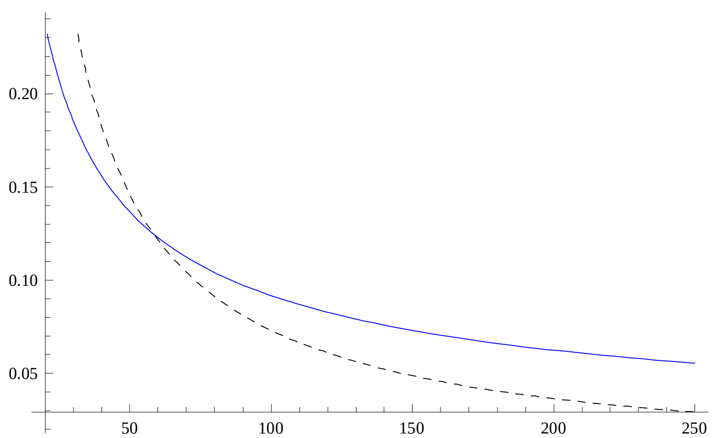

Figure 3: Solid blue - the width of the LSTUR-based 90% CI, dashed black - the width of the LUR - based 90% CI. Based on the data of Section 7.