12-1-2014

A Multivariate Stochastic Unit Root Model with an Application to Derivative Pricing

Offer Lieberman

Peter C.B. Phillips

A MULTIVARIATE STOCHASTIC UNIT ROOT MODEL WITH AN APPLICATION TO DERIVATIVE PRICING

By

Offer Lieberman and Peter C. B. Phillips

December 2014

COWLES FOUNDATION DISCUSSION PAPER NO. 1964

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

http://cowles.econ.yale.edu/

A Multivariate Stochastic Unit Root Model with an Application to Derivative Pricing*

Offer Lieberman† and Peter C. B. Phillips‡

Revised, September 18, 2014

Abstract

This paper extends recent findings of Lieberman and Phillips (2014) on stochastic unit root (SUR) models to a multivariate case including a comprehensive asymptotic theory for estimation of the model's parameters. The extensions are useful because they lead to a generalization of the Black-Scholes formula for derivative pricing. In place of the standard assumption that the price process follows a geometric Brownian motion, we derive a new form of the Black-Scholes equation that allows for a multivariate time varying coefficient element in the price equation. The corresponding formula for the value of a European-type call option is obtained and shown to extend the existing option price formula in a manner that embodies the effect of a stochastic departure from a unit root. An empirical application reveals that the new model is consistent with excess skewness and kurtosis in the price distribution relative to a lognormal distrubution.

Key words and phrases: Autoregression; Derivative; Diffusion; Options; Similarity; Stochastic unit root; Time-varying coefficients.

JEL Classification: C22

---

*Our thanks to the Editor and three referees for helpful comments and suggestions on the original manuscript.

†Bar-Ilan University. Support from Israel Science Foundation grant No. 1082-14 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. Support is acknowledged from the NSF under Grant No. SES 1258258.

1 Introduction

Unit root and local unit root time series models have attracted much attention in the last few decades, providing a wellspring of work that has been found useful in applied research in many disciplines, including finance. The prototype model

$$\begin{aligned} Y_1 &= \varepsilon_{1,} \\ Y_t &= \mu + \beta Y_{t-1} + \varepsilon_t,\ t = 2, ..., n,\ \ \varepsilon_t \overset{iid}{\sim} \left(0, \sigma_\varepsilon^2\right),\ t = 1, ..., n, \end{aligned} \tag{1}$$

has substantial flexibility and, when the autoregressive parameter β is in the vicinity of unity, data generated from the model take many plausible forms that include stationary, trend stationary, random wandering, and explosive possibilities. A key mechanism in determining the large sample limit form of the process is the invariance principle for standardized versions of partial sums of the innovations $S_{\lfloor nr \rfloor} = \sum_{t=1}^{\lfloor nr \rfloor} \varepsilon_t$, where $\lfloor nr \rfloor$ is the integer part of nr. The simplest case involves the Donsker result

$$\frac{1}{\sigma_\varepsilon \sqrt{n}} \sum_{t=1}^{\lfloor nr \rfloor} \varepsilon_t \Rightarrow W(r),\ r \in [0, 1], \tag{2}$$

where $W(r)$ is standard Brownian motion and $\Rightarrow$ denotes weak convergence, but much more general results are known to hold (e.g., Phillips, 1987a; see Giraitis *et. al.*, 2012, for a recent discussion). As is well known, the limit theory has implications for standardized versions of the output process Y_t when β is in the vicinity of unity.

To illustrate, let n be the number of subintervals into which a T-year period is subdivided, such that n/T is fixed for a given data-frequency, and let μ_A and $\sigma_{\varepsilon,A}$ denote the mean and standard deviation in annualized terms, so that

$$\mu_A = \frac{n}{T}\mu \text{ and } \sigma_{\varepsilon,A} = \sqrt{\frac{n}{T}}\sigma_\varepsilon. \tag{3}$$

Then, when $\beta = 1$, we have $Y_{\lfloor nr \rfloor} = \lfloor nr \rfloor \mu + \sum_{t=1}^{\lfloor nr \rfloor} \varepsilon_t$, and for large n this leads directly to

$$Y_{\lfloor nr \rfloor} \sim Y_n(r) = T\mu_A r + \sqrt{T}\sigma_{\varepsilon,A} W(r). \tag{4}$$

It is emphasized that even though the model (1) is written in terms of any frequency, the large sample behavior (4) is expressed in common annualized terms than involve μ_A and $\sigma_{\varepsilon,A}$. It is difficult to over-emphasize the role that this last formula plays in the literature. For instance, the celebrated Black-Scholes formula (Black and Scholes, 1973, henceforth BS) for option pricing, critically depends on the assumption that stock prices, $S(r)$, follow a geometric Brownian motion, viz.,

$$\frac{dS(r)}{S(r)} = T\mu_A dr + \sqrt{T}\sigma_{\varepsilon,A} dW(r) = dY_n(r). \tag{5}$$

A tacit assumption that leads to the limit theory embedded in (4) and (5) is that the coefficient β of Y_{t-1} in (1) is fixed and equals unity for all t. For some data sets and models, this assumption may be reasonable *on average*, but it is often likely to be restrictive. Recognition of this limitation has led to the consideration of local unit root (LUR) models where β is fixed (within an array framework) but lies in the vicinity of unity (Chan and Wei, 1987; Phillips, 1987b; Phillips and Magdalinos, 2007). A more realistic working hypothesis might relax the requirement that the coefficient be fixed and allow for some time variation and possible dependencies on other stochastic variables. Phillips and Yu (2011) explored some time variation in the localizing coefficient to collapse in financial markets and bubble migration but used a deterministic β. The stochastic localizing coefficient we use here has the form

$$\begin{aligned} Y_1 &= \mu + \varepsilon_1, \\ Y_t &= \mu + \beta_t(a;n) Y_{t-1} + \varepsilon_t, \ t = 2, ..., n, \end{aligned} \tag{6}$$

where

$$\beta_t(a;n) = \exp\left(\frac{a'u_t}{\sqrt{n}}\right) \tag{7}$$

and u_t is an $L\times 1$ vector and is the source of the variation in the autoregressive coefficient. In applications, u_t will typically stand for a vector of excess returns[1] on market indices and/or related stocks, but this need not be the case. A formal factor interpretation of (7) is possible in which the loading coefficients $a_n^* = a/\sqrt{n}$ are local to zero and the factors (observable and

[1] In other words, in applications u_t would typically be a demeaned $I(1)$ process.

unobservable) are measured through u_t, while the AR coefficient $\beta_t(a;n)$ is driven by the exponential function so that the model is a nonlinear factor formulation.

We assume that partial sums of $\eta_t = (u_t, \varepsilon_t)'$ satisfy the invariance principle

$$n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} \eta_t \Rightarrow B(r) \equiv \mathrm{BM}(\Sigma), \ \Sigma = \begin{pmatrix} \Sigma_u & \Sigma_{u\varepsilon} \\ \Sigma'_{u\varepsilon} & \sigma^2_\varepsilon \end{pmatrix}, \tag{8}$$

where $B = (B_u, B_\varepsilon)'$ is a vector Brownian motion with Σ positive definite and component $L \times L$ submatrix $\Sigma_u > 0$ and scalar $\sigma^2_\varepsilon > 0$. The parameters μ and Σ are the single-period mean and covariance matrix quantities, which are fixed for a given frequency and are to be distinguished from their annual- and T-year counterparts.

The model (6)-(7) is a multivariate version of a (single regressor) stochastic unit root (STUR) model introduced in Lieberman and Phillips (2014) and belongs to the general class of time varying coefficients (TVC) models. That paper explored the connection of the STUR model to recent developments in the literature, including similarity models, specifically, Lieberman (2010, 2012), who investigated autoregressive similarity-based models with non-stochastic regressors. The term 'similarity' originated from the theory of empirical similarity, developed in Gilboa *et. al.* (2006), and under which the value of $\beta_t(a;n)$ is dictated by the degree of similarity between Y_t and Y_{t-1}, as measured by the input u_t to the exponent of (7). The main feature of the STUR model is that for any given t, the coefficient $\beta_t(a;n)$ can be less than-, equal to-, or greater than unity, with a time specific value that is determined by u_t. We note that the random coefficient structure is at the heart of the arguments developed in Meyn and Tweedie (2005) and subsequent work on GARCH processes.

This paper derives the stochastic limit theory of the STUR model (6)-(7) and uses this limit theory to generalize the classic stochastic differential equation (sde) (5) so that it embodies the limit of (6)-(7). The special case where $a = 0$ produces the limit process (4) and so the new limit theory provides an extension of (5) to include a TVC feature. Within this framework, a further contribution of the paper is to derive the BS sde for derivative pricing and the BS price of a European call option under the new scheme. Furthermore, we provide a new asymptotic theory for estimation of all the model parameters and apply these results in the construction of option pricing formulae.

The idea of modifying the base model (5) to enhance realism is by no means new. Two main streams of extension appear in the literature. The first is the stochastic volatility (SV) model (e.g., Hull and White (1987), Heston (1993)). In that model, if the volatility process is not correlated with $W(r)$, the process is consistent with a symmetric volatility smile (Renault and Touzi (1996)), whereas if there is a negative correlation between the two, the process will be consistent with an asymmetric volatility skew, which is often claimed to be empirically better suited to stock options (see, for instance, Hull, 2009).

In the second stream of literature it is suggested to replace the standard Brownian motion driver process in (5) by a fractional Brownian motion (FBM), B^H, with a Hurst parameter H. See, for instance, Hu and Øksendal (2000) and Biagini *et. al.* (2008). While the FBM model is reported to fit certain data sets better than the base model (5), B^H has correlated increments and is not a semimartingale so that the model introduces arbitrage possibilities, classic Itō calculus is inapplicable and new methods of stochastic integration using Wick algebra are required, see Bjork and Hult (2005). In contrast, the extension based on a similarity STUR model allows for the use of standard Itō calculus and is convenient for analysis and empirical work.

In sum, the BS model (5) is simple, tractable and possesses features such as completeness and no-arbitrage but suffers limitations such as no implied volatility smile and the absence of heavy tails. These features of the BS model suggest that there is value in an extension of the model that captures its main advantages while overcoming its main empirical shortcomings.

The plan for the remainder of the paper is as follows. Section 2 develops the limit theory for the multivariate STUR model and provides the associated sde. Section 3 provides a comprehensive asymptotic theory of estimation of the model parameters, in both the $\mu = 0$ and the $\mu \neq 0$ cases. Section 4 derives the BS sde corresponding to our model and the price process associated with it. The value of a European call option under the new model is given in Section 5. A numerical section is supplied in Section 6, comprised of a simulation sub-section for the results of Section 3, and an empirical application, showing that our model is consistent with excess skewness and kurtosis, as compared with the lognormal distribution implied by the BS model. Section 7 concludes and proofs will be given in the Appendix.

2 Continuous Limit of the STUR Model

By backward substitution, the model (6) gives the following solution from initialization at $Y_1 = \mu + \varepsilon_1$,

$$Y_2 = (\beta_2 + 1)\,\mu + (\beta_2\varepsilon_1 + \varepsilon_2)\,,\ Y_3 = (\beta_3\beta_2 + \beta_3 + 1)\,\mu + (\beta_3\beta_2\varepsilon_1 + \beta_3\varepsilon_2 + \varepsilon_3)\,,$$

and generally for any $t \geq 2$,

$$\begin{aligned}
Y_t &= \left(\sum_{s=1}^{t-1}\left(\prod_{j=s+1}^{t}\beta_j\right)+1\right)\mu + \sum_{s=1}^{t-1}\left(\prod_{j=s+1}^{t}\beta_j\right)\varepsilon_s + \varepsilon_t \\
&= \left(\sum_{s=1}^{t-1}\left(\prod_{j=s+1}^{t}\beta_j\right)+1\right)\mu + \sum_{s=1}^{t-1} e^{\frac{a'}{\sqrt{n}}\sum_{j=s+1}^{t}u_j}\varepsilon_s + \varepsilon_t \qquad (9)\\
&=: h_t(\beta)\,\mu + Y_t^*, \qquad (10)
\end{aligned}$$

say. In what follows it is convenient to expand the probability space as necessary to ensure that the convergence in (8) is in probability so that $n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor}\eta_t \to_p B(r)$. This procedure, which is standard in modern asymptotic theory, enables limit variates such as the vector Brownian motion $B(r)$ (which typically escape from the underlying probability space through the action of the asymptotics) to be included in the same space as the random sequence. The space augmentation also allows for the convenient replacement of weak convergence by almost sure convergence or, as here, convergence in probability, which is another well-known device in modern asymptotic theory. For further information, readers are referred to Shorack and Wellner (1986, Theorem 4, pp. 47-48). Standardizing Y_t^* we then have the following result.

Lemma 1 *In a suitably expanded probability space as $n \to \infty$*

$$n^{-1/2}Y^*_{\lfloor nr\rfloor} \to_p e^{a'B_u(r)}\left(\int_0^r e^{-a'B_u(p)}dB_\varepsilon(p) - a'\Sigma_{u\varepsilon}\int_0^r e^{-a'B_u(p)}dp\right) := G_a(r)\,, \qquad (11)$$

and

$$\frac{1}{n}\left(\sum_{s=1}^{\lfloor nr\rfloor-1}\left(\prod_{j=s+1}^{\lfloor nr\rfloor}\beta_j\right)+1\right)\mu \to_p \mu e^{a'B_u(r)}\int_0^r e^{-a'B_u(p)}dp := H_a(r)\,. \qquad (12)$$

Using the differentials $d\left(e^{a'B_u(r)}\right) = e^{a'B_u(r)}\left\{a'dB_u\left(r\right) + \frac{1}{2}a'\Sigma_u a dr\right\}$ and

$$\begin{aligned} & d\left(\int_0^r e^{-a'B_u(p)}dB_\varepsilon\left(p\right) - a'\Sigma_{u\varepsilon}\int_0^r e^{-a'B_u(p)}dp\right) \\ = \; & e^{-a'B_u(r)}\left(dB_\varepsilon\left(r\right) - a'\Sigma_{u\varepsilon}dr\right), \end{aligned}$$

we find that $G_a\left(r\right)$ follows the sde

$$\begin{aligned} dG_a\left(r\right) &= e^{a'B_u(r)}\left(\int_0^r e^{-a'B_u(p)}dB_\varepsilon\left(p\right) - a'\Sigma_{u\varepsilon}\int_0^r e^{-a'B_u(p)}dp\right) \\ &\quad \times \left\{a'dB_u\left(r\right) + \frac{1}{2}a'\Sigma_u a dr\right\} + dB_\varepsilon\left(r\right) - a'\Sigma_{u\varepsilon}dr \\ &= G_a\left(r\right)a'dB_u\left(r\right) + dB_\varepsilon\left(r\right) + \left[\frac{a'\Sigma_u a}{2}G_a\left(r\right) - a'\Sigma_{u\varepsilon}\right]dr, \end{aligned} \tag{13}$$

which has the form of a nonlinear diffusion driven by vector Brownian motion $\left(B_u, B_\varepsilon\right)$. Observe that when $a = 0$, $G_a\left(r\right)$ reduces simply to the Brownian motion $B_\varepsilon\left(r\right)$.

It follows from (10) - (12) that for large n,

$$h_{\lfloor nr \rfloor}\left(\beta\right)\mu + Y^*_{\lfloor nr \rfloor} \sim Y_n\left(r\right) = n\mu e^{a'B_u(r)}\int_0^r e^{-a'B_u(p)}dp + \sqrt{n}G_a\left(r\right). \tag{14}$$

The first term in (14) contributes an additional drift $\left(n\mu dr\right)$ to the differential equation (13), leading to the following approximate continuous time law of motion for $Y_n\left(r\right)$

$$\begin{aligned} dY_n\left(r\right) &= n\mu dr + \sqrt{n}\left(G_a\left(r\right)a'dB_u\left(r\right) + dB_\varepsilon\left(r\right)\right. \\ &\quad \left. + \left[\frac{a'\Sigma_u a}{2}G_a\left(r\right) - a'\Sigma_{u\varepsilon}\right]dr\right). \end{aligned} \tag{15}$$

In this system the nonlinear diffusion process G_a affects both the martingale component and the drift. When $a = 0$, the system reduces to $dY_n\left(r\right) = n\mu dr + \sqrt{n}dB_\varepsilon\left(r\right)$, which corresponds in form to the classic equation (5).

3 Estimation of the Model Parameters

3.1 The case $\mu = 0$

This section develops asymptotic theory for the estimation of the model parameters. Let $\hat{a}_n$ denote the least squares estimator of a.

Theorem 2 *For the model (6)–(8) with $\mu = 0$, the asymptotic distribution of $\hat{a}_n$ is given by:*

(1)

$$(\hat{a}_n - a) \Rightarrow \frac{\int_0^1 G_a(r)\,dr}{\int_0^1 G_a^2(r)\,dr} \Sigma_u^{-1} \Sigma_{u\varepsilon}, \ \text{if } \Sigma_{u\varepsilon} \neq 0.$$

(2)

$$\hat{a}_n \Rightarrow \frac{\int_0^1 B_\varepsilon(r)\,dr}{\int_0^1 B_\varepsilon^2(r)\,dr} \Sigma_u^{-1} \Sigma_{u\varepsilon}, \ \text{if } \Sigma_{u\varepsilon} \neq 0 \text{ and } a = 0.$$

(3)

$$\begin{aligned} \sqrt{n}(\hat{a}_n - a) &\Rightarrow \frac{1}{\int_0^1 G_a^2(r)\,dr} \Sigma_u^{-1} \left\{ \{E(\varepsilon_t u_t u_t')\} a \int_0^1 G_a(r)\,dr \right. \\ &\quad \left. + E\left\{ (u_t' a)^2 u_t \right\} \int_0^1 G_a^2(r) + \int_0^1 G_a(r)\,dB_{u\varepsilon}(r) \right\}, \\ \text{if } \Sigma_{u\varepsilon} &= 0. \end{aligned} \tag{16}$$

(4)

$$\sqrt{n}\hat{a}_n \Rightarrow \frac{1}{\int_0^1 B_\varepsilon^2(r)\,dr} \Sigma_u^{-1} \int_0^1 B_\varepsilon(r)\,dB_{u\varepsilon}(r), \ \text{if } \Sigma_{u\varepsilon} = 0 \text{ and } a = 0.$$

The following properties are an immediate consequence of Theorem 2 in the case where $\mu = 0$.

Remark 1 *The estimate $\hat{a}_n$ is not consistent for a unless $\Sigma_{u\varepsilon} = 0$. Thus, endogeneity in u_t plays an important role in influencing the asymptotic behavior of $\hat{a}_n$.*

Remark 2 *Under the hypothesis $H_0 : a = 0$ and when $\Sigma_{u\varepsilon} \neq 0$, we may use Theorem 2(2). In particular, in the $L = 1$ subcase, we have*

$$\hat{a}_n \Rightarrow \frac{\rho \sigma_\varepsilon \int_0^1 B_\varepsilon (r)\, dr}{\sigma_u \int_0^1 B_\varepsilon^2 (r)\, dr}, \text{ if } \sigma_{u\varepsilon} \neq 0 \text{ and } a = 0,$$

where ρ is the correlation coefficient between ε and u.

Define the following estimates of σ_ε^2, Σ_u^2 and $\Sigma_{u\varepsilon}$

$$\hat{\sigma}_{\varepsilon,n}^2 = \frac{1}{n} \sum_{t=1}^{n} \left(Y_t - e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} \right)^2, \; vech\left(\hat{\Sigma}_{u,n}\right) = \frac{1}{n} \sum_{t=1}^{n} vech\left(u_t u_t'\right),$$

and

$$\hat{\Sigma}_{u\varepsilon,n} = \frac{1}{n} \sum_{t=1}^{n} \left(Y_t - e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} \right) u_t.$$

The next result gives the limit theory of these estimates.

Theorem 3 *For the model (6)–(8) with $\mu = 0$, the asymptotic distributions of $\hat{\sigma}_{\varepsilon,n}^2$, $vech\left(\hat{\Sigma}_{u,n}\right)$ and $\hat{\Sigma}_{u\varepsilon,n}$ are given by:*

(1)

$$\hat{\sigma}_{\varepsilon,n}^2 - \sigma_\varepsilon^2 \Rightarrow \frac{\left(\int_0^1 G_a (r)\, dr\right)^2}{\int_0^1 G_a^2 (r)\, dr} \Sigma_{u\varepsilon}' \Sigma_u^{-1} \Sigma_{u\varepsilon}.$$

(2)

$$\sqrt{n} \left(vech\left(\hat{\Sigma}_{u,n}\right) - vech\left(\Sigma_u\right) \right) \Rightarrow \xi (1),$$

where $\xi (r)$ is the Brownian motion weak limit of $\frac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor nr \rfloor} vech\left(u_t u_t' - \Sigma_u\right)$.

(3)

$$\hat{\Sigma}_{u\varepsilon,n} - \Sigma_{u\varepsilon} \Rightarrow - \frac{\left(\int_0^1 G_a (r)\, dr\right)^2}{\int_0^1 G_a^2 (r)\, dr} \Sigma_{u\varepsilon}.$$

Remark 3 *$\hat{\sigma}_{\varepsilon,n}^2$ is not consistent, unless $\Sigma_{u\varepsilon} = 0$. In the $a = 0$ case, Theo-*

rem 3(1) implies

$$\hat{\sigma}^2_{\varepsilon,n} - \sigma^2_\varepsilon \Rightarrow \frac{\left(\int_0^1 B_\varepsilon(r)\,dr\right)^2}{\int_0^1 B_\varepsilon^2(r)\,dr} \Sigma'_{u\varepsilon}\Sigma_u^{-1}\Sigma_{u\varepsilon}.$$

The restricted estimator $\hat{\sigma}^2_{\varepsilon,n,0} = n^{-1}\sum_{t=1}^n (Y_t - Y_{t-1})^2$, *is consistent for* σ^2_ε *under* $H_0 : a = 0$.

Remark 4 *In the case* $L = 1$, $\xi(1) =_d N(0, \kappa_4 + 2\sigma_u^4)$, *where* κ_4 *is the 4th cumulant of* u_t. *If* u_t *is Gaussian* $\xi(1) =_d N(0, 2\sigma_u^4)$.

Remark 5 $\hat{\Sigma}_{u\varepsilon,n}$ *is not consistent, unless* $\Sigma_{u\varepsilon} = 0$. *In the* $a = 0$ *case, Theorem 3(3) reduces to*

$$\hat{\Sigma}_{u\varepsilon,n} - \Sigma_{u\varepsilon} \Rightarrow -\frac{\left(\int_0^1 B_\varepsilon(r)\,dr\right)^2}{\int_0^1 B_\varepsilon^2(r)\,dr} \Sigma_{u\varepsilon}.$$

The restricted estimate $\hat{\Sigma}_{u\varepsilon,n,0} = n^{-1}\sum_{t=1}^n (Y_t - Y_{t-1})\, u_t$ *is consistent for* $\Sigma_{u\varepsilon}$ *under* $H_0 : a = 0$.

3.2 The case $\mu \neq 0$

Theorem 4 *For the model (6)–(8) with* $\mu \neq 0$, *the asymptotic distribution of* $\hat{a}_n$ *is given by:*

$$\sqrt{n}\,(\hat{a}_n - a) \Rightarrow \frac{\int_0^1 H_a(r)\,dr}{\int_0^1 H_a^2(r)\,dr} \Sigma_u^{-1}\Sigma_{u\varepsilon}.$$

The following remarks apply when $\mu \neq 0$:

Remark 6 *The estimator* $\hat{a}_n$ *is consistent for all* a *in contrast to the case where* $\mu = 0$. *The result is explained as follows. With no loss of generality set* $L = 1$. *The nonlinear model (6) is* $Y_t = \mu + e^{\frac{a u_t}{\sqrt{n}}} Y_{t-1} + \varepsilon_t$ $(t = 2, ..., n)$ *and may be written in linear pseudo-model form as follows*

$$\begin{aligned} \Delta Y_t &= \mu + \left\{ \frac{a u_t}{\sqrt{n}} + \frac{1}{2}\frac{a^2 u_t^2}{n} + \frac{1}{6}\frac{a^3 u_t^3}{n^{3/2}} + O_p\left(\frac{1}{n^2}\right) \right\} Y_{t-1} + \varepsilon_t \\ &= \mu + a\left[\frac{u_t}{\sqrt{n}} + \frac{a}{2}\frac{u_t^2}{n} + \frac{1}{6}\frac{a^2 u_t^3}{n^{3/2}} + O_p\left(\frac{1}{n^2}\right) \right] Y_{t-1} + \varepsilon_t + O_p\left(\frac{1}{n}\right), \end{aligned}$$

or $\Delta Y_t = \mu + aZ_t + \varepsilon_t$, *where*

$$Z_t = \frac{u_t Y_{t-1}}{\sqrt{n}} + \frac{a}{2}\frac{u_t^2 Y_{t-1}}{n} + \frac{1}{6}\frac{a^2 u_t^3 Y_{t-1}}{n^{3/2}} + O_p\left(\frac{Y_{t-1}}{n^2}\right).$$

When $\mu = 0$, $Y_t = O_p\left(\sqrt{n}\right)$. *In that case,* $Z_t = O_p(1)$ *and is correlated with the equation error* ε_t *when* $\Sigma_{u\varepsilon} \neq 0$, *which explains the inconsistency of* $\hat{a}_n$ *when* $\mu = 0$. *When* $\mu \neq 0$, *we have* $Y_t = O_p(n)$, *as shown in the proof of Lemma 1. In that case, the pseudo-regressor* $Z_t = O_p\left(\sqrt{n}\right)$ *and the stronger signal ensures that* $\hat{a}_n$ *is consistent in spite of the presence of correlation with the equation error* ε_t *when* $\Sigma_{u\varepsilon} \neq 0$. *Thus, drift in the generating mechanism plays an important role in the asymptotic properties of the estimate* $\hat{a}_n$.

Remark 7 *When* $\Sigma_{u\varepsilon} = 0$, $\sqrt{n}\left(\hat{a}_n - a\right) \Rightarrow 0$ *and the convergence rate of* $\hat{a}_n$ *exceeds* $\sqrt{n}$. *Again, the presence of endogeneity in* u_t *plays a role in the asymptotics of* $\hat{a}_n$.

Remark 8 *When* $a = 0$, $H_a(r) = \mu_0 r$ *and*

$$\sqrt{n}\hat{a}_n \Rightarrow \frac{3}{2\mu_0}\Sigma_u^{-1}\Sigma_{u\varepsilon}.$$

In the $a = 0$ *and* $L = 1$ *subcase, we have*

$$\sqrt{n}\hat{a}_n \Rightarrow \frac{3\rho\sigma_\varepsilon}{2\mu_0\sigma_u}. \tag{17}$$

Let $\hat{\mu}_n$ be the least squares estimate of μ and define

$$B(a, \Sigma_u) = H_a^*(1) - a' \int_0^1 H_a^*(r)\, dB_u(r) - \frac{1}{2}a'\Sigma_u a \int_0^1 H_a^*(r)\, dr,$$

where $H_a^*(r) = \mu^{-1}H_a(r)$.

Theorem 5 *For the model (6)–(8) with* $\mu \neq 0$,

$$\hat{\mu}_n \Rightarrow \mu B(a, \Sigma_u).$$

Remark 9 *By Theorems 4 and 6(2) below, since* $\left(\hat{a}_n, \hat{\Sigma}_u\right) \to_p (a, \Sigma_u)$ *when*

$\mu \neq 0$, the rescaled estimate

$$\mu_n^* \equiv \frac{\hat{\mu}_n}{B\left(\hat{a}_n, \hat{\Sigma}_u\right)}$$

is consistent for μ.

Remark 10 *When $a = 0$, $H_a^*(1) = 1$, leading to $\hat{\mu}_n \Rightarrow \mu$.*

In the $\mu \neq 0$ case, we define the following estimates of σ_ε^2, Σ_u^2 and $\Sigma_{u\varepsilon}$

$$\hat{\sigma}_{\varepsilon,n}^2 = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \mu_n^* - e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\right)^2, \quad vech\left(\hat{\Sigma}_{u,n}\right) = \frac{1}{n}\sum_{t=1}^{n} vech\left(u_t u_t'\right),$$

and

$$\hat{\Sigma}_{u\varepsilon,n} = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \mu_n^* - e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\right) u_t.$$

Theorem 6 *For the model (6)–(8) with $\mu \neq 0$, the asymptotic distributions of $\hat{\sigma}_{\varepsilon,n}^2$, $\hat{\sigma}_{u,n}^2$ and $\hat{\Sigma}_{u\varepsilon,n}$ are given by*

(1) $\hat{\sigma}_{\varepsilon,n}^2 - \sigma_\varepsilon^2 \Rightarrow \frac{\left(\int_0^1 H_a(r)dr\right)^2}{\int_0^1 H_a^2(r)dr} \Sigma_{u\varepsilon}' \Sigma_u^{-1} \Sigma_{u\varepsilon}$,

(2) $\sqrt{n}\left(\Sigma_{u,n}^2 - \sigma_u^2\right) \Rightarrow \xi(1)$, *and*

(3) $\hat{\Sigma}_{u\varepsilon,n} - \Sigma_{u\varepsilon} \Rightarrow -\frac{\left(\int_0^1 H_a(r)dr\right)^2}{\int_0^1 H_a^2(r)dr} \Sigma_{u\varepsilon}$.

4 A STUR Extension of the BS Model

4.1 The Price Process

A fundamental building block in the BS option price formula involves a random walk with a drift, which in the discrete case amounts to equation (1) with $\beta = 1$, i.e., a STUR process with parameter $a = 0$. The results of Section 2 suggest a generalization of the BS formula. To fix ideas, it is convenient to set the parameters as $\mu_T = n\mu$, and $\sigma_{\varepsilon,T} = \sqrt{n}\sigma_\varepsilon$, so that when $\beta = 1$, equation (4) becomes

$$Y_n(r) = \mu_T r + \sigma_{\varepsilon,T} W(r). \tag{18}$$

The subscripts A and T will be used in what follows to distinguish between annualized and T-year period quantities, respectively.

A key assumption in the BS option price model is that stock prices, $S(t)$, follow a geometric Brownian motion, viz.,

$$\frac{dS(r)}{S(r)} = T\mu_A dr + \sqrt{T}\sigma_{\varepsilon,A} dW(r) = \mu_T dr + \sigma_{\varepsilon,T} dW(r). \tag{19}$$

That is, the right side of (19) is just (18), which specializes (15) above in the case $a = 0$, thereby suggesting the latter as a suitable extension giving a geometric nonlinear diffusion limit process corresponding to a more flexible time varying coefficient discrete model. We use this extension to obtain derivative pricing formulae under weaker conditions than BS, including the price of a European call option.

Define $B^* := (B_u^*, B_\varepsilon^*)' = \Sigma^{-1/2} B$, so that B^* is vector standard Brownian motion (SBM). Write the lower triangular square root of Σ as

$$\Sigma^{1/2} := \begin{pmatrix} \left[\Sigma^{1/2}\right]_{1,1} & 0 \\ \left[\Sigma^{1/2}\right]_{2,1} & \left[\Sigma^{1/2}\right]_{2,2} \end{pmatrix} := \begin{pmatrix} \Sigma_u^{1/2} & 0 \\ \Sigma'_{u\varepsilon}\Sigma_u^{-1/2} & \left(\sigma_\varepsilon^2 - \Sigma'_{u\varepsilon}\Sigma_u^{-1}\Sigma_{u\varepsilon}\right)^{1/2} \end{pmatrix}$$
$$:= \begin{pmatrix} \left[\Sigma^{1/2}\right]_1 \\ \left[\Sigma^{1/2}\right]_2 \end{pmatrix},$$

where $\Sigma_u^{1/2}$ is the positive definite square root of Σ_u. Let $a_n = (T/n)^{1/2} a$, $\Sigma_{u\varepsilon,T} = n\Sigma_{u\varepsilon} = T\Sigma_{u\varepsilon,A}$, and $\Sigma_T = n\Sigma = T\Sigma_A$. We show in the Appendix that

$$Y_n(r) = T\mu_A e^{a_n'\Sigma_{u,A}^{1/2} B_u^*(r)} \int_0^r e^{-a_n'\Sigma_{u,A}^{1/2} B_u^*(p)} dp + \sqrt{T} G_{a_n,A}(r), \tag{20}$$

where we use $\sqrt{n} G_a(r) = \sqrt{T} G_{a_n,A}(r)$ and

$$\begin{aligned} G_{a_n,A}(r) &= e^{a_n'\Sigma_{u,A}^{1/2} B_u^*(r)} \left(\left[\Sigma_A^{1/2}\right]_2 \int_0^r e^{-a_n'\Sigma_{u,A}^{1/2} B_u^*(p)} dB^*(p) \right. \\ &\quad \left. - a_n'\Sigma_{u\varepsilon,A} \int_0^r e^{-a_n'\Sigma_{u,A}^{1/2} B_u^*(p)} dp \right). \end{aligned} \tag{21}$$

Let $b(r) = (G_a(r)\, a', 1)'$ and $b_A(r) = (G_{a_n,A}(r)\, a_n', 1)'$. Because

$$\sqrt{n} b(r)' \, dB(r) = \sqrt{n} b(r)' \, \Sigma^{1/2} dB^*(r) = \sqrt{T} b(r)' \, \Sigma_A^{1/2} dB^*(r),$$

and

$$G_a(r)\, a = \sqrt{n} G_a(r)\, a/\sqrt{n} = \sqrt{T} G_{a_n,A}(r)\, a/\sqrt{n} = G_{a_n,A}(r)\, a_n,$$

we get

$$\sqrt{n} b(r)' \, dB(r) = \sqrt{T} b_A(r)' \, \Sigma_A^{1/2} dB^*(r).$$

Thus, we may rewrite (15) as

$$\begin{aligned} dY_n(r) = & \left[T\mu_A + \sqrt{T} \left(\frac{a_n' \Sigma_{u,A} a_n}{2} G_{a_n,A}(r) - a_n' \Sigma_{u\varepsilon,A} \right) \right] dr \\ & + \sqrt{T} b_A(r)' \left[\Sigma_A^{1/2} \right] dB^*(r). \end{aligned} \tag{22}$$

We replace the standard formula (19) with (22), which leads to the following geometric nonlinear diffusion that is based on the STUR model

$$\begin{aligned} \frac{dS(r)}{S(r)} = & \left[T\mu_A + \sqrt{T} \left(\frac{a_n' \Sigma_{u,A} a_n}{2} G_{a_n,A}(r) - a_n' \Sigma_{u\varepsilon,A} \right) \right] dr \\ & + \sqrt{T} b_A(r)' \left[\Sigma_A^{1/2} \right] dB^*(r). \end{aligned} \tag{23}$$

Whereas (19) is a geometric Brownian motion, the system (23) is a geometric price process that involves the nonlinear diffusion $G_{a_n,A}(r)$ and Brownian motion driver process B^*. The system collapses to (19) when $a = 0$. So, (23) may be regarded as a process that is parametrically local to geometric Brownian motion.

Next, consider the process $G(r) = \log(S(r))$ and let $[S]_r$ denote the quadratic variation process of $S(r)$. We show in the Appendix that

$$\begin{aligned} dG(r) = & \left(T \left(\mu_A - \frac{\sigma_{\varepsilon,A}^2}{2} \right) - \sqrt{T} a_n' \Sigma_{u\varepsilon,A} \right) dr \\ & + \left(\sqrt{T} \left(\frac{a_n' \Sigma_{u,A} a_n}{2} - \sqrt{T} a_n' \Sigma_{u\varepsilon,A} \right) G_{a_n,A}(r) - T \frac{a_n' \Sigma_{u,A} a_n}{2} G_{a_n,A}^2(r) \right) dr \\ & + \sqrt{T} b_A(r)' \, \Sigma_A^{1/2} dB^*(r). \end{aligned} \tag{24}$$

When $a = 0$ and $\Sigma_{u\varepsilon} = 0$, we retain the classic formula

$$\begin{aligned} dG(r) &= \left(\mu_T - \frac{\sigma^2_{\varepsilon,T}}{2}\right) dr + \sqrt{n} dB_\varepsilon(r) \\ &= T\left(\mu_A - \frac{\sigma^2_{\varepsilon,A}}{2}\right) dr + \sqrt{T}\sigma_{\varepsilon,A} dB^*_\varepsilon(r). \end{aligned} \tag{25}$$

In this case, (25) implies that

$$\log(S(r)) - \log(S(0)) \sim N\left(\left(\mu_T - \frac{\sigma^2_{\varepsilon,T}}{2}\right) r, \sigma^2_{\varepsilon,T} r\right)$$

so that

$$S(r) = S(0)\exp\left\{N\left(\left(\mu_T - \frac{\sigma^2_{\varepsilon,T}}{2}\right) r, \sigma^2_{\varepsilon,T} r\right)\right\}. \tag{26}$$

In the Appendix we show that when $a \neq 0$, $S(r)$ satisfies

$$\begin{aligned} S(r) = \; & S(0)\exp\left\{\left(T\left(\mu_A - \frac{\sigma^2_{\varepsilon,A}}{2}\right) - \sqrt{T} a'_n \Sigma_{u\varepsilon,A}\right) r \right. \\ & + \int_0^r \left(\sqrt{T}\left(\frac{a'_n \Sigma_{u,A} a_n}{2} - \sqrt{T} a'_n \Sigma_{u\varepsilon,A}\right) G_{a_n,A}(s) \right. \\ & \left. - T\frac{a'_n \Sigma_{u,A} a_n}{2} G^2_{a_n,A}(s)\right) ds \\ & \left. + \sqrt{T} a'_n \int_0^r G_{a_n,A}(s)\, \Sigma^{1/2}_{u,A}(s)\, dB^*_u(s) + \sqrt{T}\Sigma^{1/2}_{2,A} B^*(r)\right\}. \end{aligned} \tag{27}$$

Equation (27) is the price process under the physical measure. It will be used in place of (26) when calculating BS option prices, after it is reformulated in terms of the risk neutral measure, Q, details of which are given in Section 6.

4.2 BS European Option Pricing

Following Hull (2009), let $f(r, x)$ be the price at r of a European-style derivative of a stock, such as a European call option, where $x \equiv S(r)$ is the stock price and $Z(r)$ is the price of the riskless asset. Let $(\alpha_S(r), \alpha_Z(r))$ be the associated self-financing portfolio at time r. The value of the portfolio at time

r is

$$V(r) = \alpha_S(r) S(r) + \alpha_Z(r) \gamma(r),$$

where $\gamma(r) = e^{r_{f,T} r}$ and $r_{f,T} = Tr_{f,A}$ is the period-T risk-free rate of interest. The sde corresponding to the portfolio $V(r)$ is

$$dV(r) = \alpha_S(r) dS(r) + \alpha_Z(r) r_{f,T} \gamma(r) dr, \tag{28}$$

which is the self-financing condition. We show in the Appendix that

$$\alpha_S(r) = f_x, \tag{29}$$

which is the condition in the classic case, and that

$$\alpha_Z(r) = \frac{1}{Tr_{f,A}\gamma(r)} \left\{ f_r + \frac{T}{2} f_{xx} S^2(r) b_A(r)' \Sigma_A b_A(r) \right\}. \tag{30}$$

When $a = 0$ the last condition collapses to the well-known condition

$$\alpha_Z(r) = \frac{1}{Tr_{f,A}\gamma(r)} \left\{ f_r + \frac{T}{2} f_{xx} S^2(r) \sigma^2_{\varepsilon,A} \right\}.$$

Since $V(r) = f(r,x)$, it follows that $\alpha_S(r) S(r) + \alpha_Z(r) \gamma(r) = f(r,x)$, so that using (29) and (30) we obtain

$$f_x S(r) + \frac{1}{Tr_{f,A}\gamma(r)} \left\{ f_r + \frac{T}{2} f_{xx} S^2(r) b_A(r)' \Sigma_A b_A(r) \right\} \gamma(r) = f,$$

and, finally,

$$Tr_{f,A} f_x S(r) + f_r + \frac{T}{2} f_{xx} S^2(r) b_A(r)' \Sigma_A b_A(r) = Tr_{f,A} f. \tag{31}$$

Equation (31) is the generalized BS sde for a European style derivative of a stock. When $a = 0$ the formula reduces to the well-known relationship

$$Tr_{f,A} f_x S(r) + f_r + \frac{T}{2} f_{xx} S^2(r) \sigma^2_{\varepsilon,A} = Tr_{f,A} f.$$

When $K = 1$ and $a \neq 0$, (31) becomes

$$\begin{aligned} & Tr_{f,A} f_x S(r) + f_r + \frac{T}{2} f_{xx} S^2(r) \left\{ G^2_{a_n,A}(r)\, a_n^2 \sigma^2_{u,A} \right. \\ & \left. +2 a_n \sigma_{u\varepsilon,A} G_{a_n,A}(r) + \sigma^2_{\varepsilon,A} \right\} \\ = \; & Tr_{f,A} f. \end{aligned}$$

While the sde's are not used in the sequel to price options, their derivation is of some independent interest, showing how the famous BS sde's are generalized in our model. Notably, the sde for the SV model in equation (6) of Heston (1993) includes extra terms (relative to BS) which involve partial derivatives of the asset with respect to the volatility process, whereas the above sde's involve extra terms (relative to BS) which are functionals of the new limit process $G_a(r)$ reflecting the impact of time variation in the discrete model's autoregressive coefficient.

4.3 Market Incompleteness

In model (6) the time varying coefficient $\beta_t(a; n) = \exp\left(\frac{a' u_t}{\sqrt{n}}\right)$ introduces an additional source of uncertainty in the generating mechanism. The vector u_t leads to variation in the autoregressive coefficient β_t that typically has unforecastable elements, thereby introducing an additional source of risk to investors concerning the price generating mechanism. In applications, u_t may, for instance, carry the import of economy-wide common shocks or index movements that affect returns indirectly via the generating mechanism itself rather than through the equation error shocks ε_t, although these two shocks may well be correlated. If there are shocks to the way price evolves from the previous price so that the mechanism is not a martingale and the conditional expectation is not the immediately preceding price, the market is inefficient. This inefficiency may be interpreted as a form of market incompleteness because the factors and shocks embodied in u_t comprise additional unforecastable states of nature that are uncovered in the market. These shocks involve risk beyond that of the simple equation error. In the continuous time context, the shocks are manifest in the additional random component, $\sqrt{T}\left(\frac{a_n' \Sigma_{u,A} a_n}{2} G_{a_n,A}(r) - a_n' \Sigma_{u\varepsilon,A}\right) dt$, that appears in the drift of the log price stochastic differential equation. In effect, the TVC model implies uncertainty and risk in the price process drift, producing an additional source of uncertainty/risk in investment alpha.

5 The Value of a European Call Option

At time 0 the value of a European call option maturing in T years is given by

$$C = e^{-r_{f,T}} E^Q \max\{S_T - K, 0\},$$

where Q is the risk-neutral measure (defined immediately after equation (27)). To calculate the expectation in the last equation we need to find $S_T^Q \equiv S^Q(1)$, the price at time T under Q. To clarify the notation we remark that for any random variable X, by $E^Q(X)$ or by $E\left(X^Q\right)$ we mean the expected value of X under Q. Under risk-neutral pricing, $S^Q(r)$ must be a martingale with respect to Q, so that

$$E^Q(S(r)) = E\left(S^Q(r)\right) = S(0)\, e^{r_{f,T} r}.$$

In other words, under Q the price process develops from $S(0)$ at the risk-free rate r_f. To find $S^Q(r)$, we first make the following transformation:

$$\begin{pmatrix} B_u^*(r) \\ B_\varepsilon^*(r) \end{pmatrix} = \begin{pmatrix} B_u^{*Q}(r) \\ B_\varepsilon^{*Q}(r) + \gamma r \end{pmatrix}, \gamma \in \mathbb{R},$$

where the superscript Q indicates that the SBM's are under Q. As

$$dB_\varepsilon^*(r) = dB_\varepsilon^{*Q}(r) + \gamma dr \tag{32}$$

and as

$$\begin{aligned} \left[\Sigma_A^{1/2}\right]_2 B^*(p) &= \left[\Sigma_A^{1/2}\right]_{2,1} B_u^*(p) + \left[\Sigma_A^{1/2}\right]_{2,2} B_\varepsilon^*(p) \\ &= \left[\Sigma_A^{1/2}\right]_2 B^{Q*}(p) + \left[\Sigma_A^{1/2}\right]_{2,2} \gamma r, \end{aligned} \tag{33}$$

it follows from (21) that

$$\begin{aligned} G_{a_n,A}(r) &= \gamma e^{a_n' \Sigma_{u,A}^{1/2} B_u^{*Q}(r)} \left[\Sigma_A^{1/2}\right]_{2,2} \int_0^r e^{-a_n' \Sigma_{u,A}^{1/2} B_u^{*Q}(p)} dp + G_{a_n,A}^Q(r) \\ &= \gamma \xi^Q(r) + G_{a_n,A}^Q(r), \end{aligned} \tag{34}$$

say, where $G^{Q}_{a_n,A}(r)$ is $G_{a_n,A}(r)$ with $B^{*Q}_{u}(r)$ and $B^{*Q}_{\varepsilon}(r)$ in the former replacing $B^{*}_{u}(r)$ and $B^{*}_{\varepsilon}(r)$ in the latter, respectively. We obtain from (27)

$$\begin{aligned} S^{Q}(r) &= S(0)\exp\left\{\left(T\left(\mu_A-\frac{\sigma^2_{\varepsilon,A}}{2}\right)-\sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)r\right. \\ &\quad +\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}-\sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)\int_0^r\left(\gamma\xi^{Q}(s)+G^{Q}_{a_n,A}(s)\right)ds \\ &\quad -T\frac{a_n'\Sigma_{u,A}a_n}{2}\int_0^r\left(\gamma\xi^{Q}(s)+G^{Q}_{a_n,A}(s)\right)^2ds \\ &\quad +\sqrt{T}a_n'\Sigma^{1/2}_{u,A}\int_0^r\left(\gamma\xi^{Q}(s)+G^{Q}_{a_n,A}(s)\right)dB^{Q*}_{u}(s) \\ &\quad \left.+\sqrt{T}\left(\left[\Sigma^{1/2}_{A}\right]_2B^{*Q}(r)+\left[\Sigma^{1/2}_{A}\right]_{2,2}\gamma r\right)\right\} \\ &= S(0)\,\phi^{Q}_{\gamma}(r)\,\lambda^{Q}(r), \end{aligned}$$

where

$$\begin{aligned} \phi^{Q}_{\gamma}(r) &= \exp\left\{T\mu_A-\sqrt{T}a_n'\Sigma_{u\varepsilon,A}r\right. \\ &\quad +\gamma\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}-\sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)\int_0^r\xi^{Q}(s)\,ds \\ &\quad -\gamma^2T\frac{a_n'\Sigma_{u,A}a_n}{2}\int_0^r\left(\xi^{Q}(s)\right)^2ds \\ &\quad -\gamma Ta_n'\Sigma_{u,A}a_n\int_0^rG^{Q}_{a_n,A}(s)\,\xi^{Q}(s)\,ds \\ &\quad +\gamma\sqrt{T}a_n'\Sigma^{1/2}_{u,A}\int_0^r\xi(s)\,dB^{*Q}_{u}(s) \\ &\quad \left.+\gamma r\sqrt{T}\left[\Sigma^{1/2}_{A}\right]_{2,2}\right\} \end{aligned}$$

and

$$
\begin{aligned}
\lambda^{Q}(r) = \exp\Bigg\{ & -\frac{\sigma^2_{\varepsilon,A}}{2}Tr \\
& + \int_0^r \left(\sqrt{T}\left(\frac{a_n' \Sigma_{u,A} a_n}{2} - \sqrt{T} a_n' \Sigma_{u\varepsilon,A} \right) G^{Q}_{a_n,A}(s) \right. \\
& \left. - T\frac{a_n' \Sigma_{u,A} a_n}{2} \left(G^{Q}_{a_n,A}(s) \right)^2 \right) ds \\
& + \sqrt{T}\left(a_n' \Sigma^{1/2}_{u,A} \int_0^r G^{Q}_{a_n,A}(s)\, dB^{Q*}_u(s) + \left[\Sigma^{1/2}_A\right]_2 B^{Q*}(r) \right) \Bigg\}.
\end{aligned}
\tag{35}
$$

Now, set $\gamma = \tilde{\gamma}$, such that $\phi^{Q}_{\tilde{\gamma}}(1) = 1$, and let

$$
\mu^{Q}_{\lambda}(1) = e^{-r_{f,A}T} E\left(\lambda^{Q}(1)\right).
$$

The price process under Q at $r = 1$ is defined as

$$
S^{Q}(1) = S(0)\frac{\lambda^{Q}(1)}{\mu^{Q}_{\lambda}(1)} = S(0)\, e^{r_{f,A}T} \bar{\lambda}^{Q}(1),
$$

with

$$
\bar{\lambda}^{Q}(1) = \frac{\lambda^{Q}(1)}{E\left(\lambda^{Q}(1)\right)}
$$

Evidently,

$$
E\left(S^{Q}(1)\right) = S(0)\, e^{r_{f,A}T},
$$

as required.

Note that in the $a = 0$ and $\Sigma_{u\varepsilon} = 0$ case,

$$
\phi^{Q}_{\gamma}(1) = \exp\left(T\mu_A + \sqrt{T}\gamma\sigma_{\varepsilon} \right),
$$

$$
\lambda^{Q}_{BS}(1) = \exp\left(-\frac{\sigma^2_{\varepsilon,A}}{2}T + \sqrt{T}\sigma_{\varepsilon,A} B^{Q*}_{\varepsilon}(1) \right) \tag{36}
$$

and

$$
\begin{aligned}
\mu^{Q}_{\lambda}(1) &= e^{-r_{f,A}T} E\left(\lambda^{Q}_{BS}(1)\right) \\
&= e^{-r_{f,A}T}.
\end{aligned}
$$

Hence, in this special case,

$$S^{Q}(1) = S(0)\frac{\lambda^{Q}(1)}{\mu_{\lambda}^{Q}(1)} = S(0)\exp\left\{\left(r_{f,A} - \frac{\sigma_{\varepsilon,A}^{2}}{2}\right)T + \sqrt{T}\sigma_{\varepsilon,A}^{2}B_{\varepsilon}^{Q*}(1)\right\},$$

as is well-known.

A further adjustment can be made such that the mean and variance of a transformed $\lambda^{Q}(1)$ will be the same as of the lognormal distribution (corresponding to BS pricing). To do so, let

$$\lambda^{*Q}(1) = \frac{\left(e^{\sigma_{\varepsilon,A}^{2}T} - 1\right)^{1/2}}{\sigma\left(\frac{\lambda^{Q}(1)}{E\left(\lambda^{Q}(1)\right)}\right)}\left(\frac{\lambda^{Q}(1)}{E\left(\lambda^{Q}(1)\right)} - 1\right) + 1.$$

Then $E\left(\lambda_{T}^{*Q}\right) = 1$ and $Var\left(\lambda_{T}^{*Q}\right) = e^{\sigma_{\varepsilon,A}^{2}T} - 1$, giving the mean and variance of $\lambda_{BS}^{Q}(1)$, which is valid under BS pricing. The mean- and variance adjusted price process under Q is then given by

$$S^{*Q}(1) = S(0)\exp\left(r_{f,A}T\right)\lambda^{*Q}(1). \tag{37}$$

Such a transformation is desirable because the resulting process, $S^{*Q}(1)$, differs from the price process under BS only with respect to the term $\lambda^{*Q}(1)$, vs. $\lambda_{BS}^{Q}(1)$, and these two differ only in skewness, kurtosis and higher order moments. Hull (Ch. 18, 2009) argues that the information embodied in the skewness and kurtosis measures for processes that have the same mean and variance as BS can be translated into information on volatility smiles.

In order to simulate the value of a European call option maturing in T years, we calculate $G_{a_n,A}^{Q}(r)$ over the grid $\{r = 0, 1/n, 2/n, ..., 1\}$, after simulating a vector Brownian motion driver process and noting that T corresponds to $r = 1$. We then calculate the sample mean of $\max\left\{S^{Q}(1) - K, 0\right\}$, denoted by $\overline{\max\left\{S^{Q}(1) - K, 0\right\}}$, using a large number of replications. The estimated European call option price is

$$\hat{C} = e^{-r_{f,T}}\overline{\max\left\{S^{Q}(1) - K, 0\right\}}, \tag{38}$$

or

$$\hat{C}^{*} = e^{-r_{f,T}}\overline{\max\left\{S^{*Q}(1) - K, 0\right\}} \tag{39}$$

if the refinement (37) is preferred.

6 Numerical Analysis

6.1 Simulations

We downloaded from Yahoo Finance daily data on the closing prices of Google and Nasdaq composite indexes (tickers GOOG and ^IXIC), over the period 1-2-2009 through to 11-20-2013, giving a total of 1231 observations for each series. This sample period is post the 2008 market crash, so we avoid possible issues of structural breaks in the illustrations.

With the obvious notation, we obtained estimates of the following empirical STUR model

$$\log\widehat{(Google)}_t = 0.000941 + \exp\left(\frac{4.9540\left(\Delta\log\left(Nasdaq\right)_t - 0.000713\right)}{\sqrt{1231}}\right)\log\left(Google\right)_{t-1}, \tag{40}$$

where 0.000713 is the estimated daily return of $\Delta\log\left(Nasdaq\right)_t$ over the sample period. We have also obtained the estimates $\hat{\sigma}^2_{u,A} = 0.2112$, $\hat{\sigma}^2_{\varepsilon,A} = 0.0716$ and $\hat{\rho}_{\varepsilon,u} = 0.6923$, over the same sample period.

For a hypothesis test of the form $H_0 : a = 0$, when $\sigma_{u\varepsilon} \neq 0$, we know from Remark 8 that

$$\sqrt{n}\hat{a}_n \Rightarrow \frac{3\rho\sigma_\varepsilon}{2\mu_0\sigma_u} = \frac{3n\left(\sigma_{u\varepsilon,A}\right)}{2T\left(\sigma^2_{u,A}\right)\left(\mu_{0,A}\right)}.$$

With daily data, $n/T = 252$. Under H_0, the model is a random walk with a drift and therefore the historical estimates of μ and Σ are consistent. To verify the results of Section 3 we generated 250 samples of $n = 1000, 2000, ..., 250000$ (daily) observations from a random walk model with a drift (i.e., the null model with $a = 0$). The value of μ was taken to be $\hat{\mu} = 0.000941$. The disturbance term was generated from a bivariate normal distribution with (daily) covariance matrix consistent with the historical estimates mentioned above, by dividing the annual covariance estimates by 252. For each of the 250 samples we have obtained $\hat{a}_n$ and consequently, we computed the sample average (based on the 250 samples) of

$$R_n = \frac{\sqrt{n}\hat{a}_n}{3\rho\sigma_\varepsilon/2\sigma_u\mu_0}, n = 1000, 2000, ..., 250000. \tag{41}$$

We obtained

$$\bar{R}_n = \frac{1}{250}\sum R_n = 0.9991,$$

confirming the result in Remark 8. The graph of R_n against n is provided in Figure 1.

For the same output, we also obtained the following regression results

$$(\widehat{\log|\hat{a}_n|}) = 7.328 - 0.508\log(n), \; R^2 = 0.978.$$

The implication of both Figure 1 and the last equation is that $|\hat{a}_n| \sim n^{1/2}$, corroborating the decay rate stated in Remark 8.

Moving on in a similar way except that the true value of a is $a_0 = 0.15$, we generated[2] 250 samples of $n = 1000, 2000, ..., 250000$ (daily) observations from our process, obtaining a plot of $\sqrt{n}\,(\hat{a}_n - a_0)$ against n in Figure 2, which confirms the stated $\sqrt{n}$-rate. As a further check, we ran the log-log regression for this scenario, obtaining

$$(\widehat{\log|\hat{a}_n - a_0|}) = 7.642 - 0.533\log(n), \; R^2 = 0.973,$$

again corroborating the result of Theorem 4 that $|\hat{a}_n - a_0| \sim n^{-1/2}$.

To complete this subsection, we also estimated the model with $\mu = 0$, obtaining

$$\log\widehat{(Google)}_t = \exp\left(\frac{4.9544\,(\Delta\log(Nasdaq)_t - 0.000713)}{\sqrt{1231}}\right)\log(Google)_{t-1}.$$

Using the same historical estimates for the covariance matrix as in the $\mu \neq 0$ case, we generated 100,000 replications of the right side of Theorem 2(2) with $n = 1231$ integral points. The resulting kernel density estimate is given in Figure 3. The distribution is evidently bi-modal with a local minimum at zero.

[2] In this setup the estimates of μ and of Σ are inconsistent, but are used as if they were the true values in order to assess the finite sample implications of Theorem 4.

6.2 An Empirical Application

We start with the single regressor STUR model (40), expanding it to a two-regressor application later on. The Akaike (AIC), Schwarz (SC) and sum of squared errors (SSE) values for the model (40) are -5.976, -5.968, and 0.182, respectively. On the other hand, for the model (1) the figures are -5.327, -5.319 and 0.349, respectively, and for the model (1) with $\beta = 1$ the corresponding values are -5.327, -5.323 and 0.349. Thus, in terms of selection criteria, the STUR model provides a clear improvement over the basic model.

The estimate of a in (40) is 4.954. Running with an increasing sequence of subsamples $n = 100, 200, ..., 1231$ reveals that $\sqrt{n}\hat{a}_n$ increases. Thus, using (17) we would reject the hypothesis $H_0 : a = 0$. This result, together with the selection criteria figures, justifies the use of the TVC model over (1).

For the annual risk-free rate we used the Federal Funds Rate from the Bloomberg website quoted as $r_{f,A} = 0.0007$ (0.07% per annum). The 3-month treasury yield was identical (in annual terms). Other choices of the risk-free rate, such as the 12 month treasury yield (0.11%) or the 2-year yield (0.29%), which may be better suited to use for options with longer expiration periods might also be considered. But for this empirical illustration we have used the Federal Funds Rate.

The practice in this illustration is similar to the one described in Christoffersen and Jacobs (2004, p. 1207) in the sense that first we obtain consistent estimates of the model parameters using the stock return data and consequently plug in these estimates instead of the corresponding parameters under the risk neutral measure.

Google is a non-dividend paying stock and is therefore suited to this application. We took expirations on 1-3-2014 (36 days), 1-17-2015 (415 days) and 1-15-2016 (778 days) and considered strike prices $K = 950$, 1030, 1060, 1100, 1200. The closing price for Goggle stock on 11-27-2013 was 1063.11.

For each scenario, we have calculated the BS classic call option prices (see, e.g., equation (13.20) of Hull (2009)) in Mathematica. To simulate $G_{a_n,A}(r)$, we generated a vector of two standard Brownian motions scaled by $\hat{\Sigma}_A^{1/2}$. For T, we substituted the number of days to expiration divided by 365. The prices $S^Q(1)$ were evaluated with 500 integral points and 2000 replications over simulated Brownian motions.

We remark that the estimate $\hat{a}_n$ was obtained from daily data. In order to simulate $S^Q(1)$ for any T and n, and thereby for any implied data frequency, we need to take account of the time dimensionality of the parameter. To this end, we recall that $a_n = (T/n)^{1/2}\, a$. Moreover, had we estimated (40) with data of any other frequency, with n_f observations over the same sample period, least squares estimation would have yielded

$$\hat{a}_{n,d} = \frac{n}{n_f}\hat{a}_{n,f},$$

where $\hat{a}_{n,d}$ and $\hat{a}_{n,f}$ are the least squares estimates of a based on the daily and general-frequency data, respectively. With the consistency result of Theorem 4, this means that with our daily data, for any T and n we can replace a_n by

$$\sqrt{\frac{T}{n_f}}\frac{n_f}{n}\hat{a}_{n,d} = \frac{\sqrt{Tn_f}}{n}\hat{a}_{n,d}.$$

For the illustration, we calculated (38) and (39), each with the historical ρ, denoted by $\hat{\rho}$, but also with preset values $\rho = 0$ and $\rho = 0.95$, in order to investigate the effect of the correlation on the results. Finally, we also evaluated (38) and (39) after fitting a multivariate extension to (40), viz.,

$$\begin{aligned}\log\widehat{(Google)}_t &= 0.000944 \\ &\quad + \exp\left(\frac{4.5903\,(\Delta\log(Nasdaq)_t - 0.000713)}{\sqrt{1231}}\right. \\ &\quad \left. + \frac{0.3921\,(\Delta\log(AAPL)_t - 0.001411)}{\sqrt{1231}}\right)\log(Google)_{t-1}, \quad (42)\end{aligned}$$

where Apple inc. (ticker AAPL) had an average return of 0.001411 over the sample period.

The results are presented in Table 1. A few comments are in order. As the pricing schemes only differ from each other by the terms $\lambda^{Q}_{BS}(1)$, $\bar{\lambda}^{Q}(1)$ and $\lambda^{*Q}(1)$, an investigation of their behavior will be sufficient for the explanation of the price differences. First, $\hat{C}$ is mostly larger than BS and the converse holds for $\hat{C}^{*}$. The reason for the former is that the simulated standard deviation of $\bar{\lambda}^{Q}(1)$ is larger than that of $\lambda^{Q}_{BS}(1)$. We remark that the standard deviation of $\lambda^{*Q}(1)$ is equal to that of $\lambda^{Q}_{BS}(1)$ by construction. Second, the skewness and kurtosis coefficients of both $\bar{\lambda}^{Q}(1)$ and $\lambda^{*Q}(1)$ (which are equal to each other by construction) are greater than those of

$\lambda_{BS}^{Q}(1)$. Hull (Ch 18, 2009) argues that these results are consistent with a volatility skew. Thirdly, there is no noticeable pattern of the results with respect to ρ. Finally, there is not much difference between the single-regressor and two-regressor results for small T, and a small difference for large T.

In Table 2 we investigate the moments of $\lambda_{BS}^{Q}(1)$, $\bar{\lambda}^{Q}(1)$ and $\lambda^{*Q}(1)$, where clearly, excess skewness and kurtosis of $\bar{\lambda}^{Q}(1)$ and $\lambda^{*Q}(1)$ relative to $\lambda_{BS}^{Q}(1)$ are visible. Kernel density estimates of these terms are provided in Figure 4-5, emphasizing the peakedness of $\lambda^{*Q}(1)$ relative to $\lambda_{BS}^{Q}(1)$.

Overall, it appears that the TVC feature of the model results in a superior performance over the basic model in terms of the usual AIC, SC and SSE criteria, as well as a significant $\hat{a}_n$ estimate and peaked distribution which is more consistent with empirical findings.

7 Conclusions

The time-varying coefficient model is a natural extension of the simple AR(1) model in which, at any given time period, the coefficient of the lag dependent variable can be less than-, equal to- or greater than unity depending on a vector of unobserved factors with local to zero loading coefficients. Unlike the local to unit root model in which the coefficient converges to unity as the sample size tends to infinity, in our model the effect of the stochastic coefficient does not vanish as the sample size increases. As a result, the limit process is not geometric Brownian motion but a nonlinear diffusion. The new model and limit theory provides a generalization to Black-Scholes call option pricing.

We have established asymptotic theory for estimates of the model parameters. As expected, in the driftless case with endogeneity in the factors ($\Sigma_{u\varepsilon} \neq 0$), the estimate of the localizing coefficient $\hat{a}_n$ is inconsistent. As an empirical illustration, the results were applied in the pricing of Google's call options. In this application the results show that the price process under the new model (and under the risk neutral measure) has excess skewness and kurtosis compared with the lognormal distribution.

References

Biagini, F., Y. Hu, B. Øksendal and T. Zhang (2008). *Stochastic Calculus for Fractional Brownian Motion and Applications*, Springer, Probability and Its Applications.

Björk, T. and H. Hult (2005). A note on Wick products and the fractional Black-Scholes model. SSE/EFI Working Paper Series in Economics and Finance, No. 596.

Black, F. and M. Scholes (1973). The Pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.

Chan, N. H. and C. Z. Wei (1987). Asymptotic Inference for Nearly Nonstationary AR(1) Processes, *Annals of Statistics* 15, 1050–1063.

Christoffersen, P. and K. Jacobs (2004). Which GARCH model for option valuation? *Management Science* 50, 1204–1221.

Gilboa, I., O. Lieberman and D. Schmeidler (2006). Empirical similarity. *The Review of Economics and Statistics* 88, 433–444.

Giraitis, L., H. L. Koul and D. Surgailis (2012). *Large Sample Inference for Long Memory Processes.* London: Imperial College Press.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* 6, 327–343.

Hu, Y. and B. Øksendal (2003). Fractional white noise calculus and applications to finance. *Infinite Dimensional Analysis Quantum Probability and Related Topics* 6, 1–32.

Hull, J. C. (2009). *Options, Futures, and Other Derivatives*, 7th Edition. Pearson Education International, New Jersey.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance* 42, 281–300.

Lieberman, O. (2010). Asymptotic theory for empirical similarity models. *Econometric Theory* 26, 1032–1059.

Lieberman, O. (2012). A similarity-based approach to time-varying coefficient nonstationary 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. Mimeo.

Meyn, S. P. and R. L. Tweedie (2005). *Markov Chains and Stochastic Stability*. Springer-Verlag, London.

Phillips, P. C. B. (1987a). "Time Series Regression with a Unit Root," *Econometrica,* 55, 277–302.

Phillips, P. C. B. (1987). "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika* 74, 535–547.

Phillips, P. C. B., T. Magdalinos (2007). "Limit Theory for Moderate Deviations from a Unit Root," *Journal of Econometrics* 136, 115-130.

Phillips P. C. B. and J. Yu (2011). "Dating the Timeline of Financial Bubbles during the Subprime Crisis," *Quantitative Economics,* 2, pp. 455-491.

Renault, E. and N. Touzi (1996). Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance* 6, 279–302.

Shorack, G. R. and J. A. Wellner (1986). Empirical Processes with Applications to Statistics. Wiley, New York.

Appendix

Proof Lemma 1. For $t = \lfloor ns \rfloor$ for any $s > 0$, we have

$$n^{-1/2} \sum_{j=1}^{t} \eta_j = B\left(t/n\right) + o_p\left(1\right),$$

so that

$$\begin{aligned} n^{-1/2} Y_t^* &= n^{-1/2} \sum_{s=1}^{t-1} e^{\frac{a'}{\sqrt{n}} \sum_{j=s+1}^{t} u_j} \varepsilon_s + O_p\left(n^{-1/2}\right) \\ &= n^{-1/2} e^{\frac{a'}{\sqrt{n}} \sum_{j=1}^{t} u_j} \sum_{s=1}^{t-1} e^{-\frac{a'}{\sqrt{n}} \sum_{j=1}^{s} u_j} \varepsilon_s + O_p\left(n^{-1/2}\right) \\ &= n^{-1/2} e^{\{a' B_u(t/n) + o_p(1)\}} \sum_{s=1}^{t-1} e^{\left\{-\frac{a'}{\sqrt{n}} \sum_{j=0}^{s-1} u_j - \frac{a'}{\sqrt{n}} u_s\right\}} \varepsilon_s + O_p\left(n^{-1/2}\right) \end{aligned}$$

$$
\begin{aligned}
&= n^{-1/2} e^{a' B_u(t/n)} \\
&\quad \times \sum_{s=1}^{t-1} e^{-\{a' B_u((s-1)/n) + o_p(1)\}} \left(1 - \frac{a' u_s}{\sqrt{n}} + O_p\left(n^{-1}\right)\right) \varepsilon_s + o_p(1) \\
&= e^{a' B_u(t/n)} \sum_{s=1}^{t-1} e^{-a' B_u((s-1)/n)} \left(\frac{\sum_{j=1}^{s} \varepsilon_j}{\sqrt{n}} - \frac{\sum_{j=1}^{s-1} \varepsilon_j}{\sqrt{n}} \right) \\
&\quad - e^{a' B_u(t/n)} \sum_{s=1}^{t-1} e^{-a' B_u((s-1)/n)} \left(\frac{a' u_s \varepsilon_s}{n} \right) + o_p(1). \qquad (43)
\end{aligned}
$$

Setting $t = \lfloor nr \rfloor$ and noting that $\mathbb{E}\left(e^{-a' B_u(p)}\right)^2 < \infty$, the first term on the right hand side (rhs) of (43) has limit

$$
e^{a' B_u\left(\frac{t}{n}\right)} \sum_{s=1}^{t-1} e^{-a' B_u\left(\frac{s-1}{n}\right)} dB_\varepsilon\left(\frac{s}{n}\right) \to_p e^{a' B_u(r)} \int_0^r e^{-a' B_u(p)} dB_\varepsilon(p) := G_a^*(r), \qquad (44)
$$

and the second term is

$$
\begin{aligned}
&-e^{a' B_u(t/n)} \sum_{s=1}^{t-1} e^{-a' B_u\left(\frac{s-1}{n}\right)} \left(\frac{a' u_s \varepsilon_s}{n} \right) = -a' e^{a' B_u(t/n)} \sum_{s=1}^{t-1} e^{-a' B_u\left(\frac{s-1}{n}\right)} \\
&\times \left(\frac{u_s \varepsilon_s - \Sigma_{u\varepsilon}}{n} + \frac{\Sigma_{u\varepsilon}}{n} \right) \\
&= -a' \Sigma_{u\varepsilon} e^{a' B_u(t/n)} \frac{1}{n} \sum_{s=1}^{t-1} e^{-a' B_u((s-1)/n)} + O_p\left(n^{-1/2}\right) \\
&\to_p - a' \Sigma_{u\varepsilon} e^{a' B_u(r)} \int_0^r e^{-a' B_u(p)} dp.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
n^{-1/2} Y_{\lfloor nr \rfloor}^* \to_p \quad & G_a^*(r) - a' \Sigma_{u\varepsilon} e^{a' B_u(r)} \int_0^r e^{-a' B_u(p)} dp \\
&= e^{a' B_u(r)} \left(\int_0^r e^{-a' B_u(p)} dB_\varepsilon(p) - a' \Sigma_{u\varepsilon} \int_0^r e^{-a' B_u(p)} dp \right) \\
&=: G_a(r), \qquad (45)
\end{aligned}
$$

as required for (11). Next,

$$\frac{1}{n}\left\{\sum_{s=1}^{t-1}\left(\prod_{j=s+1}^{t}\beta_j\right)+1\right\}\mu = \frac{1}{n}\left(\sum_{s=1}^{t-1}e^{\frac{a'}{\sqrt{n}}\sum_{j=s+1}^{t}u_j}+1\right)\mu$$
$$\to_p \mu e^{a'B_u(r)}\int_0^r e^{-a'B_u(p)}dp,$$

giving (12). ■

Proof Theorem 2. For the case $\mu = 0$, the objective function is

$$Q_n(a) = \sum_{t=2}^{n}\{Y_t - \beta_t(a)Y_{t-1}\}^2. \tag{46}$$

In this model, letting $t = \lfloor nr \rfloor$,

$$\frac{Y_t}{\sqrt{n}} \Rightarrow e^{a'B_u(r)}\left\{\int_0^r e^{-a'B_u(p)}dB_\varepsilon(p) - a'\Sigma_{u\varepsilon}\int_0^r e^{-a'B_u(p)}dp\right\} =: G_a(r). \tag{47}$$

Minimizing (46) with respect to a yields

$$\begin{aligned}
\dot{Q}_n(\hat{a}_n) &= -2\sum_{t=2}^{n}\{Y_t - \beta_t(\hat{a}_n)Y_{t-1}\}\dot{\beta}_t(\hat{a}_n)Y_{t-1} = 0 \\
&\Longrightarrow \sum_{t=2}^{n}\{Y_t - \beta_t(\hat{a}_n)Y_{t-1}\}u_t\beta_t(\hat{a}_n)Y_{t-1} = 0 \\
&\Longrightarrow \sum_{t=2}^{n}Y_tu_t\beta_t(\hat{a}_n)Y_{t-1} = \sum_{t=2}^{n}u_t\beta_t^2(\hat{a}_n)Y_{t-1}^2 \\
&\Longrightarrow \sum_{t=2}^{n}u_t\left(1+\frac{\hat{a}_n'u_t}{\sqrt{n}}+O_p\left(n^{-1}\right)\right)Y_tY_{t-1} \qquad (48)\\
&= \sum_{t=2}^{n}u_t\left(1+\frac{2\hat{a}_n'u_t}{\sqrt{n}}+O_p\left(n^{-1}\right)\right)Y_{t-1}^2.
\end{aligned}$$

The first term on the lhs of (48) is

$$\begin{aligned}
\sum_{t=2}^{n} u_t Y_t Y_{t-1} &= \sum_{t=2}^{n} u_t \left\{\beta_t\left(a\right) Y_{t-1} + \varepsilon_t\right\} Y_{t-1} \\
&= n \sum_{t=2}^{n} u_t \beta_t\left(a\right) \frac{Y_{t-1}^2}{n} + \sqrt{n} \sum_{t=2}^{n} u_t \varepsilon_t \frac{Y_{t-1}}{\sqrt{n}} \\
&\sim n \sum_{t=2}^{n} u_t \left(1 + \frac{a' u_t}{\sqrt{n}}\right) \frac{Y_{t-1}^2}{n} + \sqrt{n} \sum_{t=2}^{n} u_t \varepsilon_t \frac{Y_{t-1}}{\sqrt{n}} \\
&\sim n^{3/2} \int_0^1 G_a^2\left(r\right) dB_u\left(r\right) + n^{3/2} \Sigma_u a \int_0^1 G_a^2\left(r\right) dr \\
&+ n^{3/2} \Sigma_{u\varepsilon} \int_0^1 G_a\left(r\right) dr + n \int_0^1 G_a\left(r\right) dB_{u\varepsilon}\left(r\right) 1\left\{\Sigma_{u\varepsilon} = 0\right\}.
\end{aligned} \tag{49}$$

The second term on the left hand side (lhs) of (48) is

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{t=2}^{n} u_t u_t' \hat{a}_n Y_t Y_{t-1} &= \frac{1}{\sqrt{n}} \sum_{t=2}^{n} u_t u_t' \hat{a}_n \left\{\beta_t\left(a\right) Y_{t-1} + \varepsilon_t\right\} Y_{t-1} \\
&\sim \frac{1}{\sqrt{n}} \left\{ n \sum_{t=2}^{n} \left(1 + \frac{a' u_t}{\sqrt{n}}\right) \frac{Y_{t-1}^2}{n} u_t u_t' + \sum_{t=2}^{n} \varepsilon_t Y_{t-1} u_t u_t' \right\} \hat{a}_n \\
&\sim \frac{1}{\sqrt{n}} \left\{ n^2 \left(\frac{1}{n} \sum_{t=2}^{n} \frac{Y_{t-1}^2}{n}\right) \Sigma_u + \sqrt{n} \sum_{t=2}^{n} \frac{Y_{t-1}^2}{n} \left(u_t' a\right) u_t u_t' \right. \\
&\left. + n^{3/2} E\left(u_t u_t' \varepsilon_t\right) \left(\frac{1}{n} \sum_{t=2}^{n} \frac{Y_{t-1}}{\sqrt{n}}\right) \right\} \hat{a}_n \\
&\sim \frac{1}{\sqrt{n}} \left\{ n^2 \Sigma_u \int_0^1 G_a^2\left(r\right) dr + n^{3/2} E\left\{\left(u_t' a\right) u_t u_t'\right\} \int_0^1 G_a^2\left(r\right) \right. \\
&\left. + n^{3/2} E\left(\varepsilon_t u_t u_t'\right) \int_0^1 G_a\left(r\right) dr \right\} \hat{a}_n
\end{aligned}$$

$$\sim n^{3/2}\Sigma_u \hat{a}_n \int_0^1 G_a^2(r)\,dr + n\left(E\left\{(u_t'a)\,u_t u_t'\right\}\right)\hat{a}_n \int_0^1 G_a^2(r) \tag{50}$$
$$+\,n\left\{E\left(\varepsilon_t u_t u_t'\right)\right\}\hat{a}_n \int_0^1 G_a(r)\,dr.$$

Combining (49) and (50), the lhs of (48) behaves as

$$n^{3/2}\left\{\int_0^1 G_a^2(r)\,dB_u(r) + \Sigma_u\left(\hat{a}_n + a\right)\int_0^1 G_a^2(r)\,dr + \Sigma_{u\varepsilon}\int_0^1 G_a(r)\,dr\right\} \tag{51}$$
$$+\,n\left\{\left\{E\left(\varepsilon_t u_t u_t'\right)\right\}\hat{a}_n \int_0^1 G_a(r)\,dr + \left(E\left\{(u_t'a)\,u_t u_t'\right\}\right)\hat{a}_n \int_0^1 G_a^2(r)\right.$$
$$\left.+\int_0^1 G_a(r)\,dB_{u\varepsilon}(r)\,1\left\{\sigma_{u\varepsilon}=0\right\}\right\}.$$

The first term on the rhs of (48) is

$$\sum_{t=2}^{n} u_t Y_{t-1}^2 \sim n^{3/2}\int_0^1 G_a^2(r)\,dB_u(r) \tag{52}$$

and the second term on the rhs of (48) is

$$\frac{2}{\sqrt{n}}\sum_{t=2}^{n} u_t u_t' \hat{a}_n Y_{t-1}^2 \sim 2\Sigma_u \hat{a}_n n^{3/2}\int_0^1 G_a^2(r)\,dr. \tag{53}$$

Using (52) and (53), the rhs of (48) behaves as

$$n^{3/2}\left\{\int_0^1 G_a^2(r)\,dB_u(r) + 2\Sigma_u \hat{a}_n \int_0^1 G_a^2(r)\,dr\right\}. \tag{54}$$

Equating (51) to (54), we obtain

$$
\begin{aligned}
& n^{3/2}\left(\int_0^1 G_a^2(r)\,dB_u(r)+\Sigma_u(\hat{a}_n+a)\int_0^1 G_a^2(r)\,dr+\Sigma_{u\varepsilon}\int_0^1 G_a(r)\,dr\right)\\
& +n\left\{\{E(\varepsilon_t u_t u_t')\}\hat{a}_n\int_0^1 G_a(r)\,dr+(E\{(u_t'a)\,u_t u_t'\})\,\hat{a}_n\int_0^1 G_a^2(r)\right.\\
& \left.+\int_0^1 G_a(r)\,dB_{u\varepsilon}(r)\,1\{\sigma_{u\varepsilon}=0\}\right\}\\
= \; & n^{3/2}\left(\int_0^1 G_a^2(r)\,dB_u(r)+2\Sigma_u\hat{a}_n\int_0^1 G_a^2(r)\,dr\right).
\end{aligned}
$$

In the $\Sigma_{u\varepsilon}\neq 0$ case, the limit is seen to be

$$
(\hat{a}_n-a)\Rightarrow\frac{\int_0^1 G_a(r)\,dr}{\int_0^1 G_a^2(r)\,dr}\Sigma_u^{-1}\Sigma_{u\varepsilon}, if\Sigma_{u\varepsilon}\neq 0,
$$

whereas, in the $\Sigma_{u\varepsilon}=0$ case, we obtain

$$
\begin{aligned}
\sqrt{n}(\hat{a}_n-a) &\Rightarrow \frac{1}{\int_0^1 G_a^2(r)\,dr}\Sigma_u^{-1}\left\{\{E(\varepsilon_t u_t u_t')\}\,a\int_0^1 G_a(r)\,dr\right. \\
&\quad \left.+E\left\{(u_t'a)^2u_t\right\}\int_0^1 G_a^2(r)+\int_0^1 G_a(r)\,dB_{u\varepsilon}(r)\right\}, \\
\text{if }\Sigma_{u\varepsilon} &= 0.
\end{aligned}
\tag{55}
$$

Furthermore, in the $\Sigma_{u\varepsilon}=0$ and $a=0$ case, the limit is

$$
\sqrt{n}\hat{a}_n\Rightarrow\frac{1}{\int_0^1 G_a^2(r)\,dr}\Sigma_u^{-1}\int_0^1 G_a(r)\,dB_{u\varepsilon}(r)\,,\text{ if }\Sigma_{u\varepsilon}=0\text{ and }a=0
$$

which reduces to

$$
\sqrt{n}\hat{a}_n\Rightarrow\frac{1}{\int_0^1 B_\varepsilon^2(r)\,dr}\Sigma_u^{-1}\int_0^1 B_\varepsilon(r)\,dB_{u\varepsilon}(r)\,,\text{ if }\Sigma_{u\varepsilon}=0\text{ and }a=0.
$$

■

Proof Theorem 3. For part (1), we have

$$\hat{\sigma}_{\varepsilon,n} = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\right)^2$$

$$= \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - e^{a' u_t/\sqrt{n}} e^{(\hat{a}_n - a)' u_t/\sqrt{n}} Y_{t-1}\right)^2$$

$$= \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - e^{a' u_t/\sqrt{n}}\left(1 + \frac{(\hat{a}_n - a)' u_t}{\sqrt{n}}\right.\right.$$

$$\left.\left.+\frac{\left((\hat{a}_n - a)' u_t\right)^2}{2n} + o_p\left(\frac{1}{n}\right)\right) Y_{t-1}\right)^2$$

$$= \frac{1}{n}\sum_{t=1}^{n}\left(\varepsilon_t - e^{a' u_t/\sqrt{n}}\left(\frac{(\hat{a}_n - a)' u_t}{\sqrt{n}}\right.\right.$$

$$\left.\left.+\frac{\left((\hat{a}_n - a)' u_t\right)^2}{2n} + o_p\left(\frac{1}{n}\right)\right) Y_{t-1}\right)^2$$

$$= \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2$$

$$-\frac{2}{n}\sum_{t=1}^{n}\varepsilon_t e^{a' u_t/\sqrt{n}}\left(\frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + \frac{\left((\hat{a}_n - a)' u_t\right)^2}{2n} + o_p\left(\frac{1}{n}\right)\right) Y_{t-1}$$

$$+\frac{1}{n}\sum_{t=1}^{n} e^{2a' u_t/\sqrt{n}}\left(\frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + \frac{\left((\hat{a}_n - a)' u_t\right)^2}{2n} + o_p\left(\frac{1}{n}\right)\right)^2 Y_{t-1}^2. \quad (56)$$

The first term on the rhs of (56) converges in probability to σ_ε^2, whereas the second term becomes

$$-\frac{2}{n}\sum_{t=1}^{n}\varepsilon_t\left(1 + \frac{a' u_t}{\sqrt{n}} + o_p\left(\frac{1}{\sqrt{n}}\right)\right)\left(\frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + \frac{\left((\hat{a}_n - a)' u_t\right)^2}{2n}\right.$$

$$\left.+ o_p\left(\frac{1}{n}\right)\right) Y_{t-1}.$$

The dominant terms in the last expression are equal to

$$\begin{aligned}
&= -\frac{2}{n}\left(\sum_{t=1}^{n}\left(\varepsilon_t\frac{(\hat{a}_n-a)' u_t}{\sqrt{n}}Y_{t-1}+\varepsilon_t\frac{\left((\hat{a}_n-a)' u_t\right)^2}{2n}Y_{t-1}\right)\right. \\
&\quad \left.\sum_{t=1}^{n}\left(\varepsilon_t\frac{a'u_t(\hat{a}_n-a)' u_t}{n}Y_{t-1}+\varepsilon_t\frac{a'u_t\left((\hat{a}_n-a)' u_t\right)^2}{2n^{3/2}}\right)Y_{t-1}\right) \\
&= -\frac{2}{\sqrt{n}}\left\{(\hat{a}_n-a)'\int_0^1 G_a(r)\,dB_{u\varepsilon}(r)\right. \\
&\quad +\frac{1}{2}(\hat{a}_n-a)' E\left(\varepsilon_t u_t u_t'\right)(\hat{a}_n-a)\int_0^1 G_a(r)\,dr \\
&\quad \left.+a'E\left(\varepsilon_t u_t u_t'\right)(\hat{a}_n-a)\int_0^1 G_a(r)\,dr+o_p(1)\right\} \\
&\to_p 0.
\end{aligned}$$

The leading term in the third term on the rhs of (56) is seen to be

$$\begin{aligned}
\frac{1}{n^2}\sum_{t=1}^{n}\left((\hat{a}_n-a)' u_t\right)^2 Y_{t-1}^2 &= (\hat{a}_n-a)'\left(\frac{1}{n}\sum_{t=1}^{n}u_t u_t'\left(\frac{Y_{t-1}}{\sqrt{n}}\right)^2\right)(\hat{a}_n-a) \\
&= (\hat{a}_n-a)'\Sigma_u(\hat{a}_n-a)\int_0^1 G_a^2(r)\,dr+o_p(1).
\end{aligned} \tag{57}$$

Using Theorem 2(1), (57) becomes

$$\begin{aligned}
&\left(\frac{\int_0^1 G_a(r)\,dr}{\int_0^1 G_a^2(r)\,dr}\Sigma_{u\varepsilon}'\Sigma_u^{-1}\right)\Sigma_u\left(\Sigma_u^{-1}\Sigma_{u\varepsilon}\frac{\int_0^1 G_a(r)\,dr}{\int_0^1 G_a^2(r)\,dr}\right)\int_0^1 G_a^2(r)\,dr+o_p(1) \\
\Rightarrow\ & \frac{\left(\int_0^1 G_a(r)\,dr\right)^2}{\int_0^1 G_a^2(r)\,dr}\Sigma_{u\varepsilon}'\Sigma_u^{-1}\Sigma_{u\varepsilon}.
\end{aligned}$$

It follows that

$$\hat{\sigma}_{\varepsilon,n}^2-\sigma_\varepsilon^2 \Rightarrow \frac{\sigma_{u\varepsilon}^2\left(\int_0^1 G_a(r)\,dr\right)^2}{\sigma_u^2\int_0^1 G_a^2(r)\,dr}\Sigma_{u\varepsilon}'\Sigma_u^{-1}\Sigma_{u\varepsilon}.$$

For part (2), we have

$$\begin{aligned} vech\left(\hat{\Sigma}_{u,n}\right) &= \frac{1}{n}\sum_{t=1}^{n} vech\left(u_t u_t'\right) \\ &= \frac{1}{n}\sum_{t=1}^{n} vech\left(u_t u_t' - \Sigma_u + \Sigma_u\right) \\ &= vech\left(\Sigma_u\right) + \frac{1}{n}\sum_{t=1}^{n} vech\left(u_t u_t' - \Sigma_u\right), \end{aligned}$$

so that

$$\sqrt{n}\left(vech\left(\hat{\Sigma}_{u,n}\right) - vech\left(\Sigma_u\right)\right) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} vech\left(u_t u_t' - \Sigma_u\right) \Rightarrow \xi\left(1\right).$$

Let e_t be the least squares residual from the regression (6). For part (3), we have

$$\begin{aligned} \hat{\Sigma}_{u\varepsilon,n} &= \frac{1}{n}\sum_{t=1}^{n} e_t u_t \\ &= \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\right) u_t \\ &= \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - e^{a' u_t/\sqrt{n}} e^{(\hat{a}_n - a)' u_t/\sqrt{n}} Y_{t-1}\right) u_t \\ &= \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - e^{a' u_t/\sqrt{n}}\right. \\ &\quad \left.\left(1 + \frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + \frac{\left((\hat{a}_n - a)' u_t\right)^2}{2n} + o_p\left(\frac{1}{n}\right)\right) Y_{t-1}\right) u_t. \end{aligned} \tag{58}$$

The dominant terms in the last expression become

$$\frac{1}{n}\sum_{t=1}^{n}\left(\varepsilon_t-\left(1+\frac{a'u_t}{\sqrt{n}}\right)\right.$$

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

$$=\Sigma_{u\varepsilon}-\frac{1}{n}\sum_{t=1}^{n}\left(\frac{(\hat{a}_n-a)'u_t}{\sqrt{n}}Y_{t-1}u_t+\frac{\left((\hat{a}_n-a)'u_t\right)^2}{2n}Y_{t-1}u_t\right.$$

$$\left.+\frac{a'u_t(\hat{a}_n-a)'u_t}{n}Y_{t-1}u_t+\frac{a'u_t\left((\hat{a}_n-a)'u_t\right)^2}{2n^{3/2}}Y_{t-1}u_t\right)+o_p(1). \tag{59}$$

By $[\Sigma_u]_j$ we denote the jth row of Σ_u, which is also equal to the jth column of Σ_u. The first term in the brackets of (59) behaves as

$$\begin{aligned}-\sum_{j=1}^{K}(\hat{a}_n-a)_j\frac{1}{n}\sum_{t=1}^{n}u_{j,t}u_t\frac{Y_{t-1}}{\sqrt{n}} &= -\sum_{j=1}^{K}(\hat{a}_n-a)_j[\Sigma_u]_j\frac{1}{n}\sum_{t=1}^{n}\frac{Y_{t-1}}{\sqrt{n}}+o_p(1)\\ &= -\Sigma_u(\hat{a}_n-a)\int_0^1 G_a(r)dr+o_p(1)\\ &\Rightarrow -\left(\int_0^1 G_a(r)dr\right)\Sigma_u\Sigma_u^{-1}\Sigma_{u\varepsilon}\frac{\int_0^1 G_a(r)\,dr}{\int_0^1 G_a^2(r)\,dr}\\ &= -\frac{\left(\int_0^1 G_a(r)\,dr\right)^2}{\int_0^1 G_a^2(r)\,dr}\Sigma_{u\varepsilon}.\end{aligned}$$

All other terms in the brackets of (59) are negligible and therefore

$$\hat{\Sigma}_{u\varepsilon,n}-\Sigma_{u\varepsilon}\Rightarrow-\frac{\left(\int_0^1 G_a(r)\,dr\right)^2}{\int_0^1 G_a^2(r)\,dr}\Sigma_{u\varepsilon}.$$

■

Proof Theorem 4. In the $\mu \neq 0$ case, the ols estimator of μ is equal to

$$\hat{\mu}_n = \bar{Y}_n - \frac{1}{n}\sum_{t=2}^{n} e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}$$

and the plugged-in score function is

$$\begin{aligned}\dot{Q}_n\left(\hat{a}_n;\hat{\mu}_n\right) &= -\frac{2}{\sqrt{n}}\sum_{t=2}^{n}\Bigg\{\left(Y_t-\bar{Y}_n\right)-\Bigg(e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} \\ &\quad -\frac{1}{n}\sum_{t=2}^{n} e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\Bigg)\Bigg\} u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} \\ &= 0.\end{aligned}$$

This implies that

$$\begin{aligned}&\sum_{t=2}^{n}\left(Y_t-\bar{Y}_n\right) u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} \\ &= \sum_{t=2}^{n}\left(e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} - \frac{1}{n}\sum_{t=2}^{n} e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\right) u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}.\end{aligned} \tag{60}$$

It follows from Lieberman and Phillips (2014) that in this case

$$\frac{Y_t}{n} \Rightarrow \mu_0 e^{a' B_u(r)} \int_0^r e^{-a' B_u(p)} dp \equiv H_a\left(r\right).$$

Now,

$$\sum_{t=2}^{n} Y_t u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} = \sum_{t=2}^{n}\left(\mu_0 + e^{a' u_t/\sqrt{n}} Y_{t-1} + \varepsilon_t\right) u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}.$$

We have

$$\mu_0 \sum_{t=2}^{n} u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} \sim \mu_0 \sum_{t=2}^{n} u_t \left(1 + \frac{\hat{a}_n' u_t}{\sqrt{n}}\right) Y_{t-1}$$

$$\sim \mu_0 \left(n^{3/2} \int_0^1 H_a(r)\, dB_u(r) + \frac{1}{\sqrt{n}} \sum_{t=2}^{n} u_t u_t' \hat{a}_n Y_{t-1} \right)$$

$$\sim \mu_0 \left(n^{3/2} \int_0^1 H_a(r)\, dB_u(r) \right.$$

$$\left. + \frac{1}{\sqrt{n}} \Sigma_u \hat{a}_n n^2 \left(\frac{\sum_{t=2}^{n} Y_{t-1}}{n^2} \right) \right)$$

$$\sim \mu_0 n^{3/2} \left(\int_0^1 H_a(r)\, dB_u(r) + \Sigma_u \hat{a}_n \int_0^1 H_a(r)\, dr \right). \tag{61}$$

Also,

$$\sum_{t=2}^{n} u_t e^{(\hat{a}_n + a)' u_t/\sqrt{n}} Y_{t-1}^2 \sim \sum_{t=2}^{n} u_t Y_{t-1}^2 + \frac{1}{\sqrt{n}} \sum_{t=2}^{n} u_t u_t' (\hat{a}_n + a) Y_{t-1}^2$$

$$\sim n^{5/2} \int_0^1 H_a^2(r)\, dB_u(r) + \frac{1}{\sqrt{n}} \Sigma_u (\hat{a}_n + a) n^3 \sum_{t=2}^{n} \frac{Y_{t-1}^2}{n^3}$$

$$\sim n^{5/2} \left(\int_0^1 H_a^2(r)\, dB_u(r) + \Sigma_u (\hat{a}_n + a) \int_0^1 H_a^2(r)\, dr \right) \tag{62}$$

and

$$\sum_{t=2}^{n} \varepsilon_t u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} \sim \sum_{t=2}^{n} \varepsilon_t u_t Y_{t-1} + \frac{1}{\sqrt{n}} \sum_{t=2}^{n} \varepsilon_t u_t u_t' \hat{a}_n Y_{t-1}$$

$$\sim \Sigma_{u\varepsilon} n^2 \int_0^1 H_a(r)\, dr + \{E(\varepsilon_t u_t u_t')\} \hat{a}_n n^{3/2} \int_0^1 H_a(r)\, dr \tag{63}$$

$$+ n^{3/2} \int_0^1 H_a(r)\, dB_{u\varepsilon}(r). \tag{64}$$

It follows from (61)-(63) that the first component of the lhs of (60) behaves as

$$
\begin{aligned}
&\mu_0 n^{3/2}\left(\int_0^1 H_a(r)\,dB_u(r)+\Sigma_u\hat{a}_n\int_0^1 H_a(r)\,dr\right)\\
&+n^{5/2}\left(\int_0^1 H_a^2(r)\,dB_u(r)+\Sigma_u(\hat{a}_n+a)\int_0^1 H_a^2(r)\,dr\right)\\
&+\Sigma_{u\varepsilon}n^2\int_0^1 H_a(r)\,dr+\{E(\varepsilon_t u_t u_t')\}\hat{a}_n n^{3/2}\int_0^1 H_a(r)\,dr\\
&+n^{3/2}\int_0^1 H_a(r)\,dB_{u\varepsilon}(r)\\
&\sim n^{5/2}\left(\int_0^1 H_a^2(r)\,dB_u(r)+\Sigma_u(\hat{a}_n+a)\int_0^1 H_a^2(r)\,dr\right) \qquad (65)\\
&+\Sigma_{u\varepsilon}n^2\int_0^1 H_a(r)\,dr.
\end{aligned}
$$

The second component of the lhs of (60) is

$$
\begin{aligned}
\bar{Y}_n\sum_{t=2}^n u_t e^{\hat{a}_n' u_t/\sqrt{n}}Y_{t-1} &\sim n\int_0^1 H_a(r)\,dr\left(\sum_{t=2}^n u_t Y_{t-1}+\frac{1}{\sqrt{n}}\sum_{t=2}^n u_t u_t'\hat{a}_n Y_{t-1}\right)\\
&\sim n\int_0^1 H_a(r)\,dr\left(n^{3/2}\int_0^1 H_a(r)\,dB_u(r)\right.\\
&\left.+\Sigma_u\hat{a}_n n^{3/2}\int_0^1 H_a(r)\,dr\right)\\
&= n^{5/2}\int_0^1 H_a(r)\,dr\left(\int_0^1 H_a(r)\,dB_u(r)\right. \qquad (66)\\
&\left.+\Sigma_u\hat{a}_n\int_0^1 H_a(r)\,dr\right). \qquad (67)
\end{aligned}
$$

Combining (65) with (66) the lhs of (60) behaves as

$$
\begin{aligned}
& n^{5/2}\left(\int_0^1 H_a^2(r)\, dB_u(r) + \Sigma_u(\hat{a}_n + a)\int_0^1 H_a^2(r)\, dr \right. \\
& \left. - \int_0^1 H_a(r)\, dr \left(\int_0^1 H_a(r)\, dB_u(r) + \Sigma_u \hat{a}_n \int_0^1 H_a(r)\, dr\right)\right) \qquad (68) \\
& + \Sigma_{u\varepsilon} n^2 \int_0^1 H_a(r)\, dr.
\end{aligned}
$$

The first term on the rhs of (60) is

$$
\begin{aligned}
\sum_{t=2}^{n} u_t e^{2\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}^2 &\sim \sum_{t=2}^{n} u_t Y_{t-1}^2 + \frac{2}{\sqrt{n}} \sum_{t=2}^{n} u_t u_t' \hat{a}_n Y_{t-1}^2 \\
&\sim n^{5/2}\left(\int_0^1 H_a^2(r)\, dB_u(r) + 2\Sigma_u \hat{a}_n \int H_a^2(r)\, dr\right). \qquad (69)
\end{aligned}
$$

The second term on the rhs of (60) is

$$
\left(\frac{1}{n}\sum_{t=2}^{n} e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\right) \sum_{t=2}^{n} u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}.
$$

We have

$$
\begin{aligned}
\frac{1}{n}\sum_{t=2}^{n} e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} &\sim \frac{1}{n}\left(\sum_{t=2}^{n} Y_{t-1} + \frac{1}{\sqrt{n}}\sum_{t=2}^{n} \hat{a}_n' u_t Y_{t-1}\right) \\
&\sim \frac{1}{n}\left(n^2 \int_0^1 H_a(r)\, dr + n\hat{a}_n' \int_0^1 H_a(r)\, dB_u(r)\right) \\
&= n\int_0^1 H_a(r)\, dr + \hat{a}_n' \int_0^1 H_a(r)\, dB_u(r). \qquad (70)
\end{aligned}
$$

Using (61) and (70), the second term on the rhs of (60) is

$$
\left(\frac{1}{n}\sum_{t=2}^{n} e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\right) \sum_{t=2}^{n} u_t e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}
$$

$$
\begin{aligned}
\sim \quad & \left(n \int_0^1 H_a(r)\, dr + \hat{a}_n' \int_0^1 H_a(r)\, dB_u(r)\right) \\
& \times n^{3/2} \left(\int_0^1 H_a(r)\, dB_u(r) + \Sigma_u \hat{a}_n \int_0^1 H_a(r)\, dr\right). \qquad (71)
\end{aligned}
$$

Using (69) and (71), the rhs of (60) behaves as

$$
\begin{aligned}
& n^{5/2} \left(\int_0^1 H_a^2(r)\, dB_u(r) + 2\Sigma_u \hat{a}_n \int H_a^2(r)\, dr \right. \\
& \left. - \int_0^1 H_a(r)\, dr \left(\int_0^1 H_a(r)\, dB_u(r) + \Sigma_u \hat{a}_n \int_0^1 H_a(r)\, dr\right)\right). \qquad (72)
\end{aligned}
$$

Equating (68) to (72), we obtain

$$
\begin{aligned}
& n^{5/2} \left(\int_0^1 H_a^2(r)\, dB_u(r) + \Sigma_u (\hat{a}_n + a) \int_0^1 H_a^2(r)\, dr \right. \\
& \left. - \int_0^1 H_a(r)\, dr \left(\int_0^1 H_a(r)\, dB_u(r) + \Sigma_u \hat{a}_n \int_0^1 H_a(r)\, dr\right)\right) \\
& + \Sigma_{u\varepsilon} n^2 \int_0^1 H_a(r)\, dr \\
& = n^{5/2} \left(\int_0^1 H_a^2(r)\, dB_u(r) + 2\Sigma_u \hat{a}_n \int H_a^2(r)\, dr \right. \\
& \left. - \int_0^1 H_a(r)\, dr \left(\int_0^1 H_a(r)\, dB_u(r) + \Sigma_u \hat{a}_n \int_0^1 H_a(r)\, dr\right)\right).
\end{aligned}
$$

This leads to the result

$$
\sqrt{n}\,(\hat{a}_n - a) \Rightarrow \frac{\int_0^1 H_a(r)\, dr}{\int_0^1 H_a^2(r)\, dr} \Sigma_u^{-1} \Sigma_{u\varepsilon}.
$$

■

Proof Theorem 5. With $\bar{Y}_n = n^{-1} \sum_{t=2}^n Y_t$, the least squares estimator of μ is given by

$$\begin{aligned}\hat{\mu}_n &= \bar{Y}_n - \frac{1}{n}\sum_{t=2} e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} \\ &= \frac{Y_n - Y_1}{n} - \frac{1}{n}\sum_{t=2}\left(\frac{\hat{a}_n' u_t}{\sqrt{n}} + \frac{\left(\hat{a}_n' u_t\right)^2}{2n} + o_p\left(\frac{1}{n}\right)\right) Y_{t-1}.\end{aligned}$$

Now,

$$\frac{Y_n}{n} \Rightarrow \mu e^{a' B_u(1)} \int_0^1 e^{-a' B_u(r)} dr = \mu H_a^*(1)$$

and $Y_1/n = O_p\left(n^{-1}\right)$. Furthermore, using Theorem 4,

$$\frac{1}{n^{3/2}} \hat{a}_n' \sum_{t=2} u_t Y_{t-1} \Rightarrow \mu a' \int_0^1 H_a^*(r)\, dB_u(r)$$

and

$$\frac{1}{n^2} \hat{a}_n' \left(\sum_{t=2} u_t u_t' Y_{t-1}\right) \hat{a}_n \Rightarrow \mu a' \Sigma_u a \int_0^1 H_a^*(r)\, dr.$$

Hence,

$$\hat{\mu}_n \Rightarrow \mu\left(H_a^*(1) - a' \int_0^1 H_a^*(r)\, dB_u(r) - \frac{1}{2} a' \Sigma_u a \int_0^1 H_a^*(r)\, dr\right) = \mu B(a).$$

■

Proof of Theorem 6. By definition, $\mu_n^* = \frac{\hat{\mu}_n}{B(a)}$. For part (1),

$$\begin{aligned}\hat{\sigma}_{\varepsilon,n} &= \frac{1}{n}\sum_{t=1}^n \left(Y_t - \mu_n^* - e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1}\right)^2 = \frac{1}{n}\sum_{t=1}^n \left(Y_t - \mu - e^{\hat{a}_n' u_t/\sqrt{n}} Y_{t-1} + o_p(1)\right)^2 \\ &= \frac{1}{n}\sum_{t=1}^n \left(Y_t - \mu - e^{a' u_t/\sqrt{n}} e^{(\hat{a}_n - a)' u_t/\sqrt{n}} Y_{t-1} + o_p(1)\right)^2 \\ &= \frac{1}{n}\sum_{t=1}^n \Bigg(\varepsilon_t + o_p(1) \\ &\quad - e^{a' u_t/\sqrt{n}} \left(\frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} + \frac{\left((\hat{a}_n - a)' u_t\right)^2}{2n} + o_p\left(\frac{1}{n}\right)\right) Y_{t-1}\Bigg)^2. \qquad (73)\end{aligned}$$

Now,

$$-\frac{2}{n}\sum_{t=1}^{n}\varepsilon_t e^{a'u_t/\sqrt{n}}\left(\frac{(\hat{a}_n-a)'\,u_t}{\sqrt{n}}+\frac{\left((\hat{a}_n-a)'\,u_t\right)^2}{2n}+o_p\left(\frac{1}{n}\right)\right)Y_{t-1}$$
$$=-\frac{2}{n}\sum_{t=1}^{n}\varepsilon_t\left(1+\frac{a'u_t}{\sqrt{n}}+o_p\left(\frac{1}{\sqrt{n}}\right)\right)$$
$$\times\left(\frac{(\hat{a}_n-a)'\,u_t}{\sqrt{n}}+\frac{\left((\hat{a}_n-a)'\,u_t\right)^2}{2n}+o_p\left(\frac{1}{n}\right)\right)Y_{t-1}. \tag{74}$$

By Theorem 4, $\hat{a}_n-a=O_p\left(n^{-1/2}\right)$. Therefore,

$$\begin{aligned}-\frac{2}{n}\sum_{t=1}^{n}\varepsilon_t\frac{(\hat{a}_n-a)'\,u_t}{\sqrt{n}}Y_{t-1} &= -2\left(\hat{a}_n-a\right)'\int_0^1 H_a\left(r\right)dB_{u\varepsilon}\left(r\right)+o_p\left(\frac{1}{\sqrt{n}}\right)\\ &= O_p\left(\frac{1}{\sqrt{n}}\right),\end{aligned}$$

$$-\frac{2}{n}\sum_{t=1}^{n}\varepsilon_t\frac{\left((\hat{a}_n-a)'\,u_t\right)^2}{2n}Y_{t-1}=O_p\left(\frac{1}{n}\right),$$

$$-\frac{2}{n}\sum_{t=1}^{n}\varepsilon_t\frac{a'u_t}{\sqrt{n}}\frac{(\hat{a}_n-a)'\,u_t}{\sqrt{n}}Y_{t-1}=O_p\left(\frac{1}{\sqrt{n}}\right),$$

$$-\frac{2}{n}\sum_{t=1}^{n}\varepsilon_t\frac{a'u_t}{\sqrt{n}}\frac{\left((\hat{a}_n-a)'\,u_t\right)^2}{2n}Y_{t-1}=O_p\left(\frac{1}{n^{3/2}}\right).$$

It follows that (74) converges in probability to zero. The last term in (73) is

$$\frac{1}{n}\sum_{t=1}^{n}e^{2a'u_t/\sqrt{n}}\left(\frac{(\hat{a}_n-a)'\,u_t}{\sqrt{n}}+\frac{\left((\hat{a}_n-a)'\,u_t\right)^2}{2n}+o_p\left(\frac{1}{n}\right)\right)^2Y_{t-1}^2. \tag{75}$$

The leading term in the last expression is

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\left(\frac{(\hat{a}_n-a)'u_t}{\sqrt{n}}\right)^2 Y_{t-1}^2 &= (\hat{a}_n-a)' E(u_t u_t') \left(\sum_{t=1}^{n}\left(\frac{Y_{t-1}}{n}\right)^2\right)(\hat{a}_n-a) \\
&\Rightarrow \left(\frac{\int_0^1 H_a(r)\,dr}{\int_0^1 H_a^2(r)\,dr}\Sigma_{u\varepsilon}'\Sigma_u^{-1}\right)\Sigma_u\left(\int_0^1 H_a^2(r)\,dr\right) \\
&\quad \times \left(\frac{\int_0^1 H_a(r)\,dr}{\int_0^1 H_a^2(r)\,dr}\Sigma_u^{-1}\Sigma_{u\varepsilon}\right) \\
&= \frac{\left(\int_0^1 H_a(r)\,dr\right)^2}{\int_0^1 H_a^2(r)\,dr}\Sigma_{u\varepsilon}'\Sigma_u^{-1}\Sigma_{u\varepsilon}.
\end{aligned}$$

All other terms in (75) converge in probability to zero. Hence,

$$\hat{\sigma}_{\varepsilon,n}^2-\sigma_\varepsilon^2 \Rightarrow \frac{\left(\int_0^1 H_a(r)\,dr\right)^2}{\int_0^1 H_a^2(r)\,dr}\Sigma_{u\varepsilon}'\Sigma_u^{-1}\Sigma_{u\varepsilon},$$

as required.

The proof of part (2) is identical to the $\mu=0$ case. Finally, for part (3),

$$\begin{aligned}
\hat{\Sigma}_{u\varepsilon,n} &= \frac{1}{n}\sum_{t=1}^{n} u_t\left(Y_t-\mu_n^*-e^{\hat{a}_n' u_t/\sqrt{n}}Y_{t-1}\right) \\
&= \frac{1}{n}\sum_{t=1}^{n} u_t\left(Y_t-\mu+o_p(1)-e^{a'u_t/\sqrt{n}}e^{(\hat{a}_n-a)'u_t/\sqrt{n}}Y_{t-1}\right) \\
&= \frac{1}{n}\sum_{t=1}^{n} u_t\left(\varepsilon_t+o_p(1)-\left(1+\frac{a'u_t}{\sqrt{n}}+o_p\left(\frac{1}{\sqrt{n}}\right)\right)\right. \\
&\quad \times\left.\left(\frac{(\hat{a}_n-a)'u_t}{\sqrt{n}}+\frac{\left((\hat{a}_n-a)'u_t\right)^2}{2n}+o_p\left(\frac{1}{n}\right)\right)Y_{t-1}\right). \qquad (76)
\end{aligned}$$

Now, $n^{-1}\sum_{t=1}^{n} u_t\varepsilon_t \to_p \Sigma_{u\varepsilon}$ and

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} u_t \frac{(\hat{a}_n - a)' u_t}{\sqrt{n}} Y_{t-1} &= \frac{1}{n}\sum_{t=1}^{n} u_t u_t' \frac{Y_{t-1}}{n}\left(\sqrt{n}\left(\hat{a}_n - a\right)\right) \\
&\Rightarrow \left(\Sigma_u \int_0^1 H_a(r)\, dr\right)\left(\frac{\int_0^1 H_a(r)\, dr}{\int_0^1 H_a^2(r)\, dr}\Sigma_u^{-1}\Sigma_{u\varepsilon}\right) \\
&= \frac{\left(\int_0^1 H_a(r)\, dr\right)^2}{\int_0^1 H_a^2(r)\, dr}\Sigma_{u\varepsilon}.
\end{aligned}$$

All other terms in (76) converge in probability to zero. Hence,

$$\hat{\Sigma}_{u\varepsilon,n} - \Sigma_{u\varepsilon} \Rightarrow -\frac{\left(\int_0^1 H_a(r)\, dr\right)^2}{\int_0^1 H_a^2(r)\, dr}\Sigma_{u\varepsilon}$$

and the proof of the theorem is completed. ■

Proof of (20)-(21). We can write (14) as

$$\begin{aligned}
Y_n(r) &= \mu_T e^{a'B_u(r)} \int_0^r e^{-a'B_u(p)} dp + \sqrt{n} G_a(r) \\
&= \mu_T e^{a'B_u(r)} \int_0^r e^{-a'B_u(p)} dp \\
&+ \sqrt{n} e^{a'B_u(r)} \left(\left[\Sigma^{1/2}\right]_2 \int_0^r e^{-a'B_u(p)} dB^*(p)\right. \\
&\left. - \frac{1}{n} a' \left(n\Sigma_{u\varepsilon}\right) \int_0^r e^{-a'B_u(p)} dp\right) \\
&= \mu_T e^{a'B_u(r)} \int_0^r e^{-a'B_u(p)} dp \\
&+ e^{a'B_u(r)} \left(\left[\Sigma_T^{1/2}\right]_2 \int_0^r e^{-a'B_u(p)} dB^*(p) - \frac{1}{\sqrt{n}} a' \Sigma_{u\varepsilon,T} \int_0^r e^{-a'B_u(p)} dp\right).
\end{aligned} \tag{77}$$

Further, $a'B_u(r) = a'\Sigma_u^{1/2} B_u^*(r) = (T/n)^{1/2} a' \Sigma_{u,A}^{1/2} B_u^*(r)$. Hence, (77) becomes

$$
\begin{aligned}
Y_n(r) &= \mu_T e^{a_n' \Sigma_{u,A}^{1/2} B_u^*(r)} \int_0^r e^{-a_n' \Sigma_{u,A}^{1/2} B_u^*(p)} dp \\
&\quad + e^{a_n' \Sigma_{u,A}^{1/2} B_u^*(r)} \left(\left[\Sigma_T^{1/2} \right]_2 \int_0^r e^{-a_n' \Sigma_{u,A}^{1/2} B_u^*(p)} dB^*(p) \right. \\
&\quad \left. - \frac{1}{\sqrt{n}} a' \Sigma_{u\varepsilon,T} \int_0^r e^{-a_n' \Sigma_{u,A}^{1/2} B_u^*(r)} dp \right),
\end{aligned} \tag{78}
$$

giving the required result. ■

Proof of 24. By stochastic differentiation we have

$$
\begin{aligned}
dG(r) &= \frac{dS(r)}{S(r)} - \frac{1}{2S^2(r)} d[S]_r \\
&= \frac{1}{S(r)} \left\{ T\mu_A + \sqrt{T} \left(\frac{a_n' \Sigma_{u,A} a_n}{2} G_{a_n,A}(r) - a_n' \Sigma_{u\varepsilon,A} \right) \right\} S(r)\, dr \\
&\quad + \frac{1}{S(r)} S(r) \sqrt{T} b_A(r)' \left[\Sigma_A^{1/2} \right] dB^*(r) \\
&\quad - \frac{1}{2S^2(r)} \left\{ TS^2(r)\, b_A(r)' \left[\Sigma_A^{1/2} \right] \left[\Sigma_A^{1/2} \right]' b_A(r) \right\} dr
\end{aligned}
$$

$$
\begin{aligned}
&= \left\{ T\mu_A + \sqrt{T} \frac{a_n' \Sigma_{u,A} a_n}{2} G_{a_n,A}(r) - \sqrt{T} a_n' \Sigma_{u\varepsilon,A} \right. \\
&\quad \left. - \frac{T}{2} b_A(r)' \Sigma_A b_A(r)\, dr \right\} + \sqrt{T} b_A(r)' \Sigma_A^{1/2} dB^*(r).
\end{aligned}
$$

Now,

$$
b_A(r)' \Sigma_A b_A(r) = G_{a_n,A}^2(r)\, a_n' \Sigma_{u,A} a_n + 2G_{a_n,A}(r)\, a_n' \Sigma_{u\varepsilon,A} + \sigma_{\varepsilon,A}^2 \tag{79}
$$

and thus, because

$$
\frac{a'}{\sqrt{n}} \Sigma_{u\varepsilon,T} = \frac{a'}{\sqrt{n}} T \Sigma_{u\varepsilon,A} = \sqrt{T} a_n' \Sigma_{u\varepsilon,A}
$$

we get

$$
\begin{aligned}
dG(r) &= \left(T\mu_A + \sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r) - a_n'\Sigma_{u\varepsilon,A}\right)\right. \\
&\quad \left.-\frac{T}{2}\left(G_{a_n,A}^2(r)\,a_n'\Sigma_{u,A}a_n + 2G_{a_n,A}(r)\,a_n'\Sigma_{u\varepsilon,A} + \sigma_{\varepsilon,A}^2\right)\right)dr \\
&\quad +\sqrt{T}b_A(r)'\,\Sigma_A^{1/2}dB^*(r) \\
&= \left(T\left(\mu_A - \frac{\sigma_{\varepsilon,A}^2}{2}\right) - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)dr \\
&\quad +\left(\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)G_{a_n,A}(r)\right. \\
&\quad \left.-T\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}^2(r)\,dr\right) + \sqrt{T}b_A(r)'\,\Sigma_A^{1/2}dB^*(r)\,.
\end{aligned}
$$

■

Proof of 27. Using (24), we have

$$
\begin{aligned}
&G(r) - G(0) = \log(S(r)) - \log(S(0)) \\
=\;& \int_0^r \left(T\left(\mu_A - \frac{\sigma_{\varepsilon,A}^2}{2}\right) - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)dr \\
&+\int_0^r \left(\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)G_{a_n,A}(s)\right. \\
&\left.-T\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}^2(s)\right)ds + \sqrt{T}\int_0^r b_A(s)'\,\Sigma_A^{1/2}dB(s) \\
=\;& \left(T\left(\mu_A - \frac{\sigma_{\varepsilon,A}^2}{2}\right) - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)r \\
&+\int_0^r \left\{\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)G_{a_n,A}(s)\right. \\
&\left.-T\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}^2(s)\right\}ds \\
&+\sqrt{T}a_n'\int_0^r G_{a_n,A}(s)\,\Sigma_{u,A}^{1/2}(s)\,dB_u^*(s) + \sqrt{T}\Sigma_{2,A}^{1/2}B^*(r)\,,
\end{aligned}
$$

and (27) immediately follows. ■

Proof of details of Section 4.2. We must have $df(r,x) = dV(r)$ and by direct calculation

$$\begin{aligned} dV(r) &= \alpha_S(r)\left\{\left(T\mu_A + \sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r) - a_n'\Sigma_{u\varepsilon,A}\right)\right)S(r)\,dr\right. \\ &\left. +\sqrt{T}S(r)\,b_A(r)'\,\Sigma_A^{1/2}dB^*(r)\right\} + Tr_{f,A}\alpha_Z(r)\,\gamma(r)\,dr \\ &= \left\{\alpha_S(r)\left(T\mu_A + \sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r) - a_n'\Sigma_{u\varepsilon,A}\right)\right)S(r)\right. \\ &\left. +Tr_{f,A}\alpha_Z(r)\,\gamma(r)\right\}dr + \sqrt{T}\alpha_S(r)\,S(r)\,b_A(r)'\,\Sigma_A^{1/2}dB^*(r)\,. \end{aligned} \quad (80)$$

Now, in view of (23), we have

$$(d(S(r)))^2 = TS^2(r)\,b_A(r)'\,\Sigma_A b_A(r)\,dr, \quad (81)$$

and since $df(r,x) = f_r dr + f_x dS(r) + \frac{1}{2}f_{xx}(d(S(r)))^2$, we deduce that

$$\begin{aligned} df(r,S(r)) &= f_r dr + f_x\{T\mu_A \\ &+\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r) - a_n'\Sigma_{u\varepsilon,A}\right)\right\}S(r)\,dr \\ &+ f_x\sqrt{T}S(r)\,b_A(r)'\,\Sigma_A^{1/2}dB^*(r) + \frac{T}{2}f_{xx}S^2(r)\,b_A(r)'\,\Sigma_A b_A(r)\,dr \\ &= \left\{f_r + f_x\left\{T\mu_A + \sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r) - a_n'\Sigma_{u\varepsilon,A}\right)\right\}S(r)\right. \\ &\left. +\frac{T}{2}f_{xx}S^2(r)\,b_A(r)'\,\Sigma_A b_A(r)\right\}dr + \sqrt{T}f_x S(r)\,b_A(r)'\,\Sigma_A^{1/2}dB^*(r)\,. \end{aligned} \quad (82)$$

Equating the coefficients of dr and of the stochastic component in (80) and

(82) gives

$$\begin{aligned}
&\alpha_S(r)\left\{T\mu_A+\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r)-a_n'\Sigma_{u\varepsilon,A}\right)\right\}S(r)\\
&+Tr_{f,A}\alpha_Z(r)\gamma(r)\\
=\;& f_r+f_x\left\{T\mu_A+\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r)-a_n'\Sigma_{u\varepsilon,A}\right)\right\}S(r)\\
&+\frac{T}{2}f_{xx}S^2(r)\,b_A(r)'\,\Sigma_A b_A(r)
\end{aligned}\tag{83}$$

and

$$\sqrt{T}\alpha_S(r)\,S(r)\,b_A(r)'\,\Sigma_A^{1/2}dB^*(r)=\sqrt{T}f_xS(r)\,b_A(r)'\,\Sigma_A^{1/2}dB^*(r)\,.\tag{84}$$

The latter yields

$$\alpha_S(r)=f_x\tag{85}$$

which is the condition in the classic case. Using this condition in (83) we have

$$\begin{aligned}
&\alpha_S(r)\left\{T\mu_A+\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r)-a_n'\Sigma_{u\varepsilon,A}\right)\right\}S(r)\\
&+Tr_{f,A}\alpha_Z(r)\gamma(r)\\
=\;& f_r+\alpha_S(r)\left\{T\mu_A+\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r)-a_n'\Sigma_{u\varepsilon,A}\right)\right\}S(r)\\
&+\frac{T}{2}f_{xx}S^2(r)\,b_A(r)'\,\Sigma_A b_A(r)\,.
\end{aligned}$$

which implies

$$Tr_{f,A}\alpha_Z(r)\gamma(r)=f_r+\frac{T}{2}f_{xx}S^2(r)\,b_A(r)'\,\Sigma_A b_A(r)$$

and (30) follows.■

Table 1. Google Call Option Prices

n	36	36	36	36	36
K	950	1030	1060	1100	1200
BS	116.71	54.13	37.19	20.77	3.18
$\hat{C}^*\,(\rho = 0)$	115.81	52.51	35.86	20.39	4.41
$\hat{C}^*\,(\hat{\rho})$	114.58	51.75	36.04	21.37	5.13
$\hat{C}^{*MV}\,(\hat{\rho})$	114.78	51.67	35.66	20.85	4.87
$\hat{C}^*\,(\rho = .95)$	113.90	50.23	34.90	21.17	5.98
$\hat{C}\,(\rho = 0)$	117.61	56.55	40.23	24.40	6.37
$\hat{C}\,(\hat{\rho})$	115.87	55.81	40.39	25.43	7.24
$\hat{C}^{MV}\,(\hat{\rho})$	116.18	55.63	39.89	24.77	6.89
$\hat{C}\,(\rho = .95)$	115.21	55.34	40.41	26.29	8.93
n	415	415	415	415	415
K	950	1030	1060	1100	1200
BS	179.88	136.40	122.36	105.47	71.49
$\hat{C}^*\,(\rho = 0)$	180.52	137.31	123.30	106.60	72.49
$\hat{C}^*\,(\hat{\rho})$	177.49	134.75	121.00	104.32	71.22
$\hat{C}^{*MV}\,(\hat{\rho})$	173.62	130.73	117.26	101.13	69.48
$\hat{C}^*\,(\rho = .95)$	176.74	133.99	120.36	104.15	72.17
$\hat{C}\,(\rho = 0)$	182.58	139.51	125.51	108.80	74.53
$\hat{C}\,(\hat{\rho})$	180.46	137.95	124.21	107.52	74.16
$\hat{C}^{MV}\,(\hat{\rho})$	180.89	138.63	125.21	109.02	76.71
$\hat{C}\,(\rho = .95)$	177.76	135.09	121.47	105.25	73.19

Table 1. Google Call Option Prices (Continued)

n	778	778	778	778	778
K	950	1030	1060	1100	1200
BS	219.53	179.85	166.63	150.32	115.58
$\hat{C}^*(\rho = 0)$	219.30	179.44	166.20	149.77	114.94
$\hat{C}^*(\hat{\rho})$	211.52	172.02	159.01	143.12	109.87
$\hat{C}^{*MV}(\hat{\rho})$	216.59	177.57	164.53	148.29	113.82
$\hat{C}^*(\rho = .95)$	213.42	174.65	162.10	146.64	113.70
$\hat{C}(\rho = 0)$	217.67	177.75	164.51	148.08	113.33
$\hat{C}(\hat{\rho})$	217.03	177.77	164.78	148.87	115.35
$\hat{C}^{MV}(\hat{\rho})$	222.94	184.15	171.13	154.88	120.10
$\hat{C}(\rho = .95)$	221.39	182.99	170.47	154.99	121.72

Note: n is the number of days to expiration as of 11-27-2013; K is the strike price; $S(0) = 1063.11$; 'BS' is the price based on Black and Scholes's classic formula; $\hat{C}$ and $\hat{C}^*$ are based on (38)-(39); $\hat{\rho}$ is based on Nasdaq's and Google's historical volatility; The superscript MV indicates pricing under (42).

Table 2. Summary Statistics for BS- and STUR Based Simulated Data

T	Statistic	SD	Skewness	Kurtosis
36	$\lambda^{Q}_{BS}(1)$	0.084	0.253	3.114
36	$\lambda^{*Q}(1)$	0.084	0.970	4.543
36	$\bar{\lambda}^{Q}(1)$	0.095	0.970	4.543
415	$\lambda^{Q}_{BS}(1)$	0.291	0.898	4.469
415	$\lambda^{*Q}(1)$	0.291	1.124	5.332
415	$\bar{\lambda}^{Q}(1)$	0.299	1.124	5.332
778	$\lambda^{Q}_{BS}(1)$	0.406	1.285	6.073
778	$\lambda^{*Q}(1)$	0.406	1.988	12.103
778	$\bar{\lambda}^{Q}(1)$	0.421	1.988	12.103

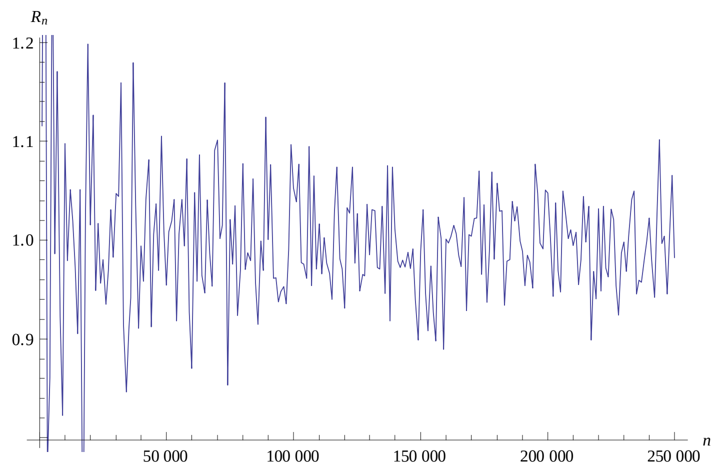

Figure 1: Plot of R_n, given in (41), against n: the $\mu \neq 0$ and $a = 0$ case.

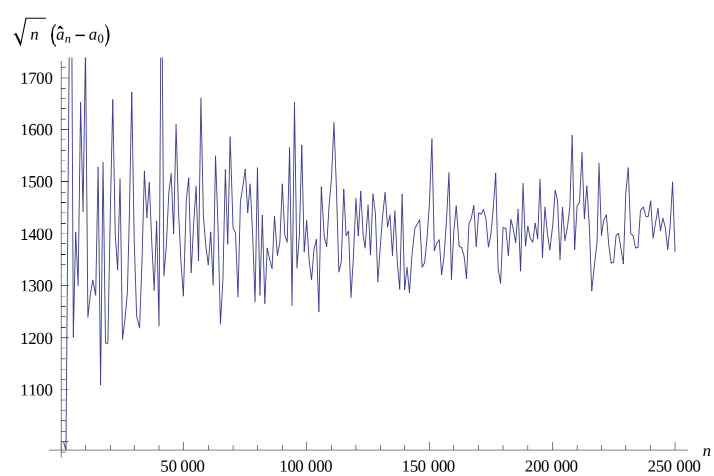

Figure 2: Plot of $\sqrt{n}\,(\hat{a}_n - a_0)$ against n when $a_0 = 0.15$: the $\mu \neq 0$ case.

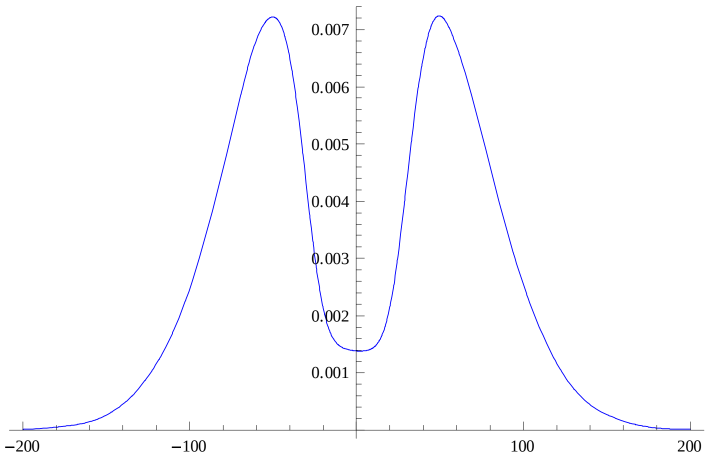

Figure 3: Kernel density estimate of the asymptotic distribution of $\hat{a}_n$ in the $\mu = 0$, $\sigma_{u\varepsilon} \neq 0$ and $a = 0$ case, with $n = 1231$, 100000 replications and historical estimates of the Google-Nasdaq data.

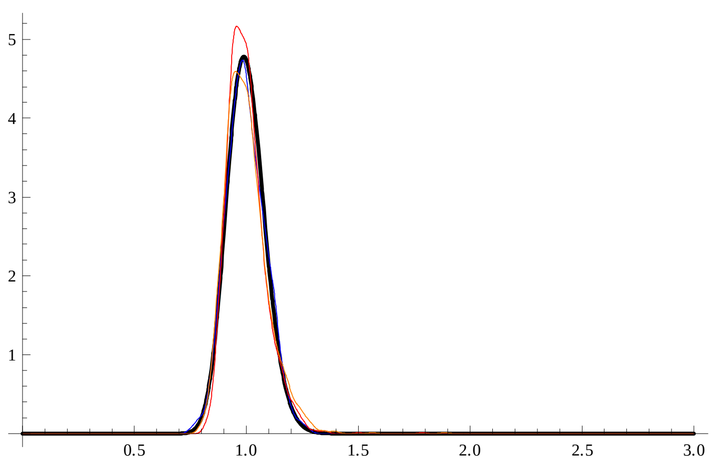

Figure 4: Kernel density estimates of the exponent terms: Theoretical $\lambda_{BS}^{Q}(1)$ (Black), Simulated $\lambda_{BS}^{Q}(1)$ (BLUE), $\lambda^{*Q}(1)$ (Red), $\bar{\lambda}^{Q}(1)$ (orange), based on the multivariate model (42) with $T = 36$ and $\rho = \hat{\rho}$.

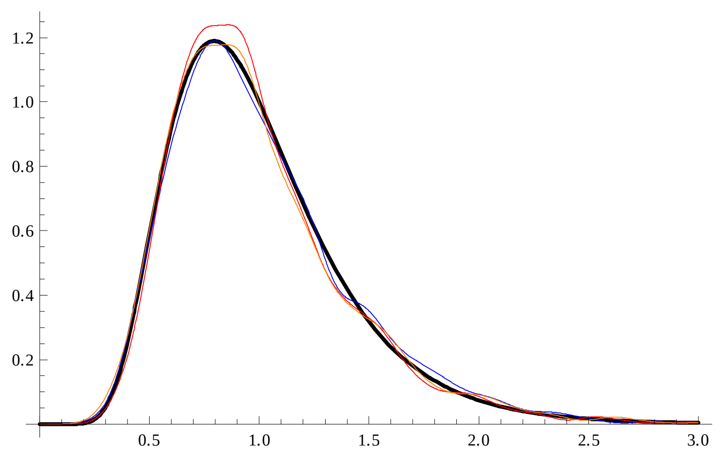

Figure 5: Kernel density estimates of the exponent terms: Theoretical $\lambda_{BS}^{Q}(1)$ (Black), Simulated $\lambda_{BS}^{Q}(1)$ (BLUE), $\lambda^{*Q}(1)$ (Red), $\bar{\lambda}^{Q}(1)$ (orange), based on the model (40)with $T = 778$ and $\rho = 0.95$.