Research Collection School Of Economics   School of Economics

A multivariate stochastic unit root model with an application to derivative pricing

Offer LIEBERMAN
Bar Ilan University

Peter C. B. PHILLIPS
Singapore Management University, peterphillips@smu.edu.sg

Citation

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

Offer Lieberman[†]and Peter C. B. Phillips[‡]

Revised, May 17, 2016

Abstract

This paper extends recent findings of Lieberman and Phillips (2014) on stochastic unit root (STUR) models to a multivariate case including asymptotic theory for estimation of the model's parameters. The extensions are useful for applications of STUR modeling and 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 substantially reduces the average percentage pricing error of the Black-Scholes and Heston's (1993) stochastic volatility (with zero volatility risk premium) pricing schemes in most moneyness-maturity categories considered.

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

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

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

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

where $W\left(r\right)$ 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

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

1

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 financial econometric applications, u_t may 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 (parametrically) local to zero and the factors (observable and 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 the vector $\eta_t = (u_t, \varepsilon_t)'$ is a strictly stationary martingale difference sequence (mds) whose partial sums satisfy the invariance principle[2]

$$n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} \eta_t \Rightarrow B(r) \equiv \mathrm{BM}(\Sigma), \ \ \Sigma = \left(\begin{array}{cc} \Sigma_u & \Sigma_{u\varepsilon} \\ \Sigma_{u\varepsilon}' & \sigma_\varepsilon^2 \end{array} \right), \quad (8)$$

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 $\sigma_\varepsilon^2 > 0$. 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.

---

[1] In other words, in applications u_t would typically be a demeaned $I(0)$ process.

[2] The invariance principle (8) holds if η_t is a strictly stationary mds with respect to the natural filtration and has finite second moment matrix Σ (Billingsley, 1968). The mds condition on η_t may be considerably relaxed, as shown in Phillips and Solo (1992), at the expense of some additional conditions concerning weak dependence and higher order moment existence that are detailed in that work.

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 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 poor fit with market data. 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 the first-order asymptotic theory of estimation of the model parameters, in both the $\mu = 0$ and the $\mu \neq 0$ cases, as well as a test statistic for the hypothesis $H_0 : a = 0$, which is based on a second order

4

asymptotic expansion for the null distribution of the least squares estimator of a in the $\mu \neq 0$ case. Section 4 derives the sde corresponding to the STUR model and the price process associated with it. The value of a European call option under the new model is given in Section 5. Some simulations are supplied in Section 6, corroborating the asymptotic results of Section 3. An empirical application follows in Section 7, showing that our model substantially reduces the average percentage error of the Black-Scholes and Heston's (1993) SV[3] pricing schemes in most moneyness-maturity categories considered. Section 8 concludes. An Online Supplement (Lieberman and Phillips, 2016) contains all proofs, further technical details, information on the data for the empirical application, and figures relevant to Sections 6-7.

2   Continuous Limit of the STUR Model

By back 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 $t \geq 2$,

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

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 $n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} \eta_t \to_p B\left(r\right)$. This procedure, which is standard in modern asymptotic theory, enables limit variates such as the vector Brownian motion $B\left(r\right)$ (which typically escape from the underlying probability space through the action of the asymptotics) to be included in the same space as the random

---

[3]Heston's (1993) SV model was estimated with the approximate MLE procedure suggested by Aït-Sahalia and Kimmel (2007), which sets the volatility risk premium parameter to zero.

sequence. The space augmentation also allows for the convenient replacement of weak convergence (denoted $\Rightarrow$) by almost sure convergence or, as here, convergence in probability, which is another well-known device in modern asymptotic theory. The results then correspond to weak convergence in the original space. 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 $n \to \infty$*

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

and

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

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

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

we find that $G_a(r)$ follows the sde

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

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

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

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

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

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

In this system the nonlinear diffusion process G_a affects both the martingale component and the drift. When $a = 0$, the system reduces to $dY_n(r) = n\mu dr + \sqrt{n}dB_\varepsilon(r)$, which corresponds in form to the classic equation (5). The process $Y_n(r)$ may be regarded as parametrically local (in terms of a) to Brownian motion with drift.

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 nonlinear least squares estimator of a which minimizes the criterion $Q_n(a) = \sum_{t=2}^n \{Y_t - \beta_t(a;n)Y_{t-1}\}^2$.

Theorem 2 *For the model (6)–(8) with $\mu = 0$, the asymptotic behavior of $\hat{a}_n$ is given by:*
   (1)

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

   (2)

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

7

(3)

$$\sqrt{n}\left(\hat{a}_n - a\right) \Rightarrow \frac{1}{\int_0^1 G_a^2\left(r\right)dr}\Sigma_u^{-1}\left\{\int_0^1 G_a\left(r\right)dB_{u\varepsilon}\left(r\right)\right.\tag{16}$$

$$\left.+\left\{E\left(\varepsilon_t u_t u_t'\right)\right\}a\int_0^1 G_a\left(r\right)dr\right\},$$

$$\text{if } \Sigma_{u\varepsilon} = 0.$$

(4)

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

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 $H_0 : a = 0$, in the $\Sigma_{u\varepsilon} \neq 0$ and $L = 1$ subcase, we have*

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

where ρ is the correlation coefficient between ε and u.

We estimate σ_ε^2, Σ_u^2 and $\Sigma_{u\varepsilon}$ by

$$\hat{\sigma}_{\varepsilon,n}^2 = \frac{1}{n}\sum_{t=2}^n\left(Y_t - e^{\hat{a}_n'u_t/\sqrt{n}}Y_{t-1}\right)^2,\ vech\left(\hat{\Sigma}_{u,n}\right) = \frac{1}{n}\sum_{t=1}^n vech\left(u_t u_t'\right),\tag{17}$$

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

and give the limit theory of these estimates in the following result.

8

Theorem 3 *For the model (6)–(8) with $\mu = 0$, the asymptotic behavior of $\hat{\sigma}_{\varepsilon,n}^2$, $\text{vech}\left(\hat{\Sigma}_{u,n}\right)$ 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) If u_t has finite fourth moments, centred partial sums of $u_t u'_t$ satisfy the invariance principle

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor nr \rfloor} \text{vech}\left(u_t u'_t - \Sigma_u\right) \Rightarrow \xi(r),$$

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

$$\Sigma_{u\otimes u} = E\left(L\left(u_t \otimes u_t - E\left(u_t \otimes u_t\right)\right)\left(u'_t \otimes u'_t - E\left(u'_t \otimes u'_t\right)\right)L'\right)$$

and L is the elimination matrix defined by

$$\text{vech}\left(u_t u'_t - \Sigma_u\right) = L\left(u_t \otimes u_t - E\left(u_t \otimes u_t\right)\right). \qquad (18)$$

(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, Theorem 3(1) implies*

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

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

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

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 $\hat{\Sigma}_{u\varepsilon,n}^0 = n^{-1}\sum_{t=2}^n (Y_t - Y_{t-1})\,u_t$ 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:

$$\sqrt{n}\,(\hat{a}_n - a) \Rightarrow \frac{\int_0^1 H_a(r)\,dr}{\int_0^1 H_a^2(r)\,dr}\Sigma_u^{-1}\Sigma_{u\varepsilon}. \tag{19}$$

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 $Y_t = \mu + e^{\frac{au_t}{\sqrt{n}}}Y_{t-1} + \varepsilon_t$ $(t = 2, ..., n)$ and may be written in linear pseudo-model[4] form as $\Delta Y_t = \mu + aZ_t + \varepsilon_t$, where

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

When $\mu = 0$, $Y_t = O_p(\sqrt{n})$. 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 seen from (12). In that case, the pseudo-regressor $Z_t = O_p(\sqrt{n})$ 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 $\hat{a}_n$.

---

[4]The linear pseudo-model is the linear model obtained by approximating a nonlinear regression model linearly in the immediate neighborhood of the true value of the parameters (see Malinvaud, 1980; Phillips, 2015).

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 $a = 0$, $H_a\left(r\right) = \mu_0 r$ and $\sqrt{n}\hat{a}_n \rightarrow_p \frac{3}{2\mu_0}\Sigma_u^{-1}\Sigma_{u\varepsilon}$. In the $a = 0$ and $L = 1$ subcase, we have*

$$\sqrt{n}\hat{a}_n \rightarrow_p \frac{3\rho\sigma_\varepsilon}{2\mu_0\sigma_u}. \tag{20}$$

In view of (20), the asymptotic distribution of $\hat{a}_n$ is degenerate under the null hypothesis $H_0 : a = 0$. In order to conduct the test, we derive a second order expansion for the distribution of $\hat{a}_n$.

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

$$\sqrt{n}\left(\sqrt{n}\hat{a}_n - \frac{3}{2\mu_0}\Sigma_u^{-1}\Sigma_{u\varepsilon}\right) \tag{21}$$

$$\Rightarrow \frac{3\Sigma_u^{-1}}{\mu_0^2}\left(\mu_0^2 \int_0^1 r dB_u\left(r\right) + \mu_0 \int_0^1 r dB_{u\varepsilon}\left(r\right) - \frac{3\mu_0}{2}\left(\int_0^1 r^2 dB_{uu'}\left(r\right)\right)\Sigma_u^{-1}\Sigma_{u\varepsilon}\right.$$

$$\left. -3\Sigma_{u\varepsilon}\int_0^1 r B_\varepsilon\left(r\right) dr + \Sigma_{u\varepsilon}\int_0^1 B_\varepsilon\left(r\right) dr\right).$$

The second order expansion given in (21) depends on nuisance parameters. Define

$$\hat{\mu}_n^0 = \frac{1}{n}\sum_{t=2}^n \left(Y_t - Y_{t-1}\right), \quad \hat{\Sigma}_{u\varepsilon,n}^{0,\mu} = \frac{1}{n}\sum_{t=2}^n \left(Y_t - \hat{\mu}_n^0 - Y_{t-1}\right)u_t, \tag{22}$$

which are consistent estimators of μ and $\Sigma_{u\varepsilon}$, respectively, under the null $H_0 : a = 0$. The estimator $\hat{\Sigma}_{u,n}$ defined in (17) is invariant to μ and is consistent by Theorem 3(2). Corollary 6 gives a test statistic which is based on (22) and Theorem 5.

Corollary 6 *For the model (6)–(8) with $\mu \neq 0$ and $a = 0$,*

$$T_n \equiv \sqrt{n}\left(\sqrt{n}\hat{a}_n - \frac{3}{2\hat{\mu}_n^0}\left(\hat{\Sigma}_{u,n}\right)^{-1}\hat{\Sigma}_{u\varepsilon,n}^0\right)$$

$$= \frac{3}{\left(\hat{\mu}_n^0\right)^2}\left(\hat{\Sigma}_{u,n}\right)^{-1}\left(\hat{\mu}_n^0\frac{\sum_{t=2}^n\left(u_tY_{t-1}\right)}{n^{3/2}} + \frac{\sum_{t=2}^n\left(u_t\varepsilon_t - \hat{\Sigma}_{u\varepsilon,n}^0\right)Y_{t-1}}{n^{3/2}}\right.$$

$$-\frac{\sum_{t=2}^n Y_{t-1}^2\left(u_tu_t' - \hat{\Sigma}_{u,n}\right)}{n^2}\hat{a}_n - \frac{2\hat{\mu}_n^0\left(\frac{(n-1)n}{2}\sum_{t=1}^{n-1}\varepsilon_t - \sum_{t=1}^{n-1}\frac{(t-1)t}{2}\varepsilon_t\right)}{n^2}\hat{\Sigma}_{u,n}\hat{a}_n$$

$$\left.+\frac{2n\bar{\bar{\varepsilon}}_n\hat{\mu}_n^0}{3}\hat{\Sigma}_{u,n}\hat{a}_n + \frac{\sum_{t=1}^{n-1}(n-t)\varepsilon_t}{n^{3/2}}\hat{\Sigma}_{u\varepsilon,n}^0 - \sqrt{n}\frac{\bar{\bar{\varepsilon}}_n}{2}\hat{\Sigma}_{u\varepsilon,n}^0 + O_p\left(\frac{1}{\sqrt{n}}\right)\right).$$

$$(23)$$

Section 6 provides simulations investigating the accuracy of (23) and an empirical application of the approach. To obtain p-values from (23) we need to simulate a random walk with a drift process and plug in values for the nuisance parameters using (17) and (22).

We continue to present results for the unrestricted estimators of μ, Σ_u, $\Sigma_{u\varepsilon}$ and σ_ε^2, in the $\mu \neq 0$ case. Let $\hat{\mu}_n$ be the least squares estimator of μ and define

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

where

$$H_a^*(r) = \mu^{-1}H_a(r) = e^{a'B_u(r)}\int_0^r e^{-a'B_u(p)}dp.$$

Theorem 7 *For the model (6)–(8) with $\mu \neq 0$, $\hat{\mu}_n \Rightarrow \mu B\left(a, \Sigma_u\right)$.*

Remark 9 *When $a = 0$, $H_a^*(1) = 1$, leading to $\hat{\mu}_n \Rightarrow \mu$.*

To obtain a consistent estimator of μ, we use $\hat{\Sigma}_{u,n}$ and $\hat{a}_n$. Rewriting $H_a^*(r)$ as $H_a^*(r) = e^{a'\Sigma_u^{1/2}B_u^*(r)}\int_0^r e^{-a'\Sigma_u^{1/2}B_u^*(p)}dp$, where $B_u^*(r)$ is a standard Brownian motion (SBM), we let

$$\hat{H}_{\hat{a}_n,n}^*(r) = e^{\hat{a}_n'\hat{\Sigma}_{u,n}^{1/2}B_u^*(r)}\int_0^r e^{-\hat{a}_n'\hat{\Sigma}_{u,n}^{1/2}B_u^*(p)}dp \qquad (24)$$

12

and have the following estimator for $B(a, \Sigma_u)$

$$B\left(\hat{a}_n, \hat{\Sigma}_{u,n}\right) = \hat{H}^*_{\hat{a}_n,n}(1) - \hat{a}'_n \hat{\Sigma}^{1/2}_{u,n} \int_0^1 \hat{H}^*_{\hat{a}_n,n}(r) dB^*_u(r) - \frac{1}{2} \hat{a}'_n \hat{\Sigma}_{u,n} \hat{a}_n \int_0^1 \hat{H}^*_{\hat{a}_n,n} dr.$$

Remark 10 *By Theorems 4 and 8(2), since $\left(\hat{a}_n, \hat{\Sigma}_u\right) \to_p (a, \Sigma_u)$, $B\left(\hat{a}_n, \hat{\Sigma}_{u,n}\right)$ is consistent for $B(a, \Sigma_u)$, and in the $\mu \neq 0$ case the rescaled estimate $\mu^*_n \equiv \frac{\hat{\mu}_n}{B(\hat{a}_n, \hat{\Sigma}_u)}$ is consistent for μ.*

In practical terms, to compute μ^*_n we need to simulate the path of the SBM $B^*_u(r)$, calculate with it $\hat{H}^*_{\hat{a}_n,n}(r)$ and $B\left(\hat{a}_n, \hat{\Sigma}_{u,n}\right)$ and form the sample average of the latter over sufficient number of iterations.

Continuing, we define the following estimates of σ^2_ε and $\Sigma_{u\varepsilon}$ in the $\mu \neq 0$ case:

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

and

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

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

(1) $\left(\hat{\sigma}^\mu_{\varepsilon,n}\right)^2 - \sigma^2_\varepsilon \Rightarrow \dfrac{\left(\int_0^1 H_a(r) dr\right)^2}{\int_0^1 H^2_a(r) dr} \Sigma'_{u\varepsilon} \Sigma^{-1}_u \Sigma_{u\varepsilon},$

(2) $\hat{\Sigma}^\mu_{u\varepsilon,n} - \Sigma_{u\varepsilon} \Rightarrow -\dfrac{\left(\int_0^1 H_a(r) dr\right)^2}{\int_0^1 H^2_a(r) dr} \Sigma_{u\varepsilon}.$

As mentioned above, the estimator defined by (17) is invariant to μ and therefore it is not included in Theorem 8. To obtain consistent estimators for $\Sigma_{u\varepsilon}$ and σ^2_ε, note that Theorem 8 implies

$$\Sigma^\mu_{u\varepsilon,n} \Rightarrow \left(1 - \frac{\left(\int_0^1 H_a(r) dr\right)^2}{\int_0^1 H^2_a(r) dr}\right) \Sigma_{u\varepsilon} = \left(1 - \frac{\left(\int_0^1 H^*_a(r) dr\right)^2}{\int_0^1 H^{*2}_a(r) dr}\right) \Sigma_{u\varepsilon},$$

13

so that $\Sigma_{u\varepsilon,n}^{*\mu} = \left(1 - \left(\int_0^1 \hat{H}_{\hat{a}_n,n}^*(r)\, dr\right)^2 / \int_0^1 \hat{H}_{\hat{a}_n,n}^{*2}(r)\, dr\right)^{-1} \Sigma_{u\varepsilon,n}^{\mu} \to_p \Sigma_{u\varepsilon}$. Finally, it follows from Theorem 8(1) that

$$\left(\sigma_{\varepsilon,n}^{*\mu}\right)^2 = \left(\hat{\sigma}_{\varepsilon,n}^{\mu}\right)^2 - \frac{\left(\int_0^1 \hat{H}_{\hat{a}_n,n}^*(r)\, dr\right)^2}{\int_0^1 \hat{H}_{\hat{a}_n,n}^{*2}(r)\, dr} \Sigma_{u\varepsilon,n}^{*\mu} \hat{\Sigma}_{u,n}^{-1} \Sigma_{u\varepsilon,n}^{*\mu} \to_p \sigma_\varepsilon^2.$$

To compute $\Sigma_{u\varepsilon,n}^{*\mu}$ and $\left(\sigma_{\varepsilon,n}^{*\mu}\right)^2$, we require simulations of $\hat{H}_{\hat{a}_n,n}^*(r)$, as described above for the estimation of μ_n^*.

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

$$Y_n(r) = \mu_T r + \sigma_{\varepsilon,T} W(r). \tag{25}$$

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.,

$$\frac{dS(r)}{S(r)} = T\mu_A dr + \sqrt{T}\sigma_{\varepsilon,A} dW(r) = \mu_T dr + \sigma_{\varepsilon,T} dW(r). \tag{26}$$

That is, the right side of (26) is just (25), 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.

14

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

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

$$\equiv \begin{pmatrix} \left[\Sigma^{1/2}\right]_1 \\ \left[\Sigma^{1/2}\right]_2 \end{pmatrix},$$

where $\Sigma_u^{1/2}$ is the positive definite square root of Σ_u and, for instance, $\left[\Sigma^{1/2}\right]_1 \equiv \left(\left[\Sigma^{1/2}\right]_{1,1} \vdots 0\right)$. 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 Supplement that

$$Y_n(r) = T\mu_A e^{a_n'\Sigma_{u,A}^{1/2}B_u^*(r)} \int_0^r e^{-a_n'\Sigma_{u,A}^{1/2}B_u^*(p)}dp + \sqrt{T}G_{a_n,A}(r), \qquad (27)$$

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

$$G_{a_n,A}(r) = e^{a_n'\Sigma_{u,A}^{1/2}B_u^*(r)} \left(\left[\Sigma_A^{1/2}\right]_2 \int_0^r e^{-a_n'\Sigma_{u,A}^{1/2}B_u^*(p)}dB^*(p)\right.$$

$$\left. -a_n'\Sigma_{u\varepsilon,A} \int_0^r e^{-a_n'\Sigma_{u,A}^{1/2}B_u^*(p)}dp\right). \qquad (28)$$

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

$$
dY_n(r) = \left[T\mu_A + \sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} G_{a_n,A}(r) - a_n'\Sigma_{u\varepsilon,A} \right) \right] dr
$$
$$
+ \sqrt{T}b_A(r)'\left[\Sigma_A^{1/2} \right] dB^*(r). \tag{29}
$$

We replace the standard formula (26) with (29), which leads to the following geometric nonlinear diffusion that is based on the STUR model

$$
\frac{dS(r)}{S(r)} = \left[T\mu_A + \sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} G_{a_n,A}(r) - a_n'\Sigma_{u\varepsilon,A} \right) \right] dr \tag{30}
$$
$$
+ \sqrt{T}b_A(r)'\left[\Sigma_A^{1/2} \right] dB^*(r).
$$

Whereas (26) is a geometric Brownian motion, the system (30) 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 (26) when $a = 0$. So, the geometric nonlinear diffusion process (30) may be regarded as a stochastic process that is parametrically local to geometric Brownian motion in the sense that when $a = 0$, the process reduces to geometric Brownian motion because the (localizing) coefficient $a_n = (T/n)^{1/2}a = 0$ in that case

Next, consider the process $G(r) = \log(S(r))$. We show in the Supplement that

$$
dG(r) = \left(T\left(\mu_A - \frac{\sigma_{\varepsilon,A}^2}{2} \right) - \sqrt{T}a_n'\Sigma_{u\varepsilon,A} \right) dr \tag{31}
$$
$$
+ \left(\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} - \sqrt{T}a_n'\Sigma_{u\varepsilon,A} \right) G_{a_n,A}(r) - T\frac{a_n'\Sigma_{u,A}a_n}{2} G_{a_n,A}^2(r) \right) dr
$$
$$
+ \sqrt{T}b_A(r)'\Sigma_A^{1/2}dB^*(r).
$$

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

$$
dG(r) = \left(\mu_T - \frac{\sigma_{\varepsilon,T}^2}{2} \right) dr + \sqrt{n}dB_\varepsilon(r) = T\left(\mu_A - \frac{\sigma_{\varepsilon,A}^2}{2} \right) dr + \sqrt{T}\sigma_{\varepsilon,A}dB_\varepsilon^*(r). \tag{32}
$$

16

In this case, (32) implies that

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

so that

$$S\left(r\right) = S\left(0\right) \exp\left\{N\left(\left(\mu_T - \frac{\sigma_{\varepsilon,T}^2}{2}\right)r, \sigma_{\varepsilon,T}^2 r\right)\right\}. \tag{33}$$

In the Supplement it is shown that when $a \neq 0$, $S\left(r\right)$ satisfies

$$S\left(r\right) = S\left(0\right) \exp\left\{\left(T\left(\mu_A - \frac{\sigma_{\varepsilon,A}^2}{2}\right) - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)r\right.$$
$$+ \int_0^r \left(\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)G_{a_n,A}\left(s\right)\right.$$
$$\left. -T\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}^2\left(s\right)\right)ds$$
$$\left. +\sqrt{T}a_n'\int_0^r G_{a_n,A}\left(s\right)\Sigma_{u,A}^{1/2}\left(s\right)dB_u^*\left(s\right) + \sqrt{T}\Sigma_{2,A}^{1/2}B^*\left(r\right)\right\}. \tag{34}$$

Equation (34) is the price process under the physical measure. It will be used in place of (33) 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 $f\left(r, x\right)$ be the price at r of a European-style derivative of a stock, such as a European call option, where $x \equiv S\left(r\right)$ is the stock price. Let $\left(\alpha_S\left(r\right), \alpha_Z\left(r\right)\right)$ be the associated self-financing portfolio at time r of the stock and the riskless asset. The value of the portfolio at time r is

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

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

$$dV\left(r\right) = \alpha_S\left(r\right)dS\left(r\right) + \alpha_Z\left(r\right)r_{f,T}\gamma\left(r\right)dr, \tag{35}$$

17

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

$$\alpha_S(r) = f_x, \tag{36}$$

which is the condition in the classic case, and that

$$\alpha_Z(r) = \frac{1}{Tr_{f,A}\gamma(r)} \left\{ f_r + \frac{T}{2} f_{xx} S^2(r) b_A(r)' \Sigma_A b_A(r) \right\}. \tag{37}$$

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

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

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 (36) and (37) 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,

$$Tr_{f,A} f_x S(r) + f_r + \frac{T}{2} f_{xx} S^2(r) b_A(r)' \Sigma_A b_A(r) = Tr_{f,A} f. \tag{38}$$

Equation (38) is the generalized BS sde for a European style derivative of a stock. When $a = 0$ the formula reduces to the well-known relationship

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

When $K = 1$ and $a \neq 0$, (38) becomes

$$Tr_{f,A} f_x S(r) + f_r + \frac{T}{2} f_{xx} S^2(r) \left\{ G_{a_n,A}^2(r) a_n^2 \sigma_{u,A}^2 \right.$$
$$\left. + 2a_n \sigma_{u\varepsilon,A} G_{a_n,A}(r) + \sigma_{\varepsilon,A}^2 \right\} = Tr_{f,A} f.$$

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 $\beta_t(a; n) = \exp(a'u_t/\sqrt{n})$ 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 imply risk beyond that of the simple equation error since the presence of u_t in the time varying coefficient $\beta_t(a; n) = \exp(a'u_t/\sqrt{n})$ means that uncertainty now enters through the equation conditional mean function (which depends on the distribution of u_t) and correspondingly affects conditional variance. In the continuous time context, the shocks are manifest in the additional random component, $\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2}G_{a_n,A}(r) - a_n'\Sigma_{u\varepsilon,A}\right)dr$, that appears in the drift of the log price stochastic differential equation. In effect, the STUR model implies uncertainty and risk in the price process drift, which is now no longer constant, thereby producing stochastic variation in investment alpha (the drift in the log price process), as well as contributing to the instantaneous conditional volatility process $\sqrt{T}b_A(r)'\left[\Sigma_A^{1/2}\right] = \sqrt{T}(G_{a_n,A}(r)a_n', 1)\left[\Sigma_A^{1/2}\right]$ through the presence of $G_{a_n,A}(r)a_n$. For the latter, It is clear that both the drift and martingale components of the sde of the conditional variance process depend in a nonlinear way on $G_{a_n,A}(r)$, which itself satisfies a nonlinear diffusion process (with a stochastic drift). In view of these complexities, it is not apparent how the present model might be mapped into a more

19

conventional stochastic volatility model specification.

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. 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 $E(X^Q)$ 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

$$E^Q(S(r)) = E(S^Q(r)) = S(0) e^{r_{f,T} r}, \tag{39}$$

see, for instance, Hull (2009, eq'n (13.18)). 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:

$$\begin{pmatrix} B_u^*(r) \\ B_\varepsilon^*(r) \end{pmatrix} = \begin{pmatrix} B_u^{*Q}(r) \\ B_\varepsilon^{*Q}(r) + \gamma r \end{pmatrix}, \gamma \in \mathbb{R}, \tag{40}$$

where the affix Q indicates that the SBM's are under Q. We have

$$dB_\varepsilon^*(r) = dB_\varepsilon^{*Q}(r) + \gamma dr \tag{41}$$

and as

$$\begin{aligned}
\left[\Sigma_A^{1/2}\right]_2 B^*(p) &= \left[\Sigma_A^{1/2}\right]_{2,1} B_u^*(p) + \left[\Sigma_A^{1/2}\right]_{2,2} B_\varepsilon^*(p) \\
&= \left[\Sigma_A^{1/2}\right]_2 B^{Q*}(p) + \left[\Sigma_A^{1/2}\right]_{2,2} \gamma p,
\end{aligned} \tag{42}$$

it follows from (28) that

$$\begin{aligned}
G_{a_n,A}(r) &= \gamma e^{a_n' \Sigma_{u,A}^{1/2} B_u^{*Q}(r)} \left[\Sigma_A^{1/2}\right]_{2,2} \int_0^r e^{-a_n' \Sigma_{u,A}^{1/2} B_u^{*Q}(p)} dp + G_{a_n,A}^Q(r) \\
&= \gamma \xi^Q(r) + G_{a_n,A}^Q(r),
\end{aligned} \tag{43}$$

say, where $G_{a_n,A}^Q(r)$ is $G_{a_n,A}(r)$ with $B_u^{*Q}(r)$ and $B_\varepsilon^{*Q}(r)$ in the former replacing $B_u^*(r)$ and $B_\varepsilon^*(r)$ in the latter, respectively. We obtain from (34)

$$
\begin{aligned}
S^Q(r) = S(0)\exp\Bigg\{ &\left(T\left(\mu_A - \frac{\sigma_{\varepsilon,A}^2}{2}\right) - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)r \\
&+ \sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)\int_0^r \left(\gamma\xi^Q(s) + G_{a_n,A}^Q(s)\right)ds \\
&- T\frac{a_n'\Sigma_{u,A}a_n}{2}\int_0^r \left(\gamma\xi^Q(s) + G_{a_n,A}^Q(s)\right)^2 ds \\
&+ \sqrt{T}a_n'\Sigma_{u,A}^{1/2}\int_0^r \left(\gamma\xi^Q(s) + G_{a_n,A}^Q(s)\right)dB_u^{Q*}(s) \\
&+ \sqrt{T}\left(\left[\Sigma_A^{1/2}\right]_2 B^{*Q}(r) + \left[\Sigma_A^{1/2}\right]_{2,2}\gamma r\right)\Bigg\} \\
= S(0)&\,\phi_\gamma^Q(r)\lambda^Q(r),
\end{aligned}
\tag{44}
$$

where

$$
\begin{aligned}
\phi_\gamma^Q(r) = \exp\Bigg\{ &\left(T\mu_A - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)r \\
&+ \gamma\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)\int_0^r \xi^Q(s)\,ds \\
&- \gamma^2 T\frac{a_n'\Sigma_{u,A}a_n}{2}\int_0^r \left(\xi^Q(s)\right)^2 ds \\
&- \gamma T a_n'\Sigma_{u,A}a_n\int_0^r G_{a_n,A}^Q(s)\xi^Q(s)\,ds \\
&+ \gamma\sqrt{T}a_n'\Sigma_{u,A}^{1/2}\int_0^r \xi^Q(s)\,dB_u^{*Q}(s) \\
&+ \gamma r\sqrt{T}\left[\Sigma_A^{1/2}\right]_{2,2}\Bigg\}
\end{aligned}
\tag{45}
$$

21

and

$$\lambda^Q(r) = \exp\left\{-\frac{\sigma_{\varepsilon,A}^2}{2}Tr\right.$$

$$+ \int_0^r \left(\sqrt{T}\left(\frac{a_n'\Sigma_{u,A}a_n}{2} - \sqrt{T}a_n'\Sigma_{u\varepsilon,A}\right)G_{a_n,A}^Q(s)\right.$$

$$\left.-T\frac{a_n'\Sigma_{u,A}a_n}{2}\left(G_{a_n,A}^Q(s)\right)^2\right)ds$$

$$\left.+\sqrt{T}\left(a_n'\Sigma_{u,A}^{1/2}\int_0^r G_{a_n,A}^Q(s)\,dB_u^{Q*}(s) + \left[\Sigma_A^{1/2}\right]_2 B^{Q*}(r)\right)\right\}. \tag{46}$$

We see that in the BS case, at $r = 1$ eq'ns (45) and (46) collapse to

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

and

$$\lambda_{BS}^Q(1) = \exp\left(-\frac{\sigma_{\varepsilon,A}^2}{2}T + \sqrt{T}\sigma_{\varepsilon,A}B_\varepsilon^{Q*}(1)\right). \tag{47}$$

Setting $\tilde{\gamma} = \sqrt{T}\left(r_{f,A} - \mu_A\right)/\sigma_{\varepsilon,A}$, we obtain $\phi_{\tilde{\gamma},BS}^Q(1) = e^{Tr_{f,A}}$. With this choice of γ, eq'n (44) reduced in this case to

$$S_{BS}^Q(1) = S(0)\,e^{Tr_{f,A}}(1)\,\lambda_{BS}^Q(1). \tag{48}$$

As $E\left(\lambda_{BS}^Q(1)\right) = 1$, (39) is satisfied, as required.

In the more general case we may rewrite eq'n (44) at $r = 1$ as

$$S^Q(1) = S(0)\left(\phi_\gamma^Q(1)E\left(\lambda^Q(1)\right)\right)\bar{\lambda}^Q(1), \tag{49}$$

where

$$\bar{\lambda}^Q(1) \equiv \frac{\lambda^Q(1)}{E\left(\lambda^Q(1)\right)}. \tag{50}$$

We can set $\gamma = \tilde{\gamma}$ such that the middle term in (49) is equal to $e^{Tr_{f,A}}$, giving

$$S^Q(1) = S(0)\,e^{Tr_{f,A}}\bar{\lambda}^Q(1) \tag{51}$$

and therefore, to obtain (39), as required. We remark that in practical terms we do not need to find $\tilde{\gamma}$ explicitly, but rather only to say that once it is

set appropriately, (39) is satisfied. To emphasize, the Q measure is defined through a shift by γ, as described by (40). Whereas in the BS case the value of $\tilde{\gamma}$ is known explicitly, it is not known in our case. Nevertheless, as described above, we do not need to know it explicitly in order to obtain the risk neutral pricing.

Continuing, once the value of $\tilde{\gamma}$ is set, in a sense a first moment matching between $\lambda_{BS}^Q(1)$ in (47) and $\bar{\lambda}^Q(1)$ in (50) is attained. In both cases the mean is equal to unity. This device is sufficient for the purpose of achieving risk neutral pricing but a possible refinement can be made, as explained below.

Noting that

$$Var\left(\lambda_{BS}^Q(1)\right) = e^{\sigma_{\varepsilon,A}^2 T} - 1,$$

let

$$\lambda^{*Q}(1) = \frac{\left(e^{\sigma_{\varepsilon,A}^2 T} - 1\right)^{1/2}}{\sigma\left(\lambda^Q(1)\right)}\left(\lambda^Q(1) - E\left(\lambda^Q(1)\right)\right) + 1. \tag{52}$$

Evidently,

$$E\left(\lambda^{*Q}(1)\right) = E\left(\lambda_{BS}^Q(1)\right) = 1 \tag{53}$$

and

$$Var\left(\lambda^{*Q}(1)\right) = e^{\sigma_{\varepsilon,A}^2 T} - 1 = Var\left(\lambda_{BS}^Q(1)\right).$$

We can use (52) to rewrite (44) as

$$S^Q(1) = S(0)\,\phi_\gamma^Q(1)\left(\frac{\sigma\left(\lambda^Q(1)\right)}{\left(e^{\sigma_{\varepsilon,A}^2 T} - 1\right)^{1/2}}\left(\lambda^{*Q}(1) - 1\right) + E\left(\lambda^Q(1)\right)\right)$$

$$= S(0)\left\{\phi_\gamma^Q(1)\left(\frac{\sigma\left(\lambda^Q(1)\right)}{\left(e^{\sigma_{\varepsilon,A}^2 T} - 1\right)^{1/2}}\left(1 - \frac{1}{\lambda^{*Q}(1)}\right)\right.\right. \tag{54}$$

$$\left.\left. + \frac{E\left(\lambda^Q(1)\right)}{\lambda^{*Q}(1)}\right)\right\}\lambda^{*Q}(1),$$

say. We now set $\gamma = \tilde{\gamma}$ such that the term in the large braces of (54) is equal

to $e^{Tr_{f,A}}$. With this choice of γ, we let

$$S^{*Q}(1) = S(0) e^{Tr_{f,A}} \lambda^{*Q}(1), \qquad (55)$$

so that, by (53),

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

as required. Compared to (51), in this case the variance of $S^{*Q}(1)$ is equal to the variance of $S^Q(1)$ under BS. Therefore, this refinement achieves a two-moment matching of the STUR-based $S^{*Q}(1)$ with the BS-based $S^Q(1)$. Higher order cumulants, are still different though.

We remark that the fact should come at no surprise that whereas in the BS case $\tilde{\gamma}$ is non-random, in our case it is. This is a result of the high non-linearity of our model. The mean and variance matching seems to yield superior accuracy in option pricing, as our numerical experiment reveals and discussed in the next section.

In order to simulate the value of a European call option maturing in T years, we calculate $G^Q_{a_n,A}(r)$ 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. Using (51), the estimated European call option price is

$$\hat{C} = e^{-r_{f,T}} \overline{\max\left\{S^Q(1) - K, 0\right\}}, \qquad (56)$$

or

$$\hat{C}^* = e^{-r_{f,T}} \overline{\max\left\{S^{*Q}(1) - K, 0\right\}} \qquad (57)$$

if the refinement (55) is preferred.

6 Simulations

In this Section we corroborate by simulations some of the main asymptotic results given in Section 3. For the case $\mu \neq 0$, we check the $\sqrt{n}$ - rate of $(\hat{a}_n - a_0)$ in both the $a_0 = 0$ and $a_0 \neq 0$ subcases. Furthermore, for the $\mu \neq 0$ case, we analyze the small sample behavior of the second order expansion based test for the hypothesis $H_0 : a = 0$, given in Corollary 6. Finally, for the $\mu = 0$ case, we give kernel density plots of the asymptotic distribution of $\hat{a}_n$ given by Theorem 2(2). Treating the the cases $\mu = 0$ and

24

$\mu \neq 0$ separately is important because, as discussed in Section 3, the rates of decay associated with the estimators are different between the cases, with $\hat{a}_n$, for instance, being inconsistent in the former case.

For all the results of the case $\mu \neq 0$, we have used the parameter values $\mu = 0.000941$, $\sigma_u^2 = 0.000177$, $\sigma_\varepsilon^2 = 0.000284$ and $\rho = 0.6923$. These values were chosen to correspond to estimates obtained in the empirical application, reported in the next Section.

To check the $\sqrt{n}$ - rate of $\hat{a}_n$ in the $\mu = 0$ and $a = 0$ subcase, as given in Remark 8, we generated 250 samples of $n = 1000, 2000, ..., 250000$ observations of $\{u_t, \varepsilon_t\}$ from a bivariate, zero mean normal distribution with a covariance matrix consistent with the parameter value-setting described above. With these, we generated Y-samples from a random walk with a drift model, which corresponds to this subcase and consequently, estimated the STUR model (6)-(7) to obtain $\hat{a}_n$.

To corroborate the result in Remark 8, we denote the ratio of the left and right sides of (20) by R_n, that is

$$R_n = \frac{2\sqrt{n}\sigma_u \mu \hat{a}_n}{3\rho\sigma_\varepsilon}, \tag{58}$$

and use the estimates of $\hat{a}_n$ to plot R_n against n, for $n = 1000, 2000, ..., 250000$. The graph of (58) is provided in Figure 1 of the supplement where it is evident that R_n converges to unity, as expected. With the same simulation estimates we obtained the following regression results

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

The implication of Figure 1 and the last equation is that $|\hat{a}_n| \sim n^{1/2}$, corroborating the $\sqrt{n}$-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 250 Y-samples with $n = 1000, 2000, ..., 250000$ (daily) observations from (6)-(7), consequently estimating $\hat{a}_n$ and obtaining a plot of $\sqrt{n}(\hat{a}_n - a_0)$ against n in Figure 2 of the supplement, 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 assess the accuracy of the second order expansion with parameter estimates, eq'n (23), we generated 5000 samples from the null model (a random walk with a drift) with $n = 1000, 10000, 100000$. For each replication we estimated the unrestricted model and recorded the value of $\hat{a}_n^{(R)}$, $R = 1, ..., 5000$. For each replication we also estimated the restricted estimates, $\hat{\mu}_n^0$, $\hat{\sigma}_{u\varepsilon,n}^0$, as well as $\hat{\sigma}_{u,n}^2$, which are consistent under the null. With these we calculated vectors of the lhs and the rhs of (23) (recalling that we have neglected terms on the rhs which are $O_p\left(n^{-1/2}\right)$). In Figures 3-5 of the supplement we provided PP-plots of the lhs of (23) against the rhs (23). Clearly, with $n = 1000$ the asymptotic approximation is poor, it improves but is still not satisfactory with $n = 10000$ and is very good with $n = 100000$.

In sum, these experiments broadly corroborate the analytical findings. To complete this subsection, we generated $100,000$ replications with $n = 1231$ of the rhs of Theorem 2(2), which gives the asymptotic distribution of $\hat{a}_n$ in the $\mu = 0$, $\sigma_{u\varepsilon} \neq 0$ and $a = 0$ case, using the same parameter values as described above. The resulting kernel density estimate is given in Figure 6 of the supplement. The distribution is evidently bimodal with a local minimum at zero.

7 An Empirical Application

We downloaded from Yahoo Finance daily data on the closing prices of Google, Apple Inc. and Nasdaq composite indexes (tickers GOOG, AAPL 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 the estimation window and is post the 2008 market crash, so we avoid possible issues of structural breaks in the illustrations. We have also obtained end of day option data from December 2013 from Historical Option Data Database (available at www.historicaloptiondata.com). Google is a non-dividend paying stock and is therefore suited to this application.

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) at the end of the sample period. We have also downloaded the 3-months U.S. treasury Bill rates from the Federal Reserve Bank of St. Louis database (https://research.stlouisfed.org), but these were identical (in annual terms) to the Federal Funds Rate. 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.

We start with the single regressor STUR model, expanding it to a two-regressor application later. We obtained estimates of the following empirical STUR model (in obvious notation)

$$\log(\widehat{Google})_t = 0.000941 \tag{59}$$
$$+ \exp\left(\frac{4.9540\left(\Delta\log\left(Nasdaq\right)_t - 0.000713\right)}{\sqrt{1231}}\right)\log\left(Google\right)_{t-1},$$

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 $\hat{\sigma}_{u,A}^2 = 0.0446$, $\hat{\sigma}_{\varepsilon,A}^2 = 0.0716$ and $\hat{\rho} = 0.6923$, over the same sample period. The Akaike (AIC), Schwarz (SC) and sum of squared errors (SSE) values for the model (59) 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.

To test the hypothesis $H_0 : a_0 = 0$, we used Corollary 6 in the following way. First, we ran our STUR regression with an intercept over subsamples of increasing sizes of the available data. The values of $\hat{a}_n$ against n are given in Table 1. Clearly, $\hat{a}_n$ is stable wrt n and is approximately equal to 5, whereas, by Remark 8, under H_0 it should be $O_p\left(n^{-1/2}\right)$ under H_0. By the results of Section 6, the second order asymptotic expansion is accurate for $n = 100000$. We therefore calculated a test value $T_{100000} = 54757$ and simulated p-value of 0.0004 for the rhs of (23). Thus, in terms of the AIC, SC, SSE and the formal test procedure, the use of the STUR model over the benchmark random walk with a drift model is well and truly justified.

We proceed to evaluate the usefulness of the STUR model in option pricing. We compared between actual option prices to those obtained by: BS, Heston's (1993) SV model[5], the STUR-based $\hat{C}$ given by (56) and using

---

[5]BS classic call option prices (see, for instance, eq'n (13.20) of Hull (2009)) were calculated in MATHEMATICA. Heston's (1993) SV model was estimated with the approximate MLE procedure suggested by Aït-Sahalia and Kimmel (2007). A detailed summary of the results is given in the Supplement to this paper.

the bivariate model (59), and the mean and variance adjusted STUR-based $\hat{C}^*$, given by (57) and using the bivariate model (59). In addition, we have evaluated (56) and (57) after fitting a multivariate extension to (59), viz.,

$$
\begin{aligned}
\log\widehat{(Google)}_t = {} & 0.000944 \\
& + \exp\left(\frac{4.5903\left(\Delta\log\left(Nasdaq\right)_t - 0.000713\right)}{\sqrt{1231}}\right. \\
& \left. + \frac{0.3921\left(\Delta\log\left(AAPL\right)_t - 0.001411\right)}{\sqrt{1231}}\right)\log\left(Google\right)_{t-1}, \quad (60)
\end{aligned}
$$

where Apple inc. had an average return of 0.001411 over the sample period. Also, as a sensitivity analysis, in order to assess the effect of the correlation parameter on the results, we have calculated $\hat{C}\left(\rho=0\right)$, $\hat{C}\left(\rho=0.95\right)$, $\hat{C}^*\left(\rho=0\right)$ and $\hat{C}^*\left(\rho=0.95\right)$. For Table 2, we took expirations as of 11-22-2013, in 28, 420 and 784 days and strike prices $K = 950, 1000, 1030, 1100,$ 1120. The closing price for Goggle stock on 11-22-2013 was 1031.89. The results are given in Table 2.

Secondly, in a similar fashion to the analysis carried out by Bakshi *et. al.* (1997), in Table 5 we report the average percentage pricing errors of the various pricing schemes over 2287 option prices with varying maturity-moneyness categories, where moneyness is defined as the ratio of the stock price to the strike price. For this comparison, we concentrated on BS, SV, $\hat{C}$ and $\hat{C}^*$.

Before we turn to the results, a few comments are in order. For T, we substituted the number of days to expiration divided by 365. The prices $S^Q\left(1\right)$ were evaluated with 500 integral points and 2000 replications over simulated Brownian motions. The estimate $\hat{a}_n$ was obtained from daily data. To simulate $S^Q\left(1\right)$ 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, recall that $a_n = \left(T/n\right)^{1/2}a$. Moreover, had (59) been estimated 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

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

We now turn to discuss the results. A comparison of SV, BS, $\hat{C}$ and $\hat{C}^*$ in Table 2 reveals that $\hat{C}^*$ is best in 9 out of the 15 cases, $\hat{C}$ is best in 3 of the cases and SV is best in the remaining 3. In view of (47), (50) and (52), as BS, $\hat{C}$ and $\hat{C}^*$ only differ from each other by the terms $\lambda_{BS}^Q(1)$, $\bar{\lambda}^Q(1)$ and $\lambda^{*Q}(1)$, an investigation of their behavior will be sufficient for the explanation of the price differences, and their summary statistics are provided in Table 3. The standard deviation of $\lambda^{*Q}(1)$ is equal to that of $\lambda_{BS}^Q(1)$ by construction and was found to be smaller than the standard deviation of $\bar{\lambda}^Q(1)$. On the other hand, the skewness and kurtosis coefficients of $\bar{\lambda}^Q(1)$, which are equal to those of $\lambda^{*Q}(1)$ respectively by construction, are greater than those of $\lambda_{BS}^Q(1)$. The equality of the skewness and kurtosis coefficients for $\bar{\lambda}^Q(1)$ and $\lambda^{*Q}(1)$, on the one hand, and the smaller standard deviation of the latter compared to the former, explains why $\hat{C}$ is mostly larger than $\hat{C}^*$. Kernel density estimates of $\lambda_{BS}^Q(1)$, $\lambda^Q(1)$ $\lambda^{*Q}(1)$ are provided in Figure 7-8 of the supplement, emphasizing the peakedness of $\lambda^{*Q}(1)$ relative to $\lambda_{BS}^Q(1)$. Concerning the sensitivity of prices with respect to ρ, for both $\hat{C}$ and $\hat{C}^*$ there does not appear to be an identifiable pattern. The results concerning the multivariate extension (60) are mixed, giving superior results for $\hat{C}$ in 6 out of the 15 cases and for $\hat{C}^*$ in 7 out of the 15 cases.

We continue to analyze the daily sample of call options prices. The data comprised 2287 prices which were divided into 9 moneyness-maturity categories. The number of contracts for each category is summarized in Table 4. For a similar division of the data, see Bakshi *et. al.* (1997, p. 2029). Average percentage pricing errors in each category are given in Table 5. Overall, $\hat{C}^*$ is best in 5 out of the 9 categories , SV is best in 3 out of the 9 categories and BS is best in the remaining category. Moreover, in some cases the $\hat{C}^*$ reduces SV's average percentage pricing errors by up to 73%. All models systematically overprice options in all categories, with the greatest bias for the out of-the-money short maturity options. Apparently, $\hat{C}^*$ outperforms $\hat{C}$ in all categories. In the short maturity categories ($n < 60$), the SV model outperforms the BS model at all moneyness categories. It also outperforms $\hat{C}$ and $\hat{C}^*$ in the out-of-money and at-the-money categories, but is underper-

formed relative to $\hat{C}^*$ in the in-the-money category. The SV model suffers the greatest bias in long maturity options ($n > 120$). $\hat{C}^*$ performs particularly well for at the money ($0.97 < m < 1.03$) and in the money ($m > 1.03$) options, it has a smallest average percentage bias for in-the-money options in all maturity categories. In addition, it outperforms the rest of the models for at-the-money options in the middle- and long maturity categories, giving in total the best results in 5 out of 9 categories. The BS model has the best results only for long maturity out-of-the-money options.

To summarize, it appears that the TVC feature of the STUR model results in superior performance over the basic BS model in terms of standard AIC, SC and SSE criteria, as well as by evidence of a significant $\hat{a}_n$ estimate and in comparison to actual market option prices. The $\hat{C}^*$ formula in particular seems to reduce the average percentage pricing errors of SV in most moneyness-maturity categories and in some of them, the reduction is up to 73%.

8   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 under endogeneity in the factors ($\Sigma_{u\varepsilon} \neq 0$), the estimate of the localizing STUR coefficient $\hat{a}_n$ is inconsistent in the driftless case but is consistent in models with drift. As an empirical illustration, the results were applied in the pricing of Google's call options. The empirical results show that the new model substantially reduces the average percentage pricing error of the Black-Scholes and Heston's (1993) SV (with zero volatility risk premium) pricing schemes in most moneyness-maturity categories.

References

Bakshi, G., C. Cao and Z. Chen (1997). Empirical performance of alternative option pricing models. *The Journal of Finance* 5, 2003–2049.

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

Billingsley, P. (1968). *Convergence of Probability Measures*. New York: J. Wiley.

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 Non-stationary AR(1) Processes. *Annals of Statistics* 15, 1050–1063.

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.

31

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. *Journal of Time Series Analysis* 35, 592–623.

Lieberman, O. and P. C. B. Phillips (2016). Online Supplement to: A multivariate stochastic unit root model with an application to derivative pricing. Available online.

Malinvaud, E. (1980). *Statistical Methods of Econometrics,* 3rd Edition. Amsterdam: North–Holland.

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. (1987b). Towards a Unified Asymptotic Theory for Autoregression. *Biometrika* 74, 535–547.

Phillips, P. C. B. (2015). Edmond Malinvaud: a tribute to his contributions in econometrics. *Econometrics Journal* 18, No. 2, A1-A13.

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 V. Solo (1992). Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

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

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

Aït-Sahalia, Y. and R. Kimmel (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics* 83, 413–452.

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

Table 1. Values of $\hat{a}_n$ against n in the STUR regression of the Google-Nasdaq data

$\hat{a}_n$	n
5.049	250
5.241	500
5.041	750
4.965	1000
4.954	1231

Table 2. Google Call Option Prices

n	28	28	28	28	28
K	950	1000	1030	1100	1120
C	86	38	19	2.2	1.15
SV	86.93	47.31	**29.02**	**5.75**	**3.16**
BS	86.82	48.67	31.46	8.50	5.41
$\hat{C}$	87.03	50.00	34.79	12.80	9.61
$\hat{C}^*$	**84.54**	**45.79**	29.90	10.03	7.06
$\hat{C}\left(\rho=0\right)$	88.88	51.80	34.17	11.23	8.23
$\hat{C}\left(\rho=.95\right)$	85.31	47.90	34.31	13.64	9.91
$\hat{C}^{MV}$	86.98	49.54	34.05	13.02	9.62
$\hat{C}^*\left(\rho=0\right)$	86.74	48.22	31.07	8.22	5.54
$\hat{C}^*\left(\rho=.95\right)$	83.95	45.26	29.53	10.24	7.33
$\hat{C}^{*MV}$	84.71	46.32	29.89	10.09	7.29
n	420	420	420	420	420
K	950	1000	1030	1100	1120
C	151.3	120.4	106	76	70.4
SV	168.34	143.02	129.44	101.96	95.11
BS	159.14	133.02	119.02	90.93	84.01
$\hat{C}\left(\hat{\rho}\right)$	159.76	132.33	120.90	96.08	84.01
$\hat{C}^*\left(\hat{\rho}\right)$	**157.67**	**131.46**	**117.82**	**88.63**	**83.54**
$\hat{C}\left(\rho=0\right)$	156.85	134.04	120.91	87.72	84.63
$\hat{C}\left(\rho=.95\right)$	161.81	134.62	120.76	94.80	84.01
$\hat{C}^{MV}\left(\hat{\rho}\right)$	158.28	131.16	120.60	90.69	88.67
$\hat{C}^*\left(\rho=0\right)$	157.66	133.34	118.69	90.16	85.49
$\hat{C}^*\left(\rho=.95\right)$	155.45	131.38	115.96	87.70	81.66
$\hat{C}^{*MV}\left(\hat{\rho}\right)$	158.79	132.40	116.26	88.43	81.84

Table 2. Google Call Option Prices (Continued)

n	784	784	784	784	784
K	950	1000	1030	1100	1120
C	195.7	166.1	151.31	121.35	114.53
SV	199.12	178.41	167.20	144.09	138.18
BS	199.18	175.16	161.96	134.50	127.45
$\hat{C}$	198.66	**171.69**	**161.44**	**130.82**	129.24
$\hat{C}^*$	**198.21**	175.19	161.89	137.25	**124.28**
$\hat{C}\left(\rho=0\right)$	197.60	175.78	159.31	132.32	128.97
$\hat{C}\left(\rho=.95\right)$	200.26	177.15	162.38	137.68	130.44
$\hat{C}^{MV}$	199.08	177.86	161.90	137.01	128.21
$\hat{C}^*\left(\rho=0\right)$	197.60	175.26	165.81	133.70	125.94
$\hat{C}^*\left(\rho=.95\right)$	194.12	164.92	160.02	132.87	121.45
$\hat{C}^{*MV}$	198.89	169.94	158.79	133.47	126.80

Note: n is the number of days to expiration as of 11-22-2013; K is the strike price; $S\left(0\right) = 1031.89$; 'BS' is the price based on Black and Scholes's classic formula; SV is the price based on Heston's (1993) SV model using the procedure suggested by Aït-Sahalia and Kimmel (2007). $\hat{C}$ and $\hat{C}^*$ are based on (56)-(57); The superscript MV indicates pricing under eq'n (60).

Table 3. Summary Statistics for BS- and STUR-Based Simulated Data

n	Statistic	SD	Skewness	Kurtosis
28	$\lambda_{BS}^{Q}\left(1\right)$	0.084	0.253	3.114
28	$\lambda^{*Q}\left(1\right)$	0.084	0.970	4.543
28	$\bar{\lambda}^{Q}\left(1\right)$	0.095	0.970	4.543
420	$\lambda_{BS}^{Q}\left(1\right)$	0.291	0.898	4.469
420	$\lambda^{*Q}\left(1\right)$	0.291	1.124	5.332
420	$\bar{\lambda}^{Q}\left(1\right)$	0.299	1.124	5.332
784	$\lambda_{BS}^{Q}\left(1\right)$	0.406	1.285	6.073
784	$\lambda^{*Q}\left(1\right)$	0.406	1.988	12.103
784	$\bar{\lambda}^{Q}\left(1\right)$	0.421	1.988	12.103

Table 4. Number of Options in Each Category

	$m < 0.97$	$0.97 < m < 1.03$	$m > 1.03$
$n < 60$	407	1157	236
$60 < n < 120$	128	181	40
$n > 120$	69	44	25

Table 5. Average Percentage Pricing Errors

		$m < 0.97$	$0.97 < m < 1.03$	$m > 1.03$
$n < 60$	SV	$-\mathbf{118.3\%}$	$-\mathbf{56.5\%}$	-5.8%
	BS	-216.7%	-94.1%	-5.9%
	$\hat{C}$	-449.7%	-225.6%	-8.1%
	$\hat{C}^*$	-285.7%	-86.9%	$-\mathbf{2.9\%}$
$60 < n < 120$	SV	$-\mathbf{49.4\%}$	-23.1%	-10.9%
	BS	-52%	-17.8%	-7%
	$\hat{C}$	-100%	-21.1%	-7.1%
	$\hat{C}^*$	-77.2%	$-\mathbf{15.6\%}$	$-\mathbf{4.6\%}$
$n > 120$	SV	-74.5%	-20.8%	-9.5%
	BS	$-\mathbf{47.1\%}$	-11.5%	-3.8%
	$\hat{C}$	-64.1%	-12.2%	-3.9%
	$\hat{C}^*$	-54.3%	$-\mathbf{9.7\%}$	$-\mathbf{2.6\%}$

Note: $\hat{C}$ and $\hat{C}^*$ denote STUR based pricing and mean and variance STUR based pricing. The average percentage pricing error is defined as $\frac{1}{n_j}\sum_{i \in j}\frac{C_i - \hat{C}_i}{C_i}$, where n_j denotes the total number of options in category j and C_i and $\hat{C}_i$ represent the market price and the estimated model price, respectively.