Empirical Implementation of a Term Structure Model with Stochastic Volatility

Fabio Fornari
Banca d'Italia, Research Department
Via Nazionale 91 00184 Roma, Italy
e-mail: forn.2168@interbusiness.it

Antonio Mele
THEMA, Université de Paris X
200, avenue de la République (Bâtiment G), 92001 Nanterre Cedex, France
e-mail: antonio.mele@u-paris10.fr

This version: April, 1999

Incomplete: Estimation results forthcoming

Abstract

We provide the empirical implementation of the term-structure model developed in Fornari and Mele (1998). This model is based on a continuous time economy exhibiting equilibrium dynamics to which most asymmetric ARCH models converge in distribution as the sample frequency gets infinite. We obtain estimates of the model's parameters that are based on an indirect inference scheme in which such convergence results are used to exploit ARCH as auxiliary models. With such estimates at hand, we implement a Crank-Nicholson - type scheme and numerically solve for the equilibrium term structure predicted by our theoretical model. We find that shocks to the short term interest rate co-move positively with shocks to volatility and that the whole term-structure of interest rates sharpens as volatility increases.

Keywords: Term structure, stochastic volatility, ARCH, estimation of stochastic differential equations, indirect inference, Crank-Nicholson

1 Introduction

ARCH models have been widely used to estimate models with time-varying volatility; see Fornari and Mele (FM) (1999) for an account. Despite the popularity of ARCH, there are yet a few rigorous theoretical sounded schemes of economic equilibrium that embed some of the statistical aspects of these models. This is particularly true in the continuous-time case. Motivated by the important links that now exist between the discrete-time nature of ARCH with the continuous-time models that are typically used in finance, FM (1998) propose to construct a class of continuous-time economies displaying equilibrium dynamics to which most asymmetric ARCH models converge in distribution as the sample frequency gets infinite.

In this paper, we implement empirically one of the resulting equilibrium models by relying, accordingly, on ARCH as direct or indirect approximation of stochastic volatility. In the first case, the moment conditions that guarantee the weak convergence of ARCH are exploited and deliver a direct, preliminary estimate of the model's parameters. In the second case, ARCH models are used as auxiliary devices in an indirect inference procedure (Gouriéroux and Monfort (1993)); in this case, we explicitely follow the strategy on the correction of the asymptotic bias due to the "discretization" of the likelihood, whose principle is to use auxiliary models that can be embedded in high frequency simulating schemes converging in distribution towards the solution of our equilibrium model.

As emphasized by Campbell, Lo and MacKinlay (1997) (p.381), the empirical properties of ARCH as approximators of continuous-time stochastic volatility processes "have yet to be explored but will no doubt be the subject of future research." This is precisely what is attempted here. Our main concern is the estimation of the term-structure model presented in FM (1998). As explained before, we rely on weak convergence conditions and make use of ARCH models as *direct* approximators of the (theoretical) diffusions of the model. In a second step, we try to correct the bias due to the discrete nature of ARCH by indirect inference. It is exactly the amount of such a bias that defines the appropriateness of ARCH as direct approximators of our theoretical model. We examine the importance of the biais on the term-structure.

2 The theoretical model

Here we succinctly recall some of the features of the term structure model in FM (1998). It is a continuous time equilibrium model with state variables—much in the spirit of Cox, Ingersoll and Ross (CIR) (1985a, b)—and a representative agent having logarithmic instantaneous felicity.[1] One prediction of the model is that the short term-interest rate is the solution of the following SDE:

$$\begin{aligned} \mathrm{d}r &= (\iota - \theta r)\mathrm{d}t + \sqrt{r}\sigma \mathrm{d}W^{(1)} \\ \mathrm{d}\sigma^{\delta} &= (\overline{\omega} - \varphi\sigma^{\delta})\mathrm{d}t + \psi\sigma^{\delta}\mathrm{d}(\rho W^{(1)} + \sqrt{1-\rho^2}W^{(2)}) \end{aligned} \tag{2.1}$$

In the remainder, we shall refer to system (2.1) as the *equilibrium data generating process* (EDGP). The natural justification for such a denomination is that system (2.1) has *not* been imposed a priori, but *derived* from a fully articulated equilibrium model with a well specified primitive measure space. Accordingly, equilibrium also predicts that the risk premia demanded by agents to compensate for the fluctuations of $W^{(1)}$ and $W^{(2)}$ are:

$$\lambda^{(i)}(r, v) \equiv \lambda_i \sqrt{r} \qquad (i = 1, 2), \tag{2.2}$$

where λ_i are two real constants, and the equilibrium price of the bond satisfies the following PDE:

$$\begin{aligned} 0 \;=\; & B_t + B_r(\iota - \theta r) + B_{\sigma^\delta}(\overline{\omega} - \varphi\sigma^{\delta}) + \tfrac{1}{2}\left(B_{rr} r\sigma^2 + 2B_{r\sigma^\delta}\psi\rho\sqrt{r}\sigma^{\delta+1} + B_{\sigma^\delta\sigma^\delta}\psi^2\sigma^{2\delta}\right) \\ & -[B_r r\sigma\lambda_1 + B_{\sigma^\delta}\left(\psi\rho\sigma^{\delta}\sqrt{r}\lambda_1 + \psi\sqrt{1-\rho^2}\sigma^{\delta}\sqrt{r}\lambda_2\right)] - rB. \end{aligned} \tag{2.3}$$

In FM (1998), we implemented a solution of the above PDE that is based on a method of iterated approximations. Such a method relies on a functional iteration of a certain benchmark affine pricing rule under the action of the associated Arrow-Debreu state price. It requires the computation of multi-dimensional integrals, and here we find that a more traditional approach based on a numerical integration of the PDE (2.3) can give results in a faster manner:[2] specifically, we make use of the Crank-Nicholson method, and provide details in appendix A.

3 Statistical inference

CIR*b* found the transition and stationary densities of $\{r_t\}_{t\geq 0}$ in the case of constant volatility, and this enabled Aït-Sahalia (1996*a*) to compare those densities with the ones obtained non-parametrically.

There is no hope to find an analytical solution of the transition density of r in our model, and such a difficulty is inherent in virtually all models with stochastic volatility. As shown by FM (1998), however, there are ARCH models that converge (in the weak-sense) towards the solution of (2.1), and this suggests using these models as auxiliary

[1]Because markets are complete in our model, the use of the representative agent is by no means restrictive; see Huang (1987).

[2]Notice that Wiggins (1987) already adopted such an approach in his seminal paper on the pricing of European-type options with stochastic volatility.

devices in indirect inference procedures. In fact, it is also such a kind of convergence results that motivated FM (1998) to look for economies supporting the equilibrium dynamics in (2.1).

The parameters of interest are in $a^+ \equiv (\iota, \theta, \overline{\omega}, \varphi, \psi, \delta, \rho)$. As suggested in FM (1998), we shall assume hereafter that the "volatility concept" δ and the asymmetry parameter γ are invariant with respect to time-scale changes. Ascertaining whether such an assumption is reasonable in practice is an open question that we leave for future research; here we propose to estimate δ and γ by fitting the P-ARCH model. We thus consider the following model:

$$\begin{array}{rcl} r_t - r_{t-1} & = & \iota_1 - \theta_1 r_{t-1} + \sigma_t \sqrt{r_{t-1}} u_t \\ \sigma_{t+1}^\delta - \sigma_t^\delta & = & \left[\omega_1 - \left(1 - \beta_1 - n_{\delta,\upsilon}((1-\gamma)^\delta + (1+\gamma)^\delta)\alpha_1\right)\sigma_t^\delta\right] \\ & & +\alpha_1 \sigma_t^\delta \left[|u_t|^\delta (1-\gamma s_t)^\delta - E\left(|u_t|^\delta (1-\gamma s_t)^\delta\right)\right] \end{array} \tag{3.1}$$

and define the ML estimator of $b^+ = (\iota_1, \theta_1, \omega_1, \alpha_1, \beta_1, \gamma, \delta)'$:

$$\widehat{b}_N^+(a_0) = \arg\max_{b^+} \mathfrak{L}_N({}_1r; b^+), \quad b^+ \in B \subset \mathbb{R}_+^7,$$

where $\mathfrak{L}_N({}_1r; b^+)$ is the likelihood function implied by the model, ${}_1r$ is the observations set and N is the sample size. The estimator of δ is thus $\widehat{\delta}$, and the estimator of ρ is obtained by plugging $\widehat{\gamma}, \widehat{\delta}, \widehat{\upsilon}$ into formula (3.5) in FM (1998). To simplify, we shall only consider normally distributed errors, i.e., $\upsilon \equiv 2$. With all preceding choices, the parameters of interest to be estimated reduce now to those in $a \equiv (\iota, \theta, \overline{\omega}, \varphi, \psi)$.

It is well-known that under standard regularity conditions such as those in assumption B1 of appendix B, one has asymptotic normality for the pseudo-maximum likelihood estimator:

$$\sqrt{N}\left[\widehat{b}_N^+(a_0) - b_0^+(a_0)\right] \xrightarrow{d} \mathrm{N}\left(0, \ddot{\mathfrak{L}}_\infty^{-1}(a_0; b_0^+(a_0)) \cdot J(a_0) \cdot \ddot{\mathfrak{L}}_\infty^{-1}(a_0; b_0^+(a_0))\right),$$

where $\ddot{\mathfrak{L}}_\infty(.)$ and $J(.)$ are defined in appendix B, and

$$b_0^+(a_0) = \arg\max_{b^+} \mathfrak{L}_\infty(a_0; b^+), \text{ the limit problem.}$$

However, the true law of r, as implied by the EDGP, say $\ell_0({}_1r)$, is such that

$$\ell_0({}_1r) \notin \{\mathfrak{L}_N({}_1r; b^+),\ b^+ \text{ varying}\},$$

and the discrete-time model is expected to behave in a way that allows for:

$$b^+(a_0) \neq a_0.$$

While we are assuming that $(\widehat{\gamma}, \widehat{\delta})$ allow for the reconstruction of their continuous-time counterparts, the indirect inference procedure can correct the preceding discretization bias. The reason why we view the preceding inequality as a "discretization bias" is that when we chop time in (3.1) by creating sequences of the form $\{\iota_h, \theta_h, w_h, \alpha_h, \beta_h\}$, and by defining a stochastic process $\{{}_h r_{hk}, {}_h \sigma_{hk}^{\delta}\}_{k=0,1,\ldots}$, solution of:

$$\left\{\begin{array}{rcl} {}_h r_{h(k+1)} - {}_h r_{hk} & = & (\iota - \theta \cdot {}_h r_{hk})h + {}_h\sigma_{h(k+1)}\sqrt{{}_h r_{hk}} \cdot {}_h u_{h(k+1)} \\ {}_h\sigma_{h(k+1)}^{\delta} - {}_h\sigma_{hk}^{\delta} & = & (\overline{\omega} - \varphi \cdot {}_h\sigma_{hk}^{\delta})h + \sqrt{\frac{h}{(m_{\delta,\upsilon} - n_{\delta,\upsilon}^2)\left((1-\gamma)^{2\delta} + (1+\gamma)^{2\delta}\right) - 2n_{\delta,\upsilon}^2(1-\gamma)^{\delta}(1+\gamma)^{\delta}}}\psi \\ & & \times\, {}_h\sigma_{hk}^{\delta}\left[\left|\frac{{}_h u_{hk}}{\sqrt{h}}\right|^{\delta}(1-\gamma s_k)^{\delta} - E\left(\left|\frac{{}_h u_{hk}}{\sqrt{h}}\right|^{\delta}(1-\gamma s_k)^{\delta}\right)\right] \end{array}\right. \tag{3.2}$$

then model (3.1) can be seen as embedded in $\{{}_h r_{hk}, {}_h \sigma_{hk}^{\delta}\}_{k=0,1,\ldots}$ (namely for $h \equiv 1$), and yet $\{{}_h r_{hk}, {}_h \sigma_{hk}^{\delta}\}_{k=0,1,\ldots}$ converges weakly (as $h \downarrow 0$) to the solution of the EDGP whenever the following conditions hold true:

$$\begin{array}{rcll} \lim_{h\downarrow 0} h^{-1}\iota_h & = & \iota & < \infty \\ \lim_{h\downarrow 0} h^{-1}\theta_h & = & \theta & < \infty \\ \lim_{h\downarrow 0} h^{-1}\omega_h & = & \overline{\omega} & < \infty \\ \lim_{h\downarrow 0} h^{-1}(n_{\delta,\upsilon}((1-\gamma)^{\delta} + (1+\gamma)^{\delta})\alpha_h + \beta_h - 1) & = & -\varphi & < \infty \end{array}$$

and

$$\lim_{h\downarrow 0} h^{-1} \cdot \left[(m_{\delta,\upsilon} - n_{\delta,\upsilon}^2)((1-\gamma)^{2\delta} + (1+\gamma)^{2\delta}) - 2n_{\delta,\upsilon}^2(1-\gamma)^{\delta}(1+\gamma)^{\delta}\right] \cdot \alpha_h^2 = \psi^2 < \infty,$$

with $\overline{\omega}, \psi^2 > 0$; see thm. 3.1 and subsection 4.2 in FM (1998).

Model (3.1) is thus *one* possible discrete-time counterpart of the EDGP, and now we use it as an auxiliary model in an indirect inference scheme with simulations drawn from (3.2). Let b the vector containing the first five elements of b^+. Consider simulating system (3.2) for small h.[3] This is accomplished by setting γ, δ to their ML estimates $\widehat{\gamma}, \widehat{\delta}$, assigning values to $a = (\iota, \theta, \overline{\omega}, \varphi, \psi)$, and drawing $\frac{{}_h u_{hk}}{\sqrt{h}}$ from the normal distribution; one eventually obtains ${}_{h,h}\widetilde{r}^{(s)}(a) = \{{}_h\widetilde{r}_{hk}^{(s)}(a)\}_{k=0}^{N/h}$, $s = 1, \ldots, S$, where S is the number of simulations. For each simulation, just retain the (N) numbers ${}_h\widetilde{r}_{hk}^{(s)}(a)$ that correspond to integer indexes of time, and estimate the auxiliary model on each series of simulated data:

$$\widehat{b}_{N,s}^{(h)}(a) = \arg\max_b \mathfrak{L}_N({}_{1,h}\widetilde{r}^{(s)}(a); b), \quad s = 1, \ldots, S, \tag{3.3}$$

where ${}_{1,h}\widetilde{r}^{(s)}(.)$ denotes the set of the interest rate with integer indexes of time at simulation s and interval h. If $\|.\|$ is a norm, the indirect estimator of a is then:

$${}_h\widehat{a}_N(a_0) = \arg\min_a \left\|\widehat{b}_N - \frac{1}{S}\sum_{s=1}^{S}\widehat{b}_{N,s}^{(h)}(a)\right\|.$$

[3] Broze, Scaillet and Zakoïan (1995) give some suggestions concerning the choice of h as a function of N.

Asymptotics for ${}_h\widehat{a}_N(a_0)$ can be obtained by adapting the arguments in Gouriéroux, Monfort and Renault (1993). One has:

$$\sqrt{N}\left[{}_h\widehat{a}_N(a_0) - a_0\right] \underset{N\uparrow\infty,\ h\downarrow 0}{\overset{d}{\rightarrow}} \mathrm{N}\left(0, \frac{S+1}{S}(V_0'V_0)^{-1}V_0'\Gamma_0 V_0(V_0'V_0)^{-1}\right), \tag{3.4}$$

where $\Gamma_0 \equiv \lim_{h\downarrow 0}\Gamma_0^{(h)}$ and $V_0 \equiv \lim_{h\downarrow 0} V_0^{(h)}$, with $\Gamma_0^{(h)}$ and $V_0^{(h)}$ defined in appendix B. Broze, Scaillet and Zakoïan (1995) (BSZ) proved the preceding result in great generality—i.e. in the case of a general diffusion in $\mathbb{R}^l$—, and to avoid bias due to the discretization step used during the simulations, the authors also suggested to take $h = N^{-d}$ with $d > \frac{1}{2}$. Notice further that our specific problem is just-identified ($\dim(a) = \dim(b)$), and a is the solution of the following five-dimensional system: $\widehat{b}_N = \frac{1}{S}\sum_{s=1}^{S}\widehat{b}_{N,s}^{(h)}(a)$. This also implies a simplification of the variance in (3.4) that will be discussed in the following section. In appendix B, we check quite easily that the conditions of BSZ ensuring (3.4) hold for the scheme proposed here.

Is an auxiliary ARCH-based criterion the only device to achieve consistent estimation of a_0? Certainly not. The following diagram illustrates the situation. It conveys the main arguments that have to be used to show (3.4).

$$\begin{array}{ccc}
\mathfrak{L}_N({}_1\widetilde{r}(a_0); b^+) & \overset{h\downarrow 0}{\Longrightarrow} & \mathfrak{L}_N({}_1r; b^+) \\
\downarrow^{N\uparrow\infty} & & \downarrow^{N\uparrow\infty} \\
\mathfrak{L}_\infty^{(h)}(a_0; b^+) & & \mathfrak{L}_\infty(a_0; b^+)
\end{array}$$

Convergence of the criterion

Suppose for instance that $\mathfrak{L}_N(.)$ corresponds to the exact likelihood that is associated with, say, an ARMA representation applied to the squared observations and the simulated data. If the solution of the approximating scheme used for the simulations converges weakly to the solution of the EDGP, then, under suitable conditions given in appendix B, one has that $\mathfrak{L}_N({}_1\widetilde{r}(a_0); b^+) \Rightarrow \mathfrak{L}_N({}_1r; b^+)$: this is so because the observation set ${}_1r$ is assumed to have been generated by the EDGP. Consistency of the indirect estimator (i.e. for small N^{-1} and h) based on the auxiliary ARMA now follows from an argument similar to the one presented in appendix B.

While there is not a theory concerning the optimal choice of the criterion, however, one would like to require that the (already misspecified) auxiliary model fulfils some basic properties. Let M_1 be a candidate auxiliary model. One property of M_1 should be that it can be embedded into another model M_h, say, the solution of which converges in distribution towards the EDGP as $h \downarrow 0$.[4] Such a choice is the most natural one, and indeed is the one that is suggested in the literature; see, for instance, BSZ, and Gouriéroux and Monfort (1996, p.119-133). In view of the convergence results in FM (1998), choosing (3.1) as an auxiliary device for estimating the parameters in (2.1) is in line with such a principle. Finally, one can consider the case in which M_h is also used as the high frequency simulation generator. This, also, appears as a reasonable choice, and is suggested in the references above, too. Notice further that this case exactly corresponds to the strategy that is being considered here.

Such a strategy has an "approximation of the likelihood function" flavor—which is simply the likelihood of M_1—: as such, it constitutes an automatic correction for the asymptotic biais implied by the approximation, as originally stressed by Gouriéroux, Monfort and Renault (1993, p.S108). Finally, it is more likely that properties that are sufficient for (3.4) to hold—such as the convergence of $\mathfrak{L}_N({}_1\widetilde{r}(a_0); b^+)$ to $\mathfrak{L}_\infty^{(h)}(a_0; b^+)$, uniformly in $b^+ \in B$; or the continuity of the partial application $a \mapsto \widehat{b}_{N,s}^{(h)}(a)$—are fulfilled in cases where $\mathfrak{L}_N(.)$ applies to discrete-time counterparts of the EDGP that are embedded in the high frequency simulation generator—as for (3.1) and (3.2)—, rather than in cases in which the criterion does not even fulfill such a requirement. However, we are unable to show that such circumstances hold in great generality.

4 Results

As noted in the previous section, we have a just-identified problem, which means that V_0 is square and invertible. (3.4) thus simplifies to:

$$\sqrt{N}\left[{}_h\widehat{a}_N(a_0) - a_0\right] \underset{N\uparrow\infty,\ h\downarrow 0}{\overset{d}{\to}} N\left(0, \frac{S+1}{S} V_0^{-1}\Gamma_0 V_0'^{-1}\right).$$

In practice, we fix $h^{-1} \equiv 81$ and use $444,905$ discretization points. With an observations set of $N = 5505$, this means in fact that $h = N^{-\delta}$, where $\delta \simeq 0.5102 > \frac{1}{2}$. Finally, with an estimate of a^+ at hand, we obtain estimates of (λ_1, λ_2) in (2.2) by calibrating (λ_1, λ_2) to an observed, target term-structure. Such a procedure has been followed recently by Aït-Sahalia (1996b).

We are currently working on applying the entire procedure to real time series. We are making use of the Aït-Sahalia (1996a, b) series (that has 5505 daily observations of a weekly interest rate). We give some succinct results concerning some experiments performed with ML figures (see table 1), and $\lambda_1 = -0.8$ and $\lambda_2 = 1$. Figure 1 depicts the term-structure in the case $r = 8\%$. The solution has been obtained by numerically

[4] While it is natural to consider Euler approximations in the case of ARCH models, one can use more general discretization schemes such as the Mil'shtein's (1976) scheme in other circumstances.

Figure 1: From bottom to top, the curves correspond to values of σ^{δ} equal to, 0.0011, 0.0022, 0.0035, 0.0050, 0.0067, 0.0086, 0.0108, 0.0133, 0.0164 and 0.0200. The average level of σ^{δ} was 0.005644.

integrating the PDE (2.3), along the lines of appendix A.Our model predicts that the term-structure sharpens when volatility increases. The result is robust to alternative realistic values of (λ_1, λ_2).

5 Conclusion

Appendix A: numerical integration of the PDE (2.3)

The method that we followed is based on the Crank-Nicholson scheme. We first discuss the case in which the state-space is compact, and then show how our framework is to be embedded in it.

THE COMPACT STATE-SPACE CASE. Suppose we are given the following partial differential equation: for $(x, y, t) \in \mathcal{O}_1 \times \mathcal{O}_2 \times [0, T)$

$$0 = f_t + a(x,y) f_x + b(x,y) f_y + c(x,y) f_{xx} + d(x,y) f_{yy} + e(x,y) f_{xy} - R(x) f, \qquad \text{(A1)}$$

where $f \equiv f(x, y, t)$, a, b, c, d, e, R satisfy the regularity conditions in FM (1998), and $\mathcal{O}_i$ ($i = 1, 2$) are compacta of $\mathbb{R}_{++}$. The boundary condition we consider is $f(x, y, T) = 1$ $\forall (x, y) \in \mathcal{O}_1 \times \mathcal{O}_2$. We approximate the derivatives involved in the preceding equation via explicit and implicit approximations. We chop the state-space into a $N \times N$ grid, and time into J units:

$$\{(x_i, y_\ell)_{i,\ell=1}^N, (t_j)_{j=1}^J\} \subset \mathcal{O}_1 \times \mathcal{O}_2 \times [0, T];$$

here $x_i - x_{i-1} = \Delta x$, $y_\ell - x_{\ell-1} = \Delta y$, $t_j - t_{j-1} = \Delta t$, with $t_1 = 0$, $t_J = T$ and $N = \frac{x_N - x_1}{\Delta x} = \frac{y_N - y_1}{\Delta y}$, where $|x_N - x_1| < \infty$, and similarly for y. We define the approximation:

$$F_{i,\ell,j} \simeq f(x_i, y_\ell, t_j)$$

(and similarly for a, b, c, d, e, R), consider the "primitive" approximations:

$$\begin{array}{lll}
(f_x)_1 & \simeq \frac{F_{i+1,\ell,j} - F_{i-1,\ell,j}}{2\Delta x} & \text{(explicit at } j) \\
(f_x)_2 & \simeq \frac{F_{i+1,\ell,j+1} - F_{i-1,\ell,j+1}}{2\Delta x} & \text{(implicit at } j+1) \\
(f_y)_1 & \simeq \frac{F_{i,\ell+1,j} - F_{i,\ell-1,j}}{2\Delta y} & \text{(explicit at } j) \\
(f_y)_2 & \simeq \frac{F_{i,\ell+1,j+1} - F_{i,\ell-1,j+1}}{2\Delta y} & \text{(implicit at } j+1) \\
(f_{xx})_1 & \simeq \frac{F_{i+1,\ell,j} - 2F_{i,\ell,j} + F_{i-1,\ell,j}}{(\Delta x)^2} & \text{(explicit at } j) \\
(f_{xx})_2 & \simeq \frac{F_{i+1,\ell,j+1} - 2F_{i,\ell,j+1} + F_{i-1,\ell,j+1}}{(\Delta x)^2} & \text{(implicit at } j+1) \\
(f_{yy})_1 & \simeq \frac{F_{i,\ell+1,j} - 2F_{i,\ell,j} + F_{i,\ell-1,j}}{(\Delta y)^2} & \text{(explicit at } j) \\
(f_{yy})_2 & \simeq \frac{F_{i,\ell+1,j+1} - 2F_{i,\ell,j+1} + F_{i,\ell-1,j+1}}{(\Delta y)^2} & \text{(implicit at } j+1) \\
(f_{xy})_1 & \simeq \frac{F_{i+1,\ell+1,j} - F_{i+1,\ell-1,j} - F_{i-1,\ell+1,j} + F_{i-1,\ell-1,j}}{4\Delta x \cdot \Delta y} & \text{(explicit at } j) \\
(f_{xy})_2 & \simeq \frac{F_{i+1,\ell+1,j+1} - F_{i+1,\ell-1,j+1} - F_{i-1,\ell+1,j+1} + F_{i-1,\ell-1,j+1}}{4\Delta x \cdot \Delta y} & \text{(implicit at } j+1)
\end{array}$$

and construct the following estimates:

$$\begin{array}{l}
\widehat{f} \equiv F_{i,\ell,j} \\
\widehat{f}_t \equiv \frac{F_{i,\ell,j+1} - F_{i,\ell,j}}{\Delta t}
\end{array}$$

$\widehat{f_x} \equiv \frac{1}{2}\sum_{i=1}^{2}(f_x)_i$
$\widehat{f_y} \equiv \frac{1}{2}\sum_{i=1}^{2}(f_y)_i$
$\widehat{f_{xx}} \equiv \frac{1}{2}\sum_{i=1}^{2}(f_{xx})_i$
$\widehat{f_{yy}} \equiv \frac{1}{2}\sum_{i=1}^{2}(f_{yy})_i$
$\widehat{f_{xy}} \equiv \frac{1}{2}\sum_{i=1}^{2}(f_{xy})_i$

We plug the preceding estimates into eq.(A1) and obtain:

$$\begin{aligned}
&\alpha_{i\ell}^{(1)}F_{i-1,\ell-1,j}+\alpha_{i\ell}^{(2)}F_{i,\ell-1,j}+\alpha_{i\ell}^{(3)}F_{i+1,\ell-1,j}+\alpha_{i\ell}^{(4)}F_{i-1,\ell,j}+\widetilde{\alpha}_{i\ell}^{(5)}F_{i,\ell,j}\\
&+\alpha_{i\ell}^{(6)}F_{i+1,\ell,j}+\alpha_{i\ell}^{(7)}F_{i-1,\ell+1,j}+\alpha_{i\ell}^{(8)}F_{i,\ell+1,j}+\alpha_{i\ell}^{(9)}F_{i+1,\ell+1,j}\\
=&\\
&-[\alpha_{i\ell}^{(1)}F_{i-1,\ell-1,j+1}+\alpha_{i\ell}^{(2)}F_{i,\ell-1,j+1}+\alpha_{i\ell}^{(3)}F_{i+1,\ell-1,j+1}+\alpha_{i\ell}^{(4)}F_{i-1,\ell,j+1}+\alpha_{i\ell}^{(5)}F_{i,\ell,j+1}\\
&+\alpha_{i\ell}^{(6)}F_{i+1,\ell,j+1}+\alpha_{i\ell}^{(7)}F_{i-1,\ell+1,j+1}+\alpha_{i\ell}^{(8)}F_{i,\ell+1,j+1}+\alpha_{i\ell}^{(9)}F_{i+1,\ell+1,j+1}]
\end{aligned}\tag{A2}$$

where

$\alpha_{i\ell}^{(1)} \equiv \frac{e_{i\ell}}{8\Delta x\cdot\Delta y}$
$\alpha_{i\ell}^{(2)} \equiv -\frac{b_{i\ell}}{4\Delta y}+\frac{d_{i\ell}}{2(\Delta y)^2}$
$\alpha_{i\ell}^{(3)} \equiv -\frac{e_{i\ell}}{8\Delta x\cdot\Delta y}$
$\alpha_{i\ell}^{(4)} \equiv -\frac{a_{i\ell}}{4\Delta x}+\frac{c_{i\ell}}{2(\Delta x)^2}$
$\widetilde{\alpha}_{i\ell}^{(5)} \equiv -\frac{1}{\Delta t}-\frac{c_{i\ell}}{(\Delta x)^2}-\frac{d_{i\ell}}{(\Delta y)^2}-R_i$
$\alpha_{i\ell}^{(5)} \equiv \frac{1}{\Delta t}-\frac{c_{i\ell}}{(\Delta x)^2}-\frac{d_{i\ell}}{(\Delta y)^2}$
$\alpha_{i\ell}^{(6)} \equiv \frac{a_{i\ell}}{4\Delta x}+\frac{c_{i\ell}}{2(\Delta x)^2}$
$\alpha_{i\ell}^{(7)} \equiv -\frac{e_{i\ell}}{8\Delta x\cdot\Delta y}$
$\alpha_{i\ell}^{(8)} \equiv \frac{b_{i\ell}}{4\Delta y}+\frac{d_{i\ell}}{2(\Delta y)^2}$
$\alpha_{i\ell}^{(9)} \equiv \frac{e_{i\ell}}{8\Delta x\cdot\Delta y}$

Next, we let

$$F_j=(F_{\cdot,1,j},F_{\cdot,2,j},...,F_{\cdot,N-1,j},F_{\cdot,N,j})'$$

$$F_{\cdot,\ell,j}=(F_{1,\ell,j},F_{2,\ell,j},...,F_{N-1,\ell,j},F_{N,\ell,j})\ \ (\ell=1,...,N).$$

Starting from the boundary condition[5]

$$F_J=\mathbf{1}_{N^2\times 1}$$

(with $\mathbf{1}_{N^2\times 1}$ being a vector of N^2 ones), eq. (A2) can be solved by backward iterating the following equation

$$F_j=L\cdot F_{j+1}\quad (j=J-1,...,1),$$

[5]Below we modify the boundary condition in order to take account of transversality conditions.

where

$$L = -\widetilde{A}^{-1} \cdot A,$$

and $\widetilde{A}, A$ are block tridiagonal matrices:

$$\widetilde{A} = \begin{pmatrix} \widetilde{A}_{11} & A_{12} & \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & ... & ... & ... & ... & \mathbf{0} & \mathbf{0} \\ A_{21} & \widetilde{A}_{22} & A_{23} & \mathbf{0} & ... & \mathbf{0} & ... & ... & ... & ... & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_{32} & \widetilde{A}_{33} & A_{34} & ... & \mathbf{0} & ... & ... & ... & ... & \mathbf{0} & \mathbf{0} \\ ... & ... & ... & ... & \ddots & ... & \ddots & ... & ... & ... & ... & ... \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & ... & A_{\ell,\ell-1} & \widetilde{A}_{\ell\ell} & A_{\ell,\ell+1} & ... & ... & ... & \mathbf{0} \\ ... & ... & ... & ... & \ddots & ... & ... & ... & \ddots & ... & ... & ... \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & ... & ... & ... & ... & ... & A_{N-1,N-2} & \widetilde{A}_{N-1,N-1} & A_{N-1,N} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & ... & ... & ... & ... & ... & \mathbf{0} & A_{N,N-1} & \widetilde{A}_{NN} \end{pmatrix},$$

$$A = \begin{pmatrix} A_{11} & A_{12} & \mathbf{0} & \mathbf{0} & ... & \mathbf{0} & ... & ... & ... & ... & \mathbf{0} & \mathbf{0} \\ A_{21} & A_{22} & A_{23} & \mathbf{0} & ... & \mathbf{0} & ... & ... & ... & ... & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_{32} & A_{33} & A_{34} & ... & \mathbf{0} & ... & ... & ... & ... & \mathbf{0} & \mathbf{0} \\ ... & ... & ... & ... & \ddots & ... & \ddots & ... & ... & ... & ... & ... \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & ... & A_{\ell,\ell-1} & A_{\ell\ell} & A_{\ell,\ell+1} & ... & ... & ... & \mathbf{0} \\ ... & ... & ... & ... & \ddots & ... & ... & ... & \ddots & ... & ... & ... \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & ... & ... & ... & ... & ... & A_{N-1,N-2} & A_{N-1,N-1} & A_{N-1,N} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & ... & ... & ... & ... & ... & \mathbf{0} & A_{N,N-1} & A_{NN} \end{pmatrix},$$

where $\mathbf{0}$ are $N \times N$ matrices of zeros and (with blanks denoting zeros)

$$A_{\ell,\ell-1} = \begin{pmatrix} \alpha^{(2)}_{1\ell} & \alpha^{(3)}_{1\ell} & & & & & & \\ \alpha^{(1)}_{2\ell} & \alpha^{(2)}_{2\ell} & \alpha^{(3)}_{2\ell} & & & & & \\ & \alpha^{(1)}_{3\ell} & \alpha^{(2)}_{3\ell} & \alpha^{(3)}_{3\ell} & & & & \\ & & & & \ddots & & & \\ & & & & & \alpha^{(1)}_{N-1,\ell} & \alpha^{(2)}_{N-1,\ell} & \alpha^{(3)}_{N-1,\ell} \\ & & & & & & \alpha^{(1)}_{N\ell} & \alpha^{(2)}_{N\ell} \end{pmatrix},$$

$$\widetilde{A}_{\ell\ell} = \begin{pmatrix} \widetilde{\alpha}^{(5)}_{1\ell} & \alpha^{(6)}_{1\ell} & & & & & \\ \alpha^{(4)}_{2\ell} & \widetilde{\alpha}^{(5)}_{2\ell} & \alpha^{(6)}_{2\ell} & & & & \\ & \alpha^{(4)}_{3\ell} & \widetilde{\alpha}^{(5)}_{3\ell} & \alpha^{(6)}_{3\ell} & & & \\ & & & \ddots & & & \\ & & & & \alpha^{(4)}_{N-1,\ell} & \widetilde{\alpha}^{(5)}_{N-1,\ell} & \alpha^{(6)}_{N-1,\ell} \\ & & & & & \alpha^{(4)}_{N\ell} & \widetilde{\alpha}^{(5)}_{N\ell} \end{pmatrix},$$

and

$$A_{\ell,\ell+1} = \begin{pmatrix} \alpha^{(8)}_{1\ell} & \alpha^{(9)}_{1\ell} & & & & & \\ \alpha^{(7)}_{2\ell} & \alpha^{(8)}_{2\ell} & \alpha^{(9)}_{2\ell} & & & & \\ & \alpha^{(7)}_{3\ell} & \alpha^{(8)}_{3\ell} & \alpha^{(9)}_{3\ell} & & & \\ & & & \ddots & & & \\ & & & & \alpha^{(7)}_{N-1,\ell} & \alpha^{(8)}_{N-1,\ell} & \alpha^{(9)}_{N-1,\ell} \\ & & & & & \alpha^{(7)}_{N\ell} & \alpha^{(9)}_{N\ell} \end{pmatrix}.$$

Finally, $\widetilde{A}_{\ell\ell}$ differs from $A_{\ell\ell}$ in that the diagonal of $\widetilde{A}_{\ell\ell}$ is composed by $\widetilde{\alpha}^{(5)}_{i,\ell}$ whereas the diagonal of $A_{\ell\ell}$ is composed by $\alpha^{(5)}_{i,\ell}$.

The final step consists in deriving limiting as well as transversality conditions that eventually place restrictions on the matrices $\widetilde{A}$ and A. This depends on the specific problem at hand: see subsect.? below for the application to our problem.

The general case. When the state-space is not as in eq. (A1)—as it usually happens in finance—, the implementation of the algorithm can only be done after a previous transformation of the original state-space. In the PDE (2.3), for instance, we shall introduce two new functions of (r, σ^δ) that take values on the compact $\mathcal{O}_1 \times \mathcal{O}_2$. A convenient choice is to set $\mathcal{O}_1 \times \mathcal{O}_2 = [0,1]^2$. Then we define $v \equiv \sigma^\delta$ and

$$\begin{array}{rcl} x(r) & = & \frac{\gamma r}{1+\gamma r}, \quad \gamma > 0 \\ y(v) & = & \frac{\beta v}{1+\beta v}, \quad \beta > 0 \end{array} \tag{A3}$$

and write

$$f(x,y,t) \equiv f(x(r), y(v), t) = B(r,v,t).$$

In terms of (f, x, y) eq.(2.3) can be expressed in exactly the same format as eq. (A1), with

$$\begin{array}{l} a(x,y) \equiv \left[\iota - \theta R(x) - R(x)V(y)^{\frac{1}{\delta}}\lambda_1\right] \frac{\gamma}{(1+\gamma R(x))^2} - R(x)V(y)^{\frac{2}{\delta}} \frac{\gamma^2}{(1+\gamma R(x))^3} \\ b(x,y) \equiv \left[\overline{w} - \varphi V(y) - \psi V(y)\sqrt{R(x)}(\rho\lambda_1 + \sqrt{1-\rho^2}\lambda_2)\right] \frac{\beta}{(1+\beta V(y))^2} - \psi^2 V(y)^2 \frac{\beta^2}{(1+\beta V(y))^3} \\ c(x,y) \equiv \frac{1}{2} R(x)V(y)^{\frac{2}{\delta}} \frac{\gamma^2}{(1+\gamma R(x))^4} \\ d(x,y) \equiv \frac{1}{2}\psi^2 V(y)^2 \frac{\beta^2}{(1+\beta V(y))^4} \end{array}$$

$$e(x,y) \equiv \psi\rho\sqrt{R(x)}V(y)^{\frac{\delta+1}{\delta}}\frac{\gamma\beta}{(1+\gamma R(x))^2(1+\beta V(y))^2}$$

where $(R,V)(.,.)$ is the inversion of (A3):

$$\begin{aligned} R(x) &= \tfrac{x}{\gamma(1-x)} \\ V(y) &= \tfrac{y}{\beta(1-y)} \end{aligned}$$

Limiting and transversality conditions. Finally we impose two kinds of conditions. The first kind of conditions concerns the limiting behavior of the PDE (2.3) when $R(x)=0, R(x)=\infty, V(y)=0, V(y)=\infty$. The second kind of conditions follows from a transversality argument, and stipulates that $\lim_{r\to\infty} B(r,v,t)=0\ \forall(v,t)\in\mathbb{R}_+\times[0,T]$. We call the first kind of conditions "limiting conditions" and the second kind of conditions "transversality conditions".

To find the restrictions on the coefficients of $\widetilde{A}$ and A that correspond to $x=0$ and $y=0$, notice that, for each $t\in[0,T)$,

$$\begin{aligned} 0 &= f_t(0,y,t)+a(0,y)f_x(0,y,t)+b(0,y)f_y(0,y,t)+d(0,y)f_{yy}(0,y,t) \quad &(\text{for } x=0) \\ 0 &= f_t(x,0,t)+a(x,0)f_x(x,0,t)+b(x,0)f_y(x,0,t)-R(x)f(x,0,t) \quad &(\text{for } y=0) \end{aligned}$$

By plugging the following asymmetric approximations in the preceding equations,

$$\begin{aligned} \widehat{f_t}(0,y,t) &\equiv \tfrac{F_{1,\ell,j+1}-F_{1,\ell,j}}{\Delta t} \\ \widehat{f_t}(x,0,t) &\equiv \tfrac{F_{i,1,j+1}-F_{i,1,j}}{\Delta t} \\ \widehat{f_x}(0,y,t) &\equiv \tfrac{F_{2,\ell,j}-F_{1,\ell,j}}{\Delta x} \\ \widehat{f_x}(x,0,t) &\equiv \tfrac{F_{i+1,1,j}-F_{i,1,j}}{\Delta x} \\ \widehat{f_y}(0,y,t) &\equiv \tfrac{F_{1,\ell+1,j}-F_{1,\ell,j}}{\Delta y} \\ \widehat{f_y}(x,0,t) &\equiv \tfrac{F_{i,2,j}-F_{i,1,j}}{\Delta y} \\ \widehat{f_{yy}}(0,y,t) &\equiv \tfrac{F_{1,\ell+1,j}-2F_{1,\ell,j}+F_{1,\ell-1,j}}{(\Delta y)^2} \end{aligned}$$

we get the following difference equations:

$$\begin{aligned} \gamma_{1\ell}^{(1)}F_{1,\ell-1,j}+\gamma_{1\ell}^{(2)}F_{1,\ell,j}+\gamma_{1\ell}^{(3)}F_{2,\ell,j}+\gamma_{1\ell}^{(4)}F_{1,\ell+1,j} &= -\tfrac{1}{\Delta t}F_{1,\ell,j+1} \\ \eta_{i1}^{(1)}F_{i,1,j}+\eta_{i1}^{(2)}F_{i,2,j}+\eta_{i1}^{(3)}F_{i+1,1,j} &= -\tfrac{1}{\Delta t}F_{i,1,j+1} \end{aligned}$$

where

$$\begin{aligned} \gamma_{1\ell}^{(1)} &\equiv \tfrac{d_{1\ell}}{(\Delta y)^2} \\ \gamma_{1\ell}^{(2)} &\equiv -\tfrac{1}{\Delta t}-\tfrac{a_{1\ell}}{\Delta x}-\tfrac{b_{1\ell}}{\Delta y}-\tfrac{2d_{1\ell}}{(\Delta y)^2} \\ \gamma_{1\ell}^{(3)} &\equiv \tfrac{a_{1\ell}}{\Delta x} \\ \gamma_{1\ell}^{(4)} &\equiv \tfrac{b_{1\ell}}{\Delta y}+\tfrac{d_{1\ell}}{(\Delta y)^2} \\ \eta_{i1}^{(1)} &\equiv -\tfrac{1}{\Delta t}-\tfrac{a_{i1}}{\Delta x}-\tfrac{b_{i1}}{\Delta y}-R_i \end{aligned}$$

$\eta_{i1}^{(2)} \equiv \frac{b_{i1}}{\Delta y}$
$\eta_{i1}^{(3)} \equiv \frac{a_{i1}}{\Delta x}$

The matrices $\widetilde{A}$ and A must thus be constrained so that the elements $\alpha_{1\ell}^{(2)}$, $\alpha_{1\ell}^{(3)}$, $\widetilde{\alpha}_{1\ell}^{(5)}$, $\alpha_{1\ell}^{(5)}$, $\alpha_{1\ell}^{(6)}$, $\alpha_{1\ell}^{(8)}$, $\alpha_{1\ell}^{(9)}$, $\alpha_{i1}^{(4)}$, $\widetilde{\alpha}_{i1}^{(5)}$, $\alpha_{i1}^{(5)}$, $\alpha_{i1}^{(6)}$, $\alpha_{i1}^{(7)}$, $\alpha_{i1}^{(8)}$ and $\alpha_{i1}^{(9)}$ enter as $\alpha_{1\ell}^{(2)} = \gamma_{1\ell}^{(1)}$, $\alpha_{1\ell}^{(3)} = 0$, $\widetilde{\alpha}_{1\ell}^{(5)} = \gamma_{1\ell}^{(2)}$, $\alpha_{1\ell}^{(6)} = \gamma_{1\ell}^{(3)}$, $\alpha_{1\ell}^{(8)} = \gamma_{1\ell}^{(4)}$, $\alpha_{1\ell}^{(9)} = 0$, $\alpha_{i1}^{(4)} = 0$, $\widetilde{\alpha}_{i1}^{(5)} = \eta_{i1}^{(1)}$, $\alpha_{i1}^{(6)} = \eta_{i1}^{(3)}$, $\alpha_{i1}^{(7)} = 0$, $\alpha_{i1}^{(8)} = \eta_{i1}^{(2)}$, $\alpha_{i1}^{(9)} = 0$ in matrix $\widetilde{A}$, and as $\alpha_{1\ell}^{(2)} = 0$, $\alpha_{1\ell}^{(3)} = 0$, $\alpha_{1\ell}^{(5)} = \frac{1}{\Delta t}$, $\alpha_{1\ell}^{(6)} = 0$, $\alpha_{1\ell}^{(8)} = 0$, $\alpha_{1\ell}^{(9)} = 0$, $\alpha_{i1}^{(4)} = 0$, $\alpha_{i1}^{(5)} = \frac{1}{\Delta t}$, $\alpha_{i1}^{(6)} = 0$, $\alpha_{i1}^{(7)} = 0$, $\alpha_{i1}^{(8)} = 0$, $\alpha_{i1}^{(9)} = 0$ in matrix A.

We derive and impose similar restrictions in the cases $x = 1$ and $y = 1$. Such restrictions were also the result of the transversality condition concerning the behavior of the price at $x = 1$: $F_{N,\ell,j} = 0$ ($\ell = 1, ..., N$, $j = 1, ..., J$) which implies starting with $F_J = \widetilde{\mathbf{1}}_{N^2 \times 1}$, where $\widetilde{\mathbf{1}}_{N^2 \times 1}$ is as $\mathbf{1}_{N^2 \times 1}$ with the exception the a zero replaces the one at positions $N, 2N, ..., N^2$.

Appendix B: standard regularity conditions and the convergence of the criterion

Assumption B1.

- $\text{plim}_N\, \mathfrak{L}_N({}_1r; b^+) = \mathfrak{L}_\infty(a_0; b^+)$, say, uniformly in $b^+ \in B$.
- $\text{plim}_N\, \frac{\partial^2 \; N}{\partial b^+ \partial b^{+\prime}}({}_1r; b^+) = \ddot{\mathfrak{L}}_\infty\,(a_0; b^+)$, say, uniformly in $b^+ \in B$. Further, $\ddot{\mathfrak{L}}_\infty\,(.)$ is invertible.
- $\left[\sqrt{T} \frac{\partial \; N}{\partial b^+}({}_1r; b^+)\right]_{b^+ = b_0^+(a_0)} \xrightarrow{d} \text{N}(0, J(a_0))$, where N(.) is a standard normal variable.

Convergence of the criterion (Sketch). We assume as in BSZ the continuity of the partial application $a \mapsto \widehat{b}_{N,s}^{(h)}(a)$, and for the case $S = 1$, we define $\widehat{b}_N^{(h)}(.) \equiv \widehat{b}_{N,1}^{(h)}(.)$ and ${}_1\widetilde{r}(.) \equiv \;{}_{1,h}\widetilde{r}^{(1)}(.)$. It is not hard to show that under conditions on $\mathfrak{L}_N({}_1\widetilde{r}(a); b)$ that parallel those in assumption B1 stated above for the direct criterion $\mathfrak{L}_N({}_1r; b^+)$, the simulated estimator in (3.3) is asymptotic normal:

$$\sqrt{N}\left[\widehat{b}_N^{(h)}(a) - b_0^{(h)}(a)\right] \xrightarrow{d} \text{N}\left(0, \ddot{\mathfrak{L}}_\infty^{(h)-1}\,(a; b_0^{(h)}(a)) \cdot J^{(h)}(a) \cdot \ddot{\mathfrak{L}}_\infty^{(h)-1}\,(a; b_0^{(h)}(a))\right),$$

where $b_0^{(h)}(a) = \arg\max_b \mathfrak{L}_\infty^{(h)}(a; b)$, the limit simulation problem, and $\ddot{\mathfrak{L}}_\infty^{(h)}\,(.)$ and $J^{(h)}(.)$ are defined similarly as $\ddot{\mathfrak{L}}_\infty\,(.)$ and $J(.)$. Now, it follows by FM (1997, 1998) and FM (1994) that the solution of (3.2): $\{{}_hr_{hk}, {}_h\sigma_{hk}^\delta\}_{k=0,1,...} \Rightarrow \{r_t, \sigma_t^\delta\}_{t \geq 0}$ (the solution of the EDGP); see thm 3.1 and section 4.2 in FM (1998). By this, an extension of a result in FM

(1994) that shows that the solution of (3.2) is unique, stationary and ergodic (for fixed h), and the uniform continuity of the criterion $\mathfrak{L}_N(.;b)$ (?), it follows that $\mathfrak{L}_N({}_1\widetilde{r}(a_0);b^+) \Rightarrow \mathfrak{L}_N({}_1r;b^+)$ as $h \downarrow 0$, and we suppose, as in BSZ, that the convergence is uniform in b^+. Finally, because $\text{plim}_N\, \mathfrak{L}_N({}_1\widetilde{r}(a);b^+) = \mathfrak{L}_\infty^{(h)}(a;b^+)$ and $\text{plim}_N\, \mathfrak{L}_N({}_1r;b^+) = \mathfrak{L}_\infty(a_0;b^+)$, uniformly in $b^+ \in B$ (both by assumption), one can easily verify that for small h, this implies $b_0^{(h)}(a_0) = \arg\max_b \mathfrak{L}_\infty^{(h)}(a_0;b) = \arg\max_b \mathfrak{L}_\infty(a_0;b) = b_0(a_0)$. This is:

$$\lim_{h\downarrow 0} b_0^{(h)}(a_0) = b_0(a_0), \tag{B1}$$

while for fixed h, it is assumed that there exists only one solution to the system $b_0^{(h)}(a) = b_0(a_0)$: this has the form $\mathcal{A}^{(h)}(a_0)$, with of course $\lim_{h\downarrow 0} \mathcal{A}^{(h)}(a_0) = a_0$. Now by proposition 6 in BSZ, one has that $\sqrt{N}\left[{}_h\widehat{a}_N(a_0) - \mathcal{A}^{(h)}(a_0)\right] \xrightarrow{d} \mathrm{N}\left(0, \Sigma^{(h)}\right)$ (for fixed h), where $\Sigma^{(h)}$ is such that $\lim_{h\downarrow 0} \Sigma^{(h)} = 2(V_0'V_0)^{-1}V_0'\Gamma_0 V_0(V_0'V_0)^{-1}$, and (3.4) follows for $S = 1$. In the preceding expressions, Γ_0 is defined as the limit of $\Gamma_0^{(h)}$ as $h \downarrow 0$, and similarly for V_0, and $\Gamma_0^{(h)}$ is the limit of $\ddot{\mathfrak{L}}_\infty^{(h)-1}(a;b_0^{(h)}(a)) \cdot J^{(h)}(a) \cdot \ddot{\mathfrak{L}}_\infty^{(h)-1}(a;b_0^{(h)}(a))$ as $N \uparrow \infty$, whereas $V_0^{(h)}$ is the limit of $\left[\frac{\partial \widehat{b}_N^{(h)}}{\partial a}(a)\right]_{a=\mathcal{A}^{(h)}(a_0)}$ as $N \uparrow \infty$. The case $S > 1$ is similar. *Q.E.D.*

References

Aït-Sahalia,Y. (1996*a*): "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, 385–426.

Aït-Sahalia,Y. (1996*b*): "Nonparametric Pricing of Interest Rate Derivative Securities," *Econometrica*, 64, 527–560.

Broze, L., O. Scaillet and J. Zakoïan (1995): "Quasi-Indirect Inference for Diffusion Processes," CORE DP 9505.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997): *The Econometrics of Financial Markets.* Princeton: Princeton University Press.

Cox, J., J. Ingersoll, and S. Ross (1985*a*): "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, 363-384.

Cox, J., J. Ingersoll, and S. Ross (1985*b*): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385-407.

Fornari, F. and A. Mele (1994): "A Two Factor Arbitrage Model with Optimal Filtering Behavior", *Statistica* 54, 293-312.

Fornari, F. and A. Mele (1995): "Continuous Time Conditionally Heteroskedastic Models: Theory with Applications to the Term Structure of Interest Rates", *Economic Notes* 25, 341-366.

Fornari, F. and A. Mele (1997): "Weak Convergence and Distributional Assumptions for a General Class of Nonlinear ARCH Models," *Econometric Reviews*, 16, 205-227.

Fornari, F. and A. Mele (1998): "ARCH Models and Option Pricing: the Continuous Time Connection," THEMA DP 9830.

Fornari, F. and A. Mele (1999): *Stochastic Volatility in Financial Markets*: *Statistical Theory, Empirical Evidence and Economic Implications*, forthcoming, Kluwer Academic Publishers.

Gouriéroux, C., A.Monfort and E.Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8, S85-S118.

Gouriéroux, C. and A.Monfort (1996): *Simulation-Based Econometric Methods*, Oxford: Oxford University Press.

Huang, C.F. (1987): "An Intertemporal General Equilibrium Asset Pricing Model: the Case of Diffusion Information," *Econometrica*, 55, 117-142.

Longstaff, F. and E. Schwartz (1992): "Interest-Rate Volatility and the Term Structure: A two Factor General Equilibrium Model," *Journal of Finance*, 47, 1259-1282.

Mil'shtein, G. (1976): "A Method of Second Order Accuracy Integration of Stochastic Differential Equations," *Theory of Probability and Applications*, 23, 396-401.

Wiggins, J. (1987): "Option Values Under Stochastic Volatility. Theory and Empirical Estimates," *Journal of Financial Economics*, 19, 351-372.

Tables

Table 1: Maximum likelihood and indirect estimates of eq. (?.?):
$\mathrm{d}r = (\iota - \theta r)\mathrm{d}t + \sigma\sqrt{r}\mathrm{d}W, \quad \mathrm{d}\sigma^\delta = (\overline{\omega} - \varphi\sigma^\delta)\mathrm{d}t + \psi\sigma^\delta{\cdot}\mathrm{d}(\rho W^{(1)} + \sqrt{1-\rho^2}W^{(2)})$

	ι	θ	$\overline{\omega}$	φ	ψ	δ	ρ	$\ln \ell$
ML	$2.386{\cdot}10^{-4}$	$2.859{\cdot}10^{-3}$	$5.139{\cdot}10^{-4}$	$5.799{\cdot}10^{-4}$	0.297	1.128	0.442	31048
Indirect						$\equiv 1.128$	$\equiv 0.442$	—