Econometric Theory, **14**, 1998, 161–186. Printed in the United States of America.

QUASI-INDIRECT INFERENCE FOR DIFFUSION PROCESSES

LAURENCE BROZE
Gremars, Université de Lille 3
and
CORE

OLIVIER SCAILLET
Université Catholique de Louvain

JEAN-MICHEL ZAKOÏAN
Université de Lille 1
and
CREST

We discuss an estimation procedure for continuous-time models based on discrete sampled data with a fixed unit of time between two consecutive observations. Because in general the conditional likelihood of the model cannot be derived, an indirect inference procedure following Gouriéroux, Monfort, and Renault (1993, *Journal of Applied Econometrics* 8, 85–118) is developed. It is based on simulations of a discretized model. We study the asymptotic properties of this "quasi"-indirect estimator and examine some particular cases. Because this method critically depends on simulations, we pay particular attention to the appropriate choice of the simulation step. Finally, finite-sample properties are studied through Monte Carlo experiments.

1. INTRODUCTION

In recent years, the theory of continuous-time stochastic processes has become an indispensable tool in modern finance. This is partly due to the elegance and tractability of the so-called Itô-calculus, which allows for very general transformations of the process dynamics. Unfortunately, it is well known that the inference associated with such continuous-time processes is a difficult problem because only discrete sampled data are available.

We thank Eric Ghysels, Christian Gouriéroux, Frédéric Jouneau, Boris Leblanc, Bernard Lejeune, Alain Monfort, George Tauchen, two anonymous referees, and the participants at ESWC 95 for useful comments. Laurence Broze gratefully acknowledges financial support from the Belgian program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. Olivier Scaillet gratefully acknowledges financial support from ICM, Belgium and the support of the grant Action de recherche concertées 93/98-162 of the Ministry of Scientific Research (Belgian French Speaking Community). He also acknowledges support from CREST, where part of this research was done. Address correspondence to: Laurence Broze, CORE, 34, voie du Roman Pays, B-1348 Louvain la Neuve, Belgium; e-mail: broze@core.ucl.ac.be.

© 1998 Cambridge University Press 0266-4666/98 $9.50

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Diffusion processes estimation is widely studied in the literature, but often availability of continuous data is assumed (see, e.g., Genon-Catalot and Jacod, 1993; Danesi, Genon-Catalot, and Laurent, 1994).

The case of discrete observations is treated by several authors. The maximum likelihood estimator has the standard properties of consistency and asymptotic normality under appropriate regularity conditions (see, e.g., Prasaka Rao, 1988). However, it is generally uncomputable except in some well-known cases where an exact discrete-time version is available (for GARCH discretization, see also Drost and Werker, 1996).

An explicit expression of the transition densities for one-dimensional diffusions was given by Dacunha-Castelle and Florens-Zmirou (1986) (for a generalization to multivariate diffusion, see Pedersen, 1995). Unfortunately it involves a stochastic integral with a Brownian bridge. Therefore the expression cannot be used directly to compute the log-likelihood.

In the general case, the transition densities are untractable, and a natural approach is to approximate the likelihood function based on a discretized version of the diffusion. A direct approximation leads to inconsistency unless the observation interval goes to zero (see Florens-Zmirou, 1989). In recent papers, Pedersen (1994, 1995) derived a sequence of approximations indexed by a discretization step smaller than the observation interval and controlled by the statistician. Consistency and asymptotic normality are proved when the sample size goes to infinity for a discretization step decreasing to zero sufficiently fast.

Several other approaches have also recently been proposed in the econometric literature to deal with discretely sampled data with fixed observation interval. Parameter estimation of continuous-time processes can be achieved through the use of moment conditions derived from properties of the infinitesimal generator (Hansen and Scheinkman, 1995). Some authors (Aït-Sahalia, 1992, 1996; Hansen, Scheinkman, and Touzi, 1994) have studied nonparametric procedures for diffusion processes that rely on kernel estimation of the stationary density.

Other methods consist of simulation based procedures, such as the simulated moment method (Duffie and Singleton, 1989) or the indirect inference (Gouriéroux et al., 1993). Starting with Lerman and Manski (1981), simulations have been used as a central point of estimation procedures. A large number of simulated estimation methods have been discussed in the econometric literature. Simulated moments estimators have been introduced by McFadden (1989) and by Pakes and Pollard (1989) in the independent and identically distributed (i.i.d.) case and have been extended to time series by Lee and Ingram (1991) and by Duffie and Singleton (1993). Simulated maximum likelihood has been studied by Van Praag and Hop (1987), Gouriéroux and Monfort (1992, 1993), and Laroque and Salanié (1993) (for a review, see Gouriéroux and Monfort, 1995).

Indirect inference procedures have been proposed by Smith (1990), Gallant and Tauchen (1996), and Gouriéroux et al. (1993). The indirect inference method needs simulations of the true model and thus has to be modified for diffusion processes as a result of their continuous-time nature. Indeed, the transition den-

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

sities are generally unknown for diffusion processes. A natural way of avoiding this difficulty is to discretize the continuous-time model and to simulate the discrete-time approximation. Unfortunately this approach induces a bias in the estimation. The aim of our paper is to derive asymptotic properties of the indirect estimator when the sample size goes to infinity, with a fixed observation step, and to study the bias induced by the simulation of an approximate discrete-time model instead of the continuous-time diffusion. This bias explains the terminology "quasi-indirect" inference. We also examine the problem of the choice of an appropriate simulation step, both from an asymptotic point of view and in small sample sizes through Monte Carlo experiments.

The paper is organized as follows. Section 2 presents the continuous-time model and the quasi-indirect inference method aimed at estimating its parameter. Asymptotic properties of this method are derived therein. For illustrative purposes, two examples will be covered: the geometric Brownian motion and the Ornstein–Uhlenbeck process. In Section 3, Monte Carlo simulations for these two processes allow us to compare the finite-sample properties of the different estimators involved in the method. Furthermore, we discuss the choice of an appropriate discretization step. Section 4 concludes. Most proofs and technical developments appear in the Appendixes.

2. THE QUASI-INDIRECT INFERENCE METHOD

Let us consider the following stochastic differential equation:

$$dY_t = \mu(Y_t;\theta)dt + \sigma(Y_t;\theta)dW_t, \tag{1}$$

where $(Y_t; t \in \mathbb{R}^+)$ is a $\mathbb{R}^d$-valued process, W is a standard r-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and parameter θ belongs to the compact parameter set $\Theta \subset \mathbb{R}^p$. The functions μ and σ from $\mathbb{R}^d \times \Theta$ to $\mathbb{R}^d$ and $\mathbb{R}^d \times \mathbb{R}^r$, respectively, are measurable, are differentiable with respect to (w.r.t.) θ, and are assumed to fulfill some Lipschitz assumptions (with Lipschitz constant independent of θ) ensuring existence and almost sure uniqueness of the solution of model (1) (see, e.g., Revuz and Yor, 1991). Therefore, the solution of (1) induces a probability measure $\mathbb{P}_\theta$ on the space of continuous functions from $\mathbb{R}^+$ to $\mathbb{R}^d$. We also assume that the diffusion term in (1) is nondegenerate and that the solution process of model (1) is stationary and ergodic. The initial value Y_0 is randomly drawn from the ergodic distribution $\mathcal{L}_0$.

As it usually occurs in practical situations, the whole path of the process is not available. We can only observe Y at T fixed, equally spaced, sampling times. Let τ be the sampling interval and

$$Y_{\underline{T}}^{\tau} = (Y_\tau, Y_{2\tau}, \ldots, Y_{T\tau})$$

the observation vector. The particular value of parameter θ in model (1), under which the observations have been generated, is equal to θ_0 assumed to be identifiable and in the interior of Θ. Hence, we associate with the unique solution $(Y_t;$

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

$t \in \mathbb{R}^+$) of equation (1) the process $(Y_{t\tau};\ t \in \mathbb{N})$ that inherits the Markovian structure of the former, as well as stationarity and ergodicity.

The application of indirect inference to diffusion processes was suggested by Gouriéroux et al. (1993). Indirect inference involves four successive steps:

1. maximization of a given criterion applied to the observation;
2. creation of simulated data sets of size T and time interval τ, using (1) as the data-generating process for all values of θ;
3. maximization of the same criterion using the simulated data;
4. calibration with respect to θ to minimize the distance between the outputs of steps 1 and 3.

However, step 2 cannot be achieved in general because the transition densities of the diffusion processes are not directly computable. A natural approach is then to use discretizations of (1), *with a fixed discretization step h smaller than the observation step* τ, as generating processes to obtain simulated data. Consequently, an asymptotic bias resulting from this fixed time interval in the simulations is involved. As we do not simulate the true model (i.e., the true probability density function [p.d.f.]), we call this method *quasi-indirect inference.*

In the following, the different steps are discussed one by one.

2.1. First Step

The first step consists of maximizing a criterion function, i.e., a function S from $\Omega \times \Theta$ to $\mathbb{R}$ of the form $S(\omega,\theta) = S_T(Y_{\underline{T}}^{\tau},\theta)$. Examples of such criteria are the log-likelihood functions, contrast functions (see Genon-Catalot, 1990), or more generally any function involving empirical moments.

Let $\hat{\theta}_T^*(Y_{\underline{T}}^{\tau})$ denote any measurable solution of the equation

$$\hat{\theta}_T^*(Y_{\underline{T}}^{\tau}) = \underset{\theta^* \in \Theta}{\operatorname{argmax}}\, S_T(Y_{\underline{T}}^{\tau},\theta^*). \tag{2}$$

Under standard regularity assumptions (see Appendix A), the direct estimator $\hat{\theta}_T^*(Y_{\underline{T}}^{\tau})$ is strongly consistent to

$$\theta_0^*(\theta_0) = \underset{\theta^* \in \Theta}{\operatorname{argmax}}\, S_\infty(\theta_0,\theta^*),$$

where $S_\infty(\cdot)$ denotes the limit criterion and it is asymptotically normal (see Appendix A). More precisely, we have

$$\sqrt{T}(\hat{\theta}_T^*(Y_{\underline{T}}^{\tau}) - \theta_0^*(\theta_0)) \xrightarrow{T\to\infty} \mathcal{N}(0,\Gamma_0^*),$$

where the matrix Γ_0^* involves derivatives of the limit criterion (the explicit form is given in Appendix A).

In general, $\theta_0^*(\theta_0) \neq \theta_0$; therefore direct estimation of the true value θ_0 based on a discretization of the diffusion provides inconsistency as shown by Lo (1988). This is illustrated in the following examples for which there are explicit expressions for the asymptotic bias.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Example 1. Ornstein–Uhlenbeck process

We consider the following diffusion equation:

$$dY_t = (\alpha - \mu Y_t)dt + \sigma dW_t, \tag{3}$$

where $(Y_t; t \in \mathbb{R}^+)$ is a $\mathbb{R}$-valued process with $Y_0 > 0$, and we aim at estimating $\theta = (\alpha, \mu, \sigma^2)$. To get stationary ergodic solutions we impose $\mu > 0, \sigma^2 > 0$. Euler discretization of the Ornstein–Uhlenbeck (OU) has the form

$$Y_{t+\tau} - Y_t = (\alpha^* - \mu^* Y_t)\tau + \sigma^*(W_{t+\tau} - W_t), \tag{4}$$

whereas the *exact* discretization of (3) is given by

$$Y_{t+\tau} - Y_t = \frac{\alpha}{\mu}(1 - e^{-\mu\tau}) - (1 - e^{-\mu\tau})Y_t + \sigma e^{-\mu(t+\tau)} \int_t^{t+\tau} e^{\mu s}\, dW_s. \tag{5}$$

Note that the likelihoods obtained from models (4) and (5) are the same up to a reparameterization. The (a.s.) limit of the estimator $\hat{\theta}_T^*(Y_{\underline{T}}^{\tau})$ of $\theta^* = (\alpha^*, \mu^*, \sigma^{2*})$ in model (4) is given by

$$\theta_0^*(\theta_0) = \left(\frac{\alpha_0}{\mu_0 \tau}(1 - e^{-\mu_0\tau}), \frac{1 - e^{-\mu_0\tau}}{\tau}, \frac{\sigma_0^2}{2\mu_0\tau}(1 - e^{-2\mu_0\tau}) \right), \tag{6}$$

and we see that all the components of $\hat{\theta}_T^*(Y_{\underline{T}}^{\tau})$ are downward biased for fixed sampling interval τ. However the bias goes to zero when τ decreases, and it is of order τ in the neighborhood of continuous-time sampling. Moreover, straightforward (but tedious)[1] calculus shows that

$$\Gamma_0^* = \frac{1 - e^{-2\mu_0\tau}}{2\tau^2\mu_0^2} \begin{bmatrix} \sigma_0^2\mu_0 + 2\alpha_0^2 & 2\alpha_0\mu_0 & 0 \\ 2\alpha_0\mu_0 & 2\mu_0^2 & 0 \\ 0 & 0 & \sigma_0^4(1 - e^{-2\mu_0\tau})^2 \end{bmatrix}. \tag{7}$$

Example 2. Geometric Brownian motion

The geometric Brownian motion (GBM) is defined as a $\mathbb{R}^+$-valued solution process $(Y_t; t \in \mathbb{R}^+)$ of

$$dY_t = \mu Y_t dt + \sigma Y_t dW_t, \tag{8}$$

where $Y_0 > 0$, which can be directly solved in

$$\ln Y_t = \ln Y_0 + (\mu - \tfrac{1}{2}\sigma^2)t + \sigma W_t. \tag{9}$$

Note that (Y_t) is not an ergodic process but that the first-order difference of $(\ln Y_t)$ is ergodic for all values of $\theta = (\mu, \sigma^2)$. Maximum likelihood estimation can then be achieved, and the estimator of θ can be easily computed. Suppose, however, that the Euler discretization is applied directly on equation (8), providing

$$Y_{t+\tau} - Y_t = \mu^* Y_t \tau + \sigma^* Y_t (W_{t+\tau} - W_t). \tag{10}$$

Comparing with the OU example, the distribution of $Y_{t+\tau}$ conditional to Y_t is misspecified in model (10), because it is normal instead of being lognormal. In

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

this simple example, too, the estimator $\hat{\theta}_T^*(Y_{\underline{T}}^\tau)$ of $\theta^* = (\mu^*, \sigma^{2*})$ is easily obtained and its asymptotic almost sure (a.s.) limit is given by

$$\theta_0^*(\theta_0) = \left(\frac{e^{\mu_0 \tau} - 1}{\tau}, \frac{e^{2\mu_0 \tau}}{\tau} (e^{\sigma_0^2 \tau} - 1) \right). \tag{11}$$

Contrarily to the previous example, both components are upward biased. This bias vanishes as τ goes to zero. Furthermore, we get

$$\Gamma_0^* = \frac{e^{2\mu_0 \tau}(e^{\sigma_0^2 \tau} - 1)}{\tau^2} \times \begin{bmatrix} 1 & \tau e^{\mu_0 \tau}(e^{\sigma_0^2 \tau} + 2)(e^{\sigma_0^2 \tau} - 1) \\ \tau e^{\mu_0 \tau}(e^{\sigma_0^2 \tau} + 2)(e^{\sigma_0^2 \tau} - 1) & \tau^2 e^{2\mu_0 \tau}(e^{4\sigma_0^2 \tau} + 2e^{3\sigma_0^2 \tau} + 3e^{\sigma_0^2 \tau} - 4)(e^{\sigma_0^2 \tau} - 1) \end{bmatrix}. \tag{12}$$

These examples could suggest using the explicit expression of the asymptotic bias to obtain asymptotically unbiased estimators. Unfortunately, this approach cannot be applied. Indeed, in general, there is no exact discretized version of the continuous-time model, and the asymptotic bias cannot be derived theoretically.

2.2. Second Step

This step involves simulated data. To approximate the diffusion model we use the following sequence of discrete-time models, indexed by the time-unit h and depending on the same parameter θ as in (1):

$$Y_{(n+1)h}^{(h)} = F_h(Y_{nh}^{(h)}; \epsilon_{(n+1)h}^{(h)}; \theta), \tag{13}$$

where $(\epsilon_{(n+1)h}^{(h)}; n \in \mathbb{Z})$ is a sequence of $\mathbb{R}^d$-valued Gaussian white noises. This sequence is chosen such that (i) for all h, the solution process is unique, stationary, and ergodic; (ii) when h goes to zero, the solution converges to the diffusion process in L^2. A particular case is the Euler discretization given by

$$Y_{(n+1)h}^{(h)} = Y_{nh}^{(h)} + \mu(Y_{nh}^{(h)}; \theta)h + \sigma(Y_{nh}^{(h)}; \theta)h^{1/2}\epsilon_{(n+1)h}^{(h)} \tag{14}$$

or the Mil'shtein (1976) discretization (for a full discussion, see Kloeden and Platen, 1992).

Let $\tau/\mathbb{N} = \{h \in \mathbb{R} | \tau/h \in \mathbb{N}\}$. For any $h \in \tau/\mathbb{N}$ and any $\theta \in \Theta$, let

$$\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta) = (\tilde{Y}_\tau^{(h)}(\theta), \tilde{Y}_{2\tau}^{(h)}(\theta), \ldots, \tilde{Y}_{T\tau}^{(h)}(\theta))$$

be a set of simulated variables drawn from model (13).

2.3. Third Step

We consider now the maximization of the criterion S_T applied to the simulated set.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Let $\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta))$ denote any measurable solution of the maximization program

$$\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta)) = \underset{\theta^* \in \Theta}{\operatorname{argmax}}\, S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta), \theta^*).$$

Under standard assumptions (see Appendix A), the simulation based estimator $\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta))$ is strongly consistent to $\tilde{\theta}_0^*(\theta, h)$ and asymptotically normal:

$$\sqrt{T}(\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta)) - \tilde{\theta}_0^*(\theta, h)) \xrightarrow{T\to\infty} \mathcal{N}(0, \tilde{\Gamma}^{(h)}(\theta))$$

(the explicit form of the asymptotic covariance matrix is given in Appendix A).

Example 1. Ornstein–Uhlenbeck process (continued)

The simulation based estimator may be easily obtained by maximizing the criterion w.r.t. $\theta^* = (\alpha^*, \mu^*, \sigma^{*2})$:

$$S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta); \theta^*) = -\frac{1}{2}\log\sigma^{*2} - \frac{1}{2}\log\tau - \frac{1}{2T\sigma^{*2}\tau}\sum_{t=0}^{T-1}(\tilde{Y}_{(t+1)\tau}^{(h)}(\theta) - \alpha^*\tau - (1-\mu^*\tau)\tilde{Y}_{t\tau}^{(h)}(\theta))^2.$$

We have

$$\tilde{\theta}_0^*(\theta, h) = \left(\frac{\alpha}{\mu\tau}(1-(1-\mu h)^{\tau/h}), \frac{1-(1-\mu h)^{\tau/h}}{\tau}, \frac{\sigma^2}{\mu\tau}\frac{1-(1-\mu h)^{2\tau/h}}{2-\mu h}\right) \tag{15}$$

and

$$\tilde{\Gamma}^{(h)}(\theta) = \frac{1-(1-\mu h)^{\tau/h}}{\tau} \times \begin{bmatrix} \sigma^2\dfrac{1+(1-\mu h)^{\tau/h}}{\mu(2-\mu h)} + \dfrac{\alpha^2}{\mu^2} & \dfrac{\alpha}{\mu} & 0 \\ \dfrac{\alpha}{\mu} & 1 & 0 \\ 0 & 0 & \dfrac{\sigma^4}{\tau}\dfrac{(1+(1-\mu h)^{\tau/h})(1-(1-\mu h)^{2\tau/h})}{\mu^2(2-\mu h)^2} \end{bmatrix}. \tag{16}$$

Example 2. Geometric Brownian motion (continued)

The criterion to be maximized is given by

$$S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta), \theta^*) = -\frac{1}{2}\log\sigma^{*2} - \frac{1}{2}\log\tau - \frac{1}{2T\sigma^{*2}\tau}\sum_{t=0}^{T-1}\left(\frac{\tilde{Y}_{(t+1)\tau}^{(h)}(\theta)}{\tilde{Y}_{t\tau}^{(h)}(\theta)} - (1+\mu^*\tau)\right)^2.$$

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

The expression of the unique maximum of the corresponding asymptotic criterion is given by

$$\tilde{\theta}_0^*(\theta,h)=\left(\frac{(1+\mu h)^{\tau/h}-1}{\tau},\frac{((1+\mu h)^2+\sigma^2 h)^{\tau/h}-(1+\mu h)^{2\tau/h}}{\tau}\right). \tag{17}$$

The expression of matrix $J^*(\theta,h)$ is given in Broze et al. (1995b). It is a nondiagonal matrix.

2.4. Fourth Step

Along the previous lines, we have seen that both the direct estimator $\hat{\theta}_T^*(Y_{\underline{I}}^{\tau})$ and the simulated estimator $\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))$ for all values of θ are consistent and asymptotically normal under some usual conditions. However, for the latter, we need stronger conditions than pointwise convergence to ensure consistency of the estimators discussed later. Therefore, we make the following assumption.

Assumption A1. For a fixed simulation step $h>0$,

$$\mathbb{P}_\theta\left[\lim_{T\to\infty}\sup_{\theta\in\Theta}\|\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))-\tilde{\theta}_0^*(\theta,h)\|=0\right]=1.$$

Because we work with general criteria, it seems difficult to avoid such an assumption. However, it should be noted that the hypothesis does not concern the model itself because the criterion that $\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))$ maximizes is under the control of the researcher (as well as the set of simulated variables). Therefore, it is in principle possible to consider only criteria resulting in estimators that are empirical processes, i.e., of the form

$$\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))=\frac{1}{T}\sum_{t=1}^{T}q^{(h)}(u_t^{(h)};\theta),$$

where $(u_t^{(h)})$ is a sequence of random variables not necessarily i.i.d. and $\{q^{(h)}(\cdot;\theta),\theta\in\Theta\}$ is a class of $\mathbb{R}^p$-valued functions, measurable in their first argument. To ensure uniform almost sure convergence of $\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))$ to $\theta_0^*(\theta,h)$, it is sufficient to impose almost sure stochastic equicontinuity of the sequence $\{\nu_T^{(h)}(\cdot);T\geq 1\}$, where $\nu_T^{(h)}(\theta)=\sqrt{T}(\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))-\theta_0^*(\theta,h))$. In Appendix B, we show how to check Assumption A1 in the first example (the second example is very similar).

To establish a link between the asymptotic (nonstochastic) limits obtained previously, we need the following assumption.

Assumption A2. There exists a sequence $h(T)\to 0$ such that

$$\mathbb{P}_{\theta_0}\left[\lim_{T\to\infty}\sup_{\theta^*\in\Theta}|S_T(\tilde{Y}_{\underline{I}}^{\tau(h(T))}(\theta_0),\theta^*)-S_T(Y_{\underline{I}}^{\tau},\theta^*)|=0\right]=1.$$

The following proposition is then directly deduced.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

PROPOSITION 1. *There exists a sequence $h(T)$ such that*

$$\mathbb{P}_{\theta_0}\left[\lim_{T\to\infty} \|\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h(T))}(\theta_0)) - \theta_0^*(\theta_0)\| = 0\right] = 1.$$

Because in practice the set of observations is finite and the simulation time interval is nonzero, it seems natural to replace the asymptotic limits by the corresponding estimators. Therefore, we have to minimize a distance between estimators based on the observation and the simulation.

DEFINITION 1. *A sequence of quasi-indirect estimators is a sequence $\hat{\theta}_T(\theta_0, h)$, $h \in \tau/\mathbb{N}$, where $\hat{\theta}_T(\theta_0, h)$ is any measurable solution of*

$$\hat{\theta}_T(\theta_0, h) = \underset{\theta\in\Theta}{\operatorname{argmin}} \|\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta)) - \hat{\theta}_T^*(Y_{\underline{T}}^{\tau})\|.$$

The first problem is the consistency of this quasi-indirect estimator. For all $h \in \tau/\mathbb{N}$, we denote $\theta_0(\theta_0, h)$ the unique solution to the equation

$$\tilde{\theta}_0^*(\theta, h) = \theta_0^*(\theta_0),$$

and we get the proposition.

PROPOSITION 2. *For any fixed value of the simulation step $h \in \tau/\mathbb{N}$, we have*

$$\mathbb{P}_{\theta_0}\left[\lim_{T\to\infty} \|\hat{\theta}_T(\theta_0, h) - \theta_0(\theta_0, h)\| = 0\right] = 1.$$

Consequently, the quasi-indirect estimator is asymptotically *biased* for fixed discretization step h.

Proof. Let us consider the criterion function

$$Q_T(\theta, h) = \|\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta)) - \hat{\theta}_T^*(Y_{\underline{T}}^{\tau})\|.$$

From previous results, we have uniform convergence on Θ of $\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta)) - \hat{\theta}_T^*(Y_{\underline{T}}^{\tau})$ to $\tilde{\theta}_0^*(\theta, h) - \theta_0^*(\theta_0)$, for all $h > 0$. Moreover, $Q_T(\cdot, h)$ is continuous on Θ. Therefore, because $\tilde{\theta}_0^*(\theta, h) - \theta_0^*(\theta_0)$ is a minimum for the single value $\theta_0(\theta_0, h)$, the proposed result holds. ■

Example 1. Ornstein–Uhlenbeck process (continued)

We have

$$\theta_0(\theta_0, h) = \left(\frac{\alpha_0(1 - e^{-\mu_0 h})}{\mu_0 h}, \frac{1 - e^{-\mu_0 h}}{h}, \frac{\sigma_0^2(1 - e^{-2\mu_0 h})}{2\mu_0 h}\right).$$

Example 2. Geometric Brownian motion (continued)

We have

$$\theta_0(\theta_0, h) = \left(\frac{e^{-\mu_0 h} - 1}{h}, \frac{e^{2\mu_0 h}}{h}(e^{\sigma_0^2 h} - 1)\right).$$

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Let us now examine the asymptotic distribution of the quasi-indirect estimator.

PROPOSITION 3. *The quasi-indirect estimator* $\hat{\theta}_T(\theta_0, h)$ *is strongly consistent to* $\theta_0(\theta_0, h)$ *and asymptotically normal*:

$$\sqrt{T}(\hat{\theta}_T(\theta_0, h) - \theta_0(\theta_0, h)) \xrightarrow{T\to\infty} \mathcal{N}(0, \overline{V}_0^{(h)}),$$

where the explicit form of matrix $\overline{V}_0^{(h)}$ *is given in Appendix C.*

Proof. See Appendix C.

From the previous results, we know that the quasi-indirect estimator based on a fixed simulation step h is strongly consistent to $\theta_0(\theta_0, h)$, and furthermore, we have

$$\mathbb{P}_{\theta_0}\left[\lim_{T\to\infty} \|\hat{\theta}_T(\theta_0, h(T)) - \theta_0\| = 0\right] = 1.$$

Thus there exists a sequence $h(T)$ going to zero as T goes to infinity such that expected consistency to θ_0 holds. However, from a practical point of view, this result is unsatisfactory because for a given sample size T, one would expect to know how to choose h as a function of T to obtain a good approximation of θ_0.

Hereafter, we analyze further asymptotic properties of the quasi-indirect inference estimator by allowing the simulation step h to vary around zero (see Appendix D for technical details and for the relevant expressions of the asymptotic covariance matrices).

If we take $h = T^{-\delta}$ with $\frac{1}{4} < \delta \leq \frac{1}{2}$, we have an asymptotic normality result of the form

$$\sqrt{T}\left(\hat{\theta}_T^*(\theta_0, h) - \theta_0 - h\left[\frac{\partial \theta_0(\theta_0, h)}{\partial h}\right]_{h=0}\right) \xrightarrow{T\to\infty} \mathcal{N}(0, \overline{\overline{V}}_0).$$

On the other hand, for smaller value of h, $h = T^{-\delta}$ with $\delta > \frac{1}{2}$, we have the usual normality result:

$$\sqrt{T}(\hat{\theta}_T^*(\theta_0, h) - \theta_0) \xrightarrow{T\to\infty} \mathcal{N}(0, \overline{\overline{V}}_0).$$

Finally, if h is not chosen sufficiently small, we have some additional terms involving higher orders terms in h in the asymptotic normality result. It is important, however, to recall that the discretization step h is controlled by the statistician. In practice it means that, from an asymptotic point of view, h should be chosen as small as the computer capacities allow. This is obvious because the limit case $h = 0$ corresponds to the possibility of simulating the diffusion itself. For finite sample, this point of view will be revisited in the next section.

Example 1. Ornstein–Uhlenbeck process (continued)

The asymptotic bias of the quasi-indirect estimator is approached by

$$\theta_0(\theta_0, h) - \theta_0 = (-\tfrac{1}{2}\alpha_0\,\mu_0, -\tfrac{1}{2}\mu_0^2, -\sigma_0^2\,\mu_0)h + o(h),$$

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

and the matrix $V_0 = \lim_{T\to\infty} V_0^{(h(T))}$ is given by

$$V_0 = \begin{pmatrix} \dfrac{1-e^{-\mu_0\tau}}{\mu\tau} & 0 & 0 \\ \dfrac{\alpha_0 e^{-\mu_0\tau}}{\mu_0}\left[1+\dfrac{1}{\mu_0\tau}\right] - \dfrac{\alpha_0}{\mu_0^2\tau} & e^{-\mu_0\tau} & \dfrac{\sigma_0^2 e^{-2\mu_0\tau}}{\mu_0}\left[1+\dfrac{1}{2\mu_0\tau}\right] - \dfrac{\sigma_0^2}{2\mu_0^2\tau} \\ 0 & 0 & \dfrac{1-e^{-2\mu_0\tau}}{2\mu_0\tau} \end{pmatrix}.$$

Example 2. Geometric Brownian motion (continued)

The asymptotic bias is given by

$$\theta_0(\theta_0, h) - \theta_0 = (-\tfrac{1}{2}\mu_0^2, \sigma_0^2(2\mu_0 + \tfrac{1}{2}\sigma_0^2))h + o(h)$$

and the matrix V_0 by

$$V_0 = \begin{pmatrix} e^{\mu_0\tau} & 2e^{2\mu_0\tau}[e^{\sigma_0^2\tau} - 1] \\ 0 & e^{(2\mu_0+\sigma_0^2)\tau} \end{pmatrix}.$$

3. MONTE CARLO RESULTS

To illustrate the indirect estimation procedure, we perform some Monte Carlo experiments.[2] Both examples are examined. We first compute the exact discretization maximum likelihood (ML) estimator to be able to compare this estimator with the misspecified estimator obtained by a ML method on a Euler discretization and the estimator obtained by the quasi-indirect estimation method based on a likelihood criterion. Closed-form expressions for both exact and Euler discretization ML estimators are derived and are used whereas a numerical optimization routine is involved in solving the minimization problem of the quasi-indirect inference method. The algorithm is initialized with the estimated values given by the ML method applied to the Euler scheme. This procedure reflects the general situation where one wants to estimate parameters of a general diffusion process (not knowing its exact discretization), then uses a simple enough auxiliary model leading to closed-form estimators without being able to compute the explicit form of the indirect estimator. We hope that the results reported hereafter closely mimic the behavior of quasi-indirect estimators in more general situations.

Although the simulation experiments were conducted for all parameters (see Broze et al., 1995b), we only report the outcome concerning the diffusion parameter σ_0. This choice is motivated by the crucial need for volatility estimation in financial econometrics.

There are two experiments conducted in the study. All simulated data are obtained from the exact discretization of either GBM or OU.

In Tables 1 and 2, we report the simulation results for 1,000 replications (see also Figure 1). The sample size is $T = 200$ or $T = 50$. The discretization step h

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

TABLE 1. Characteristics of the empirical p.d.f. of the estimators (K varying)

Method	K	h	Bias	MSE	Max.	Min.	Median
Design 1. Geometric Brownian motion: $\sigma_0 = 0.5$, $T = 50$							
MLE			−0.00936	0.00252	0.63189	0.33341	0.48882
MMLE			−0.07738	0.01186	0.79045	0.22844	0.41597
QIE	1	0.1	0.00119	0.00900	0.91000	−0.52632	0.49273
QIE	1	0.01	0.00965	0.00942	0.96502	0.29575	0.49844
QIE	2	0.1	−0.01083	0.00601	0.86070	0.29627	0.48370
QIE	2	0.01	−0.00262	0.00691	0.88194	0.31765	0.48676
QIE	5	0.1	−0.01160	0.00144	0.63439	0.37340	0.48603
QIE	5	0.01	−0.00530	0.00141	0.63162	0.40550	0.49135
QIE	8	0.1	−0.02087	0.00470	0.76032	0.31794	0.47372
QIE	8	0.01	−0.01377	0.00477	0.80781	0.32520	0.47805
QIE	10	0.1	−0.02197	0.00460	0.76357	0.30891	0.47438
QIE	10	0.01	−0.01498	0.00462	0.78439	0.31995	0.47699
QIE	15	0.1	−0.02258	0.00455	0.77223	0.30812	0.47299
QIE	15	0.01	−0.01608	0.00446	0.77271	0.32387	0.47775
QIE	20	0.1	−0.02277	0.00446	0.75965	0.30936	0.47177
QIE	20	0.01	−0.01661	0.00430	0.77159	0.32141	0.47691
Design 2. Ornstein–Uhlenbeck process: $\sigma_0 = 1.5$, $T = 50$							
MLE			0.01917	0.03267	2.15925	0.95634	1.51451
MMLE			−0.16720	0.04833	1.79786	0.86268	1.33083
QIE	1	0.1	0.04094	0.08158	5.10261	0.92422	1.51607
QIE	1	0.01	0.06312	0.14237	8.27802	0.88318	1.52880
QIE	2	0.1	0.00189	0.04073	2.50470	0.85452	1.49351
QIE	2	0.01	0.01197	0.04534	2.28038	0.86697	1.49729
QIE	5	0.1	−0.01547	0.00817	1.81240	1.18801	1.48599
QIE	5	0.01	−0.00143	0.00825	1.79166	1.23555	1.49762
QIE	8	0.1	−0.01442	0.02910	2.01227	0.94374	1.47698
QIE	8	0.01	0.00085	0.03174	2.12157	0.94060	1.49398
QIE	10	0.1	0.01566	0.02919	2.01222	0.95496	1.47521
QIE	10	0.01	0.00060	0.03112	2.12303	0.97291	1.49480
QIE	15	0.1	−0.01651	0.02844	2.00788	0.97058	1.47507
QIE	15	0.01	−0.00046	0.02999	2.05627	0.96474	1.49151
QIE	20	0.1	−0.02292	0.02894	2.14687	0.96987	1.47604
QIE	20	0.01	−0.00102	0.03056	2.14587	0.97012	1.48864

varies from $\frac{1}{10}$ to $\frac{1}{5{,}000}$. The initial value is $Y_0 = 10$. Designs 1, 3, and 4 are concerned with GBM ($\sigma_0 = 0.5$); designs 2, 5, and 6 with OU ($\sigma_0 = 1.5$). For each design, we consider three estimators obtained from (i) maximum likelihood on the exact discretization (MLE), (ii) maximum likelihood on the misspecified Euler discretization (MMLE), and (iii) quasi-indirect inference (QIE). Notice that the first one is given as a gauge although it is not feasible in the general case.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

TABLE 2. Characteristics of the empirical p.d.f. of the estimators (h varying)

Method	h	Bias	MSE	Max.	Min.	Median
Design 3. Geometric Brownian motion: $\sigma_0 = 0.5$, $T = 200$, $K = 5$						
MLE		−0.00378	0.00062	0.57566	0.41571	0.49647
MMLE		−0.06954	0.00648	0.59433	0.31766	0.42714
QIE	0.1	−0.01160	0.00144	0.63439	0.37340	0.48603
QIE	0.05	−0.00834	0.00141	0.62007	0.39196	0.49010
QIE	0.01	−0.00530	0.00141	0.63162	0.40550	0.49135
QIE	0.005	−0.00464	0.00138	0.63714	0.39671	0.49192
Design 4. Geometric Brownian motion: $\sigma_0 = 0.5$, $T = 50$, $K = 5$						
MLE		−0.00936	0.00252	0.63189	0.33341	0.48882
MMLE		−0.07738	0.01185	0.79045	0.22844	0.41597
QIE	0.1	−0.01868	0.00489	0.73457	0.28320	0.47325
QIE	0.05	−0.01526	0.00504	0.76395	0.30160	0.47749
QIE	0.01	−0.01155	0.00504	0.77836	0.32301	0.48031
QIE	0.005	−0.01115	0.00505	0.77458	0.32828	0.48187
Design 5. Ornstein–Uhlenbeck process: $\sigma_0 = 1.5$, $T = 200$, $K = 5$						
MLE		0.00433	0.00734	1.77405	1.22962	1.50383
MMLE		−0.14747	0.02664	1.55143	1.14897	1.35248
QIE	0.1	−0.01547	0.00817	1.81240	1.18801	1.48599
QIE	0.05	−0.00885	0.00840	1.83764	1.17766	1.48916
QIE	0.01	−0.00143	0.00825	1.79166	1.23555	1.49762
QIE	0.005	−0.00075	0.00820	1.84266	1.21640	1.49908
Design 6. Ornstein–Uhlenbeck process: $\sigma_0 = 1.5$, $T = 50$, $K = 5$						
MLE		0.01917	0.03267	2.15925	0.95634	1.51451
MMLE		−0.16720	0.04833	1.79786	0.86268	1.33083
QIE	0.1	−0.01330	0.03047	2.10906	0.93110	1.47759
QIE	0.05	−0.00698	0.03191	2.10805	0.93396	1.48138
QIE	0.01	0.00305	0.03502	2.14045	0.91392	1.49309
QIE	0.005	0.00569	0.03384	2.11347	0.94367	1.50209

For all estimators, we show the bias, the mean square error (MSE), and the minimum, the maximum, and the median values among the 1,000 estimator values.

Looking at the discretization step in QIE, it is clear that there is no reason to match estimators computed with observed and simulated series of the same length. An increased precision can be expected from such longer simulated series. Let $K \times T$ denote the length of the simulated series, maintaining T as the observed sample size.

Table 1 emphasizes the impact of using simulated paths longer than the sample size and gathers the outcome of a Monte Carlo experiment for the small sample case $T = 50$. Increasing K does not necessarily improve the performance of QIE in terms of bias, except for the OU case with $h = \frac{1}{100}$. This result is confirmed by examining the MSE: *an optimal choice of K appears to be* 5.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

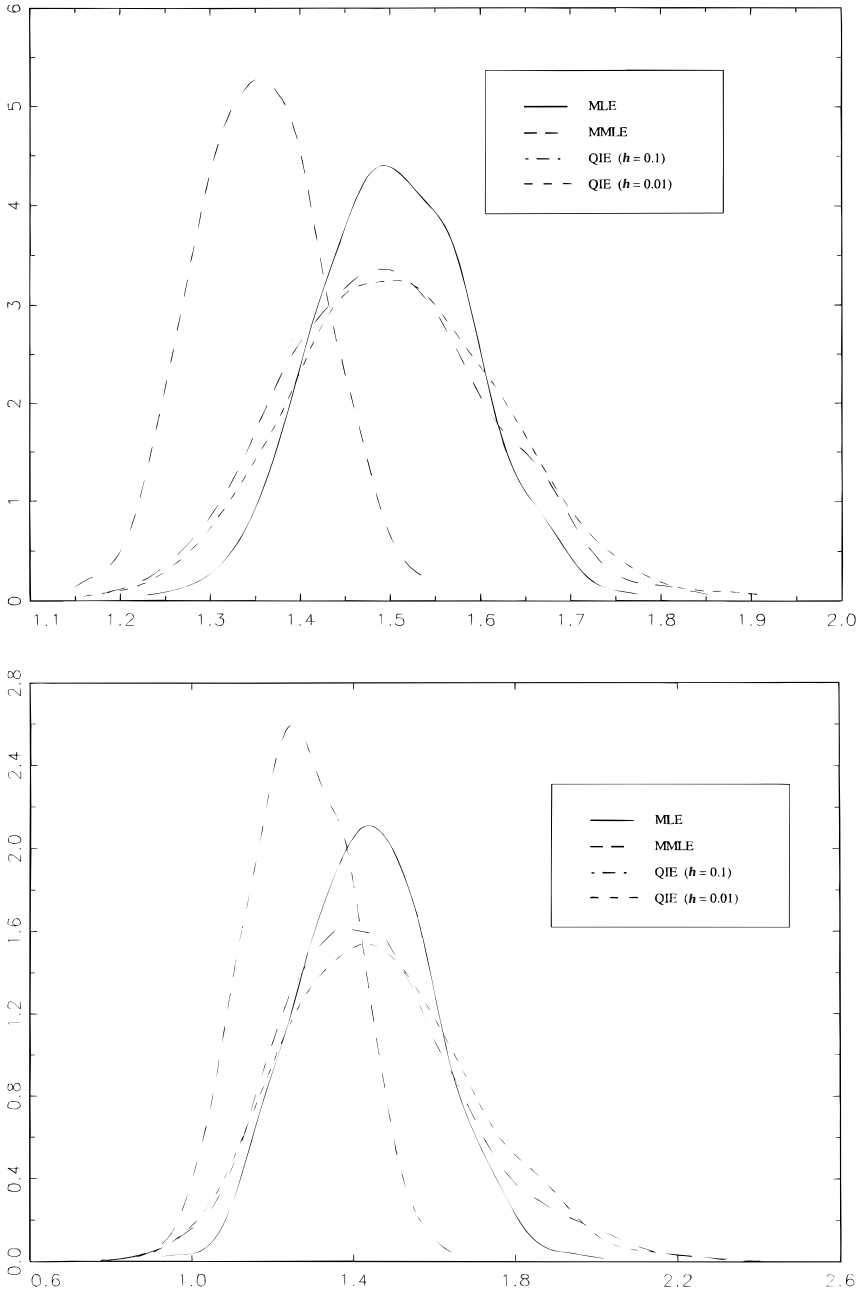


FIGURE 1. Empirical p.d.f. of the estimators of σ in the Ornstein–Uhlenbeck process for $T = 200$ and $T = 50$; $\alpha_0 = 0.3$, $\mu_0 = 0.2$, and $\sigma_0 = 1.5$.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

This value $K = 5$ has been maintained for the second experiment, reported in Table 2, where the impact of decreasing h is studied. First, for a sample size $T = 200$, we confirm the asymptotic theoretical result of the previous section: decreasing h is optimal for the bias. The MSE is improved in the GBM case and remains stable in the OU case. For the small sample case $T = 50$, decreasing h too much does not improve MSE: *An optimal choice of h appears to be* 0.01.

Finally, one should expect that MLE performs better than the other estimators, from its asymptotic properties. It is clear that applying maximum likelihood to the misspecified Euler discretization of the diffusion model provides a dramatic bias. The bias can be substantially corrected by using QIE. In fact it is seen that QIE can do even better than MLE. Similar findings have been reported by Gouriéroux et al. (1993) in the case of autoregressive processes and have been explored with high order asymptotic expansions by Gouriéroux, Renault, and Touzi (1994).

Finally, let us remark that, from simulations not reported here, the second-order discretization scheme of Mil'shtein (1976) does not lead to better results (this phenomenon was also observed by Bianchi, Cesari, and Panattoni [1993]).

It is always difficult to draw definitive conclusions from Monte Carlo experiments. Nevertheless, the following can be said. Although our theoretical asymptotic results encourage choosing an arbitrarily small discretization step, it is seen from the numerical experiments that the gain in precision does not necessarily balance the resulting computational difficulties. Moreover, a larger quantity of simulated data increases variability and alters precision. Second, the simulations clearly show that for a given discretization step increasing the length of the simulated path ($K \times T$ instead of T) does improve the performance of QIE. Choosing K too large is not recommended because it both requires impressive computational times and worsens bias and precision. Therefore, from our experience choosing $h = \frac{1}{100}$ and $K = 5$ seems a good compromise.

4. CONCLUDING REMARKS

Simulation based estimators such as indirect inference have gained considerable interest in recent years in financial econometrics and other areas. Indeed, the possibility of simulating complex models has increased our ability to estimate models hitherto infeasible to handle like diffusions and stochastic volatility models. Recent interesting developments along these lines are, e.g., Ghysels, Gouriéroux, and Jasiak (1994), Pastorello, Renault, and Touzi (1995), Engle and Lee (1992), Gallant, Hsieh, and Tauchen (1994), Gallant and Tauchen (1995), and Tauchen (1995).

This paper introduced the quasi-indirect inference method for diffusion processes and discussed its asymptotic properties. With the fixed simulation step, it is shown that quasi-indirect inference is biased even for a large sample. We proved, however, that by choosing an appropriate simulation step, sufficiently small in

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

comparison with the sample size, one obtains an asymptotically unbiased estimator. We provided explicit analytical expressions for the choice of the simulation step and conducted Monte Carlo experiments to examine the bias reduction in two commonly used examples. In addition, applications of quasi-indirect inference to real financial data can be found in Broze, Scaillet, and Zakoïan (1995a) and de Winne (1994).

Finally, it appears from our Monte Carlo experiments that the quasi-indirect estimator can have better finite-sample properties as far as bias is concerned than the exact maximum likelihood estimator.

NOTES

1. Detailed computations can be found in Broze, Scaillet, and Zakoïan (1995b).
2. Computer codes in GAUSS are available on request.

REFERENCES

Aït-Sahalia, Y. (1992) The Delta and Bootstrap Methods for Nonlinear Functionals of Nonparametric Kernel Estimators Based on Dependent Multivariate Data. Manuscript, MIT.

Aït-Sahalia, Y. (1996) Nonparametric pricing of interest rate derivative securities. *Econometrica* 64, 527–560.

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–857.

Bianchi, C., R. Cesari, & L. Panattoni (1993) Alternative Estimators of a Diffusion Model of the Term Structure of Interest Rates: A Monte Carlo Comparison. Discussion paper, University of Pisa.

Broze, L., O. Scaillet, & J.M. Zakoïan (1995a) Testing for continuous-time models of the short term interest rate. *Journal of Empirical Finance* 2, 199–223.

Broze, L., O. Scaillet, & J.M. Zakoïan (1995b) Quasi Indirect Inference for Diffusion Processes. CORE Discussion paper 9505.

Dacunha-Castelle, D. & D. Florens-Zmirou (1986) Estimation of the coefficient of a diffusion from discrete observations. *Stochastics* 19, 263–284.

Danesi, V., V. Genon-Catalot, & J.P. Laurent (1994) Estimating Continuous Time Yield Curve Models. Mimeo.

De Winne, R. (1994) The Discretization Bias for Processes of the Short-Term Interest Rate: An Empirical Analysis. Discussion paper, FUCAM.

Drost, F.C. & B.J.M. Werker (1996) Closing the GARCH gap: Continuous time GARCH modeling. *Journal of Econometrics* 74, 31–57.

Duffie, D. & K.J. Singleton (1989) Simulated Moments Estimation of Markov Models of Asset Prices. Discussion paper, Stanford University.

Duffie, D. & K.J. Singleton (1993) Simulated moments estimation of Markov models of asset prices. *Econometrica* 61, 929–954.

Engle, R. & G.G. Lee (1992) Estimating Diffusion Models of Stochastic Volatility. Discussion paper, University of California at San Diego.

Florens-Zmirou, D. (1989) Approximate discrete time schemes for statistics of diffusion processes. *Statistics* 20, 547–557.

Gallant, A.R., D.A. Hsieh, & G. Tauchen (1994) Estimation of Stochastic Volatility Models with Suggestive Diagnostics. Discussion paper, Duke University.

Gallant, A.R. & G. Tauchen (1995) Estimation of Continuous Time Models for Stock Returns and Interest Rates. Discussion paper, Duke University.

Gallant, A.R. & G. Tauchen (1996) Which moments to match? *Econometric Theory* 12, 657–681.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Gallant, A.R. & H. White (1988) *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models.* New York: Basil Blackwell.
Genon-Catalot, V. (1990) Maximum contrast estimation for diffusion process from discrete observations. *Statistics* 21, 99–116.
Genon-Catalot, V. & J. Jacod (1993) On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales IHP* 29, 119–151.
Ghysels, E., C. Gouriéroux, & J. Jasiak (1994) Market Time and Asset Price Movements: Theory and Estimation. Discussion paper, CREST.
Gouriéroux, C. & A. Monfort (1992) Simulation based inference in models with heterogeneity. *Annales d'Economie et de Statistique* 20/21, 69–108.
Gouriéroux, C. & A. Monfort (1993) Simulation based inference: A survey with special reference to panel data models. *Journal of Econometrics* 59, 5–33.
Gouriéroux, C. & A. Monfort (1995) *Simulation Based Econometric Methods.* CORE Lectures Series, Louvain-la-Neuve.
Gouriéroux, C., A. Monfort, & E. Renault (1993) Indirect inference. *Journal of Applied Econometrics* 8, 85–118.
Gouriéroux, C., E. Renault, & N. Touzi (1994) Calibration by Simulation for Small Sample Bias Correction. CREST Discussion paper, 9554.
Hansen, L. & J. Scheinkman (1995) Back to the future: Generating moment implications for continuous-time Markov processes. *Econometrica* 63, 767–804.
Hansen, L., J. Scheinkman, & N. Touzi (1994) Identification of Scalar Diffusions Using Eigenvectors. Manuscript, University of Chicago.
Kloeden, P.E. & E. Platen (1992) *Numerical Solutions of Stochastic Differential Equations.* New York: Springer-Verlag.
Laroque, G. & B. Salanié (1993) Simulation-based estimation of models with lagged latent variables. *Journal of Applied Econometrics* 8, 119–134.
Lee, B. & B. Ingram (1991) Simulation estimation of time series models. *Journal of Econometrics* 47, 197–205.
Lerman, S. & C. Manski (1981) On the use of simulated frequencies to approximate choice probabilities. In C. Manski & D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 305–319. Cambridge, Massachusetts: MIT Press.
Lo, A.W. (1988) Maximum likelihood estimation of generalized Ito processes with discretely sampled data. *Econometric Theory* 4, 231–247.
McFadden, D. (1989) A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57, 995–1026.
Mil'shtein, G. (1976) A method of second order accuracy integration of stochastic differential equations. *Theory of Probability and Its Applications* 23, 396–401.
Pakes, A. & D. Pollard (1989) Simulation and the asymptotics of optimization estimators. *Econometrica* 57, 1027–1058.
Pastorello, S., E. Renault, & N. Touzi (1994) Statistical Inference for Random Variance Option Pricing. Discussion paper, Università degli Studi Padova.
Pedersen, A.R. (1994) Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes. Discussion paper, Aarhus University.
Pedersen, A.R. (1995) A new approach to maximum likelihood estimation for stochastic differential equation based on discrete observations. *Scandinavian Journal of Statistics* 22, 55–71.
Pötscher, B. & I. Prucha (1994) Generic uniform and equicontinuity concepts for random functions: An exploration of the basic structure. *Journal of Econometrics* 60, 23–63.
Prasaka Rao, B. (1988) Statistical inference from sampled data for stochastic processes. *Contemporary Mathematics, American Mathematical Society* 20, 249–284.
Revuz, D. & M. Yor (1991) *Continuous Martingales and Brownian Motion.* Berlin: Springer-Verlag.
Smith, A. (1990) *Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models.* Ph.D. Dissertation, Duke University.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Tauchen, G. (1995) New Minimum Chi-Square Methods in Empirical Finance. Discussion paper, Duke University.
Van Praag, B.M.S. & J.P. Hop (1987) Estimation of Continuous Models on the Basis of Set-Valued Observations. Mimeo.
White, H. (1994) *Estimation, Inference and Specification Analysis.* Cambridge: Cambridge University Press.

APPENDIX A. ASYMPTOTIC PROPERTIES OF THE DIRECT ESTIMATORS

Concerning the direct estimator, we introduce the following standard regularity assumptions.

1. The partial application $S(\cdot,\theta)$ is measurable from $(\Omega,\mathcal{F})$ to $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ for all $\theta\in\Theta$ and continuous from Θ to $\mathbb{R}$ for all $\omega\in\Omega$.
2. For all $\theta_0\in\Theta$, there exists a nonstochastic limit, $S_\infty(\theta_0,\theta^*)$, such that

$$\mathbb{P}_{\theta_0}\left[\lim_{T\to\infty}\sup_{\theta^*\in\Theta}|S_T(Y_1^T,\theta^*)-S_\infty(\theta_0,\theta^*)|=0\right]=1.$$

Moreover, for all $\theta_0\in\Theta$, the function $S_\infty(\theta_0,\cdot)$ admits a unique maximum interior to Θ:

$$\theta_0^*(\theta_0)=\underset{\theta^*\in\Theta}{\operatorname{argmax}}\, S_\infty(\theta_0,\theta^*).$$

The criterion function $S_T(Y_1^T,\theta^*)$ is twice continuously differentiable w.r.t. θ^* and

(i) for all $\theta_0\in\Theta$, there exists a nonstochastic matrix, $\ddot{S}_\infty(\theta_0,\theta^*)$, such that

$$\mathbb{P}_{\theta_0}\left[\lim_{T\to\infty}\sup_{\theta^*\in\Theta}\left\|\frac{\partial^2 S_T(Y_1^T,\theta^*)}{\partial\theta^*\partial\theta^{*\prime}}-\ddot{S}_\infty(\theta_0,\theta^*)\right\|=0\right]=1;$$

(ii) $\ddot{S}_\infty(\theta_0,\theta_0^*(\theta_0))$ is invertible;

(iii)

$$\left[\sqrt{T}\,\frac{\partial S_T(Y_1^T,\theta^*)}{\partial\theta^*}\right]_{\theta^*=\theta_0^*(\theta_0)}\xrightarrow{T\to\infty}\mathcal{N}(0,J^*(\theta_0)).$$

We have then the following first-order expansion:

$$\hat{\theta}_T^*(Y_1^T)=\theta_0^*(\theta_0)+T^{-1/2}X_T^*(\theta_0)+o_p(T^{-1/2}), \tag{A.1}$$

where

$$X_T^*(\theta_0)=-\sqrt{T}\left[\frac{\partial^2 S_T(Y_1^T;\theta^*)}{\partial\theta^*\partial\theta^{*\prime}}\right]^{-1}\left[\frac{\partial S_T(Y_1^T;\theta^*)}{\partial\theta^*}\right]_{\theta^*=\theta_0^*(\theta_0)}$$

is a Gaussian random vector in $\mathbb{R}^p$ with mean 0 and asymptotic covariance matrix given by $\Gamma_0^*=[\ddot{S}_\infty(\theta_0,\theta_0^*(\theta_0))]^{-1}J^*(\theta_0)[\ddot{S}_\infty(\theta_0,\theta_0^*(\theta_0))]^{-1}$.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Let us now introduce the assumptions needed to derive the asymptotic properties of the simulation based estimator.

3. For all $\theta \in \Theta$ and all $h \in \tau/\mathbb{N}$, there exists a nonstochastic limit, $S_\infty^{(h)}(\theta, \theta^*)$, such that

$$\mathbb{P}_\theta\left[\lim_{T\to\infty} \sup_{\theta^*\in\Theta} |S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta), \theta^*) - S_\infty^{(h)}(\theta, \theta^*)| = 0\right] = 1.$$

Moreover, for all $\theta \in \Theta$, the function $S_\infty^{(h)}(\theta, \cdot)$ admits a unique maximum interior to Θ:

$$\tilde{\theta}_0^*(\theta, h) = \operatorname*{argmax}_{\theta^*\in\Theta} S_\infty^{(h)}(\theta, \theta^*).$$

The criterion function $S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta), \theta^*)$ is twice continuously differentiable w.r.t. θ^* and

(i) for all $\theta \in \Theta$, there exists a nonstochastic matrix, $\ddot{S}_\infty^{(h)}(\theta, \theta^*)$, such that

$$\mathbb{P}_\theta\left[\lim_{T\to\infty} \sup_{\theta^*\in\Theta} \left\| \frac{\partial^2 S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta), \theta^*)}{\partial\theta^*\partial\theta^{*\prime}} - \ddot{S}_\infty^{(h)}(\theta, \theta^*) \right\| = 0\right] = 1$$

uniformly in $\theta^* \in \Theta$;

(ii) $\ddot{S}_\infty^{(h)}(\theta, \tilde{\theta}_0^*(\theta, h))$ is invertible for all $\theta \in \Theta$;

(iii)

$$\left[\sqrt{T}\,\frac{\partial S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta), \theta^*)}{\partial\theta^*}\right]_{\theta^*=\tilde{\theta}_0^*(\theta,h)} \xrightarrow{T\to\infty} \mathcal{N}(0, J^*(\theta, h)).$$

Under Assumptions 1 and 3, we have, for all $\theta \in \Theta$ and all $h \in \tau/\mathbb{N}$, the following asymptotic expansion analogous to (A.1):

$$\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta)) = \tilde{\theta}_0^*(\theta, h) + T^{-1/2}\tilde{X}_T(\theta, h) + o_p(T^{-1/2}), \tag{A.2}$$

where

$$\tilde{X}_T(\theta, h) = -\sqrt{T}\left[\frac{\partial^2 S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta); \theta^*)}{\partial\theta^*\partial\theta^{*\prime}}\right]^{-1}\left[\frac{\partial S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta); \theta^*)}{\partial\theta^*}\right]_{\theta^*=\tilde{\theta}_0^*(\theta,h)}$$

is a Gaussian random vector in $\mathbb{R}^p$ with mean 0 and asymptotic variance given by

$$\tilde{\Gamma}^{(h)}(\theta) = [\ddot{S}_\infty^{(h)}(\theta, \tilde{\theta}_0^*(\theta, h))]^{-1} J^*(\theta, h)[\ddot{S}_\infty^{(h)}(\theta, \tilde{\theta}_0^*(\theta, h))]^{-1}.$$

APPENDIX B. UNIFORM CONVERGENCE OF THE SIMULATED ESTIMATOR: THE ORNSTEIN–UHLENBECK CASE

We check Assumption 1 for the Ornstein–Uhlenbeck process. The geometric Brownian motion can be analyzed very similarly using the expressions given in Broze et al. (1995b) and involves much easier computations.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

We assume that the compact parameter space Θ is included in a space of the form $\{\theta = (\alpha, \mu, \sigma^2); \bar{\mu} > \mu > \underline{\mu}, \sigma^2 > \underline{\sigma}^2\}$ where $\bar{\mu}, \underline{\mu}, \underline{\sigma}^2$ are positive constants with $\bar{\mu} < 2/h$. The first-order conditions in the maximization of $S_T(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta); \theta^*)$ can be solved easily in this case and yield estimators of θ of the form

$$\begin{cases} 1 - \hat{\mu}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))\tau = \dfrac{\overline{(YY_1)}_T - \bar{Y}_T^2}{\overline{Y^2}_T - \bar{Y}_T^2} + o_p(1), \\ \dfrac{\hat{\alpha}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\hat{\mu}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))} = \bar{Y}_T + o_p(1), \\ \hat{\sigma}_T^{*2}(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))\tau = \dfrac{1}{T}\displaystyle\sum_{t=0}^{T-1} (\tilde{Y}_{(t+1)\tau}^{(h)}(\theta) - \hat{\alpha}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))\tau \\ \qquad\qquad - (1 - \hat{\mu}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))\tau)\tilde{Y}_{t\tau}^{(h)}(\theta))^2 + o_p(1), \end{cases}$$

where $\bar{Y}_T = 1/T\sum_{t=0}^{T-1}\tilde{Y}_{t\tau}^{(h)}(\theta)$, $\overline{Y^2}_T = 1/T\sum_{t=0}^{T-1}\tilde{Y}_{t\tau}^{(h)2}(\theta)$, and $\overline{(YY_1)}_T = 1/T \times \sum_{t=0}^{T-1}\tilde{Y}_{t\tau}^{(h)}(\theta)\tilde{Y}_{(t+1)\tau}^{(h)}(\theta)$. We will note their a.s. limits $\bar{Y}$, $\overline{Y^2}$, and $\overline{YY_1}$, respectively. For simplicity, we omit the dependence in θ in these notations. The $o_p(1)$ terms correspond to quantities converging a.s. to zero when T goes to infinity, uniformly on Θ.

As the simulated process $(\tilde{Y}_{kh}^{(h)}(\theta))_{k\geq 0}$ is obtained from the use of the recursive form corresponding to the Euler discretization, $\tilde{Y}_{t\tau}^{(h)}(\theta)$ is a continuous function of θ, for all $t \geq 0$. Besides, because we work on a compact parameter space, the quantities $\bar{Y}_T$, $\overline{Y^2}_T$, and $\overline{(YY_1)}_T$ (and their limits) are uniformly bounded on Θ.

To obtain the a.s. uniform convergence of Assumption 1, it is then sufficient to prove that these terms converge uniformly to their limits and that there exists a real number $\nu > 0$ such that $\overline{Y^2}_T - \bar{Y}_T^2 > \nu$, a.s., for all $\theta \in \Theta$, at least for large T.

To check that $\sup_{\theta\in\Theta}|\bar{Y}_T - \bar{Y}| \to 0$ a.s. as $T \to \infty$, we use Corollary 4.3 in Pötscher and Prucha (1994).

i. We first show the pointwise convergence of $\bar{Y}_T$ to $\bar{Y}$. Using the recursive form corresponding to the Euler discretization, the a.s. limit $\bar{Y}$ may be written

$$\bar{Y} = \frac{\alpha}{\mu} + \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{l=0}^{t}\varphi_h^{(t-l)\tau/h}u_{(l+1)\tau}^{(h)}(\theta), \quad \text{a.s.}$$

where $\varphi_h = (1 - \mu h)$ and $u_{(l+1)\tau}^{(h)}(\theta) = \sigma\sqrt{h}\sum_{k=0}^{\tau/h-1}\varphi_h^k\epsilon_{(l+1)\tau-kh}$, or equivalently

$$\bar{Y} = \frac{\alpha}{\mu} + \lim_{T\to\infty}\frac{1}{T}\sum_{l=0}^{T-1} w_{l,T}u_{(l+1)\tau}^{(h)}(\theta), \quad \text{a.s.}$$

with $w_{l,T} = (1 - \varphi_h^{(T-l)\tau/h})/(1 - \varphi_h^{\tau/h})$.

The variables $u_{(l+1)\tau}^{(h)}(\theta)$, $l = 0,1,\ldots$ are i.i.d. with finite variance, and we have $0 < w_{l,T} < 2/1 - \varphi_h^{\tau/h}$, which leads to $1/T^2\sum_{l=0}^{T-1}V[w_{l,T}u_{(l+1)\tau}^{(h)}(\theta)] \to 0$ a.s. as $T \to \infty$, and a pointwise strong law of large numbers applies, for all $\theta \in \Theta$.

ii. Then it may directly be verified that the sequence $(\tilde{Z}_{(t+1)\tau}^{(h)}(\cdot))$, where $\tilde{Z}_{(t+1)\tau}^{(h)}(\theta) = \tilde{Y}_{(t+1)\tau}^{(h)}(\theta) - \alpha/\mu = \sum_{l=0}^{t}\varphi_h^{(t-l)\tau/h}u_{(l+1)\tau}^{(h)}(\theta)$, is a.s. asymptotically Césaro uniformly equicontinuous on Θ, i.e.,

$$A_\delta^{(h)} = \overline{\lim_{T\to\infty}}\frac{1}{T}\sum_{t=0}^{T-1}\sup_{\bar{\theta}\in\Theta}\sup_{\theta\in B(\bar{\theta},\delta)}|\tilde{Z}_{(t+1)\tau}^{(h)}(\theta) - \tilde{Z}_{(t+1)\tau}^{(h)}(\bar{\theta})| \to 0,$$

a.s. as $\delta \to 0$ where $B(\bar{\theta},\delta)$ denotes the open ball $\{\theta \in \Theta; \|\theta - \bar{\theta}\| < \delta\}$.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Indeed, we have

$$\tilde{Z}^{(h)}_{(t+1)\tau}(\theta) - \tilde{Z}^{(h)}_{(t+1)\tau}(\bar{\theta}) = \sum_{l=0}^{t} [\varphi_h^{(t-l)\tau/h} - \bar{\varphi}_h^{(t-l)\tau/h}] u^{(h)}_{(l+1)\tau}(\theta) + \bar{\varphi}_h^{(t-l)\tau/h} [u^{(h)}_{(l+1)\tau}(\theta) - u^{(h)}_{(l+1)\tau}(\bar{\theta})].$$

Hence,

$$\begin{aligned} A^{(h)}_{\delta} \leq & \varlimsup_{T\to\infty} \frac{1}{T} \sum_{l=0}^{T-1} \left[\sum_{t=l}^{T-1} \sup_{\bar{\theta}\in\Theta} \sup_{\theta\in B(\bar{\theta},\delta)} |\varphi_h^{(t-l)\tau/h} - \bar{\varphi}_h^{(t-l)\tau/h}| \right] \sup_{\theta\in\Theta} |u^{(h)}_{(l+1)\tau}(\theta)| \\ & + \varlimsup_{T\to\infty} \frac{1}{T} \sum_{l=0}^{T-1} \left[\sum_{t=l}^{T-1} \sup_{\theta\in\Theta} |\bar{\varphi}_h^{(t-l)\tau/h}| \right] \sup_{\bar{\theta}\in\Theta} \sup_{\theta\in B(\bar{\theta},\delta)} |u^{(h)}_{(l+1)\tau}(\theta) - u^{(h)}_{(l+1)\tau}(\bar{\theta})| \\ = & \varlimsup_{T\to\infty} \frac{1}{T} \sum_{l=0}^{T-1} \left[\sum_{t=l}^{T-1} \sup_{\bar{\theta}\in\Theta} |(\bar{\varphi}_h + \delta)^{(t-l)\tau/h} - \bar{\varphi}_h^{(t-l)\tau/h}| \right] \sup_{\theta\in\Theta} |u^{(h)}_{(l+1)\tau}(\theta)| \\ & + \varlimsup_{T\to\infty} \frac{1}{T} \sum_{l=0}^{T-1} \left[\frac{1 - \sup_{\bar{\theta}\in\Theta} |\bar{\varphi}_h|^{(T-l)\tau/h}}{1 - \sup_{\bar{\theta}\in\Theta} |\bar{\varphi}_h|^{\tau/h}} \right] \sup_{\bar{\theta}\in\Theta} \sup_{\theta\in B(\bar{\theta},\delta)} |u^{(h)}_{(l+1)\tau}(\theta) - h^{(h)}_{(l+1)\tau}(\bar{\theta})|. \end{aligned} \tag{B.1}$$

Now assuming that δ is chosen small enough, it can be seen that the sum in t is uniformly bounded in T and goes to zero as δ goes to zero. Moreover,

$$\frac{1}{T} \sum_{l=0}^{T-1} \sup_{\theta\in\Theta} |u^{(h)}_{(l+1)\tau}(\theta)| \leq \frac{1}{T} \left[\sum_{l=0}^{T-1} \sum_{k=0}^{\tau/h-1} |\epsilon_{(l+1)\tau-kh}| \right] \sup_{\theta\in\Theta} |\sigma\sqrt{h}|,$$

and the right-hand side of the inequality converges to a finite number when T goes to infinity. Finally, we have

$$\begin{aligned} & \sup_{\bar{\theta}\in\Theta} \sup_{\theta\in B(\bar{\theta},\delta)} |u^{(h)}_{(l+1)\tau}(\theta) - u^{(h)}_{(l+1)\tau}(\bar{\theta})| \\ & \leq \sum_{k=0}^{\tau/h-1} |\epsilon_{(l+1)\tau-kh}| \sup_{\bar{\theta}\in\Theta} \sup_{\theta\in B(\bar{\theta},\delta)} |\sigma\varphi_h^k - \bar{\sigma}\bar{\varphi}_h^k| \sqrt{h} \\ & < \lambda \sum_{k=0}^{\tau/h-1} |\epsilon_{(l+1)\tau-kh}| \end{aligned}$$

for any positive constant λ (independent of l), by choosing δ sufficiently small. Therefore the second term in the right-hand side of (B.1) is bounded by

$$\frac{2\lambda}{1 - \sup_{\bar{\theta}\in\Theta} |\bar{\varphi}_h|^{\tau/h}} \sum_{k=0}^{\tau/h-1} \varlimsup_{T\to\infty} \frac{1}{T} \sum_{l=0}^{T-1} |\epsilon_{(l+1)\tau-kh}| = \frac{\lambda\tau/h\sqrt{2/\pi}}{1 - \sup_{\theta\in\Theta} |\bar{\varphi}_h|^{\tau/h}} \quad \text{a.s.}$$

from the strong law of large numbers. It follows that $A^{(h)}_{\delta} \to 0$ a.s. as $\delta \to 0$.

To conclude on the uniform convergence of $\overline{Y}_T$, it remains to check the following uniform integrability-type condition:

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

$$\overline{\lim_{T\to\infty}}\ \frac{1}{T}\sum_{t=0}^{T-1} E\left[\left(\sup_{\theta\in\Theta}|\tilde{Z}^{(h)}_{(t+1)\tau}(\theta)|\right)^p \mathbb{1}_{\sup_{\theta\in\Theta}|\tilde{Z}^{(h)}_{(t+1)\tau}(\theta)|>M}\right]\to 0$$

a.s. as $M\to\infty$ for some $p\geq 1$.

Indeed, we have

$$\sup_{\theta\in\Theta}|\tilde{Z}^{(h)}_{(t+1)\tau}(\theta)| \leq \sqrt{h}\sum_{l=0}^{t}\sum_{k=0}^{\tau/h-1}\sup_{\theta\in\Theta}|\sigma\varphi_h^{(t-l)\tau/h+k}||\epsilon_{(t+1)\tau-kh}|$$

$$\leq \sup_{\theta\in\Theta}|\sigma|\,\frac{2}{1-\sup_{\theta\in\Theta}|\varphi_h|^{\tau/h}}\sum_{k=0}^{\tau/h-1}|\epsilon_{(t+1)\tau-kh}|.$$

Therefore it admits a finite variance. Then, the condition holds since $P[|\epsilon_{(t+1)\tau-kh}|>M]\to 0$ as $M\to\infty$.

Therefore we get

$$\lim_{T\to\infty}\sup_{\theta\in\Theta}|\overline{Y}_T-\overline{Y}|=0 \quad \text{a.s.}$$

It can be proven through similar but more tedious computations that

$$\lim_{T\to\infty}\sup_{\theta\in\Theta}|\overline{Y^2}_T-\overline{Y^2}|=\lim_{T\to\infty}\sup_{\theta\in\Theta}|\overline{(YY_1)}_T-\overline{YY_1}|=0 \quad \text{a.s.}$$

Moreover, we have

$$(\hat{\mu}^*_T(\tilde{\underline{Y}}^{\tau(h)}_T(\theta))-\tilde{\mu}^*_0(\theta,h))\tau$$

$$=\frac{[\overline{Y^2}_T-\overline{Y}_T^2][\overline{YY_1}-\overline{(YY_1)}_T+\overline{Y}_T^2-\overline{Y}^2]}{[\overline{Y^2}_T-\overline{Y}_T^2][\overline{Y^2}-\overline{Y}^2]}$$

$$+\frac{[\overline{(YY_1)}_T-\overline{Y}_T^2][\overline{Y}^2-\overline{Y^2}_T+\overline{Y}_T^2-\overline{Y}^2]}{[\overline{Y^2}_T-\overline{Y}_T^2][\overline{Y^2}-\overline{Y}^2]}.$$

Because both numerators are made of products of terms converging uniformly to zero or bounded, it remains to prove that the denominator is uniformly bounded away from zero.

It is easily seen that

$$\inf_{\theta\in\Theta}\lim_{T\to\infty}|\overline{Y^2}_T-\overline{Y}_T^2|\geq\lim_{T\to\infty}\inf_{\theta\in\Theta}|\overline{Y^2}_T-\overline{Y}_T^2|.$$

Moreover, some computations lead to the following expression:

$$|\overline{Y^2}_T-\overline{Y}_T^2|=\frac{1}{T}\sum_{k=0}^{\tau/h-1}\sum_{t=0}^{T-1}a^{(h)}_{k,t}(\theta)\sigma^2h\epsilon^2_{(t+1)\tau-kh}+o_p(1),$$

where the coefficients $a^{(h)}_{k,t}(\theta)$ satisfy $a^{(h)}_{k,t}(\theta)\geq\underline{a}$, for all h,k,t and $\theta\in\Theta$ with $\underline{a}$ a strictly positive constant. Therefore,

$$\lim_{T\to\infty}\inf_{\theta\in\Theta}|\overline{Y^2}_T-\overline{Y}_T^2|\geq\underline{a}\underline{\sigma}^2h\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{k=0}^{\tau/h-1}\epsilon^2_{(t+1)\tau-kh}\geq\underline{a}\underline{\sigma}^2\tau.$$

Hence we have $\lim_{T\to\infty}\sup_{\theta\in\Theta}(\hat{\mu}^*_T(\tilde{\underline{Y}}^{\tau(h)}_T(\theta))-\tilde{\mu}^*_0(\theta,h))=0$ a.s., from which we can deduce the uniform a.s. convergence of $\hat{\alpha}^*_T(\tilde{\underline{Y}}^{\tau(h)}_T(\theta))$ to $\tilde{\alpha}^*_0(\theta,h)$ thanks to the

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

second equation of the first-order conditions stated at the beginning of this Appendix. The uniform a.s. convergence of $\hat{\sigma}_T^{*2}(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))$ to $\tilde{\sigma}_0^{*2}(\theta, h)$ is obtained along similar lines from the third one, which completes the proof.

APPENDIX C. ASYMPTOTIC PROPERTIES OF THE QUASI-INDIRECT ESTIMATOR

The indirect estimator is defined as the solution of the minimum distance problem

$$\hat{\theta}_T(\theta_0, h) = \underset{\theta \in \Theta}{\operatorname{argmin}} \|\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta)) - \hat{\theta}_T^*(Y_{\underline{I}}^{\tau})\|.$$

The first-order conditions associated with this program are

$$\left[\frac{\partial \hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta}\right]_{\theta=\hat{\theta}_T(\theta_0,h)} [\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\hat{\theta}_T(\theta_0, h))) - \hat{\theta}_T^*(Y_{\underline{I}}^{\tau})] = 0. \tag{C.1}$$

The expansion of $\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\cdot))$ around the asymptotic value $\theta_0(\theta_0, h)$ gives, for large T,

$$\begin{aligned}\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\hat{\theta}_T(\theta_0, h)) &= \hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta_0(\theta_0, h))) + \left[\frac{\partial \hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta'}\right]_{\theta=\theta_0(\theta_0,h)} [\hat{\theta}_T(\theta_0, h) - \theta_0(\theta_0, h)] \\ &\quad + o_p(T^{-1/2}),\end{aligned}$$

and (C.1) becomes

$$\begin{aligned}\left[\frac{\partial \hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta}\right]_{\theta=\hat{\theta}_T(\theta_0,h)} \Bigg[&[\hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta_0(\theta_0, h))) - \hat{\theta}_T^*(Y_{\underline{I}}^{\tau})] \\ &+ \left[\frac{\partial \hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta'}\right]_{\theta=\theta_0(\theta_0,h)} [\hat{\theta}_T(\theta_0, h) - \theta_0(\theta_0, h)] \\ &+ o_p(T^{-1/2})\Bigg] = 0.\end{aligned}$$

Another expansion gives

$$\left[\frac{\partial \hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta}\right]_{\theta=\hat{\theta}_T(\theta_0,h)} = \left[\frac{\partial \hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta}\right]_{\theta=\theta_0(\theta_0,h)} + o_p(1)$$

and replacing in the preceding equation

$$\begin{aligned}\hat{\theta}_T(\theta_0, h) - \theta_0(\theta_0, h) = &\left[\frac{\partial \hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta} \frac{\partial \hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta'}\right]^{-1}_{\theta=\theta_0(\theta_0,h)} \\ &\times \left[\frac{\partial \hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta))}{\partial \theta}\right]_{\theta=\theta_0(\theta_0,h)} \\ &\times [\hat{\theta}_T^*(Y_{\underline{I}}^{\tau}) - \hat{\theta}_T^*(\tilde{Y}_{\underline{I}}^{\tau(h)}(\theta_0(\theta_0, h)))] + o_p(1).\end{aligned} \tag{C.2}$$

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

From equation (A.1), we know that

$$\hat{\theta}_T^*(Y_{\underline{T}}^\tau) = \theta_0^*(\theta_0) + T^{-1/2}X_T^*(\theta_0) + o_p(T^{-1/2}), \tag{C.3}$$

and similarly, from equation (A.2),

$$\hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta_0(\theta_0,h))) = \theta_0^*(\theta_0) + T^{-1/2}\tilde{X}_T(\theta_0(\theta_0,h),h) + o_p(T^{-1/2}). \tag{C.4}$$

By difference, using (C.3) and (C.4), we have

$$\hat{\theta}_T^*(Y_{\underline{T}}^\tau) - \hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta_0(\theta_0,h))) = T^{-1/2}[\tilde{X}_T(\theta_0(\theta_0,h),h) - X_T^*(\theta_0)] + o_p(T^{-1/2}). \tag{C.5}$$

Replacing in (C.2), we see that the quasi-indirect estimator fulfills the first-order expansion

$$\hat{\theta}_T(\theta_0,h) = \theta_0(\theta_0,h) + T^{-1/2}X_T(\theta_0,h) + o_p(T^{-1/2}), \tag{C.6}$$

where

$$\begin{aligned} X_T(\theta_0,h) = \sqrt{T}\Bigg[& \frac{\partial\hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta))}{\partial\theta}\frac{\partial\hat{\theta}_T^{*}(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta))}{\partial\theta'}\Bigg]^{-1}_{\theta=\theta_0(\theta_0,h)}\Bigg[\frac{\partial\hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta))}{\partial\theta}\Bigg]_{\theta=\theta_0(\theta_0,h)} \\ & \times\Bigg[\Bigg[\frac{\partial^2 S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta_0(\theta_0,h));\theta^*)}{\partial\theta^*\partial\theta^{*\prime}}\Bigg]^{-1}\Bigg[\frac{\partial S_T(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta_0(\theta_0,h));\theta^*)}{\partial\theta^*}\Bigg] \\ & -\Bigg[\frac{\partial^2 S_T(Y_{\underline{T}}^{\tau};\theta^*)}{\partial\theta^*\partial\theta^{*\prime}}\Bigg]^{-1}\Bigg[\frac{\partial S_T(Y_{\underline{T}}^{\tau};\theta^*)}{\partial\theta^*}\Bigg]\Bigg]_{\theta^*=\theta_0^*(\theta_0)}. \end{aligned}$$

The asymptotic distribution of $X_T^*(\theta_0)$ is given by a normal $\mathcal{N}(0,\Gamma_0^*)$ independent from $\tilde{X}_T(\theta_0(\theta_0,h),h)$ which converges, for $T\to\infty$, in distribution to a normal $\mathcal{N}(0,\tilde{\Gamma}^{(h)}(\theta_0(\theta_0,h)))$. Therefore, if there exists a nonstochastic matrix $V_0^{(h)}$ such that

$$\mathbb{P}_{\theta_0}\Bigg[\lim_{T\to\infty}\Bigg\|\Bigg[\frac{\partial\hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta))}{\partial\theta}\Bigg]_{\theta=\theta_0(\theta_0,h)} - V_0^{(h)}\Bigg\| = 0\Bigg] = 1,$$

and $V_0^{(h)\prime}V_0^{(h)}$ is nonsingular, the asymptotic covariance matrix of the quasi-indirect estimator is

$$\bar{V}_0^{(h)} = [V_0^{(h)\prime}V_0^{(h)}]^{-1}V_0^{(h)\prime}(\tilde{\Gamma}^{(h)}(\theta_0(\theta_0,h)) + \Gamma_0^*)V_0^{(h)}[V_0^{(h)\prime}V_0^{(h)}]^{-1},$$

where $V_0^{(h)} = \operatorname{plim}_{T\to\infty}[\partial\hat{\theta}_T^{*\prime}(\tilde{Y}_{\underline{T}}^{\tau(h)}(\theta))/\partial\theta]_{\theta=\theta_0(\theta_0,h)}$.

Various consistent estimates of the covariance matrix $\bar{V}_0^{(h)}$ may be computed along the lines described in Gallant and White (1988, Ch. 6), Andrews (1991), and White (1994, Sect. 8.3), noting as shown in Gouriéroux et al. (1993) that differentiating the first-order condition

$$\frac{\partial S_\infty^{(h)}(\theta,\tilde{\theta}_0^*(\theta,h))}{\partial\theta^*} = 0,$$

with respect to θ, the matrix $V_0^{(h)}$ may be rewritten in terms of derivatives of the limit criterion

$$V_0^{(h)} = [\ddot{S}_\infty^{(h)}(\theta,\tilde{\theta}_0^*(\theta,h))]^{-1}_{\theta=\theta_0(\theta_0,h)}\frac{\partial^2 S_\infty^{(h)}(\theta,\tilde{\theta}_0^*(\theta,h))}{\partial\theta^*\partial\theta'}\Bigg|_{\theta=\theta_0(\theta_0,h)}.$$

Therefore, $V_0^{(h)}$ can be estimated by plugging the estimate of $\theta_0(\theta_0,h)$ into the previous expression. Estimates of Γ_0^* and $\tilde{\Gamma}^{(h)}(\theta_0(\theta_0,h))$ are easily obtained along the same lines

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

because they are also expressed in terms of derivatives of the limit criterion (see Appendix A).

APPENDIX D. CHOICE OF THE DISCRETIZATION STEP

We first recall the asymptotic expansions respectively obtained for the direct, simulated, and quasi-indirect estimators (see Appendixes A and C):

$$\begin{cases} \hat{\theta}_T^*(Y_{\underline{T}}^\tau) = \theta_0^*(\theta_0) + T^{-1/2}X_T^*(\theta_0) + o_p(T^{-1/2}), \\ \hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau}{}^{(h)}(\theta)) = \tilde{\theta}_0^*(\theta,h) + T^{-1/2}\tilde{X}_T(\theta,h) + o_p(T^{-1/2}), \\ \hat{\theta}_T(\theta_0,h) = \theta_0(\theta_0,h) + T^{-1/2}X_T(\theta_0,h) + o_p(T^{-1/2}). \end{cases}$$

From similar assumptions as in Appendix A, we have the following expansions for the first-order terms in the last two expressions:

$$\begin{cases} \tilde{\theta}_0^*(\theta,h) = \tilde{\theta}_0^*(\theta,0) + h\left[\dfrac{\partial\tilde{\theta}_0^*(\theta,h)}{\partial h}\right]_{h=0} + o(h), \\ \theta_0(\theta_0,h) = \theta_0 + h\left[\dfrac{\partial\theta_0(\theta_0,h)}{\partial h}\right]_{h=0} + o(h), \end{cases}$$

and equivalently for their stochastic parts,

$$\begin{cases} T^{-1/2}\tilde{X}_T(\theta,h) = T^{-1/2}\tilde{X}_T(\theta,0) + o_p(T^{-1/2}), \\ T^{-1/2}X_T(\theta_0,h) = T^{-1/2}X_T(\theta_0,0) + o_p(T^{-1/2}). \end{cases}$$

From these asymptotic expansions, we get

$$\begin{aligned} \hat{\theta}_T^*(\tilde{Y}_{\underline{T}}^{\tau}{}^{(h)}(\hat{\theta}_T(\theta_0,h))) - \hat{\theta}_T^*(Y_{\underline{T}}^\tau) &= \tilde{\theta}_0^*(\hat{\theta}_T(\theta_0,h),0) + h\left[\frac{\partial\tilde{\theta}_0^*(\theta_0,h)}{\partial h}\right]_{h=0} + T^{-1/2}\tilde{X}_T(\theta_0,0) \\ &\quad - \theta_0^*(\theta_0) - T^{-1/2}X_T^*(\theta_0) + o_p(h) + o_p(T^{-1/2}) \\ &= \tilde{\theta}_0^*(\theta_0,0) - \theta_0^*(\theta_0) + \left[\frac{\partial\tilde{\theta}_0^*(\theta,0)}{\partial\theta'}\right]_{\theta=\theta_0}(\hat{\theta}_T(\theta_0,h) - \theta_0) \\ &\quad + h\left[\frac{\partial\tilde{\theta}_0^*(\theta_0,h)}{\partial h}\right]_{h=0} + T^{-1/2}(\tilde{X}_T(\theta_0,0) - X_T^*(\theta_0)) \\ &\quad + o_p(h) + o_p(T^{-1/2}) \\ &= \left[\frac{\partial\tilde{\theta}_0^*(\theta,0)}{\partial\theta'}\right]_{\theta=\theta_0}(\theta_0(\theta_0,0) - \theta_0 + h\left[\frac{\partial\theta_0(\theta_0,h)}{\partial h}\right]_{h=0} \\ &\quad + T^{-1/2}X_T(\theta_0,0)) + h\left[\frac{\partial\tilde{\theta}_0^*(\theta_0,h)}{\partial h}\right]_{h=0} \\ &\quad + T^{-1/2}(\tilde{X}_T(\theta_0,0) - X_T^*(\theta_0)) + o_p(h) + o_p(T^{-1/2}). \end{aligned}$$

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press

Therefore the first-order condition (C.1) becomes

$$\left[\frac{\partial\tilde{\theta}_0^{*\prime}(\theta,0)}{\partial\theta}\right]_{\theta=\theta_0}\left\{\left[\frac{\partial\tilde{\theta}_0^{*}(\theta,0)}{\partial\theta'}\right]_{\theta=\theta_0}\left(h\left[\frac{\partial\theta_0(\theta_0,h)}{\partial h}\right]_{h=0}+T^{-1/2}X_T(\theta_0,0)\right)\right.$$
$$\left.+h\left[\frac{\partial\tilde{\theta}_0^{*}(\theta_0,h)}{\partial h}\right]_{h=0}+T^{-1/2}(\widetilde{X}_T(\theta_0,0)-X_T^{*}(\theta_0))\right\}$$
$$+o_p(h)+o_p(T^{-1/2})=0.$$

The equation in the nonrandom terms writes

$$\left[\frac{\partial\tilde{\theta}_0^{*\prime}(\theta,0)}{\partial\theta}\right]_{\theta=\theta_0}\left\{\left[\frac{\partial\tilde{\theta}_0^{*}(\theta,0)}{\partial\theta'}\right]_{\theta=\theta_0}\left[\frac{\partial\theta_0(\theta_0,h)}{\partial h}\right]_{h=0}+\left[\frac{\partial\tilde{\theta}_0^{*}(\theta_0,h)}{\partial h}\right]_{h=0}\right\}=0,$$

and for the random terms

$$\left[\frac{\partial\tilde{\theta}_0^{*\prime}(\theta,0)}{\partial\theta}\right]_{\theta=\theta_0}\left\{\left[\frac{\partial\tilde{\theta}_0^{*}(\theta,0)}{\partial\theta'}\right]_{\theta=\theta_0}X_T(\theta_0,0)+(\widetilde{X}_T(\theta_0,0)-X_T^{*}(\theta_0))\right\}=0.$$

Hence we deduce for the constant term

$$\left[\frac{\partial\theta_0(\theta_0,h)}{\partial h}\right]_{h=0}=-\left[\left[\frac{\partial\tilde{\theta}_0^{*\prime}(\theta,0)}{\partial\theta}\right]_{\theta=\theta_0}\left[\frac{\partial\tilde{\theta}_0^{*}(\theta,0)}{\partial\theta'}\right]_{\theta=\theta_0}\right]^{-1}$$
$$\times\left[\left[\frac{\partial\tilde{\theta}_0^{*\prime}(\theta,0)}{\partial\theta}\right]_{\theta=\theta_0}\left[\frac{\partial\tilde{\theta}_0^{*}(\theta_0,h)}{\partial h}\right]_{h=0}\right]$$

and for the random part

$$X_T(\theta_0,0)=\left[\left[\frac{\partial\tilde{\theta}_0^{*\prime}(\theta,0)}{\partial\theta}\right]_{\theta=\theta_0}\left[\frac{\partial\tilde{\theta}_0^{*}(\theta,0)}{\partial\theta'}\right]_{\theta=\theta_0}\right]^{-1}$$
$$\times\left[\frac{\partial\tilde{\theta}_0^{*\prime}(\theta,0)}{\partial\theta}\right]_{\theta=\theta_0}(\widetilde{X}_T^{*}(\theta_0)-\widetilde{X}_T(\theta_0,0)),$$

which has as asymptotic covariance matrix

$$\bar{\bar{V}}_0=2[V_0'V_0]^{-1}V_0'\Gamma_0^{*}V_0[V_0'V_0]^{-1},$$

with $V_0=\lim_{T\to\infty}V_0^{(h(T))}$.

https://doi.org/10.1017/S0266466698142019 Published online by Cambridge University Press