Journal of Econometrics 178 (2014) 539–557

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

An asymptotic analysis of likelihood-based diffusion model selection using high frequency data[☆]

Hwan-sik Choi [a,*], Minsoo Jeong [b], Joon Y. Park [b,c]

[a] Department of Economics, Binghamton University, 4400 Vestal Parkway E., P.O. Box 6000, Binghamton, NY 13902, USA
[b] Department of Economics, Sungkyunkwan University, 25-2 Sungkyunkwan-Ro, Jongnu-Gu, Seoul, 110-745, Republic of Korea
[c] Department of Economics, Indiana University, Wylie Hall, 100 S. Woodlawn, Bloomington, IN 47405, USA

ARTICLE INFO

Article history:
Received 29 March 2012
Received in revised form
12 August 2013
Accepted 16 August 2013
Available online 7 September 2013

JEL classification:
C22
C52
E43
E47
G12

Keywords:
Diffusion
Model selection
High frequency observation
Likelihood ratio
Information criterion
Spot interest rate

ABSTRACT

We provide a new asymptotic analysis of model selection procedure that compares likelihoods of two candidate diffusion models. Our asymptotic analysis relies on two dimensional asymptotic expansions with shrinking sampling interval Δ and increasing sampling span T, and clarifies the different roles of drift and diffusion functions in the selection of diffusion models. In particular, we show that the model with superior diffusion function specification is always preferred to the competing model regardless of their drift specifications if Δ is sufficiently small relative to T. The specifications of drift functions matter only when the models have an identical diffusion specification.

1. Introduction

The model selection test of Vuong (1989) based on the Kullback–Leibler information criterion (KLIC, Kullback and Leibler (1951)) has been quite popular and considered as a standard tool to evaluate the relative adequacy of two nonnested models. To select a better model, it compares the likelihoods of two competing models evaluated at the maximum likelihood estimates of their parameters and tests for the null hypothesis of their observational equivalence. The KLIC-based test is directional in the sense that a model with higher likelihood is preferred to the other. For an example of the applications of the test, the reader is referred to, e.g., Chen and Scott (1993) or Aït-Sahalia and Kimmel (2010). The test has been extended to allow for dependency in observations. Rivers and Vuong (2002) generalize the test using a more general class of divergence measures applicable for serially correlated data, and Choi and Kiefer (2010) improve the size properties of the test by using the heteroskedasticity autocorrelation consistent (HAC) variance estimator with the fixed-b asymptotics proposed earlier by Kiefer and Vogelsang (2002, 2005).

In the paper, we consider the KLIC-based test for the selection of diffusion models. The model selection test is contrasted with the specification test, which tests whether the candidate model is correctly specified. For the specification of a diffusion model, we may use the tests developed earlier by Aït-Sahalia (1996), Hong and Li (2005) and Chen et al. (2008). The specification test, however, assumes that there is a true model and tests specifically if

☆ We are very grateful to a Co-editor and two anonymous referees for many useful comments. We would also like to thank Yoosoon Chang, Zack Miller, and seminar participants at Cornell (Economics), Kansas (Economics), Purdue (Economics and Statistics), and Texas A&M (Economics and Statistics) for helpful discussions. Most of this work was done while Choi was working in the Department of Consumer Science at Purdue University, and he is grateful for the research support from Purdue University. This work was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2013S1A5A2A01018558).

* Corresponding author.
E-mail addresses: hwansik@binghamton.edu (H.-s. Choi), mssjong@yahoo.com (M. Jeong), joon@indiana.edu (J.Y. Park).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.08.036

the candidate model is correct. In contrast, the model selection test presumes that the candidate models are incorrect and helps us to choose a better model in describing any given set of observations. We believe that the model selection test is of a more practical value than the specification test in choosing a diffusion model, since it is considered by many not as a true model but simply as a model adopted for convenience to approximate some complicated dynamics. In fact, for the Eurodollar interest rates, Aït-Sahalia (1996) rejects all parametric diffusion models tested in the paper except for one very general nonlinear diffusion model, and Hong and Li (2005) reject all parametric diffusion models commonly used in practical applications.

This paper, in particular, provides an asymptotic analysis of the KLIC-based test for diffusion model selection relying on high frequency observations. Our asymptotic analysis relies on two dimensional asymptotics with decreasing sampling interval ($\Delta \to 0$) and increasing sampling span ($T \to \infty$), which is in contrast with the conventional asymptotics requiring only sample size to increase ($N \to \infty$). A diffusion model is specified completely by its drift and diffusion functions, and the likelihood function of a diffusion model is given uniquely once the transition is obtained from the specifications of the drift and diffusion functions. Though the specifications of the drift and diffusion functions jointly determine the likelihood, their effects are of different orders of magnitudes. This is clearly unraveled in our two dimensional asymptotics. For stationary diffusions, the magnitudes of the effects on the log-likelihood ratio of the drift and diffusion specifications are respectively of orders T and $N = T/\Delta$. Loosely put, this is because the information on the diffusion function is accumulated at the rate of $N = T/\Delta$, while the corresponding rate for the accumulation of information on the drift function is of order T.

When Δ is small, the diffusion specifications play a dominant role in determining the rankings of the likelihoods. The leading term in the asymptotic expansion of the log-likelihood ratio depends solely on the diffusion specifications, and the drift specifications appear only in the second-order term with all other components consisting of the likelihoods. The second-order term becomes in general negligible compared with the leading term as $\Delta \to 0$, and therefore, the smaller Δ is, the more important the diffusion specifications relative to the drift specifications. In fact, we show that the model with a superior diffusion specification is always preferred regardless of its drift specification as long as $\Delta \to 0$. The drift specifications may influence the rankings of the likelihoods of competing models, only if the diffusion specifications are equivalent. However, even in this case, the drift specifications become unimportant if Δ is small enough relative to T so that $\Delta\sqrt{T} \to 0$. Consequently, unless the diffusion specifications are identical, the drift specifications do not affect the selection of diffusion models if Δ is sufficiently small relative to T.

The drift specifications matter if the two diffusion models have an identical diffusion specification. For the drift specifications, however, the log-likelihood ratio accumulates information at a slower rate depending only on T, not on $N = T/\Delta$, and the power of the test does not increase if, for instance, we collect observations more frequently within a fixed time span. Moreover, the rankings of the likelihoods are affected not only by the drift specifications, but also as much by the common diffusion specification and the approximation methods of transition densities in case that they do not exist in closed-form and have to be approximated. They all appear in the leading term of our asymptotic expansion for the log-likelihood ratio. Note that the KLIC obtained from the exact and approximated densities is not necessarily identical. In general, the commonly used approximations by the Euler and Milstein schemes are indeed not good enough, and do not yield the KLIC equivalent to the true KLIC. In contrast, the approximation of Aït-Sahalia (2002) based on the Hermite expansion generally provides the same KLIC as the KLIC based on the exact transition up to the second-order terms of our asymptotic expansions.

Under stationarity and other technical regularity conditions, we show that the log-likelihood ratio with an appropriate normalization has normal limit distribution. For the actual implementation of a test for model selection, we standardize the log-likelihood ratio with the Heteroskedasticity Autocorrelation Consistent (HAC) variance estimator and use the fixed-b asymptotics. This is to avoid the dependency of the test on the choice of bandwidth parameter and to more properly account for its variability under the null hypothesis. To demonstrate the finite sample properties of our standardized test statistic, we conduct a Monte Carlo study with two sets of competing models. In each set, the competing models are equivalent in the KLIC, but one set has different diffusion specifications, and the other has an identical diffusion specification. In both examples, the actual rejection rates at the 5% level are reasonably good. For the competing models with equivalent but different diffusion specifications, the actual sizes of the tests are 6%–7% at the daily sampling frequency. For the models with an identical diffusion specification, on the other hand, the rejection rates are 5%–11%.

We apply our test for the selection of a spot interest rate model. In particular, we compare the spot interest rate models introduced by Cox et al. (1985) and Ahn and Gao (1999), which are referred hereafter to as the CIR and AG models. The constant elasticity of variance (CEV) model, which nests the CIR model as a special case, is also considered. Our KLIC-based test finds that the CIR model generally outperforms the AG model for all of the spot rates we consider in the paper. We take this as the evidence that the diffusion specification of the CIR model, $\sigma(x) = \beta\sqrt{x}$ with $\beta > 0$, is superior to the diffusion specification of the AG model, $\sigma(x) = \beta x^{3/2}$ with $\beta > 0$. When we use the square root diffusion function in both candidate models to compare their drift functions, the linear drift specification $\mu(x) = \alpha_2(\alpha_1 - x)$ with $\alpha_1, \alpha_2 > 0$ of the CIR model appears to be significantly better than the quadratic drift specification $\mu(x) = \alpha_2(\alpha_1 - x)x$ with $\alpha_1, \alpha_2 > 0$ of the AG model, at least for some spot rates at all frequencies. Overall, it appears that the CIR model with the square root diffusion and linear drift functions seems to be better supported by the data as a model for the spot interest rates.

The rest of the paper is organized as follows. In Section 2, we introduce a model selection procedure that compares the likelihoods of two candidate diffusion models. The asymptotic expansions of the log-likelihood ratio are developed subsequently in Section 3 to study the limit behavior of the model selection procedure. Our expansions provide very useful information including, in particular, how the drift and diffusion specifications affect the log-likelihood ratio in the limit as the sampling interval and sampling span change. All terms in our expansions are presented explicitly in terms of the drift and diffusion functions, and their derivatives. Section 4 provides a test for model selection based on the log-likelihood ratio. The test uses the HAC variance estimator, and its limit distribution is obtained under the fixed-b asymptotics. It is demonstrated in Section 5 that our test performs reasonably well in a finite sample. We use our test to select a spot interest rate model as an empirical application, which is reported in Section 6. Section 7 concludes the paper, and all the mathematical proofs are in Appendices.

2. Diffusion model selection

Let X be a diffusion on $\mathcal{D} \subset \mathbb{R}$ given as a solution to the stochastic differential equation (SDE)

$$dX_t = \mu_0(X_t)dt + \sigma_0(X_t)dW_t, \tag{2.1}$$

where W is the standard Brownian motion, and $\mu_0(\cdot)$ and $\sigma_0(\cdot)$ are respectively the drift and diffusion functions. Throughout