Journal of Econometrics 164 (2011) 382–403

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

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

# Semi-nonparametric estimation and misspecification testing of diffusion models

Dennis Kristensen *

*Columbia University, United States*
*CREATES, United States*[1]

ARTICLE INFO

*Article history:*
Available online 14 July 2011

*JEL classification:*
C12
C13
C14
C22

*Keywords:*
Diffusion process
Kernel estimation
Nonparametric
Specification testing
Semiparametric
Transition density

ABSTRACT

Novel transition-based misspecification tests of semiparametric and fully parametric univariate diffusion models based on the estimators developed in [Kristensen, D., 2010. Pseudo-maximum likelihood estimation in two classes of semiparametric diffusion models. Journal of Econometrics 156, 239–259] are proposed. It is demonstrated that transition-based tests in general lack power in detecting certain departures from the null since they integrate out local features of the drift and volatility. As a solution to this, tests that directly compare drift and volatility estimators under the relevant null and alternative are also developed which exhibit better power against local alternatives.



## 1. Introduction

In this study, we develop semi-nonparametric estimators and misspecification tests of the so-called drift and diffusion functions in univariate diffusion models given low-frequency observations. The proposed estimators and tests provide the researcher with tools to investigate whether a given parametric specification of the drift and diffusion function is correct and allows him to test drift and diffusion specifications separately from each other. This is in contrast to existing methods found in the literature which simultaneously test correct specification of drift and diffusion terms.

Our estimation and testing procedure takes, as starting point, two classes of semiparametric diffusion models introduced in Kristensen (2010): in the first class, the drift term is known up to a finite-dimensional parameter while the diffusion term is left unspecified; in the second class, the diffusion term is of the parametric form while the drift term is unknown. Kristensen (2010) develops estimators of the parametric component for a given model in either of the two classes. We demonstrate how the unspecified term in any of these semiparametric diffusion models can be estimated nonparametrically using kernel methods. These estimators are useful as guides in the search for a correct parametric specification since they provide information about the shape of the unspecified term. In addition, the estimators help us to develop novel misspecification tests of diffusion models.

We suggest two sets of tests. First, we propose tests for a given semiparametric diffusion model against a fully nonparametric alternative. Second, tests for a fully parametric model against either of its two semiparametric alternatives are developed. Our test statistics are chosen as weighted $L_2$-distances of the estimators of the so-called transition density obtained under null and alternative, respectively. In addition, we also consider tests that directly compare drift or diffusion estimators. We explore the asymptotic properties of the tests both under null and alternative, and obtain a number of interesting results.

First, our transition-based test of a given semiparametric model against the fully nonparametric alternative is under the null first-order asymptotically equivalent to tests of fully parametric models as developed in Aït-Sahalia et al. (2009) and Li and Tkacz (2006). This is due to the fact that estimators of the transition density under the semiparametric and parametric null, respectively, both converge with parametric rate, and as such the asymptotic distributions of our test statistics are completely driven by the fully nonparametric transition density estimator. The parametric rate of the semiparametric transition density estimator appears because computation of transition densities for low-frequency observations involves integration of both the drift and the diffusion term (see e.g. Kristensen, 2008). This integration functions as an additional smoothing mechanism that speeds up convergence rate

* Corresponding address: Department of Economics, International Affairs Building, MC 3308, 420 West 118th Street, New York, NY 10027, United States.
*E-mail address:* dk2313@columbia.edu.

[1] Center for Research in Econometric Analysis of Time Series, funded by the Danish National Research Foundation.

0304-4076/$ – see front matter 
doi:10.1016/j.jeconom.2011.07.001

of the semiparametric estimator of the transition density even though it involves kernel estimators.

Second, our proposed transition-based test statistics of the fully parametric model against either of the two semiparametric alternatives converge with parametric rate under the null. This is non-standard within the class of tests based on $L_2$-distance measures of semi-nonparametric density estimators which in general converge with nonparametric rate. Instead our transition-based tests for the fully parametric null share similarities with the class semiparametric estimators and tests that exhibit parametric rate (see e.g. Andrews, 1994; Corradi and Swanson, 2005; Whang and Andrews, 1993).

Third, we study the power properties of the tests by considering their performance under contiguous alternatives. In particular, we show that, transition-based tests are not very suitable for detection of high-frequency departures in a diffusion framework. This is in contrast to density-based tests in standard, discrete-time setting. This maybe surprising result is due to the fact that local features of the drift and diffusion terms are integrated out in the computation of transition densities and so local deviations get blurred out. It should be stressed that this problem is not special to our particular tests, but is shared by all other transition-based tests of diffusion models in the literature such as Aït-Sahalia et al. (2009). As such our power analysis should be of general interest.

The lack of power against local alternatives leads us to propose two alternative tests of the parametric null against semiparametric alternatives based on direct comparison of drift and diffusion function estimators obtained under null and alternatives. We examine their asymptotic properties both under null and alternative: They converge with a slower rate than the transition-based tests, and thus are dominated by transition-based tests in terms of detecting global alternatives. On the other hand, the tests are better at detecting local deviations of drift and diffusion functions from the null, and so have better power against local alternatives. As such they complement our transition-based tests.

Finally, we conduct a higher-order analysis of the proposed tests under the null. This analysis demonstrates that first-order asymptotic distributions obtained under the null may be a poor proxy of their finite-sample distributions. We therefore propose a Markov bootstrap method that we hope will provide a better approximation of finite-sample distributions of the test statistics. This conjecture is supported by simulation results in Aït-Sahalia et al. (2009) and Li and Tkacz (2006) who propose similar Bootstrap procedures for their tests.

The proposed tests and their theoretical analysis add to a growing literature on specification testing of diffusion models. This class of models is widely used in describing dynamics of asset pricing variables such as interest rates, stock prices, and exchange rates; see for example Björk (2004) for an overview. Since economic theory imposes little restrictions on asset price dynamics, statistical techniques are usually employed in the search for a correct specification. The literature on testing diffusion model specifications can roughly be divided up into two categories depending on whether high-frequency data is assumed available or not.

If high-frequency data is observed, simple nonparametric kernel-regression estimators of drift and diffusion terms can be used to test for correct specification (Bandi and Phillips, 2005; Corradi and White, 1999; Li, 2007; Negri and Nishiyama, 2009). In principle, these tests do not rely on stationarity which is an advantage over the approach taken here. On the other hand, asymptotic properties of estimators and associated tests do rely on the time distance between observations shrinking to zero; thus, estimators and tests will potentially be severely biased if only low-frequency data is available (see Nicolau, 2003).

To avoid the bias issues associated with high-frequency based tests, alternative tests based on fixed time distance between observations have been developed. Aït-Sahalia (1996b) propose to test for correct specification using a weighted $L_2$-distance to measure discrepancies between the marginal density under null and alternative. This class of tests was originally proposed in Bickel and Rosenblatt (1973) in a cross-sectional setting; see also Fan (1994) and Gouriéroux and Tenreiro (2001). Since the test of Aït-Sahalia (1996b) is only able to detect discrepancies in the marginal density, it is not consistent against all alternatives. This observation lead to the development of tests based on transition densities since these fully characterise diffusion models.

Our transition-based tests are most related to the ones developed in Aït-Sahalia et al. (2009) and Li and Tkacz (2006) where fully nonparametric and parametric estimators of the transition density are compared. In a similar spirit, Hong and Li (2005) propose a test where transformed versions of the transition densities are compared, while Chen et al. (2009) employ empirical likelihood techniques. These tests are all designed to examine the correct parametric specification of the drift and diffusion function jointly. In contrast, we are able to test the specification of each of the two functions characterising the model separately. Our local power analysis complements the one carried out in Aït-Sahalia et al. (2009). They specify alternatives in terms of the transition densities and find that transition-based tests have the ability to detect local deviations form the null at a better rate than CvM type tests. However, given that the end goal is to test for the correct specification of drift and diffusion term, we instead specify our alternatives directly in terms of these. By doing so, we obtain some rather different power results for transition-based tests. In particular, we show that they are not able to detect local alternatives at a higher rate compared to CvM type tests. These seemingly contradictory results are due to the fact that Aït-Sahalia et al. (2009) specify their alternatives in terms of the transition density while we focus on deviations in terms of underlying drift and diffusion functions. Since, as already noted above, the transition density involves integration over the drift and diffusion function, local features in these get smoothed out in the transition density and therefore not easily detected.

Our tests based on direct comparison of the drift and diffusion function estimates under null and alternative are related to the marginal density tests of Aït-Sahalia (1996b) and Huang (1997). However, our proposed tests involve non-trivial transformations of the marginal density and its derivatives and as such are able to detect different, more natural alternatives compared to their tests.

Instead of comparing transition densities, Kolmogorov–Smirnoff (KS) type tests have been proposed by Bhardwaj et al. (2008) and Corradi and Swanson (2005) where estimators of the cumulative distribution functions (cdf's) are compared. This on one hand means that their tests converge with parametric rate under the null and as such are more powerful at detecting certain global alternatives compared to transition-based tests. On the other hand KS-type tests are known to have difficulties detecting local deviations from the null; a shortcoming that density-based tests do not suffer from (see e.g. Escanciano, 2009; Eubank and LaRiccia, 1992).

Finally, Kristensen (2010) proposes some specification tests which appear to be the only existing tests based on low-frequency data that allow for testing correct specifications of the drift and diffusion terms separately. However, Kristensen (2010) does not supply a complete asymptotic theory. Moreover, as with CvM and KS type tests, his proposed Hausmann-type tests of fully parametric models will in general have low power against local alternatives since they are based on only matching estimators of the parametric component obtained under null and under alternatives. In particular, his tests may not be consistent against all alternatives. In contrast, we base our tests on estimators of the nonparametric component under the alternative, and so expect them to enjoy better power properties.