

Continuous-Time Volatility Regression in Large Panels*

Congshan Zhang[†]

October 21, 2019

JOB MARKET PAPER

For the latest draft, click [here](#).

Abstract

This paper studies a semiparametric inference procedure for a finite-dimensional parameter in a continuous-time regression model involving high-frequency data in a large cross-section. The model concerns the relationship between a noisy dependent process and a possibly nonlinear transform of stochastic volatility over a *fixed* time span, with its coefficients allowed to depend on a set of firm-specific characteristics. The construction of the estimator involves two steps: the nonparametric recovery of stochastic volatility processes, followed by a parametric second stage that uses the volatility estimates. I show that the estimator admits a central limit theorem and provide a consistent estimator of the asymptotic conditional variance based on a factor-analytic method. The finite sample performance of the inference procedure is satisfactory in a realistically calibrated Monte Carlo setting. In a novel empirical application, I study the relationship between bid-ask spread and the spot standard deviation of asset price. The slope coefficient estimate, which measures the heterogeneous level of information asymmetry, is closely related to firm characteristics such as size, algorithmic trading proxies, and institutional ownership.

Keywords: high-frequency data; factor analysis; volatility; microstructure noise; bid-ask spread; asymmetric information.

*I am deeply grateful to my advisor Jia Li for his insightful guidance and continuous support. I have benefited from comments and advice by Tim Bollerslev, Federico Bugni, Peter Hansen, Andrew Patton, Adam Rosen, George Tauchen and Brian Weller. I would also like to thank participants at the 2019 AMES conference for helpful discussions and comments. Any remaining errors are mine.

[†]Department of Economics, Duke University; email: congshan.zhang@duke.edu.