

Between Data Cleaning and Inference: Pre-Averaging and Robust Estimators of the Efficient Price

Per A. Mykland*
The University of Chicago
Chicago, USA
mykland@pascal.uchicago.edu

Lan Zhang
University of Illinois at Chicago
Chicago, USA
lanzhang@uic.edu

Pre-averaging is a popular strategy for mitigating microstructure in high frequency financial data. As the term suggests, transaction data (say) are averaged over short time periods ranging from 30 seconds to five minutes, and the resulting averages approximate the efficient price process much better than the raw data. – In this paper, we argue that one can improve the robustness of the procedure by instead using M-estimation to pin down the efficient price in local neighborhoods. We argue that it can be analyzed in a way that is similar to pre-averaging. The methodology applies off-the-shelf to any high frequency econometric problem. – Estimating the efficient price is a form of pre-processing of the data, and hence the methods in this paper also serve the purpose of data cleaning.

Keywords: high frequency data, M-estimation, microstructure noise