

High Frequency Data – an Essential Resource

By Richard Olsen, Olsen Ltd and OANDA Corp.

Analysis of economic trends and research of economic processes requires access to quality data. Big investments have gone into building extensive databases of macroeconomic data. So far, high frequency or tick by tick market data has not been included in these efforts. Traditionally, economists have been of the opinion that short-term price fluctuations are irrelevant noise and not worthwhile to collect. Recent discoveries in finance have changed this assessment¹. High frequency data has a high information content and is indicative of long-term trends. This is a plea to expand data collection to high frequency market data as an invaluable source of information.

During the course of the past 20 years, there has been a growing interest in the study of high frequency market data. A rich structure of statistical properties was discovered. Unlike long-term macroeconomic data, which is known to have Gaussian distribution properties, high frequency market data is non-stationary and increasingly fat-tailed over shorter intervals. It has also fractal properties. It is subject to a scaling law, where the average absolute price change increases by the same percentage from a ten to a twenty minutes interval, as from a one to two hour, or a one to two days or one to two month interval. Absolute price changes exhibit long-term autocorrelation properties, where a price shock lasting for only twenty minutes for example has an impact for six or more weeks. There are other properties as well, such as a 24 hour volatility pattern, which is the result of the three time zones of trading in Asia, Europe and America.

Classical economics assumes that financial markets are homogeneous and that short-term price movements follow a Gaussian random walk. The complex structure of statistical properties discovered with high frequency data is the result of the heterogeneity of financial markets². Market participants trade with different time horizons, some take positions for only minutes, others for hours, days, weeks or months. Depending on their trading horizon, they react differently to the same news events. The heterogeneity has the effect of creating a strong temporal interdependence of price moves.

With modern mathematical tools it is possible to analyze high frequency data and extract information of long-term trends and trend changes. We propose to expand data collection and create a global tick by tick market data repository. This repository would also include 'synthetic data' for data that is not directly

¹ *An Introduction to High-Frequency Finance*. Michael M. Dacorogna, Ramazan Gençay, Ulrich Müller, Richard B. Olsen, and Olivier V. Pictet. San Diego, CA: Academic Press, 2001. 383 pp., ISBN: 0-12-279671-3

² Olsen, R. (2000). The Fallacy of the Invisible Hand, in *Visions of Risk*, edited by Carol Alexander, Pearson Education

observable, such as yield curves and volatility surfaces. Today, there exists no commercial database that fulfills this task. It has been estimated that such a project would cost anywhere between 10 and 50 Mio USD.

Its usage would be manifold. First of all, it would provide us with a detailed record of how events unfolded. For historians, economists and finance specialists, it would be an invaluable resource to understand the history of events and get a deeper understanding of the diversity of phenomena that occur. Research of market liquidity would play a prominent role. Transparent and liquid capital markets constitute a "public good", which can only be safeguarded, if we have a deeper understanding of how the markets function.

The tick data time series repository could be used to feed a global early warning system that would operate similarly to a weather forecasting system providing predictive information of financial markets and the economy as a whole. Unlike existing market analysis which is fed by macroeconomic data that has a time delay due to the sparse underlying data, the global early warning system would be online and up to date. At the same time, the data repository could fulfill straight forward tasks – it would allow financial institutions to validate the transaction prices of their complex derivatives transactions by an independent third party resource and prevent losses, such as occurred with Allied Irish Bank.