WHAT'S NEXT IN THE PHYSICS OF MARKETS ?

FRÉDÉRIC ABERGEL*

Over the recent past, a belief has emerged that the socio-economic artefacts of human behaviour could actually be modelled and understood using statistical physics. Questions such as wealth or income distribution, traffic control, epidemics or even geopolitics are now considered attainable by such tools and concepts as the kinetic theory of gases, critical point theory and the like. In particular, financial markets, being systems with many interacting agents, are sometimes believed to exhibit some universal statistical properties. It is indeed true that striking similarities between financial data and statistical mechanics have long been identified. For instance, the original work of the French mathematician Bachelier in his 1900 Thèse de doctorat, where the random walk process was actually constructed in an explicit fashion for the first time in the history of science, found its motivations and applications in the dynamics of traded assets. Treading the steps of Bachelier, a long line of economic, econometric, mathematical and, more recently, physical literature followed, and one can safely ascertain that we now have better and richer descriptions of the statistical properties of financial markets.

Of course, some practical-minded people will inevitably ask the following question: who cares? Because it is also clear that the predictability of markets has not increased over the past century. Markets behave in a probabilistic way, the occurrence of an upward or downward move is still impossible to foresee, and many identified and well-documented phenomena such as volatility clustering, leverage, autocorrelation of orders… remain useless when trying to predict the sign of the next market move.

Of course, predictability in itself is not advertised by researchers in econophysics - although it would be extremely useful. But the predictability of some statistical - as opposed to pointwise or rather, "eventwise" in the probabilistic sense - properties is very important, in that it allows statistically-based decisions that may lead, on average, to winning strategies, robust risk hedging, efficient market-making,… Therefore, it is quite understandable - and commandable - that the modelling and descriptive efforts put on financial data have not decreased but rather, thanks to the host of high frequency, "tick-by-tick" data now available for researchers, increased dramatically. However, one must admit that some of the earlier results from econophysics - the so-called "stylized facts" - are deemed only moderately interesting by keen observers of the financial markets. A primary reason is that a purely statistical study of data may not lead to any understanding of the underlying physical, economical or even psychological mechanisms. Of course, one could argue that this is not a real problem - after all, one never demanded from Boltzman that he understand the intimate nature of a gas particle; that he be able to predict the macroscopic properties from Newton's law of motion seemed, at the time, more than enough. And in fact, it may be ultimately a matter of personal taste, to decide whether the better models should remain purely statistical, or should incorporate some explicit description of the motivations of "agents". But what is absolutely clear is that correctly choosing the type of data is necessary in order to provide a realistic model of markets: the most important feature of a good model is that the physical quantities that it considers reflect the true complexity of the phenomenon one is interested in. And the stylized facts, inasmuch as they consider price data only, strongly lack descriptive as well as explanatory power.

Originally, financial time series have been limited to closing prices. In the context of high frequency data that

* Chair of Quantitative Finance, École Centrale Paris.

has emerged over the past 20 years, closing prices have been replaced, rather naturally, by traded prices. But one clearly observes that this substitution is unsatisfactory, since the traded price is the result of a complex process - generally called "price formation" in the microstructure community - the study of which should include order book data (in the case of order-driven markets) or pre-trade quotes (for the less frequent quote-driven markets). Upon studying this order book, one can observe and measure phenomena such as the competition between market takers and makers, the relationship between arrival rates, order book profile and volatility; I will describe such recent results in the sequel of this article.

Going further, one may actually find use for data that are even more specific than standard order book data. A typical, very good example of the type of information which challenges a too simplistic view of financial markets, and for which a more specific set of data is of great help for the researcher, is the apparent contradiction between the absence of autocorrelation for the returns, and the presence of positive autocorrelation in the signs of incoming market orders. This situation potentially leads to consistency issues, because the sign of a price change is very well correlated with the sign of the market order that has caused the change and hence, a purely price-based statistical description may lead to some contradictions. Well, the light has been shed on this two apparently incompatible phenomena[1] : it is now commonly admitted that the autocorrelation of incoming trade signs is a by-product of the best execution strategies of brokers - brokers split their larger orders and send them to the market in smaller chunks, creating some autocorrelation for the signs of the trades. However, these strategies have little or no impact on the autocorrelation function of the returns, because brokers closely watch the order book and avoid sending orders that eat up more than the available liquidity at the best available Bid or Offer price, thereby not affecting directly the price. Such a fine interpretation has been made possible only through a careful analysis of the identities of exchange members, identities that are made available for researchers on some exchanges and help one differentiate between brokers, high frequency traders, investors… This very interesting study would have been impossible to perform using pure price data: more data are required, of course, but also better data and, above all, a curious - if not suspicious - mindset!

Let me now give some examples of properties of high frequency financial data that have recently been studied in the Chair of Quantitative Finance, and that exemplify the minimal level of modelling required in order to obtain a good understanding of real market properties.

As a first example, I would like to emphasize the richness of the "event time" approach and its extensions to several assets.

Going back in time as far away as Clarke's original paper on the price of cotton[2], one can find in the separation between the arrival of a change in price, and the amplitude given a change, a meaningful decomposition of the price process that can naturally be extended to more general events, such as trade arrival, limit order arrival,… In fact, such a decomposition is nothing but artificial, as it clearly represents the microscopic phenomena now driving the majority of markets, namely, order-driven electronic markets. There, orders arrive in a stochastic fashion, and the price change after an order depends on the shape of the order book when the order came. This description has been recently extended to multivariate cases[3], where a multidimensional trade time has been defined. Very interestingly, one observes that the empirical distribution of returns in trade time may be safely argued to be Gaussian, which advocates for a stochastic arrival times for trade coupled with a stationary profile of the order book.

A second relevant example is the evidence of electronic, high-frequency market-making behaviour as highlighted in[4]. In fact, the interaction between the arrival rates of market orders and those of the "decomposition-evaporation" process of limit orders is a rather intricate and important phenomenon, and little was known on its empirical properties or adequate mathematical modelling. What data reveal us is that there is an acceleration of the arrival of new limit orders right after a market order, and that the side of the order book for which this acceleration takes place does not necessarily coincide with the sign of the market order. Said differently, one can think of agents acting as electronic "market makers", renewing liquidity on both sides of the order book - and thereby controlling the Bid-Ask spread - after a market order has taken some of the available liquidity. In mathematical terms, the use of vector autoregressive processes in the vein of Hawkes helps one model and measure such a behaviour.

A third aspect, which may be at first considered a bit more theoretical, is of tremendous importance when viewed in the light of the impressive number of models used in the field of derivatives pricing: it is the question of bridging the gap between continuous-time finance and microstructure modelling. Some recent results[5] confirm in a theoretical way that simple, "zero-intelligence" order book models, actually lead to continuous diffusion processes when viewed at sufficiently large time-scales. New, ongoing work is addressing the obvious extensions of those results to (exogenous) stochastic or (endogenous) local volatility models. Interesting questions such as the existence of stationary states, the speed of convergence towards them, are actually pregnant and give new insights as to the relative time scales for derivative hedging and high frequency trading. Needless to say, the fact that there is a discrepancy at the finer scales may lead to interesting practical applications[6].

As a conclusion, and to summarize the approach I advocate, I will simply point out that the understanding of financial markets is greatly enhanced when some "inside information" is included. The two traditional ingredients of econophysics - financial data and agent-based models - can be coupled in a very efficient way, thanks to the richer and more precise data sets available for researchers, and one should try to characterize the behaviour of the agents with models that relate to the full set of information contained in financial data. This seems to be the recent trend in the physics of markets, and it should definitely stay this way. ❒

References

1. P. K. Clark, *A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices, Econometrica*, **41**, 13555, (1973).
2. J-P. Bouchaud, F. Lillo, J.D. Farmer, *How markets slowly digest changes in supply and demand*, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1266681
3. N. Huth, F. Abergel, *The times change: multivariate subordination, empirical facts,* to appear in *Quantitative Finance,* http://ssrn.com/abstract=1361486
4. I.. Muni-Toke, *Market making behaviour in an electronic order book and its impact on the Bid-Ask spread*, submitted for publication, http://arxiv.org/abs/1003.3796
5. A. Jedidi, F. Abergel, *General markovian order book modelling*, to appear in the proceedings of the Econophys-Kolkata V conference
6. F. Abergel, *Intraday correlation hedging*, http://www.institutlouisbachelier.org/risk10/tb1_abergel.pdf