DATA 698: Capstone Project Proposal

Derek G Nokes & Paul Britton

9/11/2019

Context & Motivation

Despite being highly accessible for most classes of global investors, public equity markets pose some significant investment challenges. The high degree of co-movement across single stocks makes investors – particularly those following passive strategies – vulnerable to broad-based declines in equity markets. One of the simplest and most effective strategies employed by active investors to both control the risk associated with broad-based declines, and enhance performance when markets are rising, involves exploiting a well-known stylized fact of equity markets, namely that stocks that are moving strongly in a particular direction tend to continue to move in the same direction (i.e., they possess 'momentum').

Momentum-based systematic investing systems focus on identifying stocks that are moving persistently in a particular direction and taking a position to benefit from that directional movement. The long-only version of such a strategy buys stocks that are rising most persistently and exits long positions when markets reverse. To reduce the performance volatility of such a strategy, instruments that move together can potentially be grouped and bet as though they represent a single 'factor'. The inclusion of a broad set of diverse factors improves strategy performance. The construction of a well-diversified portfolio can thus be facilitated by the effective unsupervised learning of robust time series clusters. The use of diverse clusters in stock selection can potentially significantly reduce the volatility of a portfolio, accelerate the speed of compounding, and thereby significantly increasing long-term growth in wealth.

Hypothesis

Use of a fully systematic approach to managing investments enables an investor to determine the exact responses of their strategies to any conceivable set of market conditions. Sensitivity analysis can be used to identify the set of conditions under which the system will operate within acceptable bounds. The broader the spectrum of market conditions over which a trading system can perform within acceptable performance bounds, the more robust the system.

Our high-level hypothesis is that the robustness of a momentum trading strategy can be broadened, and risk-adjusted performance can be improved, through the introduction of feedback and feedforward risk controls. Feedback risk controls operate to reduce the impact of unpredictable phenomena or events on strategy performance, while feedforward controls exploit regularities in market structure to make local predictions that aid in the enhancement of strategy performance. We aim to exploit the time series characteristics of equity price returns to create momentum strategies exhibiting increased robustness and performance relative to a baseline momentum strategy. More specifically, we postulate that unsupervised learning, particularly time series clustering, can be used to design robust controls that reduce the performance volatility of our momentum strategy and increase long-term growth in wealth.

Initial Literature Review

There exists a vast literature on the empirical characteristics of financial markets - including many comprehensive surveys - documenting extensively the basic stylized facts ([1], [2], 3], [4], [5], [6], [7], [8], and [9]). Literature exists documenting different forms of the 'momentum' effect ([10], [11], and [12]), the time-varying and serially-dependent nature of return volatility [1], and the hierarchical structure of cross-dependence ([13],[14],[15], and [16]). A similarly broad literature also exists on the derivation of financial derivative sensitivities [17]. To price and risk manage products with path-dependent payoffs similar to a momentum strategy, Monte Carlo simulation is often required ([18], [19], and [20]). Despite the link between the analysis of systematic trading strategies and the analysis of replication strategies used to manufacture financial derivative products, little published work exists leveraging the findings in these two areas of research to the analysis of systematic trading strategies [21] and [22]). There is also not much literature about the design of risk controls that exploit return time series characteristics to specifically improve the performance of momentum strategies, although there is some literature documenting performance improvements resulting from the application of basic controls ([23], [24], and [25]). Some research indicates that exploiting the hierarchical structure known to exist in return correlations between financial instruments can improve portfolio performance as compared to classical mean-variance approaches to portfolio allocation [26]. Finally, there is an extensive literature on different time series clustering methods ([27], [28], and [29]).

Solution Summary

In this project, we intend to develop metrics that can quantify the evolution of the state of momentum, volatility, and co-movement in a particular universe of single stocks over time. We will then propose one or more associated feedback or feedforward controls that can be used to enhance the performance of a simple systematic momentum trading strategy and evaluate the performance of the strategy before and after the application of our controls using simulation (i.e., strategy backtesting and/or sensitivity analysis).

Research Question

Our initial broad research question is as follows:

Can the performance, robustness, and overall efficacy of a classical equity momentum strategy operating on the S&P1500 universe be improved through the use of a combination of feedback and feedforward controls? Further, can any improvements be validated by means of well-controlled, causal backtesting and/or sensitivity analysis?

We intend to significantly narrow our research question and eventual thesis once we have completed the first couple of weeks of research and the production of our initial research artifacts.

Research Plan

The rough high-level research plan for this project is as follows:

1. Create baseline equity momentum strategy simulator capable of grid-based parameter search and walk-forward testing. Use Dask or Numba to exploit GPUs or multiple CPUs.
2. Literature review of 1) known stylized facts pertaining to equity returns, and; 2) time series clustering approaches.
3. Briefly illustrate stylized facts, particularly the existence of hierarchical structure in the co-movements of stock returns. Use Random Matrix Theory (RMT) and resampling-based approaches to show statistical significance. This step may also involve diagnostic testing to demonstrate power-law scaling in return volatility and the degree distribution of different types of networks derived from return correlations.
4. Test/evaluate candidate time series clustering methods. We intend to focus on a small subset of partitioning, hierarchical, and model-based methods for our application on single stocks. We expect the research artifacts produced during this step to narrow our focus so that we can select a single approach for use in our paper.
5. Adapt base strategy simulator to test feedback or feedforward control meta-strategies.
6. Evaluate the significance of performance differences between control-enhanced strategies and the base strategy using backtesting and/or sensitivities.

Data Sources

To perform the proposed analysis briefly outlined in this proposal, we require a list of the past and current constituents of the S&P1500 index, along with corresponding business summary, sector, sub-industry, and price data. This data has been acquired from Norgate and is available through a Python API. Although alternative open data sources exist, they are significantly less complete and accurate, and would lead to significant survivorship bias in our simulation results. Instead, we propose normalizing/anonymizing our data in such a way as to adhere to our data license to ensure the reproducibility of our project from our public repository.

Scope

Despite the significant scope of this proposed work, both of our group members have considerable experience in this area and also worked together professionally on related problems for nearly 9 years. Derek has published one journal article illustrating how sensitivities can be computed using Monte Carlo techniques and employed in the evaluation of systematic trading strategies [22].

References

[1] T. Bollerslev, R. Y. Chou, and K. F. Kroner, "ARCH modeling in finance: A review of the theory and empirical evidence," *J. Econom.*, vol. 52, nos. 1-2, pp. 5–59, 1992.

[2] W. A. Brock and P. de Lima, "Nonlinear time series, complexity theory, and finance, forthcoming in g. Maddala, and c. Rao (eds.), handbook of statistics volume 14: Statistical methods in finance." New York: North Holland, 1996.

[3] N. Shephard, "Statistical aspects of ARCH and stochastic volatility," in *Time series models*, D. R. Cox, D. V. Hinkley, and O. E. Barndorff-Nielsen, Eds. Boston, MA: Springer US, 1996, pp. 1–67.

[4] C. R. Rao and G. S. Maddala, *Handbook of statistics: Statistical methods in finance*, vol. 14. 1996.

[5] A. Pagan, "The econometrics of financial markets," *Journal of Empirical Finance*, vol. 3, no. 1, pp. 15–102, 1996.

[6] R. Cont, "Empirical properties of asset returns: Stylized facts and statistical issues," *Quant. Finance*, vol. 1, no. 2, pp. 223–236, 2001.

[7] C. Gourieroux and J. Jasiak, *Financial econometrics: Problems, models, and methods*, vol. 1. Princeton University Press Princeton, NJ, 2001.

[8] J. D. Farmer and J. Geanakoplos, "The virtues and vices of equilibrium and the future of financial economics," *Complexity*, vol. 14, no. 3, pp. 11–38, 2009.

[9] G. Marti, F. Nielsen, M. Bińkowski, and P. Donnat, "A review of two decades of correlations, hierarchies, networks and clustering in financial markets," Mar. 2017.

[10] N. Jegadeesh and S. Titman, "Returns to buying winners and selling losers: Implications for stock market efficiency," *The Journal of Finance*, vol. 48. p. 65, 1993.

[11] K. G. Rouwenhorst, "International momentum strategies," *J. Finance*, 1998.

[12] T. J. Moskowitz, Y. H. Ooi, and L. H. Pedersen, "Time series momentum," *J. financ. econ.*, vol. 104, no. 2, pp. 228–250, May 2012.

[13] R. N. Mantegna - The European Physical Journal B-Condensed Matter and 1999, "Hierarchical structure in financial markets," *Springer*, 1999.

[14] M. Tumminello, F. Lillo, and R. N. Mantegna, "Hierarchically nested factor model from multivariate data," Nov. 2005.

[15] D. J. Fenn, M. A. Porter, S. Williams, M. McDonald, N. F. Johnson, and N. S. Jones, “Temporal evolution of financial-market correlations,” *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, vol. 84, no. 2 Pt 2, p. 026109, Aug. 2011.

[16] G. Caldarelli and A. Chessa, “Financial networks,” *Data Science and Complex Networks*. pp. 91–102, 2016.

[17] R. C. Merton and Others, “Theory of rational option pricing,” *Theory of Valuation*, pp. 229–288, 1973.

[18] P. P. Boyle, “Options: A monte carlo approach,” *J. financ. econ.*, vol. 4, no. 3, pp. 323–338, May 1977.

[19] M. Broadie and P. Glasserman, “Estimating security price derivatives using simulation,” *Manage. Sci.*, vol. 42, no. 2, pp. 269–285, Feb. 1996.

[20] F. A. Longstaff and E. S. Schwartz, “Valuing american options by simulation: A simple Least-Squares approach,” *Rev. Financ. Stud.*, vol. 14, no. 1, pp. 113–147, Jan. 2001.

[21] T.-L. Dao, T.-T. Nguyen, C. Deremble, Y. Lempérière, J.-P. Bouchaud, and M. Potters, “Tail protection for long investors: Trend convexity at work,” Jul. 2016.

[22] D. Nokes and L. Fulton, “Analysis of a global futures Trend-Following strategy,” *Journal of Risk and Financial Management*, vol. 12, no. 3, p. 111, Jun. 2019.

[23] P. Barroso and P. Santa-Clara, “Momentum has its moments,” *J. financ. econ.*, vol. 116, no. 1, pp. 111–120, Apr. 2015.

[24] K. Daniel and T. J. Moskowitz, “Momentum crashes,” *J. financ. econ.*, vol. 122, no. 2, pp. 221–247, Nov. 2016.

[25] J. du Plessis, J. du Plessis, and W. G. Hallerbach, “Volatility-Weighting applied to momentum strategies,” *The Journal of Alternative Investments*. 2016.

[26] M. L. de Prado and M. L. de Prado, “Building diversified portfolios that outperform out of sample,” *The Journal of Portfolio Management*, vol. 42. pp. 59–69, 2016.

[27] T. Warren Liao, “Clustering of time series data—a survey,” *Pattern Recognit.*, vol. 38, no. 11, pp. 1857–1874, Nov. 2005.

[28] S. Aghabozorgi, A. S. Shirkhorshidi, and T. Y. Wah, “Time-series clustering–a decade review,” *Inf. Syst.*, vol. 53, pp. 16–38, 2015.

[29] N. S. Madiraju, S. M. Sadat, D. Fisher, and H. Karimabadi, “Deep temporal clustering : Fully unsupervised learning of Time-Domain features,” Feb. 2018.