Int. Fin. Markets, Inst. and Money 33 (2014) 299–316

Contents lists available at ScienceDirect

Journal of International Financial Markets, Institutions & Money

journal homepage: www.elsevier.com/locate/intfin

New evidence about the profitability of small and large stocks and the role of volume obtained using Strongly Typed Genetic Programming

Viktor Manahov [a,*], Robert Hudson [b], Philip Linsley [a]

[a] The University of York, The Management School, York YO10 5GD, Heslington, United Kingdom
[b] The University of Hull, United Kingdom

ARTICLE INFO

Article history:
Received 10 February 2014
Received in revised form 22 August 2014
Accepted 26 August 2014
Available online 8 September 2014

JEL classification:
E37
G1
E3
C1
C12
D84;
C22

Keywords:
Forecasting and simulation
Small Stocks
Agent-based modelling
Artificial stock market
Genetic programming
Capital asset pricing model
Efficiency

ABSTRACT

We employ a special adaptive form of the Strongly Typed Genetic Programming (STGP)-based learning algorithm to develop trading rules based on a survival of the fittest principle. Employing returns data for the Russell 1000, Russell 2000 and Russell 3000 indices the STGP method produces greater returns compared to random walk benchmark forecasts, and the forecasting models are statistically significant in respect of their predictive effectiveness for all three indices both in- and out-of-sample. Using one-step-ahead STGP models to investigate the differences in return patterns between small and large stocks we demonstrate the superiority of models developed for small-cap stocks over those developed for large-cap stocks, indicating that small stocks are more predictable. We also investigate the relationship between trading volume and returns, and find that trading volume has negligible predictive strength, implying it is not advantageous to develop volume-based trading strategies.

1. Introduction

Researchers have devoted considerable efforts to investigate the differences in returns between small-cap stocks and large-cap stocks from both time series and cross-sectional perspectives. Some of this research focuses on the cross-section of returns and, in particular, the issue of the extent to which small-cap stocks yield higher average returns than large-cap stocks (see the seminal papers of Banz (1981) and Reinganum (1981)). Another significant part of this literature has focused on the time series of returns to ascertain whether the returns of small-cap stocks are more predictable than those of large-cap stocks. Relatedly, in the broad context of studies of market predictability, a further major debate is whether trading volume has any predictive power with reference to returns.

* Corresponding author. Tel. +44 1904325847.
E-mail address: viktor.manahov@york.ac.uk (V. Manahov).

http://dx.doi.org/10.1016/j.intfin.2014.08.007
1042-4431/$ – see front matter

This study makes a number of contributions to the existing literature by using a Strongly Typed Genetic Programming (STGP) technique to obtain new evidence about the predictability of returns for small-cap and large-cap stocks, and whether trading volume has any predictive strength. No other study has investigated either of these through the application of STGP modelling. STGP was developed by Montana (1995) and builds on genetic programming (GP). GP is a technique for enabling computer programs to build their own solutions to problems and, consequently, GP is allied to the work that computer scientists have carried out in the related fields of artificial intelligence and machine learning (Poli et al., 2008). GP draws on the idea of adaptive behaviour to make it possible for an initial group of programs to create a next generation of programs, which then create a next generation of programs and so on. Only those programs that produce the best solutions are allowed to 'breed' the next generation. In our experiment trading rules are developed through this evolutionary-based STGP technique. Hence, the trading rules evolve as the generations succeed one another. By using this approach, and allowing all trading rules to evolve incrementally, we avoid the major shortcomings related to data mining associated with many previous studies.

In the experiment returns data for the Russell 1000, Russell 2000 and Russell 3000 indices is used to represent a risky asset (Russell, 2014a,b). We find the STGP method produces greater returns compared to random walk benchmark forecasts, and the forecasting models are statistically significant in respect of their predictive effectiveness for all three indices both in- and out-of-sample. Using one-step-ahead STGP models to investigate the differences in return patterns between small and large stocks we demonstrate the superiority of models developed for small-cap stocks over those developed for large-cap stocks, indicating that small stocks are more predictable. The STGP models for the Russell 2000, small-cap, Index outperform a simple buy-and-hold strategy by greater 4% for both out-of-sample and entire sample.

When we investigate the dynamic relationship between trading volume and index returns to determine whether the level of in-sample trading is a good predictor for out-of-sample stock returns we find that trading volume has negligible predictive strength, implying it is not advantageous developing volume-based trading strategies.

The fundamental premise underlying GP is that as populations reproduce then subsequent populations should be better adapted to survive. The experiments are initially performed based on an opening population of 10,000 traders. The experiments are then re-performed using a smaller population of 1000 traders as this enables us to investigate whether population size impacts on genetic diversity. We find the trading rules that evolve from the smaller population produce lower returns in comparison to a naive buy-and-hold strategy. This suggests that the population size does impact on genetic diversity as the trading rules for the smaller are not sufficiently adapted for the development of effective forecasting models.

The remainder of the article is structured as follows: Section 2 reviews the relevant extant literature as its links to our research design; Section 3 discusses the experimental design; Section 4 presents the results; and the paper concludes with overall remarks and discusses avenues for further research in Section 5.

2. Background

In this section we review the existing literature relevant to our study. In Subsection 2.1 we review the literature looking at the relative predictability of small and large stocks. In Subsection 2.2 we discuss the existing evidence on the link between predictability and volume.

2.1. Studies related to the relative predictability of small and large stocks

A number of prior studies examine the relative predictability of small stocks and find that they are priced in a less efficient manner than large-cap stocks (see, for example, Blume et al., 1994; Avramov, 2002; How et al., 2010). Similarly several studies show that technical trading rules have greater predictive ability for small companies (Bokhari et al., 2005; Hsu et al., 2010; Shynkevich, 2012). Lo and MacKinlay (1990) uncovered a lead-lag pattern which shows that current returns of small stocks are correlated with past returns of large stocks. Several studies attempt to explain the cross-sectional returns of small and large stocks. For instance, Chordia and Swaminathan (2000) observe that large-cap stocks respond rapidly whereas, small-cap stocks respond slowly to marketwide information.

Amihud (2002) shows changes in illiquidity influence small-cap stocks more strongly, which explains time series variations in their return premiums over time. Small-cap stocks pricing inefficiencies may also be explained by the fact that such assets are less widely purchased and consequently do not receive the same level of attention. This means that they are relatively more vulnerable to information asymmetry, experiencing gradual price adjustments because news is absorbed slowly (How et al., 2010).

2.2. Studies related to the volume-return relationship

Ghysels et al. (2000) provide a review of the stylized facts related to the volume-return relationships documented in the literature. Two of these stylized facts are relevant to our study: (1) expected returns are dependent on the volume traded; and (2) there is a highly non-linear relationship between volume and stock returns. The dynamic nature of the volume-return relationship is an important indicator of whether information about trading volume is useful in improving the predictability of stock returns.

While many studies have investigated the empirical and theoretical relationship between stock returns and trading volume, a consensus has not yet been reached regarding such issues as whether the relationship is unidirectional or