Applied Financial Economics, 2006, **16**, 1095–1111

Equity style timing using support vector regressions

Georgi Nalbantov[a,*], Rob Bauer[b] and Ida Sprinkhuizen-Kuyper[c]

[a]*Erasmus Research Institute of Management, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands*
[b]*Department of Finance, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands*
[c]*Nijmegen Institute for Cognition and Information, Montessorilaan 3, Radboud University, 6525 HR Nijmegen, The Netherlands*

The disappointing performance of value and small cap strategies shows that style consistency may not provide the long-term benefits often assumed in the literature. In this study it is examined whether the short-term variation in the US size and value premium is predictable. Style-timing strategies are documented based on technical and (macro-) economic predictors using a recently developed artificial intelligence tool called Support Vector Regressions (SVR). SVR are known for their ability to tackle the standard problem of overfitting, especially in multivariate settings. The findings indicate that both premiums are predictable under fair levels of transaction costs and various forecasting horizons.

I. Introduction

There is no doubt about the importance of investment styles in modern portfolio management. The underlying rationale for this relates to a series of influential studies documenting the potential benefits of investing in stocks with fundamental commonalities or 'styles'. In the past two decades, substantial evidence surfaced suggesting that investing in portfolios of stocks with a small market capitalization and value orientation provides a premium in the long run. The 'size premium' has been first reported by Banz (1981), who found a negative relation between a firm's market capitalization and its stock performance in the USA. The extensively researched 'value premium' has been documented most prominently by Fama and French (1992, 1998) and Lakonishok *et al.* (1994). These studies showed that stocks with typical value features such as low market-to-book (M/B), low price-to-earnings (P/E) and low price-to-cash (P/C) ratios provided higher average returns than so-called 'growth' stocks, with high M/B, P/E and P/C ratios. These empirical findings induced a discussion on the source and magnitude of the value and size premium. Some studies argued that this premium is a compensation for holding stocks under relative distress, see for instance Chan and Chen (1991) and Fama and French (1993). Another view, put forward in Lakonishok *et al.* (1994) and Haugen and Baker (1996), is that stock markets lack efficient pricing ability. A third possible explanation suggested in Lo and MacKinlay (1990) is that the obtained results are due to data snooping biases. A recent review and update for the USA by Chan and Lakonishok (2004) shows that value investing still generates promising returns in the long run. Dimson *et al.* (2003) arrive at similar conclusions for the UK value premium.

The rather disappointing performance of small cap and value strategies during the 1990s has however pointed out that style consistency may not necessarily provide superior returns in any economic regime.

*Corresponding author. E-mail: nalbantov@few.eur.nl

Applied Financial Economics ISSN 0960–3107 print/ISSN 1466–4305 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/09603100500426556

A relatively small body of literature has explicitly addressed the potential benefits of style timing strategies over a style consistent approach. Although most of these papers may differ in methodology, they all rely on the notion that the cyclical behaviour of investment styles is correlated with systematic economic and technical forces, which could make the value and size premium partially predictable. Cooper *et al.* (2001) find sufficient predictability for size-sorted strategies in the USA, but weaker results for value-sorted strategies.[1] Levis and Liodakis (1999) find moderate evidence in favour of small/large rotation strategies, but less evidence for value/growth rotation in the UK. Bauer *et al.* (2004) find evidence for the profitability of style rotation strategies in Japan, but point out that moderate levels of transaction costs can already make these results less interesting in a practical context. The majority of rotation studies employ technical (or market-based) and (macro-) economic indicators. The dependent variables, either the value or the size premium, are constructed using well-known style index series.

In this study a similar approach is used by constructing the value and size premium in the USA based on S&P style indices. The sign and magnitude of both premiums will then subsequently be forecasted using a broad set of (macro-) economic and technical predictors. In contrast to the studies mentioned above, a standard multifactor model framework will not be applied. Factor models in general suffer from deficiencies intrinsic to multiple regression techniques. Most of the studies based on this methodology *ex ante* decide to construct parsimonious models to avoid the problem of over-fitting. Increasing the number of factors at some point will deteriorate the out-of-sample prediction ability of the rotation models. Levis and Liodakis (1999) for instance report empirical results based on six factors for the size spread and eight factors for the value spread. Although their regression window is expanding, thereby updating the relevance of the factors through time, it does not provide the ability to add or delete economically viable factors. In most cases the 'optimal' choice of independent variables is based on a set of statistical criteria, like adjusted R^2, the Akaike information criterion or the Schwarz criterion. These criteria are designed to correct the inclusion of factors for the increased model complexity. Potentially, numerous relevant variables are bound to be excluded as predictors.

A further complication arises from the fact that individual factors in a model are usually assumed to be independent. Most linear regression models however are likely to suffer from multicollinearity, especially when the forecasting variables are numerous and closely related. It could therefore be argued that factor models face two pivotal challenges: first, how to employ a large set of potentially relevant variables in a factor model without jeopardizing its predictive power, and second, how to incorporate possible interactions between individual variables in the course of the model-building process without deteriorating the quality of the model.[2]

Support Vector Regressions (SVR) have become a popular analytical tool following a series of successful applications in fields ranging from optical character recognition to DNA analysis (Smola and Schölkopf, 1998; Müller *et al.*, 2001). In essence, the SVR technique is used for function estimation based on a finite number of observations, just like the linear multiple-regression technique. Numerous potential applications of SVR in finance have been reported elsewhere.[3] The combination of three key features can justify *a priori* the utilization of the SVR tool in financial forecasting modelling. First, SVR behave robustly even in high-dimensional feature problems (Maragoudakis *et al.*, 2002), or in other words, where the explanatory variables are numerous, and in noisy, complex domains (Burbidge and Buxton, 2001). Second, SVR achieve remarkable generalization ability by striking a balance between a certain level of model accuracy on a given training data-set, and model complexity.[4] And third, SVR always find a global solution to a given problem (Vapnik, 1995; Smola, 1996), in sharp contrast with neural networks for instance. A general limitation of SVR is that they produce point estimates rather than posterior probability distributions of the obtained results, which follows from the fact that SVR are a nonparametric tool. Some parameters however have to be estimated in advance via a standard procedure called 'cross-validation'. This procedure, though quite computationally extensive, additionally ensures that

[1] Other related work includes Arnott *et al.* (1989), Arnott *et al.* (1992), Jacobs and Levy (1996), Copeland and Copeland (1999), Kao and Shumaker (1999), Asness *et al.* (2000), Elfakhani (2000), Mun *et al.* (2001), Ahmed *et al.* (2002), Lucas *et al.* (2002) and Mills and Jordanov (2003).
[2] See Pesaran and Timmermann (1995) and Bossaerts and Hillion (1999) for a discussion on these and related issues.
[3] See, e.g. Müller *et al.* (1997), Smola and Schölkopf (1998), Monteiro (2001), Rocco and Moreno (2003), and Van Gestel *et al.* (2003).
[4] Note that in real-world applications the presence of noise in regression estimation necessitates the search for such a balance, see Vapnik (1995) and Woodford (2001).

model selection is based on out-of-sample rather than in-sample performance.

Using SVR models will be constructed in order to predict the value and size premiums in the US stock market. The aim is to test on a preliminary level the performance of SVR, and not to engage in an extensive data-mining exercise. For that reason, the models are built on historical data of 60 months, which is a quite common horizon in the literature. Obviously, other model-building horizons can be explored, but in such a way artificially good results could emerge, falling prey to the data-mining critique. The results of the rotation strategies are compared with so-called style consistent passive strategies. Furthermore, the forecast horizon (one-, three- and six-month signals) is varied, which serves as a model-stability test, and measure the impact of a wide range of transaction costs. The empirical section shows that style rotation strategies using signals created by SVR produce outstanding results for both the value and the size premiums.

The remainder of the paper is organized as follows. Section II discusses the choice of explanatory variables and the nature of the explained variables (the proxies for the value and size premiums). Section III deals extensively with Support Vector Regressions as an analytical tool and how it can be used to predict the value and size premiums. Section IV presents the main empirical findings, and Section V concludes.

II. Data

Construction of the value and size premium series

The choice of an appropriate measure to determine the value premium is crucial. The main goal in this study is to come up with a trading strategy, which can be easily implemented in a practical context.[5] In principle, long time series data from the Center for Research in Security Prices (CRSP) can be used. Following this venue is not well suited for a low transaction cost strategy however, since there are no readily-available instruments (e.g. futures) to exercise such a trading strategy in practice. As it is expected that the rotation strategies will have a considerable turnover, the analysis is conducted on the S&P Barra Value and Growth indices (the value premium). Transaction costs are expected to be relatively low as it is possible to buy and sell futures on these indices.[6] Both indices are constructed by dividing the stocks in the S&P 500 index according to just one single attribute: the book-to-market ratio. This procedure splits the S&P 500 index into two, mutually exclusive groups of stocks and is designed to track these accepted investment styles in the US stock market. The Value index contains firms with high book-to-market ratios and conversely the Growth index firms with lower book-to-market ratios. The combination of both (market cap weighted) indices adds up to the (market cap weighted) S&P 500.

Figure 1 and Table 1 show that a strategy purely based on the value premium would have witnessed some highly volatile periods. These series are the returns of a long position in the Value index and a short position in the Growth index throughout the entire sample period ranging from January 1988 to December 2002. Monthly maximum and minimum returns of this strategy are considerably high: 9.74% and −12.02%. Summary statistics (see Table 1) reveal that the spread series exhibits excess kurtosis. The number of negative performance months of this passive value strategy is approximately 47%. The average return on an annualized basis is −0.86% with a standard deviation of 9.64%. It is therefore concluded that pure and unconditional value investing in this particular sample period has not been a very attractive trading strategy. Furthermore, it is indeed observed that there is a cyclical pattern in the behaviour of the premium. In some periods, like for instance in the last years of the previous decade, growth stocks persistently outperformed value stocks and in other periods value stocks clearly outperformed growth stocks. A good example of the latter is the crisis in Technology (and hence 'growth') stocks in the beginning of this century. A possible explanation for this phenomenon could be that the sign of the value premium is strongly connected with the business cycle and the economic regime. It is likely that value stocks – relative to growth stocks – gain more from a surge of economic activity and a sharp upward revision of sentiment, see e.g. Schwob (2000). As profit expectations turn sharply and broadly positive at the bottom of the economic cycle, profitability and earnings growth become a less scarce resource. In such an environment portfolio managers start looking for stocks with typical value features. This largely explains why value stocks generally belong to cyclical industries. Moreover, value companies tend to belong to mature sectors of the

[5] In the case of for instance the High book-to-market minus Low book-to-market (HML) series of Fama and French (1993), relatively high transaction costs can be expected as portfolios generally exhibit unacceptable liquidity features, particularly in a monthly long/short setting.
[6] In practice the maximum exposure of the trading strategy is still restricted by the liquidity features of this future.

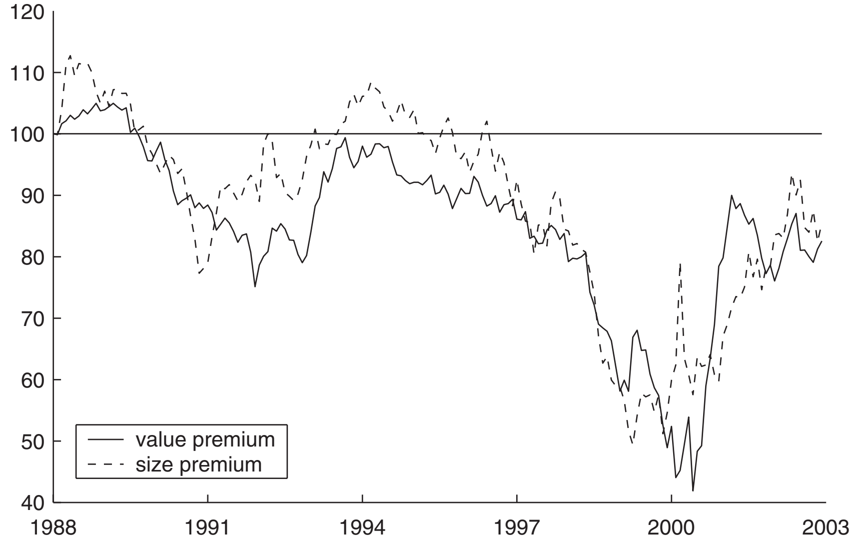

Fig. 1. Cumulative performance of the value and size premiums (1988:01–2002:12)

Table 1. Summary statistics for the value and size premiums (1988:01–2002:12). All numbers are annual data (in %) unless stated otherwise. The spread series for the value premium are computed as returns of a long/short portfolio (long S&P Barra Value index and short S&P Barra Growth index). The spread series for the size premium are computed as returns of a long/short portfolio (long S&P 500 index and short S&P SmallCap 600 index). Prior to the introduction of the S&P SmallCap 600 index in January 1994, the Frank Russell 1000 and Frank Russell 2000 indices have been used as inputs for the small-large calculations

	Value premium	Size premium
Mean	−0.86	−0.91
Standard deviation	9.64	12.04
Information ratio	−0.09	−0.08
Minimum (monthly)	−12.02	−15.71
Maximum (monthly)	9.74	16.78
Skewness (monthly)	0.06	0.27
Excess kurtosis (monthly)	3.15	4.31
% negative months	47.22	51.11

economy. These sectors generally grow and shrink with the economy, whereas growth companies can offer protection during weaker periods in the economy.

Analogously, the size premium series is created by comparing the S&P 500 index (large cap) and the S&P Small Cap 600 index.[7] The passive small-large strategy has not performed satisfactorily during the sample period used as well: a mean return of −0.91% (see Fig. 1 and Table 1). Investors that have followed this strategy have experienced even greater fluctuations than those opting for the passive value-growth strategy, as revealed by the maximum (16.78%) and minimum (−15.71%) monthly returns and the higher standard deviation (12.04%). All of these findings cast serious doubt on the wisdom of persistently favoring small stocks over large stocks in the past two decades.

Choice of the forecasting variables

Two classes of forecasting variables are introduced in this section. First, a brief overview of potential technical variables is given. Subsequently, several macroeconomic variables are addressed, which might shed some light on the behaviour of the spread series. There appears to be a striking similarity between the chronological cumulative performance of the value and size premiums (see Fig. 1), which suggests that the behaviour of both premiums might be subject to the same cyclical effects. The aim is to provide a wide range of relevant forecasting variables, but it is restricted to those claimed to be economically interpretable in the literature on this subject.

Good examples of technical factors are the lagged value and small cap spreads used by Levis and Liodakis (1999). Asness *et al.* (2000) propose two other variables of this class: the spread in valuation multiples and expected earnings growth between value portfolios and growth portfolios. Other candidates are changes in the implied volatility of the

[7] Prior to the introduction of the S&P Small Cap 600 index in January 1994, the Frank Russell 1000 and Frank Russell 2000 indices have been used as inputs for the small-large calculations.

Table 2. Variables used in the style timing models based on Support Vector Regressions

	Technical variables
LagVmG	Lagged value/growth spread
LagSmL	Lagged small/large spread
VOL	Volatility of the S&P 500
FPE	12-month Forward P/E of the S&P 500
MOM	6-month Momentum of the S&P 500
Profit cycle	Year on year change in earnings per share of the S&P 500
PE dif.	Price/earnings difference between value and growth indices, or between S&P 500 and small cap indices
DY dif.	Difference between dividend yields on value and growth indices, or S&P 500 and small cap indices
Economic variables	
Corporate credit spread	The yield spread of (Lehman aggregate) Baa over Aaa
Core inflation	The 12-month trailing change in the US consumer price index
Earnings-yield gap	The difference between the forward E/P ratio of the S&P 500 and the 10-year T-bond yield
Yield curve spread	The yield spread of 10-year T-bonds over 3-month T-bills
Real bond yield	The 10-year T-bond yield adjusted for the 12-month trailing inflation rate
Ind. prod	US industrial production seasonally adjusted
Oil price	The 1-month price change
ISM (MoM)	1-month change of US ISM purchasing managers index (mfg survey)
Leading indicator	The 12-month change in the conference board leading indicator

market, see Copeland and Copeland (1999), and price and earnings momentum in the market, see for instance Bernstein (2001), Miller *et al.* (2001), Kwon and Kish (2002).

The class of economic variables is mainly related to economic fundamentals, the business cycle and trends in corporate earnings. Examples of macroeconomic series can be found in a variety of papers on style rotation. Kao and Shumaker (1999) document the influence of industrial production, the yield-curve spread, inflation (CPI) and the corporate credit spread on the value premium. In their view, industrial production reflects the corporate earnings cycle. In periods of high corporate earnings growth, the often highly leveraged value (and small) companies profit disproportionately. The composite leading indicators (CLI) can serve as an alternative to measure the same relationship. The interest rate environment can also have a substantial impact on the sign of the value premium. A yield spread widening between long government bonds and short term T-bills will probably hurt growth companies more than value companies as their profits are based further into the future. Growth stocks have longer durations than value stocks and are therefore more interest rate sensitive. These companies will underperform most likely in a setting with steep yield curves, which implies rising interest rates in the future. In the study of Levis and Liodakis (1999) the spread series are explained by the level of inflation, changes in the short-term interest rate and the equity risk premium respectively.[8]

In Table 2 we list the variables used in the empirical analysis. In the next section we describe and discuss the non-parametric modelling tool used, namely Support Vector Regressions.

III. Methodology

This section describes the model-building tool (Support Vector Regression) and the construction of the SVR rotation models. Alongside, the qualities of SVR that justify their employment as a factor model tool are focused on.

Function estimation with SVR

Support Vector Regressions (SVR), and Support Vector Machines (SVM) in general, are rooted in Statistical Learning Theory, pioneered by Vapnik

[8] Liew and Vassalou (2000) claim that past style performance can actually function as *a forecast* for economic growth, which brings a new dimension to this literature.

(1995). In essence, SVR are just functions, named 'learning machines', of which the basic task is to 'explore' data (input-output pairs)[9] and provide optimally accurate predictions on unseen data. Extensive descriptions of SVR and SVM can be found, for example, in Smola (1996), Burges (1998), and Smola and Schölkopf (1998). Here a complete, but still compact and accessible representation of the basic SVR tool, is presented. The technical exposition follows mostly the descriptions in the abovementioned papers.

First, it should be mentioned that the standard loss function employed in SVR is the ϵ-insensitive loss function, which has the following form:

$$|y - f(\mathbf{x})|_{\varepsilon} \equiv \max\{, 0|y - f(\mathbf{x})| - \varepsilon\} \quad (1)$$

Here ϵ is predetermined and positive, y is the true target value, $\mathbf{x}$ is a vector of input variables and $f(\mathbf{x})$ is the estimated target value. If the value of the estimate $f(\mathbf{x})$ of y is off-target by ϵ or less, then there is no 'loss', and no penalty will be imposed. However, if $|y - f(\mathbf{x})| - \epsilon > 0$, then the value of the loss function rises linearly with the difference between y and $f(\mathbf{x})$ above ϵ. In practice, the actual loss associated with a given training error is equal to $C(|y - f(\mathbf{x})|_{\epsilon})$, where C is a positive constant. The term $|y - f(\mathbf{x})|_{\epsilon}$ is denoted by ξ if $y \leq f(\mathbf{x}) - \epsilon$, and by ξ^* if $y \geq f(\mathbf{x}) + \epsilon$.

The simplest case of function estimation is considered first, where there is only one input variable, x_1, one output variable, y, and l training data points, and a linear relationship between the input and output variables (see Fig. 2).

Notice that in the case of Fig. 2, the total amount of loss is equal to $C(\xi + \xi^*)$, since there are two training errors. The SVR algorithm estimates the parameters w_1 and b of the linear function $y = w_1x_1 + b$ for prespecified values of ϵ and C, ensuring that the resulting regression function achieves good generalization ability. It should not be too 'complex', but, at the same time, it should not make too many training errors. Complexity here is defined in terms of 'flatness' of the line, i.e. the smaller the slope of the line, the lower the complexity. By striking a balance between the function's complexity and accuracy on the training data in the model-construction phase, the SVR offers a solution to the common problem of overfitting.

Figure 2 considers a one-dimensional input space, i.e. there is only one independent variable. If the dimension of the input space equals n, the optimal regression function $f(\mathbf{x}) = (\mathbf{w} \cdot \mathbf{x}) + b$

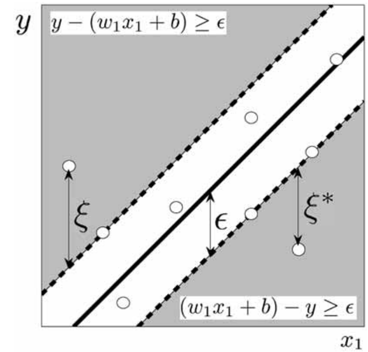

Fig. 2. An SVR solution to the problem of estimating a relation between x_1 and y. All points inside the white region in the figure are within ϵ distance from the solid, optimal regression line $y = w_1x_1 + b$, and therefore are not penalized. However, penalties ξ_i and ξ_i^* are assigned to the two points that lie inside the shaded areas (given by $y - (w_1x_1 + b) \geq \epsilon$ and $(w_1x_1 + b) - y \geq \epsilon$). The optimal regression line is as flat as possible, and strikes a balance between the area of the white region and the amount of points that lie outside this region

one is looking for, with a vector of input variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, 'weight' vector $\mathbf{w} = (w_1, w_2, \ldots, w_n)$, and the inner product $(\mathbf{w} \cdot \mathbf{x}) = w_1x_1 + w_2x_2 + \cdots + w_nx_n$. Flatness in that case is defined in terms of the Euclidean norm of the weight vector: $\|\mathbf{w}\| = (w_1^2 + w_2^2 + \cdots + w_n^2)^{1/2}$. The parameters of the linear SVR $f(\mathbf{x}) = (\mathbf{w} \cdot \mathbf{x}) + b$, i.e. $\mathbf{w}$, b, ξ_i and ξ_i^*, $i = 1,2,\ldots,l$, can be found as the unique solution of the (convex quadratic) optimization problem:

Minimize

$$\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_j^*)$$

Subject to

$$\begin{aligned} & y_i - (\mathbf{w} \cdot \mathbf{x}) - b \leq \epsilon + \xi_i \\ & (\mathbf{w} \cdot \mathbf{x}) + b - y_i \leq \epsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0 \\ & \text{for } i = 1, 2, \ldots, l \end{aligned} \quad (2)$$

The first term of the objective (minimization) function in Equation 2 deals with the complexity, and the second term deals with the accuracy (or, amount of training errors) of the model. In general, both terms cannot be minimal (or, close to zero) at the

[9] The terms 'inputs' and 'outputs' in the machine learning domain stand for the 'independent variables' and the 'dependent variables' in the finance domain.

same time. The positive parameter C determines the trade-off between the flatness of $f(\mathbf{x})$ and the amount of tolerated deviations. If C is large, some flatness could be lost in order to achieve greater training accuracy.

All points on the boundary of the ϵ-insensitive region together with the points outside that region (the training errors) are called 'support vectors'. The computation of the regression is solely based on the support vectors.

The minimization problem of Equation 2 can be represented in dual form, as a maximization problem:

Maximize

$$-\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)k(\mathbf{x}_i,\mathbf{x}_j)$$
$$+\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)y_i-\epsilon\sum_{i=1}^{l}(\alpha_i+\alpha_i^*)$$

Subject to

$$0\le\alpha_i,\alpha_i^*\le C, i=1,2,\ldots,l \text{ and } \sum_{i=1}^{l}(\alpha_i-\alpha_i^*)=0 \qquad (3)$$

where $k(\mathbf{x}_i,\mathbf{x}_j)=(\mathbf{x}_i\cdot\mathbf{x}_j)$. The application of the kernel function $k(\mathbf{x}_i,\mathbf{x}_j)$ instead of the inner product $(\mathbf{x}_i\cdot\mathbf{x}_j)$ provides for the possibility to utilize other functional forms (see below).

In SVR, the regression estimates, which result from solving Equation 3, take the form of:

$$f(\mathbf{x})=\sum_{i=1}^{l}(\alpha_i^*-\alpha_i)k(\mathbf{x},\mathbf{x}_i)+b \qquad (4)$$

The value of b can be found from the so-called Karush–Kuhn–Tucker (KKT) conditions associated with the dual optimization problem (Equation 3). The training points in the series in Equation 4 with coefficient $(\alpha_i^*-\alpha_i)$ unequal to zero are exactly the support vectors. For each training point x_i at most one of the two numbers α_i and α_i^* is unequal to zero. For the training points on the boundary of the ϵ-insensitive region holds either $0<\alpha_i<C$ or $0<\alpha_i^*<C$ and for the training errors outside the ϵ-insensitive region holds either $\alpha_i=C$ or $\alpha_i^*=C$.

Application of a kernel function transforms the original input space implicitly into a higher-dimensional input space where an optimal linear decision surface (corresponding to a non-linear decision surface in the original input space) is found. One of the most frequently applied kernels is the so-called Radial Basis Function kernel (RBF). The dimension of the feature space for the RBF is infinite, which on first sight is counterintuitive from a complexity perspective: that should lead to over-fitting. However, the literature reports very good performance of SVR using the RBF kernel (see, e.g. Burges (1998), Chang *et al.* (2001), and Müller *et al.* (2001)). Possible theoretical explanations thereof have been suggested in Burges (1998). Therefore, it appears that SVR with a RBF kernel are able to tackle the problems of overfitting effectively. For this reason this kernel is applied in this research.

The RBF kernel is defined as $k(\mathbf{x}_i,\mathbf{x}_j)=e^{-\gamma||\mathbf{x}_i-\mathbf{x}_j||^2}$, where γ is a manually adjustable parameter. The Radial Basis Function kernel is equal to 1 if $\|\mathbf{x}_i=\mathbf{x}_j\|$ and drops monotonically to zero with the Euclidean distance $\mathbf{x}_i-\mathbf{x}_j$ between the vectors $\mathbf{x}_i$ and $\mathbf{x}_j$. The greater the value of γ, the faster the function $k(\mathbf{x}_i,\mathbf{x}_j)$ decreases. So, for large values of γ the influence of a training point will be only local and the risk of overfitting will be large. So, the larger γ, the more 'complex' the radial basis function is, and the smaller the number of training errors.

Summarizing, there are three parameters ϵ, C, and γ, which have to be tuned in order to find the optimal trade-off between complexity and training accuracy of the SVR. One of the ways to find the best trade-off between these parameters is via the standard cross-validation technique, which will be explained in the subsection 3.2.

SVR style timing models

Only the construction of the value rotation model will be presented, since the size rotation model is constructed analogically.[10]

The input vectors for the SVR consist of the (historical) values for all 17 candidate explanatory factors as described in Table 2. The outputs are the corresponding differences in returns between the S&P 500 Barra Value and Growth indices. Each SVR model is trained on the data for months $t-60$ till month $t-1$ in order to predict the output for month t. In order to find the optimal model parameters ϵ, C and γ five-fold cross-validation is applied, a standard technique in machine learning (see e.g., Stone (1977) and Weiss and Kulikowski (1991)) on the training data sets of 60 months. A k-fold cross-validation procedure is utilized as follows: a given dataset is divided into k folders of equal size; subsequently, a model is built on all possible (k) combinations of $k-1$ folders, and each

[10] The software program used throughout the analysis is LIBSVM 2.4, developed by Chang and Lin (2002).

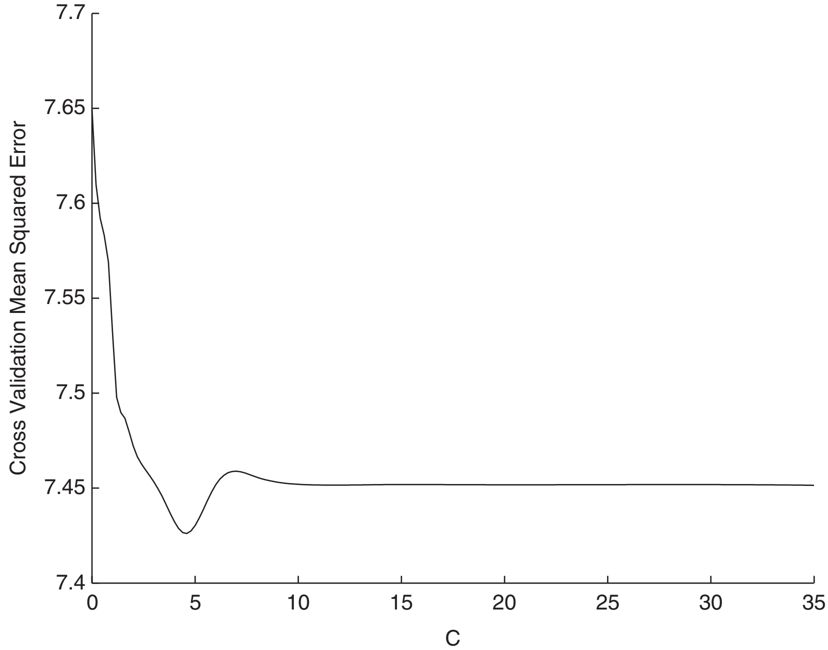

Fig. 3. Five-fold cross validation mean squared errors associated with the penalty-on-error parameter $C \in (0, 35)$ and fixed ϵ-insensitive loss function parameter (ϵ) at 1.0 and Radial Basis Function parameter at 0.007. The to-be-predicted month here is April 2000. The 'best' model is the one for which the combination of the three parameters over suitable parameter ranges produces minimal cross-validation mean squared error

time the remaining one folder is used for validation. The model that achieves minimum mean sum of squared errors on average (over the k validation folders) is considered to be the best. This best model is said to achieve minimum cross-validation mean squared error, and the parameters of this model are used in the final model for the prediction of month t.

The advantage of using a cross-validation procedure is that it ensures that model selection is based entirely on out-of-sample rather than in-sample performance. The disadvantage however is that the procedure is rather time-consuming. A tiny part of the cross-validation procedure is visualized in Fig. 3, where the vertical axis shows the cross-validation minimal squared errors for $C \in (0, 35)$, while keeping ϵ and the kernel function parameter γ fixed at 1.0 and 0.007, respectively. As suggested by the figure, the value for the minimum cross-validation mean squared error is well defined.

The predicted output, i.e. the value premium for month t is used to decide on the timing rotation strategy. A positive output will result in a signal 'Value' in which case the Value index will be bought and the Growth index sold, while a negative output will result in a signal 'Growth' with the opposite effect. In order to avoid taking decisions based on noise, an output value close to zero as a 'no signal' signal.[11]

The SVR small-large strategy is defined analogically, using S&P SmallCap 600 and S&P 500 indices.

IV. Empirical Results

In this section the main results from value-growth and small-large rotation strategies are presented using SVR with different levels of transaction costs and varying forecast horizons (one-month, three-month and six-month). Additionally, the output of an equally weighted combination of both strategies is shown. Throughout this empirical section we show returns that can be achieved when one would have been able to forecast the signal correctly each month: MAX_VG (value-growth rotation) and MAX_SL (small-large rotation). The input for the SVR model consists of 60 months of data on the whole set of 17 predetermined factors. The passive style strategies are constructed in accordance with what is expected in the literature: each month a long position is taken in the Value index and a short position in the Growth index. The passive small-large strategy consistently

[11] A range of (−0.05, 0.05) standard deviations was used relative to the average of the estimates over the training period.

buys the S&P SmallCap 600 index and sells the S&P 500 index.

Value-growth rotation strategies

Detailed results of the SVR value-growth strategy can be found in the left part of Table 3. What strikes most at first sight, is that this strategy has produced much better results than the passive strategy in the out-of-sample period starting January 1993 and ending December 2002.

Under the assumption of zero transaction costs, the SVR strategy achieves an annualized mean return of 10.30%, against a modest 0.24% respective return of the passive strategy. Combining these results with the standard deviations of returns yields (annualized) information ratios of 1.04 and 0.02, respectively. Besides, even when high transaction costs of 50 basis points (bp.) (single trip) are added into the calculations, the realized SVR-model information ratio remains quite high (0.64), and statistically significant at the 5% two-tail level. When compared to other studies on the subject, for example Bauer and Molenaar (2002) in the USA and Levis and Liodakis (1999) in the UK, the SVR results seem to demonstrate a significant improvement. The calculated Z(equality)-scores[12] provide further evidence (in the 0 bp. and 25 bp. transaction-cost environment) of a significant performance difference. In addition, the SVR strategy is able to capture 37.7%, 34.0%, and 29.3% of the return from the MAX_VG strategy under 0 bp., 25 bp. and 50 bp. transaction costs, respectively. Note that in Table 3 only the results of the MAX_VG strategy under 50 bp. transaction costs are given. Table 3 further reveals that the largest three-month and 12-month losses associated with the SVR value-growth model are substantially less than the respective losses incurred by the passive strategy. Summarizing, all of these findings can serve as an indication of robustness of the SVR strategy.

Figure 4 shows style signals associated with the SVR value-growth rotation strategy. The predominant style signal during this period is 'Growth', with some notable exceptions however. 'Value' signals have been produced mostly in 1993, in the beginning of 1994, and in the first half of 2001. Almost no 'Value' signals have been given during the periods stretching from June 1996 till August 1998, and from June 1999 till November 2000.

Figure 5 presents the realized excess returns forecasted by the basic SVR style timing strategy in the 25 bp. transaction-cost scenario. It can be seen from the figure that most of the accrued returns come out of the last four years of the sample period, which actually appears to be the most volatile.

A number of further conclusions can be drawn by examining Fig. 6. Next to the cumulative returns from the passive strategy and the SVR strategy that predicts the one-month-ahead return difference under zero transaction costs, the figure reveals the cumulative returns from two more strategies: the three- and six-month-horizon SVR strategies. The latter two strategies are constructed simply by taking the (unweighted) average of the signals produced by models constructed up to three and up to six months before any predicted month, and investing according to this combined signal. The procedure used is equal to that used in Jegadeesh and Titman (1993).

The first striking feature is that most of the cumulative returns are accrued in times of relatively higher volatility, and especially during 1993, in the beginning of 1994, and between 1999 and 2002. The magnitude of the volatility of returns can be observed by tracking the (monthly) changes in the cumulative returns of the passive strategy. Larger shocks in these series correspond to greater volatility of the value premium. A second interesting feature is that the basic one-month-horizon SVR strategy performs better than in the case of three- and six-month forecast horizons. A potential reason for this is that forecast signals produced by models built in the more distant past become less accurate than those provided by the more recent models, which are constructed using newer information.

Small-large rotation strategies

Detailed information on the small-large SVR strategy, the passive small-large strategy and the maximum attainable MAX_SL strategy can be found in the right part of Table 3.

In the out-of-sample period the passive small-large rotation strategy achieves an annual return of −1.26%. The optimal MAX_SL strategy provides an annual return of 27.04% in the 50 bp. transaction-cost scenario, which is 5.50% more than the corresponding result for the MAX_VG strategy. This reveals that the potential benefit from size rotation seems to be much greater than the one from the corresponding value-growth rotation. Table 3 shows that this extra potential can indeed be captured. Moreover, for the zero-transaction-cost

[12] Z(equality) measures the risk-adjusted performance difference between a switching Support Vector Regression strategy and the passive value-growth strategy. The Z(equality)-score is computed in a standard way (in line with, e.g. Glantz, 1992).

Table 3. Summary performance statistics for passive rotation strategies and various Support Vector Regression strategies using a one-month forecast horizon, for the period 1993:01–2002:12. 'VmG' and 'SmL' denote passive value-growth and small-large strategies, respectively. The explanatory variables are listed in Table 2. MAX_VG and MAX_SL denote the perfect foresight strategies. CV denotes a timing strategy based on Support Vector Regression cross-validation mean squared error. All numbers are annualized data unless stated otherwise. All strategies are long/short monthly positions on the style and size indices. The overall position for month $t+1$ is based on the signal produced by the optimal model based on 60 months of historical data on all explanatory factors. Transaction costs are assumed to be 0 bp., 25 bp., and 50 bp. single trip. The row '% months in Growth/Large' has to be interpreted as '% months in Growth' for the columns considering the value-growth rotation and '% months in Large' for the columns considering the small-large rotation. The interpretation of the next row is similar

	Value-growth rotation					Small-large roation				
	VmG	CV, 0 bp.	CV, 25 bp.	CV, 50 bp.	MAX_VG 50 bp.	SmL	CV, 0 bp.	CV, 25 bp.	CV, 50 bp.	MAX_SL, 50 bp.
Mean	0.24	10.30	8.30	6.30	21.54	−1.26	10.71	9.11	7.51	27.04
Standard deviation	10.95	9.95	9.90	9.90	7.83	13.00	10.92	10.92	10.96	8.87
Information ratio	0.02	1.04***	0.84***	0.64**	2.75***	−0.10	0.98***	0.83***	0.69**	3.05***
Z (equality)		2.15***	1.73*	1.30	5.00***		2.23**	1.93*	1.63	5.69***
Median	−0.11	0.33	0.31	0.30	0.50	0.05	0.45	0.44	0.43	0.79
Minimum (monthly)	−12.02	−5.51	−5.51	−5.51	−0.98	−15.71	−7.70	−7.70	−7.70	−0.96
Maximum (monthly)	9.74	12.02	11.77	11.52	11.02	16.78	16.78	16.53	16.28	16.78
Skewness (monthly)	0.01	1.21	1.18	1.13	1.61	0.20	0.77	0.76	0.74	2.63
Excess kurtosis (monthly)	2.40	2.66	2.52	2.34	3.41	4.49	4.74	4.55	4.30	11.74
% Negative months	45.83	32.50	49.17	50.00	19.17	51.67	33.33	44.07	45.83	10.00
Largest 3-month loss	−11.55	−5.90	−6.40	−6.90	−1.99	−21.63	−8.84	−8.84	−9.21	−1.18
Largest 12-month loss	−22.86	−8.07	−11.51	−15.26	2.21	−31.85	−3.21	−5.46	−7.71	8.03
% months in growth/large	0.00	53.33	53.33	53.33	45.83	0.00	45.00	45.00	45.00	51.67
% months in value/small	100.00	28.33	28.33	28.33	54.17	100.00	45.00	45.00	45.00	48.33
% months no position	0.00	18.33	18.33	18.33	0.00	0.00	10.00	10.00	10.00	0.00

Notes: *Indicates significance at the (2-tail) 10% level; **indicates significance at the (2-tail) 5% level; ***indicates significance at the (2-tail) 1% level.

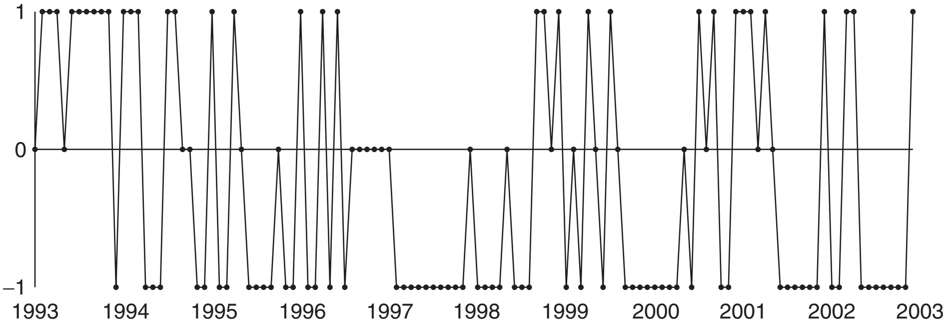

Fig. 4. Investment signals ('value' = 1, 'growth' = −1, 'no signal' = 0) produced by the SVR value-growth model investment strategy

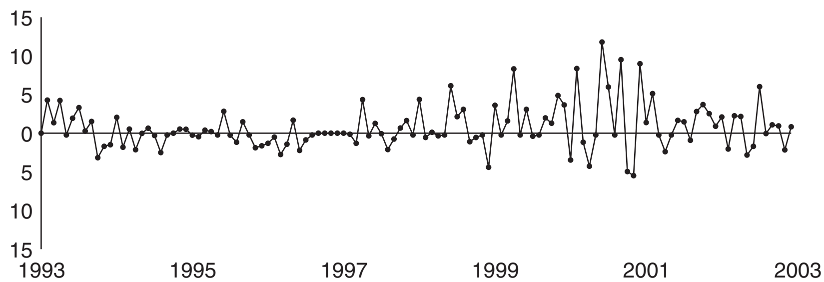

Fig. 5. Realized excess returns forecasted by the SVR value-growth investment strategy for the 25 bp. transaction costs scenario

regime, for example, the one-, three- and six-month forecast horizon small-large strategies produce 10.71%, 8.03% and 7.73% annual returns, while the respective results from the SVR value-growth strategies are 10.30%, 5.84% and 5.02% respectively.[13] As in the value-growth case, the SVR size model is able to capture roughly one-third of the maximum attainable cumulative returns under all considered transaction cost regimes.

As it turns out, the results from the robustness checks that were performed on the SVR value-growth model are also valid for SVR size rotation. Under the assumption of 50 bp. transaction costs, the realized information ratio of 0.69 from the SVR size model is significant at the (two-tail) 5% level. The SVR size strategy produces significantly different results from the passive size strategy. The largest three-month and, especially, 12-month losses from the SVR model are drastically more bearable than the ones from the passive size strategy: −8.84% versus −21.63% and −3.21% versus −31.85%, respectively, under zero transaction costs. Notice, additionally, that the one-month strategy again outperforms the longer horizon alternatives, consistent with the findings of the SVR value-growth strategy, see Fig. 7.

Simultaneous value-growth and size timing

In case investors have decided to follow both the value and size SVR strategies simultaneously at the beginning of the sample period, they would have witnessed even greater relative gains as compared to sticking only to a single type of timing (see Table 4 and Fig. 8 for details). Indicative of this are the realized information ratios of simultaneous style and size timing: 1.27, 1.06 and 0.84 under 0 bp., 25 bp. and 50 bp. single trip transaction cost regimes, all significant at the (two-tail) 1% level. Not surprisingly, these information ratios are higher than the ones associated with either style or size timing individually, as investors actually diversify the risk associated with each timing strategy. The information ratio of the passive simultaneous timing strategy is negative (−0.06). Interestingly, the largest three-month and 12-month losses associated with simultaneous investing turn out to be quite

[13] Not all these results are presented in Table 3. They are available upon request.

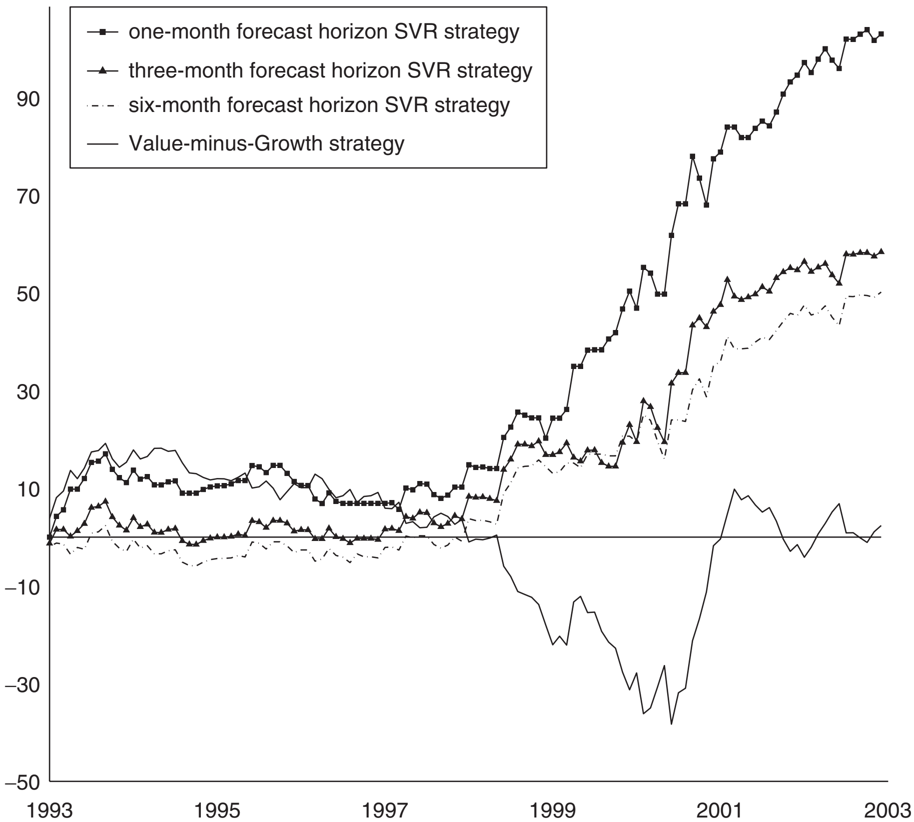

Fig. 6. Accrued cumulative returns from the passive value-growth strategy and the Support Vector Regression (SVR) one-, three-, and six-month horizon strategies for the period January 1993–December 2002, under no transaction costs. The one-month horizon strategy performs best, gaining most of its accumulated profits during turbulent times on the financial market. In such periods, the three- and six-month horizon models follow suit with a time lag, as logically expected. During relatively calmer periods, all strategies perform similarly: (-■-) one-month forecast horizon SVR strategy, (-▲-) three-month forecast horizon SVR strategy, (·–·–·) six-month forecast horizon SVR strategy, (–) Value-minus-Growth strategy

tolerable: −4.10% and −3.96%, assuming zero transaction costs. It appears that it pays to diversify the market timing strategies, at least as far as value and size timing are concerned.

Admittedly, it is expected that all of the findings are dependent on the historical model-building horizon and on the length of the trading period. Choosing to trade for a longer period could come at the expense of incurring formidable transaction costs, as noted in the Data section. Additionally, varying the length of the model-building horizon might yield a 'best' horizon that would be difficult to justify. Thus, future research could concentrate on both of these issues.

Discussion of results

Overall, the findings on the predictability of the size and the value premium corroborate the results of previous studies for the USA (Kao and Shumaker, 1999) and other mature markets, such as the UK (Levis and Liodakis, 1999) and Japan (Bauer *et al.*, 2004). This study shows that a style consistent strategy, i.e. consistently favouring value over growth and small over large, does not necessarily lead to positive returns in the long run. The proposed SVR style timing strategies, taking full advantage of information on the market and the economic cycle, are partially able to forecast the sign of both the value and the sign premium. This ability is particularly evident during volatile times, and especially during the TMT bubble and its aftermath (1998–2001). Furthermore, the transaction costs of the rotation strategies are expected to be small as futures on well-known and liquid indices are applied. This raises the possibility that institutional investors can exploit this strategy in real time.

Nonetheless, care is required in interpreting these results. Amihud and Mendelson (1986) for instance argue that the bid–ask bounce possibly creates an upward bias in reported profits from trading strategies. To remedy this effect, they suggest adjusting

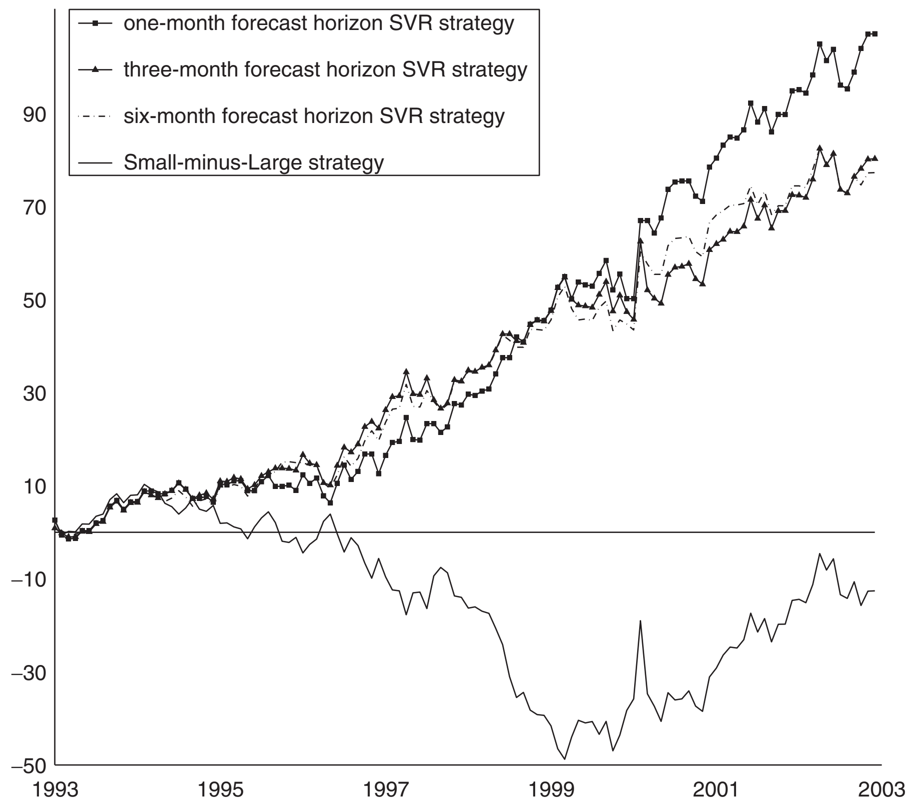

Fig. 7. Accrued cumulative returns from the passive small-large strategy and the Support Vector Regression (SVR) one-, three-, and six-month horizon strategies for the period January 1993–December 2002, under no transaction costs. The one-month horizon strategy performs best, gaining the predominant part of its accumulated profits during turbulent times on the financial market. The three- and six-month horizon models follow a very similar pattern, but perform slightly worse. During relatively calmer periods, all strategies perform similarly: (-■-) one-month forecast horizon SVR strategy, (-▲-) three-month forecast horizon SVR strategy, (·—·—·) six-month forecast horizon SVR strategy, (–) Small-minus-Large strategy

transaction prices by one-half of the bid–ask spread. This actually is equivalent to adding an extra fixed amount of transaction costs to each trade on top of what has already been assumed in the rotation strategies, provided that the bid–ask spread remains constant through time. As futures on well-known indices are used, it is likely that this effect is grossly captured by assuming relatively high transaction costs (50 bp.). Further, non-synchronous trading might induce spurious cross-autocorrelation between less frequently and more frequently traded stocks, see Campbell *et al*. (1997). This particular issue might be applicable to a greater extent to the size rotation rather than the value-growth rotation strategy, since small stocks are traded less frequently in general. The use of a forecasting horizon of three and six months however yields similar results. For such long horizons the market microstructure effects mentioned above are less likely to be relevant.[14]

V. Conclusion

This paper examines whether short-term directional variations in the size and value premium in the US stock market are sufficiently predictable to be exploited by means of a tactical timing strategy. As a forecasting tool, the so-called Support Vector Regressions (SVR) is employed. SVR have only recently been developed in the artificial intelligence field and have been rarely applied in a financial context. Using SVR, it is possible to circumvent the well-known problems of overfitting, especially in multivariate settings, in an elegant way.

[14] Additionally, strategies were looked at where higher transaction costs are assumed for small cap trades (both long and short) and short large cap trades. Results, which are available upon request, show that it is difficult to exploit the rotation strategy when using a basket of individual stocks instead of futures.

Table 4. Simultaneous passive and Support Vector Regression value-growth and small-large rotation strategies. Summary statistics for following simultaneously passive value-growth (VmG) and passive small-large (SmL) strategies on the one hand, and Support Vector Regression value-growth and small-large rotation strategies using a one-month forecast horizon, on the other, for the period 1993:01–2002:12. The explanatory variables are listed in Table 2. MAX denotes the perfect foresight value-growth and small-large rotation combined strategy. CV denotes the timing strategy based on Support Vector Regression cross validation mean squared error. All numbers are annualized data unless stated otherwise. All strategies are long/short monthly positions on the style and size indices. The overall position for month $t+1$ is based on the signal produced by the optimal model based on 60 months of historical data of all explanatory factors. Transaction costs are assumed to be 0 bp., 25 bp., and 50 bp. single trip

	Simultaneous rotation VmG plus SmL	CV, 0 bp.	CV, 25 bp.	CV, 50 bp.	MAX, 50 bp.
Mean	−0.51	10.51	8.71	6.91	24.29
Standard deviation	8.71	8.25	8.23	8.23	6.76
Information ratio	−0.06	1.27***	1.06***	0.84***	3.59***
Z(equality)		2.90***	2.43**	1.96*	7.11***
Median	0.06	0.44	0.41	0.39	0.56
Minimum (monthly)	−7.26	−3.86	−4.24	−4.61	−0.94
Maximum (monthly)	8.43	12.56	12.44	12.31	12.06
Skewness (monthly)	0.13	1.51	1.49	1.45	1.87
Excess kurtosis (monthly)	1.12	5.29	5.30	5.23	6.05
% negative months	49.17	33.33	37.50	40.83	7.50
Largest 3-month loss	−13.20	−4.10	−5.10	−6.10	−1.17
Largest 12-month loss	−26.71	−3.96	−6.34	−8.82	6.77

Notes: *Indicates significance at the (2-tail) 10% level; **indicates significance at the (2-tail) 5% level; *** indicates significance at the (2-tail) 1% level.

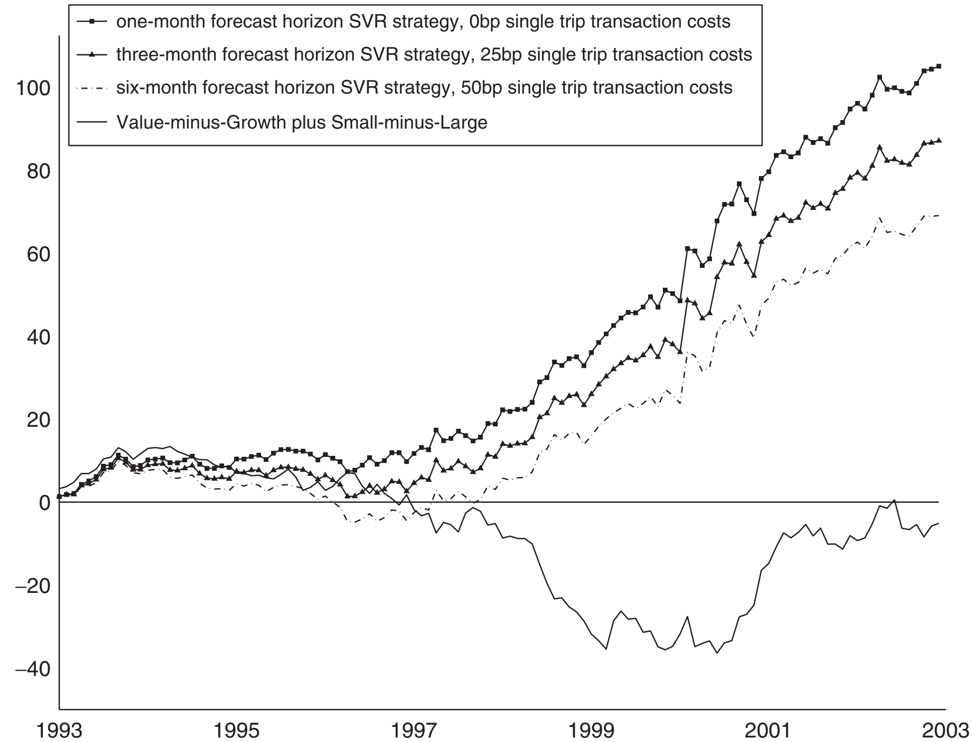

Fig. 8. Accrued cumulative returns from investing simultaneously according to the Support Vector Regression (SVR) one-month horizon value-growth and small-large strategies on the one hand, and from investing simultaneously according to the passive value-growth and small-large strategies during the period January 1993–December 2002, under different transaction costs: (-■-) one-month forecast horizon SVR strategy, 0 bp single trip transaction costs, (-▲-) three-month forecast horizon SVR strategy, 25 bp single trip transaction costs (·–·–·) six-month forecast horizon SVR strategy, 50 bp single trip transaction costs, (–) Value-minus-Growth plus Small-minus-Large

The empirical findings clearly show that both premiums are highly predictable during the trading period. This comes at odds with the mainstream literature that provides evidence for the long-term superiority of returns to value *vis-a-vis* growth and small *vis-a-vis* large stocks. After adjustment for fair levels of transaction costs this result still holds. Under high transaction cost levels, expected to be relevant in a dynamic economic environment, it is difficult in practice to obtain incremental benefits over style consistent strategies. That is why it is critical to develop timing strategies that can be implemented using index futures or low-cost trading baskets like exchange traded funds. In terms of realized information ratios, a combination of both value-growth and small-large timing produces the most interesting results.

Acknowledgements

We would like to thank ABP investments for providing the data, Evgueni Smirnov (IKAT), Roderick Molenaar (ABP Investments) and Borislav Pavlov (LIFE) for the many fruitful discussions, and Evgueni Smirnov also for many hours of using his computer.

References

Ahmed, P., Lockwood, L. and Nanda, S. (2002) Multistyle rotation strategies, *Journal of Portfolio Management*, **28**, 17–29.

Amihud, Y. and Mendelson, H. (1986) Asset pricing and the bid-ask spread, *Journal of Financial Economics*, **17**, 223–249.

Arnott, R., Dorian, J. and Macedo, R. (1992) Style management: the missing element in equity portfolios, *Journal of Investing*, **1**, 13–21.

Arnott, R., Rice, D., Kelso, C., Kiscadden, S. and Macedo, R. (1989) Forecasting factor returns: an intriguing possibility, *Journal of Portfolio Management*, **16**, 28–35.

Asness, C., Friedman, J., Krail, R. and Liew, J. (2000) Style timing: value versus growth, *Journal of Portfolio Management*, **26**, 51–60.

Banz, R. (1981) The relationship between return and market value of common stocks, *Journal of Financial Economics*, **9**, 3–18.

Bauer, R. and Molenaar, R. (2002) Is the value premium predictable in real time? LIFE Working Paper 02-003.

Bauer, R., Derwall, J. and Molenaar, R. (2004) The real-time predictability of the size and value premium in Japan, *Pacific-Basin Finance Journal*, **12**, 503–23.

Bernstein, R. (2001) *Navigate the Noise: Investing in The New Age of Media And Hype*, John Wiley and Sons, New York.

Bossaerts, P. and Hillion, P. (1999) Implementing statistical criteria to select return forecasting models: what do we learn?, *Review of Financial Studies*, **12**, 405–28.

Burbidge, R. and Buxton, B. (2001) An introduction to support vector machines for data mining, in *Keynote Papers, Young OR12* (Ed.) M. Sheppee, Operational Research Society, University of Nottingham, Nottingham, pp. 3–15.

Burges, C. (1998) A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery*, **2**, 121–67.

Campbell, J., Lo, A. and MacKinlay, A. (1997) *The Econometrics of Financial Markets*, Princeton University Press, Princetown.

Chan, K. and Chen, N.-F. (1991) Structural and return characteristics of small and large firms, *Journal of Finance*, **46**, 1467–84.

Chan, L. and Lakonishok, J. (2004) Value and growth investing: review and update, *Financial Analysts Journal*, **60**, 71–86.

Chang, M.-W., Chen, B.-J. and Lin, C.-J. (2001) EUNITE network competition: Electricity load forecasting, Winner of EUNITE world wide competition on electricity load prediction, URL http://www.csie.ntu.edu.tw/~cjlin/papers.html

Chang, C.-C. and Lin, C.-J. (2002) LIBSVM: a library for support vector machines, URL http://www.csie.ntu.edu.tw/~cjlin/libsvm

Cooper, M., Gulen, H. and Vassalou, M. (2001) Investing in size and book-to-market portfolios using information about the macroeconomy: some new trading rules, *Mimeo*, Columbia University, New York.

Copeland, M. and Copeland, T. (1999) Market timing: style and size rotation using the VIX, *Financial Analyst Journal*, **55**, 73–81.

Dimson, E., Nagel, S. and Quigley, G. (2003) Capturing the value premium in the United Kingdom, *Financial Analysts Journal*, **59**, 35–45.

Elfakhani, S. (2000) Short positions, size effect, and the liquidity hypothesis: implications for stock performance, *Applied Financial Economics*, **10**, 105–16.

Fama, E. and French, K. (1992) The cross-section of expected stock returns, *Journal of Finance*, **47**, 427–65.

Fama, E. and French, K. (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, **33**, 3–53.

Fama, E. and French, K. (1998) Value versus growth: the international evidence, *Journal of Finance*, **53**, 1975–99.

Glantz, S. (1992) *Primer of Biostatistics*, 3rd edn, McGraw-Hill, New York.

Haugen, R. and Baker, N. (1996) Commonality in the determinants of expected stock returns, *Journal of Financial Economics*, **41**, 401–39.

Jacobs, B. and Levy, K. (1996) High definition style rotation, *Journal of Investing*, **5**, 14–23.

Jegadeesh, N. and Titman, S. (1993) Returns to buying winners and selling losers: implications for market efficiency, *Journal of Finance*, **48**, 65–91.

Kao, D.-L. and Shumaker, R. (1999) Equity style timing, *Financial Analysts Journal*, **55**, 37–48.

Kwon, K.-Y. and Kish, R. J. (2002) Technical trading strategies and return predictability: NYSE, *Applied Financial Economics*, **12**, 639–53.

Lakonishok, J., Schleifer, A. and Vishny, R. (1994) Contrarian investment, extrapolation and risk, *Journal of Finance*, **49**, 1541–78.

Levis, M. and Liodakis, M. (1999) The profitability of style rotation strategies in the United Kingdom, *Journal of Portfolio Management*, **25**, 73–86.

Liew, J. and Vassalou, M. (2000) Can book-to-market, size and momentum be risk factors that predict economic growth?, *Journal of Financial Economics*, **57**, 221–45.

Lo, A. and MacKinlay, A. (1990) Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies*, **3**, 431–68.

Lucas, A., Van Dijk, R. and Kloek, T. (2002) Stock selection, style rotation, and risk, *Journal of Empirical Finance*, **9**, 1–34.

Maragoudakis, M., Kermanidis, K., Fakotakis, N. and Kokkinakis, G. (2002) Combining bayesian and support vector machines learning to automatically complete syntactical information for HPSG-like formalisms, in LREC 2002, 3rd International Conference on Language Resources and Evaluation, Vol. 1, Las Palmas, Spain, pp. 93–100. URL http://s1t.wc1.ee.upatras.gr/papers/maragoudakis10.pdf

Miller, K., Li, H. and Cox, D. (2001) US style rotation model, Industry Note, Salomon, Smith Barney.

Mills, T. C. and Jordanov, J. V. (2003) The size effect and the random walk hypothesis: evidence from the London Stock Exchange using Markov Chains, *Applied Financial Economics*, **13**, 807–15.

Monteiro, A. (2001) Interest rate curve estimation: a support vector regression application to finance, working paper, Princetown University, URL http://www.princeton.edu/~monteiro/SVM%20swaps.pdf

Müller, K.-R., Smola, A., Rätsch, G., Schölkopf, B., Kohlmorgen, J. and Vapnik, V. (1997) Predicting time series with support vector machines, in *Proceedings of the International Conference on Artificial Neural Networks* (Eds) W. Gerstner, A. Germond, M. Hasler, J.-D. Nicoud, Vol. 1327 of Springer Lecture Notes in Computer Science, Springer, Berlin, pp. 999–1004.

Müller, K.-R., Mika, S., Rätsch, G., Tsuda, K. and Schölkopf, B. (2001) An introduction to kernel-based learning algorithms, *IEEE Transactions on Neural Networks*, **12**, 181–201.

Mun, J. C., Kish, R. J. and Vasconcello, G. M. (2001) The contrarian investment strategy: additional evidence, *Applied Financial Economics*, **11**, 619–40.

Pesaran, M. and Timmermann, A. (1995) Predictability of stock returns: robustness and economic significance, *Journal of Finance*, **50**, 1201–28.

Rocco, S. C. and Moreno, J. (2003) A support vector machine model for currency crises discrimination, in *Computational Intelligence in Economics and Finance* (Eds) S. Chen and P. Wang, *Advanced Information Processing*, Springer-Verlag, pp. 171–81.

Schwob, R. (2000) Style and style analysis from a practitioners perspective: what is it and what does it mean for european equity investors, *Journal of Asset Management*, **1**, 39–59.

Smola, A. (1996) Regression estimation with support vector learning machines, Master's thesis, Technische Universität München.

Smola, A. and Schölkopf, B. (1998) A tutorial on support vector regression, NeuroCOLT2 Technical Report NC-TR-98-030, University of London, UK.

Stone, M. (1977) Asymptotics for and against cross-validation, *Biometrika*, **64**, 29–35.

Van Gestel, T., Baesens, B., Garcia, J. and Van Dijcke, P. (2003) A support vector machine approach to credit scoring, *Bank en Financiewezen*, **2**, 73–82. URL http://www.geocities.com/joaogarcia18/BANKFINVer4.pdf

Vapnik, V. N. (1995/2000) *The Nature of Statistical Learning Theory*, 2nd edn, Springer-Verlag, New York.

Weiss, S. and Kulikowski, C. (1991) *Computer Systems that Learn*, Morgan Kaufman, San Mateo.

Woodford, B. J. (2001) Comparative analysis of the EFuNN and the support vector machine models for the classification of horticulture data, in *Proceedings of the Fifth Biannual Conference on Artificial Neural Networks and Expert Systems* (Eds) N. Kasabov and B. Woodford, University of Otago Press, Dunedin, pp. 70–75.