The real-time predictability of the size and value premium in Japan

Rob Bauer[a,b,*], Jeroen Derwall[c], Roderick Molenaar[b]

[a] Maastricht University, Maastricht, The Netherlands
[b] ABP Investments, The Netherlands
[c] Erasmus University, Rotterdam, The Netherlands

Abstract

This study examines whether the short-term variation in the Japanese size and value premium is sufficiently predictable to be exploited by a timing strategy. In the spirit of Pesaran and Timmermann [J. Finance 50 (1995) 1201], we employ a dynamic modeling approach in which we explicitly allow for permutations among the determinants in order to mitigate typical data-snooping biases. Using a base set of candidate regressors, we perform an in-sample estimation of all economically sensible models. Subsequently, a ''best'' model is determined according to a selection criterion. However, whereas most studies use in-sample model selection criteria, we introduce an out-of-sample *training* period to select our models. We then implement our strategy in a second-stage out-of-sample period: the *trading* period. All stages re-occur on a monthly basis via a rolling window framework. The results confirm sufficient predictability under lower transaction cost levels. Under high transaction costs scenarios it is more difficult to obtain incremental benefits.
© 2004 Elsevier B.V. All rights reserved.

1. Introduction

Over the past decades, a wide range of influential academic studies has elaborated on the potential benefits of investing in stocks with fundamental commonalities as it would

---

* Corresponding author. Limburg Institute of Financial Economics (LIFE), Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Tel.: +31-43-3883556; fax: +31-43-3884875.
E-mail address: r.bauer@berfin.unimaas.nl (R. Bauer).

provide a premium in the long run. The perception of a ''size effect'' primarily emerged from the empirical findings of Banz (1981) who reported a negative relation between a firm's market capitalization and its stock performance in the United States. Fama and French (1992, 1998) and Lakonishok et al. (1994) showed that firms with typical ''value'' characteristics 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 ''growth'' stocks, with high M/B, P/E and P/C ratios.

Consequently, a controversy emerged on the source of the size and value premium. Among others, Chan and Chen (1991; size) and Fama and French (1993; value) argue that this premium is a reward for holding stocks of firms under relative distress. Cochrane (1999) notes that only the non-idiosyncratic component of distress is relevant since it cannot be diversified away. This interpretation implicitly relates the size and value premium to macroeconomic risk. Others attribute the observed effects to the stock market's lack of efficient pricing ability; see, for example, Lakonishok et al. (1994) or Haugen and Baker (1996). A third explanation, as initially suggested by Lo and MacKinlay (1990), is that the obtained results are due to data-snooping biases.

However, the rather disappointing performance of ''pure'' small firm and value strategies during the 1990s has pointed out that style consistency may not provide the long-term benefits initially assumed. Furthermore, there is evidence indicating that the size and value premium exhibit significant short-term (directional) variation. Because of this lack of consistency, there now is a consensus that style diversification is the optimal solution to avoid the risk associated with pure style investing, and to capture the benefits each separate investment style has to offer. At the other end of this spectrum, however, a small body of literature has explicitly addressed the potential benefits of style timing strategies over a style consistent approach. Although these papers may differ in methodology, they all rely on the opinion that the cyclical behavior of investment styles is correlated with systematic forces, which could make it partially predictable.

Most empirical work on this topic is concentrated on the well-documented markets in the United States and the United Kingdom. Levis and Liodakis (1999) find moderate evidence in favor of small/large rotation strategies, but less evidence for value/growth rotation in the United Kingdom. Consistent with their findings, Cooper et al. (2001) find sufficient predictability for size-sorted strategies to succeed, but weaker results for value-sorted strategies in the United States.[1] The Japanese stock market, however, has received little or no attention in this field. There certainly is evidence suggesting that small firm and value effects also exist in the Japanese equity market; see, e.g. Chan et al. (1993), Cai (1997), Bae and Kim (1998), Chen and Zhang (1998) and Daniel et al. (2001). The objective of this paper is to examine whether the short-term directional variation in the size and value premium in Japan is predictable and exploitable by means of a style rotation strategy. Japan's economic worries in the previous decade provide an interesting opportunity to assess the robustness and economic relevance of style timing versus style consistent strategies during periods of high economic uncertainty. Moreover, the extension

---

[1] Other related work includes Arnott et al. (1989, 1992), Jacobs and Levy (1996), Copeland and Copeland (1999), Kao and Shumaker (1999), Asness et al. (2000), Lucas et al. (2002) and Ahmed et al. (2002).

towards an unexplored sample is likely to provide substantial contributing evidence concerning the true practical significance of style rotation strategies.

The ability to predict stock returns ahead of time is the backbone of many supposedly feasible timing strategies. However, despite overwhelming evidence from academic literature to the contrary, the benefits of predictability are hardly observed in practice. As also pointed out by Cooper and Gulen (2002), the apparent predictability gap might be due to substantial biases in many reported findings that have been obtained from a setting that benefits too much from ex post knowledge. A classic example is the estimation of a single predictive model based on the entire sample period, which is not obtainable by investors in "real time". Although many papers partially tackle this critique by validating the predictive ability in an out-of-sample context, many other parameters, including the choice of predictive model, are usually determined with the benefit of hindsight. In order to obtain truly practical results with such a procedure, the assumption of a time-invariant joint significance of the determinants needs to hold which is very doubtful. Provided the empirical results in the back-testing process rely substantially on these parameters, the economic significance will be exaggerated. To mitigate the impact of "hindsight" bias, we simulate our trading strategies by means of a dynamic modeling approach in which we explicitly account for the continuous uncertainty that real-time investors face concerning the choice of optimal set of predictive variables.

Our procedure is largely an extension of the work of Pesaran and Timmermann (1995), who introduced the approach at the market return predictability level for the United States. Using a base set of forecasting variables collected from previous work or because of "a priori beliefs", we first define a universe of parsimonious models based on in-sample estimation. Following this, we allow for the selection of a "best" model according to a predefined selection criterion. However, whereas most studies use in-sample model selection criteria, we increase the likelihood of a successful forecast by introducing a 24-month *out-of-sample training* period to test and select our forecasting models. We start the implementation of the rotation strategy in a second-stage out-of-sample period, from hereon referred to as the *trading* period. All events re-occur on a monthly basis via a rolling window framework. The choice of a selection period that postdates the model estimation sample relates to the evidence of Bossaerts and Hillion (1999) who failed to find sufficient out-of-sample predictability when using conventional in-sample selection criteria.[2] We simultaneously examine two independent style rotation strategies. The first strategy attempts to anticipate directional changes in the performance of small stock portfolios relative to their large stock counterparts. The second strategy entails the timing of short-term swings in the performance difference between value stocks and growth stocks. By employing a market-neutral rotation strategy, we aim to outperform the style consistent buy-and-hold benchmark portfolio, taking into account different forecast horizons and different levels of transaction costs.

The remainder of this paper is organized as follows: in the next section, we outline the data we use to proxy for the size and value premium and to construct a base set of

---

[2] Conventional statistical criteria include the adjusted R^2, the Aikake information criterion and the Schwarz criterion.

forecasting variables. In Section 3, we elaborate on the methodology we apply to obtain real-time results. In Section 4, we discuss the empirical results. Section 5 summarizes our findings and concludes.

2. Construction of premiums and predictive variables

2.1. Choice of style indexes

To examine whether the size and value premium are predictable in real time we use the BARRA/Nikko Style Indexes to develop proxies for these premiums. We obtained monthly return data from the website of Nikko Research Institute.[3] The BARRA/Nikko Style Indexes cover all Japanese stocks listed on all the Stock Exchanges and OTC markets in Japan. The BARRA/Nikko Large Cap Index covers the upper 85% of total market capitalization. The Small Cap Index covers the remaining 15% of the total market. Within the Large Cap Index, stocks are assigned to Value and Growth Indexes. The classification of stocks into the Value Index and into the Growth Index is based on rankings according to an average rank of (a) book-to-market ratio, and (b) the average growth rate of sales for 1-year realization and prediction. One third of stocks with upper ranks are assigned to Value and stocks in the lower third range are classified into Growth. The remaining third of stocks are divided into both Value and Growth. Their weights in each index correspond to the average rank across the B/M and sales growth dimension.[4] The BARRA/Nikko indexes are rebalanced twice a year, at the end of June and December. Furthermore, all indexes are adjusted for dividends and it is assumed that dividends are reinvested.

Fig. 1 shows the cumulative return distributions of the spreads and the corresponding summary statistics for the period 1980–2001. At first glance, the small firm effect clearly lacks robustness. This is confirmed by the *t*-statistic in the table in Fig. 1: we fail to reject the null hypothesis of a zero mean for the monthly size premium. Value stocks, on the other hand, did particularly well relative to growth stocks for most of the sample period. The information ratio of the buy-and-hold portfolio, defined as the ratio of the mean return to the standard deviation, is 0.43. The *t*-statistic of 2.03 indicates the value premium is significantly positive at a 5% level. Still, improvements over a buy-and-hold strategy could have been achieved during the strong decline of the value premium at the end of the previous decade. The lack of robustness of the small firm effect clearly emphasizes the possible benefits of a style timing routine.

It is also important to mention that the premiums are not highly correlated with their counterparts in the United States, which strengthens our assumption that obtaining evidence from the unexplored Japanese market is likely to provide considerable complementary evidence on the benefits of style rotation strategies in general. Foster et al. (1997),

---

[3] http://www.nikko.co.jp/NRC/Index/style/hist.html.
[4] The BARRA/Nikko indexes should therefore not be confused with the S&P/BARRA indexes in the United States. The S&P/BARRA Value and Growth Indexes are mutually exclusive and constructed using B/M only.

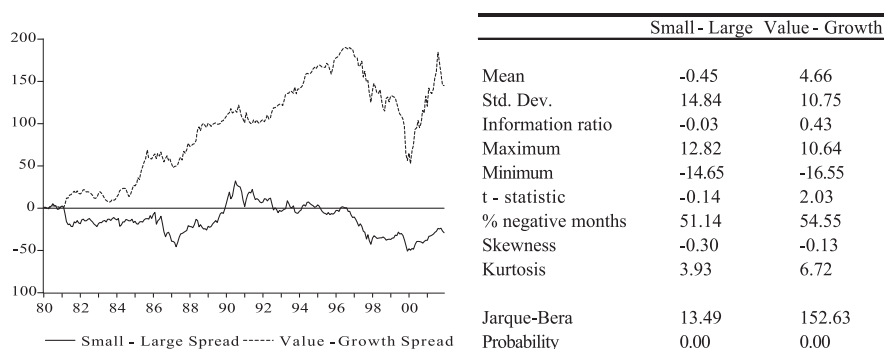|	Small - Large	Value - Growth
Mean	-0.45	4.66
Std. Dev.	14.84	10.75
Information ratio	-0.03	0.43
Maximum	12.82	10.64
Minimum	-14.65	-16.55
t - statistic	-0.14	2.03
% negative months	51.14	54.55
Skewness	-0.30	-0.13
Kurtosis	3.93	6.72
Jarque-Bera	13.49	152.63
Probability	0.00	0.00

Fig. 1. Cumulative month-to-month size and value premium. Fig. 1 shows cumulative month-to-month small/large (value/growth) return spread series during the period 1980:01–2001:12. The corresponding table presents summary statistics for each of the spreads. Mean values and standard deviations are presented on an annualized basis. The information ratio is the mean divided by the standard deviation. The *t*-statistic indicates the significance of the mean. The Jarque–Bera probability indicates the probability that the null hypothesis of a normally distributed series holds.

amongst others, note that using different data sets to examine predictability may be an ineffective approach to compile evidence if the data are highly correlated. Fortunately, the correlations between Japanese size and value premiums in this study and the S&P/BARRA spreads are 0.12 and 0.27, respectively.

2.2. Choice of base set of forecasting variables

There is a strongly developing belief that the time variation in the size premium as well as in the value premium is related to changes in the economic climate. Among others, Fama and French (1993) contended that value firms are relatively distressed companies that are more sensitive to macroeconomic factors. Cochrane (1999) furthermore commented that relative distress should not be viewed as firm-specific distress, which can be diversified away. Small firms and value firms are also said to have a higher degree of external financing risk, as these companies tend to be more leveraged, which makes them more sensitive to deteriorating economic conditions; see, e.g. Chan and Chen (1991) and Chen and Zhang (1998).[5] Perez-Quiros and Timmermann (2000) argue that small firms have a higher degree of asymmetry in their risk than large firms across periods of recession and periods of expansion and that small firms are more likely to be negatively affected by a recession and worsening credit market conditions. Consequently, the consensus is that value stocks and small caps are more positively affected by improving economic conditions, but more vulnerable than growth stocks and large caps during economic downturns. In support of this are Liew and Vassalou (2000). Using data from 10 countries, including Japan, they provided evidence of a generally positive lead-lag relation between the relative returns of small cap and/or value strategies and economic development.

---

[5] Cai (1997) found that value firms listed on the TSE between 1971 and 1992 actually had lower leverage ratios than growth companies, which gives an interesting twist to this story.

Finally, Jensen et al. (1998) showed that size and value effects in the United States are related to the monetary environment.

Unsurprisingly, previous empirical papers that attempted to predict the size premium and/or the value premium relied heavily on well-established economic state indicators, such as the popular APT factors documented in Chen et al. (1986). Levis and Liodakis (1999) use variables such as inflation, the term spread and the equity premium to forecast directional changes in the size and value premium in the United Kingdom. Kao and Shumaker (1999) use GDP forecasts, inflation, and several interest rate variables to explore the benefits of value/growth timing in the United States. Cooper et al. (2001) also use common business cycle-related variables to explore the benefits of size- and value-sorted market-neutral investment strategies in the United States. The majority of related work thus provides considerable evidence to support the notion that size and value effects are linked to the macro-economy.

Since we focus on a market that has remained unexplored in this context, we build on the knowledge outlined above and rely mainly on conventional economic variables to forecast the time variation in the size and value premium. Japanese economic state variables have been found to be essentially similar to those in other countries. For example, Hamao (1988) extended the work of Chen et al. (1986) to Japan and identified an APT-type model that includes variables such as interest rates, industrial production, inflation, the oil price and exchange rate factors. Campbell and Hamao (1989) found some predictive power for the market dividend yield and the yield spread between long- and short-term government bonds but less predictive ability for the short rate.

Specifically, our base set consists of the following 14 variables with potential predictive ability:

- DY: the dividend yield of TOPIX complemented by the dividend yield of Nikko 500. (The dividend yield data of TOPIX was only available after 1989.) Source: DataStream.
- TIBOR3m: Tokyo Interbank Offered Rate (3-month). Source: Economic and Social Research Institute (ESRI).
- DLCPI: 12-month inflation rate measured by the 12-month logarithmic change in the Consumer Price Index. Source: DataStream.
- INTSPREAD: the interest rate spread defined as the difference between the annual yield of newly issued long-term government bonds and TIBOR3m. Source: ESRI.
- GEYR: ratio of the newly issued long-term government bond yield and the market dividend yield.
- DLEAD: annual percentage change in the leading indicator. Source: ESRI.
- DIND: annual percentage change in the Index of Industrial Production (Mining and Manufacturing).
- DM1: 12-month logarithmic change in M1 money supply. Source: Bank of Japan.
- DFX: monthly logarithmic change in the Yen/Dollar exchange rate. Source: Bank of Japan.
- DOIL: 12-month logarithmic change in the Petroleum and Coal Producer Price Index. Source: DataStream.
- MKTPREMIUM: monthly return of TOPIX minus a short rate proxy. Source: DataStream.

– TOPIXVAR: return variance of TOPIX (previous 24 months). Source: DataStream.
– DAILYVAR: return variance of TOPIX (previous 30 days). Source: DataStream.
– REALYIELD: long government bond yield minus the inflation rate.

The predictive power of interest rate related variables in relation to the size premium and the value premium is derivable from various typical lines of reasoning. For instance, worsening economic conditions such as credit constraints will lead to rising interest costs that may ultimately affect the balance sheet of small companies more compared to those of large firms (Gertler and Gilchrist, 1994). Value and growth stocks usually have a different ''duration''; see, for example, Bernstein (2001). Specifically, typical growth firms retain a high degree of earnings in the short-run in order to ensure longer-term growth. Given that stocks are priced as the present value of future cash flows, value stocks and growth stocks are thus differentially sensitive to changes in the discount factor, which in turn relates their relative performance to interest rates. The choice of market related variables is mainly due to the stock market's predictive nature. The market dividend yield (DY) is known to vary in response to changes in business conditions and is able to forecast stock returns; see Fama and French (1988). Furthermore, market variability measures, such as TOPIXVAR and DAILYVAR, may signal a flight to quality stocks (i.e. large caps or growth stocks) during periods of high uncertainty in the market.[6] DIND and DLEAD explicitly track changes in the economy. Whereas the index of industrial production usually coincides with the state of the economy, the composite index of leading indicators is forward-looking. We furthermore include a measure of money growth (DM1) in the base set of variables. Changes in money supply, due to monetary policy actions for example, may affect market liquidity or economic activity and therefore future cash-flow expectations.[7] We also follow changes in the Yen/Dollar exchange rate (DFX) because exchange rate risk may be higher for large Japanese firms with more international exposure than for smaller companies that mainly operate domestically (Doukas et al., 1999).

Although these interpretations are very plausible, it should be noted that the exact nature of the information contained in these variables spans different areas and is difficult to unravel. Therefore, the interaction with style portfolio returns can be direct, indirect or non-causal. Nevertheless, as these variables are all highly accepted in a return predictability context, mostly due to their association with economic conditions, our models will be economically sensible.

A simple 72-month rolling correlation analysis between the variables and the 3-month size (value) premium, not reported graphically due to space limitations, revealed that none of the relations with the predictive variables has been stable through time. Each correlation coefficient changed sign at least once.[8] In other words, it is likely that relations reverse, disappear or restore. Reasons for this could be one or more structural breaks in the data generating processes, perhaps due to stock market bubbles or structural changes in the

---

[6] Nakagawa and Osawa (2000) also found evidence of a relation between Japanese stock market volatility and macroeconomic (industrial production) volatility.

[7] Using growth in M2 + CD, Hirata and Ueda (1998) reported that Japanese monetary aggregates are also useful in predicting recessions at shorter horizons.

[8] The graphs are available upon request.

market mechanism, or simply a highly unstable economy. This observation frequently causes real-time investors to review the optimal combination of predictive instruments with more recent data. In order to obtain practical back-testing results, any research paper on simulated investment strategies is therefore required to account for the dynamic predictive relationships observed in reality. Yet, the majority of research papers on style rotation strategies—or return predictability in general—still (implicitly) assume a time-invariant joint significance of predictive relations, making the empirical results subject to data-snooping criticism. This motivates our choice for the recursive modeling approach suggested by Pesaran and Timmermann (1995), which will be outlined in the following section.

3. Methodology

In the spirit of Pesaran and Timmermann (1995), we employ a dynamic modeling approach that is more suitable to predict the directional time variation in the size and value premium in the Japanese stock market. In econometric terms, the approach involves a recursive estimation procedure that allows for continuous permutations among the determinants in accordance with a predefined model selection criterion. Clearly, the advantage of this method is that all possible models are constantly re-estimated and re-evaluated in order to reflect an investor's continuous search for the best approximation of the "true" model based on historical information only. The purpose of this computationally intensive procedure is to minimize possible ex post data-snooping biases, which are likely to exaggerate the practical significance of our empirical results.

Given the nature of our proposed timing strategies, it suffices to forecast the *sign* of the size or value premium rather than the magnitude. This provides the opportunity to deviate from the standard Ordinary Least Squares regression procedure, which is particularly appealing considering the observed non-normality in the return series in the previous section. We therefore follow Levis and Liodakis (1999) and apply a logistic modeling approach to generate directional forecasts only. Predictions generated by such binary choice models indicate an estimated *probability* that one asset class will outperform the other. In detail, the procedure is as follows:

(1) We define a set of 14 predictive variables (x_t), described in the previous section. In order to ensure a more robust and parsimonious model specification, we constrain the number of forecasting variables in the model to a minimum of 1, and a maximum of 5. In doing so, we reduce the model universe to 3472 (out of 2^{14}) and still allow for permutations among the determinants.

(2) Using a standard Logit modeling approach of the form

$$p_{t+1} = P(y_{t+1} = 1) = \frac{\exp(\alpha + \beta x_t)}{1 + \exp(\alpha + \beta x_t)} \tag{1}$$

with $y_{t+1} = 1$ if $\text{Small}_{t+1} \geq \text{Large}_{t+1}$ (or $\text{Value}_{t+1} \geq \text{Growth}_{t+1}$), 0 otherwise, we estimate all possible variable combinations using a 72-month rolling model estimation window. To ensure a forward-looking nature of the variables and to account for delayed data
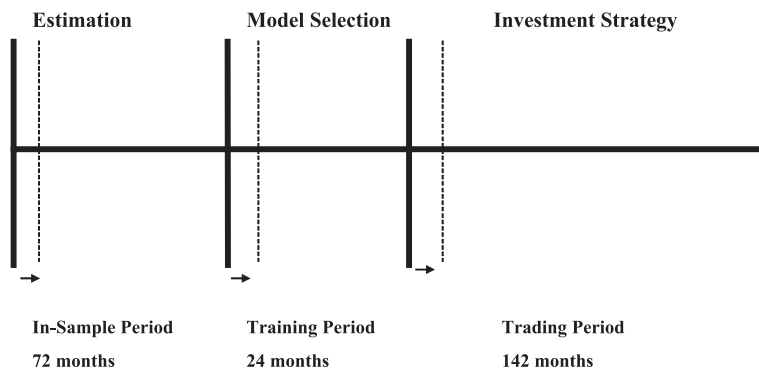
Fig. 2. Graphical representation of recursive modeling approach.

availability, all financial variables are lagged 1 month and macro-economic variables are lagged 2 months.

(3) Using a predefined selection criterion, we then select the "best" model based on the performance of the strategy in a rolling 24-month out-of-sample *training* period subsequent to the in-sample estimation period. The use of a selection period that postdates the model estimation sample relates to the evidence of Bossaerts and Hillion (1999) who failed to find sufficient out-of-sample predictability using in-sample selection criteria. The selection criterion chosen is consistent with the main purpose of our strategies, which is maximizing a switching portfolio's information ratio (IR). The IR-criterion selects models for their performance measured by the information ratio during the out-of-sample *training* period.[9] We estimate and select models for 1- and 3-month forecast horizons, and various levels of transaction costs. We explicitly incorporate transaction costs into the selection procedure, whereas most studies merely correct for transaction costs by imposing a fixed penalty following the allocation process. Having selected the best model via training, we implement our strategy in a second-stage out-of-sample period, the *trading* period.

(4) The signal generated by a model indicates the following:

– If $\hat{p}_{t+1} \geq 0.55$, we assign a "Small Cap" ("Value") signal and we go long on the Small Cap (Value) Index and short on the Large Cap (Growth) Index.
– If $\hat{p}_{t+1} \leq 0.45$, we assign a "Large Cap" ("Growth") signal and we go short on the Small Cap (Value) Index and long on the Large Cap (Growth) Index.
– If $0.45 < \hat{p}_{t+1} < 0.55$, we will assume there is no signal. We have no exposure to the indexes. This rule is applied in order to reduce the sensitivity of the model outcome.

(5) When using a 3-month forecast horizon, the portfolio construction process is more sophisticated because a forecast is then generated for a 3-month holding period. We combine monthly forecasts as follows: each month $t+2$, we build an equally weighted

---

[9] Allen and Karjalainen (1999) outline some technical trading rules based on similar *training* principles, but without using a recursive framework.

portfolio consisting of three portfolios: one created per ultimo t, one per ultimo $t+1$ and one per ultimo month $t+2$, respectively.[10]

Notice that all stages re-occur on a monthly basis. See Fig. 2 for a graphical representation of the recursive modeling approach. The out-of-sample *trading* period in which we take on active positions ranges from 1990:01 to 2001:10. Our strategies are thus implemented in a decade of high economic instability in Japan, which emphasizes the need for a back-testing procedure that explicitly accounts for model uncertainty.

4. Results

4.1. Small/large rotation

In this section, we outline the back-testing results for the small versus large cap rotation strategies. The results are based on the probability estimates generated by the selected logit models during the *trading* period. We present the outcomes for the different forecast horizons in Table 1. For the moment, transaction costs are assumed to be low: 25 basis points single trip. The main performance measure we apply to assess the performance of our strategies is the information ratio (IR). The table reports the significance of the mean return and of the IR for each strategy. Using a test proposed by Jobson and Korkie (1981), we formally compare the information ratios of the buy-and-hold portfolio and the switching portfolios. A statistically significant *Z*(equality)-score provides strong evidence of a significant performance difference between the switching portfolio and the buy-and-hold portfolio.[11]

Whereas the size consistent buy-and-hold strategy performed poorly, the small/large rotation strategies display some clear improvements under all examined scenarios. This is due to higher returns as well as lower standard deviations. When using a 1-month forecast horizon, the long–short rotation portfolio paid an annualized mean return of 5.40%. The corresponding information ratio (0.39) suggests there is also a substantial improvement on a return-to-risk basis. However, the observed *Z*-score does not confirm that the difference in information ratio is significant. The results furthermore imply that increasing the forecast horizon to 3 months pays off. The information ratio increased to 0.71, due to a substantially improved return (8.15%) and reduced volatility (11.42%).[12] The corresponding *Z*-score is significant at the 10% level and provides evidence that the switching portfolio significantly outperforms the buy-and-hold portfolio.

We define the 3-month forecasting strategy as the default small/large cap rotation strategy and focus on this particular strategy for the remainder of the analyses. Another look at Table 1 furthermore reveals that the improved performance is likely the result of a substantially reduced percentage of months with negative returns: 39% compared to 50% for the buy-and-hold portfolio. Consistent with this is the fact that longer horizon losses, as

---

[10] This approach methodologically resembles Jegadeesh and Titman (1993).

[11] Significance of the IR was derived using the approach suggested by Lo (2002). Computation of the *Z*-score is explained in Jobson and Korkie (1981).

[12] The Henriksson and Merton (1981) directional accuracy test confirms significant timing ability of models selected using a 3-month horizon: the statistic is significant at a 1% level.

Table 1
Out-of-sample performance of small/large cap rotation (assumed transaction costs: 25 bp single trip)

	Buy-and-hold	Small/large cap rotation	
		1-month horizon	3-month horizon
Performance statistics			
Mean return	− 1.11	5.40	8.15**
Standard deviation	14.76	13.89	11.42
Information ratio (IR)	− 0.07	0.39***	0.71***
Z(equality)		0.99	1.79*
Median	− 0.06	0.07	0.17
Minimum	− 12.14	− 12.82	− 8.03
Maximum	12.82	12.14	12.14
% Negative performance	50.00	46.48	39.44
Largest 3-month loss	− 30.18	− 17.39	− 14.30
Largest 12-month loss	− 41.28	− 25.65	− 20.21
Positions			
% Months in Large caps	0	44.37	40.14
% Months in Small caps	100	43.66	43.66
% Months in Cash	0	11.97	16.20

This table reports the performance of small versus large cap rotation strategies using 1- and 3-month forecast horizons. Assumed one-way transaction costs are 25 basis points. Out-of-sample trading period: 1990:01– 2001:10. Mean returns and standard deviations are presented on an annualized basis. The reported information ratio (IR) is the ratio of the mean return to the standard deviation. Z(equality) measures the significance of the difference in IR between the switching portfolios and the buy-and-hold portfolio. "Minimum" and "maximum" indicate the losses observed in a single month; the percentage of negative performance indicates the percentage of total months in the trading sample when incorrect positions were taken. We also present largest losses observed over 3- and 12-month holding periods. Positions in "large caps" indicate long–short positions in large–small, and positions in "small caps" refer to long–short positions in small–large.

  * Statistically significant at 10% level.
  ** Statistically significant at 5% level.
  *** Statistically significant at 1% level.

indicated by "largest 3-month (12-month) losses", were lowered considerably under the default strategy. Nevertheless, the default portfolio comprised long positions in small caps and short positions in large caps for the majority of the period (44%).

Fig. 3 shows month-to-month cumulative returns for the buy-and-hold portfolio and the default switching portfolio. It clearly confirms that the default strategy earned its return in quite a stable fashion, with more upswings than downswings. In contrast, the buy-and-hold strategy suffered some major declines from strong performances of large capitalization stocks during the period 1996-1998 and the "TMT-bubble" era. The rotation models partially anticipated these events, which strongly served the overall performance of the switching portfolio.

Fig. 4 shows the aggregate positions for the default strategy at each point in time. Since the default portfolio always consists of three positions resulting from three different signals, any value between − 1 and 0 or a value between 0 and 1 indicates a blend of two or more different positions at a certain point in time. The exact interpretation of the values is denoted below the graph. Overall, the frequency of rotation between small and large
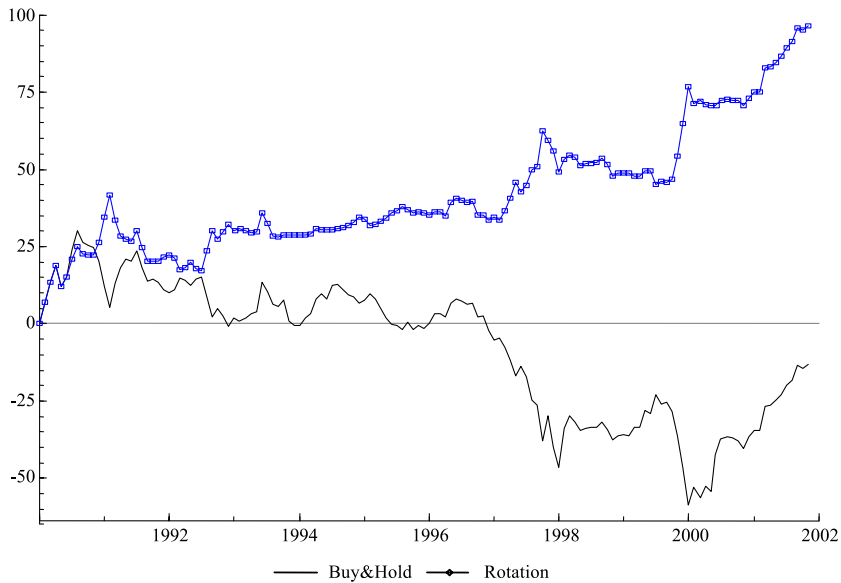
Fig. 3. Cumulative month-to-month returns for default small/large cap rotation.

caps has been quite high. Only during the last year of the trading period, the portfolio consistently comprised long (short) positions in the small (large) cap index. The fact that the portfolio frequently took on a blend of different positions may have reduced some losses from incorrect predictions, but also the gains from correct ones.

4.2. Value/growth rotation

This section reports the out-of-sample trading results for the value versus growth rotation strategies. The out-of-sample performance is summarized in Table 2. Transaction costs are still assumed low, at 25 basis point single trip.
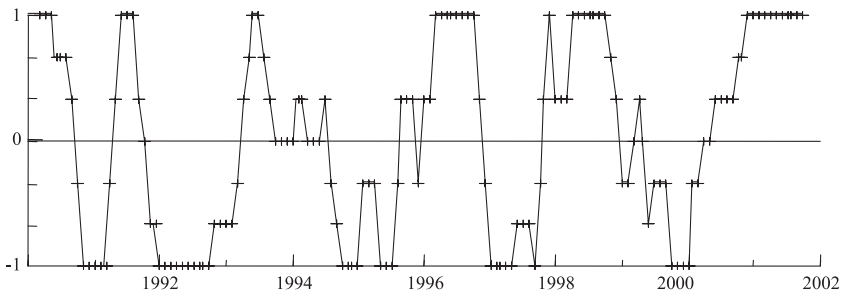

Fig. 4. Aggregate positions taken under default small/large cap rotation strategy. Note: Since the default portfolio consists of three equally weighted positions, the aggregate positions displayed in the graphs are a sum of positions divided by 3, where each position has a value of 1 whenever the generated signal indicates "small", a value of 0 when there is no signal, and a value −1 whenever the generated signal indicates "large".

Table 2
Out-of-sample performance of value/growth rotation (assumed transaction costs: 25 bp single trip)

	Buy-and-hold	Value/growth rotation	
		1-month horizon	3-month horizon
Descriptive statistics			
Mean return (%)	2.71	4.90	5.70*
Standard deviation	12.07	11.62	10.57
Information ratio (IR)	0.22***	0.42***	0.54***
Z(equality)		0.42	0.71
Median	0.23	0.12	0.12
Minimum	− 16.55	− 10.18	− 9.54
Maximum	10.64	16.55	16.55
% Negative performance	47.89	44.37	46.48
Largest 3-month loss	− 25.41	− 13.44	− 17.80
Largest 12-month loss	− 39.79	− 15.16	− 15.69
Positions			
% Months in Growth	0	38.03	31.69
% Months in Value	100	53.52	63.38
% Months in Cash	0	8.45	4.93

This table reports the performance of value versus growth rotation strategies using 1- and 3-month forecast horizons. Assumed one-way transaction costs are 25 bp. Out-of-sample trading period: 1990:01 – 2001:10. Mean returns and standard deviations are presented on an annualized basis. The information ratio (IR) is the ratio of the mean return to the standard deviation. Z(equality) measures the significance of the difference in IR between the switching portfolios and the buy-and-hold portfolio. ''Minimum'' and ''maximum'' indicate the losses observed in a single month; the percentage of negative performance indicates the percentage of total months in the trading sample when wrong positions were taken. We also present largest losses observed over 3- and 12-month holding periods. Positions in ''Growth'' indicate long–short positions in growth–value, and positions in ''Value'' indicate long–short positions in value–growth.

  *Statistically significant at 10% level.
  ***Statistically significant at 1% level.

The strategy that went consistently long on value and short on growth paid an average annualized return of 2.71%. Its corresponding information ratio has a value of 0.22. Using a 1-month forecast horizon to time the value premium provided an information ratio of 0.42. Similar to earlier reports for small/large rotation, we observe improved performance for the rotation strategy based on a 3-month forecast horizon. The strategy paid a substantially higher average annualized return (5.70%) than the passive counterpart did. Furthermore, the increase in return is not at the expense of higher volatility. With a significant information ratio of 0.54, compared to an information ratio of 0.22 for the buy-and-hold alternative, the strategy appears to provide reasonable incremental performance. The Z-scores, however, indicate that the switching portfolios do not significantly outperform the buy-and-hold strategy.

As in Section 4.1, we define results for the 3-month forecasting strategy as the default scenario. Unsurprisingly, the default strategy preferred long positions on value stocks (63%) more than long positions on growth (32%). Given that the value premium was much stronger than the size effect during the sample period, it is quite a challenge to outperform the buy-and-hold benchmark. This is also reflected in the marginal differences

in months with negative returns. Larger horizon losses remain considerably lower nonetheless. The largest observed 3- and 12-month losses were only 17.80% and 15.69%, respectively, under the default scenario, compared to 25.41% and 39.79% for the buy-and-hold portfolio.

Fig. 5 shows the time development of the cumulative month-to-month returns for the buy-and-hold portfolio and for the default rotation strategy. It evidently shows that during the period 1992–1997, the value premium was substantial and its volatility was rather low. Hence, although no incremental gains were obtained during this period, the rotation strategy actually performed quite well. Therefore, the basis of the outperformance is mainly found in the second half of the *trading period*, most notably between 1997 and 2000, during which the cumulative return gap between the rotation strategy and buy-and-hold strategy strongly widened. The returns obtained during the TMT-bubble period clearly attract the attention. The selected models quite accurately predicted the sharp return declines of value relative to growth. On the other hand, the incredible revival of value stocks that followed the puncturing of the bubble partially counterbalanced the gains obtained during previous months, because our predictive instruments initially failed to anticipate this prominent swing.

Fig. 6 shows the aggregate positions for the default value/growth rotation strategy. Recall from the previous section that a value between − 1 and 0 or a value between 0 and 1 indicates a blend of two or more different positions at a certain point in time. The rotation frequency between value and growth stock portfolios has been quite low, as has also been suggested earlier. The rate of value/growth rotation slightly increased after 1997.
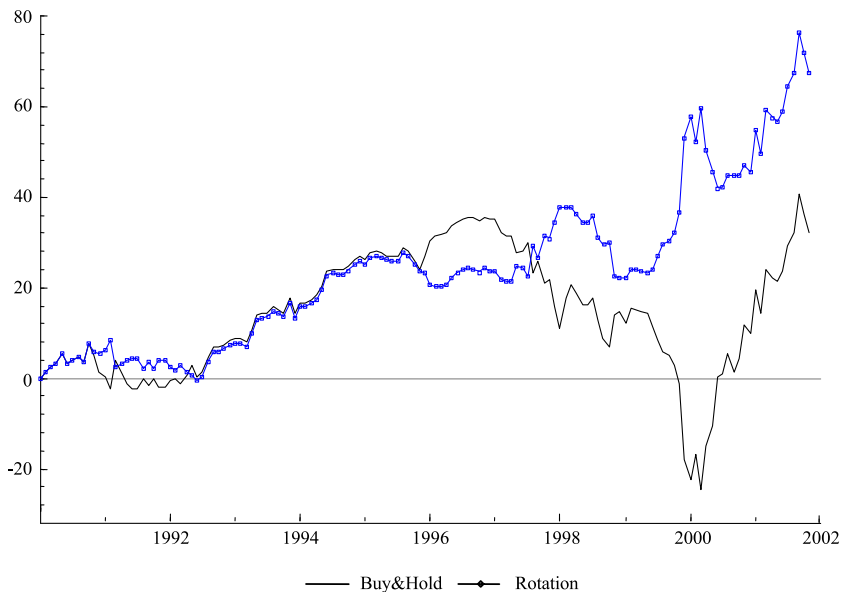

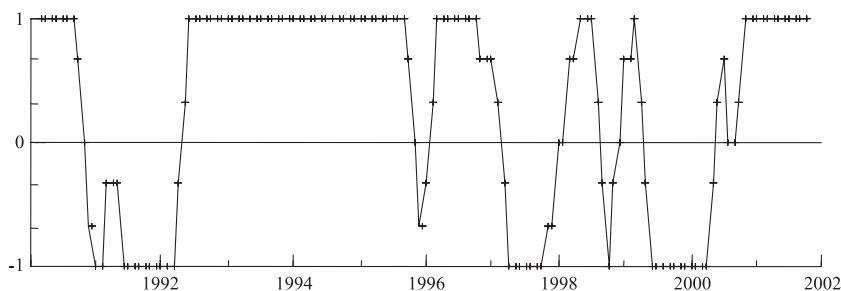Fig. 5. Cumulative returns. Value/growth rotation under default strategy.

Fig. 6. Aggregate positions taken under default value/growth rotation strategy. Note: Since the default portfolio consists of three equally weighted positions, the aggregate positions displayed in the graphs are a sum of positions divided by 3, where each position has a value of 1 whenever the generated signal indicates "value", a value of 0 when there is no signal, and a value − 1 whenever the generated signal indicates "growth".

When the robustness of the value effect started to diminish, and the TMT bubble started to develop, the portfolio was dominated by positions in growth.

4.3. Inclusion of predictive variables

To obtain more information about the relative importance of the economic variables in our base set we also report the frequency each variable was included as a determinant in the selected forecast models under the default scenarios. Fig. 7a and b shows the inclusion trough time for both small/large and value/growth rotation models. The value on the *y*-axis indicates whether a variable was included (1) or excluded (0). Inclusion percentages are reported in brackets.

At first glance, it appears that all factors provided a reasonable contribution to the overall predictive power of these particular strategies. Further examination also reveals that the default small/large rotation strategy mostly followed signals from predictive models with market related variables (MKTPREMIUM, 56%; TOPIXVAR, 42%; and DAILYVAR, 49%). Another variable included frequently is the change in M1 money supply (46%). TIBOR3, on the other hand, was underrepresented in small/large cap forecasts. Either this is due to the inappropriateness of the 3-month Tokyo Interbank Rate as a proxy for the short rate, or the short rate is simply no important predictor of size effects. Inclusion frequencies for value versus growth rotation models indicate that there is no real classification of robustness among the variables, although there was a slight preference towards DAILYVAR (49%), DY (45%) and DOIL (45%).

Although some relations appear slightly more robust over time compared to others, it is apparent that the inclusion of each of the variables varied considerably over time. The graphs thus undoubtedly indicate that in order to gain their superior relative performance the default strategies relied heavily on the ability to account for model uncertainty through permutations among the regressors. Notice that, in a real-time context, this perfectly fits the description of the Japanese financial system. It is well known that Japan's economy has been unstable during the past decades as it experienced shocks due to monetary regime switches, structural changes in the market mechanism and internal and external financial market crises. Consistent with this, our results confirm earlier suspicions that money

a

b

Fig. 7. (a) Inclusion through time of predictive variables under the default small/large style rotation strategy. (b) Inclusion through time of predictive variables under the default value/growth style rotation strategy. a and b display the inclusion of forecasting variables through time for the default rotation strategies during the trading period. A value of 1 (0) at a certain point in time implies that the particular variable was included in (excluded from) the selected model. The values in brackets report the inclusion in percentages.

managers in real time are likely to have faced a high degree of uncertainty on the optimal combination of forecasting variables, which caused them to review their perception of the "true" model frequently.

Obviously, the econometric approximation of the true (or best) forecasting model depends on the model selection methodology. Throughout the paper, we have emphasized the potential benefits of using an out-of-sample period to serve as a model selection criterion. We additionally re-examined the performance of the strategies after replacing the out-of-sample *training* period by a model selection criterion that is based on in-sample IR-maximization. A comparison of the results from this analysis, not reported here, and those

presented in this paper strongly confirms the incremental value of an out-of-sample *training* period.[13]

4.4. The effects of transaction costs

The assumption up to this point has been that the examined switching strategies are practical under low levels of transaction costs, for instance, by means of derivatives. However, since style index-based futures contracts are still unique to the United States (where there are the S&P/BARRA futures contracts), the implementation of style switching strategies under such low transaction costs is (currently) not feasible in Japan. This section therefore explores the performance of the default strategies under more appropriate transaction costs-scenarios. We assume three additional scenarios: 50, 100 and 150 basis points single trip. We are unable to account for the costs associated with the semi-annual rebalancing of the BARRA/Nikko Indexes underlying all portfolios, but we expect any bias in the Z-scores to be minor. Table 3 displays the results.

Empirical results for the default small/large rotation strategy under different levels of transaction costs are shown in columns 3–6. It is rather striking to find that increasing the level of transaction costs to 50 basis points affected the performance of the default small/large rotation strategy positively compared to the 25 basis points scenario. The information ratio increased to 0.72. One explanation for this is that the negative impact of increased transaction costs was offset by improved predictability. On the other hand, given that there are no substantial differences in long–short positions between the 25- and 50-bp scenarios, this result might as well be due to chance and should therefore be interpreted with some caution. A further increase in the level of transaction costs to 100 basis points affected the average return substantially and reduced the information ratio. The default strategy now provided an information ratio of 0.42 during the *trading* period. The corresponding Z-statistic indicates that the difference in risk-adjusted performance has become insignificant. Another increase in transaction costs to 150 basis points drastically weakened the performance of the switching portfolio. Its average return and information ratio dropped to nearly zero. Loosely speaking, this transaction cost level could be considered a break-even point between capitalization-based style switching and style consistent strategies.

Columns 7–11 in Table 3 show that the impact of transaction costs on the performance of value/growth rotation strategies is severe, which poses a challenge to investors timing style return differentials in the absence of appropriate low-cost derivative instruments. The default value/growth rotation strategy only survived an increase to 50 basis points, as it still provided an information ratio of 0.41 under this scenario. From then on, the switching portfolio underperformed the buy-and-hold portfolio due to a decreased average return. Notice, however, that the corresponding Z-scores fail to provide evidence of a significant risk-adjusted performance difference.

Although this section only reports results for 3-month return forecasts, it is also important to keep in mind that style rotation strategies based on a 1-month forecast horizon are likely to underperform the passive alternative even sooner when considering higher levels of transaction costs. Therefore, not only do our results suggest that the choice of

---

[13] Results are available upon request.

Table 3
Performance of default rotation strategies under various transaction cost scenarios

	Buy-and-hold	Default small/large rotation				Buy-and-hold	Default value/growth rotation			
		25 bp	50 bp	100 bp	150 bp		25 bp	50 bp	100 bp	150 bp
Descriptive statistics										
Mean return	− 1.11	8.15**	8.40**	4.90	0.07	2.71	5.70*	4.34	2.22	0.66
Standard deviation	14.76	11.42	11.61	11.59	11.05	12.07	10.57	10.59	11.02	11.18
Information ratio (IR)	− 0.07	0.71***	0.72***	0.42***	0.01	0.22***	0.54***	0.41***	0.20**	0.06
Z(equality)		1.79*	1.84	1.27	0.24		0.71	0.42	− 0.05	− 0.37
Median	− 0.06	0.17	0.11	0.09	− 0.02	0.23	0.12	0.01	− 0.05	− 0.19
Minimum	− 12.14	− 8.03	− 8.03	− 8.70	− 9.03	− 16.55	− 9.54	− 9.54	− 10.64	− 10.64
Maximum	12.82	12.14	12.14	12.14	11.14	10.64	16.55	16.55	16.55	16.55
% Negative performance	50.00	39.44	42.25	44.37	50.00	47.89	46.48	50.00	52.82	56.34
Largest 3-month loss	− 30.18	− 14.30	− 14.47	− 11.26	− 11.93	− 25.41	− 17.80	− 17.97	− 24.73	− 24.73
Largest 12-month loss	− 41.28	− 20.21	− 21.21	− 20.54	− 22.00	− 39.79	− 15.69	− 17.49	− 21.50	− 24.13
Positions										
% in Large (Growth)	0	40.14	40.14	37.32	37.32	0	31.69	32.39	32.39	3.80
% in Small (Value)	100	43.66	44.37	47.89	50.00	100	63.38	61.97	61.27	60.56
% Months in Cash	0	16.20	15.49	14.79	12.68	0	4.93	5.63	6.34	5.63

This table displays the performance of the default small/large and value/growth rotation strategies (1 model/3-month horizon/IR-criterion) under different levels of incurred transaction costs. Columns 2–6 show the performance of the buy-and-hold small versus large cap long–short portfolio and the results of active rotation when assuming transaction costs of 25, 50, 100 and 150 basis points single trip. Columns 7–11 are similar results for value/growth portfolios.

 * Indicates value is statistically significant at 10% level.

 ** Indicates value is statistically significant at 5% level.

 *** Indicates value is statistically significant at 1% level.

forecast horizon is a key determinant of the profitability of style switching portfolios, they also sharply contrast with many studies in which investment opportunities are usually derived from the predictability of *monthly* equity returns. In some cases, this is due to a failure of these studies to control for transaction costs arising from frequent portfolio rebalancing.

The consequences of our results are straightforward. Active style switching appears highly sensitive to the level of incurred transaction costs.[14] Given that, despite the endogenization of the choice of predictive variables, the results are still likely to be sensitive to changes in various other parameters in the modeling process (which have still been specified exogenously), it is not obvious that active portfolio switching strategies in Japan will provide substantial incremental benefits to style investors if the expenses are significant. The obtained results suggest this particularly holds for value/growth rotation, but this conclusion might be dependent on the choice of candidate determinants as well as the sample period.

5. Concluding comments

This paper examined whether short-term directional variations in the size and/or value premium in the Japanese stock market are sufficiently predictable to be exploited by means of a tactical timing strategy. Using a recursive model selection strategy based on out-of-sample model *training*, we mitigated bias resulting from ex post data snooping. Therefore, our simulation of an investor's continuous search for the optimal forecast model in real time is likely to narrow the gap between the impressive performance of simulated investment strategies reported in the academic literature and the lack thereof observed in practice.

The results from out-of-sample forecasts indicated that the size premium and the value premium are reasonably predictable in principle. Similar to Levis and Liodakis (1999) for the United Kingdom, and Cooper et al. (2001) for the United States, we provide reasonable support for small versus large, but less for value versus growth style rotation strategies in Japan. Our findings furthermore imply there are several key factors that determine the profitability of style switching strategies in the Japanese market. First, we have shown that the choice of forecast horizon is crucial to success. Confirming our initial suspicion, the results furthermore show that the ability to allow for changes in the optimal set of forecasting variables is particularly important for the Japanese market, as we observed substantial variation through time. A final key determinant of performance is the actual level of transaction costs. Our results suggest the benefits of style switching strategies are real under low transaction costs scenarios. Under scenarios with higher levels of transaction costs, however, it is difficult to obtain incremental benefits over buy-and-hold portfolios. Therefore, Japanese style rotation strategies may suffer from the lack of proper style index-based derivative products, which are currently only available in the United States.

---

[14] As a robustness check, we evaluated our strategies using a Fama and French (1993) three-factor model. The results were unaffected (if not strengthened) after controlling for multiple factor loadings.

Acknowledgements

We would like to thank the anonymous referee for valuable comments. We also thank Peter Schotman, Roger Otten, Klaas Reedijk, and colleagues at Maastricht University and ABP Investments for helpful remarks. All remaining errors are the sole responsibility of the authors. The views expressed in this paper are not necessarily shared by ABP Investments.

References

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

Allen, F., Karjalainen, R., 1999. Using genetic algorithms to find technical trading rules. Journal of Financial Economics 51, 245–271.

Arnott, R.D., Kelso, C.M., Kiscadden, S., Macedo, R., 1989. Forecasting factor returns: An intriguing possibility. Journal of Portfolio Management 16, 28–35.

Arnott, R.D., Dorian, J.L., Macedo, R., 1992. Style management: The missing element in equity portfolios. Journal of Investing 1, 13–21.

Asness, C.S., Friedman, J.A., Krail, R.J., Liew, J.M., 2000. Style timing: Value versus growth. Journal of Portfolio Management 26, 51–60.

Bae, K.H., Kim, J.B., 1998. The usefulness of earnings versus book value for predicting stock returns and cross corporate ownership in Japan. Japan and the World Economy 10, 467–485.

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

Bernstein, R., 2001. Navigate the Noise: Investing in the New Age of Media and Hype. Wiley, New York.

Bossaerts, P., Hillion, P., 1999. Implementing statistical criteria to select return forecasting models: What do we learn? Review of Financial Studies 12, 405–428.

Cai, J., 1997. Glamour and value strategies on the Tokyo Stock Exchange. Journal of Business Finance and Accounting 24, 1291–1310.

Campbell, J.Y., Hamao, Y., 1989. Predictable stock returns in the United States and Japan: A study of long-term capital market integration. NBER Working Paper 3191.

Chan, K.C., Chen, N., 1991. Structural and return characteristics of small and large firms. Journal of Finance, 451–471.

Chan, L.K.C., Hamao, Y., Lakonishok, J., 1993. Can fundamentals predict Japanese stock returns? Financial Analysts Journal 49, 63–69.

Chen, N., Zhang, F., 1998. Risk and return of value stocks. Journal of Business 71, 501–535.

Chen, N., Roll, R., Ross, S.A., 1986. Economic forces and the stock market. Journal of Business 69, 383–403.

Cochrane, J.H., 1999. New facts in finance, Center for Research in Security Prices Working Paper, 490.

Cooper, M., Gulen, H., 2002. Is time-series based predictability evident in real-time? Working Paper, Purdue University.

Cooper, M., Gulen, H., Vassalou, M., 2001. Investing in size and book-to-market portfolios using information about the macroeconomy: Some new trading rules. Working Paper, Purdue University.

Copeland, M., Copeland, T.E., 1999. Market timing: Style and size rotation using the VIX. Financial Analysts Journal 55, 73–81.

Daniel, K., Titman, S., Wei, K.C.J., 2001. Explaining the cross-section of stock returns in Japan: Factors or characteristics? Journal of Finance 55, 743–766.

Doukas, J., Hall, P.H., Lang, L.H.P., 1999. The pricing of currency risk in Japan. Journal of Banking and Finance 23, 1–20.

Fama, E.F., French, K., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–23.

Fama, E.F., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

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

Fama, E.F., French, K., 1998. Value versus growth: The international evidence. Journal of Finance 53, 1975–1999.

Foster, F.D., Smith, T., Whaley, R.E., 1997. Assessing the goodness-of-fit of asset pricing models: The distribution of the maximal R2. Journal of Finance 52, 591–607.

Gertler, M., Gilchrist, S., 1994. Monetary policy, business cycles and the behavior of small manufacturing firms. Quarterly Journal of Economics 109, 309–340.

Hamao, Y., 1988. An empirical examination of the arbitrage pricing theory using Japanese data. Japan and the World Economy 1, 45–61.

Haugen, R.A., Baker, N.L., 1996. Commonalities in the determinants of expected stock returns. Journal of Financial Economics 41, 401–439.

Henriksson, R.D., Merton, R.C., 1981. On market timing and investment performance: II. Statistical procedures for evaluating forecasting skills. Journal of Business 54, 513–533.

Hirata, H., Ueda, H., 1998. The yield spread as a predictor of Japanese recessions. Working Paper, Bank of Japan.

Jacobs, B.I., Levy, K.N., 1996. High definition style rotation. Journal of Investing 5, 14–23.

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

Jensen, G.R., Johnson, R.R., Mercer, J.M., 1998. The inconsistency of small-firm and value stock premiums. Journal of Portfolio Management 24, 27–36.

Jobson, J.D., Korkie, B.M., 1981. Performance hypothesis testing with the Sharpe and Treynor measures. Journal of Finance 36, 889–908.

Kao, D., Shumaker, R., 1999. Equity style timing. Financial Analysts Journal 55, 37–48.

Lakonishok, J., Schleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation and risk. Journal of Finance 49, 1541–1578.

Liew, J., Vassalou, M., 2000. Can book-to-market, size and momentum be risk factors that predict economic growth? Journal of Financial Economics 57, 221–245.

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

Lo, A.W., 2002. The statistics of Sharpe ratios. Financial Analysts Journal, 37–52.

Lo, A., MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3, 431–468.

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

Nakagawa, S., Osawa, N., 2000. Financial market and macroeconomic volatility: Relationships and some puzzles, Working Paper, Bank of Japan.

Perez-Quiros, G., Timmermann, A., 2000. Firm size and cyclical variations in stock returns. Journal of Finance 55, 1229–1262.

Pesaran, M., Timmermann, A., 1995. Predictability of stock returns: Robustness and economic significance. Journal of Finance 50, 1201–1228.