A RISK-SENSITIVE MOMENTUM APPROACH TO STOCK SELECTION

ABSTRACT: One of the main implications of Lo’s Adaptive Markets Hypothesis (2004, 2012, 2017) is that returns of virtually all assets can change over time. We present a local linear trend smoothing method by which this phenomenon can be captured empirically. Moreover, we introduce two localised, amended goodness-of-fit indicators capable of capturing both the direction and the continuity of recently observed price trends. Our related empirical investigation is based on a sample of 30 German blue-chip stock price series observed over a period of more than 16 years. Its results indicate that the use of these indicators as a stock-screening device can be a more useful means of identifying stocks with a superior risk/return profile than applying a conventional momentum strategy. The validity of this finding is underscored by statistical significance tests based on a Moving Blocks Bootstrap procedure.

KEY WORDS: Adaptive Markets, Local Least Squares smoothing, Moving Blocks Bootstrap

JEL CLASSIFICATION: G11, C58
1. INTRODUCTION

The Momentum Effect, i.e., the perception that stocks whose prices have been on an upward (or downward) path during the recent past are more likely than others to continue this path in the near future, is one of the most extensively studied anomalies in the equity market. Jegadeesh and Titman (1993) produced one of the first studies presenting empirical evidence for the presence of this effect. The authors find that, in the U.S. market, a strategy of sorting individual stocks by their historical performance over the most recent months, combining long positions in the best performers with short positions in the worst ones and rebalancing such a portfolio on a monthly basis, would have generated an average monthly rate of return of around 1% during the sampling period. Subsequently, many researchers have examined the time series of stock market returns in a variety of regions and for various sampling periods, looking for similar effects. Examples include, but are by no means limited to, Lee and Swaminathan (2000), Jegadeesh and Titman (2001), Rouwenhorst (1998, 1999), Chan et al. (2000); Grundy and Martin (2001); Griffin et al. (2003), Chui, Titman, and Wei (2003), and Patro and Wu (2004).

Glaser and Weber (2003) distinguish two sets of possible explanations for the momentum effect:

- Risk-based approaches rest on the assumption that financial market agents act in a fully rational manner, and that as a consequence the difference in the statistically expected returns of individual stocks captured by momentum strategies actually reflects cross-sectional variation in the extent to which these stocks are exposed to market risk. This is the viewpoint taken, inter alia, by Conrad and Kaul (1998).
- By contrast, behavioural approaches assume that the momentum effect either results from the interplay of the perceptions, reasonings, and actions of heterogenous economic agents, or reflects cognitive biases in the way these agents try to make sense of available information and act accordingly. Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Hong and Stein (1999), among others, have presented behavioural models of this kind.

If the risk-based explanation holds true, the apparent advantage of stocks with strong positive momentum should vanish as soon as differences in the risk content of individual investment alternatives are adequately accounted for. If,
on the other hand, it can be shown that stocks that displayed a particularly favourable risk/return profile in the recent past tend to outperform others in the near future, even if differences in risk are taken into account, this would lend support to the behavioural approaches.

This line of reasoning provides the starting point for our analysis. We apply a local linear smoothing procedure to time series of German blue chip stock prices (adjusted for the effects of dividends and capital measures) and introduce two variants of a localised goodness-of-fit measure with a view to simultaneously capturing the direction and the continuity of any (actual or spurious) price trends observed in the recent past. As in previous momentum-related studies, we rank individual stocks according to these newly introduced indicators and compare the risk-return profiles of the top-ranking stocks according to this scale to an equally weighted portfolio of all stocks in the sample and to those stocks with the lowest ranks, to see if there are statistically significant differences. If the ranking criteria we propose turn out to be superior to the classical momentum indicator when it comes to picking stocks with attractive risk-return profiles, this can be seen as an indication that behavioural, rather than risk-based, explanations of the momentum effect prevail.

The Adaptive Markets Hypothesis (AMH), introduced by Lo (2004) and further elaborated by the same author (Lo 2005, 2012, and 2017), provides the theoretical basis for our procedure. Due to the importance of this hypothesis for our modelling strategy, we summarize some of its main tenets in Section 2. Section 3 contains a description of the statistical techniques and measures we apply. A summary of the data used follows in Section 4. Section 5 presents our main results, and Section 6 concludes.

2. THE ADAPTIVE MARKETS HYPOTHESIS: A NEW FRAMEWORK FOR FINANCIAL MARKET ANALYSIS

The Adaptive Markets Hypothesis (AMH) can be understood as an attempt to reconcile the apparently contradictory implications of the Efficient Markets Hypothesis (EMH) on the one hand, and the various behavioural explanations for apparent deviations from the EMH, summarized, inter alia, in the survey articles by Keim (2008) and Khan (2014), on the other. Lo (2004) introduced the term ‘Adaptive Markets Hypothesis’. Borrowing from findings from various fields, including evolutionary biology and cognitive neuroscience, the author
claims that the EMH has in fact not been refuted by the empirical evidence brought forward against it. Rather, as Lo (2005) puts it,

the EMH may be viewed as the frictionless ideal that would exist if there were no capital market imperfections such as transaction costs, taxes, institutional rigidities, and limits to the cognitive and reasoning abilities of market participants.

While it is a widely held belief among economists that financial decisions are ‘rational’ in the sense that they are not impacted by emotions, Lo (2004) reaches a very different conclusion. Based on recent research in cognitive neuroscience (Grossberg and Gutowski 1987; Damasio 1994; Lo and Repin 2002; Peters and Slovic 2000; Rolls 1999), he concludes that rather than being diametrically opposite to rationality, emotions play a key role in enabling humans to evaluate the benefits and drawbacks of the different alternative actions available to them. (Damasio (1994), inter alia, provides evidence in favour of this conclusion.) Lo (2017) concludes that emotions play an important role in identifying and assessing risk, being “a reward-and-punishment system that allows the brain to select an advantageous behaviour”. Or, as the Economist magazine (2017) puts it, “If we do not fear the consequences of failure, we may act irresponsibly, just as small children need to learn to be wary of cars before crossing the road”.

Lo also explains why at times human behaviour can be irrational. Quoting research by Baumeister, Heatherton, and Tice (1994), he maintains that emotional reactions can temporarily supplant rationality, particularly in situations marked by intense levels of fear or distress. According to Sterrett (2014), regions of the brain that support emotions can react more quickly than the cognitive regions, and over the course of human evolution their ability to overtake the rational brain centres in emergency situations is the legacy of a process that helped secure the survival of humans in the face of life-threatening peril. Lo (2004) concludes that this behavioural trait may also explain why the behaviour of financial market participants exposed to extreme, adverse shocks appears irrational and erratic.

Another important building block of the AMH is the concept of “bounded rationality” introduced by Simon (1955 and 1979). According to this viewpoint, the processes of gathering and processing information and of solving complex optimization problems involve significant cost in terms of time, human effort, and (possibly) assistive technology. As Grossman and Stiglitz (1980) have
shown, once the admission has been made that information is not available for free, the efficiency of markets (in the literal sense of this term) becomes a logical impossibility:

Because information is costly, prices cannot perfectly reflect the information which is available, since if it did, those who spend resources to obtain it would receive no compensation.

Simon further argues that due to natural limitations to humans’ intellectual capacity, our actual problem-solving behaviour in real-life decision-making situations can be more realistically described as “satisficing” – an amalgam of “satisfy” and “suffice” – rather than optimizing. In Simon’s own words, as quoted in the Economist magazine (2009), this means agents “look for a course of action that is satisfactory or ‘good enough’”, and that an agent bases his decisions on “relatively simple rules of thumb that do not make impossible demands upon his capacity for thought”. On this basis, Lo (2004) likens human learning and adaptation behaviour to repeated “trial and error” processes, in the course of which patterns with unwelcome outcomes are abandoned in a way analogical to the natural selection process inherent in biological evolution. The author further concludes that as a result, “individuals develop heuristics to solve various economic challenges” that “eventually develop into approximately optimal solutions”. If, however, objectives, constraints, or environmental conditions change, irrational behaviour may result.

Lo (2004) also indicates that AMH, with its strong emphasis on evolutionary dynamics, is compatible with and closely related to two further advances in financial modelling, the agent-based approaches introduced by Arthur et al. (1997) and evolutionary game theory (see Maynard Smith 1982; Weibull 1995). A decisive feature of both these approaches is that they reject the idea embodied in the EMH that market valuation outcomes can reasonably be modelled as if they resulted from the actions taken by a single, fully informed, and rational representative agent. Rather, the authors assume that market participants differ, e.g., with regard to their risk tolerances, risk-bearing capacities, investment objectives and time horizons. Initial information endowments and investment and trading strategies also differ, but may converge as underperforming action patterns are dropped in favour of those with superior track records, and as individual agents learn to tell misleading or irrelevant pieces of information from those that are more reliable and important. In particular, the computer simulations by Arthur et al. (1997) have shown that in such a setting, depending
on the intensity with which individual agents explore alternative modes of forming expectations, markets can either settle in a rational expectations equilibrium compatible with the EMH, or can acquire “a rich psychology” in which “technical trading emerges, temporary bubbles and crashes occur, and asset prices and trading volume show statistical features... of actual market data”.

According to Lo (2005 and 2012), one of the main implications of the AMH for the patterns followed by asset price movements is that returns on virtually all asset classes may undergo cyclical variations over time. This assertion provides another key motivation for our empirical approach, which consists of fitting linear trend lines to stock price data over finite time intervals from the recent past. This way of proceeding implicitly allows for the possibility of temporary trends in stock price movements that may vary in terms of both strength and direction and can be interrupted by periods in which such movements are – or appear to be – entirely directionless and erratic. Further details are provided in the following.

3. CAPTURING VARYING TRENDS BY LOCAL LEAST SQUARES SMOOTHING

Bivariate linear regression models are based on the implicit assumption that the nature of the relationship between the conditional expectation of dependent variable \( Y \) and the explanatory variable \( X \) is representable by a straight line, and that only the slope and intercept parameters of this line are unknown and must hence be estimated from the data. In many real-world applications, however, it cannot be taken for granted a priori that this assumption is actually true. If the assumption of linearity is relaxed, the statistical relationship between a given realisation \( y_i \) of \( Y \) and the corresponding realization \( x_i \) of \( X \) can be expressed in the general form

\[
y_i = g(x_i) + u_i
\]  

(1)

where \( i \) is the index of the observation under consideration. Hastie and Tibshirani (1993) show that one way of estimating an unknown function like \( g() \) above is to use a linear function of \( X \) and a set of unknown parameters \( \alpha \) and \( \beta \), both of which are allowed to vary with the specific values taken by \( X \) in order to capture the nonlinearities potentially prevailing in the data generating process. In the above context, this would imply approximating (1) by
Here, the scalar $\epsilon_i$ represents the cumulative impact of the random disturbance term $u_i$ in (1) and any possible approximation error incurred when replacing $g(.)$ by the corresponding term in (2). Then, for any value $x^*$ that lies inside the empirically observed range of $X$, a set of related estimates can be calculated by minimizing the criterion function

$$
\Xi(\tilde{\alpha}, \tilde{\beta}) = \sum_{i=1}^{N} (y_i - \tilde{\alpha} - \tilde{\beta} \cdot x_i)^2 \cdot K_h(x^* - x_i)
$$

(3)

with respect to the parameters $\tilde{\alpha}$ and $\tilde{\beta}$. In the above equation the expression $K_h(.)$ is a kernel function, i.e., a continuous, bounded, and symmetric function that integrates to 1. An overview of some of the most commonly used kernel functions, along with a wealth of further detail, has been provided by, e.g., Härdle et al. (2004).

In our application we assume that a set of realisations of a scalar random variable $Y$ have been observed at a number of points $t = 1, \ldots, T$ in time, yielding a sample $y_t$, $t = 1, \ldots, T$ of consecutive observations. Moreover, as an explanatory variable for the movements of $Y$ over time, we intend to use the time index $t$ itself. As shown by Fan and Yao (2003, p.222), the method sketched in the paragraph above can also be applied in this setting. If the explanatory variable is the time index $t$, equation (2) can be rewritten as:

$$
y_t = \alpha(t) + t \cdot \beta(t) + \epsilon_t
$$

(2a)

Time-varying estimates of $\alpha$ and $\beta$ can then be calculated by minimizing

$$
\Xi'(\tilde{\alpha}, \tilde{\beta}) = \sum_{s=1}^{t} (y_s - \tilde{\alpha} - \tilde{\beta} \cdot s)^2 \cdot K_h(t - s)
$$

(3a)

with respect to $\tilde{\alpha}$ and $\tilde{\beta}$.

Our objective is to examine the predictive capabilities of the time-varying trends model (2). To this end, we need to put ourselves in the position of a hypothetical
observer who, during the passage of time $t$, repeatedly seeks to extrapolate from observations from the present and the recent past ($y_{t,s+1}$, with $s = 1, \ldots, h$) the likely size and direction of future movements in $Y$. A reasonably simple way of modelling this is to apply what has been referred to as a “rolling window” technique (e.g., Zivot and Wang 2006, chapter 9). In the notation used here, this is equivalent to choosing the following form for the previously unspecified function $K_h(.)$ in (3a):

$$K_h(t-s) = \begin{cases} 
0, & \text{if } s < 0 \\
(1/h) & \text{if } s \in \{0; 1; \ldots; (h-1)\} \\
0, & \text{if } s > (h-1) 
\end{cases} \quad (4)$$

Let $\hat{\alpha}(t)$ and $\hat{\beta}(t)$ denote the time-specific estimates of $\alpha$ and $\beta$ obtained by solving (3a):

$$\{\hat{\alpha}(t); \hat{\beta}(t)\} = \arg \min_{\alpha, \beta} (1/h) \sum_{s=0}^{h-1} [y(t-s) - \hat{\alpha} - \hat{\beta} \cdot (t-s)]^2 \quad (5)$$

Then, a $\tau$-period-ahead forecast $\hat{y}(t+\tau)$ of $Y$ can be based on the most recent $h$ observations as

$$\hat{y}(t+\tau) = \hat{\alpha}(t) + \hat{\beta}(t) \cdot (t+\tau) \quad (6)$$

In order to measure how closely the time path of $Y$ values observed in the time interval $[t-h+1; t]$ resemble a continuous, linear trend, a ‘localized’ version $R^2_{t,h}$ of the standard goodness of fit measure $R^2$ is applied. It owes its name to the fact that it applies only to the subset of data points observed in the above time window, and is defined as

$$R^2_{t,h} = 1 - \frac{\sum_{s=0}^{h-1} [y(t-s) - \hat{\alpha}(t) - \hat{\beta}(t) \cdot (t-s)]^2}{\sum_{s=0}^{h-1} [y(t-s) - (1/h) \sum_{s=0}^{h-1} y(t-s)]^2} \quad (7)$$
On this basis, another localized goodness-of-fit measure \( r_{t,h}^{2,*} \) can be introduced that jointly captures the direction of a trend and its degree of resemblance to a straight line. It is calculated by pre-multiplying (7) with the sign of the estimated slope parameter \( \hat{\beta}(t) \) as follows:

\[
\begin{align*}
r_{t,h}^{2,*} &= R_{t,h}^2 \text{ if } \hat{\beta}(t) > 0 \\
&= 0 \text{ if } \hat{\beta}(t) = 0 \\
&= -R_{t,h}^2 \text{ if } \hat{\beta}(t) < 0
\end{align*}
\]  
(8)

A key advantage of the above \( r_{t,h}^{2,*} \) measure is that it is confined to the interval \([-1; 1]\), which facilitates its interpretation: \( r_{t,h}^{2,*} = 1 \) implies a perfectly continual, linear upward trend, whereas \( r_{t,h}^{2,*} = (-1) \) indicates a perfectly linear downward trend. This benefit, however, contrasts with the drawback that \( r_{t,h}^{2,*} \) does not discriminate between strong and weak upward or downward trends (indicated by high absolute values of \( \hat{\beta}(t) \)). Using another modification of the localized \( R \)-square measure can circumvent this problem:

\[
r_{t,h}^{2,**} = R_{t,h}^2 \cdot \hat{\beta}(t)
\]  
(9)

The \( r_{t,h}^{2,**} \) criterion seeks to combine information on the direction, magnitude, and direction of a trend. It can thus be deemed to be more informative than \( r_{t,h}^{2,*} \) because it also accounts for the slope of the trend line. However, since there is no pre-defined upper or lower limit on this coefficient, unlike version (8) it does not lend itself to an easy qualitative interpretation.

In our application, both indicators (8) and (9) will be examined for their suitability for stock selection purposes. To this end, we apply the following procedure:
At the end of each month in the sampling period, a regression of the type described in Equations (2a) to (5) is performed for each stock in the sample. The length $h$ of the formation period in use is set to 252 trading days, i.e., approximately one year. Along with the standard 12-month momentum indicator, the $r_{t,h}^{*}$ and $r_{t,h}^{**}$ measures introduced in (8) and (9) are calculated for all candidate stocks, which are then sorted according to the outcome obtained.

On this basis, two equally weighted portfolios consisting of (a) the top 10% and (b) the bottom 10% of stocks are formed.

For each of these hypothetical portfolios, the performance in the subsequent one-month period is examined and compared to the average for all candidate stocks in the same month. This procedure is performed for each month in the sample for which a twelve-month history of past prices is available.

In order to characterise the risk–return profiles of the individual portfolios during the sampling period we use a total of five indicators, which are

(i) the average monthly rate of return,
(ii) the standard deviation of monthly returns,
(iii) the maximum drawdown, i.e., the loss, expressed as the percentage of initial wealth that an investor would have faced if they had bought a defined asset or bundle of assets at its peak price and sold at the lowest subsequent one,
(iv) the volatility-adjusted average monthly rate of return (or return over standard deviation (ROSD)) obtained by dividing (i) by (ii),
(v) the average monthly rate of return relative to the maximum drawdown (or return over maximum drawdown (ROMD)), calculated by dividing (i) by (iii).

Observed differences in the risk-return profiles of different groups of stocks may either be systematic or entirely random, at least in part. Hence, the need arises to test these differences for statistical significance. However, the very fact that the portfolios under investigation are constructed on a monthly basis using information from mutually overlapping twelve-month calibration periods introduces some serial dependence into the resulting sequence of returns, which may render the outcomes of the standard significance test based on the normal approximation unreliable. A possible solution to this problem is to use the Moving Blocks Bootstrap method pioneered by Hall (1985), Carlstein (1986),
Liu and Singh (1992), and Künsch (1989). The advantage of this procedure is that it can lead to more accurate estimates of the sampling distribution whenever there is a short-run dependence between the individual data points (see Lahiri 1990). The Moving Blocks Bootstrap consists of dividing the time series under examination into blocks of successive observations, and randomly sampling from among these blocks with replacement to generate a large number of simulated quasi-samples. By re-estimating the statistics of interest for each of these quasi-samples and evaluating the empirical cumulative frequency distribution of the individual outcomes obtained, it is possible to construct confidence intervals and perform significance tests. In line with Künsch (1989), we use overlapping blocks of data. Moreover, we ‘wrap’ the data points around in a circle prior to forming and drawing blocks, as recommended by Politis and Romano (1992). According to the last-mentioned authors, the advantage of this way of proceeding is that the blocks are automatically centred around the mean, which is an important prerequisite for the reliability of the results. We follow Hall, Horowitz, and Jing’s (1995) recommendation and set the block size to $N^{1/5}$, rounded to the nearest integer, where $N$ is the sample size.

### 4. DATA AND SOURCES

Adjusted prices for a total of 30 German blue-chip stocks were observed and collected over the period January 1, 2000 to September 30, 2017. With the main source of data being Yahoo! Finance (de.finance.yahoo.com), we sought to iron out occasional gaps and glitches in the available time series by resorting to other publicly available sources, like the financial websites finanzen.net and comdirect.de. Due to data availability and quality issues, we were unable to obtain adjusted price series for the entire set of stocks that were part of the DAX30 index during the sampling period, but had to limit our attention to the thirty stocks listed, together with some descriptive statistics, in Table 1 below.
### Table 1

<table>
<thead>
<tr>
<th>Name</th>
<th>Sampling period</th>
<th>Descriptive Statistics</th>
<th>...for daily returns</th>
<th>...for rolling one-month returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First month</td>
<td>Last month</td>
<td>Mean</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>ADIDAS</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0643%</td>
<td>1.9143%</td>
</tr>
<tr>
<td>ALLIANZ</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0037%</td>
<td>2.2434%</td>
</tr>
<tr>
<td>BASF</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0453%</td>
<td>1.8588%</td>
</tr>
<tr>
<td>BAYER</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0306%</td>
<td>2.0614%</td>
</tr>
<tr>
<td>BEIERSDORF</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0468%</td>
<td>1.7953%</td>
</tr>
<tr>
<td>BMW</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0379%</td>
<td>2.1223%</td>
</tr>
<tr>
<td>COMMERZBANK</td>
<td>2000–01</td>
<td>2017–09</td>
<td>-0.0703%</td>
<td>2.9213%</td>
</tr>
<tr>
<td>CONTINENTAL</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0846%</td>
<td>1.7537%</td>
</tr>
<tr>
<td>DAIMLER</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0127%</td>
<td>2.1671%</td>
</tr>
<tr>
<td>DEUTSCHE BÖRSE</td>
<td>2001–02</td>
<td>2017–09</td>
<td>0.0322%</td>
<td>2.2919%</td>
</tr>
<tr>
<td>DEUTSCHE BANK</td>
<td>2000–01</td>
<td>2017–09</td>
<td>-0.0196%</td>
<td>2.5765%</td>
</tr>
<tr>
<td>DEUTSCHE POST</td>
<td>2000–11</td>
<td>2017–09</td>
<td>0.0297%</td>
<td>1.8403%</td>
</tr>
<tr>
<td>DEUTSCHE TELEKOM</td>
<td>2003–01</td>
<td>2017–09</td>
<td>-0.0124%</td>
<td>2.0930%</td>
</tr>
<tr>
<td>E.ON</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0223%</td>
<td>1.9685%</td>
</tr>
<tr>
<td>FRESENIUS MED</td>
<td>2003–01</td>
<td>2017–09</td>
<td>0.0353%</td>
<td>1.8762%</td>
</tr>
<tr>
<td>FRESENIUS CARE</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0620%</td>
<td>2.4100%</td>
</tr>
<tr>
<td>HEIDELBERG CEMENT</td>
<td>2003–01</td>
<td>2017–09</td>
<td>0.0102%</td>
<td>2.4282%</td>
</tr>
<tr>
<td>HENKEL VZ</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0478%</td>
<td>1.6036%</td>
</tr>
<tr>
<td>INFINEON</td>
<td>2000–03</td>
<td>2017–09</td>
<td>-0.0203%</td>
<td>3.6398%</td>
</tr>
<tr>
<td>K+S</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0805%</td>
<td>2.4907%</td>
</tr>
<tr>
<td>LUFTANSA</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0112%</td>
<td>2.1404%</td>
</tr>
<tr>
<td>LINDE</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0355%</td>
<td>1.8069%</td>
</tr>
<tr>
<td>MERCK KGAA</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0563%</td>
<td>1.9278%</td>
</tr>
<tr>
<td>PROSIEBEN MEDIA</td>
<td>2001–05</td>
<td>2017–09</td>
<td>-0.0124%</td>
<td>3.0157%</td>
</tr>
<tr>
<td>RWE</td>
<td>2000–01</td>
<td>2017–09</td>
<td>-0.0028%</td>
<td>1.9642%</td>
</tr>
<tr>
<td>SAP</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0226%</td>
<td>2.3254%</td>
</tr>
<tr>
<td>SIEMENS</td>
<td>1999–10</td>
<td>2016–01</td>
<td>0.0115%</td>
<td>2.4137%</td>
</tr>
<tr>
<td>SALZGITTER</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0367%</td>
<td>2.5634%</td>
</tr>
<tr>
<td>THYSSENKRUPP</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0067%</td>
<td>2.3690%</td>
</tr>
<tr>
<td>VONOVIA</td>
<td>2013–07</td>
<td>2017–09</td>
<td>0.0962%</td>
<td>1.4291%</td>
</tr>
<tr>
<td>VW VZ</td>
<td>2000–01</td>
<td>2017–09</td>
<td>0.0430%</td>
<td>2.4054%</td>
</tr>
</tbody>
</table>
5. RESULTS

5.1. Risk–Return Characteristics of Individual Portfolios

Pure Momentum as Ranking Criterion

Descriptive statistics for the monthly returns on the portfolios formed using the pure momentum strategy are summarized in Table 2. At least at first sight, the results obtained seem to confirm the presence of the momentum effect: The equally weighted portfolio formed from the three best-performing stocks during the preceding 12-month formation period, labelled TOP3, outperforms the sample average (AVG) by 53 basis points during the sampling period. Conversely, the average monthly return on the BOT3 portfolio, made up of the three bottom ranking stocks in terms of historical 12-month performance, lags the sample average by 32 basis points.

![Table 2](image)

We also compare the median of monthly returns, because this is a measure of centrality that is less affected by rare observations with extremely positive or negative values. By this measure, the TOP3 portfolio slightly underperforms the sample average, but the returns on both by far exceed the median return on the BOT3 portfolio, which is close to zero. Thus, the momentum strategy owes at least some of its apparent superiority to its ability to capitalise on rare upward
jumps in individual stock prices (which are more strongly reflected by the sample mean than by the median).

However, at least in the case examined here, the perceived superiority of the historical ‘momentum winners’, at least in terms of average monthly returns, seems to be little more than a mirror image of the higher potential for loss associated with the choice of a corresponding strategy. This becomes apparent if we compare the standard deviations of monthly returns for each of the portfolios under consideration: The value obtained for the TOP3 portfolio exceeds the sample average by more than three percentage points and is also above that for BOT3 (although here the difference is substantially smaller). The indicators of downside risk shown in Table 2, which are more informative in cases where return distributions are asymmetric or heavy-tailed, point in the same direction: For the TOP3 portfolio the maximum drawdown, as well as the absolute value of the minimum monthly return and the 1% and 5% quantiles of the return distribution, exceed the corresponding values for the sample average and the BOT3 portfolio by a considerable amount. On the other hand, the BOT3 portfolio appears riskier than average with respect to all related measures employed. This leaves the impression that, at least on a risk-adjusted basis, the 12-month momentum criterion offers better guidance as to which stocks to avoid than which ones to choose.

\( r^2 * \) as the Ranking Criterion

The results obtained change considerably if the pure momentum indicator is replaced by the first of the two localized goodness-of-fit measures introduced in section 3 (i.e., \( r^2 * \)); see Table 3. In this case, the top-ranking portfolio outperforms the sample average and BOT3 in terms of both average and median returns. The difference in the average monthly returns of TOP3 and AVG is, however, smaller than in the case of the momentum strategy. This finding is consistent with the fact that the \( r^2 * \) indicator only reflects the direction and the continuity of a (spurious or actual) trend and is not sensitive to its strength. If \( r^2 * \) replaces the pure momentum as a stock selection criterion, the difference in the worst-case monthly loss between the TOP3 and AVG portfolios shrinks to 2.36 percentage points. In terms of the volatility of monthly returns, as well as measured by their 1% and 5% quantiles, the downside risk associated with TOP3 and AVG appears comparable in size, whereas the BOT3 portfolio looks far more risky. However, the most striking difference between the strategies based on these two indicators is regarding the Maximum Drawdown indicator. Here, the TOP3 portfolio outperforms the sample average by more than ten percentage points and the
BOT3 portfolio by more than thirty. The $r^{2*}$ indicator thus seems to be particularly useful, and clearly superior to the simple momentum indicator, in identifying stocks with less-than-average downside risk.

### Table 3

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>TOP3</th>
<th>AVG</th>
<th>BOT3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>-25.00%</td>
<td>-22.64%</td>
<td>-26.02%</td>
</tr>
<tr>
<td>1% quantile</td>
<td>-20.87%</td>
<td>-21.18%</td>
<td>-24.27%</td>
</tr>
<tr>
<td>5% quantile</td>
<td>-9.13%</td>
<td>-9.40%</td>
<td>-17.69%</td>
</tr>
<tr>
<td>10% quantile</td>
<td>-4.51%</td>
<td>-6.11%</td>
<td>-11.62%</td>
</tr>
<tr>
<td>25% quantile</td>
<td>-1.84%</td>
<td>-1.93%</td>
<td>-4.22%</td>
</tr>
<tr>
<td>Median</td>
<td>1.68%</td>
<td>1.43%</td>
<td>1.27%</td>
</tr>
<tr>
<td>Arithmetic mean</td>
<td>1.35%</td>
<td>0.97%</td>
<td>0.54%</td>
</tr>
<tr>
<td>75% quantile</td>
<td>4.81%</td>
<td>5.92%</td>
<td>5.86%</td>
</tr>
<tr>
<td>90% quantile</td>
<td>7.67%</td>
<td>9.61%</td>
<td>10.00%</td>
</tr>
<tr>
<td>95% quantile</td>
<td>9.29%</td>
<td>13.74%</td>
<td>13.18%</td>
</tr>
<tr>
<td>99% quantile</td>
<td>14.23%</td>
<td>21.62%</td>
<td>26.97%</td>
</tr>
<tr>
<td>Maximum</td>
<td>18.76%</td>
<td>25.86%</td>
<td>34.60%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>5.70%</td>
<td>5.89%</td>
<td>8.11%</td>
</tr>
<tr>
<td>Maximum Drawdown</td>
<td>43.72%</td>
<td>56.45%</td>
<td>76.34%</td>
</tr>
</tbody>
</table>

$r^{2**}$ as the Ranking Criterion

The results summarized in Table 4 indicate that the use of the $r^{2**}$ measure introduced in equation (9) is about as effective in separating stocks with low downside risk from those with high potential for loss as the $r^{2*}$ criterion examined above. However, the difference between the average monthly returns of the TOP3 and the AVG portfolios, which amounts to 69 basis points, is considerably larger than in the previous case (38 basis points) and also exceeds the value obtained for the momentum strategy (53 basis points). This finding fits well with the fact that $r^{2**}$ combines information on both the strength and the continuity of a trend prevailing in the formation period. In terms of most downside risk measures in use, the quality of the TOP3 portfolio constructed using the $r^{2**}$ indicator roughly matches that based on $r^{2*}$. The only important exception from this rule is the 10% quantile, in terms of which the top-ranking portfolio using $r^{2**}$ turns out to be around two percentage points riskier than the one based on $r^{2*}$. In total, however, at least judged by the descriptive statistics presented in this section, ranking stocks according to the $r^{2**}$ criterion appears to offer the most promising strategy for achieving risk-adjusted outperformance.
5.2. Statistical Significance Tests for Observed Differences

By looking at descriptive statistics alone, as we have done so far, it is impossible to judge whether the observed differences in the portfolio characteristics are entirely random or whether there are signs suggesting that they are systematic in nature. In order to address this question we use the moving blocks bootstrap procedure sketched in Section 3 to construct two-sided confidence intervals for the differences in (1) the average monthly return (AR; first row in Figure 1 below), (2) the standard deviation of monthly returns (SD; second row in Figure 1), (3) the maximum drawdown statistic (MD; third row in Figure 1), as well as (4) the average return over standard deviation (fourth row in Figure 1) and (5) the return over maximum drawdown (fifth row in Figure 1) measures. For each indicator in use we calculate the differences in the related quantities of the individual portfolios in a pairwise manner, i.e., by comparing (a) the TOP3 and BOT3 portfolios (b) the TOP3 and AVG portfolios, and (c) the AVG and the BOT3 portfolios. In Figure 1, which summarizes the results, the upper and lower limits of the estimated 99%, 95%, and 90% confidence intervals for the quantity under consideration are represented by white, light grey, and dark grey dots respectively, and the black dot in the sample represents the estimated median.

It turns out that as long as we focus on the average monthly returns in isolation, the differences observed between the individual portfolios are mostly...
statistically insignificant, as the zero line lies inside the 90% confidence intervals in eight of the nine cases under consideration. The only exception from this rule is the $r^{2**}$ indicator: Here, difference in average returns between the top-ranking portfolio and the sample average is statistically significant at the 95% level. This finding, however, is offset by the fact that significant differences in average returns cannot be found between either the TOP3 and BOT3 portfolios or between the AVG and BOT3 portfolios. It can thus be concluded that if the analyst’s only objective is to separate ‘winners’ from ‘losers’, and if risk considerations are deliberately disregarded, two of the three indicators under investigation (pure momentum and $r^{2*}$) are of little use, and in this context the statistical reliability of the third ($r^{2**}$) is, at least, limited.

However, the picture changes substantially as soon as differences in the risk indicators are taken into account. Looking at the pure momentum indicator first, we find that for both the highest-ranking and the lowest-ranking stocks, the volatility of the monthly returns significantly exceeds the sample average. On one hand, this indicates that the apparently superior performance of the momentum winners in the sample indeed came at the cost of higher risk, while on the other it seems to indicate that the assumption of a simple, monotonic relationship between expected return and risk is not tenable. If stocks are ranked using the $r^{2*}$ and $r^{2**}$ criteria the results turn out to be very different: There are no statistically significant differences between the top-ranking stocks and the sample average, whereas the lowest-ranking stocks are significantly more volatile than the others.

Turning to the maximum drawdown measure, we find that for the momentum strategy the reported differences in the related statistics for the three portfolios under consideration turn out to be statistically insignificant. Once again, the results for the $r^{2*}$ and $r^{2**}$ criteria are strikingly different: the top-ranking portfolio and the sample average are significantly less risky than the lowest-ranking portfolio if a confidence level of 95% (in the case of TOP3) and 99% (in the case of AVG) is applied. The difference in the maximum drawdown statistic between the TOP3 and AVG statistic, however, is insignificant if $r^{2*}$ and $r^{2**}$ are used as ranking criteria.

The discriminatory power of the $r^{2*}$ and $r^{2**}$ criteria is statistically most significant when it comes to identifying stocks with a favourable relationship between expected return and risk, particularly if the latter is measured by the maximum drawdown statistic. This becomes obvious when examining the last
two rows of Figure 1. In the case of the momentum indicator, both the highest-ranking stocks and the sample average significantly outperform the lowest-ranking stocks in terms of the ROSD and ROMD statistics, but the difference in ROSD and ROMD observed between TOP3 and AVG is numerically close to zero and statistically insignificant. In the case of the $r^{2*}$ indicator the differences between TOP3 and AVG are substantially higher and statistically significant at a 95% (ROMD for $r^{2*}$) or 99% (ROSD for $r^{2*}$, ROMD and ROSD for $r^{2**}$) confidence level. Statistically significant differences in terms of ROSD and ROMD also prevail between AVG3 and BOT3, as well as TOP3 and BOT3 if portfolio selection is based on $r^{2*}$ and $r^{2**}$, which further strengthens the impression conveyed above.
Although some of our results are compatible with the risk-based explanation of the momentum effect, the support that the sum of our findings lends to this approach is very limited. Rather, it seems to confirm the presumption that agents in the stock market form their expectations of future price movements by extrapolating from the recent past, and that in doing so agents do not exclusively focus on the strength and direction of perceived trends but also take their perceived stability and continuity into account. This seems to be more in line with the Adaptive Markets Hypothesis discussed in Section 2 than with the classical assumption of a monotonous relationship between risk and expected return.

6. CONCLUDING REMARKS

The probabilistic laws governing the movement of stock prices are not necessarily stable over time but may be subject to change. In line with Lo’s Adaptive Markets Hypothesis (2004, 2012, 2017), adapting to such changes can be assumed to be a gradual rather than an instantaneous process, during which agents with heterogeneous information sets, objectives, constraints, and capabilities learn by trial and error as well as imitation. If this assumption is realistic, the returns on virtually all assets can change with the passage of time. The local linear trend smoothing method presented in Section 3 is an appropriate means of modelling this phenomenon empirically. In addition, we have introduced two localised, amended goodness-of-fit indicators capable of capturing both the direction and the continuity of recently observed price trends. The results of our related empirical investigation suggest that the use of these indicators as a stock screening device probably offers a more effective way of identifying stocks with a superior risk/return profile than the application of a conventional momentum strategy. The validity of this finding is underscored by statistical significance tests based on the Moving Blocks Bootstrap procedure.

The results obtained give rise to a number of open questions and suggestions for future research. For example, the size of the historical time windows for which the local trend lines are estimated could be treated as a separate parameter which itself needs to be optimized based on past experience, rather than being stipulated a priori as in our paper. Moreover, the possibility could be explored of smoothing and extrapolating time paths of stock price series with the help of local polynomials (see, e.g., Proietti and Luati 2008), rather than using the locally linear approach we suggested. Localised versions of nonparametric test statistics for trend detection (see, e.g., the survey by Morell and Fried (2009))
may also be promising candidates for possible stock selection criteria. It would also be interesting to know whether findings comparable to the ones reported here can be made in segments of the equity market other than the one we examined.

REFERENCES


A RISK-SENSITIVE MOMENTUM APPROACH TO STOCK SELECTION


Received: January 15, 2019
Accepted: March 4, 2019