Performance of Candlestick Analysis on Intraday Futures Data

J. HENNING FOCK, CHRISTIAN KLEIN, AND BERNHARD ZWERGEL

J. HENNING FOCK
is a Ph.D. candidate in finance at the University of Augsburg in Augsburg, Germany.

CHRISTIAN KLEIN
is a Ph.D. candidate in statistics at the University of Augsburg.

BERNHARD ZWERGEL
is a Ph.D. candidate in statistics at the University of Augsburg.

Candlestick analysis is a form of technical analysis based on recognition of visual patterns in market returns. Candlestick analysis is very popular in Asia.

This analysis of the predictive power of candlestick patterns uses daily intraday DAX and Bund futures data starting January 1, 2002, and ending December 31, 2003. A mathematical definition scheme enables an objective and computerized identification of the candles.

Even without taking transaction costs into account, the results are quite poor. In most cases, the results are not significantly better than results for a benchmark with randomized transactions.

For many years technical analysis (TA) has been a contentious issue between academicians and financial practitioners. While the academic world is still discussing whether charting works, or whether it is nothing more than voodoo finance, practitioners have been using technical analysis for decades.

There is considerable empirical work on the performance and the predictive power of TA, with ambiguous or negative results. Examples are Hofmann [1973], Curcio et al. [1997], Mills [1997], Goldbaum [1999], and Dorfleitner and Klein [2002].

Research providing at least some positive findings includes Sweeney [1986], Pruitt and White [1988], Lukac and Brorsen [1990], Brock, Lakonishok, and LeBaron [1992], Pruitt, Tse, and White [1992], Balsara, Carlson, and Rao [1996], Detry and Grégoire [2000], Fernandez-Rodriguez, Gonzalez-Martel, and Sosvilla-Rivero [2000], Lo, Mamaysky, and Wang [2000], Osler [2000], Ligh, Paz, and Purvis [2002], Wong, Manzur, and Chew [2002], and Shen [2003]. A caveat is that these studies are often difficult to compare because of differences in statistical tests, time periods, technical trading rules, markets, and provision for transaction costs.

Empirical research on TA trading rules is thus problematical, with many methodological problems. In most cases historical data are used, although Ball, Kothari, and Wasley [1995] show that the use of historical data may produce an upward bias. Researchers working with historical data are also sometimes criticized for a lack of objectivity. Jegadeesh [2000] is one of the authors to broach this issue.

From a theoretical point of view, it should be impossible to earn abnormal returns with technical analysis, at least not if we accept the weak form of the efficient market hypothesis (see, e.g., Fama [1970, 1991] and Malkiel [1996]).

Although academicians have still not found a reason why TA should work, charting is widespread in the industry, as described in Tyler and Allen [1992] and Bankhofer and Hilbert [1999]. Two reasons for the enormous popularity of TA in the practical finance world is that it is simpler than "fundamental" methods and is reputed to have some forecasting power.[1]

There are many TA methods. Some people distinguish between charting (the recognition of visual patterns in charts) and market technique (the calculation of buy and sell signals using mathematical definitions).

The *candlestick technique* is based on the recognition of visual patterns that take the shape of candlesticks, a special method to see the behavior of share prices. Candlesticks or candles are very popular in Asia, and their popularity is growing in the West as well; see, e.g., Okamoto [2003].

The candlestick technique, like most other TA techniques, is often used for timing decisions, which means finding the right time to buy or sell a share. This may be the reason candlestick charting is widespread among day traders and short-term institutional investors for whom timing is more important than stock selection because of their short time horizon, as described in Okamoto [2003].[2]

An interesting note in this regard is that most of the empirical research related to TA is based on daily data; examples are Osler [2000] and Raj [2000]. One reason could be that handling intraday data may lead to several technical problems, as described in Goodhart and O'Hara [1997].

We analyze the predictive power of candlestick patterns using intraday data, a method not so far published in the finance literature, to our knowledge. We hope to lay the foundation for further research on candlestick analysis as well as research involving intraday data.

I. DATA, CANDLESTICKS, INDICATORS, AND BENCHMARK

We first explain the data we use and the way we define the candlesticks and indicators, and then our choice of benchmark.

Data

Most studies dealing with technical trading are based on daily data. Technical progress in the last few years, however, has made intraday data easily available, and current software and hardware can process much greater amounts of data. Intraday futures data are thus an obvious choice for this study.[3]

The futures with the highest turnover on the German futures exchange, Eurex, are the DAX futures on the German stock index DAX (abbreviated FDAX) and the Bund futures on German government bonds (abbreviated FGBL). We use original Eurex tick data to build continuous contracts for both futures.[4]

From these continuous contracts we construct five-minute intraday candlesticks for the sample 2002–2003 period. Because trades did not take place every second, we use the trade price closest to the beginning and end of the five-minute interval to build each candle. As a result of frequent trading, most candles covered more than four minutes and 50 seconds. Even at times of thin trading, there were no five-minute intervals without a trade.

Our software had to begin and end trading every day at the same time. Sometimes the Eurex did not start trading punctually, and a certain time delay was needed for the indicators to become available. Furthermore, there was sometimes very low volume at the end of the trading day, which would make it difficult to close actual positions. We therefore decided to start trading in the FDAX at 9:40 AM and to settle all positions by 7:45 PM. In our study, trading on the FGBL starts at 8:40 AM and ends at 6:45 PM.

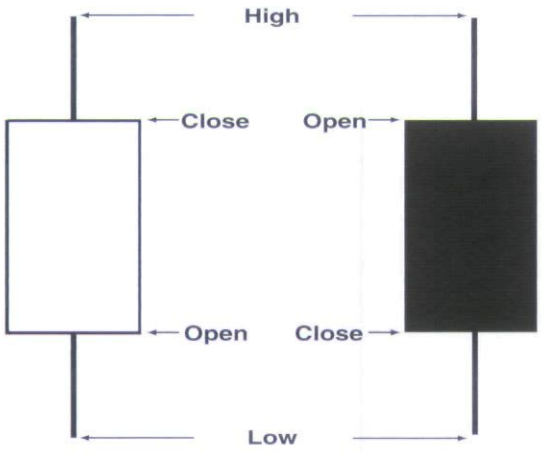
Candlesticks

The origins of candlestick charting seem to lie in Japan, where they were applied to the trading of forward contracts on rice in the 17th century. A tax on rice due to the seigneurs led to development of a market for these contracts. Soon it was also possible to sell future tax revenues forward (a type of rice forward). These rice forwards were then traded frequently, laying the foundation for the first derivatives exchange in the world, described in Schwager [1996, p. 277].

Exhibit 1 demonstrates the construction of candlestick charts. Every candle includes information on the high, low, opening, and closing prices of a particular time interval. The figurative "body" of the candle shows the difference between opening and closing price, and its length depends on this difference. If the closing price is higher than the opening price, the body is white, which signals rising prices. If the opening price is higher than the closing price, the body is black, which signals falling prices. Above and below the candle's body are the "shadows," called upper shadow and lower shadow. They depict the high and the low of the trading interval.

Although the method of candlestick analysis is quite old, there seems to be no scientific research dealing with the ability of candlesticks to forecast future market developments. Any publications dealing with this issue are in the style of popular science, presenting only the patterns and a description with no mathematical definition of the patterns and no performance study.

EXHIBIT 1
Candlestick Patterns



We find approximately 60 different candlestick patterns described, ranging from patterns with one candle to patterns with six or more candles. We rely on Nison [1991, p. 289ff.], Wagner and Matheny [1994, p. 21ff.], Landi [1997], Montassèr [2000], and Okamoto [2003] as sources for the description of the candles. As there are distinct differences among the specific patterns in these sources, we decided to use the 19 patterns that are plotted nearly identically in almost every publication on candlestick charting.

Candlestick patterns tend to be based on a few standard patterns from which a new pattern is constructed by mirroring, inverting, or adding another candle. An interesting question is whether forecasting performance is improved by extending the patterns, or whether there are other systematic connections among patterns built according to this scheme.

Exhibit 2 demonstrates this principle. The *Rain Drop* by extension becomes the *Morning Star*. It becomes the *Star* by mirroring and inverting and the *Evening Star* by mirroring, inverting, and extending. The *Rain Drop* and *Morning Star* are buy signals; the *Star* and *Evening Star* are sell signals.

Exhibit 3 pictures the patterns of all the candles in our study, and Exhibit 4 defines them by color, body, price, and proportion.

We find no mathematical definitions for the candlesticks in the popular literature, so we had to construct new definitions as a basis for scientific analysis. The colors, sizes, and proportions of the candles in this study are chosen in a subjective manner, using the charts we find in various publications.

The general problem is this subjectivity in identifying candles only visually. We need mathematical definitions to overcome this subjectivity. With the help of mathematical definitions, we lay the foundation for future research on candlesticks, much as Lo, Mamaysky, and Wang [2000] did for the analysis of charting through their systematic definition of charting patterns.

The parameters used to define the candlestick patterns are the relations of the opening (o), high (h), low (l), and closing (c) prices in the five-minute intervals. The relations are expressed by the mathematical signs >, <, =, >=, and <=. The candle's *real body size* is computed relative to the closing price as follows:

$$rb = \frac{|o - c|}{c} \qquad (1)$$

If the closing price is higher than the opening price, the body is white; if the closing price is lower, the body is black. With these parameters it is possible to define

EXHIBIT 2
Extension and Mirroring of *Rain Drop* Pattern

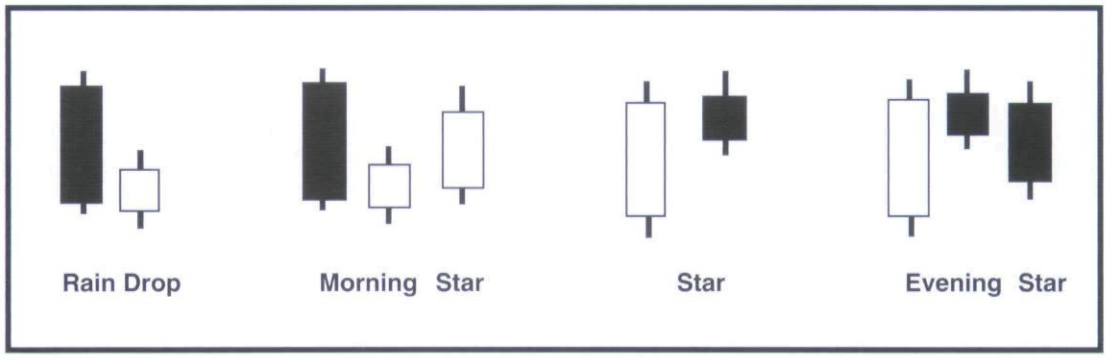
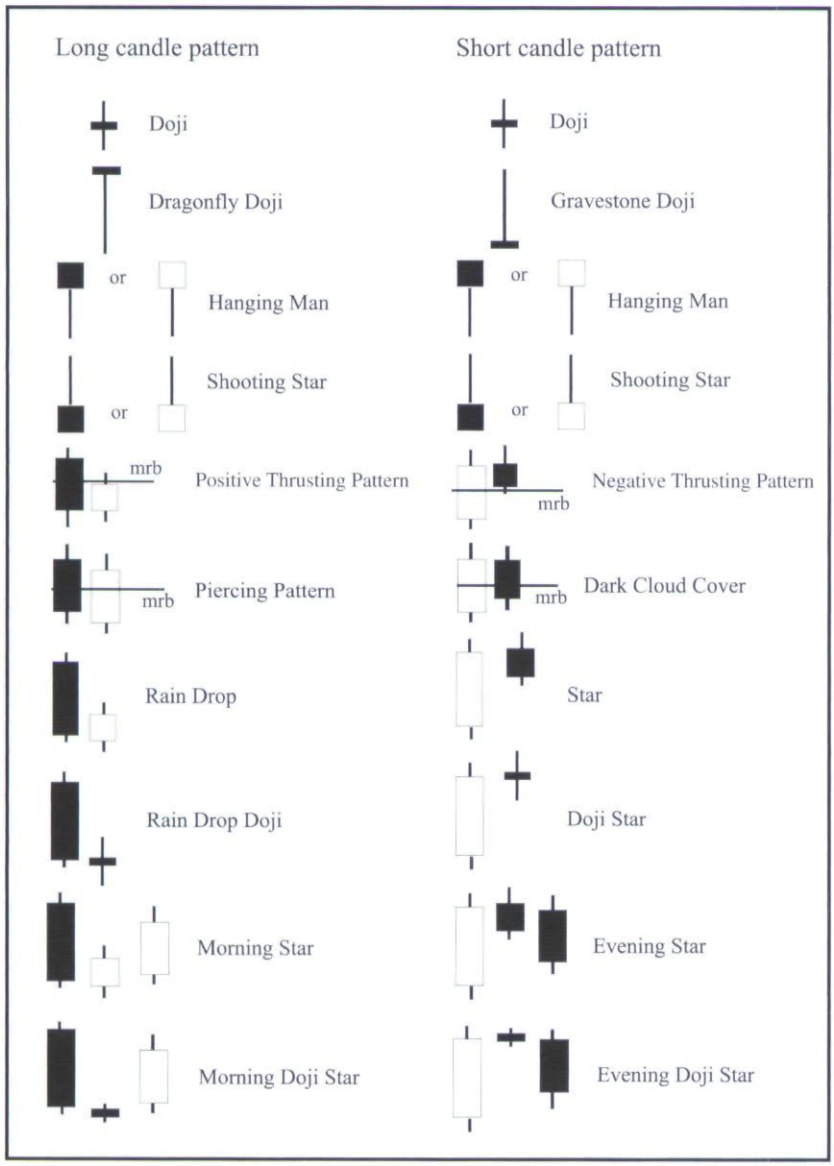


EXHIBIT 3
Visual Representation of Candles Examined



every candlestick pattern we use.

Exhibit 5 illustrates the *Rain Drop* with the four prices used to define the patterns. Patterns are defined from left to right, and numbering follows accordingly. First the candles are defined according to the parameters, color and size, and then by the proportions between the different candles.

The first candle is black with a tall real body; the second candle has a small real body, and the color is of no relevance (examples of both versions (black or white) are seen in the popular literature). The closing price of the first candle (c1) is higher than or equal to the lowest price of the second candle (l2).

We decided to classify the size of the real body (rb) into four different categories. Similar classifications are seen in the popular literature. The *doji* is the smallest real body, followed by the small, medium, and tall real body. The *doji* is the only candle that can also include candles without a real body (o = c). In a first step, the candles are defined according to this scheme, and in a second step they are translated into a numerical value for the FDAX and FGBL.

We translate the candles' verbal definitions (*doji*, *small*, *medium*, *tall*) into values for the FDAX and FGBL by using deciles of the real bodies in the 2000-2001 period (considering only real bodies larger than 0). This means that we quantify the size out of sample, right before the examination period.

The first question is whether the same deciles can be used for white and black real bodies; does the color of the candle have any influence on the absolute return distribution? We test the independence of the real body size from the candle color for the FDAX and FGBL using the Kolmogorov-Smirnov test for two independent samples. We find the candle's color has no influence on its size, so no differentiation is necessary between the deciles used for the black and white real bodies in the DAX and Bund futures. The size of the candle's real body (defined by the opening and closing prices) is indicated by rb in Exhibits 6 and 7.

The *Rain Drop* in Exhibit 3 is a good example. We see a tall first real body and a small second real body, and can translate the verbal descriptions into numbers for the FDAX (*Exhibit 6*). The tall real body should fall within the tallest 30% of real bodies, so it has to be higher than 0.1080%. The small real body ranges from the 10% to the 30% decile, which means that the size is between 0.0382% and 0.0149%.

Another problem in translating candle definitions into values for the particular futures is minimum tick sizes, which make it impossible to use the same decile scheme for the FDAX and the FGBL.[6] Under the FDAX scheme, *dojis* and

EXHIBIT 4
Definition of Candles

Name	C1		C2		C3		P1		P2			P3			
	Co	rb	Co	rb	Co	rb	Price	Pro.	Price	Price	Pro.	Price	Price	Pro.	Price
Dark Cloud Cover	w	t	b	t			o2	>=	c1	c2	>	o1	c2	<	mrb1
Doji	n.d.	d					us1	>	3*rb1	ls1	>	3*rb1			
Doji Star	w	t	n.d.	d			c1	<=	h2	us2	>	3*rb2	ls2	>	3*rb2
Dragonfly Doji	n.d.	d					ls1	>	3*rb1	us1	=	0			
Evening Star	w	t	n.d.	s	b	t	c1	<=	h2	h2	>=	o3	c3	<	c1
Evening Star Doji	w	t	n.d.	d	b	t	c1	<=	h2	h2	>=	o3	c3	<	c1
Gravestone Doji	n.d.	d					us1	>	3*rb1	ls1	=	0			
Hanging Man Black	b	s					ls1	>=	2*rb1	us1	=	0			
Hanging Man White	w	s					ls1	>=	2*rb1	us1	=	0			
Morning Star	b	t	n.d.	s	w	t	c1	>=	l2	l2	<=	o3	c3	>	c1
Morning Star Doji	b	t	n.d.	d	w	t	c1	>=	l2	l2	<=	o3	c3	>	c1
Piercing Pattern (Pos.)	b	t	w	t			o2	<=	c1	c2	<	o1	c2	>	mrb1
Rain Drop	b	t	n.d.	s			c1	>=	l2						
Rain Drop Doji	b	t	n.d.	d			c1	>=	l2	us2	>	3*rb2	ls2	>	3*rb2
Shooting Star Black	b	s					us1	>=	2*rb1	ls1	=	0			
Shooting Star White	w	s					us1	>=	2*rb1	ls1	=	0			
Star	w	t	n.d.	s			c1	<=	h2						
Thrusting Pattern (Neg.)	w	t	b	t			o2	>=	c1	c2	<	c1	c2	>=	mrb1
Thrusting Pattern (Pos.)	b	t	w	t			o2	<=	c1	c2	>	c1	c2	<=	mrb1

Color (Co): white (w); black (b); not defined (n.d.).

Real body (rb): doji (d); small (s); medium (m); tall (t).

Price: open (o), high (h), low (l), close (c), upper shadow (us), lower shadow (ls), middle of real body (mrb).

Proportion (Pro).

small candles would be identical in the FGBL; both would have no real bodies, which does not make sense. We therefore decided to shift the decile proportion by 10% for each category in the FGBL and to accept only a candle with a real body of zero as a *doji*.

Such a shift overcomes the problem, and the results should be comparable. Graphed examples of these definitions were shown to professional traders who confirmed that the naming of our candles seems correct.

After defining all 19 candle patterns for the FDAX and FGBL this way, we use a Java-based computer program to detect candle patterns in the five-minute candles over the years 2002 and 2003.

Indicators

We first test only candlesticks for significance, and then combinations of candles with other indicators. Pruitt and White [1988], Lukac and Brorsen [1990], and Pruitt, Tse, and White [1992] find better results by combining various technical analysis instruments. Many practitioners recommended applying combinations of instruments to achieve better results. We test the 19 candle patterns in combination with moving averages (MA), momentum (M), relative strength index (RSI), and moving average convergence/divergence (MACD) indicators.

Equation (2) describes the moving average as we use it. The MA is calculated over seven closing prices c_t (closing price c at time t) of the 5-minute candles, which equals a 35-minute period. The software carries out a buy if a long candle pattern appears

EXHIBIT 5

Definition of *Rain Drop*

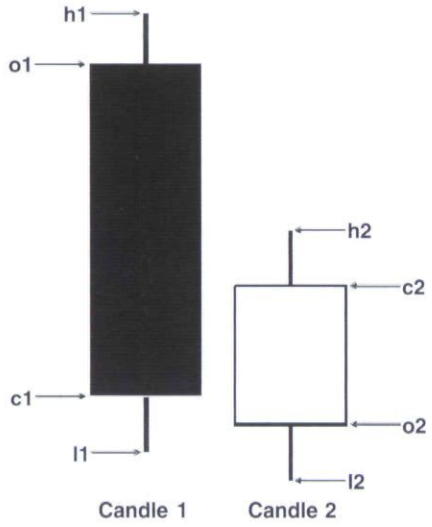


EXHIBIT 6

Size of Real Bodies for DAX Futures

Size	Decile	FDAX
Doji	rb < 10%	0% < rb < 0.0149%
Small rb	10% < rb < 30%	0.0149% < rb < 0.0382%
Medium rb	30% < rb < 70%	0.0382% < rb < 0.1080%
Tall rb	rb > 70%	rb > 0.1080%

EXHIBIT 7

Size of Real Bodies for Bund Futures

Size	Decile	FGBL
Doji	-	0%
Small rb	40% < rb	0% < rb < 0.0096%
Medium rb	40% < rb < 80%	0.0096% < rb < 0.0287%
Tall rb	rb > 80%	rb > 0.0287%

while the MA was falling (defined by the gradient of the MA), and sells if a short candle pattern appears while the MA was rising (short signal):

$$MA_7 = \frac{(c_1 + c_2 + \ldots + c_7)}{7} \tag{2}$$

The momentum indicator in Equation (3) is calculated from the closing prices of the five-minute candles. The program opens a long position if the momentum is higher than 100 when a long candle pattern appears and a short position if the momentum is lower than 100 when a short candle pattern appears:

$$M_7 = \left(\frac{c_7}{c_1}\right)100 \tag{3}$$

Computation of the relative strength index (RSI) is more complex. In Equation (4) the differences of adjacent upward closing prices in a period of seven five-minute candles are calculated. UA in Equation (5) is then the average of the differences of the upward closing prices U. The letter n represents the number of closing prices included in the RSI calculation:

$$U_t = \begin{cases} c_t - c_{t-1} \text{ if } c_t > c_{t-1} \\ 0 \quad \text{otherwise} \end{cases} \tag{4}$$

$$UA_t = \frac{U_t + U_{t-1} + \ldots + U_{t-n+1}}{n} \tag{5}$$

We do the same in Equations (6) and (7) for the downward closing prices, again over a period of seven five-minute candles:

$$D_t = \begin{cases} c_{t-1} - c_t \text{ if } c_{t-1} > c_t \\ 0 \quad \text{otherwise} \end{cases} \tag{6}$$

$$DA_t = \frac{D_t + D_{t-1} + \ldots + D_{t-n+1}}{n} \tag{7}$$

The relative strength (RS) in Equation (8) is computed using Equations (5) and (7). From the RS, the RSI is computed in Equation (9). RSI values below 40 are said to describe oversold conditions; values over 60 are said to represent overbought conditions. The program buys if a long candle pattern appears and the RSI is oversold (and sells under the reverse conditions):

$$RS_t = \frac{UA_t}{DA_t} \tag{8}$$

$$RSI_t = 100 - \left(\frac{100}{1 + RS_t}\right) \tag{9}$$

The last indicator tested with the candles is the MACD. First, exponential moving averages (MAE) are computed. MAE n_t is computed from the previous period's (MAE n_{t-1}) and the weighted current price c_t [Equation (10)]. To construct the first exponential moving

average (MAE n_n), we use the mean of the first n periods:

$$MAE\ n_t = ac_t + (1 - a)MAE\ n_{t-1} \quad (10)$$

$$a = \frac{2}{n + 1} \quad (11)$$

$$MACD = MAE5 - MAE11 \quad (12)$$

An MACD below zero means that the MAE5 is below the MAE11. A long position is opened if a long candle pattern appears during this period. An MACD above zero means that a short position is opened if a short candle pattern appears.

Benchmark

Many studies of technical trading use a buy-and-hold strategy as a benchmark. We decided not to do this because using intraday futures prices implies a short-term time horizon, which makes using a long-term buy-and-hold strategy as a benchmark pointless. We decided to test the candlesticks against a benchmark built from randomized buy signals in the underlying futures, which we believe is more appropriate for an intraday study.

There are about 60,000 five-minute candles for all futures in the two-year study period. A (0, 1) uniformly distributed random number is saved with every candle in the database. An interval with a 10% probability of a hit is chosen, and the software buys (or shorts when calculating the short benchmark) at the next opening price if a candle's stored random number lies within that interval.

The holding period for every trade in the study is 30 minutes after a position is opened. The benchmark is also built with a holding period of 30 minutes, which facilitates comparison of the results.[7]

Only one cycle is necessary to obtain enough random buys to construct the benchmark. If a position is already open, the software is not allowed to open another position if the random number lies within the interval. We follow this procedure to generate approximately 3,600 transactions in the FDAX and FGBL. From these we create the benchmark for the FDAX and FGBL against which to test the candle pattern performance.

An alternative benchmark might assume average earnings of zero. Because of the very short time intervals (less than an hour), it is likely there is no positive risk premium in the expected value at the market level. The effect of any long-term drift in prices has to be minimal compared to the stochastic variation in prices at such a short interval, so the realized drift over the sample period should have no measurable effect on the returns.

The randomized benchmarks indicate that this consideration seems correct—all benchmarks are in the range of zero. When we use the alternative benchmark of zero for the calculations, we get identical results; exceptions are mentioned in the text.

II. EMPIRICAL RESULTS

Exhibit 8 shows the results for all 19 candles. The column labeled *winning deals* provides the number of deals with positive returns after entering a long or a short position with a total of n transactions expressed as a percentage. The trades with an outcome of zero are always treated as losing deals, so the percentages of winning deals of a candle pattern that has been tested both long and short do not add up to 100%.

The next column presents the average earnings per transaction for one contract without taking any transaction costs into account. *Average earnings* is the difference in points between the buy and the evening-up multiplied by the value per point.

The standard deviations of the earnings (on a single trade) are presented in the last column. For some candle patterns (e.g., the *doji*), the literature does not specify whether we have a buy or a sell signal. In these cases we compute results for both long and short positions.

To test for significant results, we apply a t-test with the null-hypothesis H_0: $p_0 \leq 0.5$ to the winning deals. We assume that, if the patterns do not have any predictive power, winner trades should be just as likely as loser trades, and therefore about half of all trades should be winner trades. The benchmark with an average of 49.96% (long) and 47.81% (short) winning deals confirms that this assumption is realistic. The null hypothesis H_0: $\mu_{candle} \leq \mu_{Benchmark}$ is tested for the average earnings. Significant results up to 0.05 are marked with an asterisk.

The results are quite poor: In only two cases are there significant winning deals and significant average earnings. 15 of the patterns produce negative results; 16 reveal more losing deals than winning deals.[8]

Application of the zero benchmark, i.e., using a t-test with the null hypothesis H_0: $\mu_{candle} \leq 0$ produces similar results. Only the average earnings of 29.59 euros on the *Dark Cloud Cover* become significant.

One reason for these poor results may be that candle charts are said to perform better when they are used in

EXHIBIT 8
FDAX Results—Simple Deals

Name	n	Long/Short	Winning Deals	Av. Earnings €	Std. Deviation of Earnings €
Dark Cloud Cover	618	Short	50.00%	29.59	416.60
Doji	2,550	Long	49.57%	-1.58	311.38
Doji	2,550	Short	47.69%	1.58	311.38
Doji Star	565	Short	45.13%	-31.99	353.67
Dragonfly Doji	229	Long	50.22%	-4.09	247.77
Evening Star	295	Short	44.41%	-16.19	360.88
Evening Star Doji	150	Short	45.33%	-41.00	382.75
Gravestone Doji	244	Short	48.36%	-0.41	222.74
Hanging Man Black	224	Long	46.88%	-36.72	291.69
Hanging Man Black	224	Short	49.55%	36.72*	291.69
Hanging Man White	220	Long	51.82%	10.63	296.93
Hanging Man White	220	Short	45.00%	-10.63	296.93
Morning Star	301	Long	47.18%	-22.80	385.39
Morning Star Doji	140	Long	52.86%	-2.23	415.82
Piercing Pattern (Pos.)	628	Long	47.93%	-10.13	402.76
Rain Drop	1,224	Long	53.43%*	17.11*	337.27
Rain Drop Doji	520	Long	55.58%*	11.15	355.00
Shooting Star Black	232	Long	50.00%	-0.11	258.34
Shooting Star Black	232	Short	47.41%	0.11	258.34
Shooting Star White	258	Long	51.55%	6.54	285.56
Shooting Star White	258	Short	46.51%	-6.54	285.56
Star	1,275	Short	49.65%	10.81	347.61
Thrusting Pattern (Neg.)	325	Short	47.69%	-16.81	486.23
Thrusting Pattern (Pos.)	324	Long	49.38%	-42.17	479.39
Benchmark	3,637	Long	49.96%	-2.19	349.04
Benchmark	3,637	Short	47.81%	2.19	349.04

combination with other technical analysis methods. In Exhibit 9 we combine the candle chart technique with a trend indicator signal, the moving average.

Most of the results here are not much better than in Exhibit 8. The earnings of the candle patterns that do not deliver a clear buy or sell signal when used without other indicators (*doji*, *Hanging Man Black* and *White*, *Shooting Star Black* and *White*), perform better now, except for the *doji*. With one exception, these results are not significant, however.

Returns on transactions triggered by the combination of candlestick patterns with other technical indicators are presented in Exhibit 10. Again we have very diverse results. There is no clear winning strategy.

We next apply a binomial test to see whether consideration of other technical indicators improves the results. The results are of course not independent of one another; the results attributable to patterns with more than one candle are really subgroups of the different results attributable to patterns with one candle fewer. For example, the *Morning Star* is equivalent to a *Rain Drop* plus one white candle (*see Exhibit 2*). This means the transactions triggered by a *Morning Star* are all caused by a *Rain Drop* too.

To address this problem, we compare only patterns with an equal number of candles to one another. The results of the one-candle patterns are independent. Even if the patterns themselves are identical (e.g., *doji* [long] and *doji* [short]), the combination with one indicator (such as momentum or RSI) produces two non-overlapping subsamples.

Exhibit 11 shows computation of the binomial test for the one-candle pattern. A numeral 1, for example, in row *doji* and column *Momentum* means that the average return improves on using the indicator over using only the candle patterns. The 0 implies that no improvement has taken place, or there is even a worsening of returns.

The combination of one-candle patterns with the moving average (*trend*) improvements of returns in 9 of 12 cases. The probability (and consequently the p-value of the binomial test) of 9 hits in 12 trials, when the indicators have no influence on the change in performance (e.g., $p = 0.5$), is 0.054. There are no p-values lower than 0.05, which means there are no significant results.

Exhibits 12 and 13 show the results for two- and three-pattern candles. There are no significant improvements in returns.

We noted earlier that many candlestick patterns are based on only a few patterns; a new pattern is constructed by mirroring or inverting, or by adding another candle.

EXHIBIT 9

FDAX Results—Candles Combined with Trend Indicator (MA)

Name	n	Long/Short	Winning Deals	Av. Earnings €
Dark Cloud Cover	364	Short	48.63%	5.25
Doji	1,203	Long	51.12%	-7.47
Doji	1,286	Short	48.60%	-2.95
Doji Star	388	Short	45.10%	-22.42
Dragonfly Doji	106	Long	54.72%	-5.90
Evening Star	129	Short	48.84%	-9.40
Evening Star Doji	80	Short	46.25%	-21.25
Gravestone Doji	116	Short	54.31%	9.59
Hanging Man Black	117	Long	45.30%	-31.73
Hanging Man Black	99	Short	49.49%	51.64*
Hanging Man White	106	Long	56.60%	30.42
Hanging Man White	110	Short	49.09%	13.64
Morning Star	165	Long	46.06%	-39.47
Morning Star Doji	72	Long	51.39%	1.04
Piercing Pattern (Pos.)	389	Long	48.59%	-13.95
Rain Drop	825	Long	51.76%	7.00
Rain Drop Doji	355	Long	56.06%*	10.77
Shooting Star Black	115	Long	53.04%	4.02
Shooting Star Black	111	Short	50.45%	5.29
Shooting Star White	112	Long	56.25%	16.07
Shooting Star White	138	Short	46.38%	-2.99
Star	837	Short	49.22%	-1.18
Thrusting Pattern (Neg.)	207	Short	52.66%	-1.63
Thrusting Pattern (Pos.)	225	Long	48.44%	-72.61
Benchmark	1,770	Long	51.24%	-5.55
Benchmark	1,801	Short	48.92%	-1.15

Exhibit 14 shows results of testing whether forecasting ability is improved by extending the patterns, or whether there are other systematic connections between patterns built according to this scheme.

According to the popular literature on candlesticks, results should improve when a candle is added or when a *doji* is part of the pattern. Exhibit 14 divides the patterns into subgroups where the first pattern should be the weakest and the last the strongest. There are just a few improvements in the results. We are not able to find any systematic enhancement of forecasting ability.

These results indicate it is not possible to earn significant abnormal returns using the candlestick technique. Contrary to opinions expressed in the popular literature, combining the candle technique with different technical indicators does not produce a significant performance improvement.

The poor results we get in following the recommendations of the popular literature and the possibility that we have misinterpreted it—we may be inconsistent in interpreting a pattern as a buy or sell signal—prompted an ex post improvement test. We determined retrospectively whether a pattern is a buy or sell signal by the profit of the pattern without an additional indicator.

As one might expect, with this hindsight bias, all results with no indicator in Exhibit 15 are positive, if often fairly poor. Sticking to the long-short strategy as determined without an indicator and then combining it with different indicators led again to very poor, sometimes even negative, results. This supports a conclusion that our poor results are not due to possible misinterpretations but to candle patterns' lack of forecasting power.

To date no earnings calculations have considered transaction costs, which are quite high in the case of futures; a charge of about 5 euros per transaction seems realistic. So-called slippage costs must also be considered. In other words, the price used in the calculations is not likely to be the price paid in reality—the trader may not be able to conclude a transaction quickly enough to achieve this price. Nor is it clear whether the prices used for the calculations are ask or bid prices. According to LeBeau and Lucas [1992, p. 172], that can be a problem.

For the FDAX, slippage costs of one point—which corresponds to 25€—seem to be realistic. Slippage and commissions must be calculated twice (on the sell and the buy side) for each transaction, which leads to overall

EXHIBIT 10

FDAX Results—Candles Combined with Momentum, RSI, and MACD

		Momentum			RSI			MACD		
Name	L/S	n	Win. %	Earn.	n	Win. %	Earn.	n	Win. %	Earn.
Dark Cloud Cover	S	236	53.81	73.36*	226	48.67	0.33	377	48.01	12.57
Doji	L	1,268	49.13	6.40	857	50.29	-14.12	1,234	50.81	-7.15
Doji	S	1,221	47.67	7.91	911	49.18	2.63	1,316	48.56	-3.65
Doji Star	S	152	48.03	-25.82	318	44.97	-25.04	365	46.30	-27.53
Dragonfly Doji	L	114	51.75	13.16	80	55.00	-7.34	107	54.21	1.64
Evening Star	S	154	47.40	-2.68	84	47.62	-30.65	147	47.62	-6.38
Evening Star Doji	S	69	42.03	-75.18	53	50.94	-1.65	93	47.31	-13.98
Gravestone Doji	S	113	47.79	-3.87	79	50.63	-0.79	126	46.03	-1.88
Hanging Man Black	L	105	56.19	-3.21	82	47.56	-21.80	121	47.93	-31.71
Hanging Man Black	S	110	55.45	59.32*	59	52.54	59.11	103	53.40	42.60
Hanging Man White	L	112	45.54	-24.00	76	61.84*	50.00*	103	60.19*	41.14*
Hanging Man White	S	105	40.00	-46.79	70	55.71	63.21	117	51.28	16.24
Morning Star	L	138	49.28	-0.45	101	44.55	-75.12	182	48.35	-27.34
Morning Star Doji	L	66	54.55	21.21	50	54.00	25.00	87	52.87	20.26
Piercing Pattern (Pos.)	L	228	45.61	-7.51	278	49.64	-14.39	407	49.14	-8.11
Rain Drop	L	340	56.76*	23.31	591	51.95	13.20	725	52.55	15.55
Rain Drop Doji	L	134	51.49	-4.01	273	54.95	7.23	336	56.25*	14.40
Shooting Star Black	L	115	46.09	-17.50	88	56.82	13.35	117	55.56	11.65
Shooting Star Black	S	112	39.29	-29.80	82	47.56	-37.65	115	54.78	12.07
Shooting Star White	L	143	48.25	19.14	70	57.14	0.00	106	58.49*	27.12
Shooting Star White	S	109	44.04	8.14	106	46.23	-3.54	152	50.66	7.81
Star	S	383	50.91	35.12	629	49.76	3.74	742	49.60	10.88
Thrusting Pattern (Neg.)	S	101	36.63	-61.88	159	52.83	-6.21	208	51.44	-24.88
Thrusting Pattern (Pos.)	L	77	51.95	33.12	177	45.20	-91.03	225	47.11	-79.83
Benchmark	L	1,791	48.02	-1.88	1,242	50.72	-14.96	1,834	51.04	-8.80
Benchmark	S	1,776	46.17	5.41	1,266	49.45	0.57	1,802	48.61	-4.38

EXHIBIT 11

Improvement of Performance of Candle Patterns with One Candle

Name	Trend	Momentum	RSI	MACD
Doji	0	1	0	0
Doji	0	1	1	0
Dragonfly Doji	0	1	0	1
Gravestone Doji	1	0	0	0
Hanging Man Black	1	1	1	1
Hanging Man Black	1	1	1	1
Hanging Man White	1	0	1	1
Hanging Man White	1	0	1	1
Shooting Star Black	1	0	1	1
Shooting Star Black	1	0	0	1
Shooting Star White	1	1	0	1
Shooting Star White	1	1	1	1
Hits	9	7	7	9
N	12	12	12	12
p-value	0.054	0.193	0.193	0.054

EXHIBIT 12

Improvement of Performance of Candle Patterns with Two Candles

	Trend	Momentum	RSI	MACD
Hits	2	6	2	4
n	8	8	8	8
p-value	0.891	0.109	0.891	0.273

EXHIBIT 13

Improvement of Performance of Candle Patterns with Three Candles

	Trend	Momentum	RSI	MACD
Hits	3	3	2	3
n	4	4	4	4
p-value	0.250	0.250	0.375	0.250

EXHIBIT 14

Improvement of Performance of Candle Patterns (euros)

Name	No Indicator	Trend	Momentum	RSI	MACD
Rain Drop	17.11	7.00	23.31	13.20	15.55
Rain Drop Doji	11.15	10.77	-4.01	7.23	14.40
Morning Star	-22.80	-39.47	-0.45	-75.12	-27.34
Morning Star Doji	-2.23	1.04	21.21	25.00	20.26
Star	10.81	-1.18	35.12	3.74	10.88
Doji Star	-31.99	-22.42	-25.82	-25.04	-27.53
Evening Star	-16.19	-9.40	-2.68	-30.65	-6.38
Evening Star Doji	-41.00	-21.25	-75.18	-1.65	-13.98
Doji	-1.58	-7.47	6.40	-14.12	-7.15
Dragonfly Doji	-4.09	-5.90	13.16	-7.34	1.64
Hanging Man Black	-36.72	-31.73	-3.21	-21.80	-31.71
Hanging Man White	10.63	30.42	-24.00	50.00	41.14
Doji	1.58	-2.95	7.91	2.63	-3.65
Gravestone Doji	-0.41	9.59	-3.87	-0.79	-1.88
Shooting Star Black	0.11	5.29	-29.80	-37.65	12.07
Shooting Star White	-6.54	-2.99	8.14	-3.54	7.81
Thrusting Pattern (Pos.)	-42.17	-72.61	33.12	-91.03	-79.83
Piercing Pattern (Pos.)	-10.13	-13.95	-7.51	-14.39	-8.11
Thrusting Pattern (Neg.)	-16.81	-1.63	-61.88	-6.21	-24.88
Dark Cloud Cover	29.59	5.25	73.36	0.33	12.57

transaction costs of about 60€. Thus, if the costs are taken into account, we would have only a few transactions with positive earnings.

We use all the same procedures to test candlestick patterns for the Bund futures of German government bonds (FGBL). Calculated average earnings and winning trade results are even poorer than for the DAX futures. Full results are available from the authors on request.

III. SUMMARY AND CONCLUDING REMARKS

We believe we are the first to test the predictive power of candlestick patterns on an intraday basis. We use Dax futures data starting January 1, 2002, and ending December 31, 2003. To make the tests as objective as possible, we develop a mathematical definition scheme enabling a fully computerized identification of the candles in the data.

Even without transaction costs, the results are surprisingly poor. In most cases the results are not significantly better than a benchmark with randomized transactions. When we use combinations of candle patterns together with other technical indicators to try to improve the results, we get higher returns but still statistically not significant. Applying the approach to Bund futures data produces worse results than for DAX futures.

On the whole, these results are rather surprising, because many practitioners seem to be trying to forecast security prices by means of candlesticks. Of course,we tested only the most frequently used candlestick patterns. Results might be different with different underlyings and different time frames. Future research, in any case, may profit from our mathematical definition schemes of candlestick patterns.

ENDNOTES

[1]Forecasting power could also be a self-fulfilling prophecy due to its use by many market participants. See Reilly and Brown [2000, p. 873].

[2]Timing may mean the right month, week, day, hour, or even second.

[3]See Goodhart and O'Hara [1997] for an in-depth discussion of the advantages and problems of using intraday data.

[4]Data are courtesy of the Eurex. The original data also included over-the-counter (OTC) trades only cleared by the Eurex that did not arise from a normal trading process (we eliminate them).The DAX represents the 30 largest stocks with the highest volume traded on the German stock exchange. The

EXHIBIT 15

Results after Optimization (euros)

Name	No Indicator	Trend	Momentum	RSI	MACD
Rain Drop	17.11	7.00	23.31	13.20	15.55
Rain Drop Doji	11.15	10.77	-4.01	7.23	14.40
Morning Star	22.80	0.47	41.12	3.73	15.86
Morning Star Doji	2.23	5.70	21.21	33.75	39.15
Star	10.81	-1.18	35.12	3.74	10.88
Doji Star	31.99	54.72	36.73	30.53	40.13
Evening Star	16.19	16.93	35.69	1.62	25.93
Evening Star Doji	41.00	61.75	9.84	61.27	85.09
Doji	1.58	-2.95	7.91	2.63	-3.65
Dragonfly Doji	4.09	6.80	31.42	-7.34	9.12
Hanging Man Black	36.72	51.64	59.32	59.11	42.60
Hanging Man White	10.63	30.42	-24.00	50.00	41.14
Doji	1.58	-2.95	7.91	2.63	-3.65
Gravestone Doji	0.41	9.59	-3.87	8.97	-1.17
Shooting Star Black	0.11	5.29	-17.50	13.35	12.07
Shooting Star White	6.54	16.07	19.14	0.00	27.12
Thrusting Pattern (Pos.)	42.17	-28.61	67.91	-73.96	-43.43
Piercing Pattern (Pos.)	10.13	6.33	11.69	-14.39	13.86
Thrusting Pattern (Neg.)	16.81	55.92	0.06	8.42	2.46
Dark Cloud Cover	29.59	5.25	73.36	0.33	12.57

underlying for the FGBL is an ideal type of federal German bond with an interest rate of 6% and a time to maturity of between 8.5 and 10.5 years.

[5]Trading periods on the Eurex at the time of the study ran from 8:50 AM to 8:00 PM (FDAX) and from 8:00 AM to 7:00 PM (FGBL).

[6]For the FDAX the minimum tick size is 0.5 points; for the FGBL it is 0.01 points.

[7]We assume that any observable effect would become apparent after 30 minutes. Results testing different holding periods are comparable or poorer than for a holding period of 30 minutes.

[8]Patterns that are used as long and short indicators produce contrary results. That is, in this case one transaction must lead to negative returns and more losing deals.

REFERENCES

Ball, R., S.P. Kothari, and C.E. Wasley. "Can We Implement Research on Stock Trading Rules?" *The Journal of Portfolio Management*, 21 (1995), pp. 54-63.

Balsara, N., K. Carlson, and N.V. Rao. "Unsystematic Futures Profits With Technical Trading Rules: A Case for Flexibility." *Journal of Financial and Strategic Decisions*, 9 (1996), pp. 57-66.

Bankhofer, U., and A. Hilbert. "*Eine empirische Studie zur Anwendung mathematischer Methoden in der Versicherungspraxis.*" *Zeitschrift für das gesamte Versicherungswesen* (1999), pp. 611-644.

Brock, W., J. Lakonishok, and B. LeBaron. "Simple Technical Trading Rules and the Stochastic Properties of Stock Returns." *Journal of Finance*, 47 (1992), pp. 1731-1764.

Curcio, R., C. Goodhart, D. Guillaume, and R. Payne. "Do Technical Trading Rules Generate Profits? Conclusions from the Intraday Foreign Exchange Market." *International Journal of Finance & Economics*, 2 (1997), pp. 267-280.

Detry, P.J., and P. Grégoire. "Other Evidence of the Predictive Power of Technical Analysis: The Moving Averages Rules on European Indexes." Working paper, CeReFiM Université de Namur and Université Catholique de Louvain, Belgium, 2000. Available at: www.technicalanalysis.org.uk/moving_averages/DeGr01.pdf.

Dorfleitner, G., and C. Klein. "*Kursprognose mit Hilfe der technischen Analyse—eine empirische Untersuchung.*" *Journal of Financial Markets and Portfolio Management*, 16 (2002), pp. 497-521.

Fama, E.F. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance*, 25 (1970), pp. 383-418.

——. "Efficient Capital Markets: II." *Journal of Finance*, 46 (1991), pp. 1575-1617.

Fernandez-Rodriguez, F., C. Gonzalez-Martel, and S. Sosvilla-Rivero. "On the Profitability of Technical Trading Rules Based on Artificial Neural Networks: Evidence from the Madrid

Stock Market." *Economics Letters,* 69 (2000), pp. 89-94. Available at: www.fedea.es/hojas/publicaciones.html.

Goldbaum, D. "A Nonparametric Examination of Market Information: Application to Technical Trading Rules." *Journal of Empirical Finance*, 6 (1999), pp. 59-85.

Goodhart, C.A.E., and M. O'Hara. "High Frequency Data in Financial Markets: Issues and Applications." *Journal of Empirical Finance*, 4 (1997) , pp. 73-114.

Hofmann, H.W. "*Empirische Untersuchung verschiedener Anlagestrategien der Technischen Analyse anhand von 100 deutschen Standardaktien über einen Zeitraum von fünf Jahren (1967-1972) und Vergleich mit einer durch Simulation ermittelten Zufallsauswahl bzw. einer Kauf- und Halte Strategie.*" Ph.D. diss., Ludwigs-Maximilians-Universität München, 1973.

Jegadeesh, N. "Foundations of Technical Analysis—Discussion." *Journal of Finance*, 55 (2000), pp. 1765-1770.

Landi, D.A. "The Japanese Candle Stick Chart." *TraderMade International* Ltd., Technical Analysis Workshop, Frankfurt, 1997.

LeBeau, C., and D. Lucas. *Börsenanalyse mit dem Computer.* Darmstadt: Hoppenstedt & Co., 1992.

Ligh, W., N. Paz, and R. Purvis. "Market Timing: A Test of a Charting Heuristic." *Economics Letters*, 77 (2002), pp. 55-63.

Lo, A.W., H. Mamaysky, and J. Wang. "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation." *Journal of Finance*, 55 (2000), pp. 1705-1765.

Lukac, P.L., and B. W. Brorsen. "A Comprehensive Test of the Futures Market Disequilibrium." *The Financial Review*, 25 (1990), pp. 593-622.

Malkiel, B. *A Random Walk Down Wall Street: Including a Life-Cycle Guide to Personal Investing.* New York: Norton, 1996.

Mills, T.C. "Technical Analysis and the London Stock Exchange: Testing Trading Rules Using the FT30." *International Journal of Finance & Economics*, 2 (1997), pp. 319-331.

Montassèr, R.D. "The Accuracy of Candlestick Analysis Using the German Stock Market Index (DAX)." *IFTA Journal*, 2000, pp. 27-38.

Nison, S. "Japanese Candlestick Charting Techniques." New York Institute of Finance, 1991.

Okamoto, H. "Practical Utilization of Japanese Technical Charts." Presented at the TA Forum, Paris, March 22, 2003.

Osler, C.L. "Currency Orders and Exchange Rate Dynamics: Explaining the Success of Technical Analysis." Federal Reserve Bank of New York, 2000.

Pruitt, S.W., K.S. Tse, and R.E. White. "The CRISMA Trading System: The Next Five Years." *The Journal of Portfolio Management*, 18 (1992), pp. 22-25.

Pruitt, S.W., and R.E. White. "The CRISMA Trading System: Who Says Technical Analysis Can't Beat the Market?" *The Journal of Portfolio Management*, 19 (1988), pp. 55-58.

Raj, M. "Transaction Data Test of Efficiency: An Investigation in the Singapore Futures Markets." *The Journal of Futures Markets*, 20 (2000), pp. 687-704.

Reilly, F.K., and K.C. Brown. *Investment Analysis and Portfolio Management.* Mason, OH: South-Western, Thomson Learning, 2000.

Schwager, J.D. *Schwager on Futures: Technical Analysis.* New York: John Wiley & Sons, 1996.

Shen, Pu. "Market Timing Strategies that Worked." *The Journal of Portfolio Management*, 29 (2003), pp. 57.

Sweeney, R.J. "Beating the Foreign Exchange Market." *Journal of Finance*, 41 (1986), pp. 163-182.

——. "Evidence on Short-Term Trading Strategies." *The Journal of Portfolio Management*, 17 (1990), pp. 20-26.

Tyler, M., and H. Allen. "The Use of Technical Analysis in the Foreign Exchange Market." *Journal of International Money and Finance*, 11 (1992), pp. 304-314.

Wagner, G.S., and B.L. Matheny. *Trading Applications of Japanese Candlestick Charting.* New York: John Wiley & Sons, 1994.

Wong, W.-K., M. Manzur, and B.-K. Chew. "How Rewarding is Technical Analysis? Evidence from Singapore Stock Market." Working paper, National University of Singapore, 2002. Available at: www.fas.edu.sg/ecs/pub/wp/wp0216.pdf.

To order reprints of this article, please contact Ajani Malik at amalik@iijournals.com or 212-224-3205.