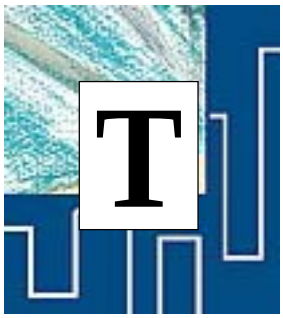
Identifying Market Trends

The trend of the market is key to most every technical approach available. But the market doesn't always trend. Here's how to use a statistical tool for determining if the market is in a trend.

by Jack Karczewski



Traders spend untold amounts of intellectual capital analyzing a market for clues about trend. We try to determine if there is a trend present, when it began, if it has reversed or broken down. The methods employed to interpret this information range from simple chart analysis to sophisticated computer algorithms. Even simple moving averages can become quite complex when the permutations and combinations of moving averages, their various lengths, crossovers and forms such as simple, weighted and exponential, are considered.

I use one analytical method in particular to identify trends. This technique gives not only the direction, but the slope, the magnitude of the error of the prediction and the reliability of the forecast of the trend. The statistical method is referred to as *linear regression*. This tool has become more readily available to traders with the current standard software.

There is a mystique that surrounds this simple tool, and many traders shun it because they don't understand its concept. There are many things that linear regression can and cannot do in explaining the concept of trend. Linear regression will not solve all of the problems of analyzing trends, but understanding the nature of the indicator and the ancillary statistics associated with its computation will provide a better understanding of market dynamics.

A LOOK AT MOVING AVERAGES

Most traders use simple moving averages, so let's take a look at the strengths and weaknesses of moving averages as a trend indicator and compare them with the linear regression method.

Moving averages are essentially a smoothing technique. By averaging past data, the moving average filters the noise associated with any time series datastream. This is true whether you are observing daily, hourly, monthly or any discrete data.

Moving averages present two fundamental problems for analysts. First, the average is just that — an average price that should be plotted or centered midway through the data interval. If a 20-day simple moving average is being used, the result is the average price of the data centered on the data 10 days back. Most technical analysis programs will shift the result to be current with the day or week. Conceptually, trends persist, and the moving average is the proxy for the current trend. As long as your current data is above the moving average and the moving average is rising, the trend is up. If the current data is below a falling moving average, the trend is down. Numerous variations of this technique exist, and many traders have employed them at one time or another.

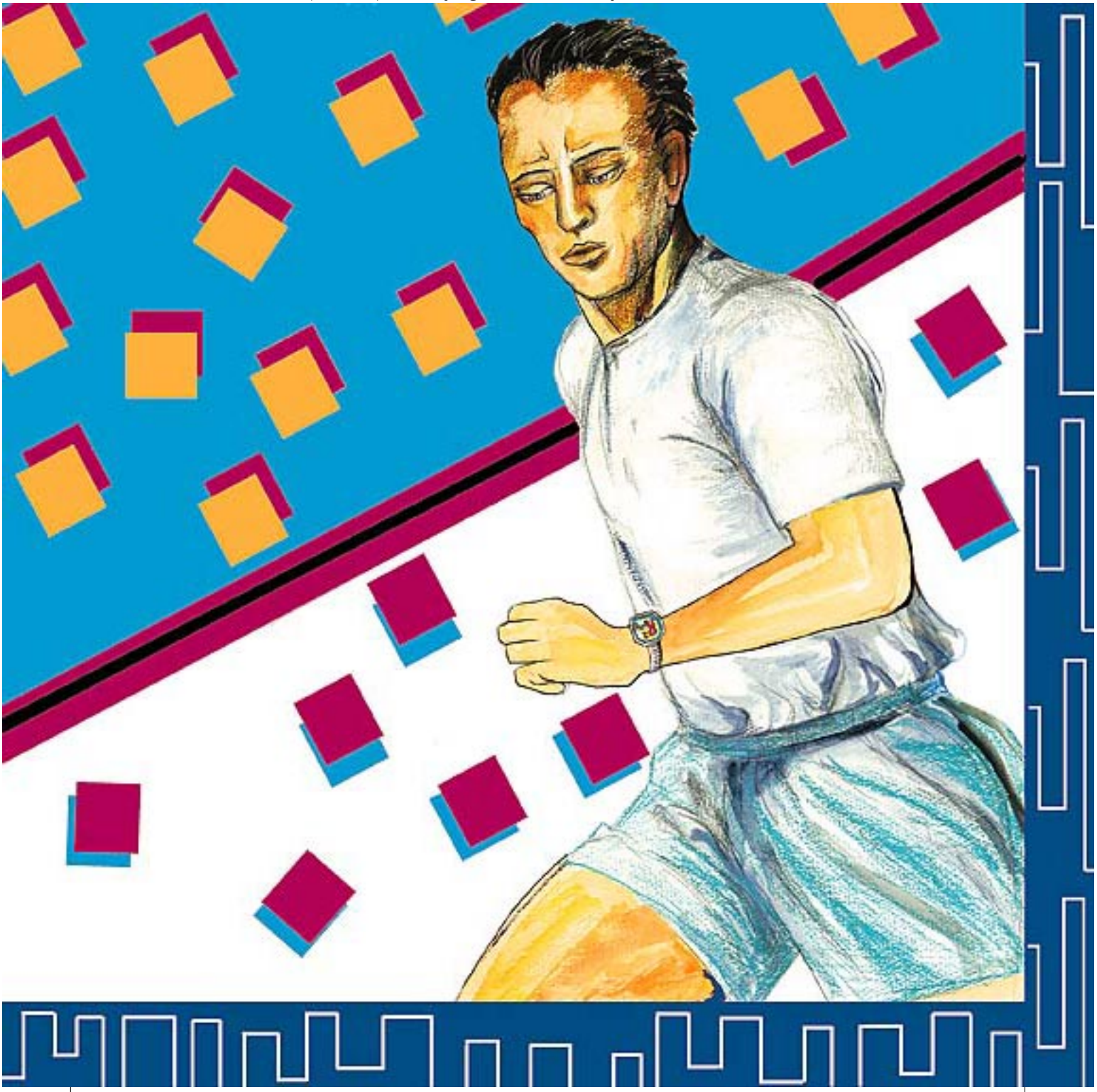
The second problem with moving averages is the arbitrary selection of the lookback period. The selection of the interval depends on the trader's requirements and the attributes of the market in question. Short intervals are responsive to market changes and retain much of the noise that the average seeks to eliminate, while long intervals eliminate much of the noise but are not particularly responsive to the market movement. Traders have devised many systems to correct these problems, but that discussion is beyond our scope here.

There are some secondary problems with moving averages that can be just as important as the primary ones. First, there is no method to measure the slope of the trend directly without resorting to some data manipulation. Second, there is no direct method to determine if the data fits the market studied (otherwise known as *goodness of fit*), and finally, there is no direct way to measure whether the prediction falls within a measurable acceptable error. These problems can be solved to some degree with additional indicators, but there is a more elegant and superior solution, and that is linear regression.

ABOUT LINEAR REGRESSION

Linear regression is a statistical technique that fits a straight line to a datastream. The datastream is an independent variable versus a dependent variable. In this case, the independent variable is time and the dependent variable is price. This data is generally viewed in a scatter diagram, but here, we use traditional price charts with just the closing price plotted on the *y*-axis, and the *x*-axis being time. A straight line is fitted so the distance is minimized between the predicted line and the data, a technique referred to as *least squares*. The name comes from the use of squaring the differences between the line and the data points. For our purposes, I will simply refer to the technique as linear regression.

Some very valuable statistics are a byproduct of linear regression analysis: *r*-squared or the coefficient of determination, the standard error of the estimate, the slope of the line and finally a prediction. While I will address each of these, I will focus on the *r*-squared as our basic tool for trend determination. The use of this readily available statistic will aid trend analysis and help determine when trend trading



CHRISTINE MORRISON

techniques should be used.

R-squared is the key statistic generated when a linear regression line is fitted. This statistic informs us how well the line explains the data; the parameters for this statistic are zero and 1. A reading of zero indicates that the dependent variable has no linear relationship to the independent variable, while an *r*-squared reading of 1 indicates that the line explained the data exactly.

High readings indicate good trends and low readings denote a nontrending or ranging market. Observing how *r*-squared behaves will give an important clue about refining our trading patterns — whether we should be using trend-following methods or range trading methods at any given time.

USING *R*-SQUARED

R-squared is used to measure the relationship of variables in an equation. Econometricians use *r*-squared to estimate how well equations or models fit the data. In multiple linear regression, each variable makes a contribution to the equation and the result of each new variable can be measured.

Observing *r*-squared as it moves through time provides us with useful trading information as the indicator can range

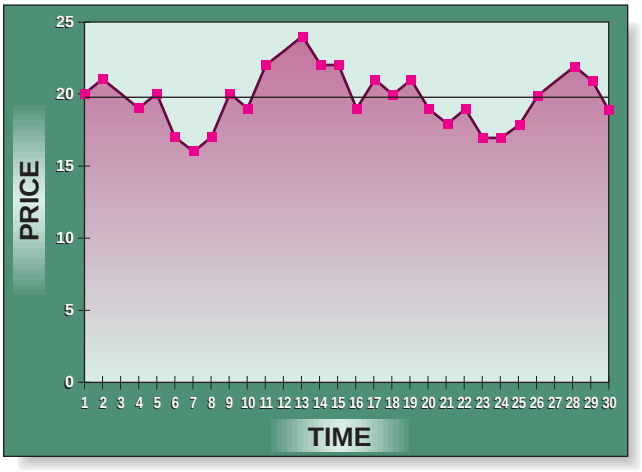


FIGURE 1: A NONTRENDING MARKET. Here's a simple close-only chart with a linear regression line plotted though the data. The price action is sideways, moving over the period of 30 observations. Relative to time, the market is not trending.

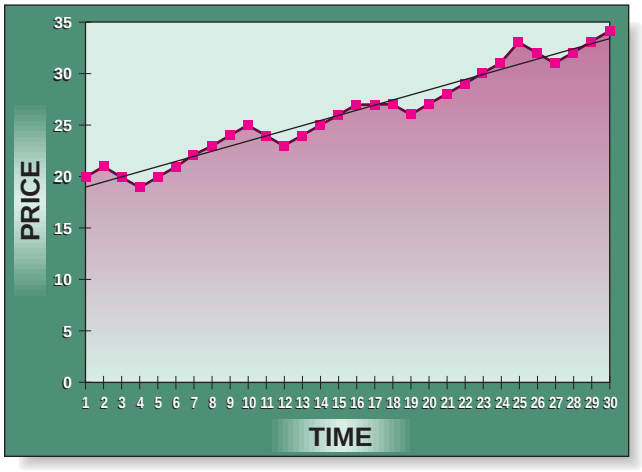


FIGURE 2: A TRENDING MARKET. Here's a chart similar to Figure 1, except the price is rising at a fairly constant pace. The r-squared indicates that the data is moving almost as a straight line over time.

from zero to 1, so indicating the degree of trend. Figure 1 shows a simple close-only chart with a linear regression line plotted though the data. The price action is sideways, moving over the period of 30 observations. The *r*-squared for the data is 0.00014, which is close to zero. This value indicates that relative to time, the market is not trending.

Figure 2 shows a similar price chart, except the price is rising at a fairly constant pace. The data has an *r*-squared of 0.95, indicating that the data is moving almost as a straight line over time. Figure 3 shows both types of periods. A trend is developing, and you can see the *r*-squared climb in value as the market trends. As the market peaks and moves into a consolidation, the *r*-squared falls in value, setting the stage for a new trend. In both situations, the *r*-squared provides a quantifiable method to measure the relationship between price and time. This relationship enables the trader to make important decisions about the trend's longevity, a departure from classical analysis of trends.

Time is the independent variable in this type of analysis; no inference should be drawn that time is the causation for the movement in the dependent variable — in this case, price. Fundamental and exogenous factors contribute to the movement of the independent variable, but this technique is not concerned with them. Like any indicator or technique, there is a certain method to applying this statistical tool.

Figure 4, the Standard & Poor's 500 index, shows a visual

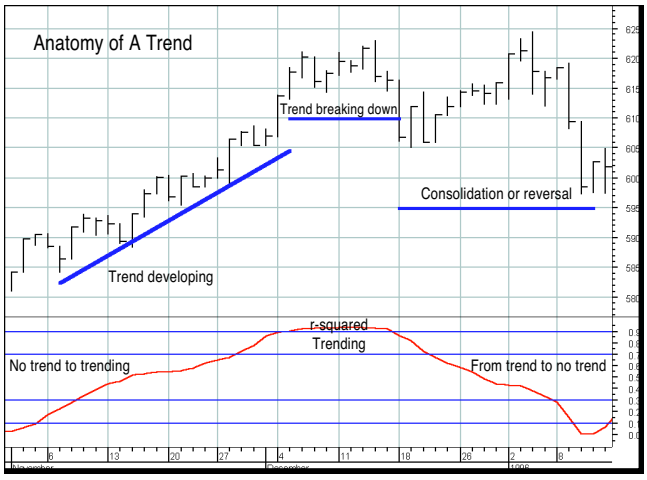


FIGURE 3: THE ANATOMY OF A TREND. This chart shows both types of periods. A trend is developing, and you can see the r-squared climb as the market trends. As the market peaks and moves into a consolidation, the r-squared falls in value, setting the stage for a new trend. In both situations, the r-squared provides a quantifiable method to measure the relationship between price and time. This relationship enables the trader to make important decisions about the trend's longevity.

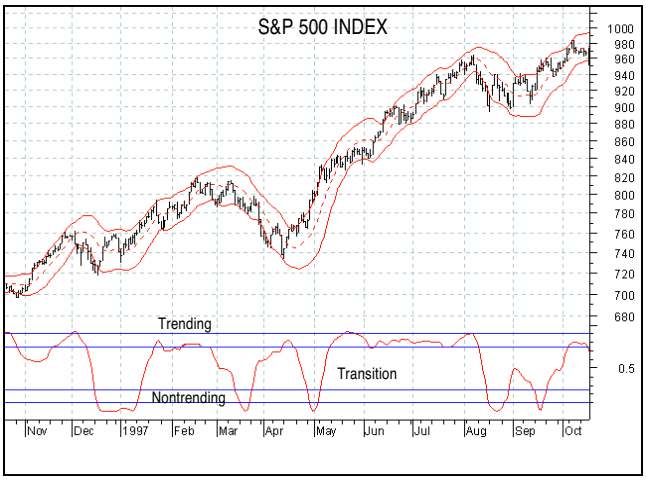


FIGURE 4: S&P 500. Here's a visual representation of r-squared plotted through time. The interval used for this indicator is 30 trading days, which was selected because it is the smallest number of observations that can be used without correcting for small samples. Longer or shorter lengths can be used, but smaller samples will need to be corrected for statistical significance, while longer lengths suffer from the same problems as moving averages; they are not responsive to short-term influences in the market. The intervals should be adjusted to the traders' time horizon.

R-SQUARED READINGS, S&P 500

R-squared	Cum. occ.	Int. occ.	Frequency	Condition	Cum freq.
0.1	488	488	22%	No	
0.2	679	191	9%	Trend	
0.3	862	183	8%		39%
0.4	1026	164	7%	Transition	
0.5	1204	178	8%		
0.6	1391	187	8%		
0.7	1618	227	10%	Trending	38%
0.8	1892	274	12%		
0.9	2174	282	13%		
1.0	2234	60	3%		

FIGURE 5: Some inferences about trading conditions may be made by studying r-squared readings. Figure 5 shows a distribution of r-squareds; these readings were taken from 30-day linear regression analysis of the S&P 500 index from August 1988 to summer 1997. Readings of greater than 0.70 describe a good fit, while readings of less than 0.30 describe a poor one. At first glance, the numbers would seem to indicate that the market does not trend as often as we would like. The statistics suggest that the S&P 500 is in a trending mode only about 35% of the time.

representation of *r*-squared plotted through time. The interval used for this indicator is 30 trading days, which is about 42 calendar days. The period of 30 days was selected because it is the smallest number of observations that can be used without correcting for small samples. Longer or shorter lengths can be used, but smaller samples will need to be corrected for statistical significance, while longer lengths suffer from the same problems as moving averages; they are not responsive to short-term influences in the market. The intervals should be adjusted to the traders' time horizon.

It should be noted that the indicator is unstable; it is very dynamic, does not remain in the same location for extended periods and frequently oscillates between zero and 1. Of further note is the manner in which it traverses the entire range from low to high and has completed these in a cyclical fashion, with the exception of spring and summer 1997.

Extreme readings are associated with changes in the markets. Once such readings are observed, they tend to reverse, though not at predictable periods. Still, it is useful to know when a market is trending. In fact, *r*-squared could be the most important variable associated with linear regression analysis; without an *r*-squared reading, the validity of the line remains undetermined.

Since linear regression is a measure of past performance, is there a method that can be used to help with trading? There is: the average length of the run can be used to play the odds in using linear regression.

THE NATURE OF *R*-SQUARED

Some inferences about trading conditions may be made by studying the nature of trends, identifiable with *r*-squared. Figure 5 is a distribution of *r*-squared readings. These readings were taken from 30-day linear regression analysis of the S&P 500 index from August 1988 to summer 1997. Readings of greater than 0.70 describe a good fit, while readings of less than 0.30 describe a poor one. While some might argue that these are arbitrary levels, they are nevertheless a point of departure. And unlike some other indicators with points that trigger action, these numbers are derived from statistical theory.

At first glance, the numbers would seem to indicate that the market does not trend as often as we would like. The statistics suggest that the S&P 500 is in a trending mode only about

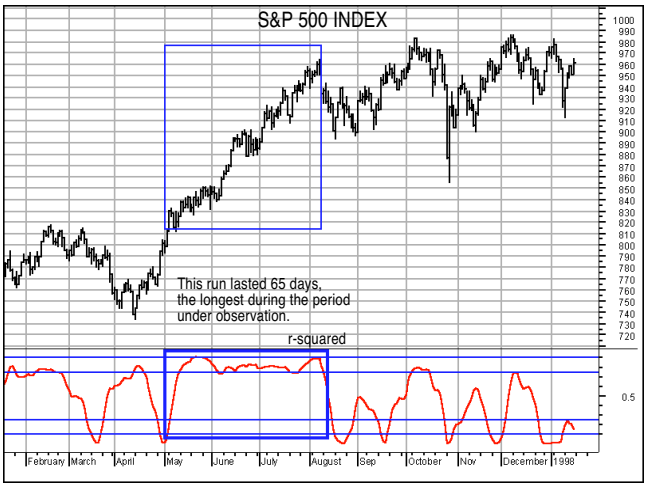


FIGURE 6: CONSECUTIVE DAYS IN A TREND. The number of days the r-squared spends at certain levels are referred to as runs. Runs are defined as consecutive days in a trend. If a market achieves an r-squared reading of greater than 0.70 and the following day has another day of greater than 0.70, that would be a run of two days. The longest run in the S&P 500 of a trending market has been 65 days; that run ended in August 1997.

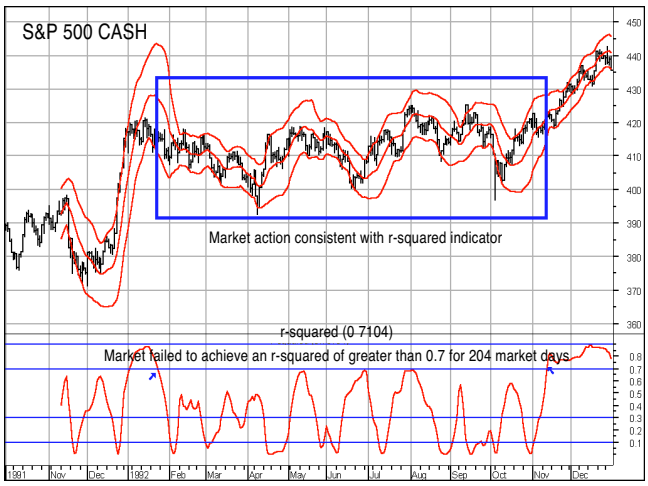


FIGURE 7: CONSECUTIVE DAYS NOT IN A TREND. Contrast that information with the longest period that a market was not in a trend. That was 204 days, which ended in November 1992. The next longest period of r-squared of less than 0.70 ended in July 1996, and that was 102 days. This information may be why many moving average systems fail.

35% of the time. If the trend is your friend, it is a very fickle one, because it doesn't stay with you very long.

While this is useful information, it is only part of the story. By observing the nature of the movement of *r*-squared, in Figure 5, we can see that it moves from nontrending to trending conditions.

RUNS

Figure 5 gives us information about the distribution of trends and nontrends, but what about the number of days the *r*-squared spends at certain levels? These are referred to as *runs*. Runs are defined as consecutive days in a trend. If a market achieves an *r*-squared reading of greater than 0.70 and the following day has another day of greater than 0.70, that would be a run of two days. The longest run in the S&P 500 of a trending market has been 65 days (Figure 6); that run ended in August 1997. The next largest runs were 55 days, which ended in June 1995, and 38 days, which ended in March 1995.

Contrast that information with the longest period that a market was *not* in a trend. That was 204 days, which ended in November 1992 (Figure 7). The next longest period of *r*-squared of less than 0.70 ended in July 1996, and that was 102 days. The trend is certainly not your friend with numbers like these. This information may be why many moving average systems fail.

This is not to suggest, however, that moving average indicators are no longer useful. Linear regression has in common with moving averages a shortcoming, and that is the difficulty in selecting the correct interval or length of the lookback period. With linear regression, it is clear when the line no longer fits the market, but with a moving average, the user must rely on the market trading below or above the moving average for that information. Figure 8 summarizes the major differences between moving averages and linear regression.

EARLY WARNING SIGNS

Since linear regression is a measure of past performance, is there a method that can be used to help with trading? There is: the average length of the run can be used to play the odds in using linear regression. STOCKS & COMMODITIES Contributing Editor Tushar Chande first proposed an indicator to observe the errors that are generated about the regression line. If a market has a very low *r*-squared reading, then this market is trading within a consolidation or congestion phase. In this instance, the usual statistical indicators are not useful, except for what they are *not* telling you. Because low or high *r*-squared readings are fleeting, perhaps the place to look for the start of trends are markets exhibiting these conditions. This supports traditional technical analysis regarding breakouts from congestion or trading ranges.

For example, observe how the errors started to occur on the positive side in Figure 9 when there was a low *r*-squared reading in the Russell 2000 index. The errors have been adjusted to reflect a percentage error from the actual. This has no significant statistical inference, but it does help in visualizing the size of errors. This technique is useful, because

LINEAR REGRESSION VS. MOVING AVERAGES

Attribute	Linear Regression	Moving Average
Indicates strength of trend	Yes	No
Direct reading of slope	Yes	No
Standardized errors	Yes	No
Interval indication (length)	No	No
Exponential (parabolic) move	No	Yes

FIGURE 8: Linear regression has in common with moving averages the difficulty in selecting the correct interval or length of the lookback period. With linear regression, it is clear when the line no longer fits the market, but with a moving average, the user must rely on the market trading below or above the moving average for that information. Figure 8 summarizes the major differences between the two.

while the linear regression gives information about the previous 30 days, the trader has to make an inference about the market going forward. Observing the errors can give you that early warning signal. The errors can be stated simply as:

((Actual-Predicted)/Predicted)) 100

ADDITIONAL TRADING AIDS

Since the *r*-squared indicator doesn't identify whether the market is in an uptrend or a downtrend, it is useful to paint the bars to mark them as one or the other if the *r*-squared reaches 0.70 or higher. This may be accomplished by plotting the slope of the regression line, but actual crossing to an up- or downtrend is more dramatic and gives a visual reference.

Figure 10 is of the March 1998 Treasury bond futures, in which the bond futures have been in an uptrend, punctuated by brief downtrends. These changes of trend are accompanied by low or zero *r*-squared readings. By definition, an *r*-squared of zero indicates that the statistical methodology indicates no trend is present. This is what technicians define as consolidation, and they can occur at tops or bottoms.

Another example is Figure 11. This is the TYX, the 30-year interest rate, disseminated by the CBOE in real time. Note the extended period that it has been in a downtrend, indicating rates are moving lower. This is a market with sharp reversals, usually caused by surprises in the monthly unemployment report. This report often sets the tone of the bond market for the coming month; careful analysis of the trend along with *r*-squared can often yield valuable information about the coming months.

Figure 12 is Ascend Communications, a stock that has fallen from a high in the low 80s to the low 20s in less than a

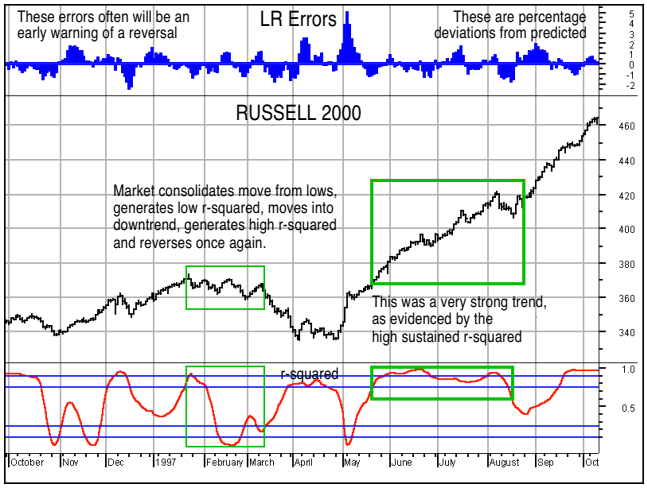


FIGURE 9: ERRORS, RUSSELL 2000 INDEX. Observe how the errors (top) started to occur on the positive side when there was a low r-squared reading in the Russell 2000 index. The errors have been adjusted to reflect a percentage error from the actual. This helps visualize the size of errors. This technique is useful, because while the linear regression gives information about the previous 30 days, the trader has to make an inference about the market going forward. Observing the errors can give you that early warning signal.

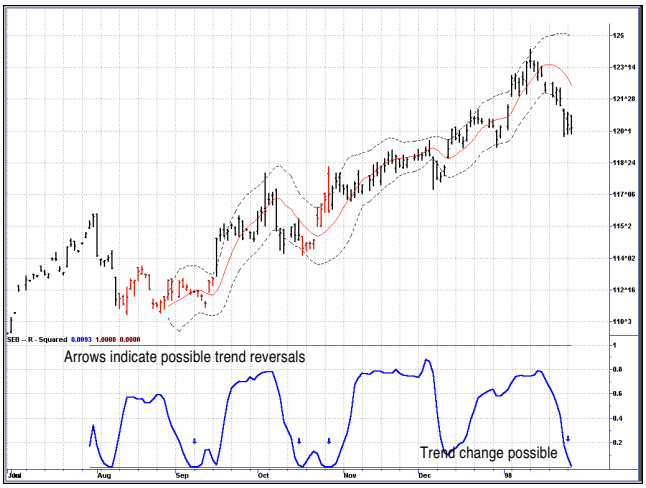


FIGURE 10: MARCH 1998 TREASURY BONDS. Here, the bond futures have been in an uptrend, punctuated by brief downtrends. These changes of trend are accompanied by low or zero r-squared readings.

year. The red bars in TradeStation represent a negative linear regression slope. See how the market reacts when the *r*-squared reading approaches zero. Look what occurred in early December when the market consolidated, and reversed from down to up. The bar changed color on what appeared to be a downtrend. The stock had a very large upmove in just a few days before settling back down again.

As of this writing, the stock has a very low *r*-squared and appears to be making another transition. This time, the stock has broken a short-term trendline and could be ready for a significant move to the upside. When a situation like this is encountered, it is best to let the market lead.

FURTHER CONSIDERATIONS

This technique is the linear part of the regression. If a market is in some power exponential move, a straight line isn't going explain that market; at least you'll be aware that the market is in a powerful move. The *r*-squared reading will, if nothing else, alert you to look to other methods for trading assis-

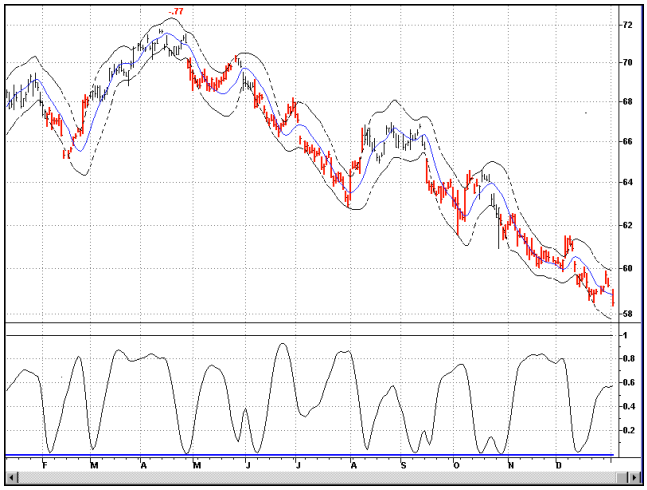


FIGURE 11: 30-YEAR TREASURY BONDS. Note the extended period that the TYX has been in a downtrend, indicating rates are moving lower. This is a market with sharp reversals, usually caused by surprises in the monthly unemployment report. This report often sets the tone of the bond market for the coming month; careful analysis of the trend along with r-squared can often yield valuable information about the coming months.

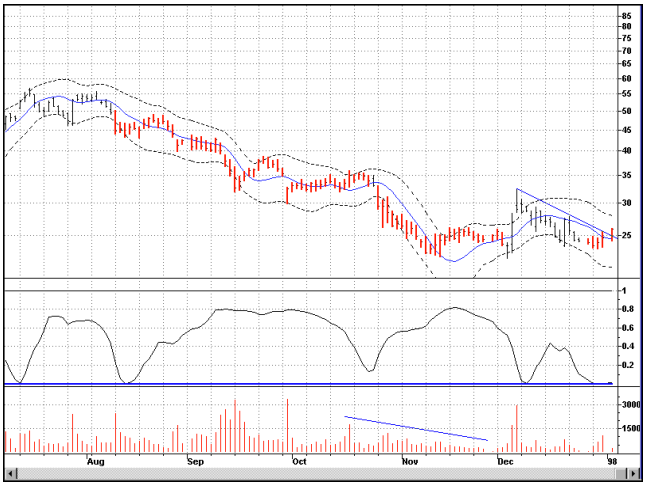


FIGURE 12: ASCEND COMMUNICATIONS. Here's a stock that's fallen from a high in the low 80s to the low 20s in less than a year. The bars, whick are red in TradeStation, represent a negative linear regression slope. See how the market reacts when the r-squared reading approaches zero. Look what occurred in early December when the market consolidated, and reversed from down to up. The bar changed color on what appeared to be a downtrend. The stock had a very large upmove in just a few days before settling back down again.

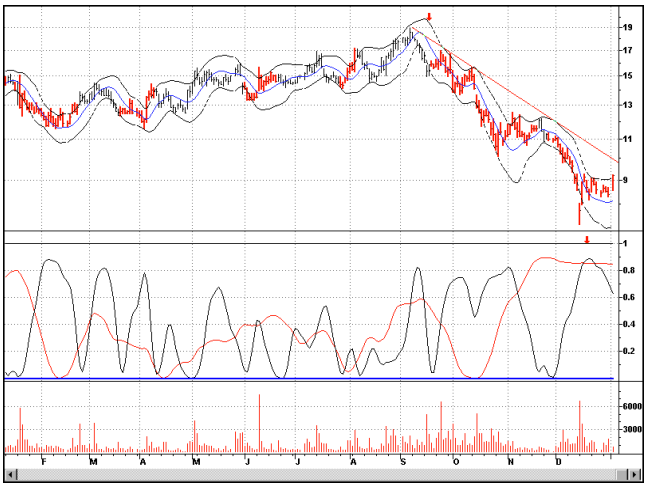


FIGURE 13: CYPRESS SEMI-CONDUCTOR. This chart illustrates the use of different lookback periods. The gray r-squared indicator is a 63-day linear regression and the black line is a 21-day linear regression. Note how the shorter-term r-squared cycles within the longer interval. Powerful moves tend to occur at high or low r-squared readings when the long and short r-squared indicators coincide. Just such a coincidence occurred in mid-December 1997, on a day when the stock was making new yearly lows.

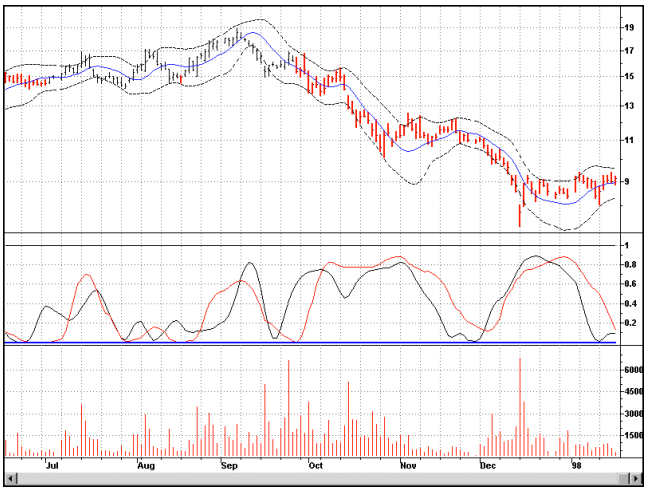


FIGURE 14: CYPRESS SEMI-CONDUCTOR, IN A DOWNTREND. The stock is clearly in a downtrend; in fact, it may have bottomed and is now in the process of moving sideways to up, which will reduce the current high r-squared reading.

tance. Multiple time periods for the linear regression might be of help, similar to the method used in moving average crossovers.

Figure 13, Cypress Semi-Conductor, illustrates the use of different lookback periods when assessing the trending nature of a stock or market. The gray *r*-squared indicator is a 63-day linear regression and the black line is a 21-day linear regression. Note how the shorter-term *r*-squared cycles within the longer interval. Powerful moves tend to occur at high or low *r*-squared readings when the long and short *r*-squared indicators coincide. Just such a coincidence occurred in mid-December 1997, on a day when the stock was making new yearly lows. It also appeared that the stock might have also had a key reversal.

Figure 14 shows the stock is clearly in a downtrend; in fact, it may have bottomed and is now in the process of first moving sideways to up, which will reduce the current high *r*-squared reading. If you choose to wait for a clear uptrend to occur, a low-risk trading opportunity may be lost. This is where a skillful trader can make a reasonable snap judgment by having useful statistics on hand. Using those statistics allows us to determine that this is not a good trending stock and so must be traded to extract maximum profits.

When using this technique, it is invaluable to study the frequency distributions of *r*-squared in order to make reasoned judgments about where the market will be headed in the future. It's true what they say: "Those who do not study history are doomed to repeat it."

Jack Karczewski is a Paine Webber broker in Scottsdale, AZ.

RELATED READING

Chande, Tushar [1992]. "Forecasting tomorrow's trading day," *Technical Analysis of* STOCKS & COMMODITIES, Volume 10: May.

†See Traders' Glossary for definition

S&C