Use of fundamental data for active investing in US equities

Howard B. Bandy, Ph.D.

NAAIM, March 2009

This paper evaluates methods of using fundamental data as input for trading US equities. It examines whether and how fundamental data can help decide what to buy, when to buy, and when to sell.

Fundamental economic data includes gross domestic product, inflation, interest rates, commodity prices, consumer sentiment, and similar items. Fundamental company data includes inventory turnover, earnings, sales, dividends, and similar items.

The focus is on techniques that may be useful to active managers.

Fundamental economic indicators

There are many economic indicators, published by many agencies. Some of those that have a significant effect on the prices of stocks, bonds, commodities, and currencies are listed below. Descriptions of the indicators, and of the schedule for release and revision of the data they report, come primarily from the websites of the reporting agencies. Additional material comes from a variety of sources accessed through the references listed in the appendix. The indicators are listed in the order of their timeliness – beginning with those reporting with the least delay following the period they cover.

- Weekly Unemployment – New filings for unemployment benefits. Published by the Employment and Training Administration, a division of the US Department of Labor. Published every Thursday, the data covers the week ending the previous Saturday, with revisions the following week.
- Consumer Confidence – Estimates of how consumers feel about jobs and the economy. The result of surveys conducted by The Conference Board, a private agency. Published monthly

on the last Tuesday of the month being reported. Revisions are rare and tend to be minor.

- ISM Manufacturing Survey – Results of surveys of purchasing managers conducted by the Institute for Supply Management (ISM), a private agency. Published monthly on the first business day after the reporting month. Seasonal adjustments are applied annually.
- ADP National Employment – An early report of employment and unemployment. The result of analysis of the payroll data processed by Automatic Data Processing (ADP), a private company. Published monthly two days before the government employment report. Minor revisions are published the following month. Annual revisions are made in March.
- Employment Situation – The government report of jobs created or lost and the unemployment rate. The result of surveys conducted by the Bureau of Labor Standards, an agency of the US Department of Labor. Published monthly on the first Friday of the month following the reporting month. The first data published is labeled "advance." One month later, along with that month's regular report, a revision labeled "preliminary" is released. Two months later, a revision labeled "final" is released. The series has its benchmark adjusted annually in June.
- Retail Sales – A report of spending on goods (not including services and not adjusted for inflation). The result of surveys of retailers conducted by the US Census Bureau, an agency of the US Department of Commerce. Published monthly about two weeks after the reporting month. Revisions are published for the next two months, and the series is re-benchmarked annually in March.
- Industrial Production and Capacity Utilization – A measure of the volume of manufacturing activity (not taking price into account). The results of data collected from companies by the Federal Reserve. Published monthly around the 15th of the month following the reporting month. Revisions are published for the next three months. The series is readjusted every fall.
- Producer Price Index (PPI) – Measures the changes in prices paid by businesses. The result of data collected from goods producers by the Bureau of Labor Statistics, an agency of the

US Department of Labor. Published two to three weeks after the reporting month. Revised once four months later. The series is revised annually in January.

- Housing Starts and Building Permits – Measures activity related to construction of new housing. The result of telephone and mail surveys conducted by the Census Bureau, an agency of the US Department of Commerce. Published monthly two or three weeks after the reporting month. Revisions are published for the next two months, and seasonal adjustments are made annually in April.
- Consumer Price Index (CPI) – The commonly used measure of inflation at the consumer level. The result of data collected by the Bureau of Labor Standards related to prices for a defined basket of goods. Published monthly, usually about two days after the PPI. No monthly revisions, but annual readjustments are computed in January. Since this series is used as the basis for annual changes in many payments, such as Social Security, some people suspect that the CPI is knowingly underestimated.
- Personal Income and Spending – A measure of the amount of money households receive and the amount they spend. The result of data collected from a variety of sources by the Bureau of Economic Analysis, an agency of the US Department of Commerce. Published monthly about four weeks after the end of the reporting month. Revisions are computed and published for the next two or more months, with annual revisions around July, and re-benchmarking every four or five years.
- Durable Goods Orders – A measure of orders received at factories. The result of data collected from manufacturing companies by the Census Bureau. Published monthly about three or four weeks after the end of the reporting month. Revisions, which can be substantial, are published one week after initial release, then for the next two months.
- International Trade in Goods and Services – A measure of exports and imports of goods and services. The result of data collected on exports by the Census Bureau and the Bureau of Economic Analysis. Published monthly the second week of the second month following the reporting month. Revisions are published monthly for several succeeding months. The

series is revised annually in June.

- Productivity and Costs – A measure of the productivity of workers who produce goods and services. The result of data collected by the Bureau of Labor Statistics from a variety of sources, including monthly payroll reports and gross domestic product. Reported quarterly about five weeks following the reporting quarter. Revisions are made monthly for three months, and as necessary as the underlying economic series are themselves adjusted.
- Gross Domestic Product (GDP) – The single number is the sum of the value of all goods and services produced in the United States. The report is broken down into detail by sector of the economy and by geographic area. It is the result of data collected from a variety of sources by the Bureau of Economic Analysis. Reported quarterly the final week of the month following the reporting quarter (final week of April for the first quarter). Revisions are reported each of the next two months, with annual revision in July and benchmark changes every five years.
- Current Account Balance – A report on the country's trade with the remainder of the world. Results of data collected by the Bureau of Economic Analysis. Reported quarterly about two and one-half months after the reporting quarter. There is one monthly revision three months later, and annual benchmark changes in June.

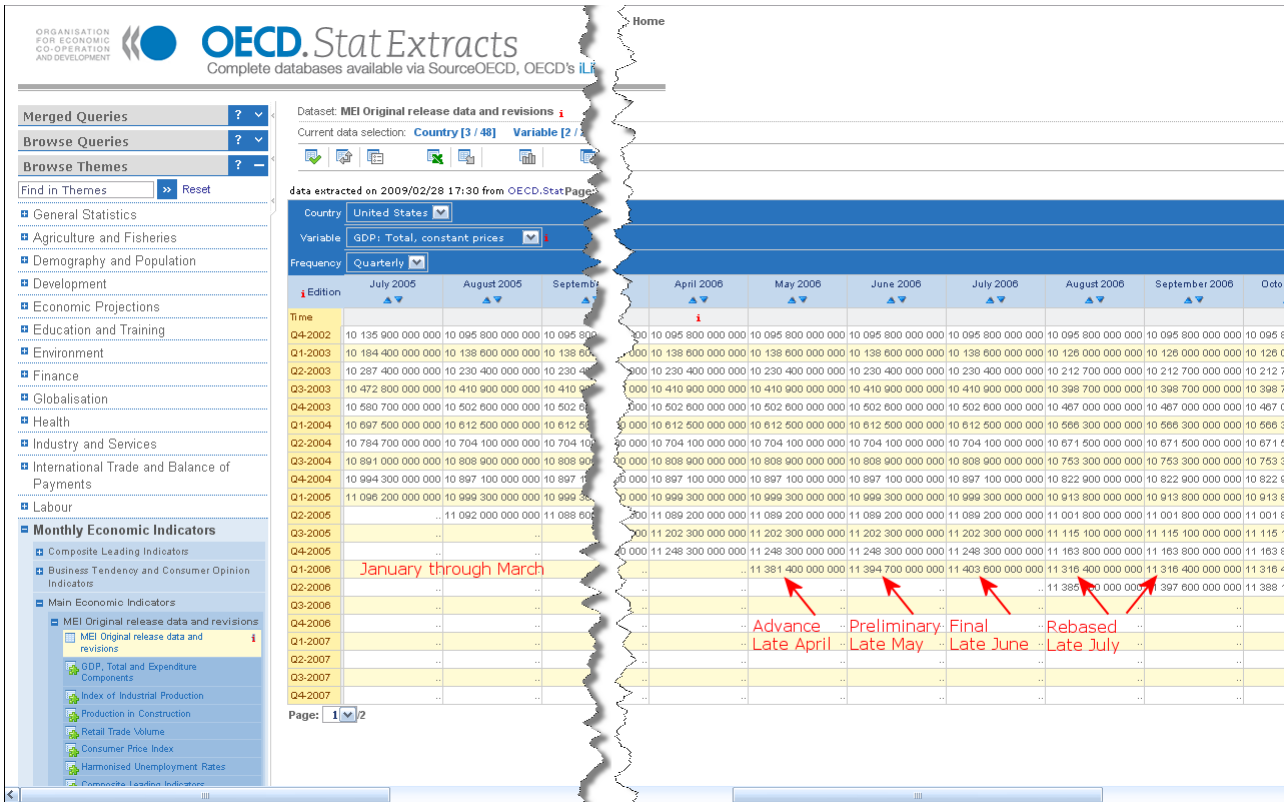
There are many more series than the sixteen described above. Some are followed closely and others widely ignored. One series that has intentionally been omitted from the list given above is the Index of Leading Economic Indicators (LEI). That index has been completely redesigned several times in the past two decades. It currently has ten components, chosen from among hundreds of available candidates to best fit the in-sample period used during its redesign. There is speculation that the most recent redesign was made to avoid having the LEI forecast an economic recession.

The following tables summarize the economic indicators described.

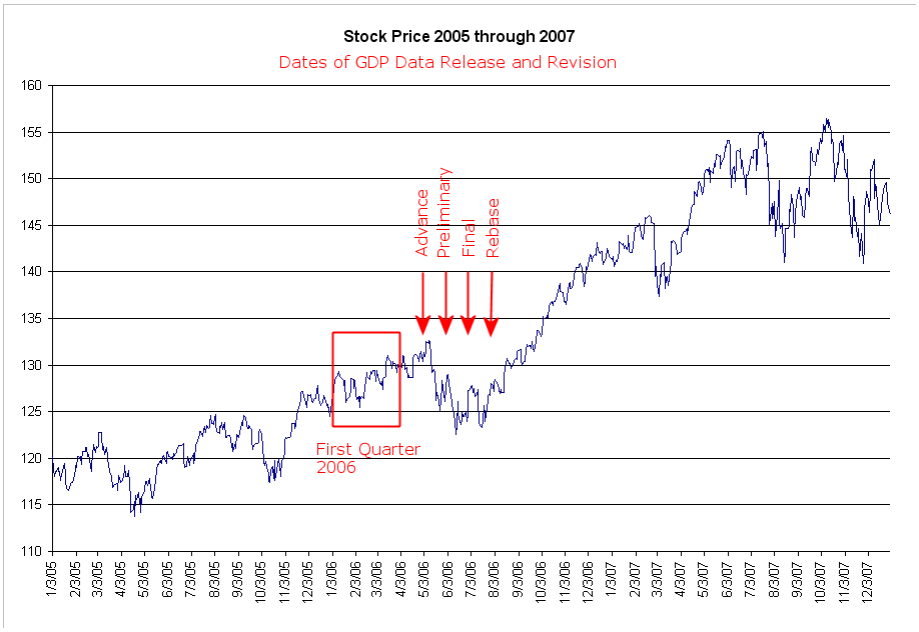
Economic Indicator	Publisher	Freq	Initial Release (Approx)
Weekly Unemployment	Employment and Training Administration	W	Thursday following
Consumer Confidence	Conference Board	M	Last Tuesday of period
Manufacturing Survey	Institute for Supply Management	M	1st day following
ADP National Employment	Automatic Data Processing	M	2 days before BLS Report
Employment Situation	Bureau of Labor Statistics	M	1st Friday following
Retail Sales	Bureau of the Census	M	14 days following
Industrial Production	Federal Reserve	M	15 days following
Producer Price Index	Bureau of Labor Statistics	M	16 days following
Housing Starts	Bureau of the Census	M	17 days following
Consumer Price Index	Bureau of Labor Statistics	M	18 days following
Personal Income	Bureau of Economic Analysis	M	4 weeks following
Durable Goods	Bureau of the Census	M	4 weeks following
International Trade	Bureau of Economic Analysis	M	6 weeks following
Productivity	Bureau of Labor Statistics	Q	5 weeks following
Gross Domestic Product	Bureau of Economic Analysis	Q	End of month following
Current Account	Bureau of Economic Analysis	Q	2 1/2 months following

Economic Indicator	Revisions	Adjustments	URL of report
Weekly Unemployment	1 week		http://www.ows.doleta.gov/unemploy/claims_arch.asp
Consumer Confidence	Rare		http://www.conference-board.org/economics/consumerconfidence.cfm
Manufacturing Survey	Rare	January	http://www.ism.ws/ismreport/index.cfm
ADP National Employment	1 month	March	http://adpemploymentreport.com/pdf/FINAL_Report_January_09.pdf
Employment Situation	1 & 2 months later	June	http://stats.bls.gov/news.release/empsit.toc.htm
Retail Sales	1 & 2 months	March	http://www.census.gov/marts/www/marts.html
Industrial Production	1 & 2 months	March	http://www.federalreserve.gov/releases/g17/Current/default.htm
Producer Price Index	4 months	February	http://www.bls.gov/ppi/
Housing Starts	1 & 2 months	April	http://www.census.gov/const/www/newresconstindex.html
Consumer Price Index	Rare	February	http://www.bls.gov/cpi/
Personal Income	1 & 2 months	August	http://www.bea.gov/newsreleases/national/pi/2009/pi1208.htm
Durable Goods	1 & 2 months		http://www.census.gov/indicator/www/m3/adv/
International Trade	1 to 4 months	June	http://www.bea.gov/newsreleases/international/trade/tradnewsrelease.htm
Productivity	1 & 3 months	Per GDP	http://www.bls.gov/lpc/
Gross Domestic Product	1 & 2 months	July	http://www.bea.gov/newsreleases/national/gdp/gdpnewsrelease.htm
Current Account	3 months	June	http://www.bea.gov/newsreleases/international/transactions/transnewsrelease.htm

The following figure illustrates the values reported for the US GDP. The highlighted reporting period is the first quarter of 2006. Note the dates and reported values for the initial report and the revisions.



The following chart shows about three years of daily price data for a stock that an active manager might trade. The box identifies the first quarter of 2006 and the annotations identify dates the GDP data covering that quarter are released, revised, and re-based. The data is taken from the figure above. The chart is intended to illustrate the lag between the period being reported and the date the final figure is available for use.



Challenges of using economic indicators

In order to be valuable, any indicator, including any economic indicator, must be:

- Timely
- Accurate
- Predictive

As is evident from the descriptions of the economic data and from the chart above, there are several issues that complicate use of fundamental economic data for trading.

1. *Timeliness related to reporting granularity.* Fundamental data is reported annually, quarterly, monthly, or weekly. Trading decisions are made monthly, weekly, daily, or intra-day. If the stock price is reported and acted upon more frequently than the economic indicator, there will be many time periods (data bars) where there is no data for the economic indicator. In order to have a value to use in calculations, the latest value that does exist will be copied forward until a new value is received. The only time that value can change is on those days (bars) when a new report is issued.

2. *Timeliness related to revision.* Economic indicators, and other fundamental data, are reported, then revised at later dates. When the historical data is retrieved from the data provider, it will usually be a series that consists of only the final revision data. In order to maintain

consistency, the data value associated with a given time period cannot be used in the trading system until after the date its final revision is published.

An alternative is to use a series that consists solely of data initially released, with no revisions or adjustments applied. The trading system would then be based on initial release data rather than final release data. Some data providers maintain a record of the initial report and all subsequent revisions, enabling the consumer of the data to create the data series that is most useful to them. That approach may work for economic series that seldom have significant revisions. But it will not work for series that are regularly heavily revised.

3. *Accuracy related to revision.* Government statistical series are regularly given annual adjustments, are re-based, and re-benchmarked. Re-basing sets a new date for the base of the index (the date it has a value of, say, 100) and adjusts all data in the series accordingly. Re-benchmarking recalculates the relationship between indicator series, adjusting those that depend on others. Any of these operations result in a revised historical data series. A trading system based on an earlier revision may not be profitable when the newly revised data is used.

4. *Accuracy related to bias.* There is a bias to any reported data. That bias is both unknown and unknowable to users of the data. Whether the report is unintentionally biased due to an innocent data preparation error or omission, or intentionally misleading, outsiders probably cannot detect the bias, its amount, or its reason. They have little alternative but to accept and use the data as reported. Bias introduces a *systematic* error into the reported statistic.

5. *Accuracy related to measurement.* The fundamental statistic reported is the result of interpretation of reports, questionnaires, and interviews. Even if there is a single correct number representing that statistic, the reported number is only one estimate of that value. If the same data and same procedures are used to recalculate the same statistic several times, the results will differ slightly from each other due to random variations in measurement and interpretation. The more subjective the question, the more random error in the answer. Preparers of these reports must be careful to avoid confusing precision with accuracy.

Measurement introduces a *random* error into the reported statistic.

6. *Predictive.* Whether the fundamental data is predictive depends on the strength of the relationship, the efficiency with which the market assimilates the information, and the insight and skill of the developer of the trading system. Remember to follow good modeling and validation practices. Keep enough data reserved for out-of-sample testing. In-sample results are always good and have no value in predicting the profitability of a system when traded on unseen data.

For all of these reasons, it is difficult to incorporate the fundamental data series with the daily or weekly price series representing the trading prices. In order to use fundamental data as a component of trading decisions, it is necessary to find surrogate data series that:

- Reflect changes in the fundamental data.
- Are reported on the same time schedule as the price series.
- Are never revised.
- Represent transactions made in public and reported through a clearing agency.

The sixteen economic indicators described above fall into just a few categories:

- Consumer – confidence, jobs, spending, housing
- Manufacturing – orders, productivity, sales
- Inflation
- Debt

Fundamental company data

Fundamental data for individual companies includes sales, earnings, dividends, book value, inventory turnovers, and similar items. Similar difficulties arise with company data as with economic indicators — granularity, timeliness, revisions, restatements, changes in company organization, and bias.

When dealing with company data, changes to the organization of the company result in changes to the data series. Explanations and footnotes may be added to annual reports, but historical data is

seldom adjusted, or even adjustable. For example, in January 2009, pharmaceutical Pfizer acquired Wyeth and gained their over-the-counter drug lines. Pfizer sold its consumer products division to Johnson and Johnson in 2006, and sold its medical device division in the 1990s. Since Wyeth no long exists as an independent company, its data is no longer maintained, but belongs in the category of inactive companies that have been discontinued, delisted, or absorbed. The Pfizer data might be adjusted by a pharmaceutical industry analyst at the request of a major client, but that data series will probably never be available for wide distribution at low cost.

The problem can become even more complex when a company contains very different divisions, each of which contributes revenue. For example, major manufacturers who operate their own captive finance divisions may have seen the majority of their earnings move from manufacturing to the financing operations. Since the company had been primarily a manufacturing company, they are probably grouped with and compared to other manufacturers. If such a company divests its financial operation, the data series will show a dramatic change, which is seldom adjusted.

The question of accuracy and trust is illustrated by the list of companies suspected or convicted of misleading investors by misstating the company's fundamental data. It includes Adelphia, Enron, Nortel, Refco, Tyco, and WorldCom.

Asset classes, indexes, and futures

There are four asset classes a manager might want to use:

- Equities
- Bonds and Notes
- Commodities
- Currencies

A search of the databases results in a list of fourteen indexes that might be related to the economic indicators and serve as surrogates for them and for the asset classes. In addition, most of them trade as futures contracts. The indexes are all sponsored by reputable companies, all have history dating from the early 1990s or earlier, all are widely followed, and all fit neatly into one of the four

asset categories. They are (using the Quotes Plus symbol):

- !CCI – Reuters CRB index
- !DXYZ – The dollar index
- !FVX – Yield on five year note
- !GLDLA – Gold, afternoon fix
- !IRX – Yield on 13 week treasury bill
- !OEX – S&P 100 index
- !OIX – CBOE Oil index
- !RUA – Russell 3000 index
- !RUI – Russell 1000 index
- !RUT – Russell 2000 index
- !SPX – S&P 500 index
- !TNX – Yield on 10 year bond
- !TYX – Yield on 30 year bond
- !VIX – Volatility index

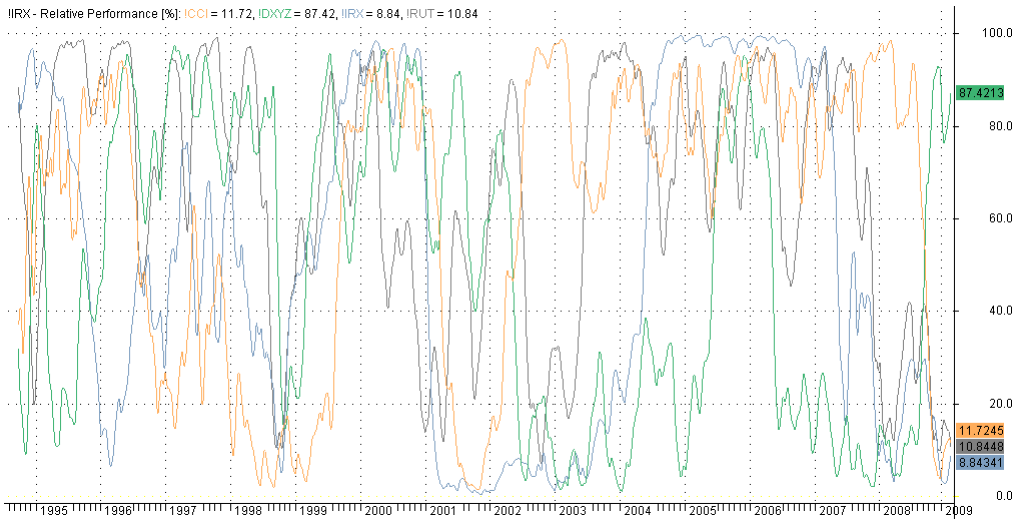
The correlation, measured over a 252 day period, of each index with every other index is displayed in the following table. Each value is the average of twelve years. The cells colored green have an R-squared value of 0.50 or more. Each column is labeled as a commodity, dollar, interest rate, or equity. There is little difference among the members of each asset class, relative to the other indexes. For example, the correlations of all the equity indexes are very similar. One index of each asset class was identified to keep.

Correlatio	Commodi	Dollar	Interest	Commodi	Interest	Equity	Commodi	Equity	Equity	Equity	Equity	Interest	Interest	Sentiment
	!CCI	!DXYZ	!FVX	!GLDLA	!IRX	!OEX	!OIX	!RUA	!RUI	!RUT	!SPX	!TNX	!TYX	!VIX
!CCI	1.00	0.31	0.26	0.36	0.24	0.26	0.36	0.26	0.26	0.27	0.26	0.25	0.25	0.22
!DXYZ	0.31	1.00	0.34	0.39	0.31	0.28	0.26	0.28	0.28	0.27	0.28	0.33	0.31	0.22
!FVX	0.26	0.34	1.00	0.29	0.35	0.34	0.27	0.34	0.34	0.34	0.34	0.90	0.76	0.27
!GLDLA	0.36	0.39	0.29	1.00	0.28	0.26	0.29	0.26	0.26	0.28	0.26	0.29	0.28	0.21
!IRX	0.24	0.31	0.35	0.28	1.00	0.26	0.27	0.26	0.26	0.27	0.26	0.33	0.30	0.21
!OEX	0.26	0.28	0.34	0.26	0.26	1.00	0.37	0.94	0.95	0.67	0.97	0.34	0.33	0.59
!OIX	0.36	0.26	0.27	0.29	0.27	0.37	1.00	0.39	0.39	0.37	0.38	0.27	0.27	0.30
!RUA	0.26	0.28	0.34	0.26	0.26	0.94	0.39	1.00	1.00	0.79	0.99	0.34	0.32	0.60
!RUI	0.26	0.28	0.34	0.26	0.26	0.95	0.39	1.00	1.00	0.75	0.99	0.34	0.32	0.60
!RUT	0.27	0.27	0.34	0.28	0.27	0.67	0.37	0.79	0.75	1.00	0.73	0.33	0.31	0.50
!SPX	0.26	0.28	0.34	0.26	0.26	0.97	0.38	0.99	0.99	0.73	1.00	0.34	0.33	0.60
!TNX	0.25	0.33	0.90	0.29	0.33	0.34	0.27	0.34	0.34	0.33	0.34	1.00	0.89	0.27
!TYX	0.25	0.31	0.76	0.28	0.30	0.33	0.27	0.32	0.32	0.31	0.33	0.89	1.00	0.26
!VIX	0.22	0.22	0.27	0.22	0.21	0.59	0.30	0.60	0.60	0.50	0.60	0.27	0.26	1.00

The correlation among the four is shown in the following figure.

Correlatio	Commodi	Dollar	Interest	Equity
	!CCI	!DXYZ	!IRX	!RUT
!CCI	1.00	0.31	0.24	0.27
!DXYZ	0.31	1.00	0.31	0.27
!IRX	0.24	0.31	1.00	0.27
!RUT	0.27	0.27	0.27	1.00

The four indexes are plotted together in a single chart, shown below. If they were to show a consistent relationship of peaks and valleys, with one leading, another following, the third following later, and the fourth last, then those four indexes would make an ideal set to be the basis of a rotational trading system. Or, perhaps be useful in creating filters for broad market timing.



Even though there is no consistent sequence of peaks and valleys, there is almost always at least one class scoring high. That is promising, but is outside the scope of this paper.

Another kind of series that can be used as a surrogate for an economic indicator is the continuous futures contract. A continuous contract is formed by joining individual front-month contracts together to form a single series.

Since the trading hours for both the indexes described above and futures contracts are approximately the same as the trading hours for stocks, the problems with timeliness and accuracy are removed.

Futures trading began with agricultural products — forward contracts in the 1850s, which evolved into futures contracts traded on exchanges by the 1870s. Other physical commodities were added over time. The big changes came in 1972 when currency futures were added, interest rates in 1975, and equities in 1982. By 1985, trading in agricultural contracts accounted for less than 25 percent of the total. Consequently, there is complete and accurate data for the prices of equities, interest rates, currencies, and commodities from 1982 on. Individual futures contract data and continuous contract data are available at low cost from data vendors such as Yahoo, Quotes Plus, Norgate Premium Data, and Pinnacle. Examples of futures contracts, some of which might be useful:

- Equities — S&P 500, NASDAQ 100, Dow Jones Industrials, Russell 2000
- Interest Rates — Eurodollars, Treasury Bills, Notes, Bonds
- Materials and Energy — Crude Oil, Copper, Lumber, Gold, CRB Index
- Currencies — US Dollar Index, Euro, Canadian Dollar
- Sentiment — Volatility Index
- Agricultural — Corn, Soybeans, Cattle, Cotton

Commodities have their own fundamentals, which are closely related to fundamental economic indicators. Data on supply, production, and consumption is available, and traders of futures contracts may be interested in that. But when the commodity data is being used as a surrogate for the economic indicators, all that is necessary is the price series.

The business cycle

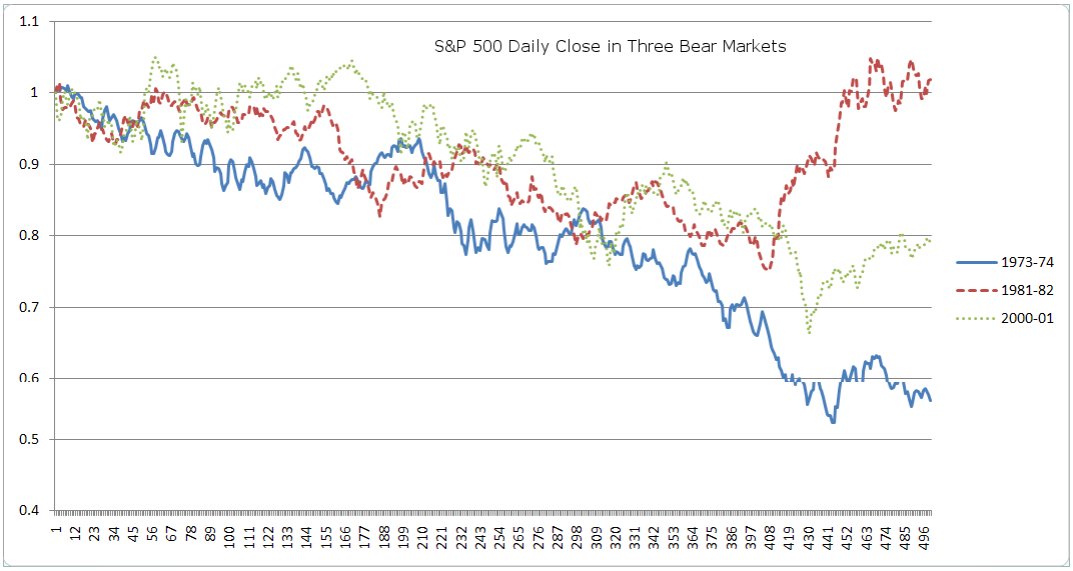
Investment advisors and publications describe the sequence that the business cycle follows. Conventional wisdom views the business cycle as repeatedly progressing through phases. Depending on their business operations, stocks tend to perform better in some phases than others.

There are two time frames for this analysis. The first time frame measures cycles relative to recession lows and the following recovery. In the 94 years since World War I, there have been 18 recessions, as defined by the National Bureau of Economic Research. The average time between recession lows is 5.2 years, with a range of 1 year to 11 years. There are too few of these events to draw general conclusions, and the holding periods are longer than fit the definition of active

management. Until the current recession, it appeared that recessions were becoming less frequent and shorter, and expansions longer. But there is too little data to be conclusive. There may be some indicators related to recessions that could be used as filters for more frequent timing and for stock selection, but those are not discussed in this paper. The following figure shows the S&P 500, plotted on a logarithmic scale, with the periods the US was officially in a recession highlighted.

Martin Pring and Sam Stovall have both written books outlining the sequence of the business cycle as it relates to recession and recovery. Both books are listed in the References. Pring's book is in print, and his website has an article describing his ideas. Stovall's book is out of print, but there are several internet sites that outline his ideas — see Investopedia, for example.

The following figure plots the S&P 500 (Yahoo Financial symbol ^GSPC) from the beginning of each of several recent recessions. Note that, even as the broad market is falling, there are significant and tradable rallies.

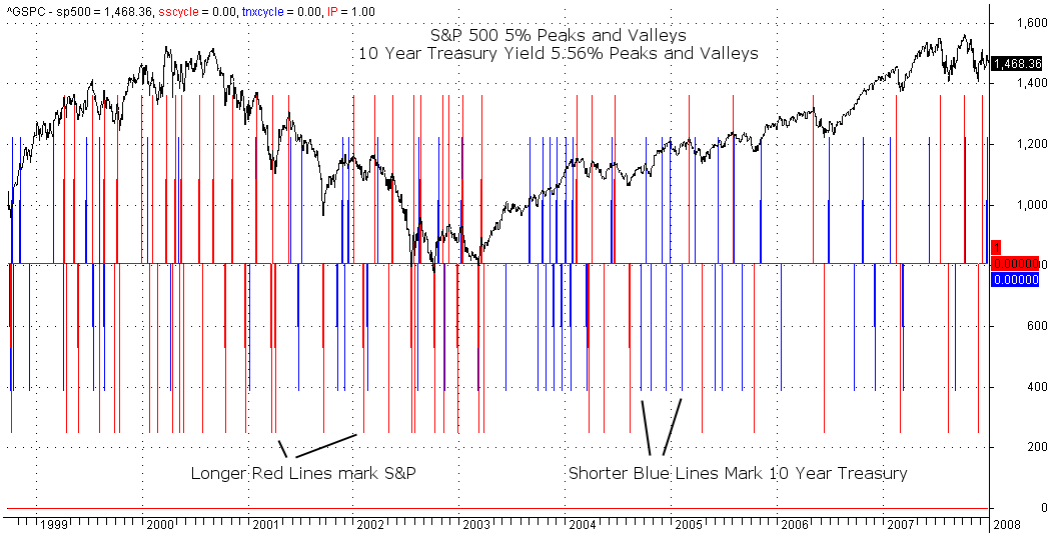


The second time frame assumes several cycles per year, and will be discussed further here.

A simple experiment will test whether two data series act in predictable phases. Define a period of time to test, say 10 years. Working with the first data series — the one that will be traded — apply a zigzag indicator and adjust the percentage parameter to correspond to the desired magnitude of the trades — say, 5 percent. This means hold a long position for any move that has at least a five percent gain but does not have a five percent drawdown. The zigzag will show the peaks and valleys. Buy every valley and sell every peak. Counting valleys, 20 valleys would correspond to the series having two complete cycles each year, on average, which corresponds to two holding periods, each an average of three months long. Working with the second data series, apply a zigzag indicator and adjust the parameter so that there are the same number of peaks and valleys as the first series. If the two series follow each other in phase, the sequence of 40 valleys will alternate — each valley from the first series is both preceded and followed by a valley from the second series. Each sequence of successive valleys by one of the series is counted as a "run." If the two series acted perfectly in phase, the combined results of the two series would result in 40 runs, each of length 1. Statistical tests, such as the runs test, can be used to measure the amount by which the two series differ from being perfectly in phase. But the goal is not to perform statistical tests, rather to develop profitable trading systems. So, if there appears to be regular alternation between valleys, write and test a trading system that identifies valleys in the second data series. Recognizing that a

valley in the second series has passed alerts the manager to be ready to take a long position in the first data series.

This technique was applied to the S&P 500 and the 10 year Treasury Note Yield (symbol ^TNX) for the period 1/1/1988 through 1/1/2008. The following figure illustrates the sequence of the peaks and valleys.



The percentage change for the S&P 500 was set to 5%. That resulted in 58 valleys over the 20 years. The percentage change necessary for the 10 year Treasury Note Yield to give the same number of valleys was 5.56%, determined by trial and error. Note that the two series do not alternate, but show runs of one series followed by runs of the other. Although this technique did not work well using the 10 Year Treasury Note, there may be other series that would work better.

Sectors as surrogates

Sector indexes or sector exchange traded funds are created by selecting individual issues that have similar characteristics. The price of the composite is a data series that reacts to economic fundamentals in the same manner as the individual component issues react. But there is less day-to-day variability. The price movements are smoother because a portion of the random noise contained in the data of the individual issues cancels each other out. A search of the database reveals a long list of exchange traded funds and indexes that might be useful.

Sector exchange traded funds

Standard and Poor's has created a set of nine select sector exchange traded funds. All 500 of the stocks that are components of the S&P 500 index are used. Each stock is also a component of the sector fund that its business fits best. The funds have between 29 and 83 components. The categories and ETFs are:

- Basic materials — XLB
- Consumer discretionary — XLY
- Consumer staples — XLP
- Energy — XLE
- Financials — XLF
- Health care — XLV
- Industrials — XLI
- Technology — XLK
- Utilities — XLU

We know it is possible to use those sector ETFs to develop profitable trading models. The question addressed in this paper is whether they can be used as surrogates for economic indicators.

Five stocks were selected from each ETF — the four stocks with the highest weighting in the ETF, and one stock that has about 2% weighting. Keep in mind that all the stocks in the S&P 500 are large cap stocks.

Daily percentage price changes were computed for each stock, each sector, and the S&P 500 for about 10 years. For each of the 45 stocks, a multiple linear regression was used to measure the relationship between the stock, its sector, and the S&P 500. This is a descriptive model, not a predictive model, so we will not be concerned with the problems that arise when regressing non-stationary times series against each other.

The form of the regression equation is:

Stock = Intercept + A * Sector + B * SP500

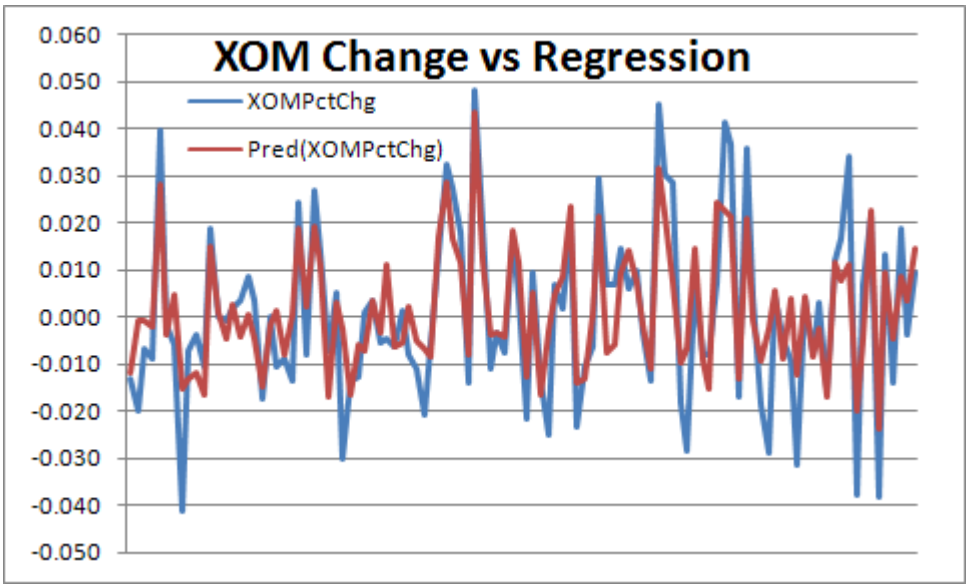
where Stock, Sector and SP500 are the daily percentage prices changes.

The regression equation for Exxon Mobile, XOM, related to the Energy Select Sector, XLE, and the S&P 500, SPY, is:

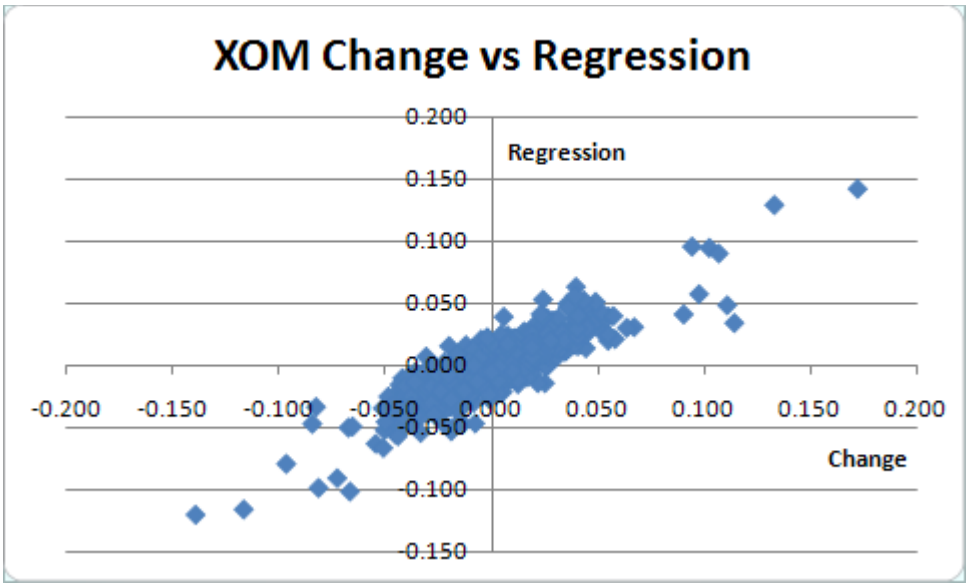
XOM = 0.00 + 0.74 * XLE + 0.14 * SPY

The regression has an R-squared of 0.737.

The following chart shows the daily price change observed (the blue line that has the greater range) and the daily price change explained by the regression (the red line) for a short, but typical, period.

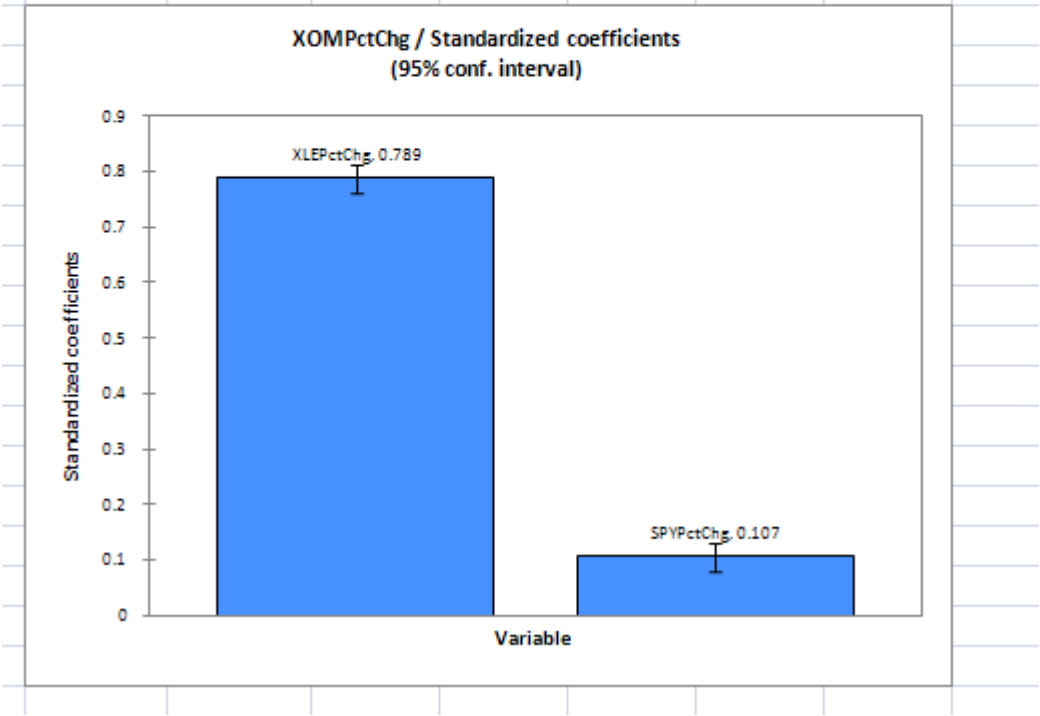


The scatter plot of the entire ten years is shown in the next chart. The observed change is plotted on the horizontal axis, the change explained by the regression on the vertical axis.

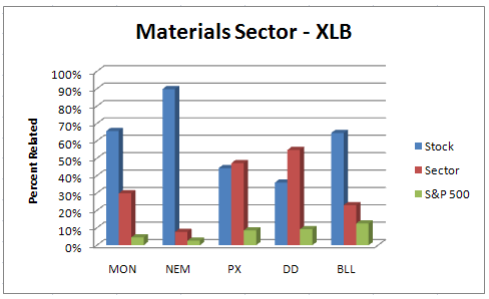


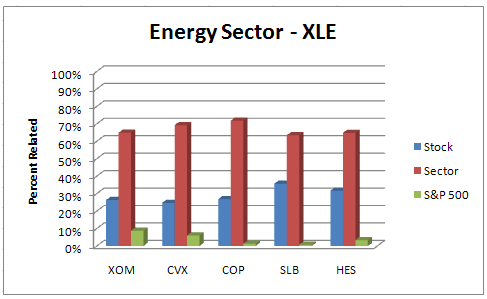
The next chart illustrates the magnitude of the coefficients associated with the sector and the broad market. The whiskers show the 95% confidence levels. The regression is highly significant.

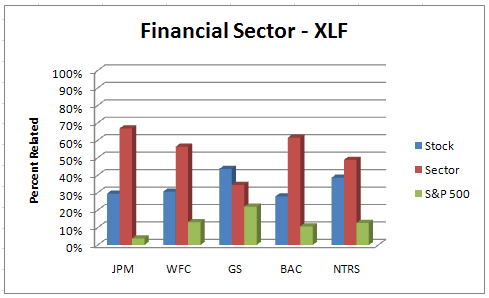
Source	Value	andard err	t	Pr > |t|	er bound (9	er bound (95%)
XLEPctChg	0.789	0.013	61.186	< 0.0001	0.763	0.814
SPYPctChg	0.107	0.013	8.275	< 0.0001	0.081	0.132

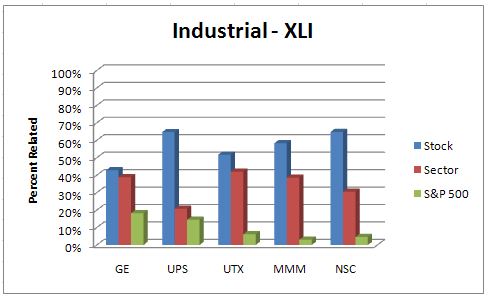


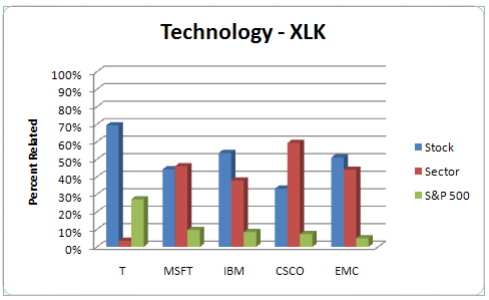
The following nine charts illustrate the relationship between the stock's price change, the price change of the sector it belongs to, and the price change of the broad market, for the 45 stocks and 9 select sectors chosen. The percentage of the stock's price change that is explained by the stock, the sector, and the broad market are shown by a group of three bars for each stock. The blue bar to the left shows the relative importance of the stock itself. The red bar in the middle shows the relative importance of the sector. The green bar to the right shows the relative importance of the broad market. The sum of the three bars is 100%. Note how some sectors strongly affect the stocks, as shown by tall red bars. Other sectors contribute less, as shown by the tall blue bars. The broad market contributes relatively little.

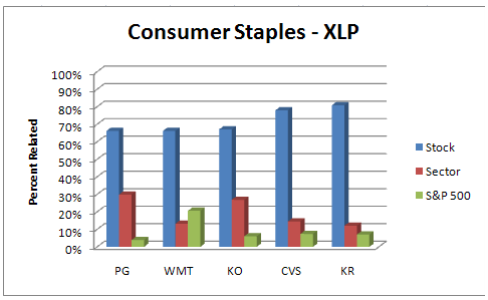

Materials Sector - XLB
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
MON
NEM
PX
DD
BLL
Stock
Sector
S&P 500

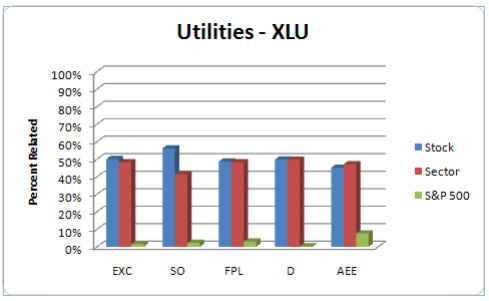

Energy Sector - XLE
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
XOM
CVX
COP
SLB
HES
Stock
Sector
S&P 500

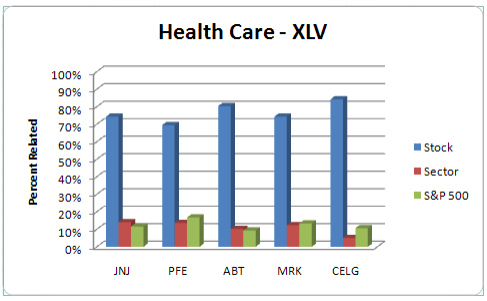

Financial Sector - XLF
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
JPM
WFC
GS
BAC
NTRS
Stock
Sector
S&P 500

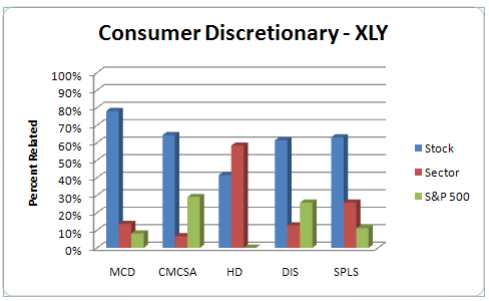

Industrial - XLI
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
GE
UPS
UTX
MMM
NSC
Stock
Sector
S&P 500


Technology - XLK
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
T
MSFT
IBM
CSCO
EMC
Stock
Sector
S&P 500


Consumer Staples - XLP
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
PG
WMT
KO
CVS
KR
Stock
Sector
S&P 500


Utilities - XLU
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
EXC
SO
FPL
D
AEE
Stock
Sector
S&P 500


Health Care - XLV
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
JNJ
PFE
ABT
MRK
CELG
Stock
Sector
S&P 500


Consumer Discretionary - XLY
Percent Related
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
MCD
CMCSA
HD
DIS
SPLS
Stock
Sector
S&P 500

For all 45 stocks, the average percentage contributions to the stock's price change are:

Stock itself	54.4%
Sector	36.3%
S&P 500	9.3%

Examination of the relationships of these stocks to their sectors shows that the nine S&P sector ETFs are too broad and too diverse to use as surrogates for economic indicators.

Other sector-based approaches

To narrow the scope of the sector, we can: create custom sectors; use sectors defined by industry codes; or use sectors defined by an agency.

Create custom sectors

To create a custom index, a population of stocks is examined. All of those in the Russell 3000, for example. The goal is to identify a number of issues, say 10 to 20, that behave the same, or close to the same, in all market conditions. An index of these is created and used to generate the trading signals. The trades taken can be a basket consisting of the stocks in the index, a few highly liquid stocks that are components of the index, or some related issue. The mathematics and mechanics of creating a custom index are straight-forward. But creating the index raises a serious issue of membership bias. The composition of the index when it is created today is almost certainly different than it would be if it had been created in the past. One way to reduce the effect of membership bias to is select the stocks that form the index as of a date in the past. Since the model validation process will involve in-sample development and out-of-sample testing, choose the date of the start of the out-of-sample period as the date for selecting the components of the index.

Use sectors defined by industry codes

Use of industry codes removes the effort of searching for issues that behave similarly, and it removes a portion of the membership bias. Instead, it relies on accurate assignment of industry codes, a task performed by a government or private agency.

The North American Industry Classification System (NAICS) is the coding system used by Federal

statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the U.S. business economy. NAICS is an agency of US Census Bureau. The NAICS codes are assigned to businesses according to their business activities. It is difficult to determine the NAICS code for a given company — retrieval by stock symbol is not available. NAICS codes are not usable for sector index creation.

The Industry Classification Benchmark (ICB) is a company classification system developed by Dow Jones and The Financial Times and the London Stock Exchange (FTSE). It is used to segregate markets into sectors within the macroeconomy. The ICB uses a system of 10 industries, partitioned into 18 supersectors, which are further divided into 39 sectors, which then contain 104 subsectors.

The Global Industry Classification Systems (GICS) is a company classification system developed by Standard and Poor's and Morgan Stanley Capital International (MSCI) Barra. GICS defines peer groups tightly and avoids grouping unlike companies together. The GICS structure is four levels — comprised of 10 sectors, 24 industry groups, 68 industries and 154 sub-industries.

GICS competes directly with ICB. The ten highest level categories are almost the same. ICB uses a "product-oriented" approach, where the category a stock is assigned is determined by what the company produces. Companies that produce goods are generally in different categories than those that produce services, although there are some companies that produce both and are difficult to classify. GICS uses a "market-oriented" approach, where the category a stock is assigned is determined by how its products are sold. Consumer staples companies, for example, include companies that provide both consumer products and services that are considered necessities.

Both systems expose the definitions of their top three levels freely. Both require a subscription to learn the definitions of the fourth and most narrowly defined categories. Both GICS and ICB require a subscription to allow retrieval of industry code through lookup of the stock ticker.

The stocks in the S&P indexes have been assigned one of the two-digit GICS sector codes. Those codes are:

10	Energy
15	Materials
20	Industrials
25	Consumer Discretionary
30	Consumer Staples
35	Health Care
40	Financials
45	Information Technology
50	Telecommunications Services
55	Utilities

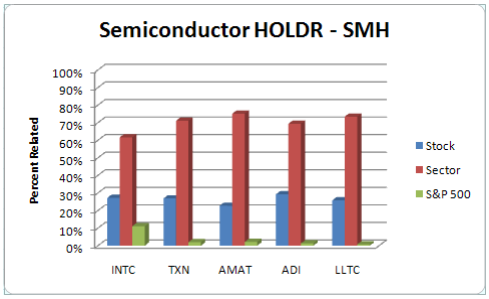
In order to be useful as surrogates for economic indicators, much finer resolution is required.

Use sectors defined by an agency

Use of an index created by an agency has the advantage that membership and survivorship issues were dealt with at the time they arose. If the components of the index are similar enough, that might work. As of March 2, 2009, all 79 of the components of the S&P Sector Select Consumer Discretionary ETF, symbol XLY, were assigned GICS code 25. But the XLY ETF included Amazon, Auto Nation, International Game Technology, New York Times, and Pulte Homes — clearly too dissimilar for this purpose. Another categorization system gives the following classification for those five stocks:

- Amazon — Technology, Internet, Retail
- Auto Nation — Retail, Miscellaneous
- International Game Technology — Travel and Leisure, Gaming
- New York Times — Media, Books and Newspapers
- Pulte Homes — Construction, Residential

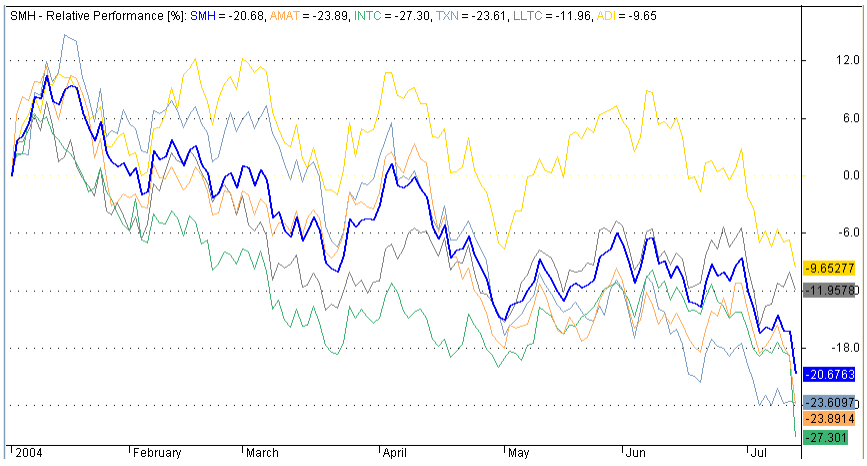
Some of the indexes and exchange traded funds are more precisely defined. For example, the HOLDRs Semiconductor ETF, symbol SMH, consists of about 18 stocks. The same regression model that was used with the nine S&P select sector ETFs was applied to daily percentage price changes for SMH and its top five holdings. The next chart shows the relative importance of the stock, the sector, and the broad market.



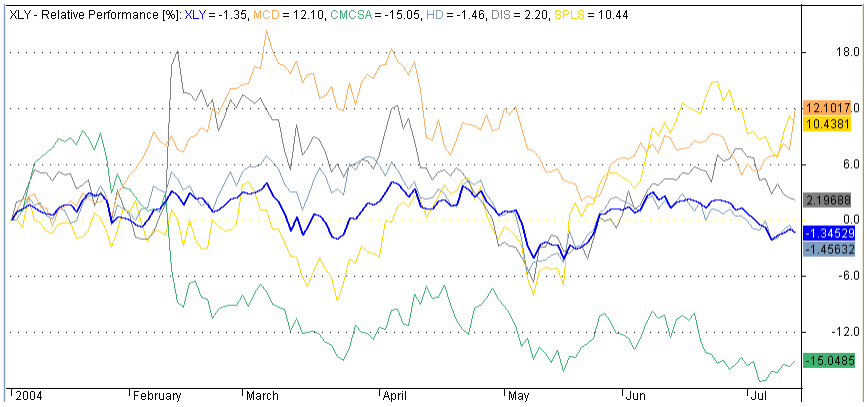
Note how strong the influence of the sector is. For all 5 stocks, the average percentage contributions to the stock's price change are:

Stock itself	26.3%
Sector	70.2%
S&P 500	3.5%

The following chart shows the relative prices of SMH (the heavier blue line in the middle) and the top five components for a few months in 2004. The SMH is very representative of the components.



Compare with the relative prices of XLY (the heavier blue line in the middle) and its top five components for the same time period.



More precisely defined sectors are easier to model, because their components tend to move in reaction to the same fundamental information. And after an accurate model has been created, trading the sector index itself or the liquid components of it are more profitable because there is less variation between the model and its components.

Summary and Conclusion

The goal of this paper was to evaluate methods of using fundamental economic indicators and fundamental company data as components of trading systems used by active managers. The conclusions are as follows:

- Economic indicators, such as those published by governmental and other agencies, are not appropriate for direct use in active trading systems for reasons involving their timeliness and accuracy.
- Company fundamental data are not appropriate for use in active trading systems for reasons involving their timeliness and accuracy.
- Selecting representatives of four asset classes — equities, bonds, commodities, and currencies — to form a trading system shows profit potential. There appears to be more potential for use in a rotational trading system than for use as surrogates for fundamental data.

- Selecting representatives of the business cycle — manufacturing, consumer, etc — to form a trading system also shows profit potential. Complications arise because of problems in defining the sector indexes. Again, there is potential for use in a rotational trading system, but not as surrogates.
- If sectors are carefully and narrowly defined, most of the price movement of individual securities is due to price movement of the associated sector.
- Surrogate data, primarily in the form of sector indexes defined and maintained by financial services companies, can be useful in active trading systems.
- Intermarket analysis, such as examination of interest rates, can be used for broad market timing, or to create filters to permit or block equity trades.
- Exchange traded funds can be used both to generate the trading signals and to take the position.
- Indexes and futures contracts can be used in trading systems of their own or as surrogates of economic indicators.

References

AAII — American Association of Individual Investors. Stock Screener Pro provides historical fundamental company data and a screening program. Subscription. http://www.aaii.com/store/sipro/index.cfm

American Economics Association. Extensive information of interest to economists. In particular, the Resources For Economists page, http://rfe.org/ with links to journals, articles, and data. Much of the content is free. http://www.aeaweb.org/index.php

Automatic Data Processing. Publishes the monthly ADP National Employment Report. http://adpemploymentreport.com/

Bandy, Howard. *Introduction to AmiBroker*. Blue Owl Press. 2008.

Bandy, Howard. *Quantitative Trading Systems.* Blue Owl Press. 2007.

Baumohl, Bernard, *The Secrets of Economic Indicators*, Second Edition, Wharton School Publishing, 2008.

BEA — Bureau of Economic Analysis. An agency of the US Department of Commerce. Publishes the monthly personal income report. Also provides other historical economic fundamental data. Free. http://www.bea.gov/

Capital IQ, a division of Standard & Poor's. Detailed data on public and private companies. Subscription. https://www.capitaliq.com/Main3/ourproducts_platform_fundamentals.asp

Compustat, a division of Standard and Poor's. Current and historical fundamental data on current and inactive companies. Subscription. https://www.compustatresources.com/support/index.html

Compustat. The top three levels of GICS. http://mi.compustat.com/docs-mi/help/gics_info.htm#utils

Conference Board. Publishes the monthly Consumer Confidence Index and Index of Leading Indicators. Mostly subscription. http://www.conference-board.org/

Department of the Census, an Agency of the US Department of Commerce. Publishes the monthly retail sales report. http://www.census.gov/

Economagic. Historical economic fundamental data. Free. http://www.economagic.com/

Edgar. Current and historical fundamental company data. Security and Exchange Commission filings. Subscription. http://www.edgar-online.com/

Federal Reserve. http://www.federalreserve.gov/

Federal Reserve Bank of New York. Reports related to monetary policy. Free. http://www.newyorkfed.org/

ICB. Click the "structure" link to open a pdf with the four level ICB structure defined and codes listed. http://www.icbenchmark.com/icb_structure.html

GICS definitions in Excel file format. http://www.mscibarra.com/products/gics/structure.jsp

Global Financial Data. Historical end-of-day OHLCV price data for active and delisted US stocks, company fundamental historical data, economic fundamental historical data. Subscription. https://www.globalfinancialdata.com/

Hemscott, a division of Morningstar. Current and historical fundamental data on current and delisted companies. Subscription. http://www.hemscottdata.com/database.html

ICB — Industry Classification Benchmark. Information about the ICB. Access to list of classification by company is by subscription. http://www.icbenchmark.com/

ICB — Industry Classification Benchmark. Wikipedia entry giving the definitions for industry, supersector, and sector. http://en.wikipedia.org/wiki/Industry_Classification_Benchmark

Institute for Supply Management. Publishes the monthly manufacturing survey. http://www.ism.ws/

Investopedia. Article on sector rotation giving Sam Stovall's industry rotation sequence. http://www.investopedia.com/articles/trading/05/020305.asp

Investopedia. Definitions of GICS and ICB classification methods. http://www.investopedia.com/articles/stocks/08/global-industry-classification-industrial-classification-benchmark.asp

MSCI Barra — GICS definitions in Excel file format. http://www.mscibarra.com/products/gics/structure.jsp

NBER — National Bureau of Economic Research. Definition of recession and list of recessions and depressions in the US. http://www.nber.org/cycles/

NAICS — North American Industry Classification System. An agency of the US Census Bureau. Definitions of the NAICS industry codes. http://www.census.gov/eos/www/naics/index.html

Norgate Premium Data. End-of-day historical and daily update data for US and Australian equities, futures, and foreign exchange. Subscription. http://www.premiumdata.net/

OECD – Organization for Economic Coordination and Development. Statistics, economic, and social data for 30 global countries. Free. http://www.oecd.org/

Pinnacle Data. End-of-day historical and daily update data for indexes and commodities. Subscription. http://www.pinnacledata.com/products.asp

Price Data. Historical intra-day OHLCV price data for US stocks and futures. Subscription. http://www.grainmarketresearch.com/gm.cfm

Pring, Martin. *The Investor's Guide to Active Asset Allocation*. McGraw-Hill. 2006.

Pring, Martin. Understanding the Business Cycle. http://www.pring.com/articles/article8.htm

Quotes Plus. End-of-day historical and daily update data for equities, mutual funds, exchange traded funds, commodities. Subscription. http://quotes-plus.com/joomla/index.php

Rimes. Coordinator and redistributors of fundamental data on companies. Subscription. http://www.rimes.com/Default.aspx?xml=mainmenu.xml&node=1.1

Standard and Poor's. GICS pages. Extensive information about categorization of stocks. http://www2.standardandpoors.com/portal/site/sp/en/us/page.topic/indices_gics/2,3,1,7,0,0,0,0,0,0,0,0,0,0,0,0,0.html

Standard and Poor's. Select Sector pages. Descriptions of the sectors and their component lists. http://www.sectorspdr.com

Stovall, Sam. *Sector Investing*. McGraw-Hill. 1996. Out of print.

Thomson Reuters. Extensive information about companies. Subscription. http://www.thomsonreuters.com/

US Bureau of Labor Statistics, an agency of the US Department of Labor. Historical economic fundamental data, much of it in both tabular and graphical format. Free. http://stats.bls.gov/

Worden. Publishers of the TeleChart 2007 trading system platform. The database includes fundamental data. Subscription. http://www.worden.com/default.aspx

Yahoo Finance. Historical end-of-day OHLCV price data for US and world stocks, general company news, current company fundamental data. Free. http://finance.yahoo.com/