Re-Thinking And Strategically Using Investment Factors

Marc Gerstein *3/19/18*

Whether by design or happenstance, the Fama French research led to a world in which factors, such as Value, Size, etc. are seen as being important in stock selection. That's not necessarily false, but it's drastically over-simplified.

No single factor ever has any meaning to an investor. That said, from time to time, various other things may line up such as to create the illusion that a single factor works. But the difference between what's real and can be relied upon in the future versus what is specific only to the exact historic time period studied is the intuition behind the use of the factor(s). We can't just ask: Did the factors work? We also have to ask: Should they have worked? Could one have expected them to work before having studied them? Are there logical financially sensible reasons why they can be expected to work? If we can't answer questions like these in the positive, then the best we can glean from empirical studies are time specific report cards that do not necessarily mean anything outside the confines of the sample period that was studied.

We also have to be clear on what we mean when we say a factor does or does not "work." It's not just about return. Risk counts too. And consideration of risk is a two-part inquiry. What level of risk can or should a particular investor assume based on his or her particular temperament, circumstances, etc.? And what is the market's attitude toward risk (sometimes, the market rewards risk-taking and other times, assumption of greater risk is punished). The latter does not mean every investor should keep changing his or her risk tolerance based on his or her assessment of current market conditions (though many can and do). It does, however, mean that an assessment of whether or not a strategy works depends not necessarily on how well it does but on how it does relative to what one has a right to expect given a particular market environment.

The goal here is to test core widely-recognized factors, not just individually but in combinations and to explain what we do and what we see in terms of why things make sense financially. And if we're going to consider sensibility, then we can't look at any factor until we logically justify doing so. So before looking at test results, I'll start from step one and derive each of the three factors I'll study from absolute inviolable financial logic combined with practical implementation considerations.

Valuing Financial Assets - The PROCESS Is Certain

When it comes to identifying genuine factors, we know, absolutely positively, how financial assets should be priced. The challenges we face are in how confidently or credibly we can come up with the necessary inputs.

Sometimes, it's easy. We know, for instance, how to value a $1.00 bill. There's no uncertainty. Hence we can be sure that if we try to sell a $1.00 bill for $1.25, we'll attract no offers while we would, if we cut the price to $0.75, be overwhelmed with bids.

Other assets are more challenging, such as a $1,000 bond that pays $50 a year in interest for the next 20 years and then matures (principal get paid back). We know the cash receipt dates. We know the amounts of the cash receipts. But we have to make assumptions about creditworthiness. That's usually do-able, but it can introduce an element of debate. And we have to make tougher assumptions about an appropriate rate of investment return. That, too, can be done with some degree of credibility; few today would suggest we assume a 25%

return, and few would propose a 0.0025% return. But even as we converge toward more reasonable assumptions, there will still be room for reasonable minds to differ than would be the case with valuing a $1.00 bill. Note, though, the nature of the debates about creditworthiness and required return. Nobody would disagree with the process. All the arguments would address the inputs.

Stocks are much harder because we can't say what cash we'll wind up getting back, when we'll get cash, and what the rate of return should be. Debate regarding inputs can be so intense and speculative as to make it appear as if stocks can't be reliably valued at all. But that's not so. We know, with 100% certainty, the method of valuation. For stocks, as for everything else, it's the present value of all expected future cash receipts. Because stocks don't have a formal maturity date or known resale date, we have to adapt to the infinite series, something that is accomplished by the Gordon Dividend Discount Model (DDM):

$P = D/(R - G)$

- Where,
  - D is the present dividend,
  - R is the required rate of return, and
  - G is the expected infinite Growth rate of the dividend.

Any factors we use, to be legitimate (i.e. to justify use in the unknowable future, which may or may not vary from what was observed in history-based study periods), must derive logically from that formulation.

Valuing Stocks - The Challenge Is In The Inputs

What makes for so much angst among stock investors is the reality that each of those required inputs is incredibly iffy at best. We can make plausible assumptions about dividend amounts and the timing of receipts (which is why there is something special about genuine income stocks) but many stocks pay no dividends, yet obviously are valuable. Also, most of us are clueless when it comes to what we'll get when we sell and when we'll sell. As for the required rate of return, that's comes off as a complete free-for-all guessing game.

But however difficult the inputs, we should never lose track of the process.

Re-Casting The DDM In A Useful Way

Before tackling the inputs, let's re-cast the DDM in a manner consistent with the way we know the market works in the real world.

We'll do that by substituting E, Earnings, for D, dividends.

A couple of generations ago, such a step might have seemed absurd. But our investment culture has, for better or worse (an interesting debate in and of itself) reached the point where investors act as if all corporate earnings accrue directly to shareholders but that shareholders willingly allow management to retain a lot of it (or all of it) for reinvestment in the business. If we can swallow that, we can re-cast the DDM as:

P = E/(R - G).

Having done this, we can use basic algebra to derive an ideal valuation formula:

P/E = 1/(R - G)

We can derive similar formulations for ideal P/S, P/B, P/CF, P/FCF, etc. but for the sake of clarity, I'll just use P/E. (By the way, academicians often like to use P/B and do so in a way that supports a descending sort; i.e, BM, or Book-to-Market. All the concepts discussed here, however, hold true for BM, although we'd need to makes some input and rhetorical adjustments.)

Variations aside, that P/E formulation pretty-much blows away conventional Fama French type orthodoxy: We now see that there is no logical reason to assume a low P/E stock is preferable to a high P/E stock. Any P/E, however low or high, is good or bad depending on how 1/(R - G) shakes out.

The algebra tells us that as G (a negative number in the denominator) rises, P/E rises (assuming all else is equal). That's what leads us to the by-now-well-known PEG ratio, which is on the right track in that it recognizes the relationship between P/E and growth, but is incomplete because it omits the R factor.

The algebra also tells us that as R (a positive number in the denominator) rises, P/E falls (assuming all else is equal). We see this in action all the time: General levels of interest rates are a big component of R (the rate of return) and everybody knows rising interest rates tend to depress stock values while falling rates tend to increase valuations. But there's more here than just interest rates. Higher risk investments require an additional premium. This is the company specific component of R. Assuming all else remains equal, as company risk rises, P/E falls and vice versa.

Coping With Inputs And Deriving Real-World Factors

Earnings

E is the easiest input factor to address. We can look up the latest reported earnings. And for many companies, we have estimates of near-term future earnings.

Required Return

R is a complex factor that involves considerations relating to the market as a whole and other considerations that are company-specific.

The Capital Asset Pricing Model (CAPM) is a good schematic for the way to analyze R. The formula is:

R = RF + (B * RP)

- where

  - RF is the risk-free interest rate,

  - RP is the expected risk premium that induces one to bypass zero risk and instead invest in a risky asset class (in this case, stocks), and

  - B, popularly known as Beta, is an indication of how risky an individual asset (i.e. a specific stock) is relative to its overall asset class (the stock market).

Note that as famous as this model is, academicians have since expanded it to express R in relation not just to the risk-free rate and the equity-market risk premium but also to other factors. Sometimes they add economic type factors. Many quants use fundamental factors, and that's the way a lot of academic research is organized:

R = A + (B1 * F1) + (B2 * F2) + (B3 * F3) . . . . etc.

- where

  - A is a constant (the famous "alpha" many crave - a degree of return separate and distinct from the operation of the ordinary factors)

  - F is a Factor. In the CAPM, there is only one factor, the market risk premium. Other fundamental factors added later include Value, Momentum, Size, and Quality. If you see the phrase "risk factor," you're dealing with one of these F inputs.

  - B is a company-specific measure of the sensitivity of the stock to the factor. Academicians refer to this as a "factor loading."

If you're statistically inclined and recognize this as a multiple regression formula, you're spot on. That's how Fama french and other come up with their factors. The regresses historically observable returns against a whole lot of things and identified those that met tests of statistical significance. And this is how they lost sight of how P/E, or value in general, really works (and how they also lost sight of how the other factors are supposed to work). Those studies are not designed to explore or capture the 1/(R - G) that makes ideal P/E what it is. And this, by the way, is how the nefarious practice of "data mining" penetrated and infected the field of equity research.

In order to keep it real, to work with Value as it's supposed to work, as the DDM tells us it must work, I'm going to eliminate the expansions and stick with the basic RF + (B * RP) formulation. If it was good enough for the Noble Committee to award a prize based on it (it was and they did), then it ought to be good enough for us.

Since I'm not forecasting the market, I'm not going to work with RF or RP. I'll just keep them in mind as I observe the world and recognize that:

(1) Changes in interest rates will impact the stocks, along with the rest of the market, through changes in RF, and that

(2) Changes in general investment community appetite for risk will influence the stocks in general (either market-wide or a subset of the market that includes the stock in question) through changes in RP. As an example of the latter, if investors feel optimistic and decide to

adopt a generally "risk on" stance, smaller companies may benefit from a decline in the RP (risk premium) investors demand for such issues. That could allow a small cap stock to rally simply because its small, even if there is no change in anything about the specific company.

Investors alway need to keep their eyes on RF and RP (and be aware of how different types of stocks may merit different RPs at different points in time) as a matter that is separate and apart from their view about individual companies.

The one thing with which we can actually work when we develop stock-selection strategies is B, the company-specific item. As with R in general, B has a positive sign within its portion of the denominator of the overall P/E faction. So as B rises, ideal P/E will shrink all else being equal. Conversely, as B falls, ideal P/E will rise.

Because the CAPM comes to us from academic quantitative finance, we should naturally expect a purely statistical input for B, Beta. And indeed, it is actually computed through regression, with B being the "factor loading" that applies to the RP "risk factor." Hence Beta's generally accepted definition as the measure of a stock's risk (volatility) relative to the market with 1.00 meaning equally risky, .75 meaning 75% as risky as the market, 1.12 meaning 12% more risky than the market, etc. (Actually, statisticians really define risk as statistical "volatility," but more often it's expressed as the square root of volatility,"standard deviation," in order to work with numbers that look apples-to-apples with returns.

As investment strategists/stock pickers who have access to lots of fundamental data, we're not bound to that definition of Beta, which was created by and for the convenience of mathematicians. There's nothing inherently wrong with convenience if we can have confidence that it can work for us. But in practice in this area, we cannot necessarily make that assumption. Traditional Beta, is a purely statistical "report card" on how a stock traded in the past (whatever measurement period was used for the computation). Maybe that is consistent with the inherent riskiness of the stock. Maybe it's not. Weird things happen in the world all the time.

In fact, observation shows that the Beta numbers often produce ridiculous results. If you look at enough individual cases, you'll see many in which frighteningly risky companies have very low, or possibly even negative Betas because there were extreme stock price movements that charted their own courses (in response to the ebb and flow of news and expectations) with little or no regard to general market action.

As a fundamental investor, I'd rather bypass statistical orthodoxy and focus on what risk really is.

I'm fine working with the notion of volatility. (Some say we should look only at downside volatility since nobody is averse to extreme gains. But that can be dangerously misleading. The company characteristics that cause gains to be extreme are the same as those that would cause losses to be comparably extreme should events unfold differently. We're better off looking at the inherent volatility/riskiness of the business, rather than locking in on the way events happened to unfold in a particular historic period and naively assuming they'll continue to pan out that same way in the future.)

So rather than counting on historical share volatility relative to the market to persist and be the only thing worthy of consideration, I'd prefer to add and/or substitute a hefty dose of data that supports assumptions about future profit volatility (and consequent future share price volatility). Those are the kinds of things picked up by the Portfolio123 "Basic: Quality" ranking system

(and other ranking systems one might create based on the quality theme). The higher the Quality rank, the more willing I am to presume lesser volatility and risk in the future.

So I'm going to use Q (the "Basic: Quality" ranking system) as a proxy for B, the only company-specific component of R, RF + (B*RM), with which I can work. But in order to model on Portfolio123, I also have to stay consistent in terms of the ascending/descending protocol. B is presented within R and CAPM in such a way that lower numbers suggest lower risk. But when I work with Q on Portfolio123, I have to assume higher numbers (higher ranks scores) equate to lower risk.

That means that all else being equal, fair P/E moves up or down based on increases or decreases in Q (Agin, since we're picking stocks as opposed to doing market forecasts, I'm assuming we'll only look at company factors).

This tells us something important: Many see Value (pursuit of low tallies for P/E and other valuation ratios) as a conservative strategy. That's 180-degrees wrong. This is not just a matter of the DDM-PE formulation. It's common sense and widely observable in all walks of life. Protection from risk is valuable and we pay for it. We do so with health insurance, auto insurance, homeowners insurance, premiums attached to hedging strategies, etc. If two stocks have identical growth expectations, you should expect to pay a premium for the risk protection afforded by the stock's ties to the more conservative business.

This provides us with a key strategic goal: Look for situations, anomalies, where we believe a stock is less risky than the level of risk that's "baked into" the stock's current valuation by Mr. Market. We can implement this through strategies that link high-value rank scores with high quality-rank scores. The interesting part is to test and discover how high we should go in each area. Do we need the best of the best in both V and Q? Or can we focus more on eliminating the worst of the worst in one or both respects, or some sort of combination. That's the kind of testing that is likely to be powerful for us.

Growth

In pure DDM terms, the assumed growth rate is supposed to be a growth rate stretching through infinity. That's a problem. Most of us have more than enough trouble projecting one quarter into the future. There is no hope that anyone anywhere can ever come up with a plausible defensible infinite-growth assumption.

That blows away pretty much every published Discounted Cash Flow (DCF) model you may see, since the infinite-growth input accounts for a large portion of the final answer. Hence my revised label for DCF: Discounted Cash Fluff. (For what it's worth, mathematicians tell us that the infinite growth rate has to be a very small number because we must assume that any company will decelerate into an advanced state of business maturity between now and infinity. Yeah right. What they really should be saying is that the infinite growth assumption needs to be a very small number to make darn sure we don't plug in a G that's higher than R, something that would produce a negative valuation.)

OK. So even if we can get comfortable with R, the infinite G conundrum nails down the fact that we absolutely positively cannot work with DDM or the P/E = 1/(R - G) formulation by plugging numbers into a spreadsheet. But that doesn't mean we abandon it. We use it instead as a strategic roadmap that tells us to look for anomalies, buy or sell opportunities, in which P/Es seem too low or high in light of future growth prospects.

So in working with the value factor, we won't actually plug in Growth numbers per se. We'll work with indications of growth expectations. We won't get hung up on infinity. We'll instead think in terms of (1) too optimistic, (2) reasonable, or (3) too low relative to what's "baked into" the stock's current valuation. So similar to the way we work with Q, we can model toward high valuation rank scores (or low ratios) compared with high scores in some sort of Growth rank. (and again, we need not always look for best of the best. We can explore and test elimination of the worst, etc.)

Many will, as a matter of first impression, assume we can and should use historical growth rates as inputs into DDM-type valuations and strategies, or a ranking system based on such factors. That might work (companies tend to evolve, rather than zig and zag day in and day out). But it's risky because it requires us to assume that historical trends will persist into the future. As often as we see it happening, we also know that things do change. (Can you remember the day when the paradigm for Apple switched from perennial super-growth powerhouse to "Oh my, can they really boost demand for iPhone X and when will they come up with the next new super-product?" In fact, Apple is one of a gazillion examples of the notion of corporate life cycle (as companies age, growth rates tend to shrink) and consider countless qualitative factors.

When it comes to future Growth, we can extrapolate from the past and if we're lucky (if enough companies are not shifting positions in their respective lifecycles), studies might actually produce good-looking results. But rather than count on just this, we can also consider clues that tap into the (often-qualitative) thinking of the investment community by looking at Sentiment (manifestations of what investors are saying about how they feel about a company's future) and/or Momentum (evidence of how investors actually acted in response to how they feel about a company's future.). These aren't specific G assumptions. They are, instead, proxies for great, pretty good, OK, lackadaisical or rotten and as long as we translate them into numbers (as we do with Sentiment and Momentum ranking systems or rules), we can work with them. But which proxy or set of proxies should we use? We can all come up with and test our own answers. Below are the results of my tests using the following:

- The Portfolio123 pre-set "Basic: Growth" ranking system
- The Portfolio123 pre-set "Basic: Sentiment" ranking system
- The Portfolio123 pre-set "Basic: Momentum" ranking system
- A 50-50 combination of the "Basic: Growth" and "Basic Sentiment" ranking systems
- A 50-50 combination of the "Basic: Growth" and "Basic Momentum" ranking systems
- A 50-50 combination of the "Basic: Momentum" and "Basic Sentiment" ranking systems
- A single-factor Portfolio123 ranking system using the mean projected LT EPS Growth rate

For each raking system I test, I'll use the PRussell 3000 Universe and set up screens that identify a particular ratings bucket:

- To screen for the Best bucket, I'll use Rating("xxx")>66
- To screen for the Middle bucket, I'll use Between(Rating("xxx"),34,66)
- To screen for the Best bucket, I'll use Rating("xxx")<34

For each screen, I did a ten year (2/20/08 - 2/20/18) backtest with 13-week rebalancing.

I also ran, for each screen, a rolling backtest covering that same 10-year period, assuming 13-week holding periods, and that new portfolios would start each week).

Here are the results of tests I did using various raking systems to serve as a proxy for the Growth input I need for the DDM formulation.

Growth - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best G	11.46%	19.52%	0.59%	3.75%	7.55%	-8.20%
Mid G	10.63%	20.73%	-0.47%	3.29%	7.62%	-8.27%
Worst G	8.92%	24.69%	-2.98%	3.17%	8.30%	-10.54%
Best - Worst	**2.54%**	**-5.17%**	**3.57%**	**0.58%**	**-0.75%**	**2.34%**

Sentiment - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best S	13.59%	20.45%	2.20%	3.73%	8.13%	-8.03%
Mid S	10.95%	20.85%	-0.20%	3.32%	7.8%	-8.64%
Worst S	6.45%	23.42%	-4.99%	2.67%	7.54%	-10.35%
Best - Worst	**7.14%**	**-2.97%**	**7.19%**	**1.06%**	**0.59%**	**2.32%**

Momentum - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best M	10.70%	18.32%	0.63%	2.79%	6.87%	-8.13%
Mid M	12.23%	19.90%	1.29%	3.42%	7.7%	-8.01%
Worst M	7.83%	27.84%	-4.70%	3.5%	8.88%	-10.86%
Best - Worst	**2.87%**	**-9.52%**	**5.33%**	**-0.71%**	**-2.01%**	**2.73%**

Sentiment/Growth Combination - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best S/G	13.05%	19.60%	2.02%	3.53%	7.79%	-7.84%
Mid S/G	11.54%	21.00%	0.29%	3.44%	7.86%	-8.36%
Worst S/G	6.67%	24.42%	-4.97%	2.74%	4.81%	-10.77%
Best - Worst	**6.38%**	**-4.82%**	**6.99%**	**0.79%**	**2.98%**	**2.93%**

Sentiment/Momentum Combination - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best S/M	12.93%	18.75%	2.31%	3.29%	7.44%	-7.83%
Mid S/M	11.04%	20.56%	0.05%	3.5%	7.64%	-8.33%
Worst S/M	6.96%	26.16%	-5.14%	3.07%	8.27%	-10.88%
Best - Worst	**5.97%**	**-7.41%**	**7.45%**	**0.22%**	**-0.83%**	**3.05%**

Momentum/Growth Combination - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best M/G	11.25%	18.18%	1.04%	2.99%	7.04%	-7.89%
Mid M/G	12.18%	20.37%	1.02%	3.38%	7.65%	-8.08%
Worst M/G	7.09%	27.12%	-5.13%	3.3%	8.64%	-11.04%
Best - Worst	**4.16%**	**-8.94%**	**6.17%**	**-0.31%**	**-1.60%**	**3.15%**

Long-Term EPS Growth Rate Projection - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best LTGr	11.05%	21.78%	-0.43%	3.44%	8.05%	-8.97%
Mid LTGr	11.89%	20.93%	0.51%	3.58%	7.95%	-8.18%
Worst LTGr	8.00%	22.40%	-3.07%	2.64%	7.3%	-9.91%
Best - Worst	**3.05%**	**-0.62%**	**2.64%**	**0.80%**	**0.75%**	**0.94%**

There's something to be said for all of these approaches. And, of course countless variations of all could be created many of which will be usable. As a choice for going forward with this study, I selected the "Basic: Sentiment" ranking system to serve as the proxy for G in the DDM-based strategic framework.

Testing the VSQ (Value-Sentiment-Growth) Framework

Again, I don't literally use the P= E/(R - G) or P/E = 1/(R - G) formulas. Instead, I use the framework as a logical basis for the proposition that an investment strategy should look for stocks that are mis-priced (i.e., stocks for which valuation is higher or lower than it deserves to be) based on Growth prospects (i.e., Sentiment) and/or Risk (i.e. based on Quality).

- I would aim to Buy stocks with valuations that deserve to be higher than they are because the data suggests Growth expectations (Sentiment) are favorable and/or because the data suggests Expected Future Beta/Risk is low (i.e., Quality is high).

- I would aim to Sell stocks with valuations that deserve to be lower than they are because of the data suggests Growth expectations (Sentiment) are weak and/or because the data suggests Expected Future Beta/Risk is high (i.e., Quality is low).

It's also important to remember what we're looking for in terms of test results. We may want the best of the best, but we should be aware that there is room in our strategic thinking to benefit of eliminating the worst of the worst.

The rolling tests are especially vital because of the way they allow us to separate up-market and down-market performance. This is important because we must always be aware of risk and market conditions. There cannot be any single strategy that will work for all kinds of investors at all times. Different investors' attitudes toward risk differ and the market rewards or punishes increased risk-taking at different times.

Testing Stylistic Purity

We'll start by looking at at each of the three ranking systems, Basic: Value, Basic: Sentiment, and Basic: Quality, in isolation, as is typically done in Fama French type factor studies.

We'll start with the Value factor:

Value Only - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best V	13.70%	24.73%	1.27%	4.22%	9.15%	-8.93%
Mid V	9.48%	21.12%	-1.57%	3.06%	7.52%	-8.86%
Worst V	7.87%	19.59%	-2.35%	2.43%	6.79%	-9.23%
Best - Worst	**5.83%**	**5.14%**	**3.62%**	**1.79%**	**2.36%**	**0.30%**

At first glance, the famous Value factor seems to work. Annual return and alpha are best for the best-ranked group and worst for the worst-ranked group. Volatility is high for the higher ranked stocks, but that's OK: The alpha shows we get more than enough return to compensate for the extra standard deviation. Besides, higher volatility is good when it means gains and the association of higher returns with higher standard deviation suggests we're getting that good combination.

This, however, is an example of why the rolling tests are so important. They can help us see when good results are influenced by good luck in having selected the right start-to-end period. And considering we've had such a prolonged bull market, it's easy to "stumble" into a favorable start-to-end test period.

The rolling tests here show us that value did, indeed, work well during up-market periods, but was trivially useful at best when times were tough. This confirms that Value is not a

conservative strategy. The Value table showed that whatever we may have gained in margin of safety (the Holy Grail sought by traditional advocates of vale as a conservative strategy) etc. was eaten away by the results of being invested often in bad companies during bad periods.

Now let's look at Sentiment.

Sentiment Only - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best S	13.59%	20.45%	2.20%	3.73%	8.13%	-8.03%
Mid S	10.95%	20.85%	-0.20%	3.32%	7.8%	-8.64%
Worst S	6.45%	23.42%	-4.99%	2.67%	7.54%	-10.35%
Best - Worst	**7.14%**	**-2.97%**	**7.19%**	**1.06%**	**0.59%**	**2.32%**

That's impressive generally but especially so when we look at the rolling tests.

The latter suggest that when times are tough, we shouldn't run for cover with value as many suggest (see above); go, instead, with what the Street likes, or at least avoid what the Street hates. Interestingly, though, when times are good, the growth proxy is not a big deal; a minor benefit. Apparently, when we're in a bull market, growth expectations tend to be favorable across many stocks and it can be hard to separate the wheat from the chaff.

The next table confirms the appropriateness of using the Quality rank as a risk indicator.

Quality Only - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best Q	12.17%	18.62%	1.65%	3.42%	7.36%	-7.09%
Mid Q	11.69%	21.09%	0.35%	3.48%	7.93%	-8.38%
Worst Q	7.11%	25.33%	-4.83%	2.81%	8.17%	-11.51%
Best - Worst	**5.06%**	**-6.71%**	**6.48%**	**0.61%**	**-0.81%**	**4.42%**

The basic tests show that the standard deviations line up impressively along with annual return and alpha. But the rolling test is much more interesting. In bull markets, we get punished a bit for conservatism. It's in the down markets where "risk off" pays. This is consistent with what we expect.

By the way, a word about the so-called size factor: It's a branch of the Quality factor. Size means economies of scale (better coverage of fixed costs) and typically better business portfolio diversification (even within supposedly single-industry companies which still show a lot of variation in terms of specific products and services, types of customer/clients, regional exposure, etc.). These translate to enhanced profit stability which tends to translate into enhanced stock stability. In terms of P/E = 1/ (R - G) strategizing, size can be used in addition to or in lieu of Q, and use of Q alone will give you the risk benefits similar to that of a large-cap approach; i.e., Q means big companies and/or smaller companies that act/feel big.

Combining Two-Style Approaches

Logic tells us two factors should be better. If we look at P/E alone, there may be and often are enough weak decision makers out there (those who don't really look at their numbers) to allow 1/(R - G) to line up in such a way as to make it look as if P/E alone is working. But why rely on luck. Why not try to boost our probabilities of success looking for stocks that are good in all respects.

Besides, not only does that comport with common sense and the basic P/E = 1/(R - G) logic, double filtering also reduces the number of stocks in each bucket and brings us closer to an investable result set. The first round of single-style tests divided the PRussell 300 universe into three buckets of approximately 1,000 stocks each. Now, each of those buckets will be divided into three second-level buckets based on a second ranking system with each bucket now holding about 333 stocks.

We'll start our two-style inquiry by looking at Value-Growth, or as we implement it, Value Sentiment. I starts out looking for favorably valued stocks; i.e. those with generally low ratios. But I then go further and focus on growth prospects (through the indirect Sentiment proxy) by distinguishing the best third (the ones for which expectations are most optimistic), a mainstream middle group, and the potential dogs (the bottom third in which value ratios are probably low because they deserve to be low, i.e. because growth prospects are least favorable). What I've done here is modify the standard value factor by using the 1/(R - G) concept to identify bullish anomalies (situations where valuation ratios are low despite favorable expectations regarding growth) and also to identify probable probable value traps.

Value-Sentiment- Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best V/Best S	18.09%	23.86%	5.44%	4.6%	9.22%	-7.72%
Mid V/Mid S	10.22%	20.07%	-0.51%	3.07%	7.39%	-8.48%
Worst V/ Worst S	3.75%	20.28%	-6.33%	1.64%	6.20%	-10.54%
Best - Worst	**14.34%**	**3.58%**	**11.77%**	**2.96%**	**3.02%**	**2.82%**

The ability to make this distinction is an important element of what separates those who succeed at value investing from those who don't.

So far, we've ignored R, required return. Let's look now at how that combines with Value. This is a muddy item since, as we saw with the CAPM, it has a lot of it has to do with the market and not with characteristics of the individual company. I'm strategizing only based on the latter since I want to separate market forecasting from stock selection (stocks I expect to be better selections in whatever kind of market we get).

For this round of testing, I'm going to again start with a favorably valued (best one-third) subset of the PRussell 3000 universe and then run a second set of filters to distinguish between equally-sized higher-quality (least risky), middle, and lower quality (riskiest) buckets.

Value-Quality - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best V/Best Q	13.15%	22.22%	1.44%	3.85%	8.24%	-7.84%
Mid V/Mid Q	10.99%	19.79%	0.26%	3.19%	7.38%	-7.98%
Worst V/ Worst Q	5.60%	23.97%	-5.44%	2.15%	7.25%	-11.47%
Best - Worst	**7.55%**	**-1.75%**	**6.88%**	**1.70%**	**0.99%**	**3.63%**

We see that generally, this VQ approach OK, as DDM theory suggests. There is the potential to generate some Alpha by looking for stocks with low relative valuation ratios for which the market has misjudged the relative level of company risk. But the differences are not as consistent and as stark as what we saw with VS. That, too, makes sense. The link between changes in G and changes in P/E appear much more clean and direct than the link between changes in Q and changes in P/E; the latter are watered down by the roles played by RF and RP.

There's also the return-risk continuum itself. When we choose to reduce risk or increase quality, we aren't necessarily looking to boost returns. More likely we're looking to diminish risk, and we tend to do so, if not as a permanent choice, than if/when we expect market weakness.

We see the impact of this watering down in the 10-year standard deviation, which is not the lowest one we see. But the rolling down-market results show that we are, indeed, getting the downside benefit we seek when times are tough. This reminds us that there really is no such thing as a single best-for-all-times best-for-all-investors strategy. Risk profile is an important driver of any investment decision.

Now, before going on to combine all the VSQ elements, let's look at the remaining two-headed strategy: QS (Quality Sentiment). There isn't necessarily a theoretical compulsion to do so: To ignore value (the relationship between how a stock is priced and what it's worth based on the things that drive PE) seems so . . . well, you know ("wrong?"). But there is the Shiller-Lee Value-Noise thing about which I've written. (Recall that we can't ever really expect to see Price equal to Value. The messy real world is such that P = V + N, price equals value plus noise (rather than P = V.)

Favorable sentiment is as good a way as any to represent potential increase bullish noise. We saw at the beginning that a pure sentiment strategy can work on its own, and we saw too with VS that Sentiment and Value make for a nice combination. Let's now see if we can add something to Sentiment-Noise by supplementing it with Quality (Risk) only.

Sentiment-Quality Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	**Stan. Dev. %**	**Annual Alpha %**	**All Markets**	**Up Markets**	**Down Markets**
Best S/Best Q	12.74%	18.60%	2.15%	3.53%	7.52%	-7.12%
Mid S/Mid Q	11.11%	20.50%	0.11%	3.49%	7.9%	-8.27%
Worst S/ Worst Q	1.74%	26.52%	-9.77%	1.84%	8.29%	-11.04%
Best - Worst	**11.00%**	**-7.92%**	**11.92%**	**1.69%**	**-0.77%**	**3.92%**

The short answer is that this can be productive. But as with VQ, we must recognize the way in which Q waters down and pulls in a direction opposite that of pure S. As with VQ, we see Q adding a healthy does of downside protection to S. But the up-market rolling results are intriguing: If you've ever had a feeling that bull markets are friendly to genuine garbage, well, let's say we see here some empirical support for that idea. Or we could express it another way: When the bull is raging, it may pay to buy hated stocks because sooner or later the bears will throw in the towel.

Anyway, here again we see there is no good-for-all-seasons good-for-all-investors answer. If you want to eliminate value and do noise only, it may pay, if you feel nervous ignoring value, to temper your approach with a bit of Quality.

The Trifecta — The Full-Blown VSQ Combination

I implement the three-prong test through the following steps.

1. I sort the PRussell 3000 universe and separate it into three equally-sized buckets (each having approximately 1,000 stocks) based on the Portfolio123 Basic: Value ranking system.

2. I sort each 1,000 value bucket into three equally sized secondary buckets (each having approximately 333 stocks) based on the Portfolio123 Basic: Sentiment ranking system.

3. Finally, I sort each 333 value-sentiment bucket into three equally sized tertiary buckets (each having approximately 111 stocks) based on the Portfolio123 Basic: Quality ranking system.

4. Next I run basic 10-year backtests and rolling tests (13-week holding periods; each rolling test starts 1 week after the previous one) for each of the 27 VSQ groups.

5. The table below shows test results for the following 111 stock buckets

    A. Best Value - Best Sentiment - Best Quality
    B. Mid Value - Mid Sentiment - Mid Quality
    C. Worst Value - Worst Sentiment - Worst Quality

Statistically speaking, it would be nice to see Bucket A always being better than Bucket B, and Bucket B being better than Bucket C.

Realistically, though, we have to recognize that Quality is apt to pull in the opposite direction to Value and Sentiment, emphasizing risk reduction rather than return maximization. And we know to expect different things in up and down markets; when the bull is stampeding, risk reduction is likely to feel like a bucket of cold water poured on one's head, rather than a boost.

So here, now, are the VSQ results.

Value-Sentiment-Quality - Russell 3000

	Start to Finish 10-Year Test			10-Year Rolling Tests (new 13 wk. tests starts every week) Average of all sets of 13-week returns		
	Annual Return %	Stan. Dev. %	Annual Alpha %	All Markets	Up Markets	Down Markets
Best V/Best S/Best Q	15.10%	22.47%	3.18%	4.13%	8.46%	-7.41%
Mid V/Mid S/ Mid Q	9.69%	19.41%	-0.67%	2.91%	6.92%	-7.8%
Worst V/ Worst S/ WorstQ	2.01%	23.70%	-8.56%	1.18%	8.14%	-11.69%
Best - Worst	**13.09%**	**-1.23%**	**11.74%**	**2.95%**	**0.32%**	**4.28%**

Generally speaking Best-Best-Best is the way to go. But it's in the down markets where the VSQ Trifecta really shines, as Quality softens the impact of the bear.

When the market is rising, we can can make good cases for going true to form (best-best-best) or playing a contrarian risk-tolerant game (worst-worst-worst). Or, if we're generally bullish, we can go VS and let a good market take us off the hook for ignoring Q.

Implications For Strategy Design

The easiest way to guide yourself toward a potentially successful strategy is to thoughtfully address V, S and Q.

These studies were pretty-much binary with regard to all three. Either a style was considered (on) or ignored (off). Actual investable strategies needn't do that. It's perfectly proper to decide you want a lot of this less of that, etc. You can use a single-style ranking system to sort a universe pre-filtered based on one or two other styles. You can use a two-style ranking system (where each style may be equally or unequal as you choose) and a set of eligibility filters. Or you can use a single 3-style ranking system with or without any filtering beyond such basics as liquidity. (And of course you can come up with your own ways to express V, S and Q, and even come up with a different growth proxy). The possible ways you can strategize are endless.

But whatever you do, it's always a good idea to think through how your choices fit within a VSQ (or VGQ) framework. If you do that, you're much more likely to wind up with a model that is capable of giving you satisfying live performance. It's also likely to spotlight the return-risk tradeoffs and induce you to make choices that are most consistent with your situation and your views of the market.

Also, awareness of the respective impacts of V, S/G and Q will free you from the needless complexities of wondering what's in and what's out.

For example, Value has to work. Saying it doesn't work makes no more sense than paying $1.25 for a $1.00 bill. So let's suppose you have a Value model that's been underperforming. If you can accept that Value absolutely positively must work at all times, you'll find it easier to identify constructive steps that can address your situation.

- Did you proxy for and address G as well as you could have?
- Is there something beyond tour control that is causing G expectations to deflate across the board? (An economic slowdown, perhaps?)
- Or is it that the interest-rate component of R (or the risk premium for equities) that is rising and deflating value across the board?
    - Is the market right to re-value stocks?
    - Or is the market overestimating the impact of changes in RF and under-estimating prospective G?

Isn't it better to think in terms of questions such as these, and make thoughtful decisions to change course or be patient, than to throw your hands up and passively say "Oh darn, Value is out of favor."

There is no easy answer; if there was one, anybody who took a statistics class and has a spreadsheet could become rich enough to buy their own country. It's not just about studying factors or even combinations. It's about understanding how each factor influences stock prices, how they interact with one another to exert combined influence, and understanding what messages are being delivered when you observe how factors are performing.