The Golden Age of Quant

Eric H. Sorensen, Ph.D.

August 2019
The Circles of Quant

The veteran reporter asked during a recent live news telecast: “So, Eric, is the Golden Age of Quant Over? Everybody has the same data, same models...” My response was (and is) in essence, “Not at all, in fact it’s just beginning.” During the interview, I elaborated by drawing (in the air with my right hand), the “Three Circles of Quant” (discussed below and referenced throughout this treatise).

The ‘Golden Age’ question stirred mental ashes at that moment. My mind transported to the article published a decade earlier regarding “the future of active equity management.” It was the aftermath of the last financial crisis, and the dramatic decline in equity values that followed, combined with some discontent across quantitative strategies (Sorensen, 2009). We highlighted two requirements to a healthy quantitative investing landscape. Managers needed to continue embracing (1) “an innovative and diverse a set of approaches across the active investment arena”; and (2) the necessity to “differ from each other in the ways they access and leverage information in terms of their information sources, information processing and information implementation.”

On the one hand, we noted that “Catchy yet misleading phrases and fads ... will also come and go, with the next batch of marketing-driven buzz.” This manifested in the subsequent decade with labels such as “big beta”, “smart beta”, and “machine learning” to name a few. On the other hand, we saw it imperative that we persist in one of the primary motivations driving quantitative solutions: Discovery and the pleasure of accomplishing enlightening empirical research. Dollars are the other motivation, often achieved by copying/marketing the discovery. Taken to an all too often extreme, however, dollars converging on discovery equals crowding. And crowding ultimately fosters capitulation, as in 2008-09.¹

Figure 1: The Permanent Yet Dynamic Circles of Quant

In Figure 1, the three circles are inherently emblematic and are essential foundations of true systematic quantitative investing. They have always existed as elements of success, although their tri-part magnitude and relative importance varies as strategies evolve. In 1986 and 1987, we developed the E-model at Salomon Brothers. Ours was the first U.S. Equity factor model to be deployed with S&P 500 securities as an indexation alternative by several U.S. pension funds (Sorensen, 2009). In 1986, we had all three circles, albeit with fewer pixels.

The Southern Circle Domain in financial markets at that time (and following) was conducive to asset price corrections to fair value. There were few exogenous government influences to distort relative prices. Free market stock investors methodically drove relative prices to converge to value, in contrast to the
government moving whole markets with politically motivated economic shocks or liquidity lunges. (This is in stark contrast to the 2008–2018 period when unprecedented central governmental intervention distorted natural capital market-driven risk and return relationships.)

The Western Circle Data in the 1980s was anything but massive. The only data available in machine-readability were corporate financial information provided by Compustat®, and later, World Scope®.

The Eastern Circle Digital largely comprised multiple regression and dividend (earnings) discount models. From 1977 to 1987, this author published dozens of academic journal articles using this machinery in research projects that required modeled estimates of financial asset prices. These regression-based factor models amounted to “machine learning” before machine learning got its new name. The E-model was an early expert system that refined and estimated a John Burr Williams valuation proposition (J.B. Williams, 1938). The model solved for fair value as a function of normalized earnings, growth expectations and risk “factors” for firm-specific characteristics. For example, earnings volatility and balance sheet leverage are to relative P/E for a corporate stock as credit rating is to yield for a corporate bond — more risk means lower value (price), ceteris paribus. By use of cross-sectional regression, the factors enabled estimates for adjusting the model’s discount rate in a present value of future earnings formulation. The model literally solved for fair value — what price should converge to in equilibrium in a reasonable time.

The E-model was the first econometric factor model deployed with real institutional pension dollars invested in the broad market (the S&P 500 list of stocks) seeking to outperform the “passive” index. It succeeded. Beginning in 1987 and continuing into the mid-1990s, several of Salomon Brothers corporate clients benefited by portfolios weighted according to the “E-Model Tilt.” The model consistently beat the cap-weighted S&P 500 index.

Credit to Financial Economists

Was this the Golden Age of Quant? No. Now it is just leaving the adolescent stage. The 1980s were, rather, the embryonic age of applied quant. The 1970s were the Golden Age of theoretical quant born of the path-breaking research by financial economists and academicians. Prior to the 1970s, formal finance curriculum was largely devoted to two questions: 1) What are the relevant accounting identities for corporate financial decision making? and 2) What does a financial institution (bank) do?

Then it began. A rich legacy of financial markets research at universities commenced to shape and influence the evolution of modern portfolio management. These advances were not dissimilar to modernization achievements in other academic disciplines, such as scientists’ contributions to medicine, communication, digital technology or space travel. We often summarize the following major research themes as: 1) You should earn returns consistent with risk taken, a.k.a. Asset Pricing Equilibria; 2) You should earn no more than Expected Returns, a.k.a. Asset Pricing Competition; 3) But, should You beat the historic annualized 9.3% Lorie and Fisher finding, it must be an anomaly, (a.k.a. Asset Pricing Behavior), 4). You should use modern machinery to model future price behavior, a.k.a. Asset Pricing Econometrics (Quant).
The early pricing research was descriptive. As the financial markets matured past the 1970s, the descriptive became normative. The Capital Asset Pricing Model (CAPM) and optimal portfolio theory with the separation of alpha and beta became institutionalized in dramatic fashion. Now, forty-five years later, the price weighted index remains a feature story. Flows into the “optimal portfolio” still served to inhibit and stymie the creative potential for material diversity for active managers. Around the time of the E-Model Tilt, quantitative long-only equity took on a new name, index enhanced strategies, with highly successful firms such as Well Fargo (later BGI, now Blackrock) and State Street Investors. As “active quant” grew, so did the commonality of portfolios. Quantitative equity began to follow an all-too-recurrent pattern that has historically plagued other successful strategies. The plague has a pattern: creative brings on copied. In time, copied brings crowded, finally leading to illiquid capitulation. During 2006–2007, U.S. equity quants found themselves in this trap — common brains, benchmarks and betas leading to common bets — ultimately needing to be commonly unwound. Ouch.

However, such challenges lead to opportunity for the creative practitioners. Figure 2 is one version of a sequence of creative advancement. (The circles of Figure 1 are constantly moving and reshaping and always present.)

Over this evolution, we witness the ballooning of the left circle (data) and the deepening of the right circle (digital). In the middle of Figure 2, we see “big data” and “smart beta.” These are great examples of evolution — the survival toil of quant.

Our observations regarding data, on the one hand, are: 1) “big” alone is not a sufficient condition for added value — “big” but not “smart” (or causal) is potentially spurious; and 2) “big” brings perverse results if everyone uses it. On the other hand, “smart” data may be sufficient. Smart data, among other things, is not commercially ubiquitous, is given to reasonable investment horizons and is rich in fundamental intuition. We recommend the terms “big beta” (meaning excessive price weighted portfolios) and “smart data” (connoting more than just big).

Figure 2: Survival of the Quants

Shown for illustrative purposes only. Source: PanAgora; presentation at the Sagard Holding Ecosystem Conference, Eric H Sorensen, June 12, 2019
The Western Circle: Data — Today’s Alternatives

Consider Figure 3. There lie the keys to quality data. To master today’s western circle, the figure points to several key data requirements: 1) The data have to have greater potential in areas that have higher price dispersion, such as biotech and pharma (top bar in Figure 3 vs. utilities at bottom); 2) relevant corporate financial data are increasingly in the public domain, rendering them lessor predictive power today; 3) useful data should be proprietary, fundamentally intuitive or causal and have a reasonable horizon for the prediction (months, not days, in most cases); and 4) the prediction should not relate directly to price change, but to fundamentals that will influence future price change.

Figure 4 illustrates examples of the newer data required — with advanced specifications — to compete in today’s investment arena. This is driven by the evolution of asset category in the valuation of today’s firms. We estimate that today, market capitalization of the average firm is split between approximately 70% intangible, vs. 30% tangible, in contrast to the reverse (30/70) from the 1970s. An example from Figure 4 is modeling value with P/E ratio now replaced by proprietary modeling items like patents and balance sheet asset values.
With focused effort and people power we can implement true data science. This not only requires skilled computational and systems strengths, but also domain knowledge dictating where to look. Figure 5 presents some web-based results from researchers at Rice University. They examined the top 30 professional investor users of the EDGAR (SEC Filings) data system. PanAgora has been using alternative data for over a decade and was #2 in total downloads since 2003 and #1 in the diversity of filings sourced.

### Figure 5: Big Data Revolution

**Top 30 Users of EDGAR since 2003**

The table reports download statistics by form type and end-of-sample assets under management for the top 30 hedge fund users by total downloads (distinct within a month).

<table>
<thead>
<tr>
<th>Firm Name</th>
<th>Total Downloads</th>
<th>10-K/10-Q</th>
<th>8-K</th>
<th>13F</th>
<th>13D</th>
<th>13G</th>
<th>AUM (MM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Renaissance Technologies Corp.</td>
<td>4,016,439</td>
<td>3.5%</td>
<td>2.0%</td>
<td>85.5%</td>
<td>0.2%</td>
<td>0.2%</td>
<td>0.9%</td>
</tr>
<tr>
<td>PanAgora Asset Management Inc.</td>
<td>3,969,668</td>
<td>24.2%</td>
<td>21.6%</td>
<td>10.9%</td>
<td>4.9%</td>
<td>15.0%</td>
<td>7.2%</td>
</tr>
<tr>
<td>Blackrock</td>
<td>3,704,596</td>
<td>3.5%</td>
<td>0.7%</td>
<td>92.4%</td>
<td>0.1%</td>
<td>0.1%</td>
<td>0.2%</td>
</tr>
<tr>
<td>Hutchin Hill Capital, LP</td>
<td>3,044,555</td>
<td>3.7%</td>
<td>4.7%</td>
<td>8.8%</td>
<td>0.0%</td>
<td>0.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Tradenx Inc.</td>
<td>2,068,017</td>
<td>11.2%</td>
<td>5.6%</td>
<td>41.4%</td>
<td>5.0%</td>
<td>11.0%</td>
<td>15.2%</td>
</tr>
</tbody>
</table>

*Source: Crane, Crotty and Umar, (2018)*

You can also buy big data. A recent study by Citigroup is very enlightening regarding the business of “big” data. And big it is. It is big in dollars and big in space. Much of it must be stored and processed in the Cloud. The providers gather and aggregate the data from broad sources, such as satellite imagery, credit cards, and clicks, to name a few. Rumor has it that the number of self-identified big data firms total in the hundreds worldwide and reportedly generate an aggregate billions of dollars in revenue annually. We caution, however, there are caveats to purchasing from aggregators: 1) if it is costly, the portfolio results may not pay net of expense; and 2) if it is readily available, usage can drive down any relative advantage arbitrage profits.

**The Eastern Circle: Digital—Today’s Alternatives**

Artificial intelligence (AI) and machine learning (ML) are common labels in today’s quant arena, just as they are in a variety of other domains and endeavors. Their acceptance as modern portfolio tools represents major advancements from the E-model days. In earlier times, pejorative accusations of “data mining” and “black box” were damming. Now they are darling and often laudatory. For example, large institutions, such as BlackRock, are embracing the buzz and undertaking a “… venture in AI [artificial intelligence], a whole group of people working on developing computer-based investing. And that’s truly a computer saying, Buy this. Sell that.”

Consider Alan Turing 1950’s definition of AI. Basically, if the computer can achieve (or otherwise predict) in a way that is indistinguishable from a human, it is AI. Many applications of good old-fashioned AI (GOFAI) indeed conform to the definition. In 1962, Arthur Samuel programmed a computer to play...
checkers that began to win championships. We have other examples in human-equivalence in chess, driverless cars, facial recognition and so on. These GOFAI examples operate in applications with two conditions that make them significantly different than the stock market price, as in the BlackRock quotation above. These applications operate in a world of finite outcomes (choices); and a world of rules.

Stock prices follow Brownian motion. There are infinite outcomes, and the path to the outcomes is mostly stochastic. Basically, price has no memory and no rules. AI using modern machine learning to correlate hordes of input factors (many that are not intuitive) to explain price movements is not functional, but rather folly.

However, machine learning techniques can be quite useful in investing, as are multiple linear regression as well as classification and regression trees. These are “machine learned” estimators, along with other classical statistical approaches, such as linear and nonlinear, univariate and multivariate.

In years past, undergraduate courses were called statistics and graduate courses were called econometrics. Today, we have classes in machine learning. In fairness, there is a distinction with ML in contrast to classical econometrics. They invoke computationally intensive recursive algorithms with the goal that the machine seeks to alter (improve) the predictive model again and again without any redesign by a human. The learning seeks to minimize the forecast error with an objective similar to maximum likelihood. This is the differentiator compared to, say, multiple regression, that arrives at one set of parameter estimates and subsequently requires the designer (econometrician) to revise the input specifications and check for an improved “goodness of fit” statistic. Classical econometrics are therefore structured, or supervised, in the machine learning lingo, in that the data vectors have labels for data items that vary in value across observations.

ML can be structured (supervised) or unstructured (unsupervised). With unstructured algorithms, there are no prior specifications of relationships and no assumptions; just let the data seek and speak. For examples, K-means, a popular unsupervised machine learning method, guesses and checks for error minimization, then guesses and guesses and guesses again. The downside of this method is often illogical clusters with classification associations that may well be spurious, and therefore not predictive as time passes. Another popular class of machine learning approaches, neural nets, have come in and out of favor over the past 70 years, transitioning from neural science in earlier years to computer science more recently. Interestingly, the recent boom in the computer game industry has driven a renewed interest in neural nets.

Deep learning models are multi-layer neural nets where the intermediate artificial neuron layers are used for feature transformation and extraction, with or without recursive feedback loops. Deep learning became possible with the recent advent of powerful computing devices (e.g., GPUs) and abundant and readily available data, but they need to be applied judiciously to be value additive. When applied lacking intuition, they can be hard to interpret, over-fitted and render spurious findings.

In addition to K-means and neural/deep networks, there are dozens of modeling approaches that fall under the ML label. An example is Random Forests. Many years ago, we estimated classification and regression trees (CART) for a variety of investment predictions (Sorensen and Miller, 2000). The tree is structured and exposed to a data set that might
be used in a linear regression. In contrast to regression, the model literally builds an econometric decision tree with a hierarchy of “if-then” rules. Human intervention is necessary to re-estimate the tree seeking improved predictions. In the essence of ML, software from Random Forests creates hundreds of trees with randomized inputs. The intense computing can produce over 100 trees with hundreds of exogenous variables and thousands of observations in less than a second. The goal is recursive error minimization by converging on a consensus of results across myriad trees.

Figure 6 has the two broad areas of ML that we find useful in quant investing. In addition to models that extend classical statistics, a key area that shows promise is natural language processing (NLP). NLP recognizes patterns in language often using word-embedded techniques that translate words and phrases into high-dimensional vector space. This digitalization of language enables the machine to cluster, categorize and comprehend vast amounts of information.

**Figure 6: Machine Learning for Investing: Advantages**

- **Statistical Forecasting**
  - Regression – Linear and Logistic
  - Artificial Neural Nets
  - Classification Trees
  - Random Forests
  - And many others

- **Natural Language Processing (NLP)**
  - Applies algorithms to identify/extract natural language
  - Abides by rules for unstructured language data
  - Converts words into “Word to Vec,” a format that computers understand
  - It’s been used in linguistics for years
  - Google Search algorithms are an example

Until a few years ago, linguists were the primary users of NPL apparatus. In addition, until recently, *Bag of Words* was the common application counting positive and negative words. Recently there has been tremendous advancement in translation applications. For example, Google has pioneered applications for association and classification for business purposes.

PanAgora’s research and portfolio teams have deployed a variety of NLP applications to generate return signals. Figure 7 illustrates one of many examples. Shown here is a project to classify and categorize stocks using NLP techniques. There are studies that document the return enhancement potential through rotating portfolios with industry and sector commonalities over time, (Sorensen and Burke, 1986); (Moskowitz and Grinblatt, 1999). These have been done with standard GICS data. A second approach is to classify pseudo industries using time-series econometrics. Stocks that trade with higher correlation are candidates for same-cluster categorization. A third approach illustrated here represents that Amazon, Google and Apple are more closely associated than the GICS would define. How do we know? We know because the NLP apparatus machine reading thousands of media and corporate reports reveals that they are often discussed in close proximity. Among other applications, this is a forecast of future changes in GICS.

For illustrative purposes only. Source: PanAgora
Figure 7: Machine Learning: Momentum

- Standard industry classifications become stale – arbitrage
- Machine learned clusters anticipate changes in GICS

![GICS 2550: Consumer Disc. – Retailing](source)

<table>
<thead>
<tr>
<th>GICS 2550: Consumer Disc. – Retailing</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Reilly Automotive</td>
</tr>
<tr>
<td>Home Depot</td>
</tr>
<tr>
<td>Tiffany &amp; Co</td>
</tr>
<tr>
<td>GAP</td>
</tr>
</tbody>
</table>

![GICS 4510: I.T.- Software & Services](source)

<table>
<thead>
<tr>
<th>GICS 4510: I.T.- Software &amp; Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mastercard</td>
</tr>
<tr>
<td>Visa</td>
</tr>
</tbody>
</table>

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
</tr>
<tr>
<td>Netflix</td>
</tr>
<tr>
<td>Facebook</td>
</tr>
<tr>
<td>Google</td>
</tr>
<tr>
<td>Salesforce</td>
</tr>
<tr>
<td>Microsoft</td>
</tr>
</tbody>
</table>

ML Cluster

Source: Shown for illustrative purposes; PanAgora, Chen and Li (2018)

Figure 8 (see next page) presents a particularly challenging task for NLP. The goal is quantifying investor sentiment toward A-share stocks in China. The challenges are as follows: 1) Chinese is not an alphabet language; 2) there are 50,000 pictorial characters (and growing); 3) there are no spaces between words; and 4) words are understood by the context of a set of adjacent characters (Chen, Lee, and Mussalli, 2019).

With respect the stock market, 80% of the volume is driven by retail investors (as opposed to institutional), and it is a *hype* market. Sentiment is the driver, and sentiment views are available on blogs. In China, sentiment expressions in blogs are hampered by considerable government censorship.

Embedded word algorithms and neural nets enable sentiment detection in “reading” the blogs. Words are converted to vectors, taking order into consideration and model identification of a cohort of slang pairings, (Chen, Lee, and Mussalli, 2019). The interesting part is that Chinese is a tonal language. Spoken words with radically different meanings can sound highly similar. Due to censorship, for example, the blogger that does not like a stock may choose character(s) for “spicy chicken” because it is phonetically associated with the character for “rubbish” or “garbage.” Chen, Lee, and Mussalli also point out that an avid A-share blogger may write that a firm is “spicy chicken” to communicate it is really “rubbish.” This is a way to obviate the government impulses to encourage (censorship). The character for “river crab” looks like “harmony,” a government euphemism for you better fall in line and conform. This application is a semi-supervised approach to identify slang usage on blogs across hundreds of opinion platforms.
Figure 8: NLP Example: A-Share Investor Blogs

- Chinese stock forums, heavy retail trading and strong government censorship allow a model for investor sentiments.

Contrasted with western languages, Chinese languages have many unique challenges:

- Different syntactic structure
- Segmentation: no word space
- Vocabulary: Larger and evolving
- 50,000 characters

Quant, in essence, is an expert system. One of the last bastions of fundamental pickers was the “CEO interrogation.” It amounts to face-to-face interviews to examine body language, evasive speech, tone and so on. Now our quants have the tools to do it even better (this is real AI). Figure 9, from Yu, Chen, and Lee, 2019, lists some of the metrics we now use to gauge management. (For years we have ranked managements with quant metrics using profiles and board compositions, compensations and the like.) We now use machines (using ML and NLP) to “detect deception,” “calibrate optimism” and “monitor realism.” (We are working on “body language” next.)

Figure 9: Executives Talk Too Much or Too Little

- NLP Data for Corporate Communications
  - Corporate filings including footnotes (SEC)
  - Transcripts of Events (Analysts’ Calls) – FactSet, Capital IQ [S&P]
  - Press releases and articles

- NLP Applications
  - Management Deception (Supervised) – evasive, negative, hostility, selective memory, refusal (e.g., “We really don’t have that number.”)
  - Deception + Negative Events – Optimism + Events
  - Consistency vs. Complexity
  - Correctness vs. Ambiguity Scores
  - Ownership Scores (“I”, “my” vs. “his,” “their”...)

For illustrative purposes only. Source: PanAgora; Yu, Chen and Lee (2019)

The figure above illustrates how syntax trees for the sentence “Everything would have been all right if you hadn’t said that” in two languages...
Wherefore ESG?

Yes, quants can do ESG too. Actually, quant is the only way to do it. But what is it?

In the last century, we had “tobacco-free” and “South African-free” portfolios. What do we have now? We have a cocktail comprising a dreamy 12-year bull market, mixed with a socialistic cultural trend in the Western World. So, let’s accommodate that with quant.

Figure 10 maps an optimal compromise between social adherence and dollar ($) adherence. Many of our proprietary factors rate and quantify ESG metrics such as carbon, diversity, and governance, to name a few. We find that integrating these metrics can actually improve alpha in many instances, but certainly not all (Chen and Mussalli, 2018). Depending on the client identification of ESG values, we can engineer an optimal tradeoff between making money with stocks and making peace with the world. The left quadrant of figure 10 illustrates that we can achieve both; many ESG factors enhance return over a reasonable horizon. Beyond some limit, the devotion to values deflates the goal of “value” in a return dimension.

First, we have quantified and generated an ESG score for every company (Chen, Mussalli and Zweibach, 2018). Second, we have used ML by “reading” thousands of corporate transcripts and reports to score the firm’s own self-assessment as to ESG responsibility. Third, we have the MSCI answer of their widely followed ESG scores universally available. We have a score for each company on two dimensions that are based on the following questions: 1) What do the objective quant data say about the firm’s ESG rating? And 2) What do the NPL modeling techniques and reading of public corporate communications say about the management’s self-awareness of ESG? Add the consensus MSCI score and we have a quant survival encore.

Figure 10: Optimization of Return and ESG

Here are the advantages:

- A model to construct portfolios reflecting asset owner’s values and return alpha can be developed
- Ability to inform asset owners of the exact tradeoff between alpha return and ESG goals
- One can map out an “efficient frontier” for investors

A Comment on Construction: Alpha vs. Beta — The Olympics of Quant

Portfolio construction has also evolved since the 1980s. Big Beta pronounces the dominance of price-weighted indices on the global scene. Around the time of the last financial crisis, we were recommending divergence from communal strategies such as big beta. We recommended a “shift from some of what we were calling alpha to beta, and some of what we were calling beta to alpha.” Admittedly, too many talking beta-to-alpha and vice versa seemed vague, or worse, sacrilegious.

The labels of smart beta or alternative beta literally represent that shift. However, shouldn’t
we think of this as an alpha beta mix? It is instructive to note how quantitative equity systems seeking alpha are being somewhat displaced by what is broadly called smart beta. Effectively, fee pressures on traditional approaches are affording quants the opportunity to maneuver beyond the “alpha” and “beta” separation paradigm.

Earlier, with the E-model (and “index enhanced” strategies that soon followed) selected top-rated stocks had attractive factor exposures, on average. With the growth of indexation, quants had to constrain beta to be equivalent to a price-weighted index, yet express stock overweights with the highest weighted average factor scores. This scoring is analogous to the ancient Olympic Games. According to historians, the first matches were likely dated 776 B.C. Initially, there was one event — running the “stadium race,” a 200-meter sprint inside the stadium.

Time saw the addition of four more events, giving rise to the pentathlon. Participants competed in five events comprising running, jumping, discus throwing, javelin throwing and wrestling. The highest weighted average score determined the Olympic winner. This is analogous to the ranking process in a five-factor quant model.

Smart beta has an analogous parallel, but now it is the Modern Olympics, in which teams comprise small groups of athletes highly focused on non-diversified, specific skills like running, jumping, throwing, etc. The discus thrower need not possess attributes required of the 100-meter sprinter (and vice versa). In today’s alternative beta process, the stocks (compared here to competing athletes) are not akin to Jim Thorpe, who won the pentathlon in 1912. The multi-factor alternative beta construction is to hold “clusters” of stocks with specific exposures such as value, quality, volatility, etc.

Let us assume for illustration that the track coach has to choose one method to field the team: select generalists or specialists. The two choices are individual athletes (each with all around strengths); or sub-teams of athletes (with homogenous strengths within). This is the quant equity metaphor, whereby you either populate a team with: 1) single stock multi-factor selection processes; or 2) alpha-beta mix processes with multi-factor exposures comprising single-factor tilted sleeves. The advantage of the latter is being able to merge alpha and beta, like in the old days (pre-1970).

Forthcoming, there will be synthesis with the freedom to mix alpha and beta quantitatively without being limited by some artificial “tracking error” to a bloated, price-weighted instrument (synthetic or ETF). The narrow objective to beat the market will be displaced by diversify my overall portfolio through differentiation of processes. This will require longer investor horizons with a goal of wealth accumulation over multiple time periods vs. what provided the biggest return last period. The recent growth in risk parity portfolios is evidence of this shift, in that the result is more wealth and lower volatility than traditional allocations, such as 60/40 stock-bond.

The Way Forward—Domain, Data, and Digital

Quantitative strategies will continue to reshape and advance the three circles of domain, data and digital through time. They will bring a much richer diversity of solutions that include across-asset class premia and within-asset class portfolio enhancement. Two battlefronts of innovation will separate the better quantitative investors from other adherents: 1) the fight for
superior information processes (data), such as smart alternative data that is creative and un-abused; and 2) the fight for innovation in expert security modeling and efficient portfolio construction (digital). That is, the researchers as always will seek better ingredients and better recipes. Implementation will hopefully come with fewer constraints such as smart alpha-beta blending.

Lastly, we should not underestimate the need for deep domain knowledge and experience. Will the Chief Data Scientist (CDS) become the new Chief Investment Officer (CIO)? (Representing only the left circle?) Will the Chief Engineer (CE) be promoted to CIO? (Representing just the right circle)? Either would only be a partial solution. In one instance it is “garbage in – garbage out.” In the other, it is “pure physics can be hazardous to your wealth.” Superior domain knowledge is always irreplaceable. The winners will be managers who thrive on creative discovery and position themselves at the intersection of domain, data and digital, the three circles of successful investing.
References


Legal Disclosures

This material is solely for informational purposes and shall not constitute an offer to sell or the solicitation to buy securities. The opinions expressed herein represent the current, good-faith views of the author(s) at the time of publication and are provided for limited purposes, are not definitive investment advice, and should not be relied on as such. The information presented in this article has been developed internally and/or obtained from sources believed to be reliable; however, PanAgora Asset Management, Inc. (“PanAgora”) does not guarantee the accuracy, adequacy or completeness of such information. Predictions, opinions, and other information contained in this article are subject to change continually and without notice of any kind and may no longer be true after the date indicated. Any forward-looking statements speak only as of the date they are made, and PanAgora assumes no duty to and does not undertake to update forward-looking statements. Forward-looking statements are subject to numerous assumptions, risks and uncertainties, which change over time. Actual results could differ materially from those anticipated in forward-looking statements. This material is directed exclusively at investment professionals. Any investments to which this material relates are available only to or will be engaged in only with investment professionals. There is no guarantee that any investment strategy will achieve its investment objective or avoid incurring substantial losses.

The information presented is based upon the hypothetical assumptions discussed in this piece. Specific assumptions: risk is allocated across and within sectors equally. Certain assumptions have been made for modeling purposes and are unlikely to be realized. No representation or warranty is made as to the reasonableness of the assumptions made or that all assumptions used in achieving the returns have been stated or fully considered.

Investment examples are shown for illustrative purposes only and should not be considered a recommendation to purchase or sell any security. It should not be assumed that future recommendations/transactions will be profitable or will equal the performance of the securities described. Each account is managed individually. Accordingly, there can be no assurance that any security discussed will be held in any particular client’s portfolio at the time you review this report or otherwise. As with any investment there is a potential for profit as well as the possibility of loss.

PanAgora is exempt from the requirement to hold an Australian financial services license under the Corporations Act 2001 in respect of the financial services.

PanAgora is regulated by the SEC under U.S. laws, which differ from Australian laws.

Past performance is not a guarantee of future results.

---

1 See examples of the “Four C’s” of a strategy life cycle --- create, converge, capacity and capitulation, (Sorensen, 2009).
3 In due time “discovery” became “dollarized” by soon to be very large buy-side firms.
4 Volumes of published research are readily available, starting with pioneers like Bill Sharpe, Gene Fama and many others.
5 For a narrative history of how theory became normative see (MacKenzie, 2006).
6 I am quite sure that any old quants whose practice may have spanned the 1974 – 2020 period have different takes on this. However, there are few that occupy that span.
7 Our first encounter with smart data dates to the early 1970s. A major industry of the Pacific Northwest was forest products companies, such as Weyerhaeuser and Georgia Pacific to name a few. One innovative analyst determined that during periods of rising interest rates (building slump) vs. falling rates (building boom), a small (vs. large) lumber inventory separated the winners from the losers. Consequently, the analyst hired a helicopter pilot to routinely fly him directly over the lumber yards adjacent to lumber mills, as well as active logging sites, to assess the potential inventory levels. The helicopter data was proprietary and intuitively causal. Most importantly it worked, providing valuable company insights. Forty-five years later, the contemporary version of helicopter data is satellite imaging of shoppers’ parked cars. Today’s modern space technology
version is big and it seems causal. However, it may fail if everyone has access to it. Ubiquitous is neither proprietary nor innovative which means it fails to be “smart.”

10 See the example of an original “expert system” for the WWII Battle of Britain, (Sorensen, 2009).
11 We recently listened to an AI expert point to an application that few humans can do at all. Listen to a person’s talk over the phone and then predict the person’s face. Our investor ML parallel is to “read” the person’s speech and predict what is in the heart—shareholder compassion or self-aggrandizing.
12 After writing, (Sorensen, 2009) we had occasions to address quantitative audiences at various conferences. Our prescriptions were: 1) move from tinkering to real innovation, 2) move away from commonalities, 3) move some alpha to beta; and, 4) move some beta to alpha, 5) move from tracking to volatility, and 6) move from then cap-weighted crowd.
13 See (Qian, 2005, 2006) and (Sorensen and Alonso, 2015).