**TITLE**

Bridging The Gap Between Machine Learning, Fundamental Analysis and Crowdsourcing

**ABSTRACT**

…………

**INTRODUCTION**

Background

The investment management industry is facing numerous challenges and pressures. While this has always been the case, the current dynamics are unique in many aspects. Data availability and complexity is growing, the stock market is becoming increasingly efficient, management fees are being compressed and competition continues to rise in the form of active and passive investment products and solutions. On one side of the industry are fundamental investment managers that are failing to adequately embrace the technological innovation in the form of AI and data science tools, even though in many cases the companies they are investing in are doing so. And on the other side are quantitative investment managers that are scouring the data universe for new sources of alpha generation. Rarely do fundamental and quantitative managers bridge the gap between their expertise in a collaborative and effective manner which leverages the availability of data and data science tools as both are entrenched in their respective camps for numerous reasons thus we believe a chasm remains.

Tools Needed

Regardless of what an investment manager’s strategy and style is, in order to survive and thrive in this dynamic and challenging landscape, investment firms need tools that are: systematic, understandable, adaptable, complementary, efficient, scalable, multidimensional, innovative and effective.

Systematic tools enable the minimization of biases during the investment process, many of which are driven by emotions. Human’s are often irrational beings and the topic of behavioral biases has been well studied and established (ex. confirmation bias, anchoring, saliency and framing), thus being able to leverage tools to minimize the adverse consequences of such biases is deemed favorable. By having systematic tools, discipline to one’s investment philosophy and process is increased. Human’s are also prone to herding due to various internal and external pressures, thus having more disciplined steps in one’s process will result in increased likelihood of challenging the status quo and consensus views when such are not aligned with the current data. This in turn will make it harder to explain away new and pertinent information which is able to communicate what is not what should be. Also related to the inherent discipline brought by systematic tools, repeatability is able to be achieved. Consistency is a major key in investment management and you can’t have consistent results without consistent processes. Then if the process and performance are deemed repeatable by clients, not only confidence is built but also so is trust.

Understandability is also a necessary component of success in this industry. Trust must be not only created with clients but also maintained and is deemed to be directly correlated with understanding of one’s philosophy and process. Though computational power is a well accepted fact by both investors and clients, there is often a skeptical and distrusting perception toward such technology which can be due to understandability. It is important for the manager to have a deep understanding first and foremost and then be able to communicate this to such clients. All of which also ties in with the aforementioned comments around repeatability. Optimal success is achieved when trust and understanding are built and maintained thus when circumstances ebb and flow such bedrock can help improve the likelihood of lengthening such business relationships.

Though human psychology is not prone to change, the world around us all is. As such, investment managers must be able to adequately adapt to changes in the industry’s dynamics as well as those of the companies being invested in. It takes a great deal of processing power to be able to discern when to change as the inputs sufficiently warrant such change. It is not feasible for humans to be able to process the amount information that computers can thus leaning on the latter is ideal when it comes to developing adaptability in one’s investment process. As Mark Twain said, “History doesn’t repeat itself but it often rhymes”, and so it is deemed beneficial to be able to adapt to current change by reading the tea leaves of the past which computers are able to attempt to accomplish. Additionally, back testing is widely prevalent in the investment management industry, though one shortcoming of such tests is their lack of adaptability which data science tools help to overcome.

In the investment management industry, the word “change” often has a negative connotation. This is understandable given the many stories of managers being pressured by clients as well as the overall market to change their decision making in the heat of the battle which turned out to be the exact wrong time to do so. Thus developing a strategy and tools beforehand and sticking with it regardless of the circumstances is deemed to be difficult but optimal for long term results. Managers need tools therefore that complement and build onto one’s process, but do not uproot the core tenets and foundation. Clients choose to invest with managers due to what is communicated about the present and more so the past, thus there is a balancing act that must be achieved by adapting to changes well all the while making sure such adaptations are complementary to what is already in place.

Being able to sufficiently cover all the stock’s in a universe which can equate to a few thousand and up to greater than ten thousand is daunting to put it lightly. Some may segment the universe and focus only on those with predetermined criteria while others use screening tools to quickly narrow down the universe without going into much detail or depth. Ideally, an investment manager will have tools that are efficient enough to adequately cover the entire selected universe which enables one to see not only the forest but also sufficient details about the trees. An important aspect of such efficiency is the ability to identify key data sources and using various tools to draw out information from such sources which in turn can reduce the information noise intake and draw out what is and is not important in an efficient manner. This type of crowdsourcing represents both a benefit as well as a significant challenge as one must have the capability to determine which sources of information to tap.

Related to the aforementioned efficiency is the concept of scalability. Fees are being compressed in the investment management industry and are likely to remain to be for the foreseeable future thus reducing costs is of the utmost importance without sacrificing the quality of service provided. By having scalable tools, one can expand the scope of their process without losing the execution ability. Crowdsourcing is one example of means to enhance scalability within an investment process.

Multidimensionality is a complex concept in principle and practice. In the investment management industry, breaking down silos between departments and roles is deemed to be favorable. Portfolio managers, analysts, traders and other team members need to all be on the same page in many regards. Often they are not due to incentive structures that are not aligned thus leading to less than optimal results. By gathering pertinent information and leveraging data science tools, such multidimensionality can be addressed thus bridging the gap between various parties. Costs are also reduced if you can have people, tools, models, etc. that interact with each other and cover more bases which in turn requires less overhead.

Innovation is at the heart of technology. To thrive in the future, investment managers must not fight technological change but rather embrace it. Technology is either a tailwind or headwind. There is a dichotomy prevalent in the industry in that investment managers often invest in innovative companies because of such innovation and the potential it brings but they themselves are slow to embrace it. Data science is a relatively new and quickly changing field but due to such underlying characteristics it is also deemed to be quite helpful for those applying its tools to stay at the forefront of innovation as it pertains to their investment process.

At the end of the day, the investment management industry is primarily about winning which for this study is deemed to primarily applicable to those investment managers seeking to outperform a set benchmark. As such, developing tools to aid in the decision making process for investment managers is the end goal of this research, and such effectiveness will be put to the test in a scenario analysis for the years 2020 and 2021 with the former being well known to be challenging time for the entire world, the stock market included, due to the global COVID pandemic.

Potential Solutions

In light of the aforementioned pressures being faced by the investment management industry and tools’ characteristics needed to address such challenges, a few potential solutions are as follows. Additional fundamentally trained people can be hired in efforts to garner more breadth and depth. While this would bring more experience to an investment process, it would also present challenges in the form of additional costs, more biases, less scalability and other unfavorable aspects. On the other hand, hiring purely quantitatively trained people would address many of the aforementioned needs but also has its own shortcomings. As often is the case, such individuals don’t have sufficient fundamental background and more importantly experience thus resulting in less than optimal application of their knowledge. Instead of using innovative tools to draw out insights from foundational understandings, their research is predicated on scouring the universe for sources of outperformance which results in a high risk of data mining. Based on these potential solutions there remains an identified chasm. For each side has different mindsets, different experiences and different schooling, which in turn makes it difficult to create a bridge that leverages both sides’ strengths.

This chasm is a likely explanation as to why data science tools are not being embraced as much as they ought to be given their potential. A 2019 study by the CFA Institute that surveyed portfolio managers confirmed this and found that only 10% had used any artificial intelligence or machine learning in their investment process. Then for quantitative managers to embrace more of the fundamental side of the equation would likely require a step back from relying solely on their technological tools and delving deeper into the fundamentals which given the industry pressures would not be a well received proposition by most.

Therefore, if all of this is embraced then a gap remains that must be addressed. In this study, we will seek to bridge this chasm by leveraging the fundamental side of the equation which incorporates domain knowledge and select financial outputs, crowdsourcing of pertinent and powerful information, and the technological side of the equation which incorporates innovative data science tools and methods primarily in the form of machine learning and natural language processing. Not only will this research seek to develop tools to address the aforementioned needs, but also will seek to provide suggestions and recommendations as to how they can be applied.

**LITERATURE REVIEW**

The literature review focuses on four principal topics: usage of quantitative tools and machine learning in the investment industry, research on machine learning in the investment industry, research regrading analyst revisions and research regarding the momentum phenomena in the stock market.

* + Current usage of machine learning in the investment industry
    - BofA Quant Primer
      * Quantitative investing is getting increasingly complicated and more crowded per this research. Discussed is that their quantitatively oriented clients use three times the number of factors today than they did twenty five years ago. Also, the popularity of quantitative investing has increased likely at the expense of fundamental investing. They note that the market inefficiencies in modern times is being driven by the shift in capital toward shorter term strategies relying on shorter term data and on access to better, faster and larger stores of data. Additionally, jobs advertised for data scientists and quantitative analysts outnumber those for fundamental analysts by a factor of eight. Quants are increasingly focused on real-time data feeds, AI, big data and machine learning.
      * This research serves to support the notion that quantitative investing is a growing field but is mainly predicated on short term trading investment opportunities and are not rooted in the underlying principles of fundamental investing.
    - 10 reasons ML Funds Fail
      * This research notes the high rate of failure in quantitative finance, and particularly so in financial machine learning. A few are successful but this is a rare outcome for reasons that the researchers detail in their report. One of the noted pitfalls / reasons for failure is related to the THE SISYPHUS PARADIGM. This is premised on portfolio managers making investment decisions that do not follow a particular theory or rationale thus improvement and adapting is rarely achieved. These portfolio managers do not naturally work well as a team. Then wherever this formula is applied to quantitative or ML projects, it has not produced favorable results. “The boardroom’s mentality is, let us do with quants what has worked with PMs. Let us hire 50 PhDs and demand that each of them produce an investment strategy within six months. This approach tends to backfire, because each PhD will frantically search for investment opportunities and eventually settle for (1) a false positive that looks great in an overfit backtest or (2) standard factor investing, which is an overcrowded strategy with a low Sharpe ratio, but at least has academic support.”
      * This research serves to highlight several of the shortcomings of quantitative finance, in particular many of the shortcomings previously addressed related to biases, emotions, discipline, repeatability, efficiency, understandability and repeatability.
    - ERP ML
      * This research from Empirical Research Partners details how they have begun incorporating machine learning into their analytical processes. Their algorithm draws from a library of 100 variables that they believe to generate alpha. They note a balance attempted to be struck between thinking they know how things work and observing what is actually going on. Noted is how machine learning can help especially when the reality differs from its precedents.
      * This research supports the strategy being implemented in this analysis whereby underlying beliefs and views serve as the basis for the information that is then fed into machine learning models.
    - AI Pioneers in Investment Management
      * Highlighted in this research is the technological transformation taking place in the investment industry due to the adoption of artificial intelligence (AI) and big data applications. They found that relatively few investment professionals are currently exploiting AI and big data applications in their investment processes. According to a CFA Institute survey, most portfolio managers continue to rely on Excel and desktop market data tools with only 10% of portfolio manager respondents having used AI/ML techniques in the past 12 months. Identified in this research are five major hurdles to successful adoption of AI and big data in investment processes: cost, talent, technology, leadership vision, and time.
      * This research discusses the potential of AI in investment management as well as the challenges and the lack of utilization currently. This supports the notion that such tools and capabilities are far from fully being utilized and exploited thus those that can leverage them are likely at an advantage.
    - Machine Learning in Equity
      * This research discusses the use case for ML in financial markets, though from a very skeptical viewpoint.
      * The numerical nature of financial markets makes market forecasting and portfolio construction a good use case for machine learning (ML), a branch of artificial intelligence (AI). Over the past two decades, a number of academics worldwide (mostly from the field of computer science) produced a sizeable body of experimental research. Many publications claim highly accurate forecasts or highly profitable investment strategies. At the same time, the picture of real-world AI-driven investments is ambiguous and conspicuously lacking in high-profile success cases (while it is not lacking in high-profile failures). We conducted a literature review of 27 academic experiments spanning over two decades and contrasted them with real-life examples of machine learning-driven funds to try to explain this apparent contradiction
      * the number of AI funds and their assets under management (AUM) are extremely low compared to the size of the industry.
      * Ambiguous definition, investing vs trading, trading costs / paper profit
      * This research brings a very skeptical view to ML in financial markets. Sentiment and opinions similar to these likely is slowing the adoption speed which in turn may be providing an opportunity for those willing and able to exploit the power of AI.
    - EPS Forecast – growth investing
      * This research represents an example whereby a sell-side research firm, UBS, using the machine learning model, xgboost, to predict future earnings growth.
      * We introduce a model that seeks to identify stocks with strong growth at low risk. Companies with the highest growth are also typically very risky. Investors looking for high-growth stocks need to weed out these high-risk names. One approach is consider earnings uncertainty (measured as the coefficient of variation of analysts' EPS forecasts) and remove the most uncertain names from our universe. This should introduce a quality tilt to our growth strategy because high-quality companies have fewer unpredictable issues affecting their earnings and are more transparent about how their business is going.
      * We used a forecast earnings growth model combined with a forecast earnings uncertainty screen to construct a new defensive growth strategy. We have used xgboost, a popular machine learning algorithm, to create our earnings growth forecasts and earnings certainty forecasts. This approach allows us to capture non-linear relationships between our features and the factor we are trying to forecast as well as interactions between our features.
      * This research serves as an example of how machine learning can be applied to predict key fundamental outputs.
  + Academic research on machine learning in the investment industry
    - A Machine Learning View on Momentum and Reversal Trading
      * This research attempts to analyze the momentum and reversal phenomenon in stock markets by using machine learning. In the study, various machine learning techniques, including the Decision Tree (DT), Support Vector Machine (SVM), Multilayer Perceptron Neural Network (MLP), and Long Short-Term Memory Neural Network (LSTM) were explored and compared. The experimental results demonstrated that these machine learning approaches, especially the SVM, are beneficial for capturing the relevant momentum and reversal effects, and possibly building profitable trading strategies.
      * This research serves to highlight the possibility of using machine learning in order to analyze and model perceived stock market phenomena.
    - A Backtesting Protocol in the Era of Machine Learning
      * This research notes the power of yet danger with misapplying machine learning techniques and how such misapplication can lead to disappointing results. One of their recommendations is to carefully structure the machine learning problem so that the inputs are guided by a reasonable hypothesis. Another is to refrain from tweaking one’s model.
      * This research supports the notion that the human element is important in leveraging the power of machine learning models by starting with an underlying hypothesis and then using data supporting this hypothesis to be fed into machine learning models. Additionally, as previously discussed about discipline there will likely be times when there is temptations to tweak a model, likely due to poor performance, which this research would argue against.
    - Fundamental Analysis via Machine Learning
      * This research examines the efficacy of machine learning in forecasting corporate earnings. The researchers found that machine learning models, especially those that accommodate nonlinearities are powerful and helpful in this regard. Additionally noted is the suggestion that machine learning models can uncover economically sensible relationships between historical financial information and future earnings, thus can be of high value.
      * This research supports the belief that the combination of fundamental investment principles and machine learning are potentially powerful and effective.
    - Machine Learning-Based Financial Statement Analysis
      * This paper explores the use of machine learning methods on financial statement analysis. Such models were deemed to be helpful wit Random Forest models producing the most accurate forecasts and neural-network based models performing relatively better for predictions of extreme market reactions.
      * This research also supports the previous notion of generating favorable insights and results from the combination of machine learning and financial statement analysis.
    - Machine Learning for Stock Selection
      * In this research the authors describe some of the basic concepts surrounding machine learning and provide a simple example of how investors can use machine learning techniques to forecast the cross-section of stock returns while limiting the risk of overfitting. Such overfitting is argued to be predicated by the inclusion of only individual equity characteristics and not any macro variables.
      * The overfitting discussion supports this research in that there are no macro variables included but rarely only individual financial metrics included in the analysis.
    - Predicting Profitability Using Machine Learning
      * This research utilized random forest models to predict directional changes in five profitability measures. Their results suggest that machine learning methods offer better predictive performance than traditional regression-based methods.
      * The connection between machine learning and profitability metrics is deemed interesting as the primary historical fundamentals incorporated in this research pertains to profitability.
  + Academic research regarding analyst revisions
    - The Information Content of Financial Analysts Forecasts of Earnings
      * This paper assessed the information content of revisions in financial analysts’ forecasts of earnings by analyzing the relationship between the direction of these revisions and stock price behavior. This research serves as one of the foundational research reports for this stock market phenomenon. The results supports the notion that market reaction to the changes in forecasts is relatively slow and provides opportunities for abnormal returns to investors.
      * Though this research is somewhat dated, the underlying principle regarding analyst estimate revisions remains a well-regarded stock market phenomenon which will be further addressed in this report.
    - Herding
      * This research is not financial market related but rather addresses the social behavior of herding and how its application can be broadly applied. The authors address the mechanisms of transmission as well as the patterns of connection between herding agents.
      * Herding and the study of social behavior is deemed to be intriguing and aligns with this study and the use of crowdsourcing and studying the revision patterns of analysts to garner insights of a given company relative to others.
    - Analyst’s stock views and revision actions
      * This study seeks to make the following contributions research. First, it introduces the construct of analysts’ stock views. Knowing this stock view is useful for investors to make informed decisions when analyst produce inconsistent revision signals. Second, this study provides methodological contributions by illustrating how to properly identify the stock views. Third, the empirical results support that revision-consistency contains informative value, which can help investors identify more effective revisions and reap the economic benefits from analysts’ outputs.
      * The results of this study serve as a basis to not only look into analyst estimate revisions in general but also the possibility of gleaning further insight based on the consistency or lack thereof for such estimates.
    - Financial Analysts’ Forecasts of Earnings
      * This research evaluates the quality of analysts' forecasts as surrogates for the market expectation of earnings and compares it with that of prediction models commonly used in research. Results indicate that prediction errors of analysts are more closely associated with security price movements, suggesting that analysts' forecasts provide a better surrogate for market expectations than forecasts generated by time-series models. The study also identifies factors that might contribute to the performance of the financial analysts' forecasts.
      * This research, though dated back to the late 1970’s, also provides a historical perspective of the concept of crowdsourcing for information as they assessed the use of analyst forecasts as surrogates for the market’s expectation of future earnings.
    - Herding among security analysts
      * This paper sought to show that the buy or sell recommendations of security analysts have a significant positive influence on the recommendations of the next two analysts. This influence can be traced to short-lived information in the most recent revisions. This indicates consensus herding consistent with models in which analysts herd based on little information.
      * This research serves to support this study’s use of analyst revisions data and the changes in such revisions over various time periods.
    - The behavioral basis of sell-side analysts herding
      * This research assesses behavioral bias of sell-side analysts as they tend to move away from the prevailing consensus as their confidence increases. And as their confidence falls, they herd toward the prevailing consensus. Herding is an unconscious social behavior originating from the primitive portion of our brain. This herding impulse is advantageous to our survival by demonstrating that we are the same as the crowd for reasons of self-preservation, to avoid rejection and to defuse an excuse to attack. Sell-side analysts perform an economically useful service by incorporating information in markets which are not perfectly informationally efficient. Herding is also economically rational given analysts’ career concerns: being wrong when everyone else wrong is preferable to being wrong on your own.
      * This research supports the notion that sell-side analysts perform a useful service to the marketplace and one should study them as a group rather than as individuals and thus not trying to determine which one is correct or not with their estimates.
  + Academic research regarding momentum phenomena in the stock market
    - Momentum Strategies
      * This research seeks to relate the predictability of future returns from past returns to the market’s underreaction to information, focusing on past earnings news. Past return and past earnings surprise each predict large drifts in future returns after controlling for the other.
      * This research supports the inclusion of momentum / price trend information in this analysis.
    - The role of analyst forecasts in the momentum effect
      * This research evaluates the extent to which sell-side equity analysts can facilitate market efficiency. Their study finds that analysts can provide value-relevant signals to investors by picking up indicators of momentum. The ability to identify under or over-valued stocks suggests that analysts are important information intermediaries in the price-continuation momentum effect.
      * This paper supports the notion of momentum in the markets as well as the value-relevant signals provided to the marketplace by the actions of sell-side analysts.

Hypothesis at the end of literature review

Based on this literature review the researchers hypothesize that while machine learning and crowdsourcing in the form of analyst revisions data is well known and documented, the combination of the two coupled with fundamental analysis and understanding is lacking and thus represents and opportunity worth further analysis. This review also supports the notion of not following what many others are doing due to the inherent efficiency and competition within the market. This competition is primarily in the short-term which also supports this research which will seek to develop tools for more medium term horizons. Given the lack of widespread adoption of AI, there is still opportunities related to being an early adopter. It is the researchers opinion that humans should determine the inputs to the models and that a well defined strategy must be put in place early and especially before challenging times when the temptation to tweak will heighten. Also, analyst estimates and revisions are informative and potentially value-additive thus should be analyzed thoroughly. Lastly, momentum is also prevalent in the stock market and sell-side analysts play a role in such phenomena.

**METHODS**

ANALYTICAL FOUNDATION

There are many ways to analyze and value a given company, though for this study the following formula and its underlying components will serve as the basis.

Price = valuation metric X financial output

This formula states that a given stock price can be estimated based on a chosen valuation metric multiplied by a financial output that aligns with such valuation metric. For example, one example of this is the following formula which states that the stock price is estimated to be the forward price-to-earnings ratio multiplied by the forward earnings per share estimate.

It is undeniable that all of the aforementioned components are important and how one goes about taking them into account can differ greatly. Some may focus on the growth potential and thus mainly on the forecasted financial output, which understandably label them as growth managers. Technical analysis is predicated on solely focusing on the price movement of a given stock to determine its attractiveness. Value managers often focus first and foremost on the valuation component and its attractiveness or not when analyzing a given stock. This study is premised on the underlying belief that each component is valuable and worthy of analysis but have different underlying attributes thus should be analyzed accordingly. More specifically, it is believed that valuation is a measure that oscillates over time and is likely to gravitate towards the mean. On the other hand, financial outputs (such as earnings and revenues) may oscillate for some companies more than others but are less likely to gravitate toward the mean but instead trend over time. Then price takes both into account but is a fallout of the two rather than the driver. Mathematically, it is impossible to make an investment decision based on the analysis and forecast of all three components thus this study seeks to cover all the bases by “betting on” two of them and “respecting” the third. Specifically, the researchers will create a strategy and tools that seek to bet on valuation and financial output all the while respecting the price.

DATA

The data for this study was sourced from FactSet, a financial data and analytics company. Through access gained from the SMU library, the researchers were able to download data via FactSet’s excel add-in. The starting universe chosen was the constituents of the Russell 1000 Index, which serves as a benchmark for many U.S. large capitalization investment managers. Companies in the real estate sector were removed given they typically trade on different metrics such as funds from operations (FFO) instead of earnings per share (EPS). Then this universe was reduced down to 652 by only included those companies which have financial data going back to the end of 2009. Weekly data was collected spanning from the end of 2009 to the end of the June 2021.

EXPLANATORY VARIABLES

The data downloaded can be categorized into the following: valuation, analyst revisions, analyst EPS estimates, analyst revenue estimates and price.

The measurements gathered and/or calculated can be categorized as follows: growth / trend, volatility, absolute level and relative level.

**Valuation**: To garner an encompassing and well rounded assessment of each company’s valuation, the researchers included the following valuation metrics: Price to Forward 12M EPS, Dividend Yield, Price to Book and Enterprise Value to Sales. By factoring in several valuation metrics, the researchers believe that such analysis sufficiently accounts for all the financial statements and their related financial outputs and underlying information. The researchers then include the absolute value as well as the changes in such metrics over various time periods and several measure of volatility of these metrics.

**Analyst Revisions**: As discussed in the literature review, there is presumed to be information in the estimates and revisions of sell-side analysts that cover a given company. In this section, the direction of revisions is accounted for as well as the second-derivative of such revisions over various time periods.

**Analyst EPS Estimates**: Sourced from the same analysts as above, these estimates are for the earnings per share (EPS) estimates for a given company during the next 12 month period. With these estimates the researchers calculated numerous metrics which account for the growth, second-derivative growth / revision and volatility.

**Analyst Revenue Estimates**: Similar to the EPS estimates and the associated calculated metrics, but this represents the revenue estimates instead.

**Price**: The stock price for each stock was gathered and then several metrics to assess the trend of each was calculated.

Upon gathering and calculating the aforementioned metrics, the researchers then calculated the relative percentile (0-100) of each metric relative to this universe at that specific point in time. This was premised on being able to compare one stock to the entire universe based on every metric to determine its attractiveness or not.

In total there were 370 explanatory variables.

RESPONSE VARIABLES

There were 9 different response variables included in this research which will each be modelled and incorporated into the recommended and tested strategy.

**Valuation**: This study includes the following valuation metrics: Price to Forward 12M EPS, Dividend Yield, Price to Book and Enterprise Value to Sales. For each of these metrics the researchers calculated the 3 month revision which computes the percentage change in the metric during the next 3 month period. Then this calculation was compared on a relative basis such that a percentile was computed. As such, there were four different metrics to be modeled.

**Analyst EPS / Revenue Estimates**: Analyst EPS estimates and analyst revenue estimates were modeled similarly except for the underlying metric. First, the next 3 month’s growth was measured and then compared to the universe to get a percentile. Then a 3 month revision (i.e. second derivative growth) was measured and then compared to the universe to get a percentile.

**Price**: Given the recommended application of price for this research will be to respect the trend, the researchers will create a 3 month binary classification model which measures whether or not a stock in this universe outperforms the universe median during this time period. The median was chosen instead of the equal weighted Russell 1000 index given the superior returns for the median of this research’s universe thus was deemed to be more conservative in nature. Though the Russell 1000 Index is the predominant index used by US large cap managers which is the target audience for this study, this research removes the individual weighting aspect of this index thus can be accounted for later by users of the models built.

MODELLING ANALYSIS

VALUE

Hypothesis: The researchers hypothesize that market participants are more so focused on the next best opportunity rather than market timing. Valuation is an measure that is believed to oscillates over time thus is subject to mean reversion forces. Determining which stocks are most likely to mean revert in the greatest magnitude (i.e. low to high) provide opportunities to generate relative outperformance.

Strategy

Details

Findings

Conclusion / recommendation

GROWTH

Hypothesis: As mentioned above, the researchers hypothesize that market participants are more so focused on the next best opportunity rather than market timing. EPS and revenue growth is believed to be rewarded in the marketplace as the market rewards the companies that are growing fastest as well as those whose revision growth (i.e. second derivative growth) is superior.

Strategy

Details

Findings

Conclusion / recommendation

PRICE

Hypothesis: Given the aforementioned strategy by the researchers whereby the valuation and growth metrics will drive the stock selection and the price trend will be respected, the hypothesis is to develop a classification model which assess the likelihood of outperformance for a given stock in the selected universe for the same time period as the valuation and growth models.

Strategy

Details

Findings

Conclusion / recommendation

RESULTS

Visualization

2020 SCENARIO ANALYSIS

Strategy

**RESULTS**

* Summarize model building and visualization efforts
* Summarize scenario analysis
* Accept, reject or modify hypothesis based on the results
* Clarify study scope

**DISCUSSION**

* Discuss modeling efficacy
* What unexpected insights did we garner
* Discuss takeaways – wins and losses
* What challenges did we face
* Make recommendations as to how our research can be applied in the industry
* Discuss aspects where others can go deeper with our research
  + Valuation insights
    - Having more valuation information vs peers using DS to understand best way to determine relative value
  + Insider transactions
  + More technical variables
  + More macro variables that directly or indirectly affect companies

**CONCLUSION**

* Overall summary of how this research is useful
* Revisit opening statements regarding the state of the industry and addressing of needs

**REFERENCES**

* ……