The ESG Decision Tree

By Herb Blank

Over the past six months, I have led the team that developed the Thomson Reuters Corporate Responsibility Ratings. Our primary resource in developing, constructing, and maintaining the ratings is ASSET4, a Thomson Reuters business that provides objective, relevant and systematic environmental, social and governance (ESG) information based on 250+ key performance indicators (KPIs) and 750+ individual data points along with their original data sources. Since its founding in 2003 and acquisition in 2009, Asset4 has been recognized globally as a premier source of ESG data. More than 100 analysts use their experience to collect relevant, comparable (companies often report in different units, scopes and formats) and up-to-date information utilizing publicly available sources (e.g. annual reports, NGO websites, CSR reports).

Asset4 classifies these data into categories within each major pillar. The Thomson Reuters Corporate Responsibility Ratings follow this convention in data aggregations. For example, the environmental pillar consists of three category groupings: emission reduction, product innovation, and resource reduction. The governance pillar has five categories: board functions, board structure, compensation policy, shareholders policy, and vision-and-strategy. The social pillar is the most complex with seven categories: community, diversity, employment quality, health-and-safety, human rights, product responsibility, and training-and-development.

Utilizing models we have based upon the primary ESG data underlying the aforementioned 15 categories, we are now able to provide environmental, social, governance and composite ESG ratings and rankings on over 4600 public companies worldwide. This universe is expected to expand by approximately 300 companies per year moving forward.

In order to make these universally comparable baselines, we adopt the lowest common denominator approach. All data are quantitative. No subjective assessments or overrides are used. No public companies are eliminated or penalized for populating industries considered “bad” by some constituencies or for producing products that some consider detestable. Similarly, companies that have been involved in environmental, social or governance controversies will only find their scores affected within the pillar where the controversy occurs and even there according to objective metrics that are applied uniformly.

Satisfying the goal of providing standardized ESG ratings required a huge amount of data, data cleansing, data analysis, research, and quantitative modelling. We worked closely with the team at Asset4 for the better part of a year to bring these ratings to fruition.

In attempting to transform data into ratings that were objective and meaningful, we arrive at the following conclusions:

1. There can be no definitive and universally accepted right or wrong way to weight and model the Key Performance Indicators, or KPIs, collected and measured by Asset4. That said, hard decisions had to be made in order to produce
deterministic ratings. The data speak volumes but grouping them and interpreting them properly require a massive amount of data massaging and reconciliation of exceptions.

2. The ratings are designed to provide the most appropriate peer-to-peer comparisons. At the same time, we endeavor to avoid over-fitting so the relationships remain robust over time. To accomplish this, each Asset4 pillar is handled and modelled differently. Environmental KPIs tend to be very global-industry-specific. Alternatively, corporate governance practices are best benchmarked by region. Our attempts at getting more granular by investigating region-specific models within each industry-specific environmental model led to preliminary results with little stability from year-to-year so this pursuit was abandoned. The same was true in trying to further break down the region-specific governance models to make them more industry specific.

3. The social practices pillar was the most challenging of the three. Product-responsibility and health-and-safety practices were best benchmarked by industry sector but employment quality and community citizenship practices were most differentiated by region, and human rights issues are benchmarked universally.

4. Each KPI is scored within each industrial, regional, or universal model between zero and one. Denominators for each metric KPI are calculated accordingly. We also classify each KPI in terms of “polarity” meaning whether a higher score was “bad” or “good.” If it is classified as “bad”, it needs to be subtracted from one. Our modelling efforts and weighting coefficients are driven by analysis of the data distributions. KPIs that seem to repeat the same information are weighted less. KPIs that are only reported by a relative handful of companies are generally weighted less than those reported by at least 20% of industry or regional peers. After much research and deliberation, we have decided to treat non-reporters of a KPI identically to the worst reported KPI within the peer group.

5. Policy indicators are weighted less than observed practices. With some exceptions, Boolean (Yes =1/No=0) variables were generally given less weighting than reported metrics. That said, metrics where the grouping of responses tend to be clustered tightly get lower weightings than highly-dispersed metrics. For example, within the environmental pillar, hard metrics related to emissions and usage of non-renewable resources together constitute 45% of that pillar while policy-driving statements combined for only 5% of that pillar’s weight. In the governance pillar, vision and strategy KPIs have lower weights than key metrics related to shareholder rights, board structure, and disparities in firm compensation packages. Within the social pillar KPIs related to measurable product-responsibility and health-and-safety metrics carry higher weightings than diversity-and-opportunity policy drivers.
6. Each KPI weighting is checked against academic literature, where applicable, for consistency of results. If our statistical analysis shows no differentiation for something that has been documented as a key variable, we wish to make sure that we do not underweight that KPI due to a temporal aberration. The fact that we now have more than six years of Asset4 data has helped to provide us with an increasing amount of stability in this regard but we consider this an evolving process to be checked at each ratings reconstitution.

This entire process produces three numeric values for each company screened. These are:

1. **Score.** Every company is scored from 0 to 100 for each pillar. These scores are driven by Asset4 data, which in turn is driven by company financial reporting. For current scores, the most recent year available is used with the fiscal year clearly delineated. The scores are calibrated to be robust over time while also be relative to each company’s peer group.

2. **Percentile Rank.** Based on a company’s raw scores as defined above, percentile ranks are calculated for all companies screened.

3. **Ratings.** Finally, the raw scores are z-scored for each pillar and the z-scores are divided by the highest z-score and multiplied by 100 to derive company ratings. The ratings are comparable across pillars and centered to provide the most accurate assessment of a company’s environmental, social, governance and combined ESG practices.

Distinct time series of ratings for each pillar for every company provides bases for many potential applications. These include: input for risk factor models; customized peer group analysis; loss mitigation policies; compliance; due diligence; and many different types of strategic analysis. It also allows the subscriber to reformulate overall ratings based upon his or her own viewpoint since our ratings equally weight each pillar. These are objective building blocks – tools that can be deployed as needed and tailor-made to apply overlays such as negative screening as desired.

I believe the products of these labors now constitute fair, objective, and replicable methods for baseline comparisons of ESG corporate responsibility for each pillar separately and in combination. I make no claims that these ratings are “better” or more robust than other ESG ratings or assessments. I can say that they provide objective standards for comparison that are being made available to all.
REFERENCE SOURCES


8. Burnett, Royce; Skousen, Christopher; Wright, Charlotte; “Eco-Effective Management: An Empirical Link between Firm Value and Corporate


11. De, Indrani and Clayman, Michelle, “Are All Components of ESG Scores Equally Important?” NYSSA Financial Professionals Post, July 2010


