StarQube

INVESTMENT DECISION IN A BOX

StarQube ESG #3

Timestamping ESG data and models
The importance of point-in-time data

Timestamping ESG data and models

In this article, we will endeavor to demonstrate why it is essential to timestamp your ESG data and models in order to have a dynamic ("point-in-time") view of them.

Beforehand, let us note that the ESG matter is far from being fixed, but is on the contrary extremely moving; moving on various levels:

- *Societal*: society's perception of good environmental, social or governance practices can change rapidly – for example, the outbreak of war in Ukraine may have changed the way we look at the arms industry.
- *Regulatory*: the ESG regulatory framework applicable to asset managers is still under construction; at the beginning of 2022, for example, the debate focused on whether or not nuclear power should be included in the European taxonomy.
- *Data availability*: the ESG reporting obligations applicable to issuers are also constantly evolving; ESG data is still very incomplete but will necessarily be enriched over the next few years.

But back to the timestamp of ESG data: these, like macroeconomic data or company results (1) are published with a lag, and (2) may be subject to revisions.

For example, the Pfizer Group published on August 27, 2021 greenhouse gas emissions of 2,005,730 $mtCO_2eq$ for the year 2020, revised to 1,350,000 $mtCO_2eq$ on October 1, 2021.

This lag in the publication of ESG data and the possibility of revisions raise two difficulties for asset management companies:

1. *Reporting*: based on our example, the ESG score of a portfolio including Pfizer Group financial instruments (equities or bonds), presented as of December 31, 2020, can potentially have three different values depending on whether it is calculated at the beginning of 2021 (based on emissions for the year 2019), in September 2021 (based on the first figure published for 2020) or at the end of 2021 (based on the revised 2020 figure).
2. *Backtesting*: in the absence of "point-in-time" data allowing to know precisely what information was available at each instant "t" of the past, a quantitative researcher who wishes to use the Pfizer stock to backtest an investment strategy has two options (none of which is satisfactory):
   - Option 1: apply a safety margin by considering that ESG data are published with a large lag of e.g. one year. Consequence: a probable deterioration in the quality of backtests and the likely rejection of strategies that are nevertheless relevant.
   - Option 2: consider that the last accessible data (that published on October 1, 2021) was available from its value date (December 31, 2020) and then present overoptimized backtest results that cannot be reproduced in reality.

What we have just described for an individual issuer also applies to ESG scores published by specialized rating agencies, which can also:
1. Publish their ESG scores with a lag in relation to their value dates.
2. Revise their historical scores to account for updates, raw data revisions, or because they are refining their scoring methodology.

Beyond raw ESG data, it is just as essential to timestamp (i.e. version) ESG models. Indeed, most asset managers develop proprietary ESG models – by combining the ESG scores from rating agencies or by building their models directly from raw data collected from issuers (for the most advanced).

However, these models are evolving and will continue to evolve rapidly over the next few years:
1. To reflect changes in the society's environmental, social and governance priorities (modification of raw ESG data weightings).
2. To take into account changes in regulations and likely methodological standardization.
3. To incorporate new ESG data fields that will be made available.

These changes in methodology pose the same problems for asset managers as the revisions of raw ESG data:
1. *Reporting*: the asset management company must be able to recalculate identically and a posteriori the past ESG scores of its portfolios, even if the internal ESG model has been adjusted in the meantime.
2. *Backtesting*: in the backtesting of ESG investment strategies, it will be necessary, in the future, to be able to chain ESG models that may have been developed over the years and according to the evolution of knowledge of the ESG analysts.

In summary, without a precise timestamp of their ESG data and without a versioning of their models:
1. Asset managers are necessarily confronted with practical difficulties in their reporting. These difficulties are obviously multiplied among large asset managers who potentially manage hundreds of portfolios across vast investment universes.
2. It is impossible to properly backtest investment strategies.

WHAT STARQUBE OFFERS

The StarQube database is natively point-in-time, which means that:
1. All data is written in the database with a date of availability (their timestamp) in addition to their value date.
2. Revised data is never overwritten. New data is created in the database with an identical value date but a new timestamp.
3. All data models (e.g. ESG models) are versioned and can be chained in time; it is also easy to duplicate a model to build a new one by modifying a few parameters.

Therefore:
1. *Reporting*: it is always possible to reconstruct a report identically and a posteriori (e.g. recalculate the ESG score of a portfolio at an earlier date), even when the underlying data or the models have made subject to revisions in the meantime – without effort on the part of the user.
2. *Backtesting*: the StarQube backtesting tool can – by construction – only use data as it was available on the date of the calculations, without risk of over-optimization (look-ahead) bias and again, without user intervention.

LINK TO FULL ARTICLE

About StarQube

Founded in 2013, StarQube develops an innovative and modular solution for asset management companies based on two pillars. The **data management** pillar industrializes the collection, cleansing and organization of all types of data useful for the investment process within a centralized NoSQL database. The **portfolio construction and management** pillar offers modules to analyze the research universe, build proprietary risk models, create model indices or portfolios, model and backtest investment strategies, optimize and rebalance portfolios. Graphical interfaces make it possible to view, analyze and manage portfolios using customizable screens to display the information which is relevant according to the investment style.

Contact

StarQube
Rue des Corps-Saints 4
1201 Genève

Sales contact: Guillermo Albiñana Arias
Phone: +33 6 52 33 80 33
Mail: guillermo.albinana@starqube.com
Web: www.starqube.com
LinkedIn: https://www.linkedin.com/company/starqube