RAISE PARTNER

Thematic Intelligence Generates Alpha in an Index Tracking Strategy

White Paper

CID

Executive Summary

CID and Raise Partner collaborate to combine "Thematic Intelligence" and "Robust Optimization" to build thematic portfolios by leveraging CID's thematic signals as an additional input to Raise Partner's portfolio optimization solution.

In this white paper, we describe how we applied this innovative approach to build thematic portfolios exposed to Digital Healthcare and Battery Recycling Technology, while controlling the tracking error of a strategy. We illustrate that an optimization of such thematic portfolios shows financially rewarding metrics in terms of overperformance and controlled risk, while ensuring exposure to the selected themes.

CID and Raise Partner decided to address the market with a combined solution for their clients, automating the data exchange between their solutions. Beyond the sample themes and investment strategies addressed in this white paper, a wide range of thematic investment ideas can be implemented with this new combined solution.

Feel free to reach out at contact@raisepartner.com or sales@cid.com.

Introduction

Thematic investing has become increasingly popular in recent years [1]. Besides criteria such as sectors, regions, and defined market caps, investors and asset managers now include information about expected business transformation and trends to build high-performing and diversified portfolios [2] [3] [4] [5].

Innovation and emerging themes are often interdisciplinary and result from intensive collaboration between experts from various industry sectors (e.g., Healthcare and IT). Likewise, innovating companies emerge from multiple industry sectors, or very specific niches. For asset managers, it does not suffice to select securities and stocks using standard industry taxonomies (e.g., GICS, SIC, NACE …) or fundamental and quantitative data to identify innovation-relevant investment targets. In addition to insights from market analysts and quants, asset managers need to incorporate the expertise both from business experts and academic research to explore trends and assess innovation with regards to its future impact [6] [7].

To align such macro research insights with fundamental and quantitative analysis, asset managers need to close the “Insights Gap”: Is a theme relevant? Is a theme attractive for investment? Who are the thematic leaders? This process, however, is very time consuming as asset managers need to directly compare significant volumes of qualitative data (such as company strategies reported and commented on in the news) with fundamental and quantitative data sources (such as revenue, market cap, and price performance).

In this whitepaper, we describe an approach driven by **Thematic Intelligence** to assess company exposures to innovative themes that expedites the time-consuming process of asset allocation in thematic investing.

Additionally, we demonstrate that the optimization of sample thematic portfolios (*Digital Healthcare, Battery Recycling and Technology*) reveals financially rewarding metrics. Applying this process to S&P500 companies, our back testing analysis proves excess return of higher value. Consequently, this approach provides high quality asset allocation for investment portfolios. It enables asset managers to research and construct portfolios faster, including the capability to analyze the overall performance of what started as a new investment strategy.

Company media exposure to trending themes

To support asset managers to optimally close the "Insights Gap", we derived a ranked list of companies exposed to innovative themes. Such a ranking does not only reveal theme-relevant companies, but it additionally provides insight into thematic leaders. The ranking was determined by automatically integrating information from news via open web crawling to provide company exposure scores related to the theme and fundamental data (e.g., revenue). We enriched relationship information to include suppliers' and partners' theme exposures in the ranking. The overall process was optimized to minimize the number of manual steps required.

Themes applied for back testing

We used the following trending topics to evaluate our approach:

Digital Healthcare

Digital Healthcare is an interdisciplinary market with roots in IT and healthcare. The goal is to enhance traditional healthcare services and procedures with software and hardware solutions. This enables for more precise and personalized as well as more scalable healthcare solutions. Various aspects such as telemedicine, augmented reality or treatment recommendations based on data and questionnaires are part of Digital Healthcare.

Battery Recycling and Technology

Battery recycling is an activity that aims to reduce the number of batteries being disposed of as municipal solid waste. Disposing batteries by the same process as regular household waste has raised concerns over soil contamination and water pollution because they contain various rare-earth elements, heavy metals, and toxic chemicals. Regaining such valuable elements requires a variety of technologies and makes this theme interesting for investment.

Finding signals of theme exposure

A direct signal of theme exposure exists when a news article specifically mentions a company exposed to a topic.

We used **Affinity**, CID's Thematic Intelligence platform, to determine monthly company media exposures to trends and themes:

- **ThemeAI**, the AI brain of our system maps themes, signals, and events onto entities like companies and people, powered by CID's unique NLP, Deep Learning and Graph technology. It crawls relevant and interesting news for asset managers. Here, international, national, and local news are monitored to assure a broad coverage of business-related events. These news reports are processed and categorized along a set of defined criteria, such as:
    - Corporate reports (e.g., earnings announcements, product launches)
    - Human resource matters (e.g., employee or location changes)
    - Transactions (e.g., investments, funding, mergers and acquisitions, IPOs)
    - Business risk (e.g., legal disputes, cyber attacks, trade conflicts, environmental risk)
- Asset managers or analysts use **ThemeScape** to establish Machine Learning models for qualitative themes by tagging pre-processed news according to the desired theme focus (Machine Learning approach text classification).
- Leveraging these models, **ThemeScore** identifies investable companies and scores their theme exposure, allowing to detect thematic leaders.

Additionally, **Affinity** provides several options to fine-tune the scoring to support tailored asset management workflows and investment strategies. Here, we applied the following settings:

- We aggregated exposures from news based on suitable date ranges to determine sustainable transformations and to avoid

irrelevant media spikes. We observed that a good aggregation period is 28 days.

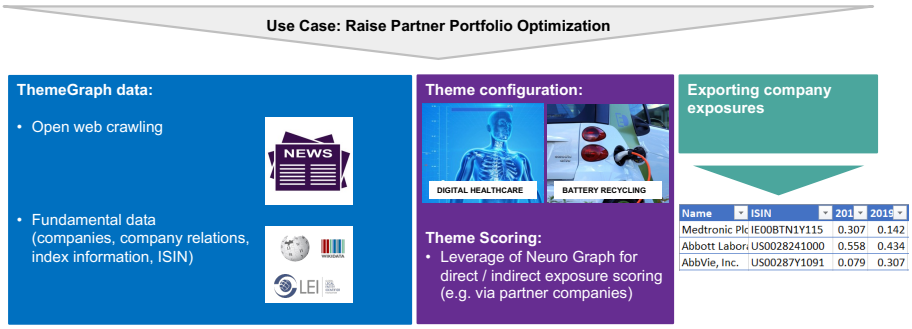
- We provided exposure scores over a period of two years to secure sufficient data for a robust back testing.
- We focused on a subset of S&P 500 companies to compare the quality of different back-tested thematic portfolios easily.

Expanding the portfolio of companies with CID's ThemeGraph

Direct signals provide scores for companies that are directly mentioned in the news. This approach, however, does not account for hidden champions, which are often not present in the media but have a significant market position in a supply chain. CID's **Affinity ThemeGraph** provides a network of companies to enrich the exposure ranking by factoring in exposure scores from linked companies including customers, partners, and suppliers.

Graph theory is applied to "transfer" scores through the graph. The benefit of this approach is it includes different levels of indirect company exposures in the eventual exposure values. Indirect exposure scores are normalized on a scale between 0 and 1 to allow comparisons between companies.

CID and Raise Partner analyzed relationships between the directly exposed companies and potentially hidden champions. We revealed potential performance based on thematic exposures and informed the portfolio construction process.

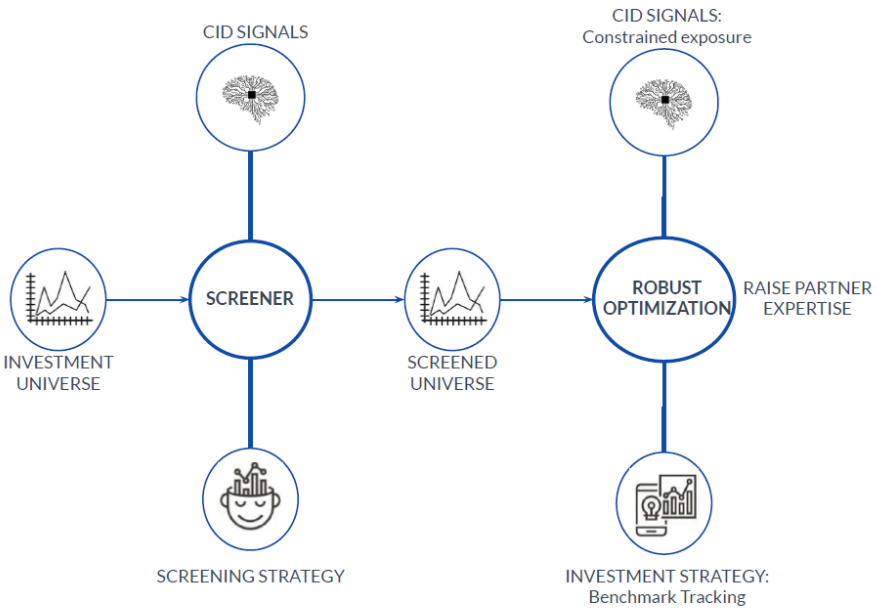


Aggregated CID theme scores in Raise Partner optimization tools

Once we derived theme exposure scores from Affinity, we intended to identify the best way to leverage such information within the portfolio optimization tools developed by Raise Partner and prove its value-add.

There are two complementary ways of using this information, as shown in the following graph:

S&P HealthCare Index performance against S&P Index performance



The first straightforward idea was to screen the investment universe, to exclude less exposed stocks (compared to a defined threshold), and to define a portfolio following a strategy leveraging the exposure scores provided by CID. This way, we reduced the investment universe to companies with higher theme exposure. However, this approach might have excluded stocks potentially interesting from other aspects (e.g., over performance, limitation of risk, etc.).

We therefore explored a second type of optimization combining multiple criteria in the company selection process: a defined theme

exposure threshold and a measure to control and reduce risk. In detail, we tracked a benchmark and defined a minimum theme exposure at the portfolio level simultaneously. This was achieved by minimizing the tracking error between the final portfolio and its benchmark with a linear constraint on the aggregated theme score at the portfolio level. The optimization problem we solved is the following:

$$\min_{\omega} TE(\omega)$$

$$\omega^T \lambda \geq min\_exposure$$

where:

- TE(ω) is the tracking error between the optimized portfolio and the benchmark
- ω is the vector of optimized weights computed by the optimizer
- λ is the vector of the chosen exposures at the instrument level that we wanted to control
- min_exposure is the minimum exposure score we defined for the optimized portfolio

Backtesting of individual portfolios

To highlight the value added to portfolio construction and optimization by theme exposure scores, we computed different out-of-sample backtests from January 1st, 2019, to July 19th, 2021, leveraging two theme exposure data series provided by CID.

The mechanism of out-of-sample backtesting enabled us to check the robustness of a strategy over time and see how it performed during the past years. The idea was to reallocate the portfolio every three months according to the solution of the described optimization problem and using the CID exposure data (between two rebalancing dates the stock selection did not change). To estimate parameters as inputs to the simulation, we used a one-year rolling, fully out-of-sample time window for the backtesting process.

We worked with two themes, Digital Healthcare and Battery Recycling. The results described in the following sections are the ones obtained from a minimization of the tracking error ("TE") between our given universe and the chosen benchmark - S&P 500 - with linear constraints on the thematic exposures. It is important to note that none of these optimizations had any target return constraint.

Even though the theme exposure scores did not reveal the same companies, we managed to combine them into a single investment universe and to perform our optimization in the combined universe of both themes. We also performed a standard index tracking approach without any input from CID, used as a reference portfolio to measure the value added by the theme exposure constraint.

Results evaluation – Digital Healthcare

Raise Partner backtested different levels of constraints on the given exposures for a combined company universe for both themes (404 stocks). Exposure scores remained stable until the next exposure rebalancing date. For the remainder of stocks with no exposure score at a specific time, we assumed it to be null.

For the sake of brevity, we present the results obtained for two different backtests (on top of the standard index tracking approach):

a) Backtesting with a 0.4 constraint on indirect Digital Healthcare exposure
b) Backtesting with a 0.35 constraint on Direct Digital Healthcare exposure

Therefore, the optimization problems we solved are the following:

$$\begin{matrix} \min_\omega TE(\omega) \\ \omega^T \lambda_{\text{Indirect\_DH}} \geq 0.4 \\ \omega^T \mathbf{1} = 1 \end{matrix} \qquad (1)$$

$$\begin{matrix} \min_\omega TE(\omega) \\ \omega^T \lambda_{\text{Direct\_DH}} \geq 0.35 \\ \omega^T \mathbf{1} = 1 \end{matrix} \qquad (2)$$

Ex-post statistics

The following ex-post risk and performance indicators were computed on the backtest:

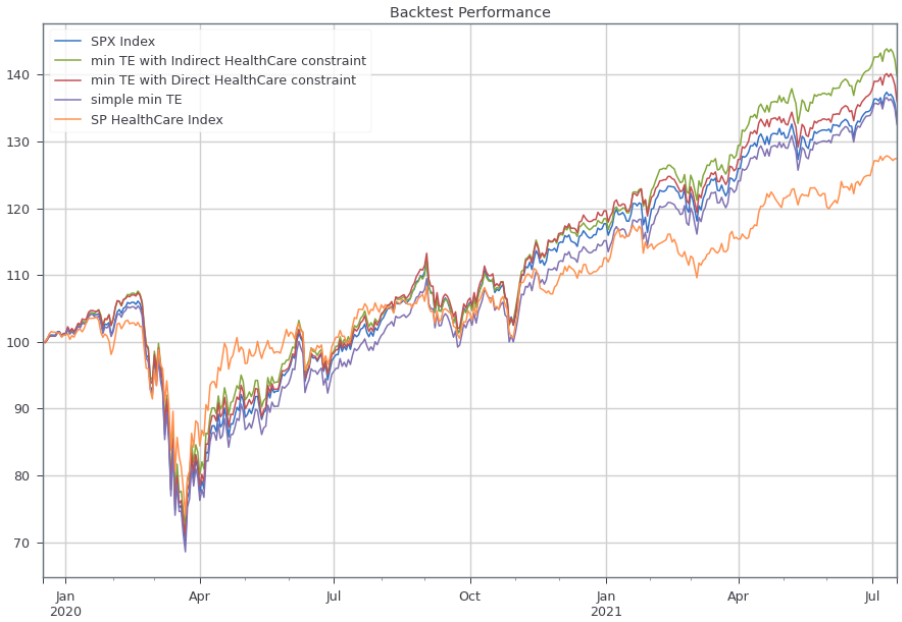
- Annualized tracking error of the strategies vs. the benchmark (S&P 500)
- Annualized excess return compared to the benchmark
- Information ratio, when positive, which is the ratio between the excess return and the tracking error

The following table shows the value of these indicators for these two backtests and our reference strategy, a simple minimization of the tracking error without any CID exposure input.

	TE	Excess Return	IR
Min TE	**1.83%**	**-0.18%**	**<0**
Min TE with Indirect HC Constraint	**2.91%**	**2.84%**	**0.977**
Min TE with Direct HC constraints	**2.8%**	**1.36%**	**0.49**
S&P HealthCare	**11.8%**	**-3.45%**	**<0**

These statistics show a significant improvement in the information ratio due to the increase in excess return (from the signals) and the control of the tracking error (from the optimization).

The following graph shows the performance of the different backtests described above.



Backtest performance against the S&P500 Index performance

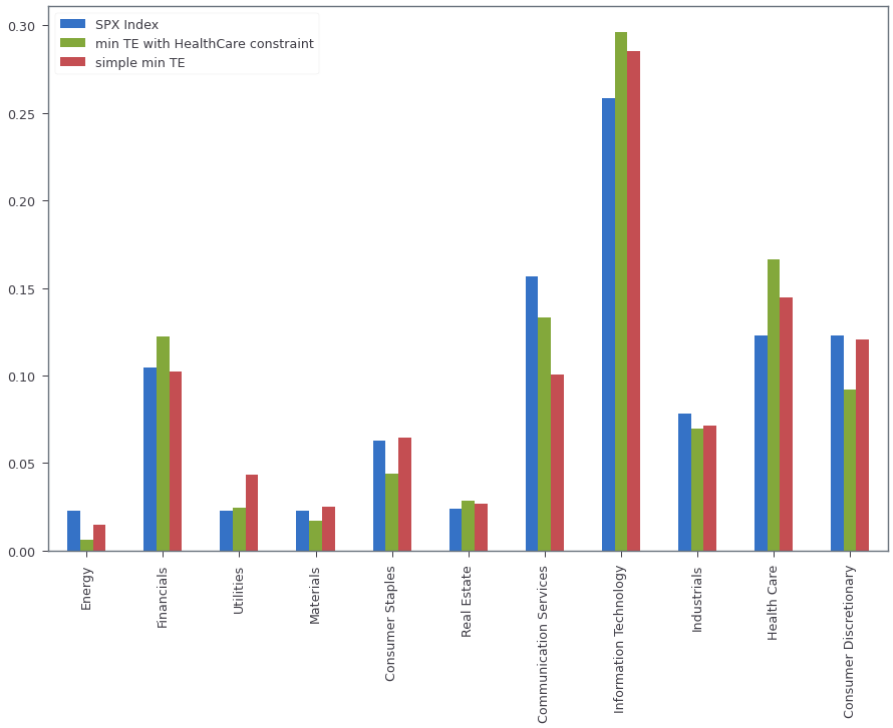
The backtests using constraints on CID exposures outperform the reference portfolio (S&P500 graph in blue), even though no target return constraint was set. It can be relevant to compare the obtained back-tested performance to the S&P HealthCare Index (an index aggregating all companies belonging to the "HealthCare" GICS sector) because this index could have been an alternative allocation choice for this thematic investing process. The graph shows that this index does not outperform the reference portfolio and – based on its components – provides no sector diversification.

The goal was to create thematic portfolios following a benchmark without degrading performance. These first results show that the CID exposure data provided valuable information concerning innovation

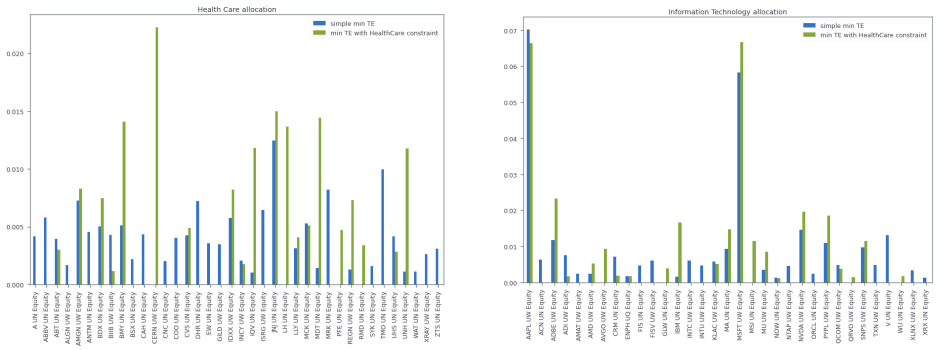
themes that outperformed the average market, and that we successfully integrated this information into Raise Partner optimization models.

Sector Allocation

It is also very interesting to observe the effect of the constraints on CID theme exposures on the sector allocation.



Allocation of stocks form a sector point of view



Stock picking effect (inside the Healthcare and the Information Technology sectors)

These graphs show the robustness of our optimization tools when constraints must be met. The first graph shows the allocation aggregated by sector for the last optimization of the reference backtest and the one with indirect Digital HealthCare exposure constraint and the allocation of the S&P 500 at the same date. This shows that Raise Partner optimization tools did not completely change the allocation from a sectorial point of view, ensuring a certain diversification. Looking at the allocation within each of these sectors, we notice that the optimizer has an impact, indeed, and supports stock picking. The two graphs above illustrate this process in two different sectors, proving that the HealthCare sector was not the only one impacted by the optimization.

For instance, inside the HealthCare sector, Cerner Corporation, LabCorp, and Medtronic were picked while Thermo Fisher Scientific and Merck & Co were dropped.

It is important to see that the optimizer managed to keep a low tracking error while enhancing the exposure to Digital HealthCare without completely distorting the allocation.

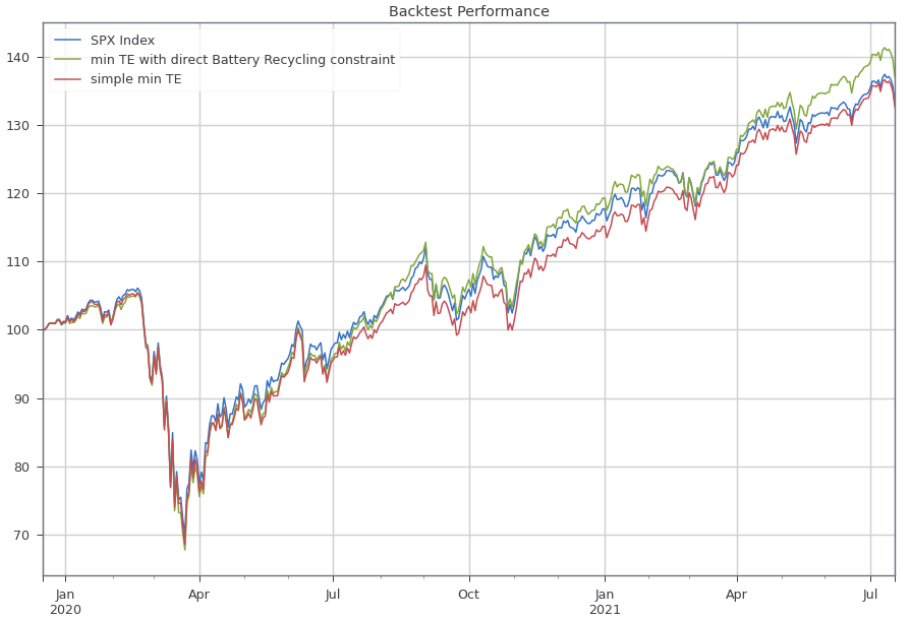
Results evaluation – Battery Recycling

The same backtesting process was applied to Battery Recycling using the same parameters (rolling estimation window, start date and end date, frequency of allocation) but with constraints on the corresponding theme exposures. The results shown in the following are the ones obtained for a minimization of the tracking error between our portfolio and the benchmark in the same investment universe with both, a constraint of 5% direct exposure to Battery Recycling and no constraint.

$$\min_{\omega} TE(\omega)$$

$$\omega^T \lambda_{BR} \geq 0.05$$

$$\omega^T \mathbf{1} = 1$$



As we can see on this backtested performance graph, the strategy exposed to Battery Recycling outperforms the SP500 Index, without any target return constraint.

	TE	Excess Return	IR
Min TE	**1.83%**	**-0.18%**	**<0**
Min TE with Direct BR Constraint (5%)	**3.52%**	**1.83%**	**0.52**

Once again, we can see an improvement of the Information Ratio when performing a minimization of the tracking error under a Battery Recycling exposure constraint.

Exposure volatility

We observed that the Battery Recycling exposures provided by CID were volatile, which could be a problem with regards to data consistency. Stable signals are indeed preferred when backtesting, so that the results are coherent over time. To tackle this issue, and to avoid a higher turnover rate in our allocation, we decided to compute "smoothed" exposure scores to work with. Basically, we computed a three-month rolling exposure average to flatten any anomalies. These smoothed exposures were introduced to the optimization problem by replacing the vector λ by the average values:

$$\min_{\omega} TE(\omega)$$

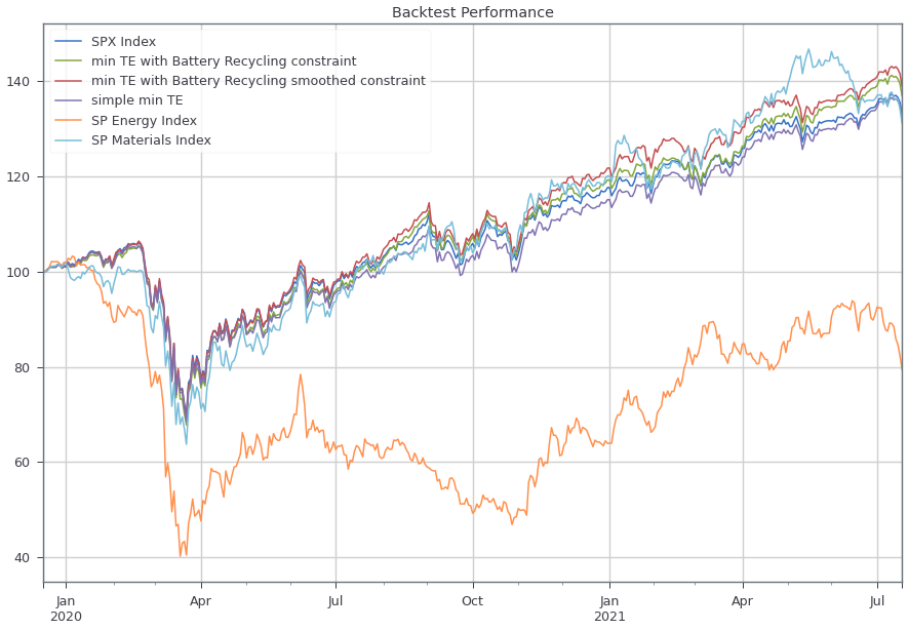
$$\omega^T \lambda_{smoothed_B R} \geq 0.05$$

$$\omega^T \mathbf{1} = 1$$

We applied the same backtesting process as described in the previous section using these smoothed exposures to the Battery Recycling theme. We set the minimum exposure at 5%, too. The investment universe was still composed of the same stocks as before.

The first graph presents the backtested performance compared to the S&P 500 index and two other indices, the S&P Materials Index (an aggregation of all stocks of the S&P 500 index belonging to the "Materials" GICS sector) and the S&P Energy index (an aggregation of all stocks of the S&P 500 index belonging to the "Energy" GICS sector).

The latter two indices are alternatives for a thematic investing strategy without sector diversification.



	TE	Excess Return	IR
Min TE	1.83%	-0.18%	<0
Min TE with direct BR constraint (5%)	3.52%	1.83%	0.52
Min TE with direct smoothed BR constraint (5%)	3.11%	2.67%	0.86
SP Energy Index	35.97%	-22.41%	<0
SP Materials Index	14.17%	0.36%	0.03

This smoothing process outperforms the benchmark by 2.67% while controlling the tracking error and maintaining Battery Recycling theme exposure.

Combination of direct Digital Healthcare and Battery Recycling constraints

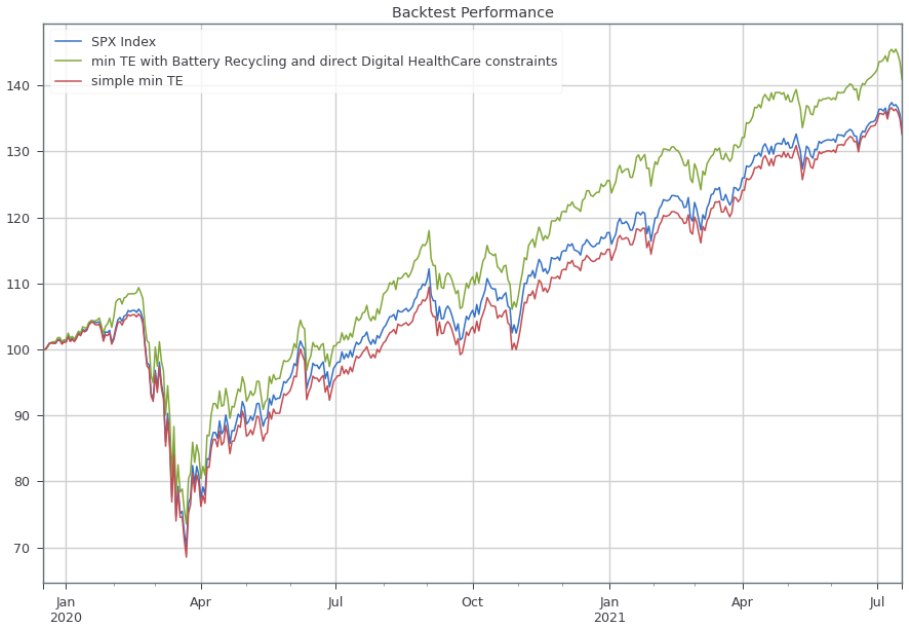
The flexibility of the Raise Partner optimization tools allowed us to combine both themes in a single optimization process. We chose to constrain the optimized portfolio with a minimum direct exposure of 35% to Digital Healthcare and a minimum direct exposure of 5% to Battery Recycling. The optimization problem we solved at each rebalancing date remained the same: we minimized the tracking error between our portfolio and the S&P 500. To avoid high turnover rates, we set the exposure constraint on the smoothed exposures, computed along the same process as described in the previous sections.

$$\min_{\omega} TE(\omega)$$

$$\omega^T \lambda_{smoothed\_Direct\_HC} \geq 0.35$$

$$\omega^T \lambda_{smoothed\_BR} \geq 0.05$$

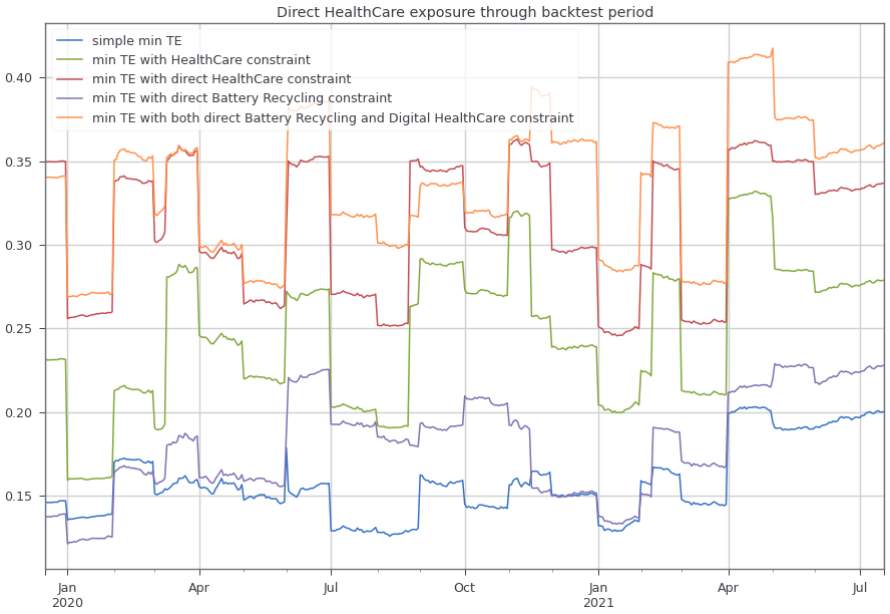
$$\omega^T \mathbf{1} = 1$$

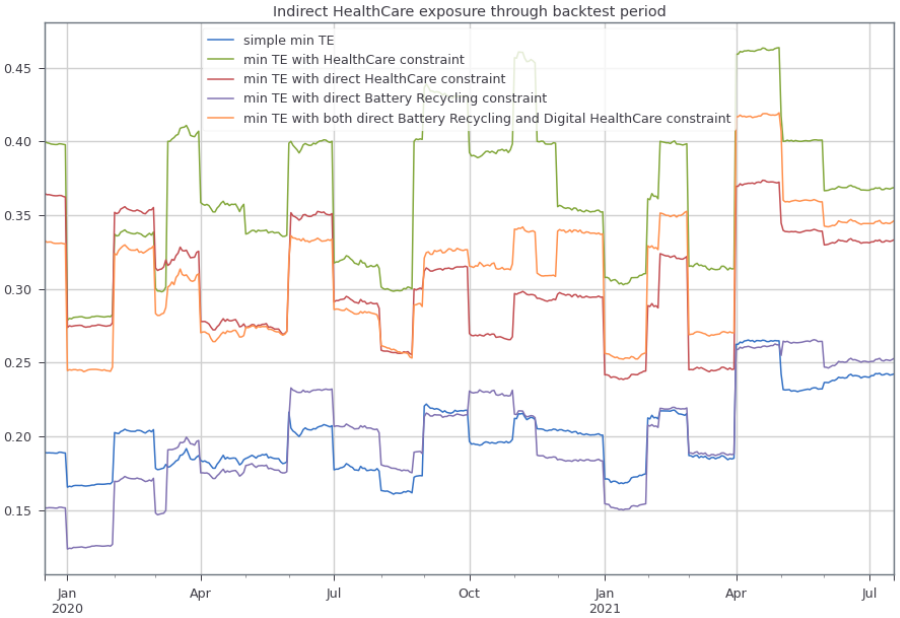


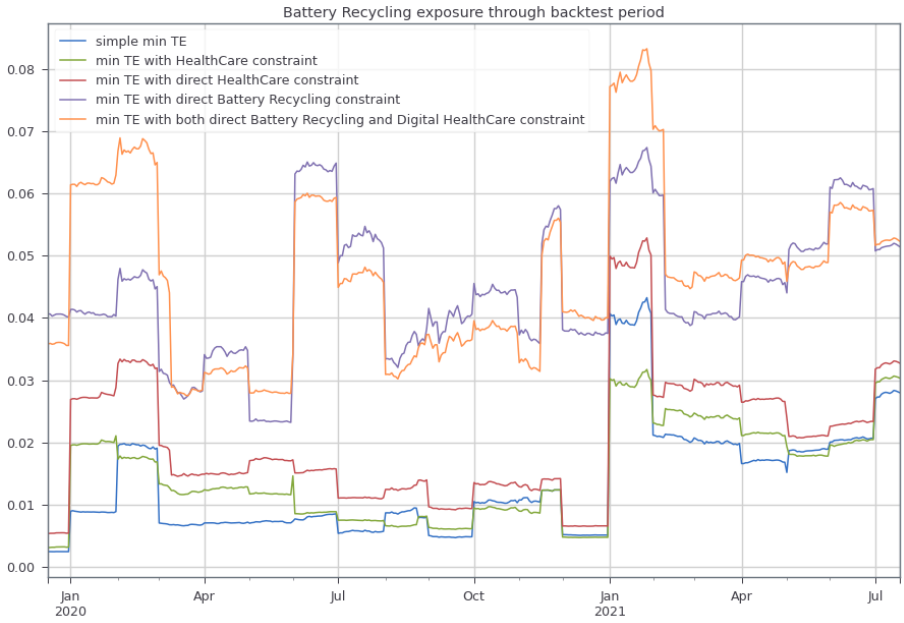
	TE	Excess Return	IR
Min TE	1.83%	-0.18%	<0
Min TE with direct BR constraint (5%) and direct HC constraint (35%)	4.01%	3.56%	0.88

Comparison of theme exposures over time

The following graphs show the exposure to the two themes we worked with over the backtesting period.







These graphs show that the simple minimization of the tracking error (in blue) is the least exposed to the themes. Thanks to the theme exposure scores provided by CID and the flexibility of the Raise Partner optimization tools, we proved that it is possible to create a multi-theme portfolio and control risk.

The increase of the information ratio was not a goal we intended to achieve but it can be interpreted as a performance improvement with regards to the innovative themes we chose.

Conclusion and outlook

Raise Partner and CID collaborated on enhancing a thematic portfolio construction process. The partners applied the example of an index tracking strategy to demonstrate the strengths and possibilities of combining their areas of expertise. We leveraged company exposures to innovative and potentially profitable themes computed by CID and provided these scores into Raise Partner optimization tools. We managed to build a simple showcase illustrating the construction of a thematic allocation while controlling a specific risk measure.

Bottom line, we enrich a portfolio construction process with theme building media exposure calculation and portfolio optimization capabilities. We enable an asset manager to back-test the performance of individual themes across industries and to fine-tune the portfolio optimization along specific fund characteristics. This may include:

- For Theme Building, we may combine trending themes with additional dimensions such as ESG
- For exposure calculation, we could include only specific types of network companies such as suppliers or partners
- For portfolio optimization, we could start with the stock screening based on exposure values. We would exclude companies from the investment universe with a score lower than a defined threshold. We would adjust the smoothing of exposure scores to optimize turnover costs.
- We would diversify constraints by including additional parameters such as transaction costs or turnover constraints, sector exposure constraints or quantitative ESG constraints.

These approaches demonstrate the flexibility of our process and software for an asset manager to adapt all parameters to specific fund requirements.

Annex - Robust Optimization Techniques by Raise Partner

Convexity of the objective function and constraints are crucial in optimization: usually, only convex optimization problems can be solved efficiently. Convexity also increases the robustness of the optimization problem: when inputs (covariance matrices, etc.) are slightly modified, only slight changes are observed in the output optimal portfolio. It is a crucial property since inputs are highly uncertain and changes in portfolios have a cost (transaction costs).

A lack of convexity can occur when the investment universe is large compared to the number of observations. In this case, an optimization problem is used to convexify the covariance or correlation matrix by increasing its definite positivity; given the original matrix (non-convex or not convex enough), this problem finds the closest matrix with the target convexity and respecting other types of constraints.

$$min_{X \in S^n} ||X - X_0||_F$$

sc

$$X \geq \epsilon Id_n$$

$$< C_i; X >= b_i \; \forall \; i \; \in 1 \ldots neq$$

$$< C_i; X > \leq \; b_i \; \forall \; i \; \in 1 \ldots nineq$$

with,

- $||.||_F$: the Frobenius norm
- S^n : set of square symmetric matrices of size n
- X_0 : the original covariance matrix (can be definite positive or too slightly)
- C_i and b_i : matrices and vectors used for linear constraints
- ϵ : value of the expected smallest eigenvalue of X

Using this calibrated variance/covariance matrix in the optimization model ensures the stability of the optimal portfolio.

Raise Partner

Founded in 2001, Raise Partner is a leading independent specialist in investment risk analysis and portfolio optimization software. We deliver risk/ESG smart analytics and optimization models to asset servicers, asset managers, wealth managers, insurers, investment bankers, sales, and advisors.

Smart Risk APIs are the result of 20+ years of R&D in market risk modelling, cross asset class portfolio analysis and robust portfolio optimization. Its modular and scalable technology makes it a unique partner for global financial actors to deliver high added value services.

Gabrielle Merrien
Jolan Banckaert
Sophie Echenim

contact@raisepartner.com

CID

CID's Thematic Intelligence platform Affinity allows wealth & asset managers to create new investment strategies, energize existing ones and create new products and do it quickly, easily and in near real time. Asset managers use Affinity to research, build and rank new investment strategies; make fast, accurate allocation and rebalancing decisions; and develop and sell more targeted, customized, thematic investment products.

Dr. Christian Brettschneider
Dr. Jörg Dallmeyer
Max Lörch
Alexander Stumpfegger

sales@cid.com

Bibliography

[1] ETF [Online]. Available: https://www.etf.com/sites/default/files/images/Magazines/ETF%20Report/2021/04/etf_report_january2020.pdf.

[2] Vontobel, "Thematic investment shows rapid growth. Are investors ready?," 07 06 2021. [Online]. Available: https://am.vontobel.com/.

[3] Credit Suisse, "Investment Outlook 2022," [Online]. Available: https://www.credit-suisse.com/microsites/investment-outlook/en.html.

[4] Allianz [Online]. Available: https://www.allianzgi.com/en/investable-themes/the-china-way/china-insights/nine-themes-for-investing-in-chinas-transformation.

[5] Defiance [Online]. Available: https://www.defianceetfs.com/the-thematic-etf-boom/.

[6] MSCI, "Thematic Experts," [Online]. Available: https://www.msci.com/thematic-investing/theme-experts.

[7] The Singularity Group [Online]. Available: https://singularity-group.com/about/expert-advisory-board/.