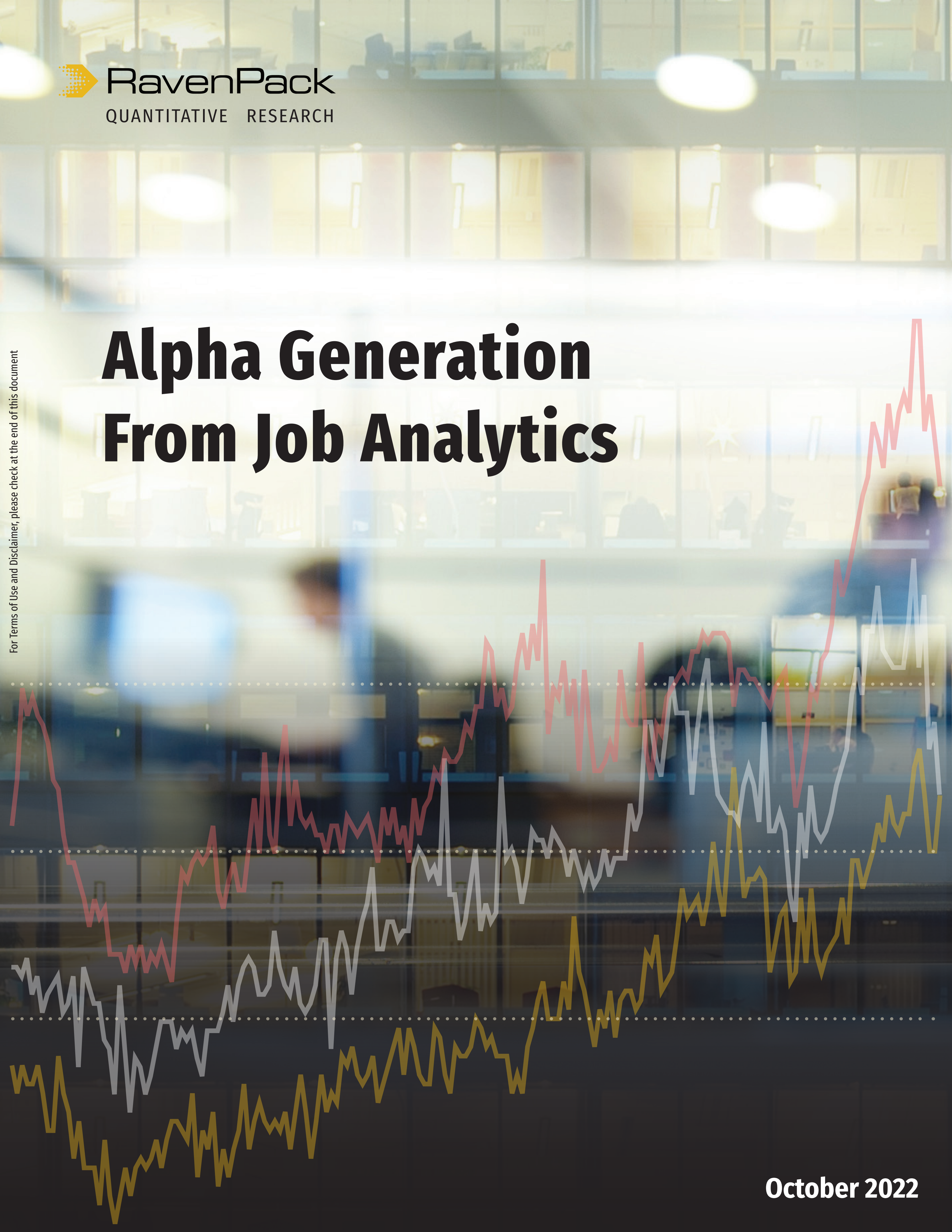


RavenPack
QUANTITATIVE RESEARCH

Alpha Generation From Job Analytics

October 2022

Executive Summary

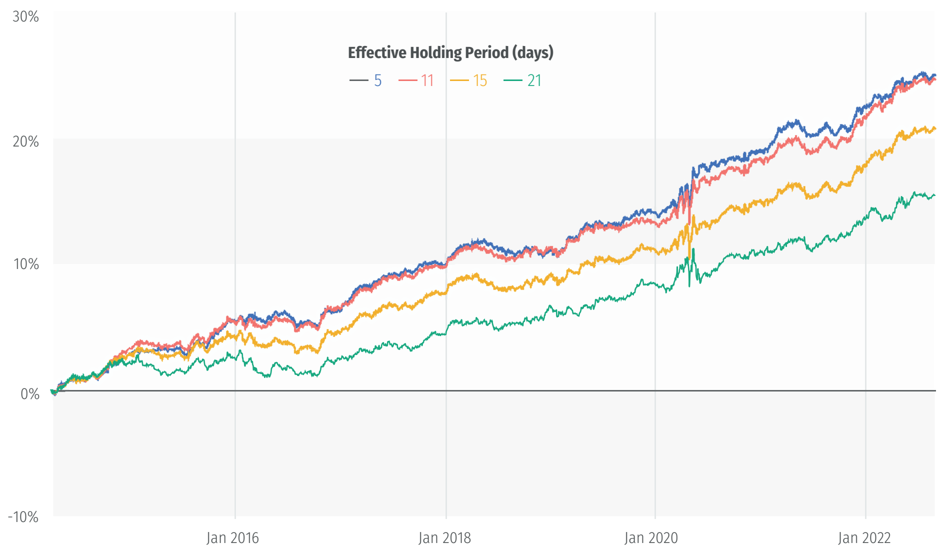
Human capital is a key factor in value creation. Therefore, changes in hiring can provide valuable insights into a company's performance and investment prospects. In this paper, we demonstrate how investors can benefit from job posting data using RavenPack Job Analytics powered by LinkUp. We find that:

(i) Monthly hiring growth is positively correlated with future stock performance. Additionally, companies with higher hiring growth that hire in similar locations outperform. A long/short sector-neutral portfolio delivers an Information Ratio of 1.1 and Annualized Returns of 2.9% with a weekly holding period.

(ii) A strategy focused on the monthly growth of the most demanded positions for each sector generates better performance than the overall growth strategy. The resulting portfolio achieves an Information Ratio of 1.4 and Annualized Returns of 3.0% with a weekly holding period.

(iii) Job descriptions carry orthogonal alpha. Companies with a stable distribution of soft skills outperform others, leading to a strategy with an Information Ratio of 0.9 and Annualized Returns of 2.0% with a monthly holding period.

(iv) The combined strategy shows robust performance over different trading horizons. With an effective holding period of two weeks, the strategy delivers an Information Ratio of 1.7 with Annualized Returns of 2.9%. The Information Ratio remains above 1.0 with a monthly holding period.



Cumulative log returns for the final combined portfolio using various smoothing windows, resulting in different holding periods. Results for U.S. mid/large-cap companies.

RavenPack Research Team

Peter Hafez
Chief Data Scientist

Ricard Matas
Head of Data Science Research

Francisco Gomez
Senior data Scientist

Alan Liu
Data Scientist

research@ravenpack.com
www.ravenpack.com

Americas: **+1 646 277-7339**
EMEA & APAC: **+44 (0) 20 3290 7395**

Table of Contents

1. Introduction 4
2. Data description 5
3. Portfolio construction & backtesting 7
4. Hiring growth 8
5. Most demanded positions 10
6. Skills similarity 12
7. Human capital & job requirements 14
8. Combined strategy 15
9. Conclusions 17
Appendices 18
References 19
Terms of Use 20

1. Introduction

There is a strong link between human capital and business success. A company's growth or decline in hiring can, therefore, provide valuable insights into its performance and investment prospects.

In recent years, a number of studies have explored the relationship between human capital and financial performance. ROUEN (2020) showed that the future value of a company's personnel expenses (PE), a proxy for human capital investment efficacy, is positively associated with characteristics of human-capital intensive firms, but found that market participants fail to fully factor in human capital investment. GUTIERREZ (2019) found that changes in the number of job postings at an organization are positively correlated with future investment performance and that this relationship is stronger when postings are driven by growth rather than replacement. QIN (2019) showed that turnover is negatively associated with future financial performance but that the negative relationship disappears when turnover is very low, suggesting that a certain amount of turnover is beneficial. More recently, CHUCHU (2022) found that high administrative intensity is associated with lower employee turnover and higher employee job satisfaction, but that administrative intensity is negatively associated with future firm performance.

In the past, most studies in this space have focused exclusively on job posting numbers, as analysis of job descriptions is onerous. In this paper, we present RavenPack Job Analytics powered by LinkUp, which analyzes job postings directly from employer websites using state-of-the-art Natural Language Processing (NLP) technology, and provides data on positions, locations, and job descriptions.

In **SECTION 2**, we provide a brief overview of RavenPack Job Analytics. **SECTION 3** outlines our portfolio construction and backtesting methodology. In **SECTIONS 4**, **5**, **6**, and **7**, we showcase four trading strategies using job openings, hiring locations, and job descriptions for the U.S. mid/large-cap universe to demonstrate how this data can provide consistent and idiosyncratic alpha. In **SECTION 8**, we present a combined strategy using the previous four methodologies as building blocks. Finally, we provide general conclusions.

2. Data description

RavenPack Job Analytics data is sourced from LinkUp, the global leader in real-time job market data. LinkUp captures job postings directly from employer websites daily and includes more than 200 million jobs listings from over 60,000 companies across more than 195 countries since 2007. From a volume perspective, the data is weighted towards the U.S. market as almost 80% of the postings originate from U.S. companies, as shown in **FIGURE 1**. The data, however, captures the global hiring activity of these firms.

The data set covers the U.S. mid/large-cap[1] universe reasonably well at around 80% over the recent years. This area of the market represents 67% of total postings. Coverage of the U.S. small-cap and European mid/large-cap universes is closer to 60%, but increasing over time. The volume bias towards the U.S. is not merely due to superior coverage but also a result of large-cap companies engaging in more hiring activity.

RavenPack Job Analytics powered by LinkUp includes information on the position listed in each job posting. Positions are classified into 23 major groups, 98 minor groups, and 459 broad groups under the O*NET-SOC 2019 occupations[2] structure. This refined structure makes the jobs data easier to digest.

RavenPack job posting documents contain a machine-generated title covering employer, position, and location data, which delivers information in a clear and concise manner. **FIGURE 2** shows the paragraph distribution of job postings from U.S. companies over time. Prior to 2013, most job posting documents contain no paragraphs (only titles) because descriptions were not collected.

Powered by LinkUp

RavenPack Job Analytics is powered by LinkUp, the global leader in real-time job market data. LinkUp sources job postings directly from employer websites daily, and includes more than 200 million jobs listings from over 60,000 companies across more than 195 countries since 2007.

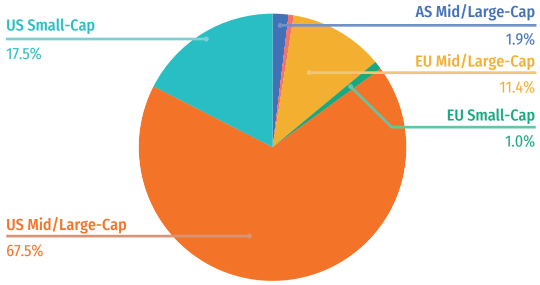


FIGURE 1: Percentage of job postings per trading universe. *Source: RavenPack, October 2022*

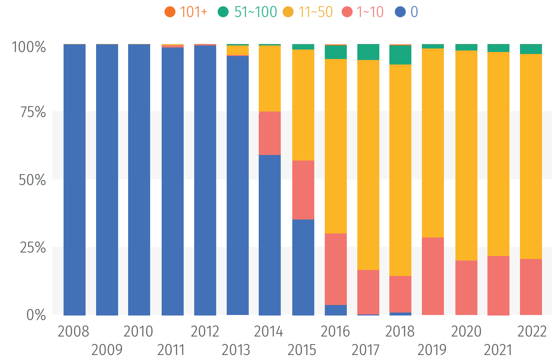


FIGURE 2: Paragraph distribution of job posting documents for U.S. mid/large-cap companies per year. *Source: RavenPack, October 2022*

1 The global universe consists of the top 3,000 companies in the U.S. and the top 7,000 companies worldwide ex. U.S., based on market capitalization. The large-, mid-, and small-cap categorizations worldwide ex. U.S. are achieved by splitting companies into groups based on their market cap percentile (p), i.e. p > 87.5 for large/mid-caps and p ≥ 87.5 for small-cap. In addition, we apply some stricter stability requirements on how companies move from one group to another, adding ±2.5 to each of the threshold values. The U.S. market cap categorizations are achieved by assigning top 1,000 ranked companies by market cap to large/mid- and 1,001-3,000 companies to small-cap groups.

2 More details can be found here: O*NET-SOC Taxonomy at O*NET Resource Center (onetcenter.org)

Job descriptions contain an abundance of information; however, processing it requires a considerable amount of effort. RavenPack Job Analytics provides a solution to process jobs data more efficiently and systematically. RavenPack Job Analytics also provides information related to job descriptions, skills required, and benefits, as shown in **FIGURE 3**. Qualifications describe employers' requirements of candidates and contribute about 70% of total volume of entity detections within the jobs taxonomy. Qualifications can be divided into six different groups: skills, personality traits, knowledge, experience, education and abilities. Skills can be further split into other subcategories such as technical skills or soft skills.

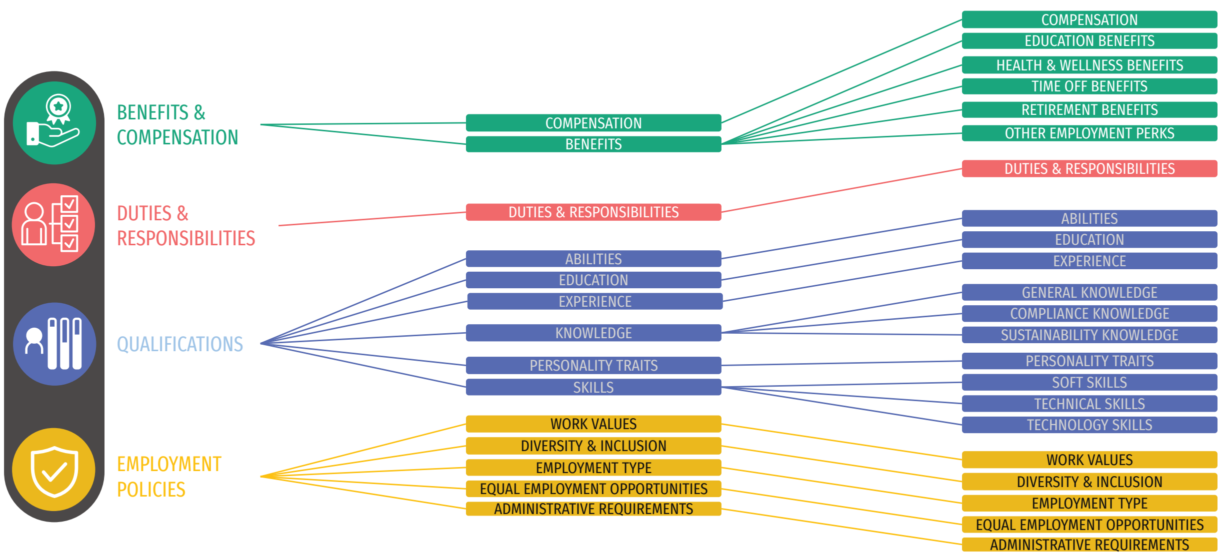


FIGURE 3: Jobs taxonomy hierarchy from less granular to more granular themes (left to right). *Source: RavenPack, October 2022*

3. Portfolio construction & backtesting

FIGURE 4 shows the company coverage of the U.S. mid/large universe over time. Since 2013, coverage has been strong at more than 70% of the universe. With this in mind, and the fact that the collection of job descriptions started in 2013, our strategies are based on out-of-sample backtests starting in April 2014 for the U.S. mid/large universe.[3] To remove potential duplicates of job postings within the same day, we only count job posting events with event similarity days and title similarity days greater than or equal to 1.0.

We construct long/short sector-neutral portfolios using the following methodology:

- Given a certain feature, we rank companies cross-sectionally and then select the top/bottom 20% for each sector.[4]
- Sector portfolios are equal-weighted, however, the relative weights between sectors are proportional to the number of stocks selected in each sector. For example, the allocation to a sector portfolio with 12 companies would be twice the allocation of a sector portfolio with only 6 companies.

All portfolios are sector and dollar neutral, with 50% long and 50% short exposure. Strategies are rebalanced daily to ensure path-independence. In subsequent sections, we investigate the impact of smoothing down weights over a range of windows to reduce turnover and assess the signal decay.

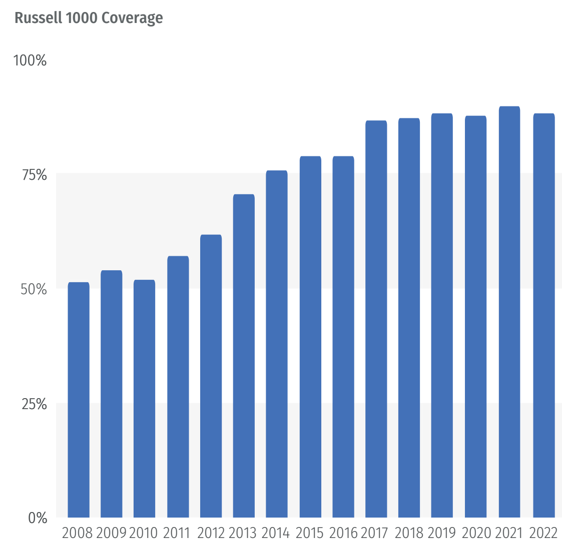


FIGURE 4: Coverage for U.S. mid/large-cap companies per year.
Source: RavenPack, October 2022

3 In some cases we need five quarters of a year for feature construction so even if we use data starting from January 2013 we are only able to build the strategies from April 2014.

4 Results are robust over similar threshold selection, as shown in Appendix A. The optimal selection, due to the balance between signal strength (for higher thresholds) and portfolio breath (for lower thresholds), may depend on the actual use case.

4. Hiring growth

In line with previous studies, we find that changes in the number of job postings are positively correlated with companies' future performance. In this section, we explore a simple strategy based on monthly growth of job postings within the U.S. mid/large-cap universe.

If the total number of jobs being posted by company k on each day t is $N_k(t)$, each company's monthly hiring growth on day t is calculated as the percentage change between the jobs posted over the last 30 days relative to the prior 30-day period :

$$Monthly\ Growth_k(t) = \frac{\sum_{\Delta=0}^{-29} N_k(t-\Delta) - \sum_{\Delta=-30}^{-59} N_k(t-\Delta)}{\sum_{\Delta=-30}^{-59} N_k(t-\Delta)} \quad (1)$$

Using the monthly growth formula above, we rank companies cross-sectionally each day. We then go long the top 20% and short the bottom 20% within each sector, constructing a sector-neutral portfolio with 50% long and 50% short exposure, as described in **SECTION 3**. **FIGURE 5** shows the cumulative log returns of the strategy from April 2014 to August 2022. Performance is robust, with Annualized Returns of 1.96% and an Information Ratio of 0.89.

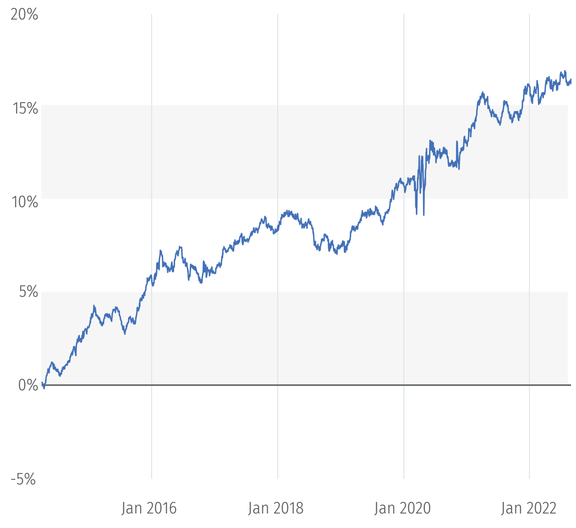


FIGURE 5: Cumulative log returns of the monthly hiring growth strategy for the U.S. mid/large-cap universe, from April 2014 through August 2022. *Source: RavenPack, October 2022*

4.1 Location enhancement

GUTIERREZ (2019) showed that the relationship between changes in the number of job postings and future performance is stronger when postings are associated with growth rather than replacement, because replacement is not supported by an increase in demand. Similarly, we argue that the relationship is stronger when growth is driven by demand, rather than the exploration of new markets – where investment returns are not guaranteed.

Companies with stable growing demand tend to show similar hiring growth across existing locations. By contrast, growth driven by exploration of new markets tends to increase the hiring in specific, new locations. We capture this concept by measuring the distributional change of the hiring location over time using the cosine similarity of the hiring location vectors. The location vector $\{N_{k,s}(t)\}_s$ stores the number of job postings for each location s at the state level[5] (if the company is not hiring in a given state, the corresponding vector dimension is equal to zero).

5 We look at hiring locations globally. We use States in the US and provinces in other countries.

To align with the monthly growth features we constructed earlier, we aggregate daily location hiring volumes at a monthly frequency, $NS_k(t)$:

$$\overline{NS}_k(t) \equiv \left\{ \sum_{\Delta=0}^{-29} N_{k,s}(t-\Delta) \right\}_s .$$

The cosine similarity between the current and previous month's location vectors is then calculated as the dot product divided by the modulus of each vector:

$$Location\ Similarity_k(t) = \frac{\overline{NS}_k(t) \cdot \overline{NS}_k(t-30)}{|\overline{NS}_k(t)| \times |\overline{NS}_k(t-30)|} \quad (2)$$

The location similarity is bound between zero and 1.0 due to the non-negative values in the location vectors. The similarity is higher when the hiring distribution across all locations exhibits smaller changes between the two periods. It is worth noting that changes in relative volumes between different locations have an impact on similarity scoring, while the overall shifts in hiring volume have no effect.

If our initial assumption stands, we should see a divergence in typical market behavior when conditioning hiring growth on location similarity. We analyze the market reaction by performing an event study whereby the initiation of long positions in companies within the top 20% hiring growth bracket mark the event.

FIGURE 6 shows the average excess return (in basis points) around the event dates (day zero) when splitting companies based on median location similarity within their respective sectors (top/bottom 50% similarity). The companies with higher location similarity clearly outperform those with lower similarity, consistent with our assumption.

Having verified our assumption, we seek to enhance the basic hiring growth strategy by double sorting on location similarity. For each sector, we keep the bottom 20% hiring growth leg the same as before, but only go long half the companies in the top 20% hiring growth leg — the half corresponding to the higher 50% location similarity (the red line in **FIGURE 6**). This means that the allocations of the remaining long positions will also increase since we want to ensure maximum exposure. The Annualized Returns increase more than 45%, improving from 1.96% to 2.85%, while the Information Ratio rises 20% to 1.09. **FIGURE 7** shows the comparison of cumulative log returns between the raw hiring strategy ("Raw Growth") and the strategy enhanced by location similarity ("Growth Enhanced by Location"), demonstrating consistent outperformance over time.

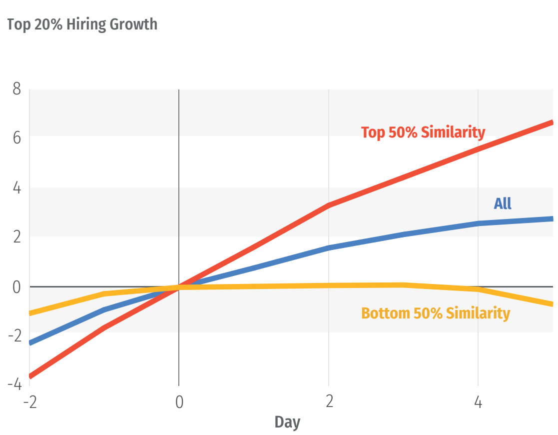


FIGURE 6: Average price reaction for the top 20% growth companies per sector, conditioned on the location similarity (top/bottom 50%) for U.S. mid/large-cap universe, from April 2014 through August 2022. *Source: RavenPack, October 2022*

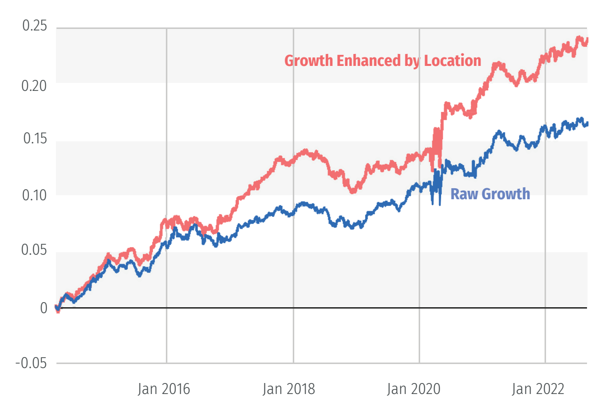


FIGURE 7: Cumulative log returns of the monthly hiring strategy with location similarity enhancement for the U.S. mid/large-cap universe, from April 2014 through August 2022. *Source: RavenPack, October 2022*

5. Most demanded positions

The growth strategy presented above is sector-neutral, so the performance of the overall portfolio can be viewed as a linear combination of the smaller sector dollar-neutral portfolios. **FIGURE 8** shows the Information Ratio of the basic hiring growth strategies across different sectors. Performance is quite varied, ranging from an Information Ratio of close to 1.0 in the best performing sector, to negative values in the worst performing sectors.

It is counterintuitive that certain sectors generate negative performance, suggesting that companies with greater hiring growth underperform on average (note that we are looking at relatively short timescales). However, there are many factors being ignored in this simplified interpretation which can impact sectors differently. For example, we are unable to differentiate real employee growth from turnover. CHUCHU (2022) found evidence that high administrative intensity is associated with lower employee turnover, but with lower future firm performance, which suggests that the composition of positions is also important.

Each sector is unique when it comes to hiring, and in many cases growth in preferred positions can deliver better performance than general hiring growth. For example, lawyers and judges are preferred by the Commercial Services sector, as they are closely related to profit generating units. Using only growth coming from these positions would result in a specific hiring growth strategy with an Information Ratio of 1.0, compared to the close-to-zero Information Ratio for the overall growth strategy (**FIGURE 8**). One could potentially select the most suitable candidates for each sector to enhance performance. However, in order to avoid manual selection, we propose a systematic way to select positions per sector over time that only looks at past information, even if this does not necessarily reflect the optimal choice all the time. We do so by focusing on the most demanded positions for each sector:

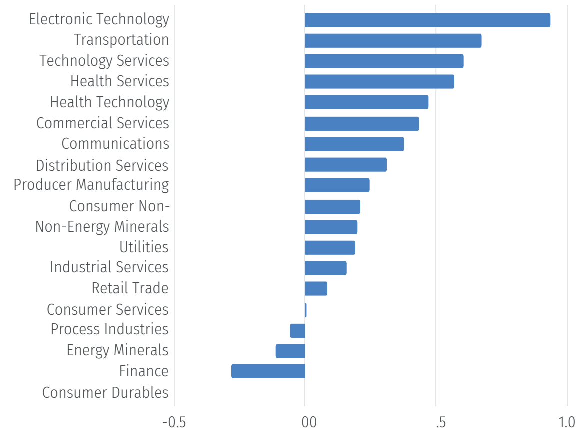


FIGURE 8: Information Ratio for raw hiring growth strategies by sector for the U.S. mid/large-cap universe, from April 2014 through August 2022. *Source: RavenPack, October 2022*

- We use the minor groups within the O*NET Occupation taxonomy as our target position clusters.
- We construct monthly hiring growth features for each of the minor groups (much like the Equation 1 but only counting job hiring for positions within each group).
- We build a hiring growth strategy for each of the position groups. This is sector neutral based on top/bottom 20% hiring growth.
- At the beginning of each year, we select the minor group that results in the largest average portfolio size in the prior year for each sector. We consider this the most demanded position within the sector. If we obtain multiple positions after the portfolio size filter, we choose the position with the highest Information Ratio that year.[6]
- We aggregate the selected strategies for each sector into a final portfolio.

6 Note that in many cases — except when we have multiple positions at the same level — we would not need to build a portfolio for the most demanded position selection, just a count of the number of companies that are hiring for that position group within each sector.

The final strategy is based on the daily top/bottom 20% company selection for hiring growth in specific positions for each sector. The strategy is sector-neutral and rebalanced daily. **FIGURE 9** shows the performance comparison between the overall hiring growth strategy ("Total Growth") and the hiring growth strategy selecting the most demanded position group ("Top position growth"). The most demanded positions strategy exhibits enhanced robustness over time and delivers a significant improvement in performance, with an Information Ratio of 1.43 (versus 0.89), and Annualized Returns of 3.01% (versus 1.96%).

In certain sectors, the top positions selected are quite stable. For example, "Computer Occupations" are consistently selected in the Technology Services sector while "Engineers" are always selected in Producer Manufacturing. "Financial Specialists" are selected about 90% of the time in the Finance sector. The strategy's outperformance relative to the overall hiring growth strategy is likely due to the most common positions being associated with core business operations.

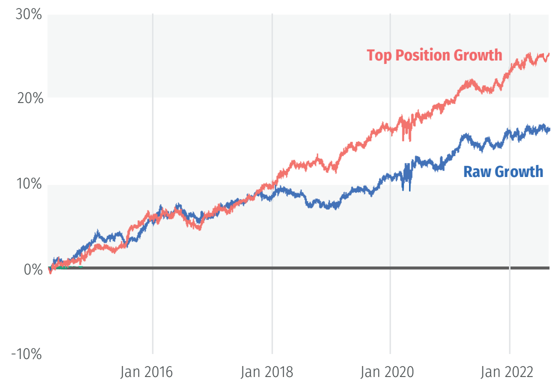


FIGURE 9: Cumulative log returns of most demanded hiring growth strategy for the U.S. mid/large-cap universe, from April 2014 through August 2022. *Source: RavenPack, October 2022*

6. Skills similarity

Hiring positions data can provide insights into business strategy. A shift in corporate strategy can result in a change to the employment structure and consequently lead to new hiring patterns that can be observed when analyzing positions data. For example, a company may shift from a focus on research and development to a focus on revenue growth and consequently increase its sales and marketing staff. Or, it may undergo structural changes requiring more HR-related and managerial roles. We may not always successfully identify these changes if the analysis is based on a specific position taxonomy, however. One reason for this is that job titles sometimes change while duties remain similar. For example, it is common to use trending job titles (e.g. "Data Scientist") that do not reflect the real nature of the position (e.g. a sales person with some technical knowledge). Therefore, job positions can cause a composition change in hiring but may not reflect a genuine shift in corporate strategy. Occasionally, they may also miss real changes. Here, we propose an alternative way to identify the hiring composition changes using skills detected in the text of job postings.

Ravenpack Job Analytics powered by LinkUp includes valuable information extracted from job descriptions, including the required skills, and therefore provides more detailed insights into the real nature of each position. In particular, we put our focus into the Soft Skills that we detect within the job description. We could also include hard skills, such as technology skills, in the analysis; however, we are aiming for a coarse job

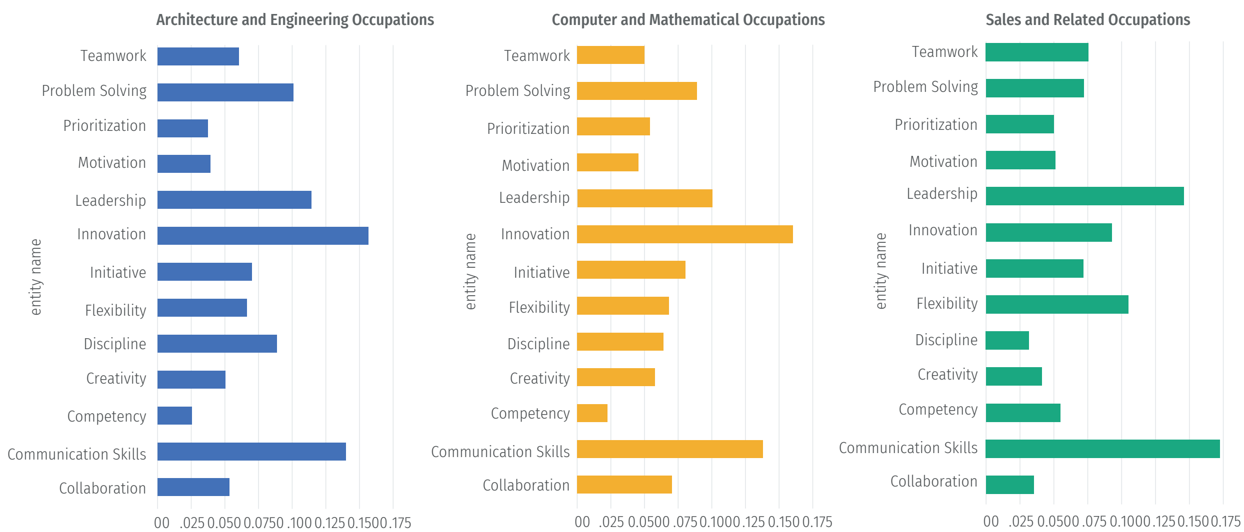


FIGURE 10: Distributions of top soft skills across three different position groups based on job postings from January 2022 through June 2022. *Source: RavenPack, October 2022*

classification that enables us to identify broad business changes and we want to avoid focusing on very specific skills. Companies often use new technologies, but this doesn't necessarily mean there has been any underlying business change (for example, new programming languages or cloud services becoming trending).

Soft skills reflect more general abilities such as innovation, communication skills, and leadership. By contrast, hard skills are more specific in nature (technical and technology skills). There are only around 100 soft skills, compared to 2,000 hard skills. However, even with a smaller, more stable composition, soft skills represent about 50% of the total volume of skills detected. Given their coverage and stability, soft skills are ideal to evaluate broad hiring position changes.

The most frequently mentioned soft skills are "communication skills" and "customer service". However, due to the different nature of positions, the composition of desired soft skills can be quite different. For example, the top soft skills within architecture and engineering are "innovation" and "communication skills", while in sales and related occupations they are "customer service" and "merchandising". These different distributions offer an alternative way to cluster positions and identify different hiring trends.

FIGURE 10 shows the distribution of top soft skills for three different position groups. The distributions of the first two groups (Architecture and Engineering, Computer and Mathematical) are quite similar, while the third group is meaningfully different. This is expected as the first two position groups are more research oriented, and therefore require innovation and problem solving skills. These skills are less represented in sales-related occupations.

We evaluate changes in the soft skills required over time using the same similarity metrics we used for locations (Equation 2). Given that we are assessing slower scale business changes, in line with the slower moving distributions of soft skills, we use quarterly aggregates for the soft skills vector and measure changes year over year (the most recent 90 days versus the same period in the previous year).

If $\{N_{k,f}(t)\}_f$ represents a vector of volumes for all soft skills f, where volume is the number of job postings from company k where we detected each skill on day t (if the company is not hiring for a given soft skill, the corresponding vector dimension is equal to zero); the quarterly aggregates are:

$$\overline{NF}_k(t) \equiv \left\{ \sum_{\Delta=0}^{-89} N_{k,f}(t-\Delta) \right\}_f.$$

then the cosine similarity between current and previous year's skills vector is calculated as:

$$Skills\ \ Similarity_k(t) = \frac{\overline{NF}_k(t) \cdot \overline{NF}_k(t-365)}{|\overline{NF}_k(t)| \times |\overline{NF}_k(t-365)|}$$

Based on the daily ranking of this similarity metric, we go long the top 20% companies and short the bottom 20% companies in each sector, constructing sector-neutral portfolios within the U.S. mid/large-cap universe. **FIGURE 11** shows the cumulative log returns of this strategy from April 2014 to August 2022. Annualized Returns are 1.88% while the Information Ratio is 0.86. Notably, turnover is only 3.72%, even with daily rebalancing, which leads to effective holding periods of more than one month.

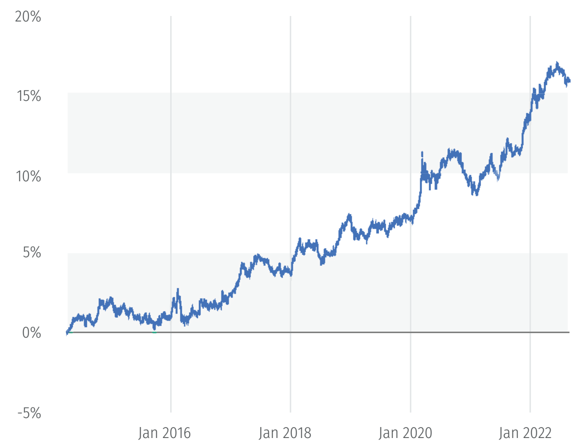


FIGURE 11: Cumulative log returns of the soft skills strategy for the U.S. mid/large-cap universe, from April 2014 through August 2022. *Source: RavenPack, October 2022*

7. Human capital & job requirements

SCHULTZ (1960) argues that economic growth can only occur if physical capital and human capital rise together, and that human capital is the factor most likely to limit growth. Corporations also require both physical capital and human capital in order to prosper. The modern human capital theory focuses on a worker's experience and skills, including factors such as education, training, intelligence, skills, and loyalty. OZTURK (2008) found that education raises both productivity and creativity, while promoting entrepreneurship and technological advances. Although it is difficult to accurately value the human capital of an organization, we try to build an approximation by assessing the requirements embedded within the job descriptions.

We find evidence that higher requirements for education is an indicator of better future performance:

- Companies with higher hiring growth in positions requiring a Bachelor's degree (or higher) outperform others within the same sector. A strategy that goes long the top 20% companies with higher growth in these positions and short the bottom 20%, measured month to month, yields a portfolio with Annualized Returns of 1.64% and an Information Ratio of 0.79.
- Additionally, hiring more PhD-level candidates is a positive signal. A strategy that goes long the top 20% companies with a higher percentage of PhD-level requirements over the previous month and shorts the bottom 20% delivers a portfolio with Annualized Returns of 1.85% and an Information Ratio of 0.76.

The education level only looks at the candidates' abilities from a single perspective. We could use other themes within the qualifications group in the RavenPack Jobs taxonomy to obtain a more complete overview of candidate requirements (qualifications contain six different themes, as shown in **FIGURE 3**).

If every single term under the qualifications topic represents a desirable asset, then the unique qualification count is an approximation of the overall requirement level for the candidates, and an estimation of the future employee value. Based on this assumption, we employ the following strategy:

- First, we count the unique terms within the qualifications taxonomy group for each company, using a monthly rolling window (30 days). This is an aggregate for all jobs being posted by that company over that period.
- We then construct a month-over-month change signal, as the net change in the current count relative to 30 days prior. A positive change means that the total qualification bar has increased compared to the past, as has future employee value.
- With this net count change, we go long the top 20% companies and short the bottom 20% within each sector, constructing a sector-neutral portfolio within the U.S. mid/large-cap universe.

Following these steps, the resulting portfolio delivers an Information Ratio of 1.07, with Annualized Returns of 2.01%, as depicted in **FIGURE 12**. The way we measure qualification growth is affected by the increases in hiring, as it is more likely that a growing company will also be hiring for a wider set of qualifications. Therefore, it is unsurprising that such a strategy is fairly correlated to raw hiring growth. There likely are better ways to disentangle these effects. For example, we could normalize the qualification count based on the total number of postings. However, we will leave this for future research.

FIGURE 12: Cumulative log returns of the qualifications change strategy for U.S. mid/large-cap universe, from April 2006 through August 2022. *Source: RavenPack, October 2022*

8. Combined strategy

So far we have introduced four main strategies using information on hiring growth, location similarity, most demanded positions, soft skills similarity, and unique change in qualifications. The corresponding cumulative log returns are shown in **FIGURE 13**. While the qualifications strategy exhibits similar performance to the hiring growth strategy due to similar aggregation windows and the analogous underlying nature between them, we find a relatively low correlation between the rest. Indeed, the soft skills similarity strategy exhibits moderately negative correlation with the others in recent years, especially during the 2020 coronavirus outbreak and the second half of 2021.

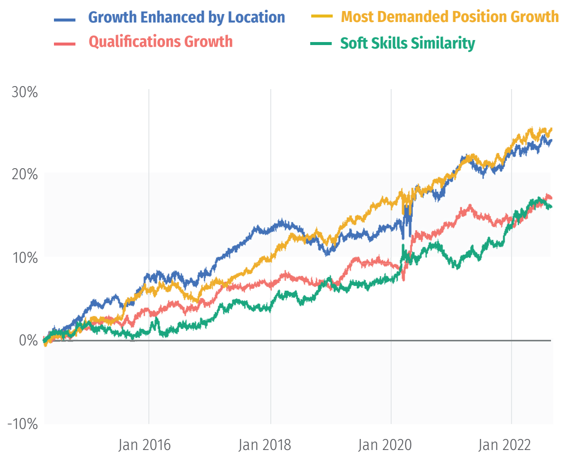


FIGURE 13: Cumulative log returns of all strategies for the U.S. mid/large-cap universe, from April 2014 through August 2022. *Source: RavenPack, October 2022*

Given the orthogonality between the different signals, we combine all strategies equally to form the final portfolio.[7] Predictably, the combined strategy has a higher Information Ratio of 1.67 along with Annualized Returns of 2.93%.

To further investigate the signal decay, we apply different moving average windows to constrain the turnover (applied directly to the final strategy allocations). **FIGURE 14** shows the cumulative log returns of the final strategy across different effective holding periods. The portfolios deliver consistent value over the whole period with a relatively low decay. **TABLE 1** summarizes the main performance statistics across different smoothing windows. With an effective holding period of two weeks, the strategy delivers an Information Ratio of 1.64 with Annualized Returns of 2.73%.

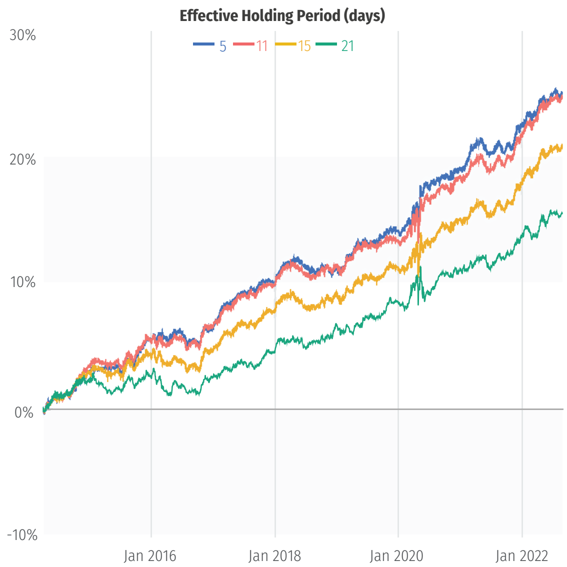


FIGURE 14: Cumulative log returns of the combined strategy across different smoothing windows, resulting in different holding periods. *Source: RavenPack, October 2022*

To examine the alpha profile of our strategies, we evaluate the portfolios using a traditional risk factor model.[8] As shown in **TABLE 2**, we observe low factor exposures when adjusting for the contribution attributed to traditional factors. The resulting adjusted performance is very similar to that of the original strategies

7 To do this, we simply average the allocations of the four strategies, followed by a daily cross-sectional rescaling to ensure 50% long and 50% short exposure (otherwise we may be slightly underallocated due to strategies sometimes trading in opposite directions).

8 We employ an in-house risk-factor model that uses an exponentially-weighted least squares dynamic regression of 10 common factors (Growth, Quality, Yield, Profitability, Investment, Market, Low Vol, Low Size, Momentum and Value). The raw vs adjusted comparison is carried out from December 2008 to September 2021.

across different trading horizons, demonstrating that RavenPack Job Analytics signals are a source of robust alpha generation (see more details on factor exposure in **APPENDIX B**). Moreover, the results indicate that controlling for traditional factors can improve performance by lowering volatility and thereby improving the Information Ratios.

Smoothing Window	Annualized Returns (%)	Information Ratio	Portfolio Size	Turn-over(%)	Effective Holding Period (days)
1	2.97	1.67	588	16.8	5
4	2.93	1.67	709	8.9	11
8	2.47	1.42	735	6.5	15
20	1.84	1.05	764	4.6	21

TABLE 1: Performance of final combined strategy with different smoothing windows, from April 2014 through August 2022.
Source: RavenPack, October 2022

Effective Holding Period (days)	Annualized Returns		Information Ratio	
	Raw	Factor-Adjusted	Raw	Factor-Adjusted
5	2.93%	2.46%	1.58	1.61
11	2.83%	2.23%	1.55	1.51
15	2.39%	1.92%	1.32	1.23
21	1.86%	1.36%	1.02	0.89

TABLE 2: Comparison of actual versus risk-adjusted performance for the final combined strategy across several holding periods. Results from April 2015 through August 2022.
Source: RavenPack, October 2022

9. Conclusions

In this study, we showcased the capabilities of RavenPack Job Analytics powered by LinkUp by demonstrating how to construct different strategies using information contained in job postings. The strategies exhibit robust performance across different trading horizons, ranging from days up to one month.

We began by building a simple strategy based on the monthly growth of job openings. We then showed how companies with higher growth that hire in similar locations outperformed their peers. The resulting Annualized Returns and Information Ratios increased from 1.96% to 2.85% and from 0.89 to 1.09 respectively.

Next, we proposed a systematic way to select the most demanded positions for each sector. With this, we constructed a more targeted hiring growth strategy, which generated Annualized Returns of 3.01% and an Information Ratio of 1.43.

We found that soft skills mentioned in job descriptions reveal valuable information in relation to the composition of hiring positions. Companies with an unstable distribution of soft skills – indicating changes in company structure or business strategy – underperformed sector peers. A trading strategy based on the distributional similarity of soft skills delivered an Information Ratio of 0.86, with effective holding periods longer than a month.

Education is important in human capital valuation. We found that companies with a larger number of job openings requiring PhD-level, or more hiring growth requesting Bachelor's degree or higher education levels, outperformed peers within the same sector. Both strategies delivered Information Ratios above 0.7. Using the full qualifications taxonomy, we were able to evaluate candidates' ability more comprehensively. We proposed a simple count of distinct qualifications as a way to reflect the quality of candidates and approximate a company's future employee value. The strategy based on the net change in qualifications delivered an Information Ratio of 1.07 and Annualized Returns of 2.01%.

Finally, we took advantage of the significant orthogonality between the different strategies when combined into a single portfolio. With an effective holding period of two weeks, the combined strategy delivered an Information Ratio of 1.67 and Annualized Returns of 2.93%, with low traditional factor exposures across all trading horizons.

Appendix A

In addition to the top/bottom 20% selection used for the paper results, we also test our strategies based on different top/bottom percentages (we keep a single threshold; however, we could also consider asymmetric behavior). Performance is robust across different percentiles. The optimal selection, given the signal strength (for higher thresholds) and portfolio breath (for lower thresholds) trade-offs, may depend on the actual use case.

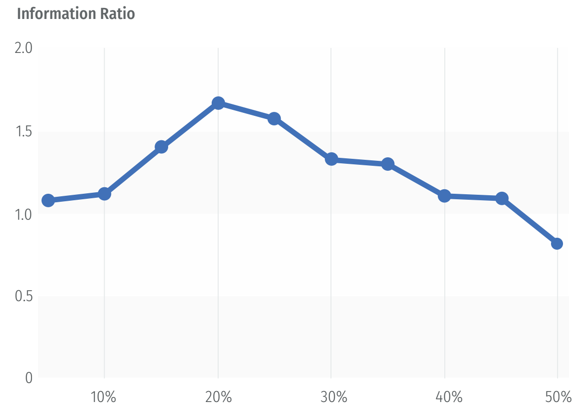


FIGURE 15: Information Ratio of the final combined strategy using different percentage thresholds when selecting top/ bottom companies. *Source: RavenPack, October 2022*

Appendix B

The final combined strategy exhibits low correlation across traditional factors.[9] **FIGURE 16** shows the levels of exposure over the 10 common factors included in our model across different effective holding periods.

The average exposure is mostly bound within the 10% range, with sporadic observed values of around 20% at the two-sigma level of significance.

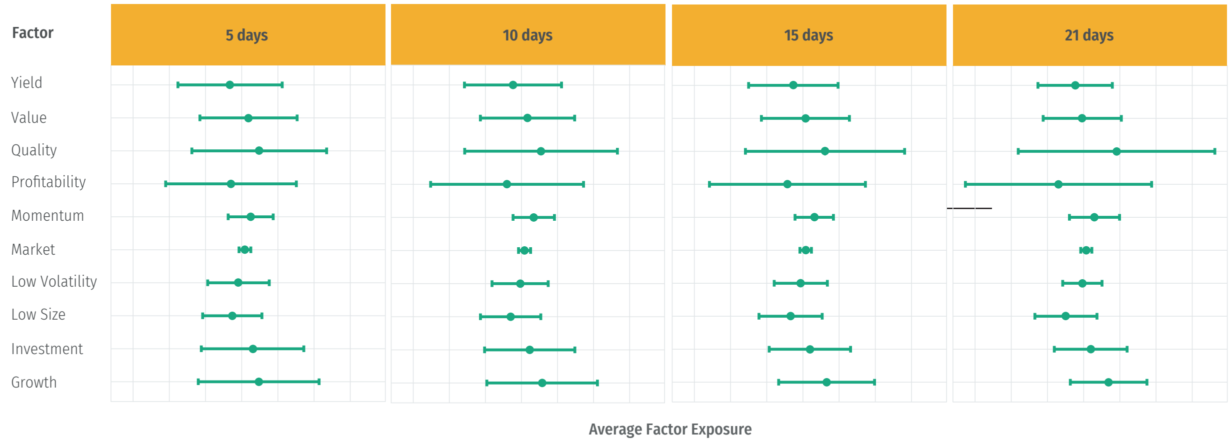


FIGURE 16: Average factor exposures of the final combined strategy across different effective holding periods. The bars indicate a two-sigma interval of confidence for daily exposures. *Source: RavenPack, October 2022*

9 We employ an in-house risk-factor model that uses an exponentially-weighted least squares dynamic regression of 10 common factors (Figure 16). The raw vs adjusted comparison is carried out from April 2015 to August 2022.

References

[1] ROUEN, ETHAN AND REGIER, MATTHIAS, *The Stock Market Value of Human Capital Creation* (March 3, 2020). Harvard Business School Accounting & Management Unit Working Paper No. 21-047, Available at SSRN: https://ssrn.com/abstract=3707710 or http://dx.doi.org/10.2139/ssrn.3707710

[2] GUTIERREZ, ELIZABETH F. AND LOURIE, BEN AND NEKRASOV, ALEXANDER AND SHEVLIN, TERRY J. AND SHEVLIN, TERRY J., *Are Online Job Postings Informative to Investors?* (July 16, 2019). Management Science, forthcoming., Available at SSRN: https://ssrn.com/abstract=3342669 or http://dx.doi.org/10.2139/ssrn.3342669

[3] LIANG, CHUCHU & LOURIE, BEN & NEKRASOV, ALEXANDER & XIA, SHIJUN. (2022). *Too Much of a Good Thing? Administrative Support Staff, Employee Turnover, Innovation, and Firm Performance.*

[4] GUTIÉRREZ, ELIZABETH & LOURIE, BEN & NEKRASOV, ALEXANDER & SHEVLIN, TERRY. (2020). *Are Online Job Postings Informative to Investors?.* Management Science. 66. 2801–3294. 10.1287/mnsc.2019.3450.

[5] OZTURK, ILHAN, *The Role of Education in Economic Development: A Theoretical Perspective.* Available at SSRN: https://ssrn.com/abstract=1137541 or http://dx.doi.org/10.2139/ssrn.1137541

Terms of Use

RavenPack is the category leader in Alternative Data for the corporate and financial sectors since 2003. We develop AI technology to process content at scale, including one of the largest training sets for NLP applications. Our clients include the most successful hedge funds, banks, and asset managers in the world.

For more information about RavenPack, including trials, please contact us at:

Sales Team: sales@ravenpack.com

Website: www.ravenpack.com

Technical Support: support@ravenpack.com

General enquiries: info@ravenpack.com

Americas: +1 646 277-7339
EMEA and APAC: +44 (0) 20 3290 7395