Alpha Products

# NFP Forecasting Using LinkUp's Jobs Dataset

April 2025

# Executive summary: NFP forecasting using LinkUp's jobs dataset

1) LinkUp's jobs datasets, sourced from company websites detects labor market trends in real-time, helping forecast US NFP [non-farm payroll]
2) Over the last 2 years, LinkUp's US NFP forecasts beat consensus [Reuters consensus] 65% of the time [see slide 11, 12], with lower forecast errors
3) Trading signals developed using LinkUp forecasts have generated an aggregate positive 1-day return average of 10% in FX (major currency pairs) on the day of NFP release, for the last 25 NFP releases, confirming the economic value of such dataset [see slide 13]
4) Both Long and Short FX trading signals achieve 67% success rate with avg success returns 2x failure returns. [see slide 13]
5) Asides from high top-level correlations, at sector level too, LinkUp jobs datasets tracks NFP sector dynamics closely [see slide 6, slide 7, slide 8]
6) Details of the data transformation, model selection and back test methodology and model statistics at the end [see slide 15 onwards]

# Team

## Toby Dayton

LinkUp CEO

**Toby Dayton** served as CEO of LinkUp since 2007 having joined the business in 2001. Now part of GlobalData following the acquisition in 2024, Toby's focus is growing GlobalData's Signals business, leveraging the company's wide array of proprietary, alternative datasets covering 20+ verticals and 10 horizontal datasets including LinkUp's flagship jobs dataset.

## Jon Norberg

LinkUp CSO

**Jon Norberg** has been Chief Strategy Officer at LinkUp since 2021, focused on developing the most valuable and dynamic data solutions for clients across the financial markets.

Prior to LinkUp, Jon held senior investment roles in institutions and family offices investing in hedge funds, private equity, and real estate

## Nic Gustafson

LinkUp Solutions Manager

**Nic Gustafson** is LinkUp's Solutions Manager, focused on maintaining a dataset of 300m+ job openings and working with clients to refine and implement data solutions.

Prior to LinkUp, Nic worked as an analyst at 3M.

## Adarsh Jain, CFA

Director, Financial Markets

**Adarsh Jain** CFA is Director of Financial Markets, focused on building deep insights and alpha generating products using vast suite of proprietary alternative datasets and AI/ML models

Prior to GlobalData he has worked at UBS, HSBC and D.E Shaw, in their equity and macro research teams

## Subhadri S. Mallik

Practice Head, Financial Markets

**Subhadri** is Practice head of Financial Markets, focused on developing machine-learning models to predict macroeconomic/sector/ company signals.

Prior to GlobalData he has worked at UBS and CRISIL in their research departments.

# About LinkUp's Jobs Dataset

| | | |
|---|---|---|
| **4,900,000**<br>ACTIVE JOBS | **40,000**<br>COMPANIES ACTIVELY HIRING | **8,500+**<br>PUBLIC COMPANY TICKERS |
| **300,000,000+**<br>HISTORICAL JOBS | **78,000+**<br>HISTORICAL COMPANIES | **195**<br>COUNTRIES |

## Distinct Competence

- **Real Time labor demand.** LinkUp delivers daily data on millions of job openings from tens of thousands of companies around the globe.
- **Historical data.** LinkUp's employment dataset is the job openings archive of record, dating to 2007, further back than any competitor.
- **Uncorrupted sourcing.** LinkUp scrapes job listings directly from company hiring websites, while our competitors source from job aggregators burdened with duplicate and expired listings.
- **Hierarchical view.** Our data aggregates at the macro, sector, company, and regional levels.

## Data Practices & Delivery

- **Exclusively scrape publicly available information from employer career pages.**
- **Scrape frequency: Tier 1 Companies (Public):** Every 24 hours | **Others:** Every 48 hours.
- **Data Latency:** Typically, T+1 days for Tier 1 Companies.
- **Maintain a scrape log** to track breaks/updates.

## Biases

- Data is overweight U.S.-based job listings (62% of active jobs).
- Data gaps:

  - Jobs that are not posted on company websites.

  - Internal job postings not visible to the public.

# Key LinkUp jobs indicators used in the study

## Key LinkUp Jobs indicators used in the study:

**Active Jobs:**

Total count of unique job listings that were available and posted for at least one day during a given month across all sectors.

This correlates very closely with US JOLTS [Job Openings and Labor Turnover Survey]: 90%+ correlation, and covers 55% of US of job openings aggregated under JOLTS

**Posted Jobs:**

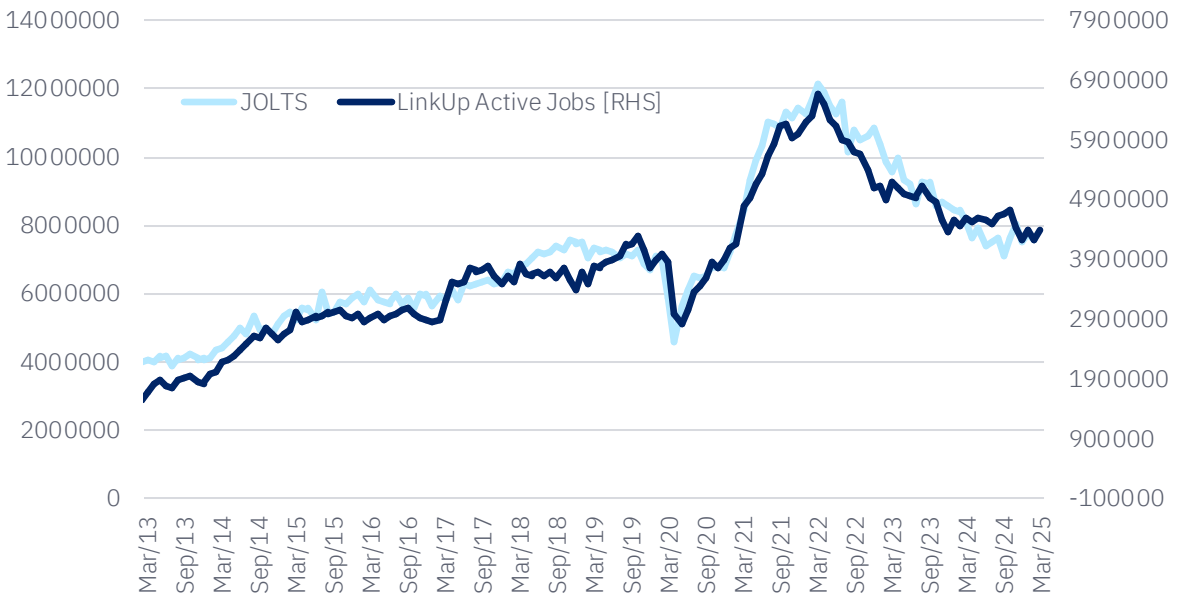
Job listings newly created during the given month. This shows labor demand and leads relationship with NFP, as job postings convert into job closures about 40-50 days out.

**Closed Jobs:**

Job listings removed during the given month. This is a close proxy for hiring done during the month, as job listings get removed post-hiring. A coincident indicator of corporate labor demand.

# LinkUp's Active Jobs indicator closely tracks U.S. Job Openings [JOLTS]

## Chart: LinkUp US Job openings Vs US JOLTS




Note: excluding govt jobs

## Key points

- Covers 55% of all US job vacancies
- Correlation with US JOLTS by 98%

**LinkUp Active Jobs:**

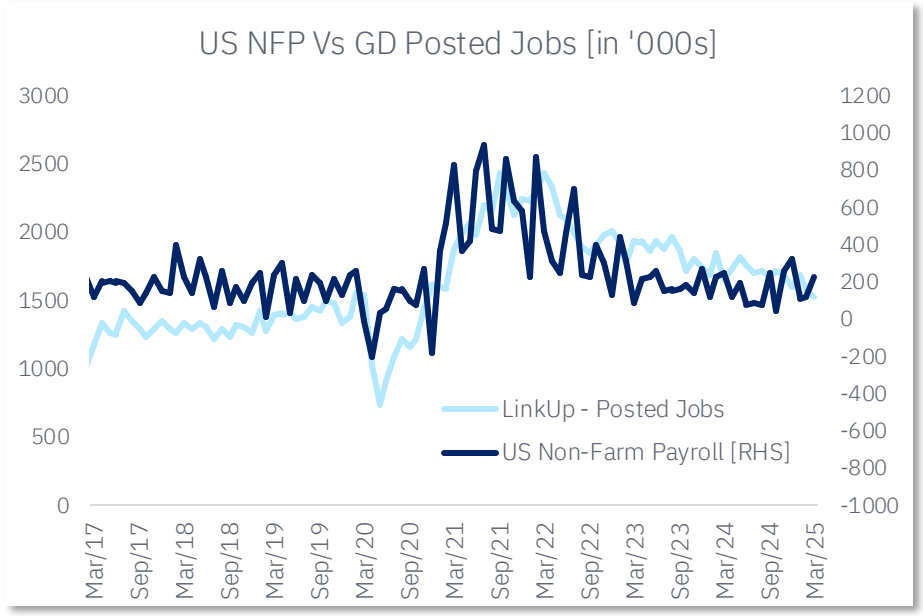
Total count of unique job listings that were available and posted for at least one day during a given month across all private sectors.

**JOLTS (Job Openings and Labor Turnover Survey):**

U.S. labor market indicator published by the **Bureau of Labor Statistics (BLS).** Measures job openings, via survey sent to ~21,000 business establishments in the US and provides snapshot of job opening trends across the US

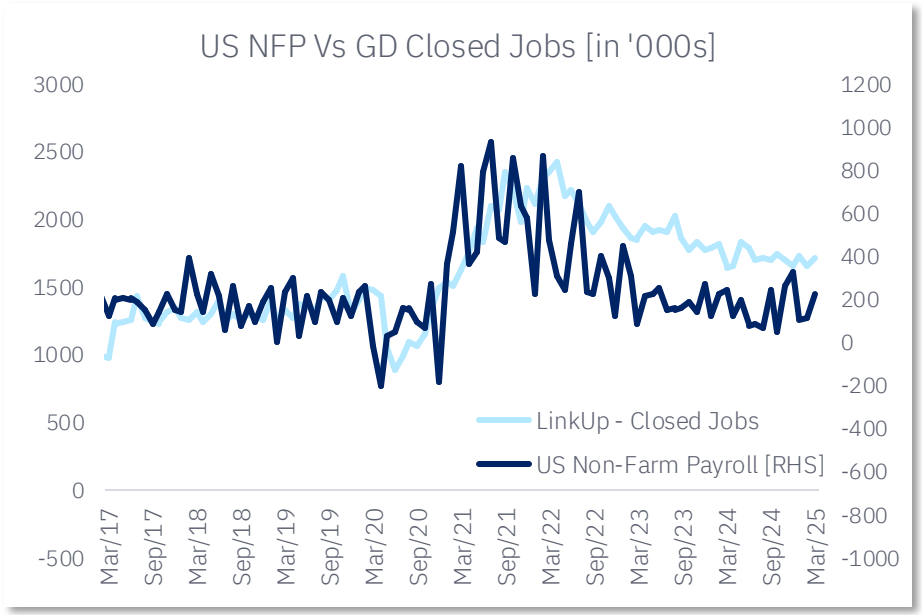
# Posted & Closed Job indicators track NFP data well, ahead of government release.

## Chart: US NFP Vs LinkUp Posted Jobs [in '000s]




**LinkUp Posted Jobs:** Job listings newly created during the given month.

## Chart: US NFP Vs LinkUp Closed Jobs [in '000s]




**LinkUp Closed Jobs:** Job listings removed during the given month.

**Note:** All series have been smoothed for extreme values during Covid

# Closed jobs also track NFP closely at the Sector Level


**Note:** All values in the chart are normalized using minmax transformation on 12-month moving average of NFP [Final release data] and LinkUp Closed jobs data

# Our sectoral distribution mirrors BLS reports

| % share of LinkUp closed Jobs and BLS, across sectors | | | |
|---|---|---|---|
| **Sectors** | **LinkUp Closed Jobs** | **BLS** | **Delta** |
| Healthcare | 17.3 | 13.8 | 3.5 |
| Retail | 16.7 | 9.4 | 7.3 |
| Accomodatin & Food | 9.4 | 8.6 | 0.8 |
| manufacturing | 8.5 | 7.7 | 0.8 |
| Public Admin | 8.1 | 18.3 | -10.2 |
| Professional services | 6.4 | 6.6 | -0.2 |
| Banking | 6.1 | 4.1 | 2 |
| Admin Support | 4.7 | 5.5 | -0.8 |
| Education | 4.1 | 2.4 | 1.7 |
| Technology | 3.7 | 1.8 | 1.9 |
| Other services | 3.4 | 3.6 | -0.2 |
| Unspecified | 2.5 | 0 | 2.5 |
| Transport | 2.1 | 4.1 | -2 |
| Wholesale Trade | 2.1 | 3.7 | -1.6 |
| Real Estate | 1.8 | 1.5 | 0.3 |
| Construction | 1.2 | 5 | -3.8 |
| Arts, Entertainment & Rec | 0.8 | 1.6 | -0.8 |
| Utilities | 0.5 | 0.4 | 0.1 |
| Oil & Gas and Mining | 0.2 | 0.4 | -0.2 |
| Mgmt. of companies | 0 | 1.6 | -1.6 |
| **Total** | **100** | **100** | **0** |

LinkUp overweight on Healthcare and Retail...

... and underweight on Public admin...

... But otherwise very close to sector composition of BLS

**Note:** BLS figures are proportion of the total Non-Farm Payrolls as of the latest release | LinkUp figures are average % shares over last 6 months

# Forecasting NFP using jobs dataset

# LinkUp forecasts beat consensus 65% of the time, with lower forecast error

Forecast success rate: 2 years

| | First Release |
|---|---|
| LinkUp | 65% |
| Consensus | 35% |
| Delta | 1.9x |

2x more beats

Avg Forecast errors: last 2 years

| | First Release |
|---|---|
| LinkUp | 68K |
| Consensus | 75K |
| Delta | 7K |

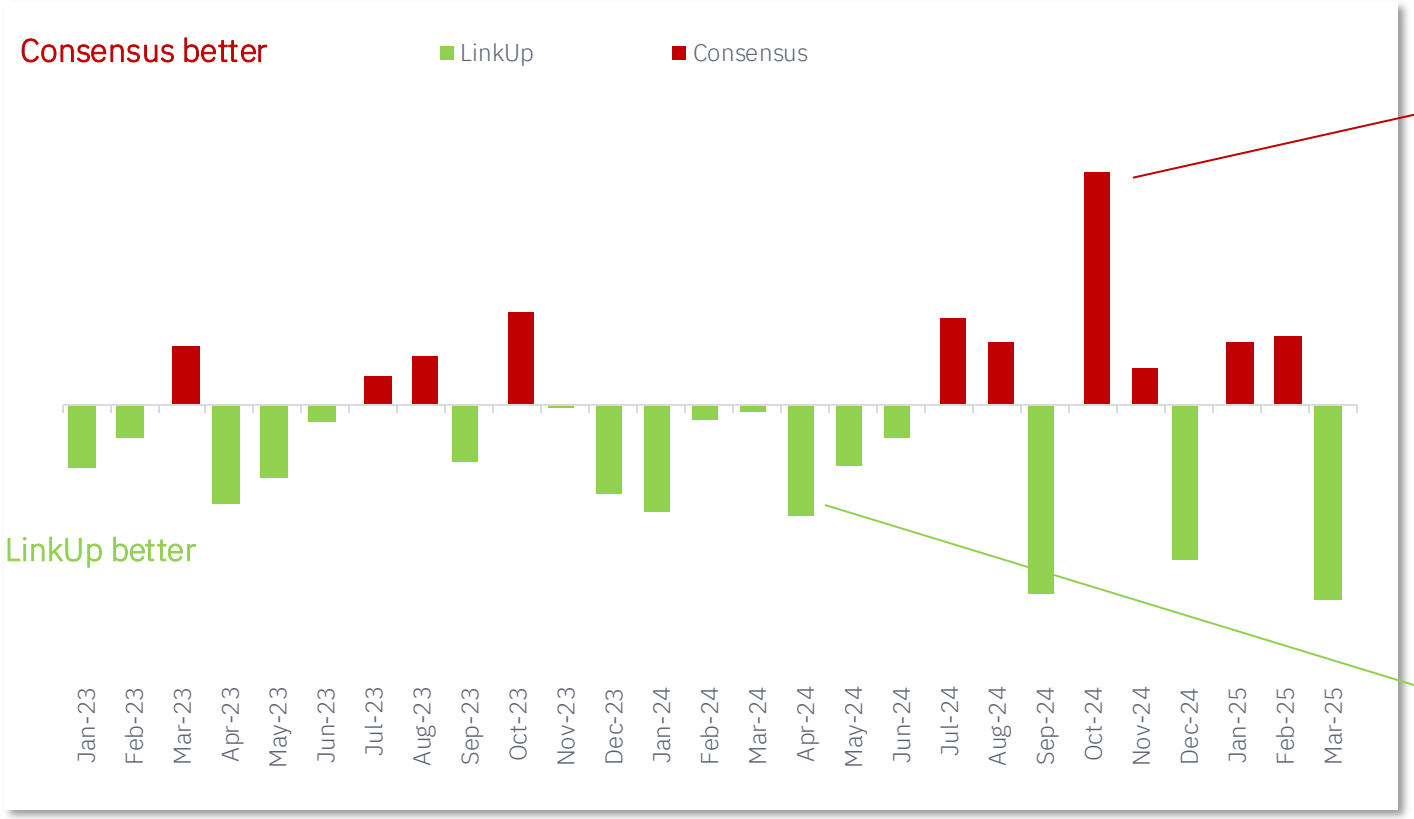
7K lower errors

**Note:** We have considered only the first release numbers which generally gets released in the first Friday of a month.

| Months | NFP First Release | NFP Final Release | LinkUp Forecast | Consensus | Best Forecaster |
|---|---|---|---|---|---|
| Jan-23 | 517 | 444 | 209 | 185 | LinkUp |
| Feb-23 | 311 | 306 | 218 | 205 | LinkUp |
| Mar-23 | 236 | 85 | 211 | 239 | Consensus |
| Apr-23 | 253 | 216 | 218 | 180 | LinkUp |
| May-23 | 339 | 227 | 218 | 190 | LinkUp |
| Jun-23 | 209 | 257 | 200 | 225 | LinkUp |
| Jul-23 | 187 | 148 | 211 | 200 | Consensus |
| Aug-23 | 187 | 157 | 222 | 170 | Consensus |
| Sep-23 | 336 | 158 | 192 | 170 | LinkUp |
| Oct-23 | 150 | 186 | 215 | 180 | Consensus |
| Nov-23 | 199 | 141 | 217 | 180 | LinkUp |
| Dec-23 | 216 | 269 | 204 | 170 | LinkUp |
| Jan-24 | 353 | 119 | 221 | 180 | LinkUp |
| Feb-24 | 275 | 222 | 206 | 200 | LinkUp |
| Mar-24 | 303 | 246 | 203 | 200 | LinkUp |
| Apr-24 | 175 | 118 | 201 | 243 | LinkUp |
| May-24 | 272 | 193 | 208 | 185 | LinkUp |
| Jun-24 | 206 | 87 | 209 | 190 | LinkUp |
| Jul-24 | 114 | 88 | 208 | 175 | Consensus |
| Aug-24 | 142 | 71 | 184 | 160 | Consensus |
| Sep-24 | 254 | 240 | 212 | 140 | LinkUp |
| Oct-24 | 12 | 44 | 201 | 113 | Consensus |
| Nov-24 | 227 | 261 | 186 | 200 | Consensus |
| Dec-24 | 256 | 323 | 219 | 160 | LinkUp |
| Jan-25 | 143 | 111 | 194 | 170 | Consensus |
| Feb-25 | 151 | 125 | 186 | 160 | Consensus |
| Mar-25 | 228 | 117 | 209 | 135 | LinkUp |

Note: on NFP first release

# Lower forecasting errors compared to consensus




Note: on NFP first release

Consensus better Example [Oct-2024]

Actual: 12k
LinkUp Forecast: 201k
Consensus: 113k

LinkUp Abs Error: 201k – 12k = 189k
Consensus Abs Error: 113k – 12k = 101k

LinkUp Error – Consensus Error = 189k – 101k = +88k

LinkUp better Example [Apr-2024]

Actual: 175k
LinkUp Forecast: 201k
Consensus: 243k

LinkUp Abs Error: 201k – 175k = 26k
Consensus Abs Error: 243k – 175k = 68k

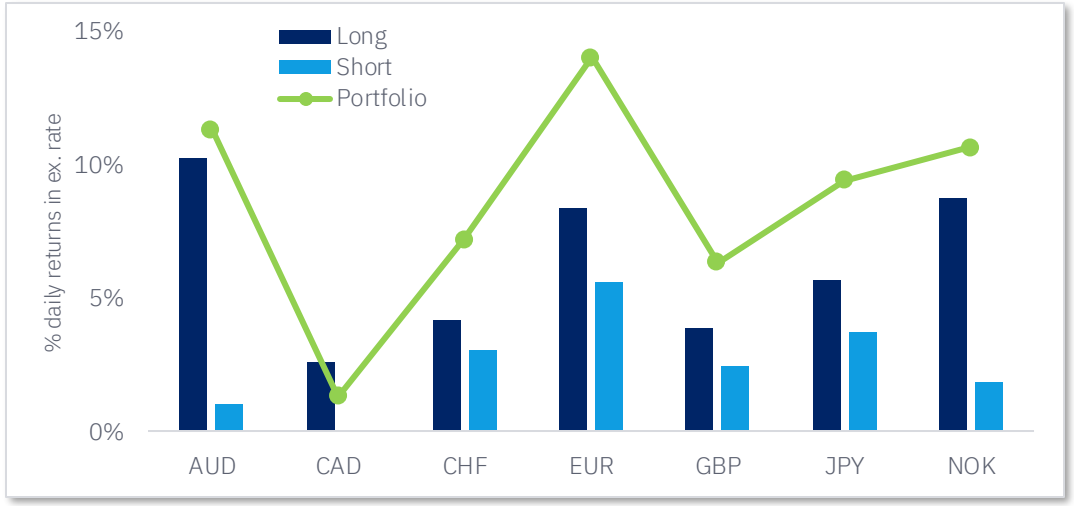
LinkUp Error – Consensus Error = 26k – 68k = +42k

**Note:** Delta is computed by taking difference between forecasts of LinkUp / Consensus and first NFP data release

# LinkUp Forecasts Generated positive 1-day alpha in FX and equity markets

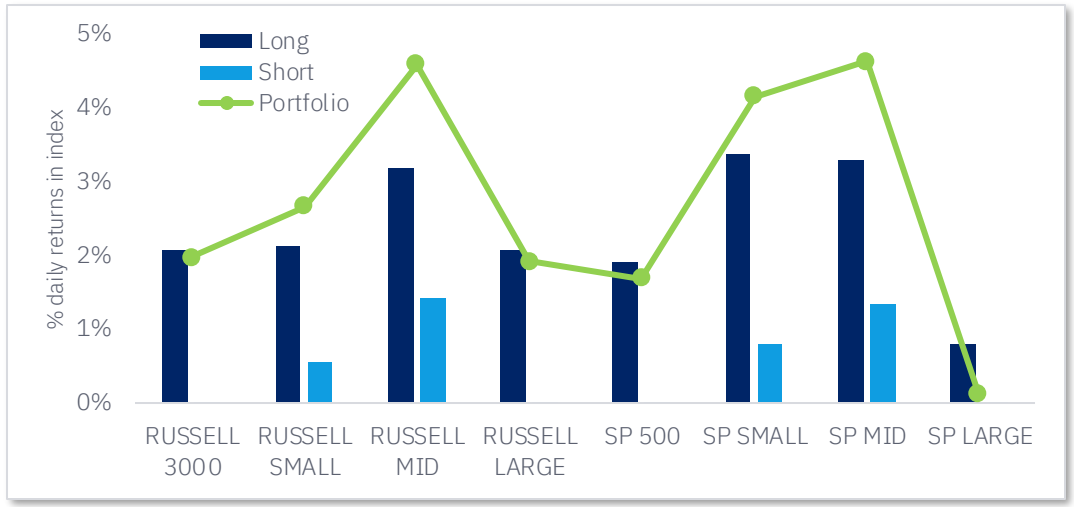
Long (Buy) when LinkUp forecast is >=10% above consensus
Short (Sell) when LinkUp forecast is 10% below consensus

## Aggregate 1-day FX return from all trading signals, across FX pairs




| | No. of signals | Signal counts % | Success % | Avg Success returns (bps) | Avg Failure returns (bps) |
|---|---|---|---|---|---|
| Long | 18 | 67% | 66.7% | 73 bps | 41 bps |
| Short | 9 | 33% | 66.7% | 52 bps | 25 bps |

## Aggregate 1-day Equity return from all trading signals, across market cap indices




| | No. of signals | Signal counts % | Success % | Avg Success returns (bps) | Avg Failure returns (bps) |
|---|---|---|---|---|---|
| Long | 18 | 65.4% | 64.7% | 99 bps | 140 bps |
| Short | 9 | 34.6% | 33.3% | 82 bps | 67 bps |



Note: on NFP first release

# Modeling approach / methodology

# Methodology:

- Explanatory variables: Initial unemployment claims (prior month), Posted jobs (and lags), Closed jobs (and lags), Recession outlier (dampen Covid period extremes)
- LinkUp Posted jobs and Closed jobs are taken from the 20th of the prior month to the 19th of the reference month. For example 2024-11-19 counts listings from 2024-10-20 to 2024-11-19. The series is ran on the 27th of the forecast month.
- Seasonally adjust LinkUp variables to match NFP series which is seasonally adjusted
- Normalise all variables using z-score to standardise difference in scales
- Training a **Ridge model** on NFP against seasonally adjusted features **[Model training between 2008 to 2022]**
- Tune and choose the optimal hyperparameter (lambda = 100) using cross-validation
- Produce out-of-sample month forecasts for the last 2 years

# Model specifications

## Model Performance

| Metric | Value |
|---|---|
| Regularization (Lambda) * | 100 |
| Condition Number ** | 15.9 |
| Test RMSE | 0.06 |
| Test MAE | 0.04 |

## Model coefficient

| Feature | Coef. | Weight (%)*** |
|---|---|---|
| Recession Dummy | - 0.06 | |
| Init. Claims (lag 1) | - 0.04 | 33% |
| Closed jobs (lag 0) | + 0.02 | 20% |
| Closed jobs (lag 1) | - 0.03 | 25% |
| Posted jobs (lag 2) | + 0.002 | 1% |
| Posted jobs (lag 3) | + 0.02 | 21% |

## Explanatory variables

**Initial Claims:**

US Department of Labor releases weekly initial claims and refers to the number of individuals who filed for unemployment benefits for the first time. These claims are considered a leading indicator of labor market health, with a negative correlation, i.e., an increase in initial claims indicates weaker future hiring trends.

**Closed Jobs:**

Job listings removed during the given month. This is a close proxy for hiring done during the month, as jobs listing get removed post-hiring. A co-incident indicator of corporate labor demand and has large weight in the first release model.

**Closed Jobs: Lags**

We note some mean reversion in closed jobs from the previous month, as a 1-month lag co-efficient has a negative sign, meaning closed jobs from previous month have downward adjustment from excess if any in previous month. However, at longer lag durations, it continues to have a positive effect.

**Posted Jobs:**

Job listings newly created during the given month. This shows latent corporate demand for labor and leads the relationship with NFP, as jobs postings convert into job closures over 2-3 month period.

**Posted Jobs: Lags**

Posted Jobs 2-month and 3-month lag is important driver of current month NFP, given the lead relationship as noted above.

**Recession Dummy:** Mainly used to dampen the extreme values during Covid

Notes:
* Lambda is the penalty factor that Ridge uses to shrink the coefficients; this value was computed from k-fold cross-validations and range of nos. between 0.00001 and 100
** Condition number is an indicator of multicollinearity; a 10-30 value suggests moderate collinearity still exists, but the model can be relied upon
*** Weights were calculated exc. Recession Dummy and absolute values were used and normalised on absolute coef. totals

# Why use Ridge Model?

A simple linear regression model of NFP regressed on LinkUp's jobs variables shows they all are statistically significant. However, since they are correlated, multi-collinearity renders model coefficients unstable.

Ideally, we want to make use of all LinkUp jobs variables, as lead/lags interact in complex ways to impact labor demand.

We therefore use the Ridge model, which helps reduce multi-collinearity by shrinking coefficients to stabilize estimates and reduce variance, making the model more robust while using all relevant variables.

And compared to more complex AI models, Ridge model coefficients are interpretable. While they don't produce sparse models, they do however provide a good balance between performance and interpretability

# Further research

1. Additional non-labor market variables like ISM, PMI and other demand parameters can be used to further improve the NFP forecasts.

2. Given LinkUp's coverage is global, in the future we will extend macro labor forecasts across G-10 countries

3. Our initial assessment and working hypothesis is that certain sectors (like Pharma for example) with specific job roles (business dev, M&A) carry leading information into future business performance and consequently share price performance. This is under study.