Alpha Score Full Report

Executive Summary

This report documents the final live-safe alpha score built from:

1. Frozen transcript sentiment model 69.
2. Frozen technical model.
3. A sector breadth overlay derived from the company alpha score.

The final company alpha score is:

```
sentiment_component = 100 / (1 + exp(-model_69_primary_score))

alpha_score =
    0.40 * sentiment_component
  + 0.60 * technical_raw_score
```

This formula is live-safe because it does not require knowing the percentile rank of a transcript among other companies reporting on the same day. It can be calculated as soon as a transcript has been scored and the latest technical score is available.

The strongest company-level backtest result is in the high raw-score bands. For the live-safe 40/60 alpha score, the 80-100 score bucket had average 12-month abnormal return of +84.2% and median abnormal return of +72.5% over the filtered validation sample. This result is promising but high-skew; it should not be interpreted as a repeatable expected return.

The sector breadth indicator is also useful. The best sector version is:

```
sector_breadth_80 = percentage of scored companies in a sector with alpha_score >= 80
```

The best trend-adjusted sector score is:

```
sector_score =
    50% * rank_across_sectors(sector_breadth_80)
  + 50% * rank_across_sectors(sector_breadth_80 - sector_prior_average_breadth_80)
```

This combined sector score had pooled Spearman correlation of +0.40 with forward sector/event abnormal returns and high-minus-low tercile spread of +11.5%.

Scope And Intended Use

The alpha score is intended to rank individual companies after earnings transcripts. It is not a complete portfolio construction system. It does not include position sizing, risk limits, transaction costs, liquidity constraints, borrow costs, or delisting corrections.

Recommended use:

- Company ranking after earnings/transcript events.
- Watchlist prioritization.
- Portfolio idea generation.
- Sector breadth confirmation.

Not recommended as-is:

- Standalone position sizing.
- Standalone sector ETF timing.
- A fully institutional backtest without additional survivorship, delisting, liquidity, and slippage controls.

Data Sources

Transcript Sentiment

Frozen model 69 files:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\model transfer
```

Primary frozen 3-year score file:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\model transfer\model_69_scores\model_69_scores_frozen_last_3y.csv
```

Important columns:

Column	Meaning
`symbol`	Company ticker.
`call_date`	Transcript date.
`entry_date`	Entry date used in return backtest.
`entry_close`	Entry close used in return backtest.
`abret_12m`	12-month stock return minus SPY return.

Column	Meaning
ret_12m	12-month absolute stock return.
russell3000_proxy	Whether event is in the current Russell proxy universe.
price_ge_5_at_entry	Whether entry close is at least $5.
model_69_domain	Whether event is in the frozen model scoring domain.
primary_score	Frozen post-sector-normalized raw sentiment score.

Technical Model

Technical model transfer folder:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer
```

Source-of-truth frozen model:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer\technical_frozen_model.pkl
```

Human-readable weights:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer\technical_frozen_model_weights.json
```

Production scoring script:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer\score_technical_frozen.py
```

Technical price data:

```
C:\Users\Admin\OneDrive\Cowork\Investing\data warehouse\prices
```

Universe:

```
C:\Users\Admin\OneDrive\GPT\wave\universe\universe_current.csv
```

Final Blended Outputs

Main final alpha event file:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\live_safe_sentiment_technical_blend_event_scores.csv
```

Main live-safe 40/60 score band result:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\live_sigmoid_blend_40s_60t_raw_band_returns.csv
```

Sector breadth files:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\sector_breadth_level_vs_trend_summary.csv
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\sector_breadth_level_vs_trend_tercile_returns.csv
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\sector_breadth_level_vs_trend_cohort_spreads.csv
```

Return Definition

For transcript-based alpha score tests:

1. Signal date is the transcript `call_date`.
2. Entry price is the next available trading-day close after the call date.
3. Exit price is the first available close at least 365 days after entry.
4. Absolute return:

```
ret_12m = exit_close / entry_close - 1
```

1. Abnormal return:

```
abret_12m = stock_ret_12m - SPY_ret_12m
```

Filters:

- Current Russell proxy only.
- Entry price at least $5.
- Event must be in model 69 scoring domain.
- Technical score must be available.
- 12-month return must be observable.

Final live-safe alpha validation sample:

```
2023-04-24 to 2025-04-14
15,969 events
2,156 symbols
```

Technical Score Construction

The technical model uses only price and volume information available on or before the technical score date.

The frozen technical production score is:

```
technical_raw_score = m13_build_winner_blend_raw
```

It is a 0-100 score and is calculated as:

```
m13_build_winner_blend_raw =
    mean(
      m01_12_1_momentum_score,
      m08_momentum_trend_score,
      m11_logit_top20_score,
      m12_ridge_abnormal_score
    )
```

The four components are:

```
m01_12_1_momentum_score
  = percentile_rank(z_mom_252_21)

m08_momentum_trend_score
  = percentile_rank(mean(
      z_mom_252_21,
      z_ret_126d,
      z_ret_63d,
      z_sma200_ratio,
      z_dist_52w_high
    ))

m11_logit_top20_score
  = percentile_rank(frozen logistic probability of top-20% 12M abnormal return)

m12_ridge_abnormal_score
  = percentile_rank(frozen ridge prediction of 12M abnormal return)
```

The technical model was trained/frozen using build observations through 2020-12-31. Validation signal dates are after 2020-12-31.

Frozen technical verification:

```
matched_rows: 51,124
max_abs_raw_diff: 1.42e-14
max_abs_percentile_diff: 1.42e-14
```

This means the frozen artifact reproduces the original technical score calculations to floating-point precision.

Technical Standalone Backtest

Standalone technical validation period:

```
2021-03-31 to 2025-03-31 signal dates
```

Validation quintile results:

Quintile	Events	Avg abnormal 12M	Avg absolute 12M	Median abnormal 12M	Win vs SPY
Q1	6,627	-12.1%	+1.9%	-14.1%	31.0%
Q2	6,637	-9.4%	+4.6%	-12.0%	33.6%
Q3	6,633	-4.8%	+9.1%	-11.6%	36.3%
Q4	6,636	-4.7%	+9.3%	-10.9%	37.5%
Q5	6,642	+2.2%	+16.2%	-10.9%	40.4%

Q5 minus Q1 abnormal spread:

```
+14.3%
```

Quarterly cohort robustness:

```
13 of 17 validation cohorts had positive Q5-Q1 spread.
Average cohort spread: +12.0%
Median cohort spread: +14.1%
```

Sentiment Score Construction

Model 69 produces a raw post-sector-normalized sentiment score:

```
model_69_primary_score
```

This score is z-like and can be positive or negative. For the live alpha score, it is mapped to a stable 0-100 scale using a logistic transform:

```
sentiment_component = 100 / (1 + exp(-model_69_primary_score))
```

This avoids same-day transcript percentile ranking. That matters because, in live use, not all companies have reported yet.

Examples:

Raw sentiment score	Sentiment component
-2.0	11.9
-1.0	26.9
0.0	50.0

Raw sentiment score	Sentiment component
+1.0	73.1
+2.0	88.1
+3.0	95.3

Final Company Alpha Score

Final production formula:

```
alpha_score =
    0.40 * sentiment_component
  + 0.60 * technical_raw_score
```

Where:

```
sentiment_component = 100 / (1 + exp(-model_69_primary_score))
technical_raw_score = m13_build_winner_blend_raw
```

The 40/60 weighting was selected because:

- Technical was the stronger standalone validation signal.
- Sentiment added event-specific information that pure price/volume cannot know.
- 40/60 had the strongest Q5-Q1 spread in the simple weight sweep.
- 50/50 had slightly higher Q5 average return, but 40/60 had better separation and win rate.

Weight sweep on validation transcript events:

Sentiment weight	Technical weight	Q5 abnormal	Q5-Q1 abnormal	Q5 absolute	Q5 win vs SPY
75%	25%	+0.7%	+13.4%	+21.0%	38.6%
60%	40%	+2.1%	+16.1%	+22.4%	40.0%
50%	50%	+2.5%	+17.2%	+22.9%	40.4%
40%	60%	+2.3%	+18.0%	+22.6%	40.6%
25%	75%	+1.8%	+17.9%	+22.1%	40.3%

The final live-safe formula uses the raw calibrated components, not same-day percentile ranks.

Final Company Alpha Backtest

Live-safe 40% sentiment / 60% technical score bands:

Alpha score band	Events	Symbols	Avg score	Avg abnormal 12M	Median abnormal 12M	Avg absolute 12M	Win vs SPY
0-20	44	38	18.3	-19.8%	-19.9%	-1.0%	13.6%
20-40	4,260	1,330	33.3	-15.8%	-16.9%	+3.9%	26.2%
40-60	8,326	1,978	49.2	-10.5%	-13.7%	+9.3%	33.4%
60-80	2,937	1,080	67.3	+6.7%	-7.5%	+26.5%	44.0%
80-100	402	208	86.2	+84.2%	+72.5%	+103.8%	69.4%

Interpretation:

- The score is strongly directional.
- Scores below 60 underperformed SPY on average.
- Scores from 60-80 were positive on average but had negative median abnormal returns.
- Scores above 80 were very strong but concentrated and high-skew.

Top-bucket concentration check:

```
80-100 bucket: 402 events / 208 symbols
Top 10 tickers contributed about 40% of total top-bucket abnormal return.
Winsorized top-bucket abnormal return: about +59%.
```

The top bucket is not purely one or two names, but the result is materially helped by large winners.

Sector Breadth Indicator

The company alpha score can also be aggregated into a sector signal.

The best simple sector indicator is:

```
sector_breadth_80 =
    count(companies in sector with alpha_score >= 80)
    / count(scored companies in sector)
```

Sector-quarter inclusion rule:

```
at least 20 scored company events in the sector-quarter
```

Pure sector breadth result:

Breadth tercile	Sector-quarters	Avg breadth >=80	Avg abnormal 12M	Median abnormal 12M	Avg absolute 12M	Sector win vs SPY
Low	32	0.4%	-13.3%	-14.3%	+6.5%	15.6%
Mid	24	1.9%	-7.7%	-9.1%	+12.2%	16.7%
High	32	3.9%	-4.3%	-2.4%	+15.6%	34.4%

High-minus-low abnormal spread:

```
+9.0%
```

Cohort robustness:

```
7 of 8 sector cohorts positive
median high-minus-low spread: +7.2%
```

Sector Breadth Level Versus Trend

Two sector breadth ideas were tested:

1. Compare sector breadth versus other sectors at the same time.
2. Compare sector breadth versus the same sector's prior history.

Pure trend was weak. The best result combined relative breadth and within-sector improvement:

```
sector_prior_average_breadth_80 =
    average of sector_breadth_80 in prior quarters only

sector_breadth_improvement =
    sector_breadth_80 - sector_prior_average_breadth_80

sector_score =
    0.50 * percentile_rank_across_sectors(sector_breadth_80)
  + 0.50 * percentile_rank_across_sectors(sector_breadth_improvement)
```

Comparison:

Sector signal	Spearman vs 12M abnormal	High-low spread	Positive cohorts
Cross-sector breadth level	+0.38	+9.0%	7 / 8
QoQ change in breadth	-0.04	+0.6%	5 / 7
Current breadth minus prior average	+0.24	+7.6%	6 / 7
Level + QoQ combo	+0.31	+10.0%	6 / 7

Sector signal	Spearman vs 12M abnormal	High-low spread	Positive cohorts
Level + prior-average combo	+0.40	+11.5%	7 / 7

Recommended sector breadth formula:

```
sector_score =
    50% * rank_across_sectors(% alpha_score >= 80)
  + 50% * rank_across_sectors(current % alpha_score >= 80 - sector prior average)
```

Sector score tercile results for the recommended level + prior-average combo:

Sector score tercile	Sector-quarters	Avg signal	Avg abnormal 12M	Median abnormal 12M	Avg absolute 12M	Sector win vs SPY
Low	28	27.3	-13.8%	-14.8%	+5.1%	14.3%
Mid	21	52.9	-9.2%	-9.9%	+9.8%	9.5%
High	28	83.0	-2.4%	-2.0%	+17.0%	39.3%

This is better than pure trend and modestly better than pure cross-sector breadth.

Robustness Assessment

What is encouraging:

- Technical standalone validated over 17 quarterly cohorts, with 13 positive Q5-Q1 spreads.
- Transcript-event 50/50 sentiment/technical blend had positive Q5-Q1 spread in 10 of 10 validation cohorts.
- The live-safe raw alpha score is strongly monotonic by raw score band.
- The frozen technical artifact reproduces the original scores exactly.
- Sector breadth has positive predictive value, especially when breadth is high versus other sectors and improving versus sector history.

Main concerns:

1. Current-universe survivorship bias.

The universe is based on a current Russell proxy. Historical Russell membership and delisted companies are not fully represented.

1. Iteration/data-snooping risk.

Many model versions were tested. The final formula is simpler and frozen, but the research path still benefited from seeing validation results.

1 Short final alpha validation window.

The live-safe blend validation sample covers 2023-04-24 to 2025-04-14. This was a strong momentum/growth/AI regime, which may favor the technical component.

1 Top-bucket skew.

Top alpha-score returns were helped by large winners such as APP, PLTR, ROOT, MSTR, CVNA, and similar names. Winsorized results remain positive, but the magnitude should be haircut.

1 Returns exclude full implementation frictions.

The test uses close-to-close returns. It does not fully model slippage, transaction costs, taxes, borrow constraints, liquidity caps, portfolio turnover, or dividends unless already embedded in the local price series.

Confidence

My confidence assessment:

Claim	Confidence
High alpha scores contain real predictive information	Medium-high
The score is useful for ranking post-earnings stocks	Medium-high
The exact historical return magnitude will repeat	Low
The alpha score can support portfolio idea generation	Medium
The score alone is enough for position sizing	Low
Sector breadth is useful as a tilt/confirmation signal	Medium

Practical confidence:

```
65-70% confidence that the alpha score contains real signal.
35-45% confidence that the backtested magnitude survives a fully institutional-quality test.
```

Reproduction Commands

Score Technical Model

```
python "C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer\score_technical_frozen.py" --score-date
2026-04-24
```

Recreate Technical Model Transfer Files

```
python "C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer\freeze_technical_model.py"
```

Existing Blend Files

The final live-safe blend event file already exists:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\live_safe_sentiment_technical_blend_event_scores.csv
```

The formulas used to produce it are documented in the appendices below.

Appendix A: Company Alpha Formula

Final live alpha formula:

```
sentiment_component = 100 / (1 + exp(-primary_score))

alpha_score =
    0.40 * sentiment_component
  + 0.60 * technical_raw_score
```

Where:

```
primary_score = frozen model 69 post-sector-normalized raw sentiment score
technical_raw_score = frozen technical m13_build_winner_blend_raw
```

Live interpretation:

Alpha score	Interpretation
< 40	Weak
40-60	Neutral/low conviction
60-80	Positive but mixed
80-90	Strong
90-95	Very strong
> 95	Exceptional, but likely high-skew/high-risk

Appendix B: Technical Feature Pseudocode

```
for symbol in universe:
    prices = load_price_history(symbol)
    prices = prices[prices.date <= score_date]

    if len(prices) < 253:
        skip symbol

    close = latest_close

    ret_5d = close / close_lag_5 - 1
    ret_21d = close / close_lag_21 - 1
    ret_63d = close / close_lag_63 - 1
    ret_126d = close / close_lag_126 - 1
    ret_252d = close / close_lag_252 - 1

    mom_252_21 = close_lag_21 / close_lag_252 - 1

    sma50_ratio = close / sma50 - 1
    sma200_ratio = close / sma200 - 1
    ma_cross = sma50 / sma200 - 1
    dist_52w_high = close / rolling_252d_high - 1

    vol21 = stdev(daily_returns_21d) * sqrt(252)
    vol63 = stdev(daily_returns_63d) * sqrt(252)
    vol252 = stdev(daily_returns_252d) * sqrt(252)

    volume_surge = avg_volume_21d / avg_volume_252d - 1
    dollar_volume_log = log(close * avg_volume_63d)
    rsi14_norm = (rsi14 - 50) / 50

    risk_adj_mom = mom_252_21 / vol252
    mom_volume_confirmed = mom_252_21 * (1 + clip(volume_surge, -0.5, 1.0))
```

Appendix C: Technical Scoring Pseudocode

```
panel = compute_raw_technical_features(universe, score_date)

for feature in raw_features:
    median = cross_section_median(panel[feature])
    mad = cross_section_median(abs(panel[feature] - median))

    if mad is valid and mad > 0:
        z = 0.6745 * (feature - median) / mad
    else:
        z = standard_z_score(feature)

    z = clip(z, -5, 5)
    z = z * feature_direction
    panel["z_" + feature] = z

m01_raw = z_mom_252_21
m01_score = percentile_rank(m01_raw)

m08_raw = mean(
    z_mom_252_21,
    z_ret_126d,
    z_ret_63d,
    z_sma200_ratio,
    z_dist_52w_high
)
m08_score = percentile_rank(m08_raw)

m11_raw = frozen_logit.predict_proba(z_features)[:, 1]
m11_score = percentile_rank(m11_raw)

m12_raw = frozen_ridge.predict(z_features)
m12_score = percentile_rank(m12_raw)

technical_raw_score = mean(m01_score, m08_score, m11_score, m12_score)
technical_percentile_score = percentile_rank(technical_raw_score)
```

For live alpha blending, use:

```
technical_raw_score
```

not `technical_percentile_score`.

Appendix D: Sentiment Scoring Pseudocode

```
transcript = load_transcript(symbol, call_date)
primary_score = frozen_model_69_score(transcript)

sentiment_component = 100 / (1 + exp(-primary_score))
```

Important:

```
Do not convert sentiment into a same-day reporting percentile for live use.
```

The live score must work when only one company has reported.

Appendix E: Alpha Blend Pseudocode

```
for transcript_event in scored_transcripts:
    symbol = transcript_event.symbol
    call_date = transcript_event.call_date

    primary_score = transcript_event.primary_score
    sentiment_component = 100 / (1 + exp(-primary_score))

    technical_row = latest_technical_score(
        symbol=symbol,
        score_date <= call_date
    )

    technical_component = technical_row.technical_raw_score

    alpha_score = (
        0.40 * sentiment_component
        + 0.60 * technical_component
    )
```

Appendix F: Sector Breadth Pseudocode

```
for each quarter:
    events = transcript_events_in_quarter
    events = events with valid alpha_score and sector

    for each sector:
        sector_events = events[events.sector == sector]

        if len(sector_events) < 20:
            skip sector-quarter

        sector_breadth_80 =
            count(sector_events.alpha_score >= 80) / len(sector_events)

        sector_prior_average_breadth_80 =
            average of this sector's sector_breadth_80
            over prior quarters only

        sector_breadth_improvement =
            sector_breadth_80 - sector_prior_average_breadth_80

    breadth_level_rank =
        percentile_rank_across_sectors(sector_breadth_80)

    breadth_improvement_rank =
        percentile_rank_across_sectors(sector_breadth_improvement)

    sector_score =
        0.50 * breadth_level_rank
        + 0.50 * breadth_improvement_rank
```

Appendix G: Backtest Pseudocode

```
for event in transcript_events:
    if not russell3000_proxy:
        skip

    if entry_close < 5:
        skip

    if model_69_domain is false:
        skip

    call_date = event.call_date
    entry_date = next_trading_day_after(call_date)
    entry_close = stock_close(entry_date)

    exit_date = first_trading_day_on_or_after(entry_date + 365 days)
    exit_close = stock_close(exit_date)

    stock_return = exit_close / entry_close - 1

    spy_entry = SPY_close(entry_date)
    spy_exit = SPY_close(exit_date)
    spy_return = spy_exit / spy_entry - 1

    abnormal_return = stock_return - spy_return

    alpha_score = compute_alpha_score(event)
```

Then group by raw score bands:

```
bands = [0, 20, 40, 60, 80, 100]

for band in bands:
    avg_abnormal_return = mean(abnormal_return)
    median_abnormal_return = median(abnormal_return)
    avg_absolute_return = mean(stock_return)
    win_rate_vs_spy = mean(abnormal_return > 0)
```

Appendix H: Key Output Files

Company alpha:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\live_safe_sentiment_technical_blend_event_scores.csv
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\live_sigmoid_blend_40s_60t_raw_band_returns.csv
```

Technical model transfer:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer\TECHNICAL_MODEL_HANDOFF.md
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer\technical_frozen_model.pkl
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\technical\model transfer\score_technical_frozen.py
```

Sector breadth:

```
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\sector_breadth_level_vs_trend_summary.csv
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\sector_breadth_level_vs_trend_tercile_returns.csv
C:\Users\Admin\OneDrive\GPT\Sentiment analysis\blended\sector_breadth_level_vs_trend_cohort_spreads.csv
```