Using Earnings Call Transcripts to Predict Stock Performance

. . .

Jonathan Khalfayan, Justin Kahl, Santiago Rodriguez, Matthias Schmitz, Sam Sklar, Juan Pablo Villarreal

Roadmap

1. **Obtain Earnings Call Transcripts**
2. Apply Natural Language Processing algorithm
3. Run regressions to weight keywords
4. Develop scoring algorithm to determine when to buy/sell
5. Run backtest evaluation of algorithm

Earnings Calls

Stock Criteria

- **Small**
  - Between the 25th and 75th percentile based on their market capitalization
- **Cheap**
  - Below 25th percentile EV to Ebitda ratio
- **Highly Leveraged**
  - Above mean LT Debt/EV

Binary Variables

LT debt year 1 > LT debt year 2

Prior year returns below median:
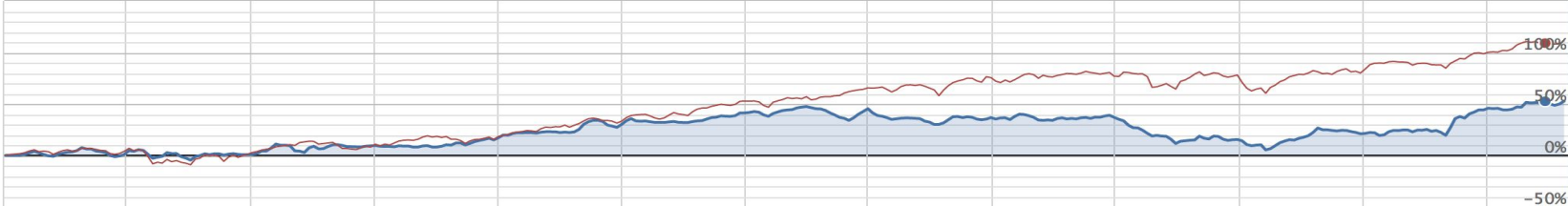
% Revenue Growth > % Asset Growth

High Gross Profit/Assets relative to LT Debt/Assets

Benchmark and Relative Performance

Total Returns	Benchmark Returns	Alpha	Beta	Sharpe	Sortino	Volatility	Max Drawdown
52.4%	111.5%	0.01	0.53	0.57	0.85	0.13	-30.4%

Cumulative performance: ■ **Algorithm** 52.52% ■ **Benchmark (SPY)** 109.02%   Week of Mar 26, 2017 | Week | Month | All |

Weekly returns $16,262

Transactions

Jan 2011 Jul 2011   Jan 2012   Jul 2012   Jan 2013   Jul 2013   Jan 2014   Jul 2014   Jan 2015   Jul 2015   Jan 2016   Jul 2016   Jan 2017

2012   2013   2014   2015   2016   2017

Data Collection

- Earnings call transcripts from THOMSON REUTERS, manually scraped

```
709  •FEBRUARY 09, 2017 / 4:00PM, CDE - Q4 2016 Coeur Mining Inc Earnings Call
710  CORPORATE PARTICIPANTS
711  Courtney Lynn Coeur Mining, Inc. - VP, IR & Treasurer
712  Mitch Krebs Coeur Mining, Inc. - President, CEO
713  Frank Hanagarne Coeur Mining, Inc. - SVP, COO
714  Hans Rasmussen Coeur Mining, Inc. - SVP, Exploration
715  Peter Mitchell Coeur Mining, Inc. - SVP, CFO
716
717  CONFERENCE CALL PARTICIPANTS
718  Joseph Reagor ROTH Capital Partners - Analyst
719  Chris Thompson Raymond James - Analyst
720  Mark Mihaljevic RBC Capital - Analyst
721  Chris Terry Deutsche Bank - Analyst
722  Jean-Paul Tsotsos BMO Capital - Analyst
723  Craig Johnston Scotia Bank - Analyst
724  Brett Levy Loop Capital - Analyst
725
726  PRESENTATION
727  Operator
728  Good morning, and welcome to the Coeur Mining, Inc., fourth quarter and yearend 2016 financial results conference call.
     (Operator Instructions)
729  Please note this event is being recorded.
730  I would now like to turn the conference over to Courtney Lynn. Please go ahead, ma'am.
731
732  Courtney Lynn - Coeur Mining, Inc. - VP, IR & Treasurer
733  Thank you, and good morning. Welcome to Coeur Mining's fourth quarter and full year 2016 earnings conference call. Our results
     were released
734  after yesterday's market close, and a copy of the press release and slides for today's call are available on our website.
735  Before we get started, I would like to remind everyone that our press release and some of our comments on the call include
     forward-looking
736  statements from which actual results may differ. Please review the cautionary statements included in our press release and
     presentation, as well
737  as the risk factors described in our 10-K.
738  I'll now turn it over to Mitch Krebs, President and Chief Executive Officer.
```

Data Collection

Next Steps:

1. Incorporate a larger number of equities in the dataset.
   a. Available extensive dataset still pending exploitation
2. Enhance current dataset
   a. Fill in gaps of non-existent data

Roadmap

1. Obtain Earnings Call Transcripts
2. **Apply Natural Language Processing algorithm**
3. Run regressions to weight keywords
4. Develop scoring algorithm to determine when to buy/sell
5. Run backtest evaluation of algorithm

Language Processing

Methodology:

1. Categorize words from .txt file into their corresponding fiscal quarter
2. Account for discrepancies in length of transcript, case, and punctuality
3. Count the frequency of specific words in each quarter
4. Write csv files of organized data for regression analysis

```
1  Quarter,cost,prudent,long-term,short-
   term,capital,reduce,revenue,efficient,synergies,growth,sales,marketing,save,structure,negotiate,pay,tax,rates,demand,supply,refina
   nce,refinancing,margin,turnover,enhance,reduction,lean,cutting,cut,yield,deleverage,dividend,value,repurchase,Number Of Words
2  [q1 2017],0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
3  [q4
   2016],192.215281115,0,57.6645843345,9.61076405574,144.161460836,19.2215281115,211.436809226,0,0,461.316674676,19.2215281115,0,9.61
   076405574,9.61076405574,0,0,38.443056223,76.8861124459,19.2215281115,0,0,86.4968765017,153.772224892,0,0,48.0538202787,0,0,0,0,0,0
   ,28.8322921672,48.0538202787,1
4  [q3
   2016],86.5005135968,0,21.6251283992,0,210.845001892,10.8125641996,129.750770395,10.8125641996,5.4062820998,367.627182786,10.812564
   1996,0,0,21.6251283992,0,5.4062820998,124.344488295,59.4691030978,10.8125641996,16.2188462994,0,10.8125641996,243.282694491,0,5.40
   62820998,32.4376925988,5.4062820998,0,0,0,0,0,81.094231497,21.6251283992,1
```

```python
import csv
listOfPhrases = []
listOfWords = []
previous = ""
label = ""
listOfQuarters = []
with open('IDT.txt','r') as f:
    for line in f:
        for word in line.split():
            word = word.lower()
            if word[len(word)-1] == ',' or word[len(word)-1] == '.':
                word = word[:len(word)-1]
            phrase = previous+" "+word
            if previous == "q1" or previous =="q2" or previous == "q3" or previous == "q4":
                if word[0]=="2" and word[1]=="0":
                    label = "["+previous+" "+word+"]"
                    if listOfQuarters.count(label) == 0:
                        listOfQuarters.append(label)
                        print(label)
            listOfPhrases.append(phrase+" "+label)
            listOfWords.append(word+" "+label)
            previous = word
```

Language Processing

Next Steps:

1. Account for different tenses of words.
   a. Potentially use global vectors for word representation (GloVe research group at Stanford)
   b. Multi-Dimensional Matrixes with similar words mapped at neighboring locations within the Matrix, word frequency calculated for set of words.
2. Investigate two word keyphrases such as "Pay Down"

Word Selection Criteria

- We started analyzing the frequencies of 26 keywords that could possibly predict future growth.
- The presence of words like "cost-cutting," "deleveraging," and "growth" that we believe will lead to these small, highly-leveraged, companies to produce higher returns than the benchmark

- For every company's earnings call transcript:
  - Frequency = # of appearances of keyword within specified fiscal quarter

Data

Screenshot table containing frequency values for one company for desired fiscal quarters and keywords:

Quarter	cost	prudent	long-term	short-term	capital	reduce	revenue	efficient	synergies	growth	sales	marketing	save	structure	negotiate	pay	tax	rates	demand	supply
[q1 2017]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[q4 2016]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[q3 2016]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[q2 2016]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[q1 2016]	0	0	0	0	1	3	1	0	0	4	36	1	0	0	0	0	0	0	3	0
[q4 2015]	0	0	0	0	0	0	2	0	0	7	27	2	0	0	0	0	0	0	2	0
[q3 2015]	1	0	1	0	0	0	0	0	0	3	41	2	0	1	0	0	0	0	5	1
[q2 2015]	0	0	0	0	2	0	3	0	0	18	32	0	0	1	0	0	0	0	2	3
[q1 2015]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
[q4 2014]	0	0	0	0	0	0	1	0	0	17	37	3	0	1	0	0	0	0	4	6
[q3 2014]	0	0	1	0	0	0	1	1	0	8	27	0	0	1	0	0	0	0	2	3
[q2 2014]	0	0	1	0	0	0	1	1	0	7	27	1	0	2	0	0	0	1	2	4
[q1 2014]	1	0	3	0	1	0	1	0	0	6	37	2	0	1	0	1	0	0	6	2
[q4 2013]	1	0	1	0	7	0	1	0	0	2	22	0	0	1	0	0	2	0	4	3
[q3 2013]	3	0	0	0	10	0	0	0	2	4	30	0	0	0	0	0	0	0	1	1
[q2 2013]	1	0	0	0	1	0	0	0	0	2	44	0	0	0	0	0	0	2	3	3
[q1 2013]	0	0	0	0	4	0	1	0	0	3	29	0	0	0	0	0	1	0	3	4
[q4 2012]	5	1	0	0	1	0	1	0	0	10	41	0	0	2	0	1	3	0	3	4
[q3 2012]	0	1	1	0	6	0	1	0	0	10	33	0	0	1	0	0	3	0	1	3
[q2 2012]	1	0	0	0	4	0	0	0	0	1	27	0	0	0	0	0	1	0	4	1
[q1 2012]	4	0	1	0	0	0	0	0	0	4	34	0	0	0	0	0	0	0	3	5
[q4 2011]	5	1	1	0	9	0	0	1	0	9	35	1	1	1	0	0	3	1	2	0
[q3 2011]	5	0	0	0	6	1	0	0	0	10	31	0	1	0	0	0	3	2	3	0
[q2 2011]	6	0	1	0	4	0	1	1	0	6	31	0	0	0	0	0	1	2	1	1
[q1 2011]	1	0	1	0	2	1	0	0	0	9	34	2	0	0	0	0	0	0	4	4
[q4 2010]	4	0	0	0	3	1	0	0	0	10	44	2	0	0	0	0	0	3	1	4

Roadmap

1. Obtain Earnings Call Transcripts
2. Apply Natural Language Processing algorithm
3. **Run analysis to weigh keywords**
4. Develop scoring algorithm to determine when to buy/sell
5. Run backtest evaluation of algorithm

Key Word Analysis

Word	Score	Number of Occurrences
Save	1.37	122
Dividend	.894	739
cutting	.805	92
cut	.772	473
short-term	.754	333
pay	.720	594
demand	.685	2053
prudent	.682	195
reduction	.657	1107

Next Steps

Roadmap

1. Obtain Earnings Call Transcripts
2. Apply Natural Language Processing algorithm
3. Run regressions to weight keywords
4. **Develop scoring algorithm to determine when to buy/sell**
5. Run backtest evaluation of algorithm

Algorithm

Insight: By using the correlations found in the Earnings Call Transcript analysis, we can score companies by quarter as new transcripts are released.

Methodology:

- When high frequencies of key terms occur (cost cutting, dividend, reduce, save) increment score.
- Add additional binary values for whether long term debt has decreased from last year (signalling debt paydown) as well as increasing Gross Profit/Assets ratio.
- Buy companies with highest scores that also meet binary metrics. Try on Quarterly vs Yearly holding period

Roadmap

1. Obtain Earnings Call Transcripts
2. Apply Natural Language Processing algorithm
3. Run regressions to weight keywords
4. Develop scoring algorithm to determine when to buy/sell
5. **Run backtest evaluation of algorithm**

Stock Performance

- To prevent look-ahead bias, we will trade with one quarter lag.
- For instance, we will act upon data obtained from Q1 2010 at the beginning of Q3 2010.

Hypothesis and Predictions

- Investing in small, cheap, highly-leveraged stocks with improving asset turnover and estimated debt paydown that we filter by keywords such as "cost-cutting" will lead to a strategy that will outperform the S&P 500 Index.

Thank You!