

LGO Edge Fund

Jeremy Hare and Shai Dekel





Project Overview

PIOTROSKI MODEL RETURN:

23%

1. Interpretable
2. Easy Data Access
3. Underutilized!

**High F-Scores = Higher Returns
Low F-Scores = Higher Losses**



F-Score Details



Profitability

1. Return On Assets	ROA (Income before Extraordinary items/Assets)
2. Operating Cash Flow	CFO (Cash Flow from Operations)
3. Change in ROA	ROA YoY
4. Accruals	CFO/Assets vs ROA

Leverage

5. Change in Leverage	Long-Term Debt / <u>Avg Total Assets YoY</u> (negative)
6. Change in Current ratio	Current Assets / Current Liabilities YoY
7. Number of shares Issued	Common shares issued



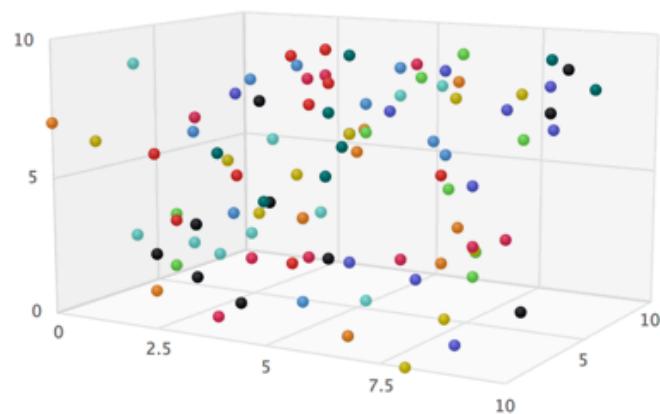
Operating Efficiency

8. Change in Gross margin	Sales – Cost of Goods Sold / Sales YoY
9. Change in Asset Turnover Ratio	Total Sales / Total Assets YoY



Architecture Overview

User Interface



Database

Focused on S&P500 companies only
over 5,100 financial and stock observations over last 11 years (2005-2016)

- Basic Company Info
- Annual Financial Info
- Stock Return Info

Analysis

Links User Interface to Database and outputs information based on the inputs from user

Linear Regression Model includes:
A) Piotroski's standard method
B) Jeremy/Shai evaluation method
C) Different holding strategies



Data Mining

- Compustat - North America
 - Financial Release Data – All companies in S&P500
 - Fiscal Year Financials Between 2005-2016
 - ~10,000 observations
- Center for Research in Security Prices (CRSP)
 - Monthly stock and total S&P500 performance
 - ~100,000 observations

Linking the two databases (CSV files) took some computing power and patience



Data Scrubbing

Started with a process and set up an algorithm

Excluded data that met any of the following criteria:

- Missing data from Compustat or CRSP
- Non-continuous data (year-to-year)
- First/Last year in database (no F-Score/return info)
- Removed anomalies that didn't match reality (e.g. 40000% stock return)

With clean data - F-Score ratios were calculated for each stock observation and linked to the stock performance a year later



Analysis

Split data by size of firm (Revenue based)

Split by F-Score (low - 0-2, mid 3-6, high 7-9)

Linear Regression in JavaScript

Ran additional analysis in R (for fun):

Linear Regression

Logistic Regression

CART

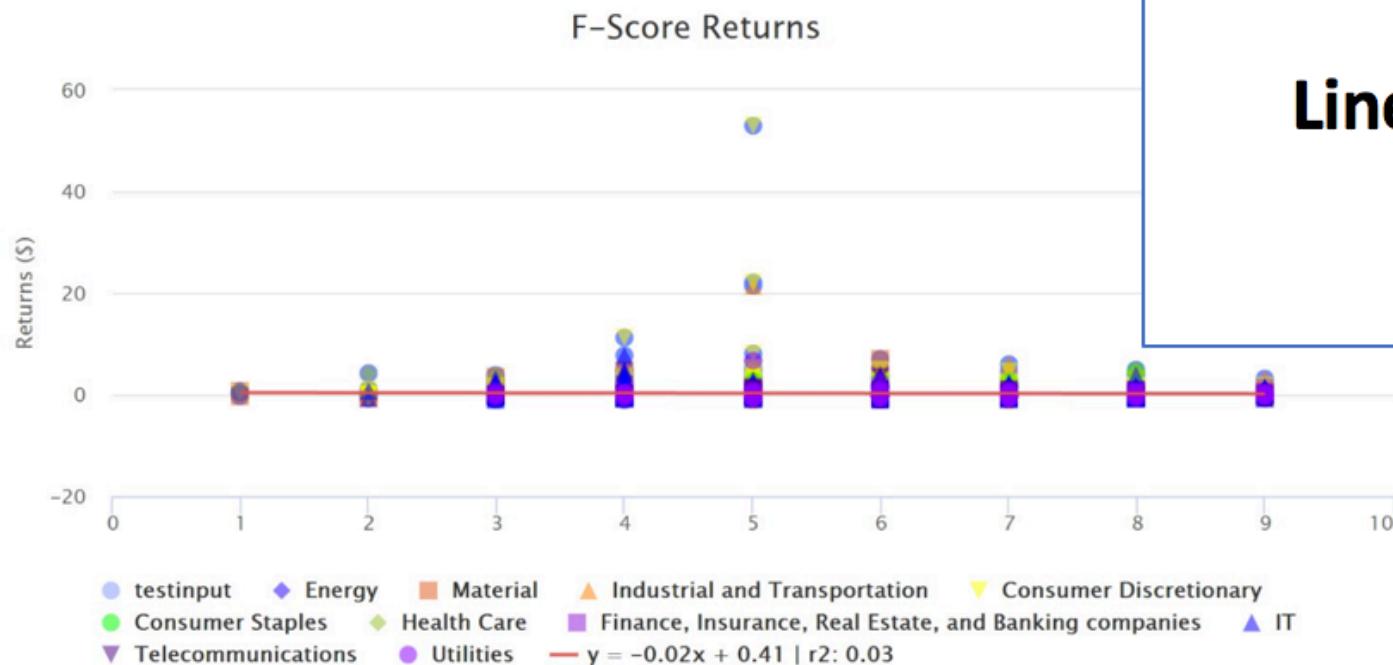
Random Forest



F-Score vs. Returns

Little Correlation between F-Score and Returns

**Linear Regression on Test Set =
 R^2 of 0.03**





And Not Significant...

Companies with almost the exact same F-Score metrics had drastically different returns....

	ROA	CFO	ROA Change	Accruals	Leverage	Current Ratio	Gross Margin	Asset Turnover	Shares	F-Score	Return	Positive Return	Number in Cluster
1	89%	99%	26%	92%	37%	46%	43%	33%	88%	5.50	35%	100%	1,463
2	88%	98%	31%	90%	37%	44%	43%	38%	87%	5.60	-24%	0%	824

Companies with the best F-Score and the worst F-Score had very similar returns....

	ROA	CFO	ROA Change	Accruals	Leverage	Current Ratio	Gross Margin	Asset Turnover	Shares	F-Score	Return	Positive Return	Number in Cluster
3	59%	86%	9%	82%	12%	20%	17%	18%	63%	3.70	14%	61%	593
4	99%	100%	96%	97%	83%	74%	89%	85%	97%	8.20	13%	64%	1,028



Results



We could not reproduce Piotroski's results

Found no statistical significance of F-Score to returns

- Insult to injury - Return of the low F-Scores was higher than high F-scores (opposite of what we wanted to see)

No predictive power for CART / Random Forest



Key Takeaways

We think we know what happened...

S&P500 is 80% of the market cap of publicly traded U.S. Companies

- *Stock performance may be less dependent on financial results*
- *May already be built into price expectations for year*

“...Differentiation is weak among the largest firms ... Thus, the improvement in returns is isolated to firms in the bottom two-thirds of market capitalization.” - J. Piotroski

Next steps is to analyze the data on the companies in the lower two-thirds of traded market cap, rather than the top 10%!