ANALYSIS OF U.S. STOCK MARKET FLUCTUATION IN 2020

Phoenix Wang

01
OVERVIEW

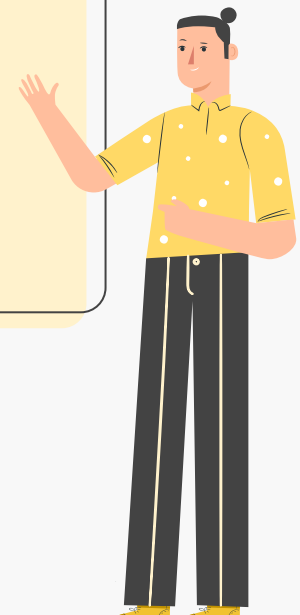
02
DATA
PREPARATION

03
FLUCTUATION
ANALYSIS

04
CONCLUSIONS

OBJECTIVES

- Explain the variation in U.S. stock returns in 2020
  - Initial COVID shock (January to March)
  - Market recovery (April to December)
- Determine the characteristics of
  - The best-performing stocks
  - The worst-performing stocks



02.
DATA PREPARATION

Types of Data
Sources of Data
Cleaning and Preprocessing

DATA SOURCES

WRDS COMPUSTAT

- Accounting data
- Entire database
- Fiscal Year 2018 & Fiscal Year 2019

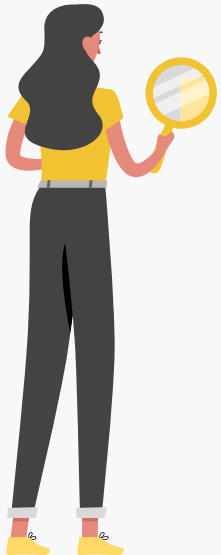
WRDS CRSP

- Stock & risk data
- Entire database
- January 2019 - December 2020

YAHOO FINANCE

- Stock data
- Business descriptions
- Ad hoc searches

PRE-PROCESSING STOCK DATA



LACK OF DATA

Drop observations that are

- Missing valid monthly returns for all 12 months in 2020
- Missing tickers

EXCESSIVE DATA

For 19 stocks with two sets of returns,

- Cross examine prices (with YFinance)
- Drop the set that does not match

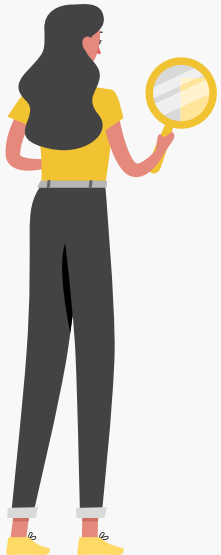
RETURN = 0

- Spot check the prices of stocks with monthly returns of 0

MISSING BETAS

- Replace with the average beta of each SIC industry in 2019

PRE-PROCESSING ACCOUNTING DATA



NO STOCK DATA

Drop companies that

- Do not have stock data in 2020
- Do not have accounting data for Fiscal Year 2018 & 2019

REVENUE = 0

- Drop companies that had no revenue in 2019

CLASSIFICATIONS

- **GIC Standard:**
  11 sectors, 24 groups, 69 industries
- **NAICS code:** 20 sectors
- **SIC code:** 12 groups

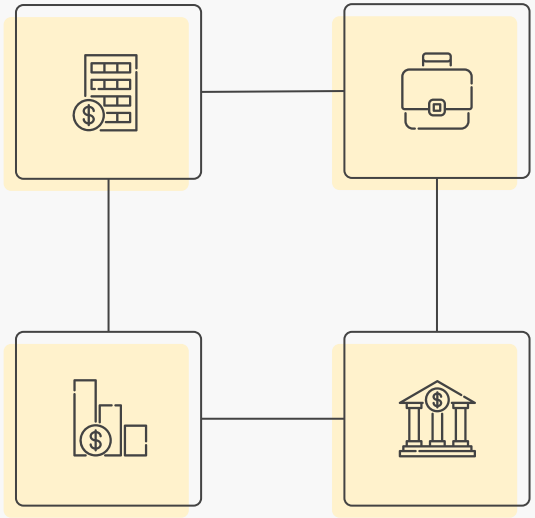
FINANCIAL RATIOS

- Use mostly data from 2019
- Employ data from 2018 to calculate average Assets & average Equity

MAIN DATAFRAME

OVERVIEW

4,130 unique tickers
4,044 business descriptions



INDUSTRY INDICATORS

GICS (11-24-69)
NAICS (20)
SIC (10)

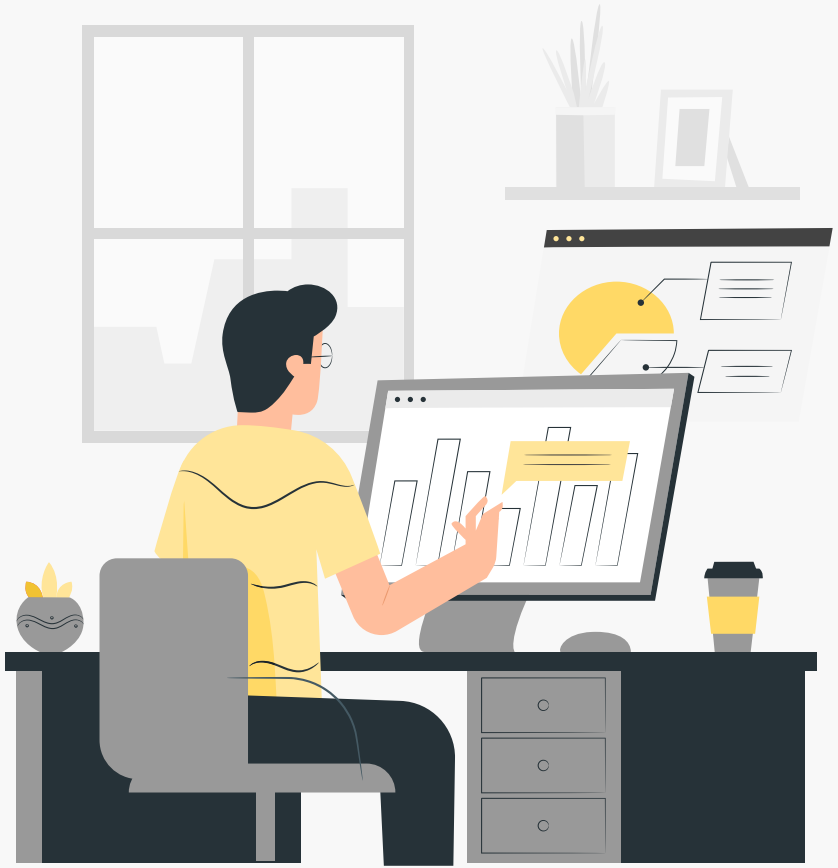
STOCK DATA

3-month return (Jan-Mar 2020)
9-month return (Apr-Dec 2020)
Market beta (2019)

ACCOUNTING DATA

15 financial ratios

- ROA, ROS, ROE, Asset Turnover, Gross Margin, Cash to Asset, Quick Ratio, Sales per Employee...



03. FLUCTUATION ANALYSIS

What are the average returns for each industry?
Do the highest/lowest return industries make economic sense?
What is the explanatory power of industry-fixed effect regressions?

01. Does Industry Explain the Differences in Returns?

Fluctuation During COVID Shock

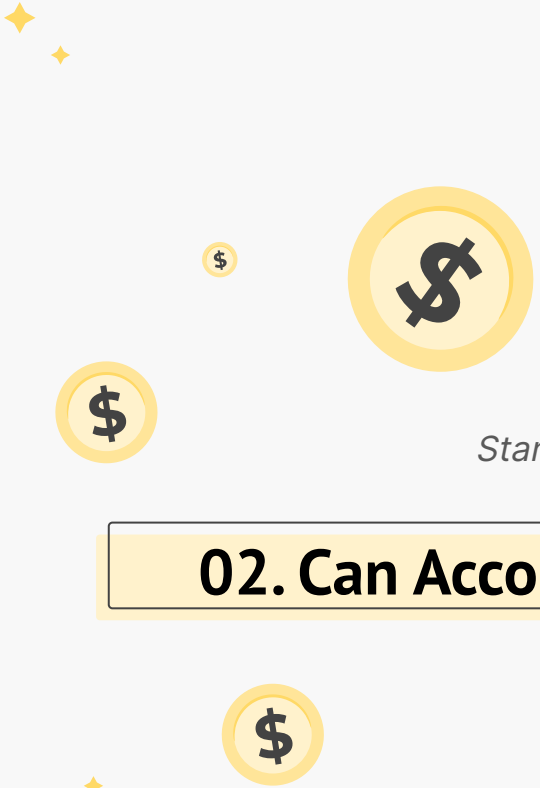
Examine missing values → Create dummy variables → Regress returns on dummies

RetEarly2020	Adj. R^2	Highest (Avg. Return)	Lowest (Avg. Return)
11 GICS sectors	0.075	Health Care (-0.14)***	Energy (-0.55)***
24 GICS groups	0.078	Food & Staples Retailing (-0.05) Health Care Equipment & Services (-0.1)***	Energy (-0.55)
69 GICS industries	0.095	Food & Staples Retailing (-0.056) Health Care Equipment & Services (-0.1)***	Energy Equipment & Services (-0.65)***
19 NAICS sectors	0.04	Educational Services (-0.12) Utilities (-0.2)***	Mining, Quarrying, and Oil and Gas Extraction (-0.54)***
10 SIC groups	0.03	Agriculture, Forestry and Fishing (-0.22) Services (-0.24)***	Mining (-0.54)***

Fluctuation During Market Recovery

Examine missing values → Create dummy variables → Regress returns on dummies

RetLate2020	Adj. R^2	Highest (Avg. Return)	Lowest (Avg. Return)
11 GICS sectors	0.055	Consumer Discretionary (1.74)***	Utilities (0.38)**
24 GICS groups	0.072	Automobiles & Components (2.3)***	Food & Staples Retailing (0.32) Insurance (0.37)**
69 GICS industries	0.094	Automobiles (3.76)***	Gas Utilities (0.04) Insurance (0.37)**
19 NAICS sectors	0.04	Other Services (3.38)*** Retail Trade (1.86)***	Educational Services (0.31) Utilities (0.34)**
10 SIC groups	0.036	Retail Trade (1.7)***	Nonclassifiable (0.43) Finance, Insurance and Real Estate (0.52)***



Stand at the beginning of 2020 and observe 2019 accounting characteristics

02. Can Accounting Characteristics Explain the Variation?

Explaining Fluctuation with Ratios

Calculate financial ratios → Address outliers & NA → Regress returns on ratios

- ROA, Asset Turnover, ROS, ROE, Equity Multiplier, Asset Intensity, Gross Margin, Cash to Asset, Cash Ratio, Quick Ratio, Long-Term Debt to Assets, Sales per Employee, and Altman's T1, T2, T3

	RetEarly2020	
	Adj. R^2	**Significant Ratios**
All 15 ratios	0.039	(+) Gross Margin (-) ROS, Long-Term Debt to Assets, Sales per Employee
10 (significant & higher-R^2)	0.037	(+) T1 (-) Long-Term Debt to Assets, 1/Sales per Employee
5 suggested	0.031	(+) Cash to Asset (-) Long-Term Debt to Assets, 1/Sales per Employee

Explaining Fluctuation with Ratios

Calculate financial ratios → Address outliers & NA → Regress returns on ratios

- ROA, Asset Turnover, ROS, ROE, Equity Multiplier, Asset Intensity, Gross Margin, Cash to Asset, Cash Ratio, Quick Ratio, Long-Term Debt to Assets, Sales per Employee, and Altman's T1, T2, T3

	RetLate2020	
	Adj. R^2	**Significant Ratios**
All 15 ratios	0.062	(+) ROS, Long-Term Debt to Assets (-) Asset Turnover, Gross Margin, Cash to Asset, T3
6 (significant & higher-R^2)	0.048	(+) Asset Turnover (-) T3
5 suggested	0.014	(+) Cash to Asset, Long-Term Debt to Assets, 1/Quick Ratio, 1/Sales per Employee

Explaining Fluctuation with Ratios

	RetEarly2020			RetLate2020		
	Adj. R^2	Intercept	Coefficient	Adj. R^2	Intercept	Coefficient
Return on Assets	0.014	-0.3089***	-0.2249***	0.023	0.855***	-1.1152***
Asset Turnover	0	-0.3052***	0.0011	0.014	0.7011***	0.3138***
Return on Sales	0.018	-0.3119***	-0.0219***	0.002	0.8877***	-0.0279***
Return on Equity	0.008	-0.3041***	-0.0676***	0.011	0.881***	-0.3049***
Equity Multiplier	0.005	-0.2688***	-0.009***	0.005	1.025***	-0.036***
Asset Intensity	0	-0.3054***	0.0011	0.014	0.7024***	0.3289***
Gross Margin	0.002	-0.3***	0.0082***	0.007	0.9637***	-0.00005***
Cash to Assets	0.019	-0.3431***	0.3404***	0.001	0.8545***	0.3835**

**Highlight in darker shade = Regressed on the inverse of the ratio*

Explaining Fluctuation with Ratios

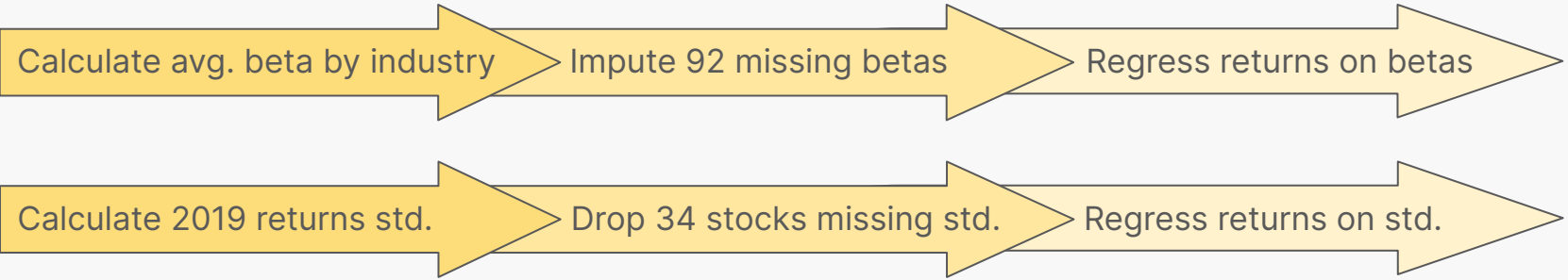
	RetEarly2020			RetLate2020		
	Adj. R^2	Intercept	Coefficient	Adj. R^2	Intercept	Coefficient
Cash Ratio	0.01	-0.325***	0.0294***	0.001	0.9337***	-0.0019***
Quick Ratio	0.002	-0.2786***	-0.0225**	0.006	0.7784***	0.1397***
Long-Term Debt to Assets	0.013	-0.2496***	-0.2066***	0.002	0.8312***	0.3035**
Sales per Employee	0.012	-0.3306***	6.5973***	0.003	0.8364***	14.1525**
T1	0.01	-0.3106***	0.1034***	0.006	0.8689***	0.314***
T2	0.013	-0.3115***	-0.0262***	0.018	0.8463***	-0.1217***
T3	0.016	-0.3026***	-0.2579***	0.023	0.8875***	-1.1947***

**Highlight in darker shade = Regressed on the inverse of the ratio*

What is the explanatory power of 2019 market betas?
What is the explanatory power of the standard deviation of 2019 returns?

03. Do Pre-COVID Risk Measures Explain the Variation?

Explaining Fluctuation with β and σ

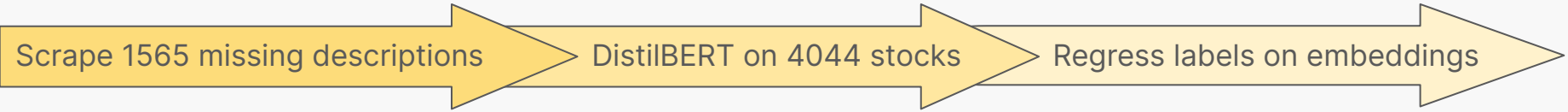


	RetEarly2020			RetLate2020		
	Adj. R^2	**Intercept**	**Coefficient**	**Adj. R^2**	**Intercept**	**Coefficient**
2019 Market Beta	0.005	-0.2587***	-0.0398***	0	0.8775***	0.0318
2019 Returns Std.	0	-0.3047***	0.0431	0.026	0.6569***	1.9352***

Use DistilBERT model to predict whether stocks experienced high/low returns

04. Can Business Descriptions Explain 2020 Performance?

Predicting Performance with Descriptions



BetterEarly2020 & **BetterLate2020**

= 1, if 3- or 9-month return is in the top 35% (higher than 65% of all stocks)

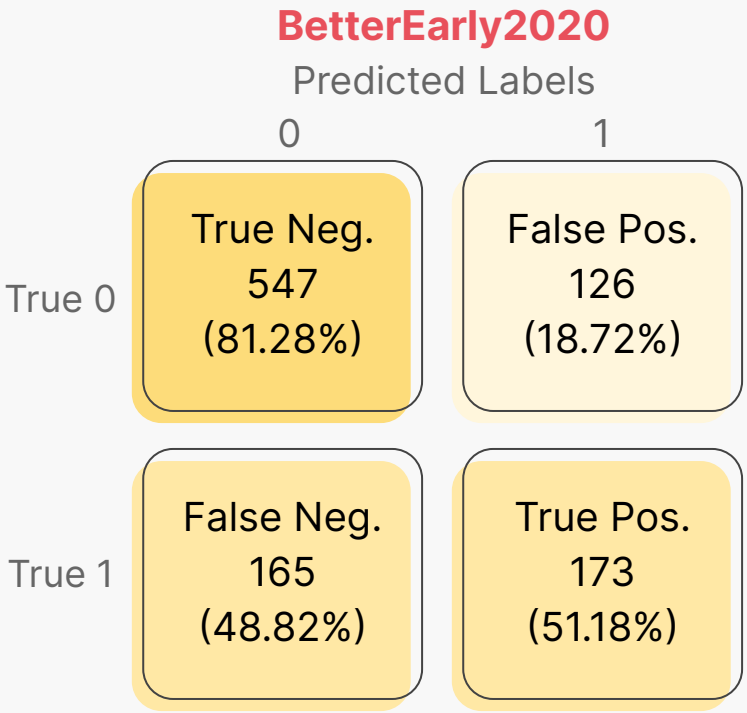
= 0, otherwise

→ **1416** stocks are labeled 1 and **2628** stocks are labeled 0

→ **75%** in training set and **25%** in test set

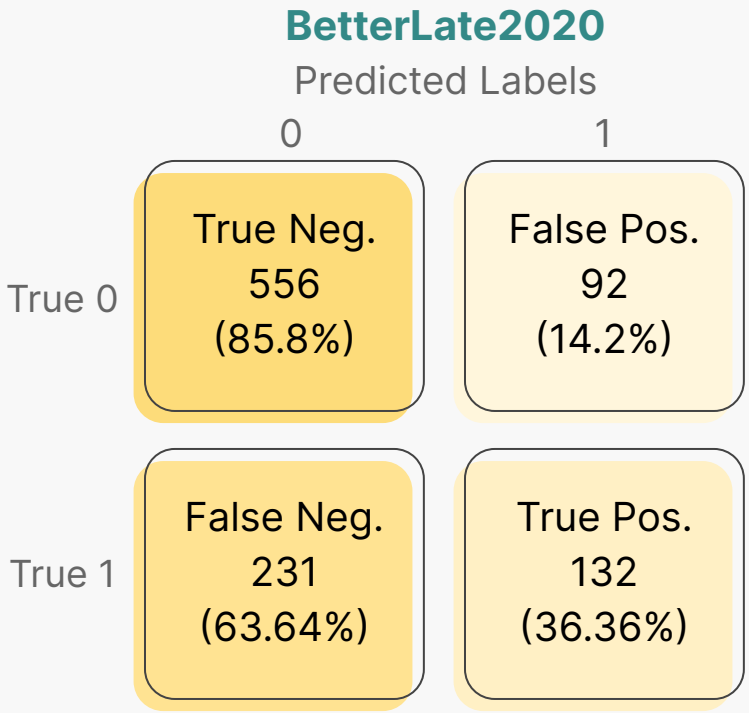
- Use embeddings created by DistilBERT to train two logistic regression models to predict whether a stock performs better than most in 2020

Predicting Performance with Descriptions



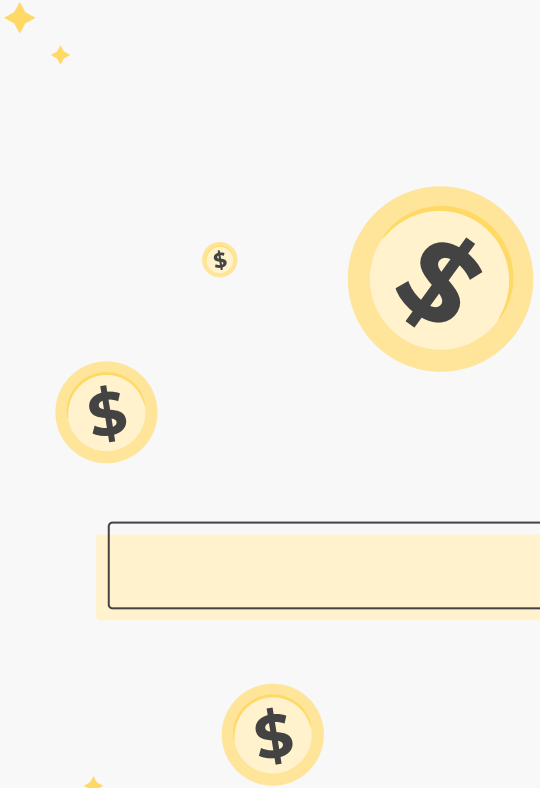
Accuracy = 0.71

(Random guess accuracy = 0.54)



Accuracy = 0.68

(Random guess accuracy = 0.55)



What happens if we utilize all the aforementioned explanatory variables?

05. Putting Everything Together

Explaining Fluctuation (Linear Reg.)

- 3,967 stocks
- Variables: 24 GICS industry groups, 2019 market beta, 2019 return std., financial ratios, description

	RetEarly2020 Adj. R^2	RetLate2020 Adj. R^2
GICS + all 15 ratios	0.088	0.128
GICS + ROA + ROE + Long-Term Debt to Asset + description	0.115	0.189
+ Sales to Employee	0.115	0.189
+ beta	0.117	0.189
24 GICS dummies + 768 description embeddings	**0.44**	**0.399**

Explaining Fluctuation (Logistic Reg.)

- 3,967 stocks (2,578 in training set; 1,389 in test set)
- Variables: 24 GICS industry groups, 2019 market beta, 2019 return std., financial ratios, description

BetterEarly2020

Predicted Labels

	0	1
True 0	True Neg. 526 (82.7%)	False Pos. 110 (17.3%)
True 1	False Neg. 190 (53.37%)	True Pos. 166 (46.63%)

Accuracy = 0.7

(Random guess accuracy = 0.56)

BetterLate2020

Predicted Labels

	0	1
True 0	True Neg. 559 (83.43%)	False Pos. 111 (16.57%)
True 1	False Neg. 180 (55.9%)	True Pos. 142 (44.1%)

Accuracy = 0.7

(Random guess accuracy = 0.55)

04.
CONCLUSIONS & CHALLENGES

CONCLUSIONS

- **GICS** has higher explanatory power than NAICS and SIC
  - Least impacted by COVID shock: **Service** industry (especially **Health Care**)
    Most impacted by COVID shock: **Energy** industry
  - Strongest rebound: **Automobiles & Retail** industries (consumer-facing)
    Slowest recovery: **Utilities** industry
- Characteristics of stocks that performed better in 2020
  - **More cash** (to cover current liabilities) at the end of 2019
  - **Higher long-term debt**
  - **Higher reliance on labor**
    - Less sensitive to market swings (lower beta)→ better 3-month return
    - More volatile (higher 2019 return std.) → better 9-month return
- **Business descriptions** provide good predictions about 2020 stock performance

CHALLENGES & FUTURE WORK

- Is there a better way to impute missing accounting data?
- Are there better financial ratios we can use?
- Is there a better way to impute missing market betas?
- Is there other textual data we can utilize?

THANK YOU

Q&A