How do Seasons Impact Markets?

Why Do Prices Go Up and Down?
Strong evidence exists which tie movements in equity and commodity market prices to outside, external world events. We discuss how governments impact prices of these financial products through the release of economic indicators. We also tie releases to seasonal cycles, which also appear to exert influence upon prices.

How We Approached the Study
Two separate studies took place. First, we studied long term seasonal cycles, then another study, which sampled four months of trading of a risk asset, where multiple economic announcements were made. In our historical study we looked at price over time, as divided by seasons.

Looking at an Example
In our study of a risk asset, we observed the flow of funds into buy and sell orders on one exchange. We then collected the data on the economic releases and synchronized them with the risk asset, to study the correlations between buy and sell order flows, and the price of the asset.

Insights We Gained from the Study
Seasons matter. We found that during certain seasons, distinct price changes occur. Also, while observing the behavior of a single financial asset, we observed correlated activity of the sales of that asset, on its price. This occurred both before and after highly consequential economic announcements were made by authorities in the United States.

Recommendations and Follow Through
When investing in the stock market, investors may get higher returns while the market is more volatile. This effect tends to be stronger in Q4, and hence investors should pay extra attention to market volatility in Q4.

In our study on QQQ and SPY, which replicates the performance of Nasdaq-100 and S&P 500 respectively, seasonality suggests that the stock prices fall ~2% at the end of Q1.

Cheuk Lung Yau steveyau@umich.edu
Stefan Bund bund@umich.edu
Scarlett Liang zhujunl@umich.edu

The Basics of Markets

Across the world, markets are where individuals and institutions can buy and sell things. These are popular things to sell:

- commodities, like wheat, cotton, and gold
- portions of companies, called stock, or equities
- money itself, like Japanese Yen or the US Dollar
- IOU's, or debts owed by others
- digital currencies, like Bitcoin, which are exchanged for national currencies
- shoes, used good and personal items

For each item, a place must exist to sell them. Personal items such as used clothing can be sold on websites like EBay. Stock in companies are sold on stock exchanges, and digital currencies are bought and sold on exchanges.

For stocks, commodities, cryptocurrencies and commodities, exchanges post tables of data that describe the orders which are pending, in their system. These orders are for selling or buying, and they tend to describe the price which is needed, to settle the transaction, and the quantity for sale.

These lists of buy and sell orders are called order books. Digital exchanges allow trades to happen at fast speed, and these order books are updated many times a second. This means that exchanges post data on how orders will execute, before they execute.

This creates an opportunity to learn how the appearance of orders may determine the price of a financial product, in the near future.

Vocabulary We Use

1. **Capitalization**. The measure of how much money is placed in an order, waiting on an exchange to be settled.
2. **Metaorders**. Describes the kind of order waiting to be finished, whether it is a buy or a sell order, on an exchange.
3. **Bid**. An order to buy something.
4. **Ask**. An order to sell something.
5. **Midpoint**. A synonym for the price. A midpoint exists where buying begins and selling ends. The price represents where buyers and sellers currently agree on a price, which is a median between where buying end, and selling begins.
6. **Order book**. A digital computer program where orders are stored. Orders wait here to be matched when prices meet their specified amount and price point.
7. **Quarter**. A continuous set of three months, during a year. Usually four exist in a year, and they map to **seasons** exactly.
8. **Market**. The place where an exchange takes place, usually described by a symbol, representing what is sold there, ie GOLD.
9. **Exchange**. A computer which matches buy and sell orders to customers, attempting to buy or sell a financial asset.
10. **Price**. The price of an asset is a temporary result of a settled transaction, where buyer and seller were matched via exchange. Price is an outcome of locating a midpoint between sell orders, and buy order.
11. **Seasonality**. In this paper, seasonality refers to predictable change over a long-term cycle (e.g. a quarter)
12. **Flows**. Flows are capitalization. They represent the funds which are reserved by order book orders.

Data Sources and References

Data sources We Employed

1. **Coinbase Exchange Web RPC API**

Provides limit order book data on a per-second basis, in the public websocket feed, WebSocket Overview | Coinbase Cloud. We used the API to generate high-frequency data on limit order books.

2. **Yahoo Finance streaming API**

Contains Open, Close, High, Low (OHLC) data on financial instruments, with a duration and interval we specify.

3. **Trading Economics Website**

Contains data on economic announcements, http://tradingeconomics.com

4. **Rad Disco / Stefan Bund public data repository, Github**

For aggregated order book data, ratios of buy/sell meta orders, capitalization flows. github, repository, https://github.com/stefanbund/m1

Prior Literature

1. Eggers, Ellison, Seok Lee. *The economic impact of recession announcements*. Journal of Monetary Economics, 120 (2021), 40-52. https://doi.org/10.1016/j.jmoneco.2021.03.002
2. Choi, Jayaraman. *Is reversal of large stock-price declines caused by overreaction or information asymmetry: Evidence from stock and option markets.* Journal of Futures Markets, volume 29, Issue 4, p. 348-376. https://doi.org/10.1002/fut.20360

Analytic Frameworks

1. Steve (seasonal_decompose / YFinance)
2. Scarlett (statsmodels / YFinance)
3. Stefan (Coinbase API)

Visualization Frameworks

1. Seaborn
2. Matplotlib
3. Altair

Collaboration Technologies Used

1. Github, repository: https://github.com/stefanbund/m1
2. Google collab
3. Slack
4. Zoom
5. Email

Images Used for Seasonality

https://towardsdatascience.com/finding-seasonal-trends-in-time-series-data-with-python-ce10c37aa861

Data Manipulation

Phase I: Establishing Seasonal Patterns in Markets

In our initial phase, we downloaded sets of price data, in various intervals via Yahoo Finance's Python API.

This process allowed us access to raw price and volume data, and we transformed the API feed into discreet csv files, to limit the use of memory, enhance the storage of the data, and allow for researchers to port large volume market data across the notebooks.

Phase 2: Combing Economic Announcement Data with Market Data

Using a four month sample of limit order book data, we aggregated the ratio of sell order data and buy data. This occupied roughly 1,000 lines of data per day, as the system aggregated the limit order book every several seconds.

The limit order book was compiled through one csv per day, into a global data frame.

With each economic indicator announcement, we created one data frame representing the state of the market before and after the announcement. Assuming that the announcement created conditions both attracting and repelling investment, we constructed dataframes to analyze correlations between sell and order capitalization both before and after the announcement.

More details about this are presented in the following slide.

Phase 3: Investigating Seasonal Risk

In the last phase, we go back to downloaded data from Yahoo Finance's Python API and read major index fund into into dataframe.

We then calculated daily standard deviation using each quarter's index funds' data

Also calculated return of the quarter, by using (last day closing price - first day closing price)/ first day closing price

Calculated normalized returns using:

$$\frac{return - \min(return)}{\max(return) - \min(return)}$$

Normalized standard deviation is then calculated using a similar equation. The linear regression on seasonality and plots in the last analysis are based on the normalized variables above.

Data Manipulation: Connecting Capitalization to Economic Announcements

Objective One: Get a Snapshot of Market Activity

In order to sample the active market, our data pipeline had to cover these basic issues:

1. acquire the capitalization of buy orders,
2. acquire the capitalization of sell orders,
3. capture price, volume and the time of the snapshot

Capitalization is an aggregate of an order's price multiplied by its volume, delivering a dollar value of an order. For each type of order (buy/sell), a sum is provided, per snapshot of the market.

The application to do this was written in node.js between the years 2019 and 2022. It runs continuously, acquiring the above market statistic every five seconds.

It uses the Coinbase exchange, and connects via API over an https websocket.

The Application's Continuous Cycle
This application must ingest approximately 13,000 data points for each snapshot, roughly once per 5 second interval.

For each snapshot of the market, it writes one line to a csv file, for that day. This file is entitled with the day of record, and each row is captured with a UNIX epoch timestamp.

In our team data, these application files are located inside the `lob_caps` directory, next to the data they create.

Objective Two: Combine Market Data with Economic Announcement Data

The midpoint of the order book is the settled price. Thus, the column, MP is the midpoint. The collection of data every five seconds, with corresponding ask and bid quantities, is associated with the price, at that moment. *The common date key is needed to help integrate the economic indicators with the numerous order book entries.*

However, the format of dates in the economic indicator announcements is human readable date-time. The order book data is stored as UNIX epoch. While iterating the economic indicators, the dates must be transformed so that we may fetch capitalization data associated with the announcement.

The Challenge, Integrating Two Data Sources By A Date Key

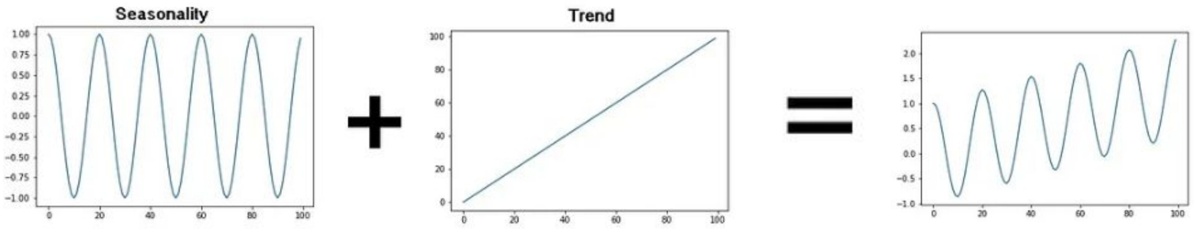
The UNIX timestamp must then be formatted into the python version of the epoch, which is a 13 number value. Given the fact that most of the order book values were written by a UNIX system, their dates appear with only 10 values.

We make up for this discrepancy by adding three zeroes to the end of each date, which enables the announcements to correlate to market events. However, the distance of the three zeroes tends to create a ten minute difference between the release of the indicators, and the exact market events.

Combining Our Data Sets By Date
Once the two files could join on a common key, the two datasets could integrate. We were able to locate the flows of money before, and after the announcement.

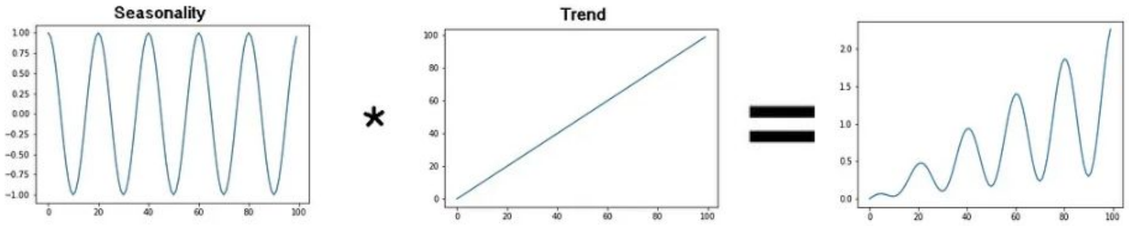
How Seasons Influence Prices

There are two types of seasonality - additive and multiplicative. They are differentiated by whether the amplitude of seasonality changes over time. It is additive if the amplitude remains the same.



We can model additive seasonality by adding seasonality to the trend line : Y[t] = T[t] + S[t] + e[t]

Seasonality is multiplicative if the amplitude changes over time.



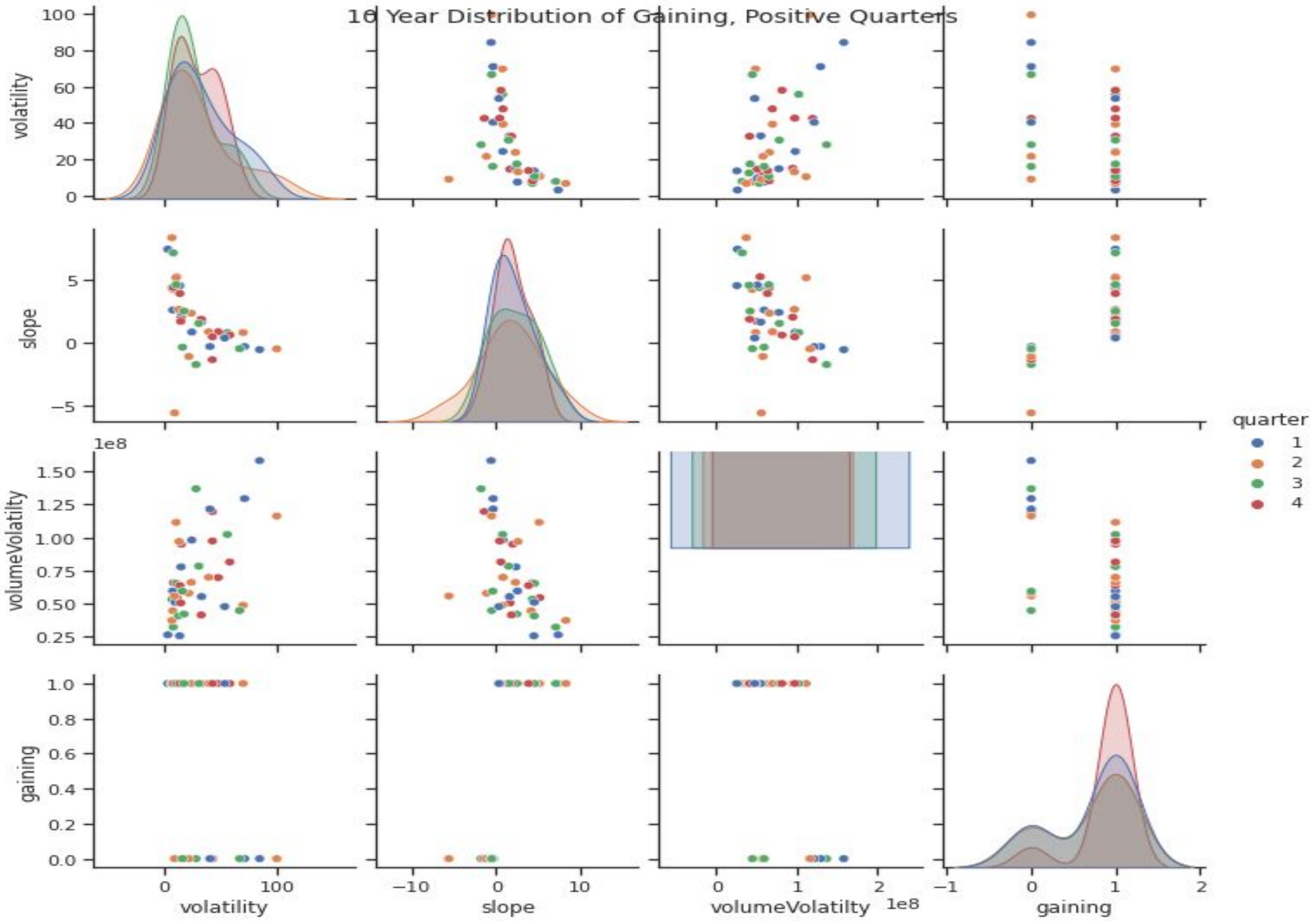
Multiplicative seasonality multiplies seasonality by the trend line : Y[t] = T[t] *S[t] *e[t]

Here, we will be using multiplicative model for our seasonality decomposition as the seasonality changes over time.
Below is our result :

Seasonality index represents the change of the stock price to the average stock price. An index of 1.02 represents an increase of 2%. An index of 0.98 represents a decrease of 2%.

The price tends to drop ~2% at the end of Q1. For the other quarters, the price change is too small to be accountable.

Which Seasons Normally Gain the Most?



The Distribution of Positive, or Gaining Seasons

We collected ten years of price data on the QQQ and SPR indexes. We created dummy values in our data which indicated whether a three month quarter (or season) resulted in a positive, or gaining value. We then counted the prevalence of 'gaining' versus non-gaining seasons, and organized the data by season.

We found that distinct differences in the distribution of 'gaining' seasons existed in our data.

In the pair-plot at the left, the four seasons are expressed in terms of the occurrence of 'gaining' seasons. The red shape (Q4) is distinctly among the majority of gaining seasons, historically.

In the following slide, we analyze a Fall season, attributed to the fourth quarter of the year, and study the ratios of buy and sell orders, by value.

In a Losing Season, Sell Order Capitalization Correlates to Price Change

Limit Order Books: Where orders go, to wait for execution
In any given exchange, a limit order book exists to store a record of orders, as they are placed, and as they wait for execution. This means that a public record of all waiting orders exists, for the public to read.

We have digested this list of pending orders, and aggregated their content into simple sums, for both buying and selling orders. This provides a distribution of values which suggests overweight buying, or selling activity is ahead for the exchange.

We observe that the amount of orders of the sell variety are closely aligned with the path the price takes, in the Fall season, 2022.

Strong Correlations Exist Between Sell Order Flows and Declining Prices During Pessimistic Economic Announcements

In an environment of pessimistic economic announcements, where inflation rose, consumer prices and consumer sentiment declined, sell order flows accompanied declining prices closely. The highly correlated price to sell order flows suggest a strong relationship between these two variables.

In our study, we looked at how the flow of funds into buy and sell orders resulted in potential predictive conditions. We analyzed the correlation between order book metaorders and settled price, using snapshots of order books in intervals of several seconds, or up to 1,000 times per day, over a four month period.

For markets before and after economic announcements, sell order capitalization correlated to prices before announcements, with 75% probability, and after announcements, with 97% probability.

This seasonal behavior aligned with the trends of economic announcements. However, this downtrend operates within the outlier distribution within typical Fall seasonality.

No Matter the Season, Volatility Correlates to Profitability

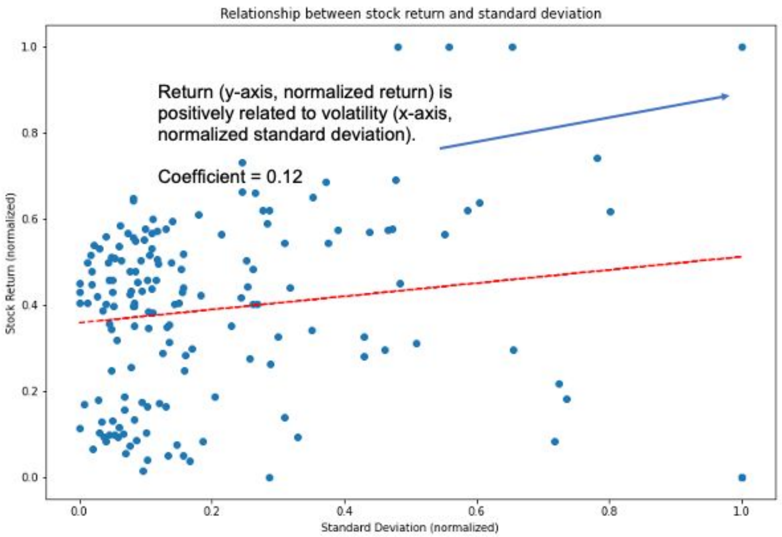
Calculating pearson correlation of normalized volatility vs return

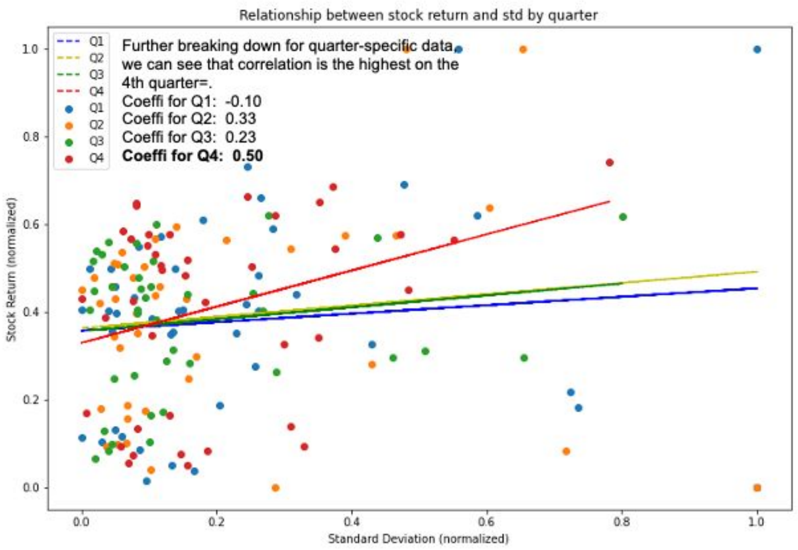
While it is widely believed that in order to get high returns, one need to invest in high risk stocks. It is less commonly understood that whether investing in more turbulent time periods leads to better returns.

To answer this question, we here looked at quarterly volatility and returns of 4 main index funds. Volatility is defined as normalized standard deviation of each day's closing price.

Looking at overall relationship, we observed slightly positive correlation between risk and returns, indicating on average, investors may get higher returns while the market is more volatile [*Left Chart 1*].

By further breaking down to look at quarter specific data, we found that the correlation is higher on the 4th quarter, indicating risk taking closer to the end of year may be a good investment idea [*Right Chart*].






```
                            OLS Regression Results
==============================================================================
Dep. Variable:            norm_return_   R-squared (uncentered):          0.422
Model:                             OLS   Adj. R-squared (uncentered):     0.417
Method:                  Least Squares   F-statistic:                     94.79
Date:                 Mon, 20 Feb 2023   Prob (F-statistic):           3.70e-17
Time:                         02:32:32   Log-Likelihood:                -45.071
No. Observations:                  131   AIC:                             92.14
Df Residuals:                      130   BIC:                             95.02
Df Model:                            1
Covariance Type:             nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
x1             0.9380      0.096      9.736      0.000       0.747       1.129
==============================================================================
```

Relationship between risk and return is further validated by linear regression model

In addition to pearson correlation, we also trained a one variable linear regression on normalized return ~ normalized standard deviation.

The result shows 0.422 r square, and a positive coefficient. P value is significant which helps confirm that the observed relationship is statistically significant.

Coefficient of 0.938 indicates that a 1x increase in normalized standard deviation can lead to 0.938x more normalized return.

Statement of Work

Cheuk Lung Yau (steveyau@umich.edu)

Initiated the study of seasonality with an analysis of the amplitude and magnitude of seasonal price behavior. Established the tools to identify the volatility of a season.

Integrated ten years of QQQ, SPY index data into a single time-series study. Applied a discovery algorithm to determine the amplitude, multiplicative scale factor per season.

Identified how seasonality delivered an expected risk factor.

Stefan Bund (bund@umich.edu)

Studied correlations between economic announcements, limit order book flow types (buy and sell) within the envelope of one season.

Integrated a four month data collection effort with public economic announcement data, during the same time period.

Identified how volatility was a potentially associated with the direction and pessimism within the announcements.

Created the node.js application necessary to capture market data, from Coinbase.

Performed an analysis of the probability that any given season will result in positive price gain, and the associated pair plots on slide 7.

Zhujun ‘Scarlett’ Liang (zhujunl@umich.edu)

Studied how volatility and return on investment correlate across the seasons.

Integrated ten years of pricing data of standard index funds on public exchanges such as VIX and QQQ while identifying volatility.

Established the potential association between volatility, return within the seasonal pattern. Built linear regression model to validate the hypothesis that return is relevant to volatility.

Cheuk Lung Yau steveyau@umich.edu
Stefan Bund bund@umich.edu
Scarlett Liang zhujunl@umich.edu