Social Networks
Final Project
Spring 2019

Yash Goel
Sakada Lim
John Reynolds
Jing Xia

Incentives to Cheat on Economic Announcements

1. Hypothesis:

There is a large incentive to "cheat" around financial announcements such as non-farm payrolls, fed minutes, central bank policy, natural resource data, etc. The individuals familiar with this information have an incentive to leak the information to "financial agents" who can then act on this data before before its publicly announced. We can verify this hypothesis fast and cheaply by measuring the gains that a "financial agent" could make by prepositioning in the securities most impacted by the announcements. If there is no measurable difference between an agent who his prior knowledge of an event relative to an agent who could acquire then we can refute the hypothesis. If we can show there are gains to be made from acting on the knowledge beforehand then we can attempt to identify the scope of products and timings of optimal prepositioning.

2. Background:

While most insider trading headlines and enforcement activity center around stock trading, the insider market for macro trading/investing goes relatively unnoticed. Macro investing centers trading and investing in financial products that are impacted by macro economic events. These will include Interest Rate Futures, Foreign Exchange, Equity Futures, and Commodity Futures(Oil, Gold, Natural Gas, Crops, LiveStock, etc). Investors can tap into Futures Markets anywhere in the world to easily place bets on anticipated directional changes in any of the before mentioned contract types.

The largest price movements are typically centered around Central Bank Policy Updates and US Non-Farm Payrolls. Non-Farm Payrolls are released the first friday of every month at exactly 8:30 am (EST) and will provide the change in US Employment for the prior month. It is viewed as the most important number of economic outlook. Later in the month the Federal Reserve releases monthly FOMC minutes which outline current Federal Reserve Monetary Policy. FOMC minutes are released at exactly 2:00pm on one Wednesday of the Month.

There has long been suspicions that the important economic numbers and Fed Policy minutes are being leaked. Helping to fuel the suspicions is the cottage industry of economic forecasting

firms run by academic economists. There are strong ties between academic economists and current Federal Reserve Officials, due the the large number of Economists who rotate between academia and the Fed. Questions have been raised as to whether trading firms paying for the forecasting services are actually paying for a forecast with large uncertainty or are they paying for access to economists who have strong ties the the Federal Reserve.

3. Data:

For analysis we used publicly available events calendar, Futures Tick Trade data from the CME and Twitter for President Trump's communications. Futures were chosen for the analysis because they trade around the clock for 5.5 days out of the week and are highly liquid. We focused this study on Equity Futures (ES-E-mini S&P 500) , Short Term Interest Rate Futures (TU-2year Future) and Long Term Interest Rate Futures(UB-Ultra Long Bond). The analysis covers 2014 until April of 2019. Tick data was chosen over lower frequency to ensure consistency in the results and ensure maximum coverage of events as many events (non-farm payrolls for example) occur outside normal hours of market operation. Note: In general futures Tick data is not freely available, but can be purchased. Tick data for commercial purposes can be extremely expensive (historical tick data costs can well exceed $100,00). Researchers can find firms willing to give tick level data.

4. Testing the Hypothesis

To test the hypothesis we want to examine the following questions:

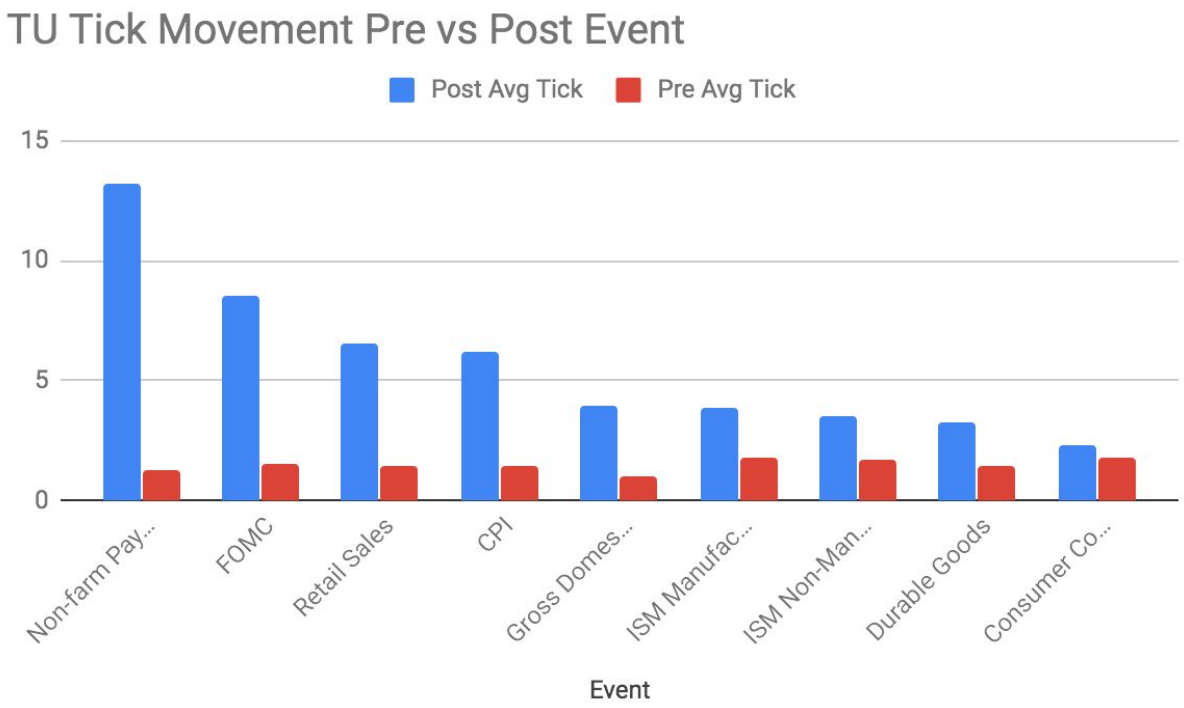
1. Do announcements present opportunity?
2. Can pre market movements help in predicting the direction of an event?
3. If we had the announcement data could we benefit from it?

4.1 Do announcements present opportunity?

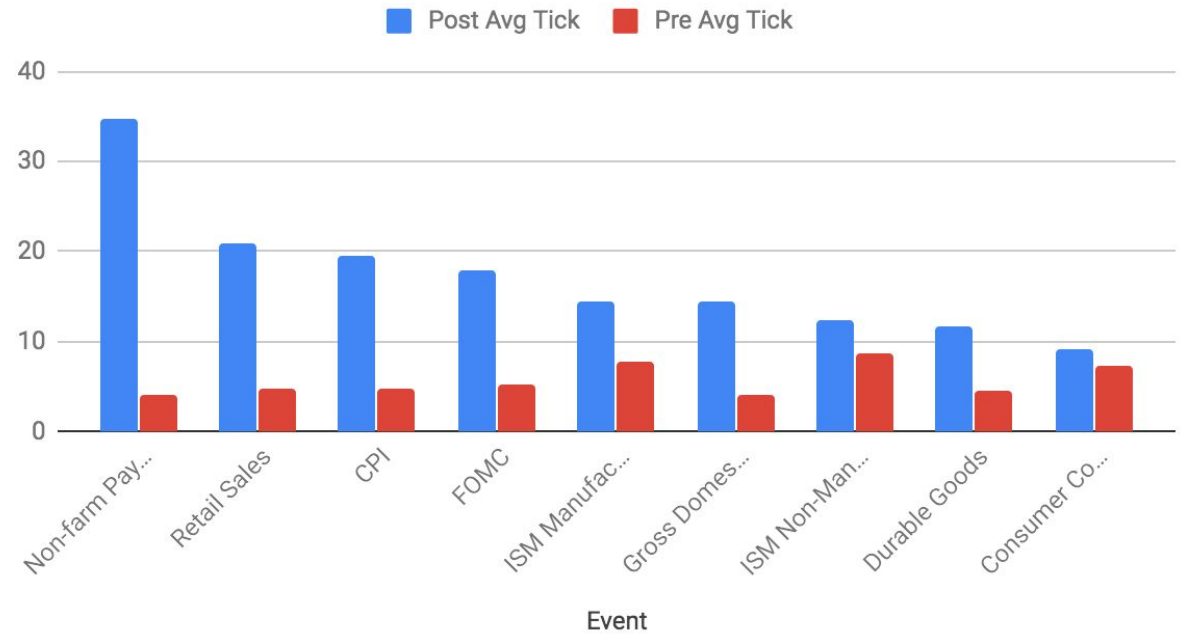
We will start with the premise there can be no crime without opportunity. To determine "opportunity" we will compare the magnitude of the market movement pre-event (for this study we chose a ten minute window and hour before each measure event) compared to post event ten minute window. Below are charts and tables comparing the pre-Event Windows and post-Event windows for nine of the most impactful economic announcements.

The short answer: events present extremely outsized opportunities. It's possible to make short term trades and gain outsized profits given the time period. The data supporting the opportunity size unambiguous. The analysis data is presented in the tables and graphs below. The average tick movement across all events is approximately 3.5x larger than the average tick movement in prices across a similar time window without and event while the **trading volume is on average 6.5x larger and can be over 20x larger.** What's staggering is **maximum price movement in ticks can be over 29x the market average.** In the sample conducted over 99% of the tick market moves were greater than the non-event tick average. In

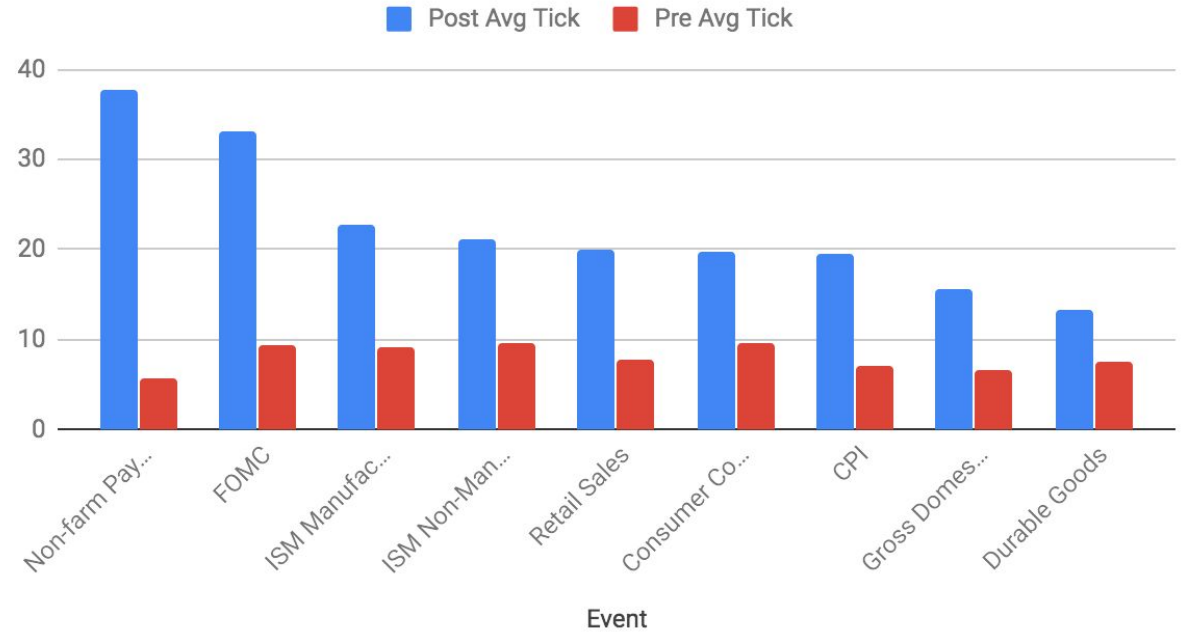
general the most impactful events in this study are Non-Farm Payrolls with moves on average 8x the market and Federal Open Market Committee (FOMC)Rates Decision with moves over 5x the market. The largest single movement in the study on an ES Contract was due to (FOMC) announcement where the ES contract moved a staggering 199 ticks in 9 minutes, compared to a 7 tick move during the average window.



UB Tick Movement Pre vs Post Event



ES Tick Movement Pre vs Post



covers approximately 58 events from Jan 2104 to April 2019

Event Tick Movement UB (Max,Mean), (No Event Mean) and Traded Volume Ratio

Symbol	Event	N	Event Max	Even Mean	No Event	Vol Ratio
UB	Non-farm Payrolls	58	80	35	4	16.4
UB	Retail Sales	58	46	21	5	8.5
UB	CPI	58	44	20	5	8.0
UB	Gross Domestic Product	46	29	14	4	5.9
UB	FOMC	54	45	18	5	3.6
UB	Durable Goods	56	29	12	5	4.2
UB	ISM Manufacturing	57	31	14	8	2.1
UB	ISM Non-Manufacturing	55	43	12	9	1.7
UB	Consumer Confidence	53	32	9	7	1.4

Event Tick Movement TU (Max,Mean), (No Event Mean) and Traded Volume Ratio

Symbol	Event	N	Event Max	Mean	No Event	VolumeR
TU	Non-farm Payrolls	58	35	13	1	17.1
TU	FOMC	54	45	9	2	5.4
TU	Retail Sales	58	14	7	1	9.0
TU	CPI	58	13	6	1	8.5
TU	Gross Domestic Product	46	10	4	1	6.7
TU	Durable Goods	56	8	3	1	4.4
TU	ISM Manufacturing	57	10	4	2	2.7
TU	ISM Non-Manufacturing	55	12	3	2	2.1
TU	Consumer Confidence	53	12	2	2	1.8

Event Tick Movement ES (Max,Mean), (No Event Mean) and Traded Volume Ratio

Symbol	Event	N	Max	Mean	No Event	VolumeR
ES	Non-farm Payrolls	58	138	38	6	14.3
ES	FOMC	54	199	33	9	3.6

ES	CPI	58	194	20	7	4.5
ES	Retail Sales	58	194	20	8	5.0
ES	ISM Manufacturing	57	65	23	9	7.5
ES	Gross Domestic Product	46	49	16	7	4.1
ES	ISM Non-Manufacturing	55	67	21	10	6.6
ES	Consumer Confidence	53	68	20	10	5.9
ES	Durable Goods	56	61	13	7	2.6

4.2 Can pre-market moves help in predicting the direction of the event?

To test this premise we exame a few different models to determine how much (if at all) pre-event returns can help in predicting post event returns. We evaluated the prediction of post-event return using two models and condition on different criteria, all pre-event returns and pre-event returns that were 1 standard deviation greater than the mean pre-event move. The first model predicts the post event return using a standard linear regression, the second model is a classification model to just predict the post move direction (up or down). We choose support vector machine because it has stricter criteria than a simple logistic regression.

The reasoning on **"Does someone know something?".**

Since not every economic event results in large post announcement market moves, we decided to check to see if larger than normal pre event market moves were any different in predicting post event moves. A priori in absence of any information, there would no reason to believe that there would be any predictive difference in the ability of higher premarket moves to be able to predict post event market moves.

Findings

We were surprised to find that when the pre-event market move is greater than 1 standard deviation from the mean pre-event market move, the predictive power of the pre-market move is significantly better compared to all market moves. **A speculator has much better odds of making a profit when the pre-event movement is larger than usual!** This was actually surprising to find. Also the ability to predict post market returns from pre-market returns using all data for each event was close to zero, while the ability to predict post even direction fared better. The results are summarized below.

Models and Results for 4.2

Model 1: linear regression of post-event return on pre-event return

$$R_t = \beta R_{t-1} + \varepsilon_t$$

Model 2: support vector machine logistic regression of post-event direction using pre-event return

$$D_t = \sigma(\beta R_{t-1}) + \varepsilon_t$$

$$where\ \sigma(x)\ is\ the\ sigmoid\ function \quad and\ D_t = 1 if\ R_t > 0\ and\ D_t = 0\ if\ R_t < 0$$

Explanation of numbers in below tables

The analysis in the table below compares two scenarios:

Scenario 1:

Can we predict the post event market move based on the pre-market move conditioning on all data.

$$E[R_t | R_{t-1}] = \beta R_{t-1} + \varepsilon \quad (Equation\ 1)$$

$$E[D_t | D_{t-1}] = \sigma(\beta R_{t-1}) + \varepsilon \quad (Equation\ 2)$$

Scenario 2:

What is the prediction if we just we larger pre-market moves?

$$E[R_t | R_{t-1},\ R_{t-1} > \sigma_{R_{t-1}}] = \beta R_{t-1} + \varepsilon \quad (Equation\ 3)$$

$$E[D_t | D_{t-1},\ R_{t_{-1}} > \sigma_{R_{t-1}}] = \sigma(\beta R_{t-1}) + \varepsilon \quad (Equation\ 4)$$

Where

$R_t = (P_t - P_{t-1})/P_{t-1}$

$P_t - Price\ of\ Futures\ Contract\ at\ time\ t$

$D_t = 1\ if\ R_t > 0\ ,\ 0\ if\ R_t < 0$

$\sigma(\beta R_{t-1}) - sigmoid$

$\sigma_{R_{t-1}} - standard\ deviation\ of\ pre-market\ returns$

$R_p - Mean\ Return\ during\ pre\ event\ window$

$R_{[p|R>\sigma]} - Mean\ Return\ of\ prevent > \sigma_{R_{t-1}}$

$R_E - Return\ of\ market\ post\ event$

$Acc(D_t) - Accuracy\ of\ SVM - Acc(1)$

$Acc(D_t | R_{t_{-1}} > \sigma_R) - Acc(2)$

$r^2 - r\ squared$

Columns for the tables

$I - Instrument$

$Event - Economic\ Event$

$Ratio - R_{[p|R>\sigma]}/R_p$

$Acc(2) - Accuracy\ of\ Scenario\ 2 - Equation\ 4$

$Acc(1) - Accuracy\ of\ Scenario\ 1 - Equation\ 2$

$N(2) - number\ of\ samples\ Scenario\ 2$

$N(1) - number\ of\ samples\ Scenario\ 1$

$r^2(2) - r\ squared\ of\ Scenario\ 2 - Equation\ 3$

$r^2(1) - r\ squared\ of\ Scenario\ 1 - Equation\ 1$

Predicting ES post return and post move direction, using all data and a one standard deviation move.

I	$Event$	$Ratio$	$Acc(2)$	$Acc(1)$	$N(2)$	$N(1)$	$r^2(2)$	$r^2(1)$
ES	CPI	3.8	0.75	0.57	4	58	0.05	0.00
ES	Consumer Confidence	0.4	0.63	0.51	8	53	0.17	0.10
ES	Durable Goods	9.7	0.71	0.66	7	56	0.06	0.01
ES	FOMC	5.9	0.75	0.65	4	54	0.62	0.07
ES	Gross Domestic Product	3.4	0.78	0.57	9	46	0.03	0.01
ES	ISM Manufacturing	1.7	0.56	0.56	9	57	0.02	0.02
ES	ISM Non-Manufacturing	0.6	0.67	0.71	6	55	0.05	0.00
ES	Non-farm Payrolls	2.5	0.60	0.50	5	58	0.47	0.00

Predicting TU post return and post move direction, using all data and a one standard deviation move.

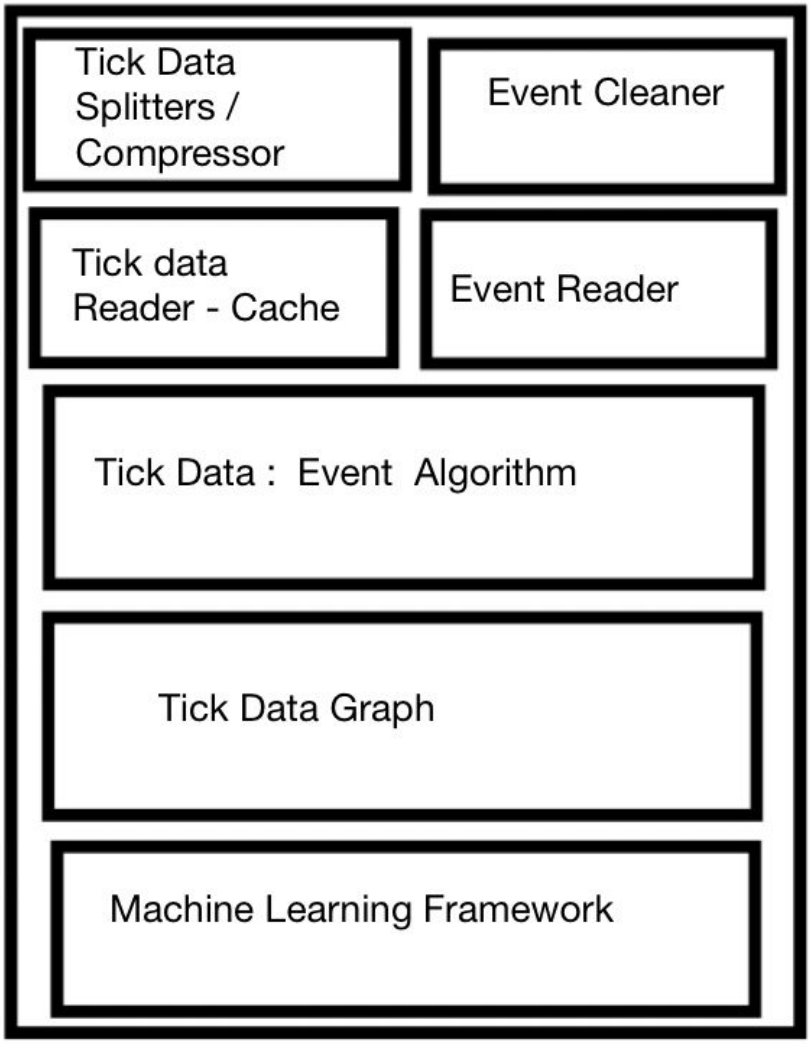
I	$Event$	$Ratio$	$Acc(2)$	$Acc(1)$	$N(2)$	$N(1)$	$r^2(2)$	$r^2(1)$
TU	CPI	2.3	0.70	0.60	10	58	0.01	0.00
TU	Consumer Confidence	263.3	0.50	0.70	2	53	1.00	0.00
TU	Durable Goods	4.1	0.60	0.64	10	56	0.03	0.00
TU	FOMC	2.4	0.83	0.54	6	54	0.25	0.01
TU	Gross Domestic Product	42.8	0.57	0.78	7	46	0.01	0.01
TU	ISM Manufacturing	1.8	0.67	0.61	9	57	0.35	0.04
TU	ISM Non-Manufacturing	0.8	0.67	0.73	6	55	0.00	0.01
TU	Non-farm Payrolls	4.1	0.71	0.50	7	58	0.64	0.08
TU	Retail Sales	2.4	0.82	0.57	11	58	0.02	0.04

Predicting TU post return and post move direction, using all data and a one standard deviation move.

I	$Event$	$Ratio$	$Acc(2)$	$Acc(1)$	$N(2)$	$N(1)$	$r^2(2)$	$r^2(1)$
UB	CPI	21.1	0.75	0.50	8	58	0.26	0.02
UB	Consumer Confidence	21.7	0.75	0.62	4	53	0.13	0.01
UB	Durable Goods	3.5	0.63	0.55	8	56	0.25	0.00
UB	FOMC	0.9	0.67	0.52	9	54	0.03	0.02
UB	Gross Domestic Product	0.0	0.57	0.52	7	46	0.03	0.00
UB	ISM Manufacturing	2.1	0.70	0.54	10	57	0.07	0.00
UB	ISM Non-Manufacturing	6.5	0.67	0.60	6	55	0.16	0.00
UB	Non-farm Payrolls	4.6	0.78	0.62	9	58	0.10	0.02
UB	Retail Sales	5.2	0.80	0.53	10	58	0.01	0.00

6. The Software Stack

The diagram below represents the software stack developed for this analysis which could turn into a product to lease to middle-tier financial firms. Many firms we have encountered do not have any in house capability to split, compresee, cache and then deliver tick data to a client program for analysis. The software takes in several GB of raw trade tick data from the CME, splits and compresses it so it can be quickly looked up via data time in a client analysis tool via the cache reader or used as input to simple machine learning framework developed for this project.



Software Stack

7. Software Features

The adage is “you can google anything”. In doing the analysis we discovered how difficult it is to answer historical based questions and find out what market reactions were on particular dates. For example we can asked the following questions. In addition it's also difficult to align events timing to the market.

What happened to the Market on Non-farm Payrolls in February 2019?

The search will come back with the March Number (as February employment numbers are released in March 2019, but we want the Market Reaction to the payroll report released in February 2019).

We can ask different variants of the question

1. Equity Market Reaction to Non-farm Payrolls in February 2019?
2. Bond Market Reaction to Non-farm Payrolls in February 2019?
3. UB Futures Contract Market Reaction to Non-farm Payrolls in February 2019?
4. ES Futures Contract Market Reaction to Non-farm Payrolls in February 2019?
5. How much money would a speculator make on 1000 contracts off NFP in Feb 2019?
6. What was the difference between Feb and March Jobs Report?

Web search engines will not produce satisfactory results on even easier variants of these questions and the further in time away the more difficult it is to find out what happend.
The software developed has the following features

- Create compressed Tick Trade Archive from Raw Exchange Data
- Take exchange tick trade data make it fast to load for any particular date across a range of contracts.
- Create Events Repository with Event , Event Date, Market Expectation and Actual Number
- Locate event based on a close date, map the market reaction and expectations and plot
- Show the profit loss of speculators taking long/short positions in contracts
- Compare Events based on Tick Movements
- Apply Machine Learning
    - Assess pre-market moves of post event moves,
    - Assess market expectations vs actual on market moves

It would be interesting to take the software developed from this project and turn it into a service to look up historical market reactions.

8. Applications to other markets

While we focused on US Futures Markets and Macroeconomic Events, the software stack developed can be applied to any exchange traded markets around the world. For instance we could use the methodology on earnings announcements on US Equities, international central banks announcements on their impact currency valuation and general foreign economic reporting and its impact on local equities and fixed income markets.

9. Conclusion

We have shown that economic announcements present outsized opportunity in terms of market volatility and trading volume compared to non-event market periods. We have show that higher

than average pre-event market moves are better predictors of the post event price direction and returns compared to the average pre-event return. This lends some credence to the adage that "someone knows something." The software developed for this problem, stumbled on the problem of the difficulty in using search engines to answer things from an historical context.and starts to address the issue of its not really easy to use search engines to find things in a historical context. The work can possibly be used to
In addition we have created the genesis of a product that can be used to analyze historical events impact on the markets.