# Register

## Study Information

### Title

*Provide the working title of your study. It may be the same title that you submit for publication of your final manuscript, but it is not a requirement.*

> How do different type of traders react to shocks?

### Authors

> ceyda ustun

### Description

*Please give a brief description of your study, including some background, the purpose of the study, or broad research questions.*

> In my study, my focus market will be stock market. In the automated stock markets, buyers and seller don't arrive simulatenously (this is the friction of the market) and liquidity provision is the essential function of markets where buyers and sellers don't arrive simultaneously. There are liquidity providers- called as LPs- (intraday intermediaries or market makers) to fill the gap arising from the imperfect syncronization between the arrivals of buyers and sellers.
>
> However, the provision of continuous market presence is costly. For that reason, LPs choose to maintain equilibrium risk exposure but that may be too low to offset large but temporary liquidity imbalances. In the event of large enough selling pressure, the liquidity on the buy side can only be obtained after a price drop that's large enough to compensate reluctant LPs for taking on additional risky inventory. In these times, theory suggest that LP's risk bearing capacity is overwhelmed and they may become unwilling to accumulate more inventory without large price concessions. This may even trigger a market crash w/o fundamental shocks.
>
> I will use an exogenous, big selling pressure event happened in Turkish Stock Market as an identification strategy. and  will try to answer the below questions emprically.
>
> - How does sudden excess selling pressure affect LP's trading pattern (inventory holding

dynamics wrt price changes) in the stock market.

- Do LPs stabilize the stock market during the crash?

## Hypotheses

*List specific, concise, and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here. If a specific interaction or moderation is important to your research, you can list that as a separate hypothesis.*

- Statistical relationship between LPs inventory change and price change is significantly changing in case of high selling pressure??

- Statistical relationship between LPs inventory change and price change is different for HFT LPs and nonHFT LPs? (In theory, faster intraday traders are able to "snipe" stale orders of slower market participants. So that, inventories of faster intraday traders may positively comove with price changes at high frequencies)

- In theory, LPs supposed to stabilize the market price. I want to check whether this holds in case of market stress (very related to the first one)

# Design Plan

## Study type

*Please check one of the following statements*

Other

## Blinding

*Blinding describes who is aware of the experimental manipulations within a study. Mark all that apply.*

No blinding is involved in this study.

## Is there any additional blinding in this study?

*Blinding (Other)*

## Study design

*Describe your study design. Examples include two-group, factorial, randomized block, and repeated measures. Is it a between (unpaired), within-subject (paired), or mixed design? Describe any counterbalancing required. Typical study designs for observation studies include cohort, cross sectional, and case-control studies.*

I have exogenous, relatively big shock from swap market to stock market. This shock - exogenous high seling pressure- will be used as an identification strategy.

This is temporary government measure taken in TRYUSD swap market on March 27 just before the local elections. In brief, it created a high selling pressure in the stock market on March 27. $168 million valued BIST30 stocks sold by foreign accounts (≈ 1% of stock market value). This selling pressure leads to second highest fall of 5.86% in BIST30 index of Turkish Stock Exchange.

One thread to my identification strategy  is that I need assume that this shock made no change in fundamentals of the stocks. (index turned to its initial values after that day)

no file selected

## Randomization
*If you are doing a randomized study, how will you randomize, and at what level?*

# Sampling Plan

## Existing Data
*Preregistration is designed to make clear the distinction between confirmatory tests, specified prior to seeing the data, and exploratory analyses conducted after observing the data. Therefore, creating a research plan in which existing data will be used presents unique challenges. Please select the description that best describes your situation. Please see https://cos.io/prereg for more information.*

Registration prior to analysis of the data

## Explanation of existing data
*If you indicate that you will be using some data that already exist in this study, please describe the steps you have taken to assure that you are unaware of any patterns or summary statistics in the data. This may include an explanation of how access to the data has been limited, who has observed the data, or how you have avoided observing any analysis of the specific data you will use in your study.*

I have two weeks of tradebook and orderbook data. The data interval is March 18-29, 2019. I have seven trading days before the event. Every transaction realized in microseconds.

- intermediary IDs
- stock IDs
- account IDs

- order type, price, volume and value
- trade type, price, volume and value
- buy/sell indicator
- user connection type (HFT, Colocation, Fix etc.)
- date and update timestamps
- change reasons of orders
- foreign vs domestic account (intermediaries not having retail banking operations in Turkey is assumed to be foreign - this info will not be used for testing my hypothesis, I use this to see how much foreign accounts sold)

## Data collection procedures

*Please describe the process by which you will collect your data. If you are using human subjects, this should include the population from which you obtain subjects, recruitment efforts, payment for participation, how subjects will be selected for eligibility from the initial pool (e.g. inclusion and exclusion rules), and your study timeline. For studies that don't include human subjects, include information about how you will collect samples, duration of data gathering efforts, source or location of samples, or batch numbers you will use.*

Both datasets are from the Stock Exchange's database.

no file selected

## Sample size

*Describe the sample size of your study. How many units will be analyzed in the study? This could be the number of people, birds, classrooms, plots, interactions, or countries included. If the units are not individuals, then describe the size requirements for each unit. If you are using a clustered or multilevel design, how many units are you collecting at each level of the analysis?*

After I classify the LP trading accounts, I will do seperate analysis for these accounts (HFT LPs and nonHFT LPs). I expect roughly 60-70 (not sure) account for each stock where 15-20 of them will be HFT LPs.

In terms of transactions, these accounts are almost doing the half of the tradebook everyday. For an arbitrary stock, it makes roughly 150.000 transactions within a day.

## Sample size rationale

*This could include a power analysis or an arbitrary constraint such as time, money, or personnel.*

## Stopping rule

*If your data collection procedures do not give you full control over your exact sample size, specify how you will decide when to terminate your data collection.*

# Variables

## Manipulated variables

*Describe all variables you plan to manipulate and the levels or treatment arms of each variable. This is not applicable to any observational study.*

no file selected

## Measured variables

*Describe each variable that you will measure. This will include outcome measures, as well as any predictors or covariates that you will measure. You do not need to include any variables that you plan on collecting if they are not going to be included in the confirmatory analyses of this study.*

My outcome variable will be an indicator for their inventory holding dynamics. So, it could be an inventory change (from tradebook) for some time frequency or price and signed quantity pairs (from orderbook) for each pair of trader- time-stock to show their demand.

My covariates can be, for example, the actual market price of the stock, lagged market price of the stock, lagged inventory holding, volatility (a proxy for risk), actual inventory holding of the LPs, past cumulative profit of the LP (if calculable - proxy for marked-to-market equity and thus for inventory limit). Lastly, covariates can be combined to grab the effect of sudden selling pressure on the LHS variable. For instance, in the case of selling pressure dummy, cross term (dummy *inventory) can be used..

no file selected

## Indices

*If any measurements are going to be combined into an index (or even a mean), what measures will you use and how will they be combined? Include either a formula or a precise description of your method. If you are using a more complicated statistical method to combine measures (e.g. a factor analysis), you can note that here but describe the exact method in the analysis plan section.*

no file selected

# Analysis Plan

## Statistical models

*What statistical model will you use to test each hypothesis? Please include the type of model (e.g. ANOVA, multiple regression, SEM, etc) and the specification of the model (this includes each variable that will be included as predictors, outcomes, or covariates). Please specify any interactions, subgroup analyses, pairwise or complex contrasts, or follow-up tests from omnibus tests. If you plan on using any positive controls, negative controls, or manipulation checks you may mention that here. Remember that any test not included here must be noted as an exploratory test in your final article.*

I have one  candidate baseline model in mind simply to examine the comovement between LP's inventory and stock price changes (no causal inference since inventories and stock prices are jointly determined) by using the tradebook data.

The dependent variable is the change in holdings of HFT LPs (do the same then for nonHFT LPs). RHS will be lagged change in inventory, lagged inventory and some predetermined lags of stock price change scaled by stock's tick price for before event day observations (maybe only first week) and  event day dummy interacted with all RHS variables and lastly hourly time fixed effect ((do the same then for nonHFT LPs) .

or VAR model by using the orderbook data (will be checked).

no file selected

## Transformations

*If you plan on transforming, centering, recoding the data, or will require a coding scheme for categorical variables, please describe that process.*

I will need to create new categorical variables for trader types by using both tradebook and orderbook data.

There is no info about the type of traders (like LPs) in the data set since stock market has no formally required liquidity providers. For this reason, I need to classify the traders by looking their directly observed individual inventory and trading volume patterns. The most salient characteristics of LPs are
* consistently buying and selling throughout a trading day - relatively high trading volume
* low levels of end of day inventory.
* intraday mean-reversion in inventory holdings.

 The idea is that if an account is classified as LP in the first week, it will remain in the same category for the event day. I will further seperate LPs into HFTs and nonHFTs by looking their connection type to the Stock Exchange.

I calculated the below variables.

- trader volume scaled by market trading volume
- end of day net position scaled by trader volume
- Number of trade transaction per minute
- Number of order transaction per minute
- Intertrade duration
- Interorder duration
- Share of the limit orders and market orders

## Inference criteria
*What criteria will you use to make inferences? Please describe the information you'll use (e.g. specify the p-values, Bayes factors, specific model fit indices), as well as cut-off criterion, where appropriate. Will you be using one or two tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this?*

## Data exclusion
*How will you determine which data points or samples if any to exclude from your analyses? How will outliers be handled? Will you use any awareness check?*

## Missing data
*How will you deal with incomplete or missing data?*

## Exploratory analysis

*If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time.*

# Other

## Other

*If there is any additional information that you feel needs to be included in your preregistration, please enter it here. Literature cited, disclosures of any related work such as replications or work that uses the same data, or other context that will be helpful for future readers would be appropriate here.*

Continue editing (/h8rje/drafts/5ddb934de15e44000c1dc27c/)                    Register