BUILDING A GROWTH-FRIENDLY TAX ENVIRONMENT

The 4th Belt and Road Initiative Tax Administration Cooperation Forum

— Improving Tax Environment

Tbilisi Georgia 11-13 September 2023

BUILDING A GROWTH-FRIENDLY TAX ENVIRONMENT

Technology helps to simplify tax compliance

Name: Ashot Muradyan
Title: Deputy Head
Jurisdiction/Organization: State Revenue Committee of Republic of Armenia

Classical approach to the main functions of Tax Administration

- Fiscal — Collecting money to the State Budged
- Control — To find and punish fraud taxpayers
- Services — To deliver services to taxpayers

Approach is drastically changed today

- Be “user friendly”, deliver as much services, as possible
- Prevent tax fraud, use only targeted control
- As a result, rise money to the State Budged

Armenian Tax Administration is now going in following directions

- Big Data Analyses
- Artificial Intellect and Machine Learning
- Forecasting Tax Fraud
- Targeting of control
- Behavior Investigation

AI&ML for:

01 ML model for detecting tax fraud

02 Tracking imported products over market chain

03 Effect of audit on behavior of taxpayers

PART ONE

ML model for detecting tax fraud

The model is expected to allow

- More **accurate** identification of fraudulent taxpayers that current system wouldn't identify

1

- More **precise** audit (i.e. increase "success" rate of the audit)

2

3

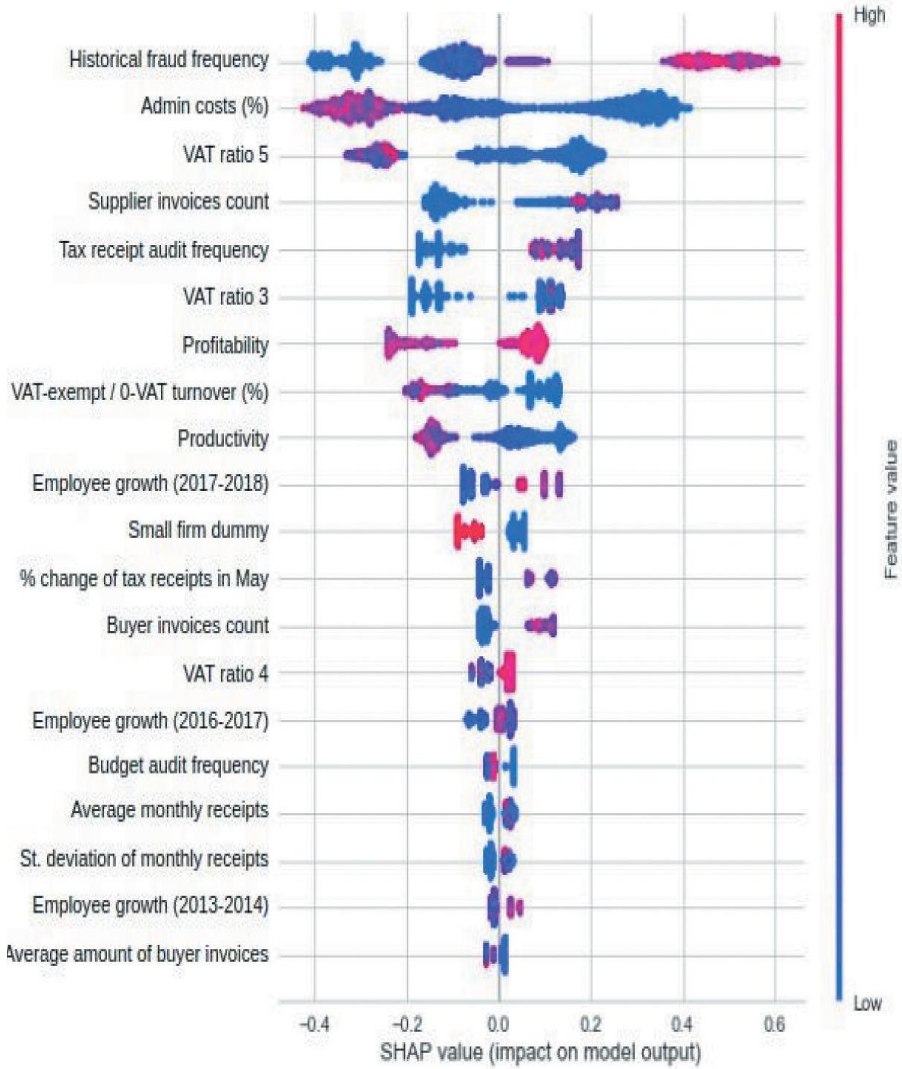
- **Automatic** extraction of rules describing fraudulent behavior

1. **Development of fraud model:** the fraud model considers only taxpayers that were previously audited and considered as fraudulent or not (as otherwise we would not have any information on their evasion).
2. Selection of the best model based on performance metrics - among Logistic Regression, Decision Trees, Random Forest, Gradient Boosting
3. **Use Lift score to focus on high probability fraud:** the model is then used to extract the TINs that have very high likelihood of being fraudulent
4. **Extraction of risk factors:** the fraud model is then used to extract factor, which allow to identify taxpayer groups that have higher propensity to tax evasion.
5. Application of network data.

Table 4. Audit model predictions and Fraud model actual accuracy by deciles.

	Audit model results*			Fraud model results	
Prediction probability decile	Number of taxpayers	Decile probability average	Audited companies in 2020	Fraud detected	Audit accuracy by deciles
1	580	0.677	1125	696	62%
2	138	0.265	386	178	46%
3	49	0.097	236	93	39%
4	23	0.04	48	21	44%
5	7	0.019	5	1	20%
6	0	0.01	66	23	35%
7	1	0.008	78	22	28%
8	0	0.008	190	63	33%
9	0	0.007	134	20	15%
10	0	0.007	273	36	13%
Total taxpayers:	798		2541	1153	45%

*Audit model presents predictions only for comprehensive (budget) audit plans.



PART TWO

Tracking imported products over market chain

Main problem. Tracking a certain set of imported goods - from importing taxpayer to final consumer using structured (codes) and non structured (textual info) contained in invoices, tax receipts and other trade documents available in SRC database.

The experiment was done on 2021 data. The main resources used to solve the task were

- Import data with labeled product codes (11 digits), product definitions
- Tax receipt data with human labeled 4-digit codes and product definitions
- Invoice data (including acc) with product descriptions only (no categorization provided)

ML algorithms used

YAKE for keyword extraction

- frequency based unsupervised algorithm
- domain independent
- language agnostic

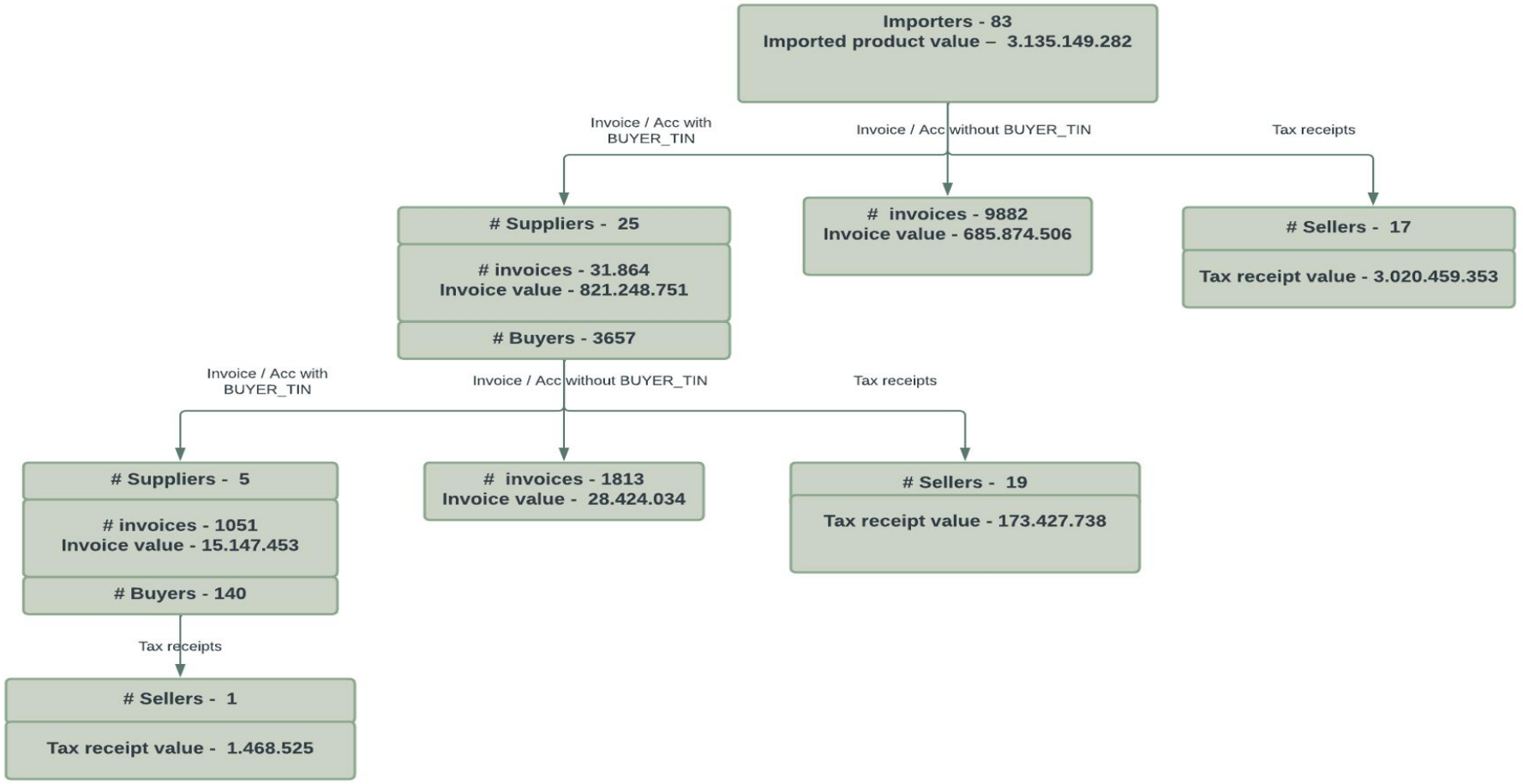
Bidirectional Encoder Representations from Transformers (BERT) for text embedding

- transformer architecture
- contextualized embeddings
- widely used in the literature

Product categorization based on ensembling 3 similarity scores

- Intersection of keywords in armenian, Similarity of keywords in english, Similarity of full description

Experiment results for “Rice” Product



PART THREE

Effect of audit on behavior of taxpayers

1. Focus on taxpayers operating with turnover tax and providing tax receipts
2. Only Yerevan Taxpayers
3. Only Tax receipt audits in 2021

Given that in the previous stage of the project we assisted SRC to assign physical address to each tax receipt machine installed in Yerevan using ML, it is now possible to conduct spatial analysis.

Research Questions:

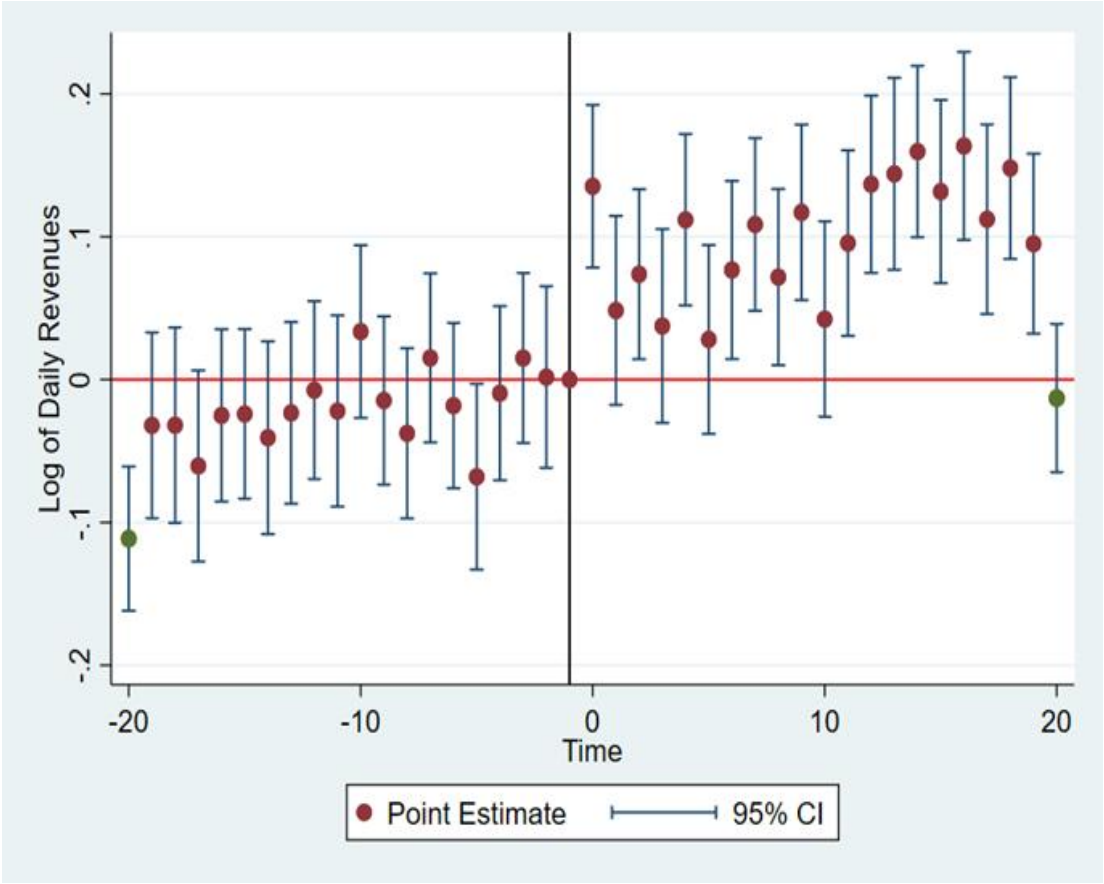
1. Estimate the permanent and temporary effects of audits on own behavior
2. Estimate the permanent and temporary effects of audits on immediate neighbor behavior

Steps implemented

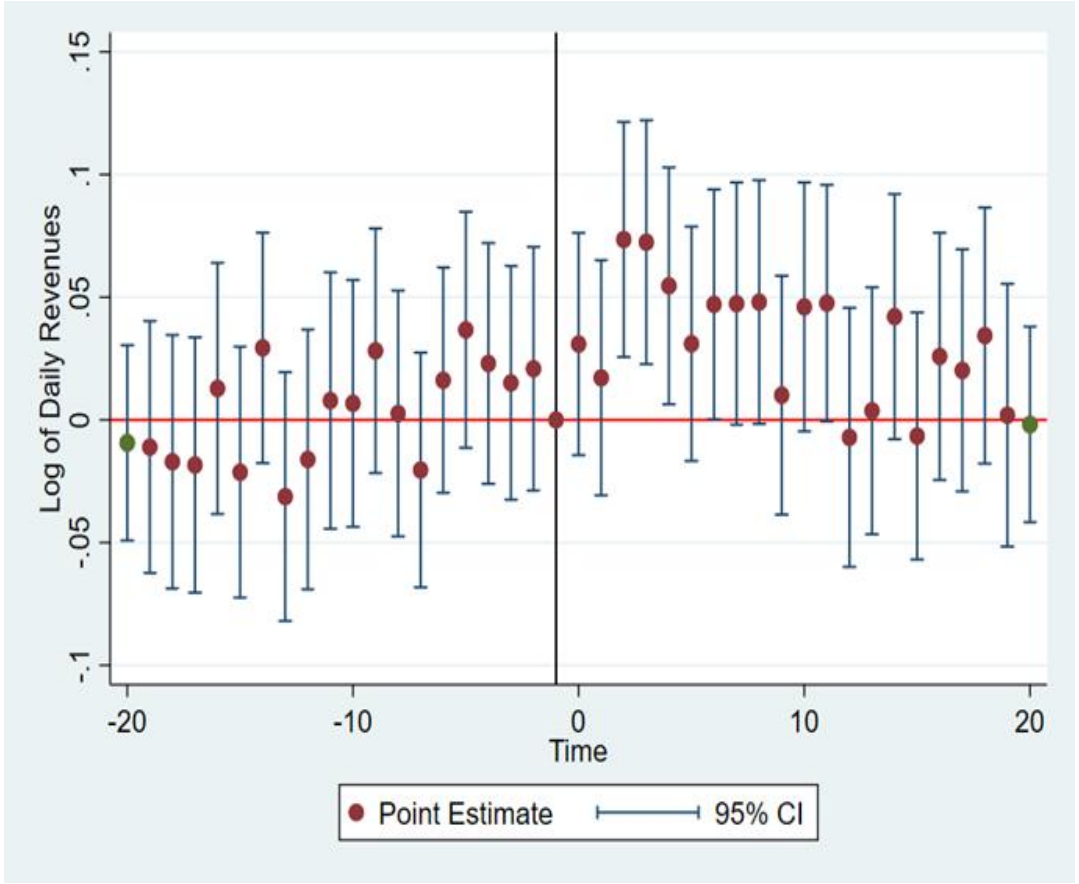
1. Using address information retrieve latitude and longitude
2. Obtain the ranked order of up to 10 nearest neighbors within the radius of 200 meters
3. Merge the obtained dataset with audit data and turnover data (daily observations) to obtain around 13000*250= 3,250,000 observations
4. Apply event study approach to identify the effect of the audit
   - Changes in the reported turnover with leads (prior) and lags (after) the event (audit)
   - Take into account that audit can be effective (resulting in fraud identified) and ineffective
   - Control for firm level individual fixed effects, time trend, weekdays and holidays.

Some results

Effect of ineffective audit on reported turnover:



Effect of neighbor audit on non-audited

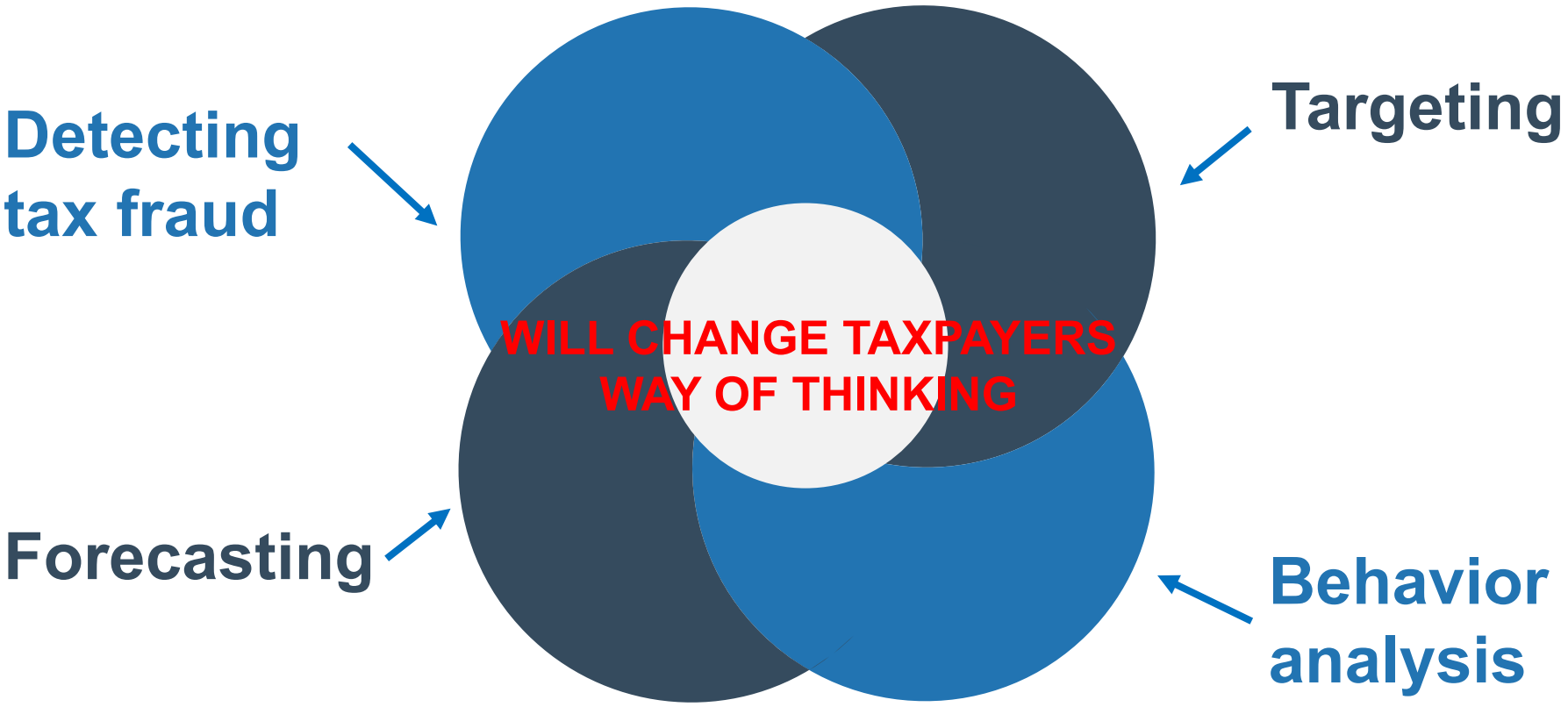


POST SCRIPTUM

Taxes in XXI century

In XXI century everything will be done by

ONE	TWO	THREE
BIG DATA	ML	AI


Detecting tax fraud
Targeting
WILL CHANGE TAXPAYERS WAY OF THINKING
Forecasting
Behavior analysis

AND THEN

There are no taxes → There are contributions

There are no taxpayers → There are partners

BUILDING A GROWTH-FRIENDLY TAX ENVIRONMENT

THANKS