A Behavioral Insights Approach to Nudging Taxpayers in Brazil

Special Secretariat of Federal Revenue of Brazil

1st Fiscal Region (RF01)
7st Fiscal Region (RF07)
Center of Excellence in Artificial Intelligence(CEIA)
Daniel Belmiro
Luis Fernando Kauer
Leon Solon Silva
Jorge Jambeiro Filho

Jorge Jambeiro Filho – Jorge.jambeiro@rfb.gov.br
Tax Auditor of the Special Secretariat of Federal Revenue of Brazil
Head of the Center of Excellence in Artificial Intelligence(CEIA)
CEIA / DISAD / COTEC / RFB

Receita Federal

Behavioral Insights Approach

- Using insights from behavioral sciences
  - psychology, cognitive science, social science, etc.
- To nudge people into making better decisions
- Find empirically what actually affects people's decisions
- Already used inside and outside Brazil
- First experience within the Special Secretariat of Federal Revenue of Brazil

Receita Federal

Variations of the behavioral approach

- Traditional variation of the behavioral insights approach
  - Send different types of letters to taxpayers
    - Traditional letter
    - Letter focused on Simplification
    - Letter focused on Social norms
    - Letter focused on Consequences
    - Etc
  - Find out which letter works better
  - From that point on, use the best letter
- Predefined groups variation
  - Separate taxpayers in groups
  - Find which letter is better for each group
  - From that point on, use the best letter for each group
  - Requires knowledge of how to form the groups

The Machine Learning Variation

- Divide taxpayers in
  - One group for each type of letter
    - including the traditional letter
  - one group for the ML driven letters
- Send different types of letters to taxpayers
  - The ML group should not receive any letters at this point
- Observe the response of each taxpayer
- Build a dataset including
  - Values for several features for each taxpayer
  - The type of letter sent to each taxpayer
  - The outcome (what the taxpayer did)
- Train a machine learning algorithm to predict the outcome
- Ensure the algorithm outputs calibrated probabilities for each possible outcome
- Use the algorithm to predict the outcome for each taxpayer in the ML driven group
  - Vary the type of letter
  - Calculate de return expectation for each type of letter
  - Choose the best letter for this individual taxpayer
- Send the best letter for each taxpayer
  - maximum return expectation
- Observe the response of each taxpayer
  - Find out if choosing letters as described is actually better than sending the best letter overall to every taxpayer.

Traditional experiments

- No segmentation of Taxpayers
- All run in the 7th Fiscal Region
- The types of letters varied, but not much:
  - Social Norms,
  - Social Norms
  - Loss Aversion
  - Emotional
  - Traditional letter
- The in the least successful experiment the best letter was 20% better the traditional letter.
- In the most successful it was 33% better.

Predefined groups experiment

- Run in the 1st Fiscal Region
- We sent letters to 2.489 small companies that earned more during the pandemic than in previous years
- Four types of letters
  - Social Norms,
  - Social Norms plus Simplification,
  - Loss Aversion plus Consequences
  - Traditional letter
- Four groups of taxpayers
  - Low tax evasion risk
  - Moderate tax evasion risk
  - Medium tax evasion risk
  - High tax evasion risk

- Results were different for each group
  - High risk taxpayers
    - Alternative letters were worse
    - Loss of 8.22%.
  - Medium risk taxpayers
    - Social norms and simplification was better
    - Gain of 30.96%
  - Moderate risk taxpayers
    - Social norms and simplification was better
    - Gain of 14.90%
  - Low risk taxpayers
    - Loss aversion was better 41.15%
    - Gain of 41.15%

Machine Learning experiment

- Run in the 1st Fiscal Region
- We sent letters to 1.510 purchasers of rural products from natural people suppliers
  - They are obligated to collect a specific social contribution
- Five types of letters
  - Reminders and appointments, Simplification, Social Norms, Loss Aversion, Traditional
- We had just five groups
  - We has not defined the correct protocol for the experiment at this point
- The letter sent to each taxpayers was chosen at random
  - Because the test set is not so big, some letter end up with better taxpayers than others

- Responses were collected
  - Best results overall: Reminders and appointments
  - 2% better than the traditional letter
  - This result was by far the worse in all experiments
  - We had chosen this experiment to apply machine learning, before seeing this result
- A dataset was built anyway
  - Type of letter
  - Outcome:
    - Paid more taxes
    - Didn't paid more taxes
  - Other attributes: size of the company, age of the company, adherence to prior tax compliance programs, expected amount of taxes to be paid, etc.

Machine Learning experiment

- Dataset split in train and test sets
- Machine learning algorithms trained to predict the outcome
  - Random Forest
- Probabilities were calibrated using Sklearn CalibratedClassifier on top of the RandomForest Classifier
- For each taxpayer in the test set we made 5 predictions, each one considering a different type of letter
- The best letter overall according to predictions was the traditional letter
  - This means that the algorithm considers that “Reminders and appointments” was just the luckiest letter
- Predictions for the policy of sending the best letter for each taxpayer
  - 2.7% better than traditional letter
  - An increase of 2.7% would mean a lot in absolute values
  - However, such a small difference in percentage, can easily be due to overfitting
- We plan a larger experiment before trying the policy in practice.