SALES RETURNERS REPORT

Stefano Bonomi Boseggia - 3129696
Bumin Kagan Cetin - 3128786
Ian Ronk - 3256506
Ossama Tchina - 3137796

Our problem

Reduce returning costs

However:

People want to return items without paying shipping costs

Consumers are more likely to shop elsewhere if they view a return policy change as unfair

Adressing the issue of serial returning

SENTIMENT ANALYSIS

We believe that the reasons as to why products are returned can be found from customer reviews, hence, we developed a sentiment analysis algorithm to search for these reasons.
The process:

- Make the reviews clean and readable
- Skim through and select the ones with low rating
- Keeps and plot the most relevant words
- The managers decide which to rely on and asks the algorithm to search for the review where these words were used
- The manager reads them and make up its mind on the reasons behind sales return

PRICE DISCRIMINATION

Will there be an increase in price discrimination of returning items? Currently, big difference in which items get send back, but is this really the case and is the data correct? Confirmatory check whether prices are as listed in the dataset:

- Get the SKUs
- Retrieve Amazon search results
- Get price of the item (in local currency)
- Convert price
- Compare price to the price in dataset
- Results
- Ethics? Delete data

ALGORITHM PRECISION

We thought it might be very fruitful to improve the current recommendation algorithms to further utilize the customer data in order to improve the accuracy and precision of the algorithm for each specific customer.

- The idea is that if the customers are given better recommendations, the probability that they like the product will increase. Consequently, this will increase their likelihood of returning the product.

BLOCK POLICY

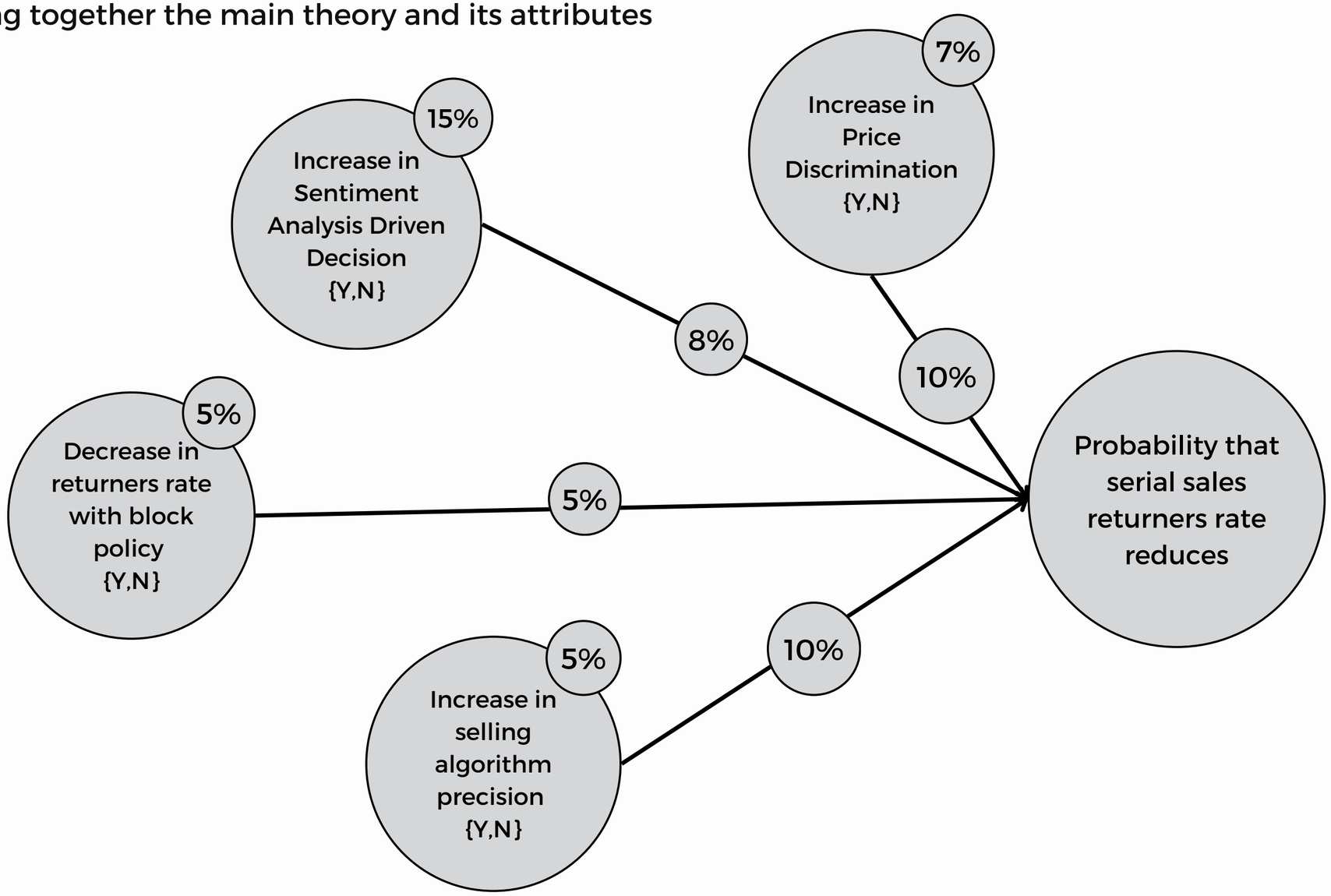
Main idea: punishment to deter unwanted behaviour

Persisting in serial returning

Blocked from returning at all

Main Theory

Putting together the main theory and its attributes

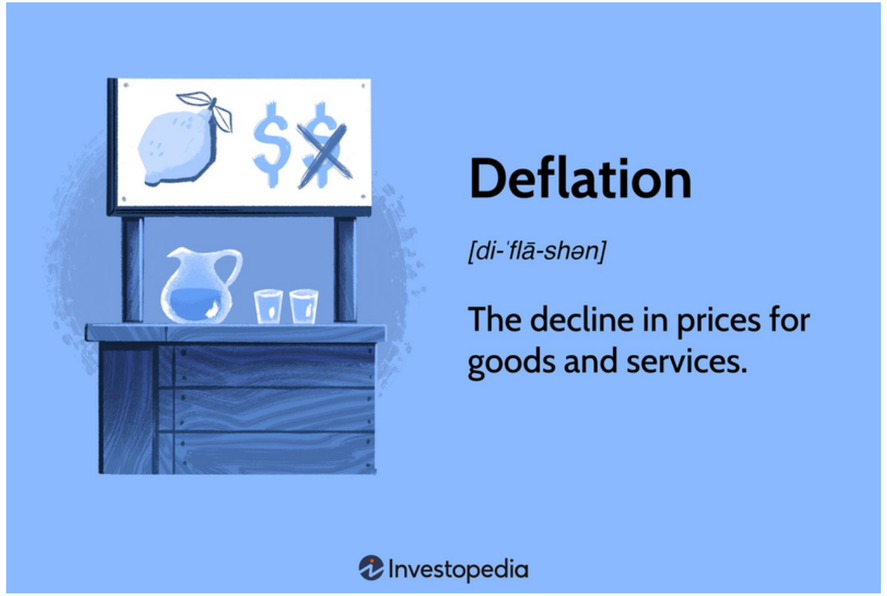


INCREASE IN E-COMMERCE SHARES

- Possible consequential increase in the rate of serial returns by the increase in the shares of products that are sold online rather than in the physical stores.
- 35% increase with a 60% probability of happening

DEFLATION

- Possibility of deflation. It will result in lower demand for products and higher pickiness.
- An increase of 30% the rate of serial returns with a probability of 25%.

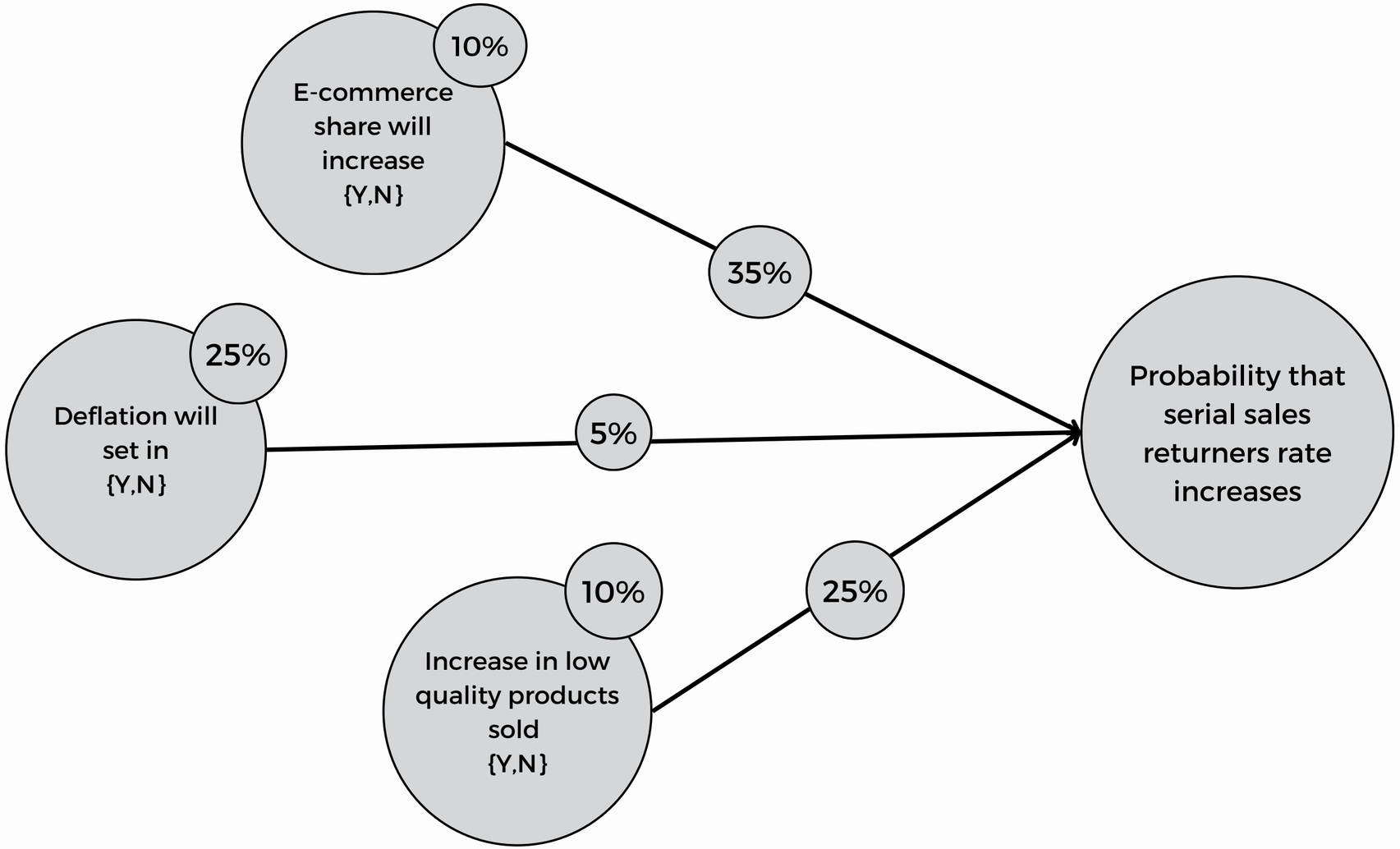


INCREASE IN LOW QUALITY PRODUCTS SOLD

- Possible decrease in the overall quality of certain products in the economy
- An increase of 25% with a probability of 10%

Alternative Theory

Putting together the alternative theory and its attributes



PRIOR

The Main Theory VS The Alternative Theory

THE UNCONDITIONAL EXPECTED VALUE:

$$V = \omega V_{\Theta} + (1-\omega)V_{\bar{\Theta}}$$

9.9%

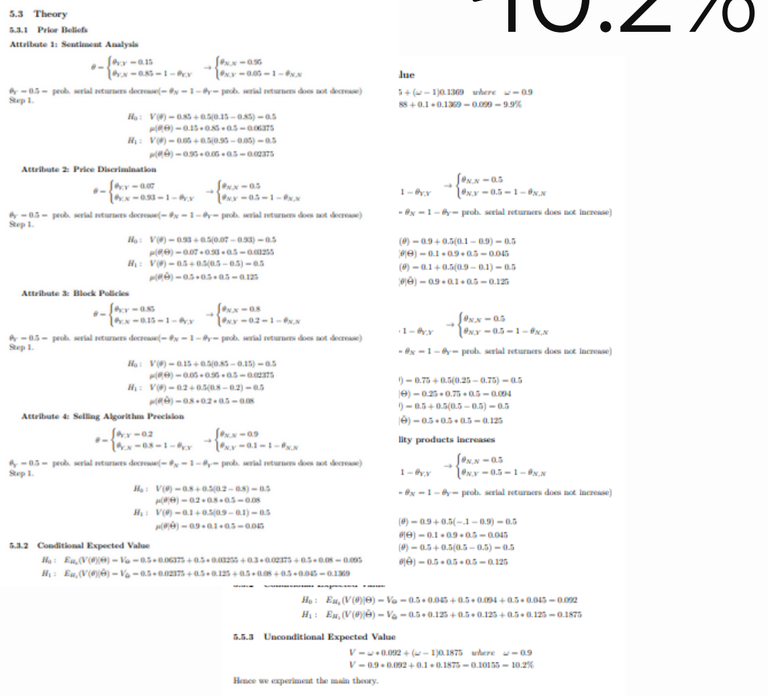
10.2%

We test for
the Block Policy of
the Main Theory

Hypothesis to test:

$\Delta\theta > 0$

$\theta(y,y) > \theta(y,n)$



5.3 Theory

5.3.1 Price Beliefs

Attribute 1: Sentiment Analysis

$$\theta = \begin{cases} \theta_{Y,Y} = 0.15 \\ \theta_{Y,N} = 0.85 = 1 - \theta_{Y,Y} \end{cases} \rightarrow \begin{cases} \theta_{N,N} = 0.95 \\ \theta_{N,Y} = 0.05 = 1 - \theta_{N,N} \end{cases}$$

$\theta_Y = 0.5 =$ prob. serial returners decrease($= \theta_N = 1 - \theta_Y =$ prob. serial returners does not decrease)
Step 1.

$$H_0: \ V(\theta) = 0.85 + 0.5(0.15 - 0.85) = 0.5$$
$$\mu(\theta|\Theta) = 0.15 * 0.85 * 0.5 = 0.06375$$
$$H_1: \ V(\bar{\theta}) = 0.05 + 0.5(0.95 - 0.05) = 0.5$$
$$\mu(\theta|\bar{\Theta}) = 0.95 * 0.05 * 0.5 = 0.02375$$

Attribute 2: Price Discrimination

$$\theta = \begin{cases} \theta_{Y,Y} = 0.07 \\ \theta_{Y,N} = 0.93 = 1 - \theta_{Y,Y} \end{cases} \rightarrow \begin{cases} \theta_{N,N} = 0.5 \\ \theta_{N,Y} = 0.5 = 1 - \theta_{N,N} \end{cases}$$

$\theta_Y = 0.5 =$ prob. serial returners decrease($= \theta_N = 1 - \theta_Y =$ prob. serial returners does not decrease)
Step 1.

$$H_0: \ V(\theta) = 0.93 + 0.5(0.07 - 0.93) = 0.5$$
$$\mu(\theta|\Theta) = 0.07 * 0.93 * 0.5 = 0.03255$$
$$H_1: \ V(\bar{\theta}) = 0.5 + 0.5(0.5 - 0.5) = 0.5$$
$$\mu(\theta|\bar{\Theta}) = 0.5 * 0.5 * 0.5 = 0.125$$

Attribute 3: Block Policies

$$\theta = \begin{cases} \theta_{Y,Y} = 0.85 \\ \theta_{Y,N} = 0.15 = 1 - \theta_{Y,Y} \end{cases} \rightarrow \begin{cases} \theta_{N,N} = 0.8 \\ \theta_{N,Y} = 0.2 = 1 - \theta_{N,N} \end{cases}$$

$\theta_Y = 0.5 =$ prob. serial returners decrease($= \theta_N = 1 - \theta_Y =$ prob. serial returners does not decrease)
Step 1.

$$H_0: \ V(\theta) = 0.15 + 0.5(0.85 - 0.15) = 0.5$$
$$\mu(\theta|\Theta) = 0.05 * 0.95 * 0.5 = 0.02375$$
$$H_1: \ V(\bar{\theta}) = 0.2 + 0.5(0.8 - 0.2) = 0.5$$
$$\mu(\theta|\bar{\Theta}) = 0.8 * 0.2 * 0.5 = 0.08$$

Attribute 4: Selling Algorithm Precision

$$\theta = \begin{cases} \theta_{Y,Y} = 0.2 \\ \theta_{Y,N} = 0.8 = 1 - \theta_{Y,Y} \end{cases} \rightarrow \begin{cases} \theta_{N,N} = 0.9 \\ \theta_{N,Y} = 0.1 = 1 - \theta_{N,N} \end{cases}$$

$\theta_Y = 0.5 =$ prob. serial returners decrease($= \theta_N = 1 - \theta_Y =$ prob. serial returners does not decrease)
Step 1.

$$H_0: \ V(\theta) = 0.8 + 0.5(0.2 - 0.8) = 0.5$$
$$\mu(\theta|\Theta) = 0.2 * 0.8 * 0.5 = 0.08$$
$$H_1: \ V(\bar{\theta}) = 0.1 + 0.5(0.9 - 0.1) = 0.5$$
$$\mu(\theta|\bar{\Theta}) = 0.9 * 0.1 * 0.5 = 0.045$$

5.3.2 Conditional Expected Value

$$H_0: \ E_{H_0}(V(\theta)|\Theta) = V_{\Theta} = 0.5 + 0.06375 + 0.5 * 0.03255 + 0.3 + 0.02375 + 0.5 * 0.08 = 0.095$$
$$H_1: \ E_{H_1}(V(\theta)|\bar{\Theta}) = V_{\bar{\Theta}} = 0.5 + 0.02375 + 0.5 * 0.125 + 0.5 * 0.08 + 0.5 * 0.045 = 0.1369$$

lue

$5 + (\omega - 1)0.1369 \quad where \quad \omega = 0.9$
$88 + 0.1 + 0.1369 = 0.099 = 9.9\%$

$1 - \theta_{Y,Y} \rightarrow \begin{cases} \theta_{N,N} = 0.5 \\ \theta_{N,Y} = 0.5 = 1 - \theta_{N,N} \end{cases}$

$= \theta_N = 1 - \theta_Y =$ prob. serial returners does not increase)

$(\theta) = 0.9 + 0.5(0.1 - 0.9) = 0.5$
$\theta|\Theta) = 0.1 * 0.9 * 0.5 = 0.045$
$(\theta) = 0.1 + 0.5(0.9 - 0.1) = 0.5$
$\theta|\bar{\Theta}) = 0.9 * 0.1 * 0.5 = 0.125$

$\cdot 1 - \theta_{Y,Y} \rightarrow \begin{cases} \theta_{N,N} = 0.5 \\ \theta_{N,Y} = 0.5 = 1 - \theta_{N,N} \end{cases}$

$= \theta_N = 1 - \theta_Y =$ prob. serial returners does not increase)

$) = 0.75 + 0.5(0.25 - 0.75) = 0.5$
$|\Theta) = 0.25 * 0.75 * 0.5 = 0.094$
$) = 0.5 + 0.5(0.5 - 0.5) = 0.5$
$|\bar{\Theta}) = 0.5 * 0.5 * 0.5 = 0.125$

lity products increases

$1 - \theta_{Y,Y} \rightarrow \begin{cases} \theta_{N,N} = 0.5 \\ \theta_{N,Y} = 0.5 = 1 - \theta_{N,N} \end{cases}$

$= \theta_N = 1 - \theta_Y =$ prob. serial returners does not increase)

$(\theta) = 0.9 + 0.5(-.1 - 0.9) = 0.5$
$\theta|\Theta) = 0.1 * 0.9 * 0.5 = 0.045$
$(\theta) = 0.5 + 0.5(0.5 - 0.5) = 0.5$
$\theta|\bar{\Theta}) = 0.5 * 0.5 * 0.5 = 0.125$

$$H_0: \ E_{H_0}(V(\theta)|\Theta) = V_{\Theta} = 0.5 * 0.045 + 0.5 * 0.094 + 0.5 * 0.045 = 0.092$$
$$H_1: \ E_{H_1}(V(\theta)|\bar{\Theta}) = V_{\bar{\Theta}} = 0.5 * 0.125 + 0.5 * 0.125 + 0.5 * 0.125 = 0.1875$$

5.3.3 Unconditional Expected Value

$$V = \omega * 0.092 + (\omega - 1)0.1875 \quad where \quad \omega = 0.9$$
$$V = 0.9 * 0.092 + 0.1 * 1875 = 0.10155 = 10.2\%$$

Hence we experiment the main theory.

Survey

COVARIATES :

- GENDER
- AGE
- CULTURAL BACKGROUND
- EDUCATION
- AVG MONTHLY SPENDING

A/B TESTING

A/B testing

CONTROL

TREATMENT

RANDOMIZATION:

→ through Qualtrics

→

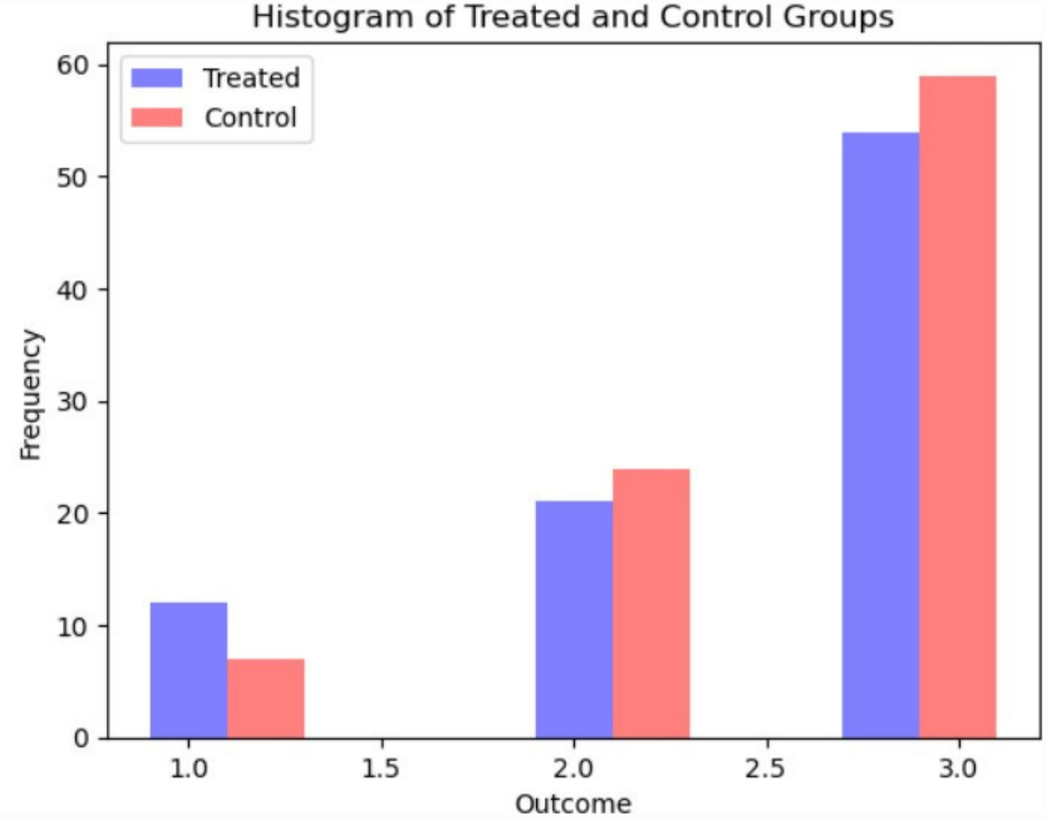
Asia - 48%
Europe - 45%
Africa - 9%
America - 2%
Oceania - 0%

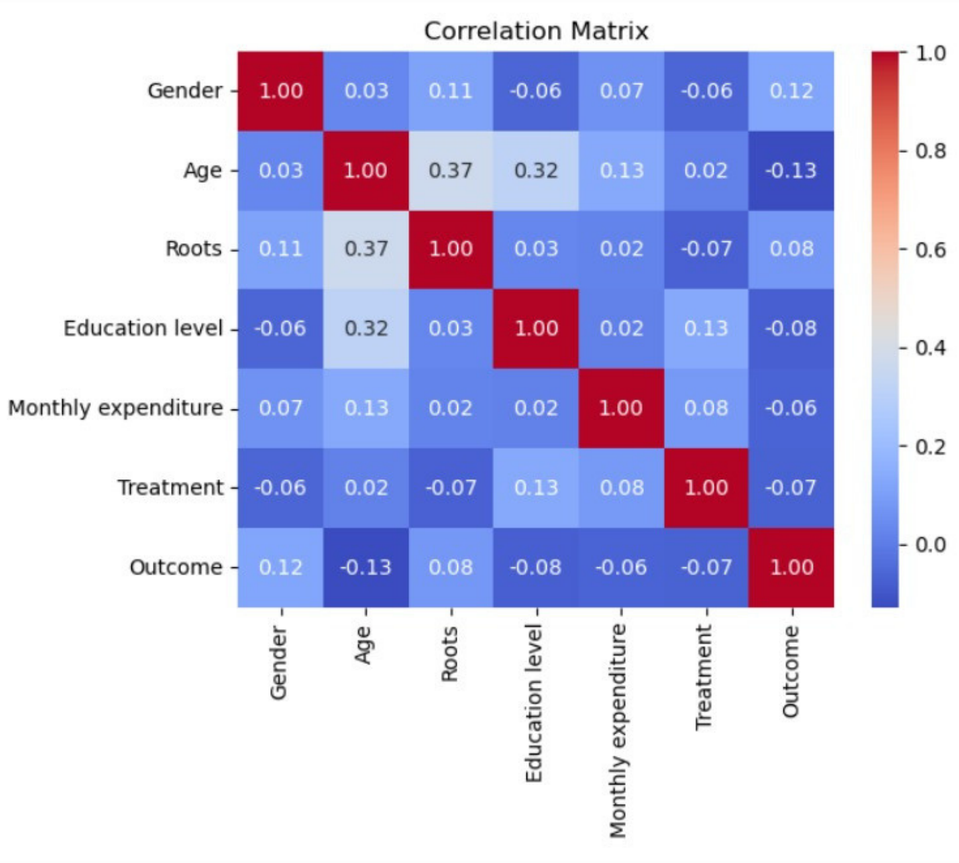
Male - 56%
Female - 44%

Data analysis and Preprocessing

Through python script we cleaned and ordered the data, eliminated the NaN values and obtained a dataset of around 100 observations (from 250 respondent) and 6 covariates.

We then plot some graph to inspect the data distribution visually.





Regression

A multinomial logistic regression have been used.

Treatment coefficient for Outcome = 3 as a proxy for our updated teta

Only Q1 for Outcome=3 is a statistically significant coefficient for %5 significance level.

Need for more observations and more covariates.

```
Optimization terminated successfully.
         Current function value: 0.831451
         Iterations 6
Multinominal Logistic Regression Results:
+-----------+---------+---------+--------+-------+--------+--------+
|     0     |    1    |    2    |   3    |   4   |   5    |   6    |
+-----------+---------+---------+--------+-------+--------+--------+
| Outcome=2 |  coef   | std err |   z    | P>|z| | [0.025 | 0.975] |
|   const   | 1.9012  |  1.317  | 1.444  | 0.149 | -0.680 | 4.482  |
|    Q1     | -0.4496 |  0.256  | -1.753 | 0.080 | -0.952 | 0.053  |
|    Q2     | 0.4279  |  0.585  | 0.731  | 0.465 | -0.719 | 1.575  |
|    Q3     | -0.1820 |  0.312  | -0.583 | 0.560 | -0.794 | 0.430  |
|    Q6     | 0.1010  |  0.400  | 0.252  | 0.801 | -0.683 | 0.885  |
|   Q9_1    | -0.0014 |  0.002  | -0.701 | 0.483 | -0.005 | 0.002  |
| Treatment | -0.6381 |  0.590  | -1.082 | 0.279 | -1.794 | 0.517  |
| Outcome=3 |  coef   | std err |   z    | P>|z| | [0.025 | 0.975] |
|   const   | 1.8984  |  1.205  | 1.576  | 0.115 | -0.463 | 4.260  |
|    Q1     | -0.4888 |  0.228  | -2.140 | 0.032 | -0.936 | -0.041 |
|    Q2     | 0.7192  |  0.538  | 1.337  | 0.181 | -0.335 | 1.773  |
|    Q3     | 0.1982  |  0.282  | 0.702  | 0.483 | -0.355 | 0.752  |
|    Q6     | -0.0214 |  0.359  | -0.059 | 0.953 | -0.725 | 0.683  |
|   Q9_1    | -0.0013 |  0.002  | -0.748 | 0.454 | -0.005 | 0.002  |
| Treatment | -0.4861 |  0.539  | -0.902 | 0.367 | -1.542 | 0.570  |
+-----------+---------+---------+--------+-------+--------+--------+
```

Results

After updating our prior beliefs, the expected value of the main theory (9.7%) remains lower than the one one of the alternative one (10.2%). If we have to stop here, we would certainly implement the alternative theory. We are aware that continuing experimenting is the best idea eventually.

THE UNCONDITIONAL EXPECTED VALUE:

$$V = \omega V_{\Theta} + (1 - \omega) V_{\tilde{\Theta}}$$

9.7%

10.2%

5.6 Updating Beliefs

5.6.1 Attribute 3: Block Policy

$$\theta = \begin{cases} \theta_{Y,Y} = |-0.4861| \\ \theta_{Y,N} = 0.5139 = 1 - \theta_{Y,Y} \end{cases} \rightarrow \begin{cases} \theta_{N,N} = 0.8 \\ \theta_{N,Y} = 0.2 = 1 - \theta_{N,N} \end{cases}$$

$$\begin{aligned} H_0: \quad & V(\theta) = 0.51 + 0.5(0.49 - 0.51) = 0.5 \\ & \mu(\theta|\Theta) = 0.49 * 0.51 * 0.5 = 0.12495 \\ H_1: \quad & V(\theta) = 0.5 \\ & \mu(\theta|\tilde{\Theta}) = 0.08 \end{aligned}$$

5.6.2 Conditional Expected Value

$$H_0: \quad E_{H_0}(V(\theta)|\Theta) = V_{\Theta} = 0.5 * 0.06375 + 0.5 * 0.03255 + 0.5 * 0.012485 + 0.5 * 0.08 = 0.058$$

$$H_1: \quad E_{H_1}(V(\theta)|\tilde{\Theta}) = V_{\tilde{\Theta}} = 0.5 * 0.02375 + 0.5 * 0.125 + 0.5 * 0.08 + 0.5 * 0.045 = 0.1369$$

5.6.3 Unconditional Expected Value

$$V = \omega * 0.058 + (\omega - 1)0.1369 \quad where \quad \omega = 0.5$$

$$V = 0.5 * 0.058 + 0.5 * 0.1369 = 0.097 = 9.7\%$$

We go for Alternative Theory

Amazon & C.: sta finendo l'era dei resi gratis?

di Riccardo Staglianò

Fotografia di Luigi Narici/Agf

GALAPAGOS Sempre più compagnie di ecommerce cominciano a farsi pagare le restituzioni

24 NOVEMBRE 2023 ALLE 18:40

1 MINUTI DI LETTURA

Conclusion & Discussion

If we had more/infinite budget we would have improved the following aspects:

of covariates

Increasing the number of covariates would have helped us in better restoring randomization, checking balances and unveiling heterogeneity.

of attributes

More attributes would have certainly increased the precision of our model (also its complexity but we are assuming we have a huge budget).

of experiments

More experimentation would have helped in reducing the uncertainty around our main theory.

THANK YOU

For Your Attention

Q&A