# Contact timing in telemarketing campaigns

## Machine Learning for econometrics

Eléa Bordais, Jade Sillere, Eliott Von-Pine, Patryk Wiśniewski

Tuesday, April 1, 2025

# Table of contents

# Introduction

**Objectives**: Identify the factors that influence a customer's likelihood of subscribing to term deposits during a marketing campaign.

**Contribution**: Identify the impact of the time of the call on the likelihood of subscribing.

**Dataset**: Marketing dataset from a Portuguese banking institution of a marketing campaign from May 2008 to November 2010

**Methods**: Compare traditional causal inference methods with modern machine-learning based approaches

# Data overview

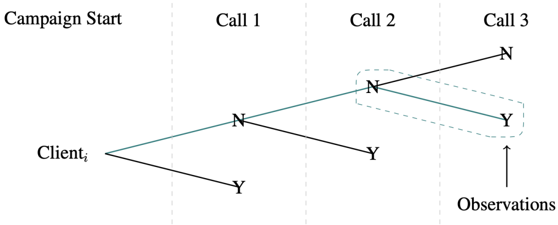


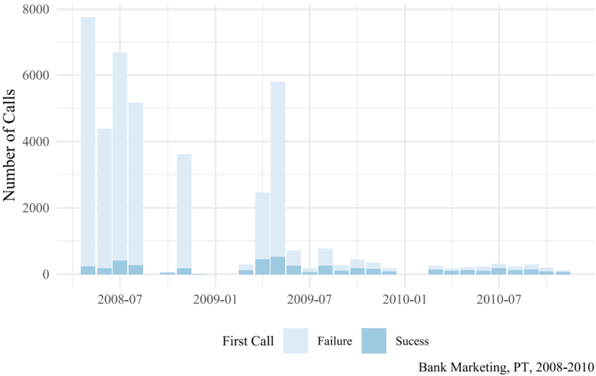

Figure: Campaign process and observed data

- Two types of contact : Inbound calls and outbound calls
- Most follow-up calls are at the request of the client

## Overview of Variables

- **Outcome:** client subscribed to term deposit? (Yes or No)
- **Client variables:** age, job, marital status, education, loans, etc.
- **Contact variables:** communication type, month and day of last call, duration, number of contacts, info about previous campaign
- **State of the economy variables:** employment rate, consumer price index, consumer confidence index, 12-month euribor rate, etc.

# Outcome and Campaign Organization




- Outcome variable: Imbalance of "no" (88.7%)

# Macroeconomic context

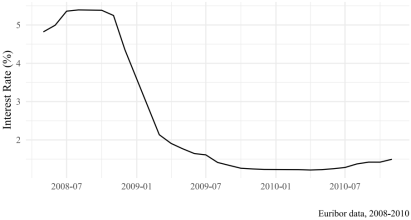



Figure: Euribor 12mo rates

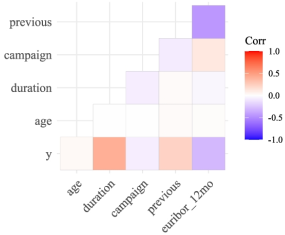



Figure: Correlation

A priori negative relation between interest rates and outcome, why?

# Time of Week: Naive approach

| Share | Monday | Tuesday | Wednesday | Thursday | Friday |
|---|---|---|---|---|---|
| Failure | 0.901 | 0.882 | 0.883 | 0.879 | 0.892 |
| Success | 0.099 | 0.118 | 0.117 | 0.121 | 0.108 |
| Count | 8514 | 8090 | 8134 | 8623 | 7823 |

Table: Share of successful calls

- Calls close to uniformly distributed across days
- Higher success probabilities in the middle

# Timing of contact

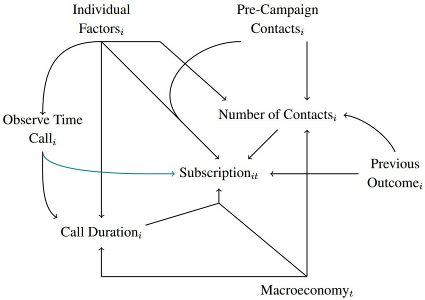
- Hypotheses:
  - Calls in the middle of the week are more successful
  - Calls at the end of the month are more successful
- Idea:
  - Leisure/Business mindset: in the middle of the week people are more rational, planning financials
  - Catching-up/deadlines on Mondays/Fridays
  - Employees receive wage at the end month

# Timing of contact - PICO Formulation

| Component | Description |
|---|---|
| **Population** | Customers from a Portuguese bank, contacted as part of a telemarketing campaign for a term deposit product |
| **Intervention** | Receiving a call during the middle of the week (Tue, Wed, Thur) / at the end of the month ($>$ 16th) |
| **Comparison** | Customers contacted another day |
| **Outcome** | Subscribing or not to the term deposit (probability) |

# Directed Acyclic Graph

Figure 6: Relation Between Call-Time and Take-Up




- **Post-treatment:** *duration*
- **Unconfoundedness:** Bank calls uniformly across days
- **Overlap:** Treated and control have comparable characteristics

# Data pre-processing and imputation

- Recoded missing values
- Filtered for clients only contacted a single time during the campaign
- Converted the numeric pdays value to 3 categories (never contacted, contacted within a week and contacted beyond a week)
- Imputation using the MICE approach

# Imputation Method: `mice` Package

Assuming missing values are MAR (Missing At Random), I used the `mice` package with default imputation functions to generate 5 imputed datasets:

- `pmm`: Predictive Mean Matching (numeric data)
- `logreg`: Logistic Regression (binary data, 2 levels)
- `polyreg`: Polytomous Regression (unordered categorical, $>2$ levels)
- `polr`: Proportional Odds Model (ordered categorical, $>2$ levels)

The number of imputations is supported by literature (Bennet, 2001; Hawthorne and Eliott, 2005 and Royston et al.).

# Estimators

- Simple Logit
- Double Machine Learning - PLM and IRM
  - Random forest + tuning hyperparameters
  - Ensemble (for PLM only) using the following estimators for the classifiers/regressors:
    - Random Forest, Logit (classifier only), Neural Net, linear model (regressor only), XGBoost (regressor only)
- Causal Random Forest

# Moment of the week

Table 3: Change in take-up in middle of the week

| Statistic | C-FOREST | MLPLR | MLPLR-ENS | MLIRM | LOGIT |
|---|---|---|---|---|---|
| Estimate | 0.0125 | 0.0122 | 0.0087 | 0.0103 | 0.0151 |
| Std Errors | 0.0047 | 0.0047 | 0.0041 | 0.0048 | 0.0048 |
| p-value | 0.0079 | 0.0101 | 0.0314 | 0.0309 | 0.0015 |
| CI 95% | 0.0033-0.0217 | 0.0029-0.0215 | 0.0008-0.0167 | 0.0009-0.0196 | 0.0057-0.0244 |

- Statistically significant results in all models
- Largest coefficient with logit, smallest with PLR-ENS

# Moment of the Month

Table 4: Change in take-up end of month

| Statistic | C-FOREST | MLPLR | MLPLR-ENS | MLIRM | LOGIT |
|---|---|---|---|---|---|
| Estimate | 0.0117 | 0.0136 | 0.0079 | 0.0139 | 0.0148 |
| Std Errors | 0.0061 | 0.0060 | 0.0054 | 0.0061 | 0.0050 |
| p-value | 0.0533 | 0.0242 | 0.1447 | 0.0223 | 0.0032 |
| CI 95% | -0.0002-0.0237 | 0.0018-0.0255 | -0.0027-0.0186 | 0.0020-0.0259 | 0.0050-0.0246 |

- Coefficients not always significant
- Again, largest coefficient with logit, smallest with PLR-ENS

# Moment of the Week (Heterogeneity)

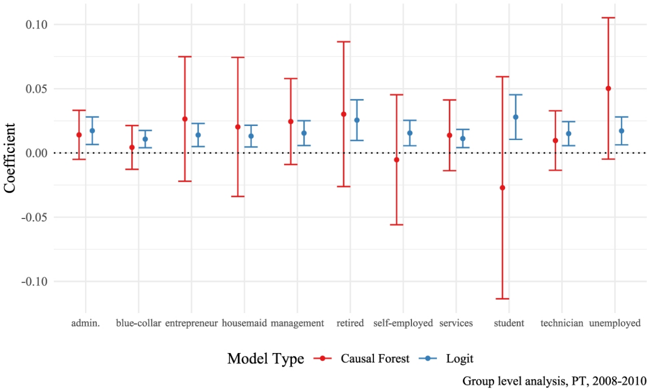



Figure 7: Change in take-up in middle of the week

# Moment of the Month (Heterogeneity)

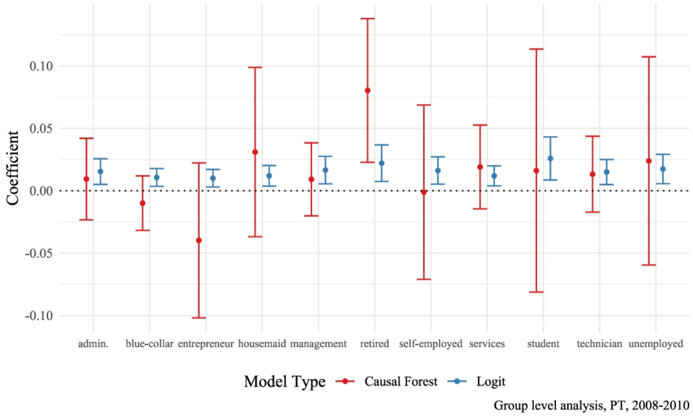
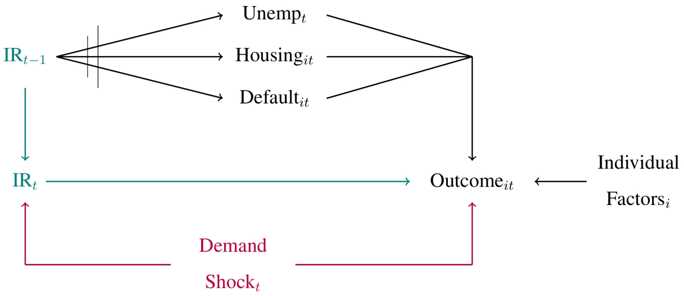



Figure 8: Change in take-up at the end of the month

# DAG, PICO and Macro

Figure 9: Interest Rates and Take-Up




Interest rates are prices, so they relate both to demand and supply.

# Interest Rate Effects

Table 5: Results from IV estimation

| Statistic | LPM2SLS-IV | DML-IV | Probit-IV |
|---|---|---|---|
| Estimate | 0.0473 | 0.0327 | 0.0185 |
| Std Errors | 0.0051 | 0.0028 | 0.0046 |
| p-value | 0.0000 | 0.0000 | 0.0001 |
| Compute (s) | 0.0577 | 39.5599 | 0.2065 |

- LPM likely over-estimating the effects

- DML-IV less rigid than Probit-IV

# Conclusion

Take home message:

- Timing in marketing matters not only *who?* but *when?*
- Careful about endogeneity when working with prices

Some of the limits of what we've shown:

- Individuals that do not pickup are not observed
- Simulation exercise needed to better compare methods
- External validity not guaranteed
- Instrument validity