

# **Improving contact rates in the field through analysis of linked Census survey data**

Folasade Ariyibi

# Overview

---

- Introduction
- Matching Process
- Characteristics of Non-contacts
- Paradata
- Multivariate Analysis
- Examples of Profiles
- Conclusion

# Introduction

---

- Response rates have been declining over time
- Constraints used in weighting may account for some non-response bias but not all
- Weighting based on incorrect assumptions may lead to bias results
- Better to try and get more representative responses as oppose to compensating for non-response

# Introduction

---

- The ONS carry out a Census every 10 years
- 2011 Census response rate – 94% overall
- At the same period as the Census, other social surveys are run in field.
- Six social surveys were selected for inclusion in this study.
- Census Non-Response Link Study (CNRLS)

# Introduction

---

CNRLS involves matching Census data with social survey data to analyse patterns of non-response. This allows for:

- Analysis of potential bias in social surveys
- Opportunity to inform field force of better collection strategies (contact and co-operation)

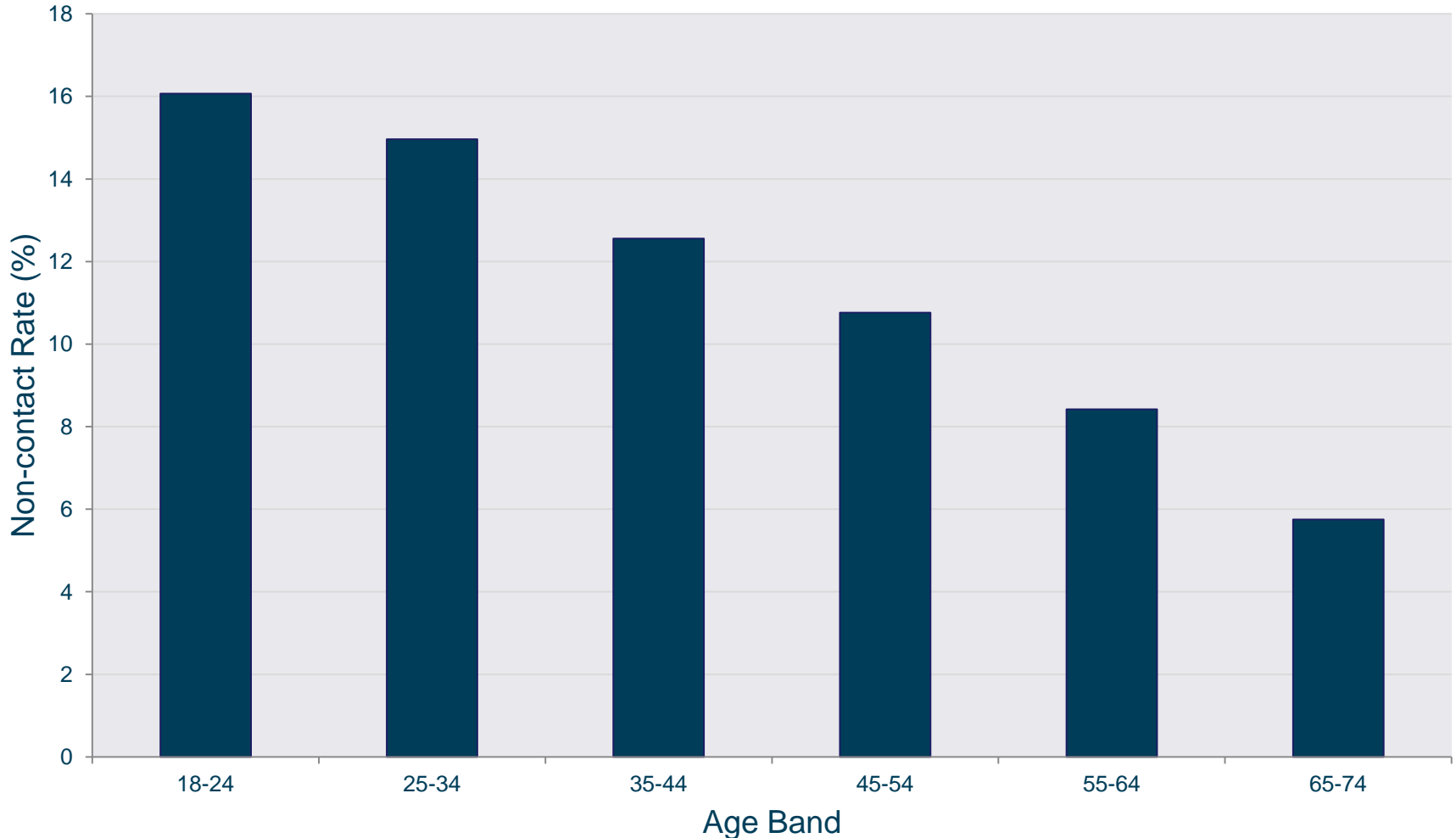
# Matching Process

---

- Matching carried out by Census-matching staff in ONS
- Matching at household and individual level
- Links survey-census addresses around 2011 Census date
- Match rates 94%
- Analysis restricted to Wave 1 for panel/longitudinal surveys

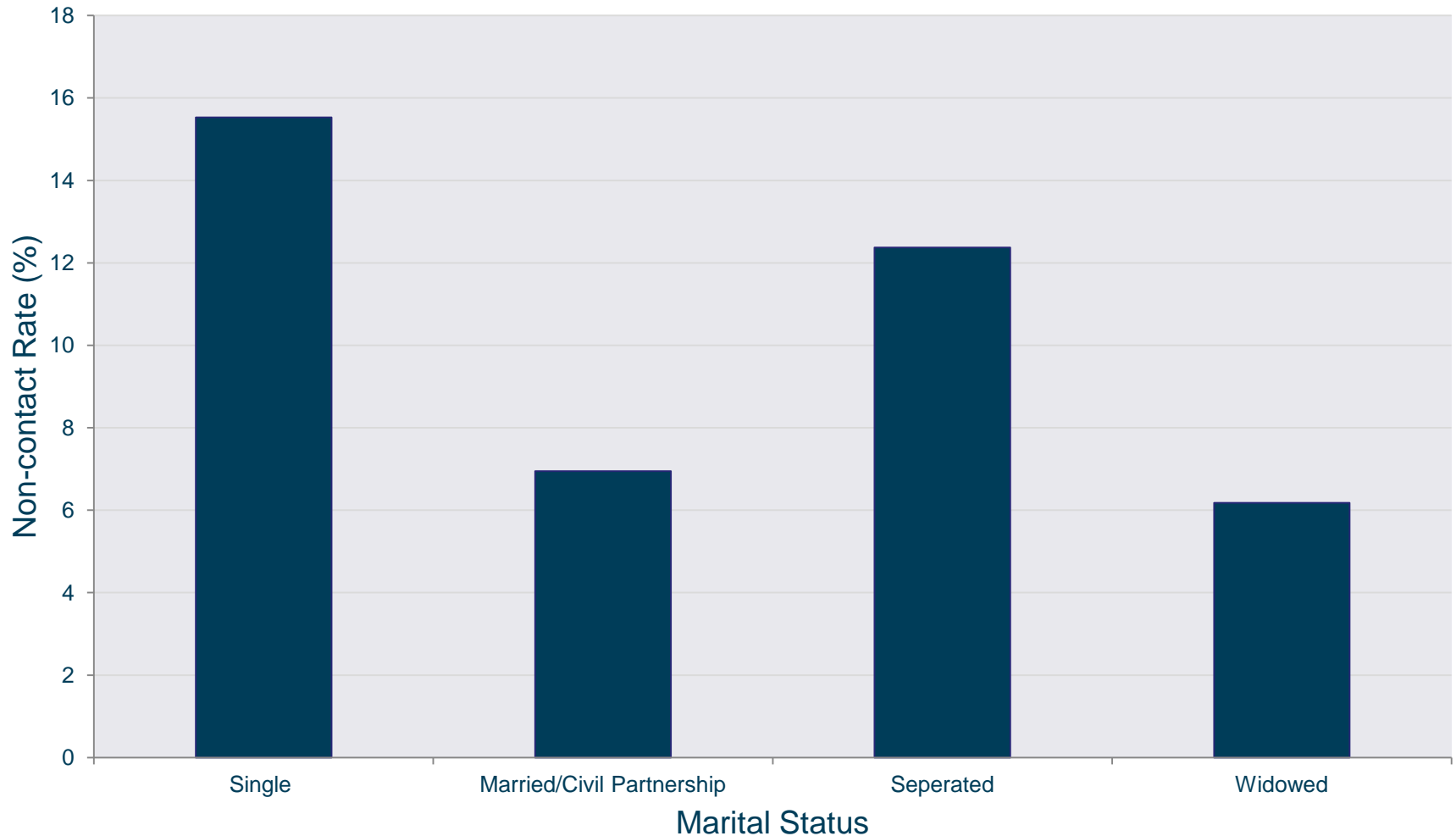
# Characteristics of Non-contacts

## Non-contact Rates by Age



# Characteristics of Non-contacts

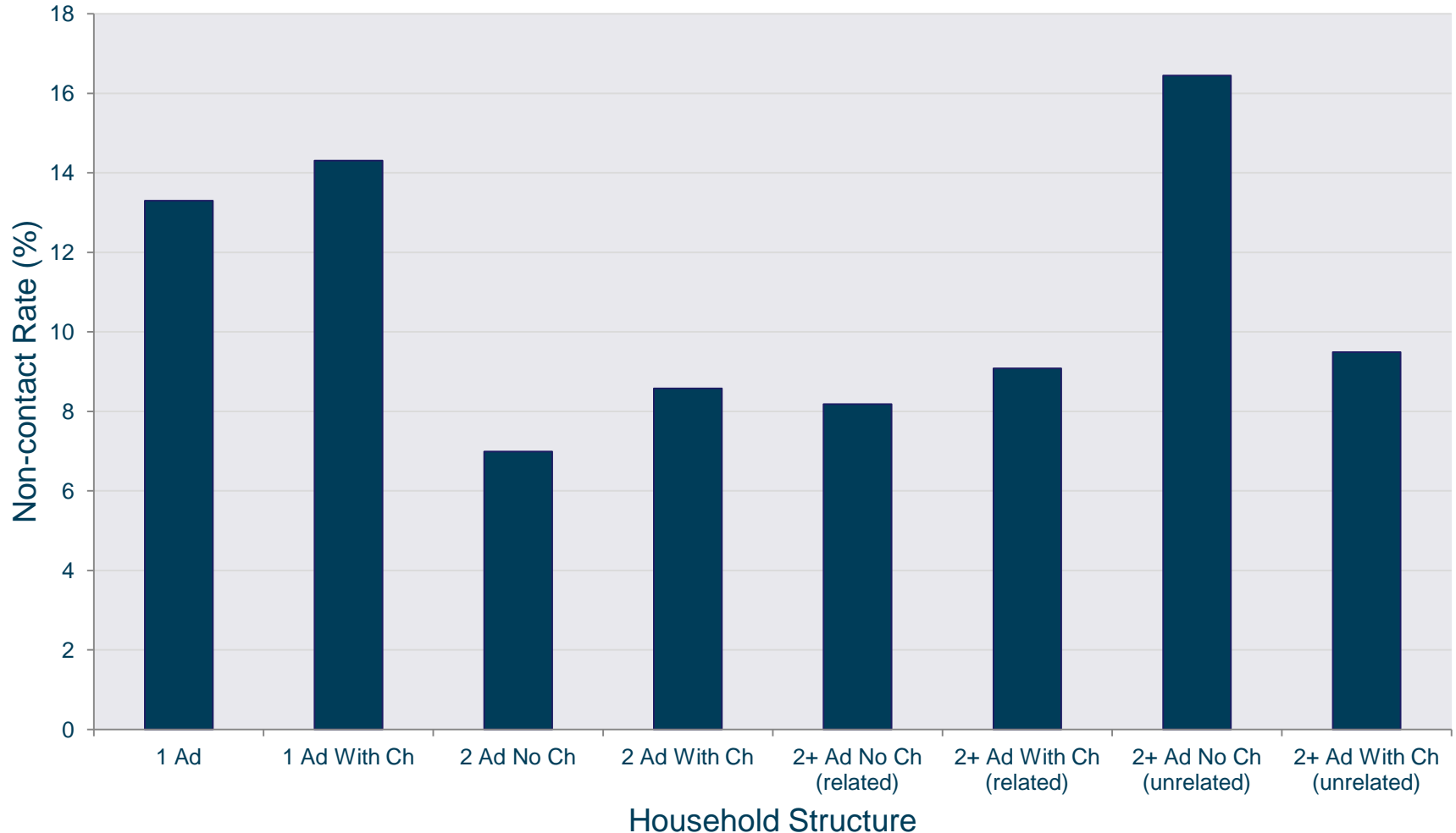
## Non-contact Rates by Marital Status





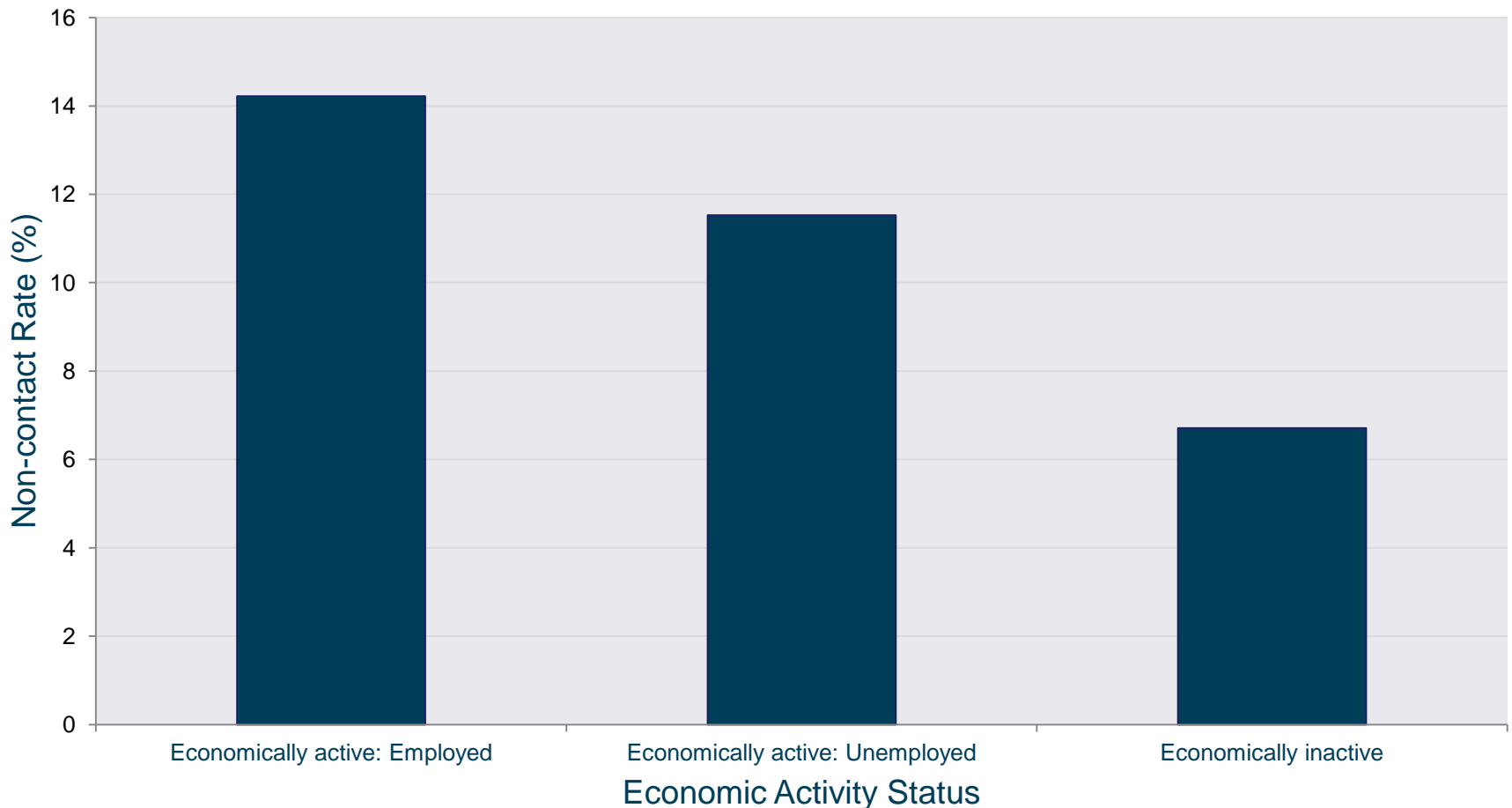
# Characteristics of Non-contacts

## Non-contact Rate by Household Structure



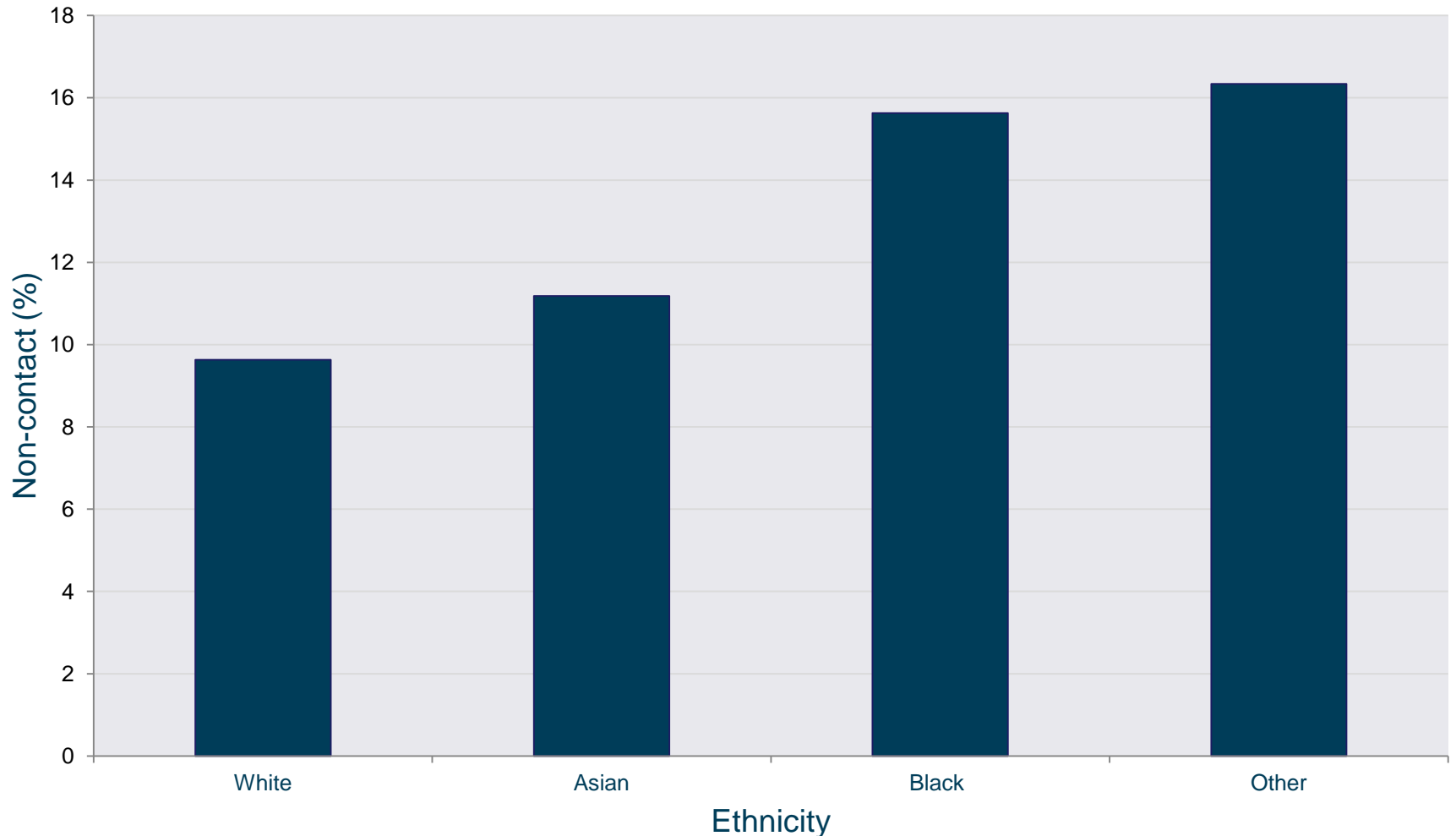
# Characteristics of Non-contacts

## Non-contact Rate by Economic Activity Status



# Characteristics of Non-contacts

## Non-contact Rate by Ethnicity



---

**But how will this help the field  
force?**

# Paradata

---

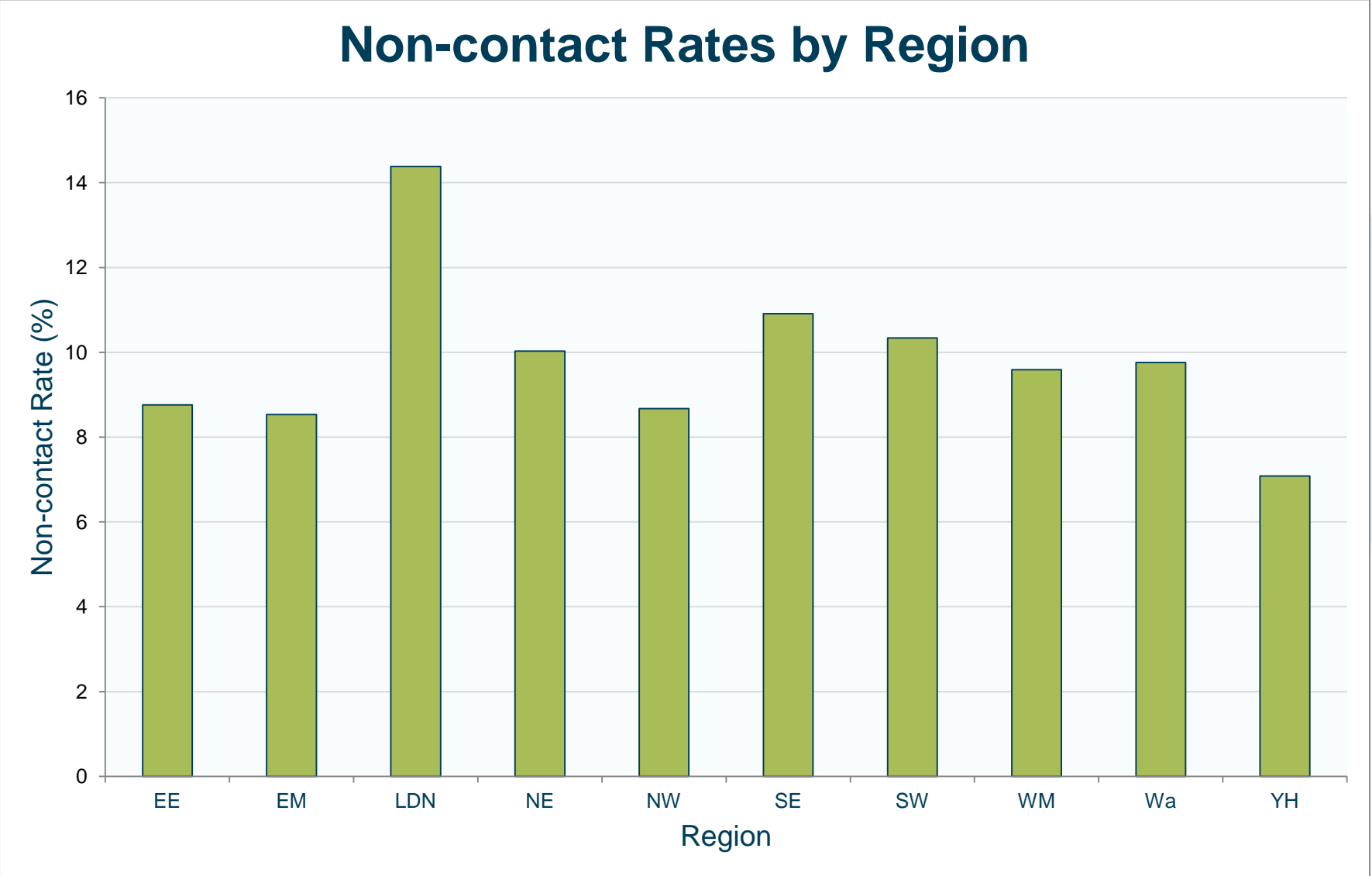
Field interviewers tasked with collecting additional data about sampled addresses.

The following variables were considered for use in the paradata analysis:

- Region (frame data)
- Calling Day
- Calling Time
- Type of Accommodation
- Call Duration
- Total Number of Attempts
- Number of Rooms

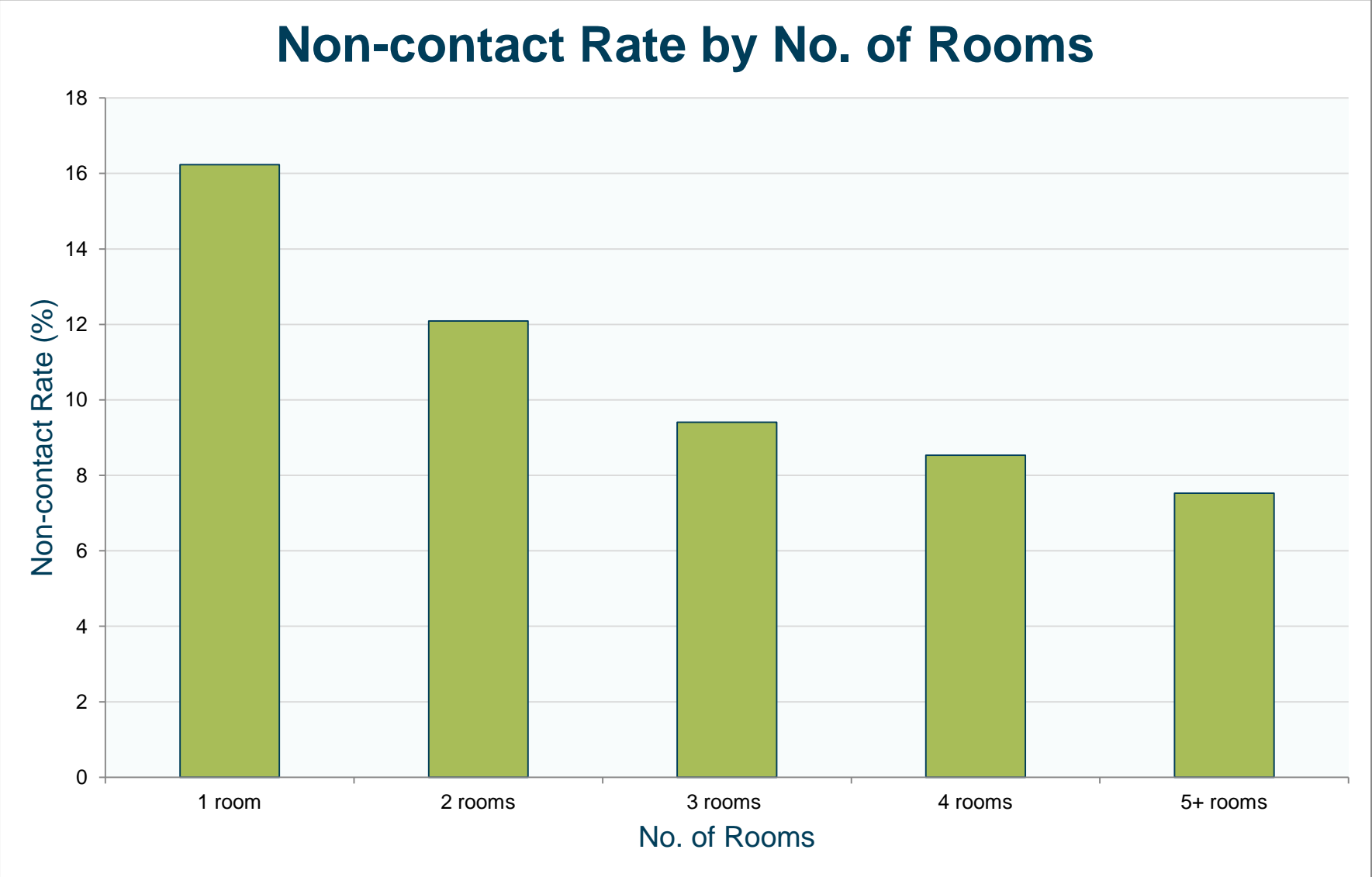
# Paradata

## Non-contact Rates by Region



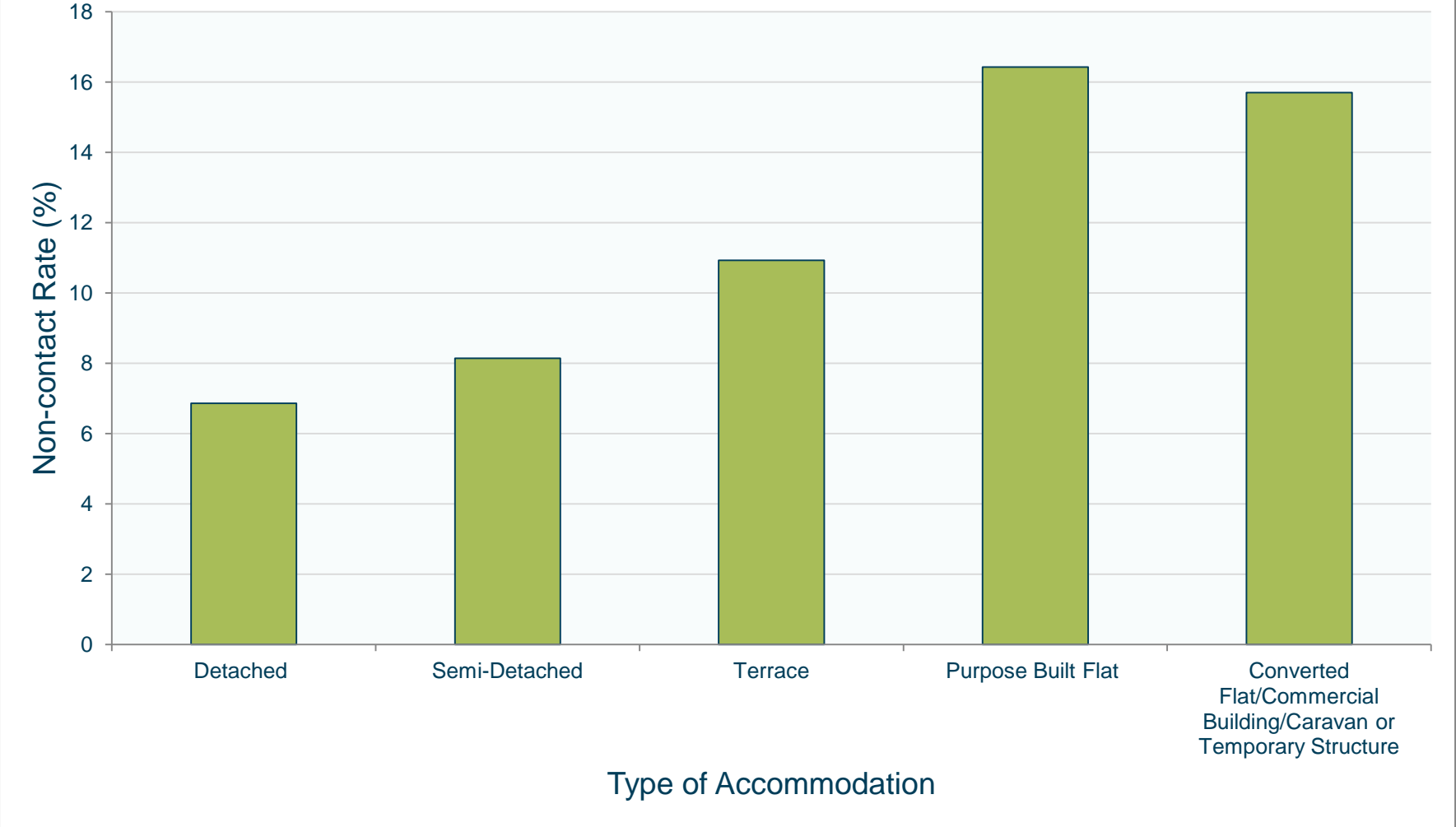
# Paradata

## Non-contact Rate by No. of Rooms



# Paradata

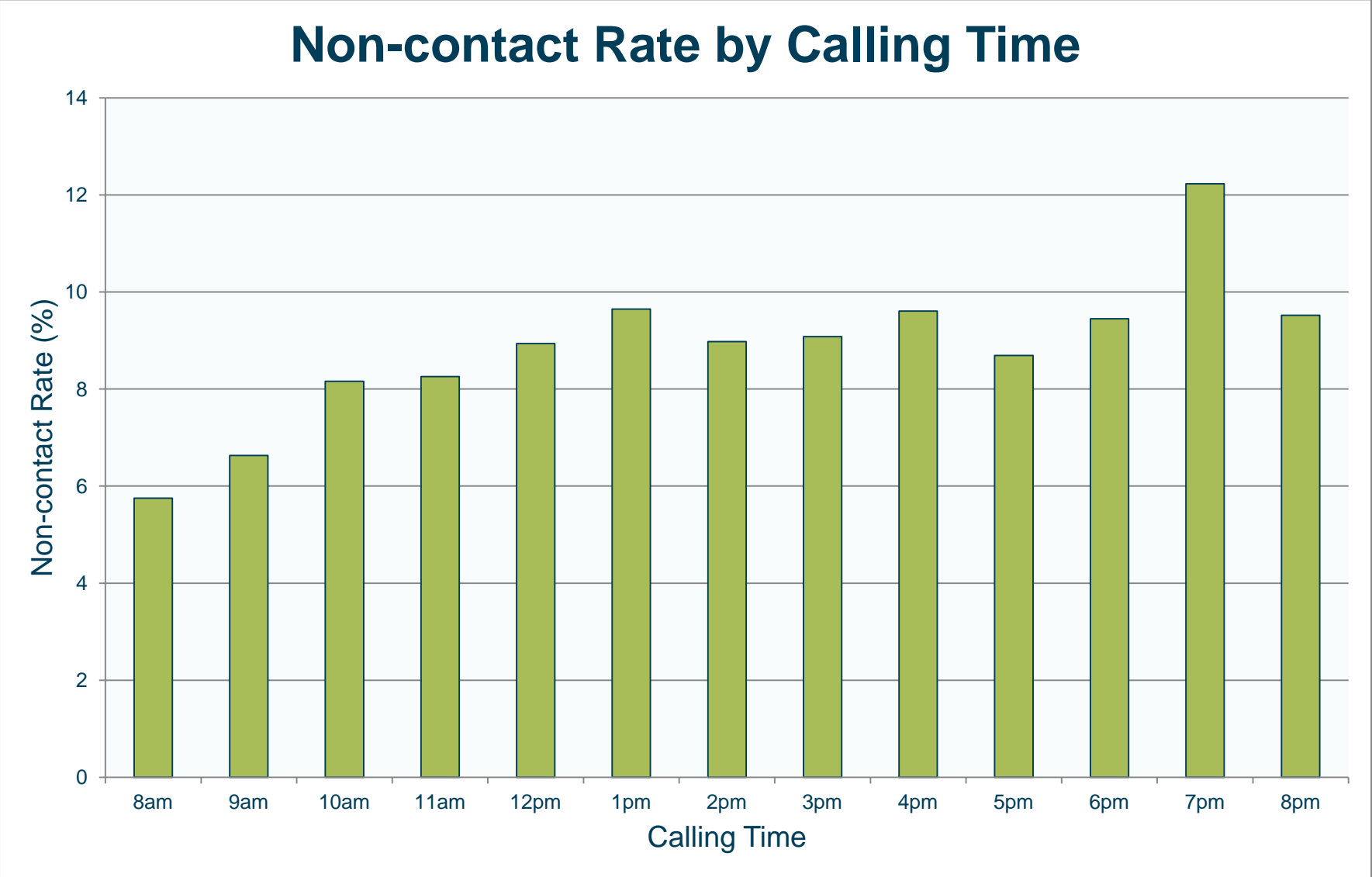
## Non-contact Rate by Type of Accommodation





# Paradata

## Non-contact Rate by Calling Time



# Multivariate Analysis

---

- Logistic regression used to consider the joint predictive power of a set of covariates on non-contact
- Binary outcome (contact/non-contact)
- Modelled using household and person characteristics
- Some classes were collapsed to simplify model
- Model fitted using stepwise regression

# Logistic Regression Findings

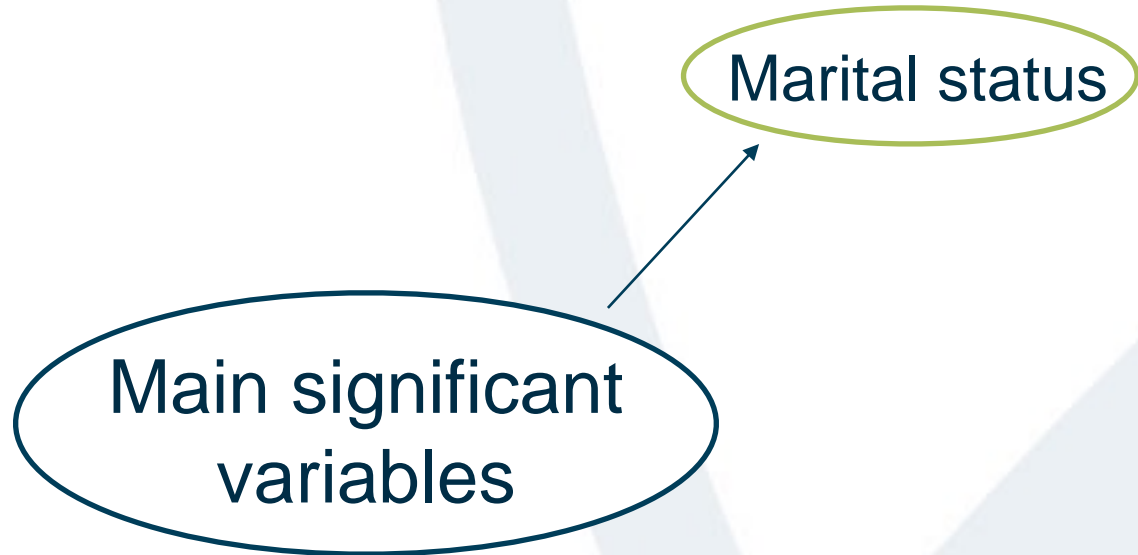
---



Main significant  
variables

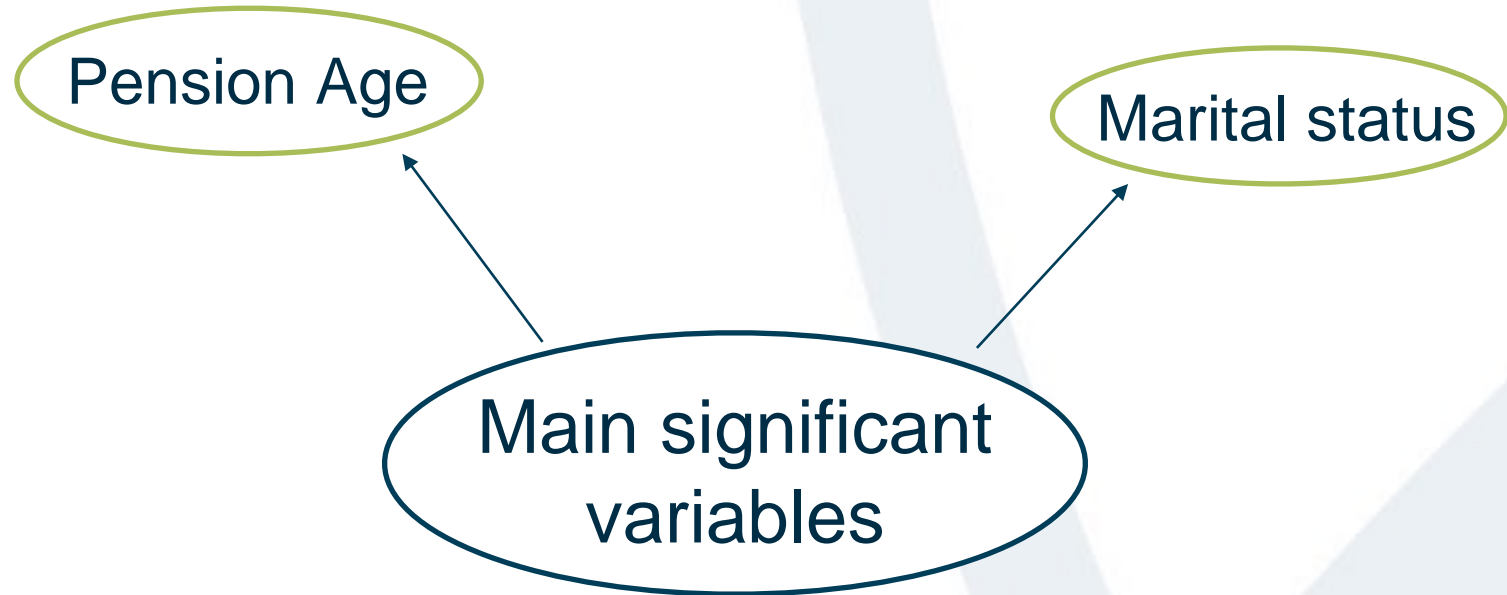
# Logistic Regression Findings

---



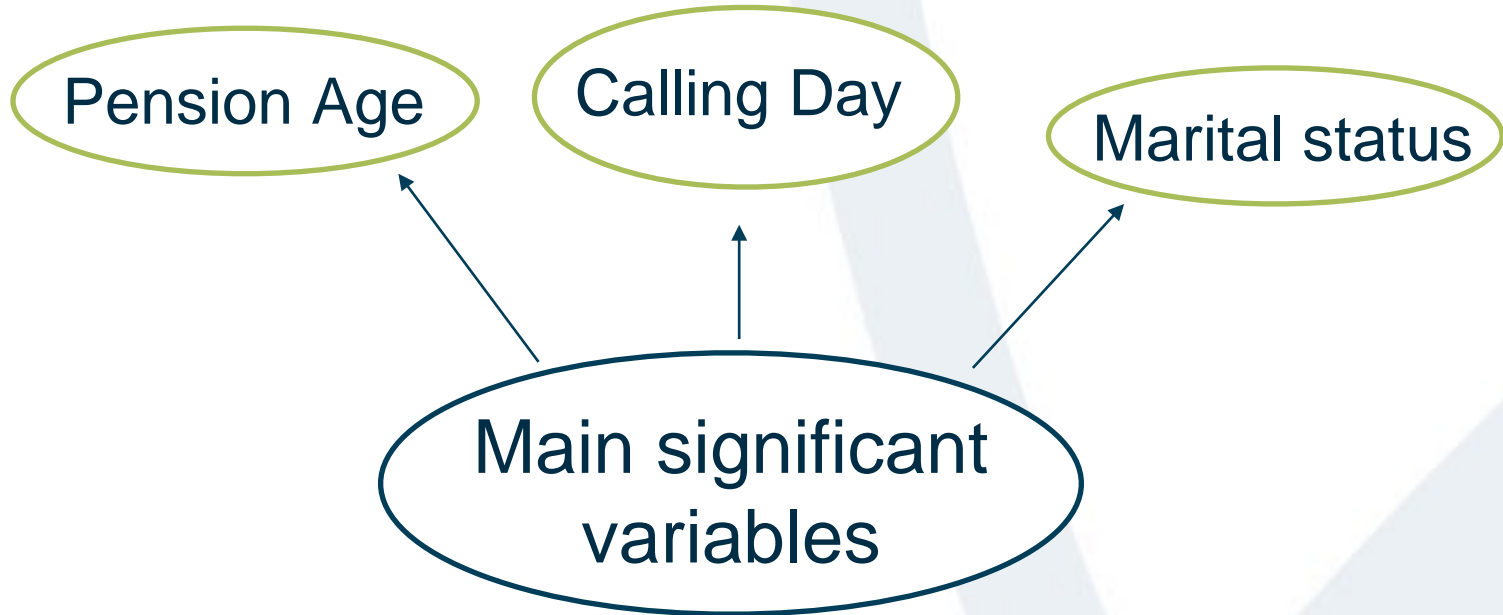
# Logistic Regression Findings

---



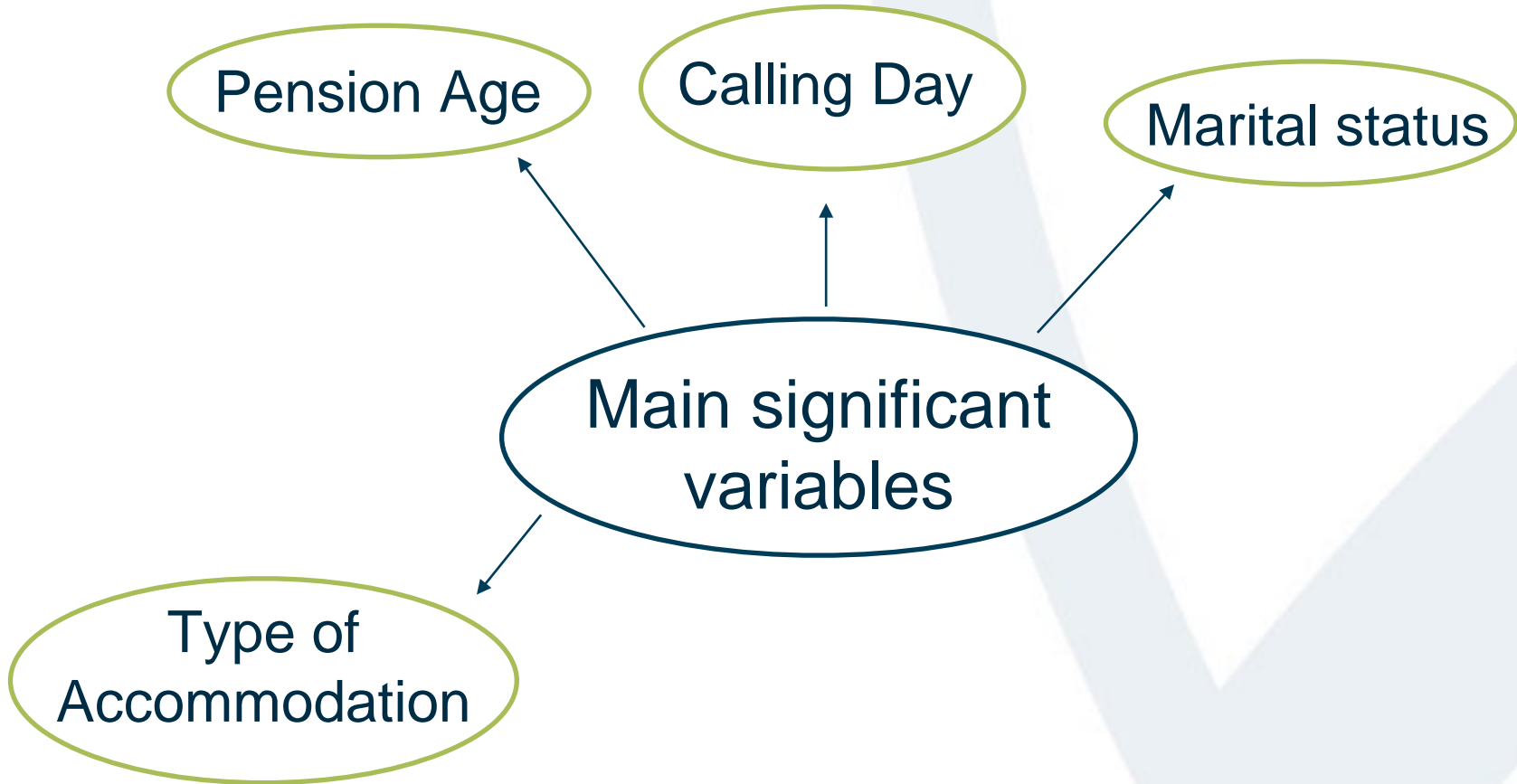
# Logistic Regression Findings

---



# Logistic Regression Findings

---



# Logistic Regression Findings

---

Pension Age

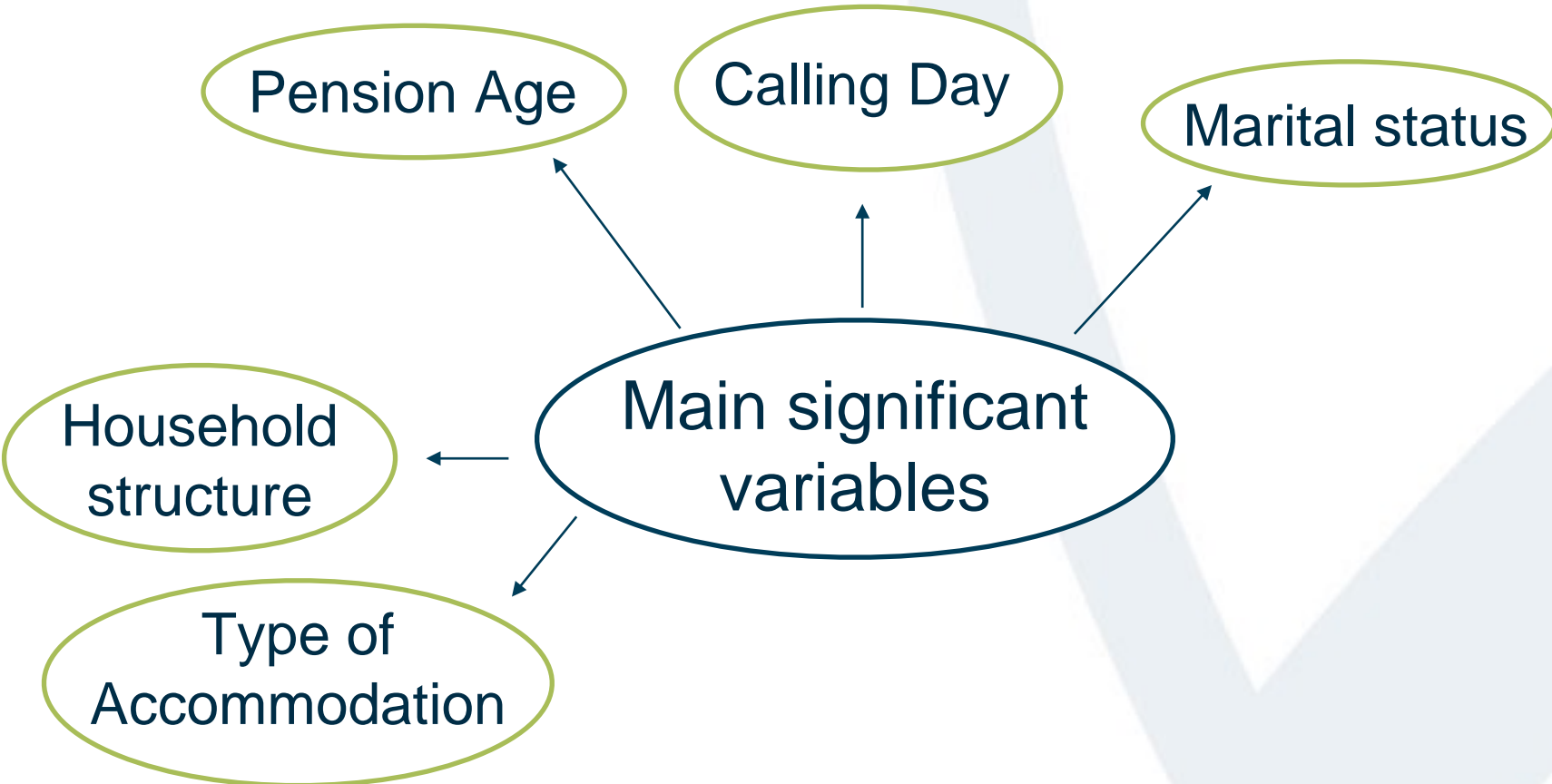
Calling Day

Marital status

Main significant variables

Household structure

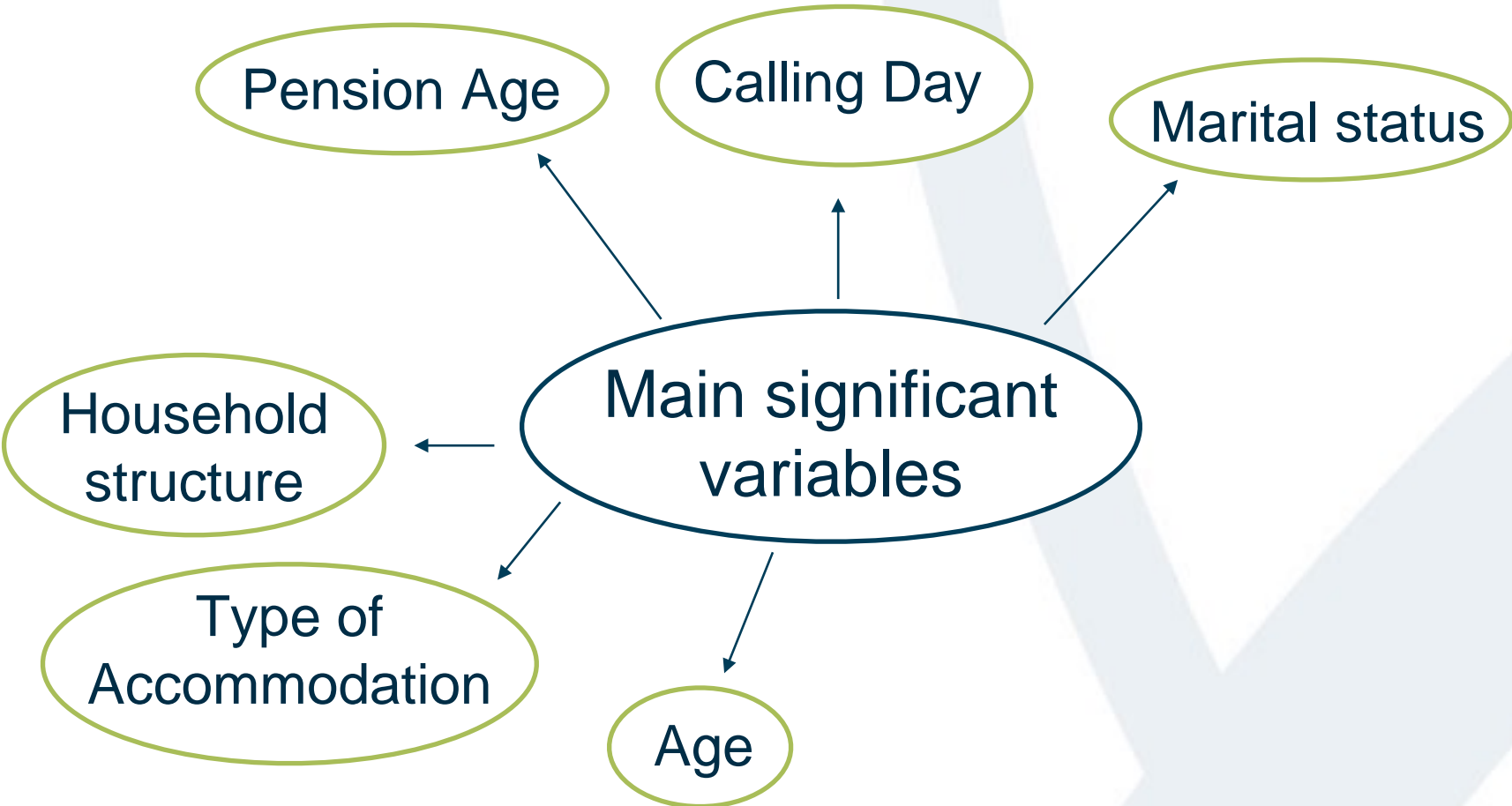
Type of Accommodation





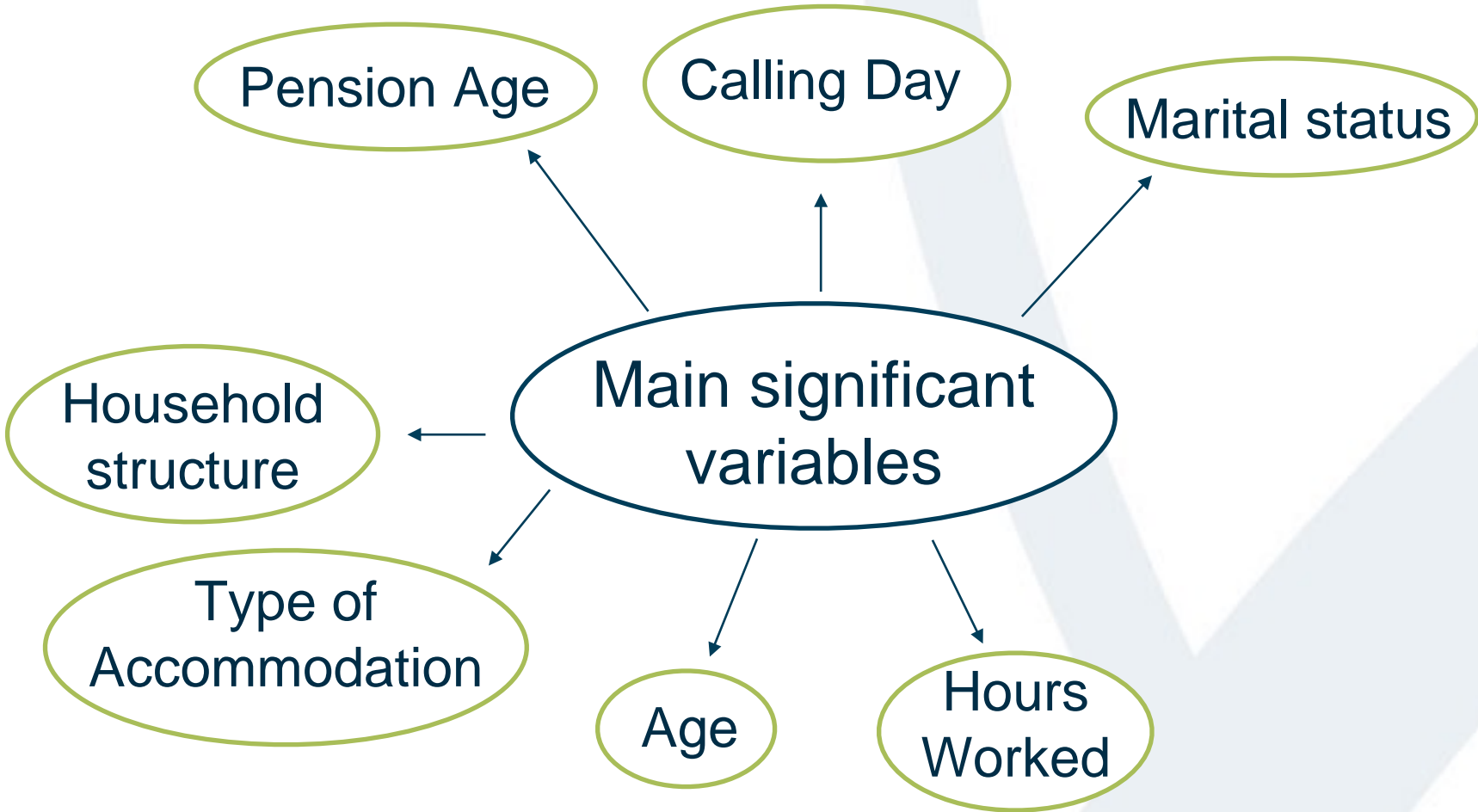
# Logistic Regression Findings

---



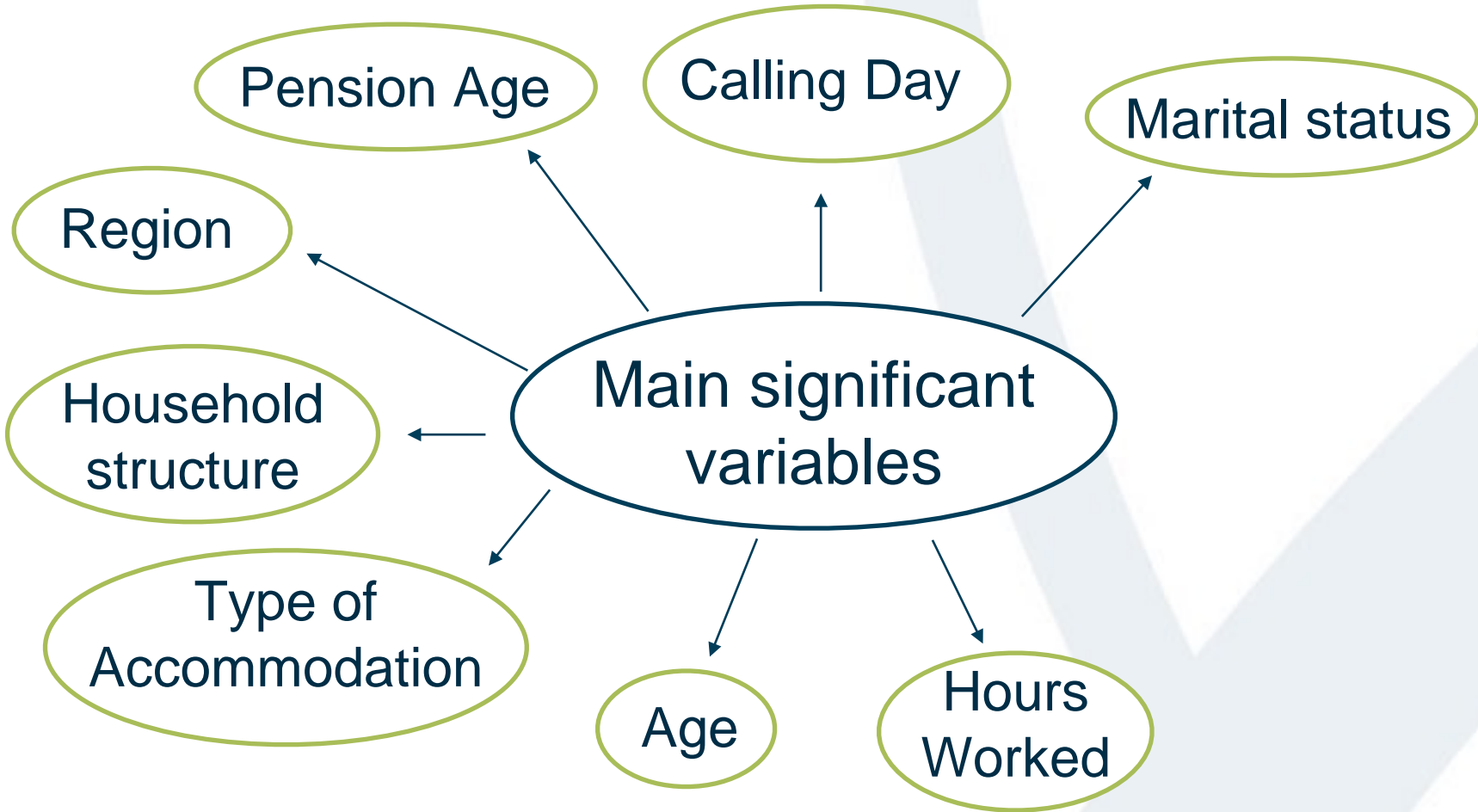
# Logistic Regression Findings

---



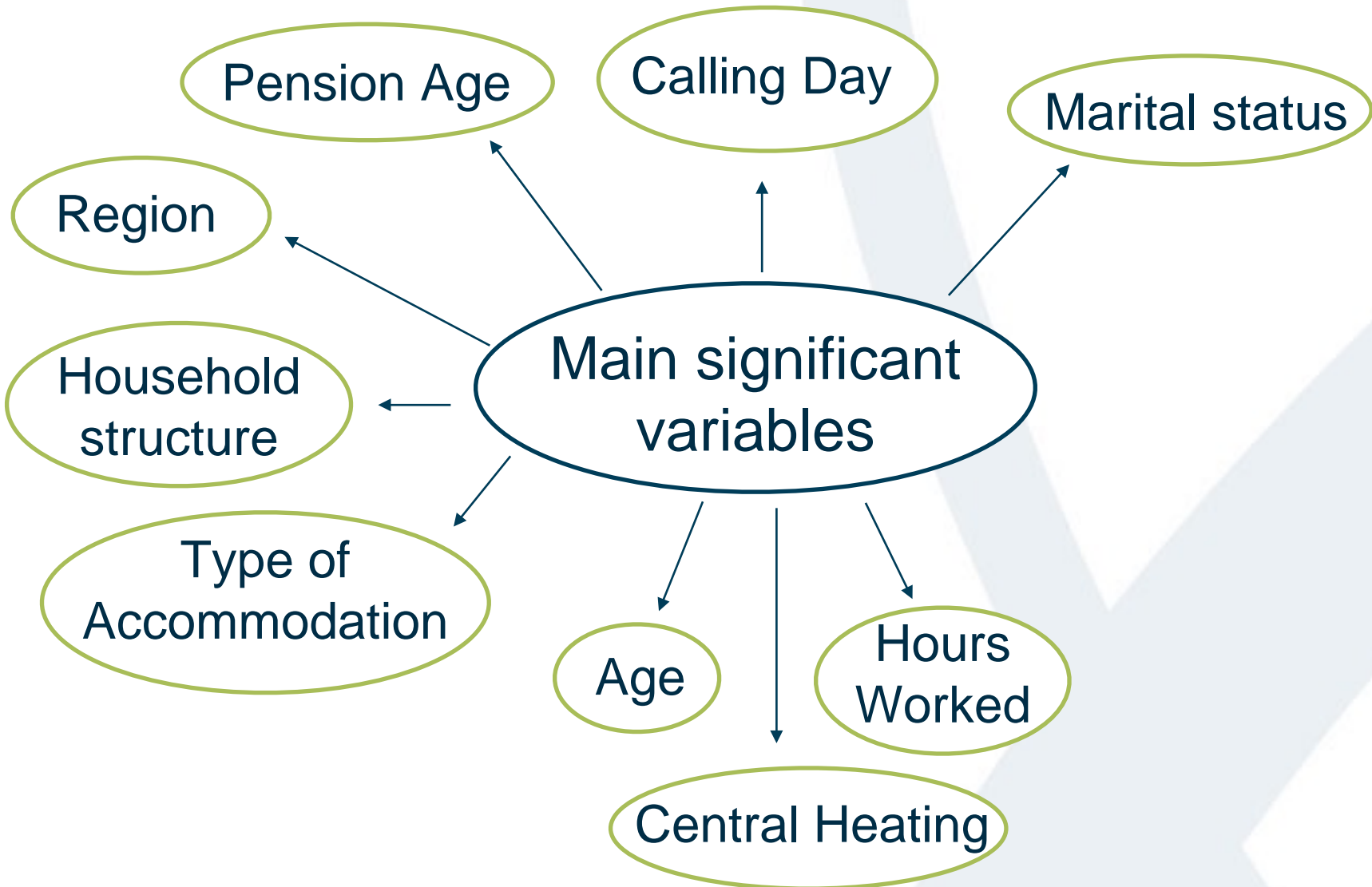
# Logistic Regression Findings

---



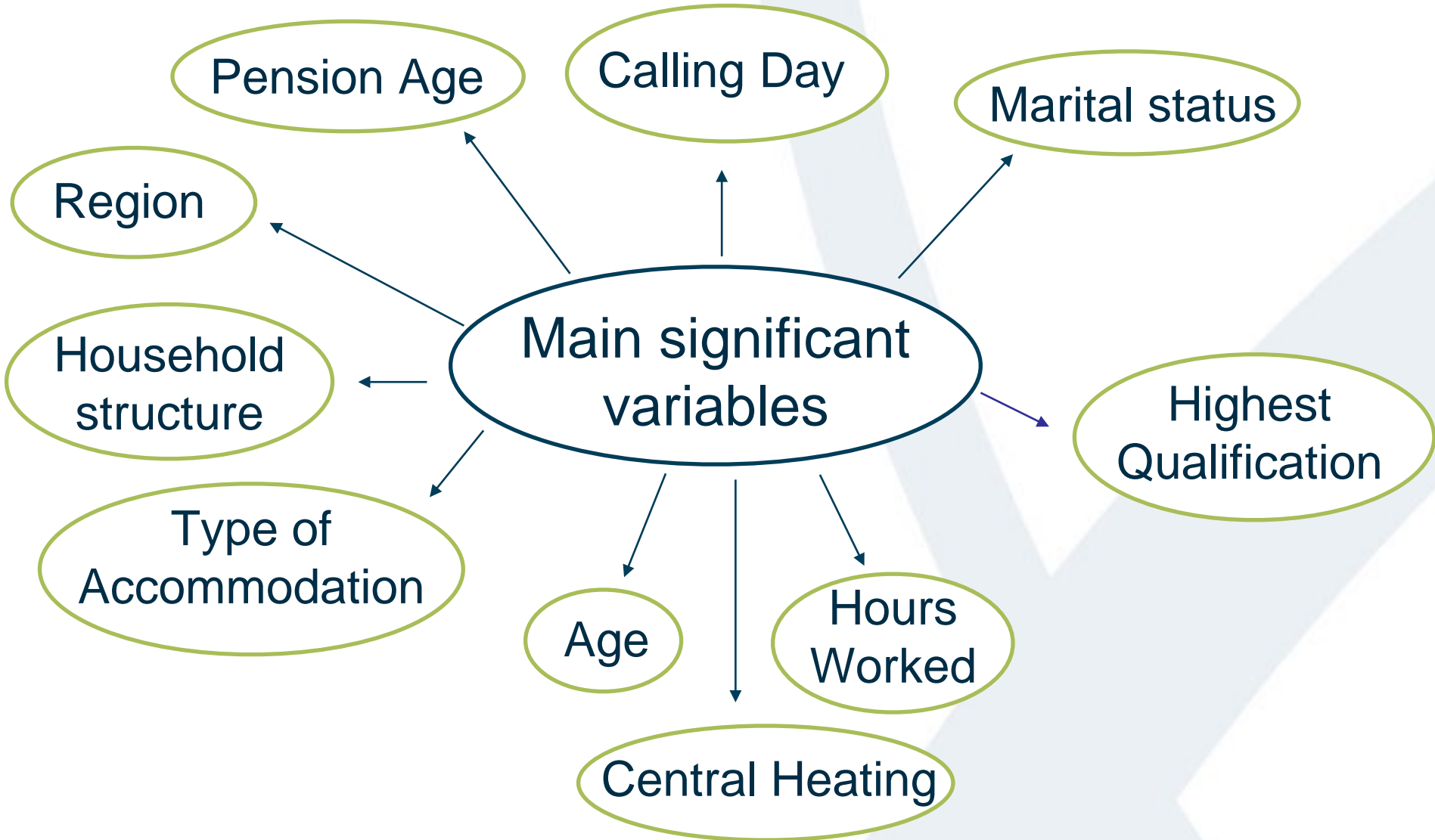
# Logistic Regression Findings

---



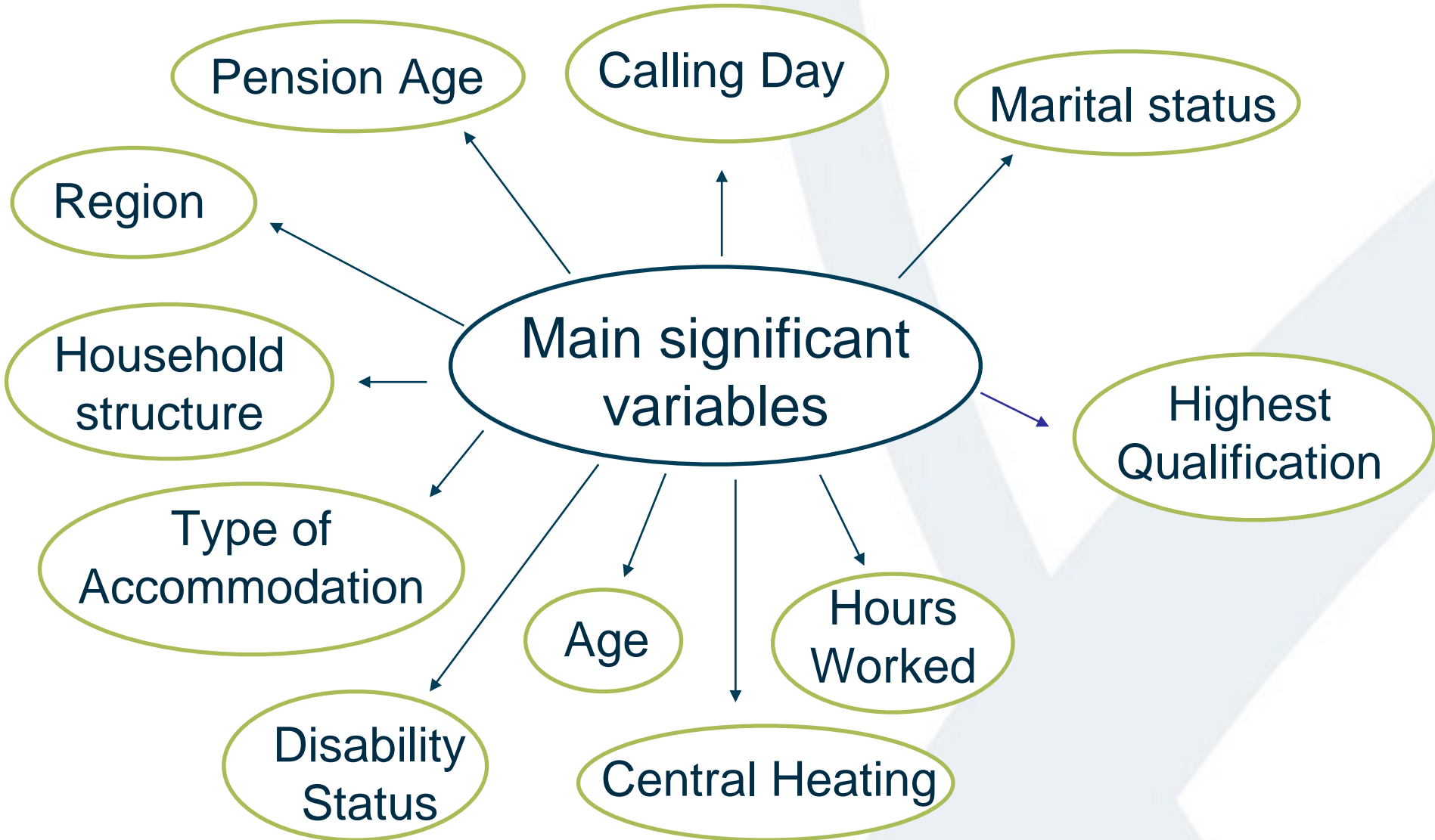
# Logistic Regression Findings

---



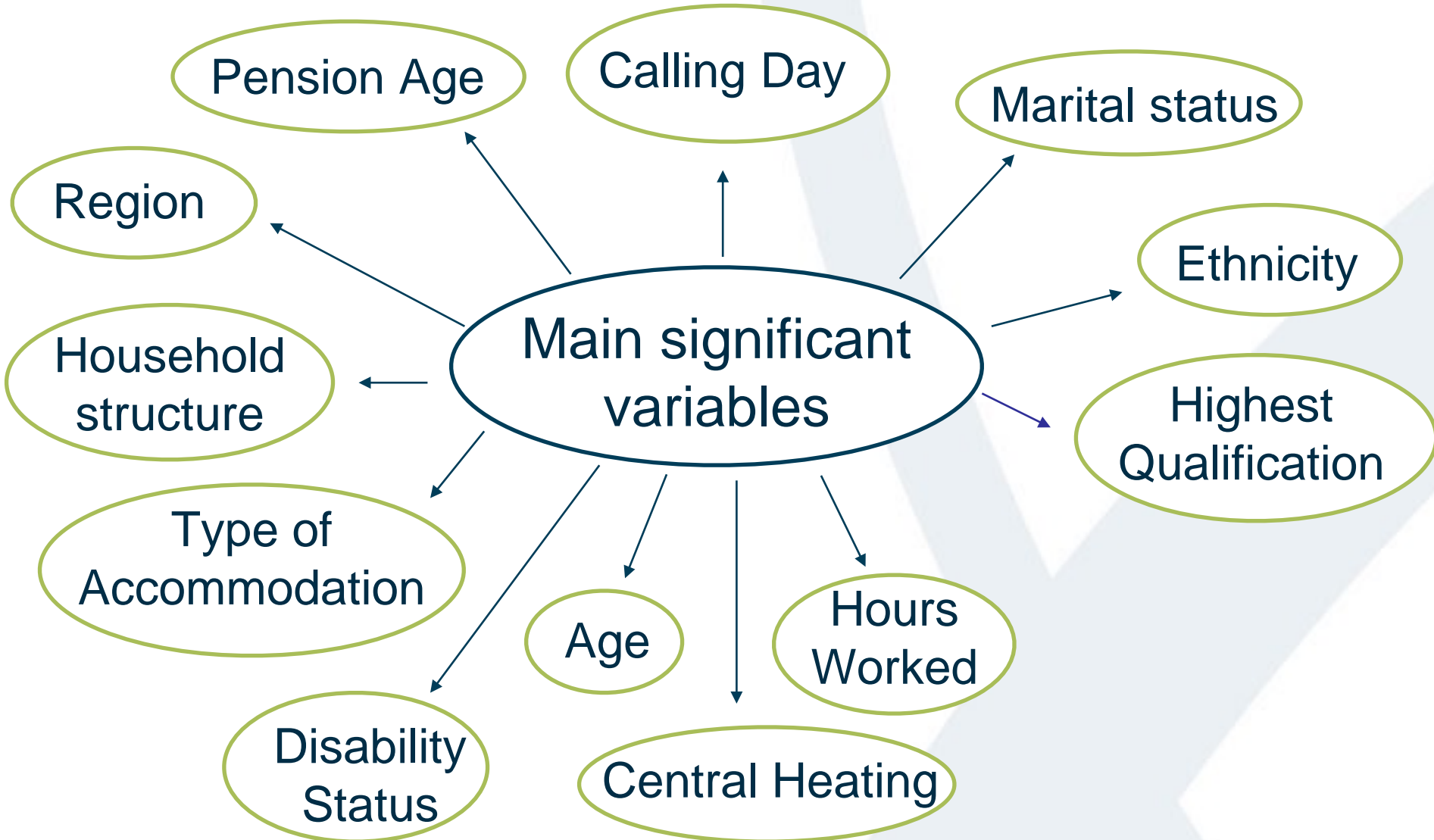
# Logistic Regression Findings

---



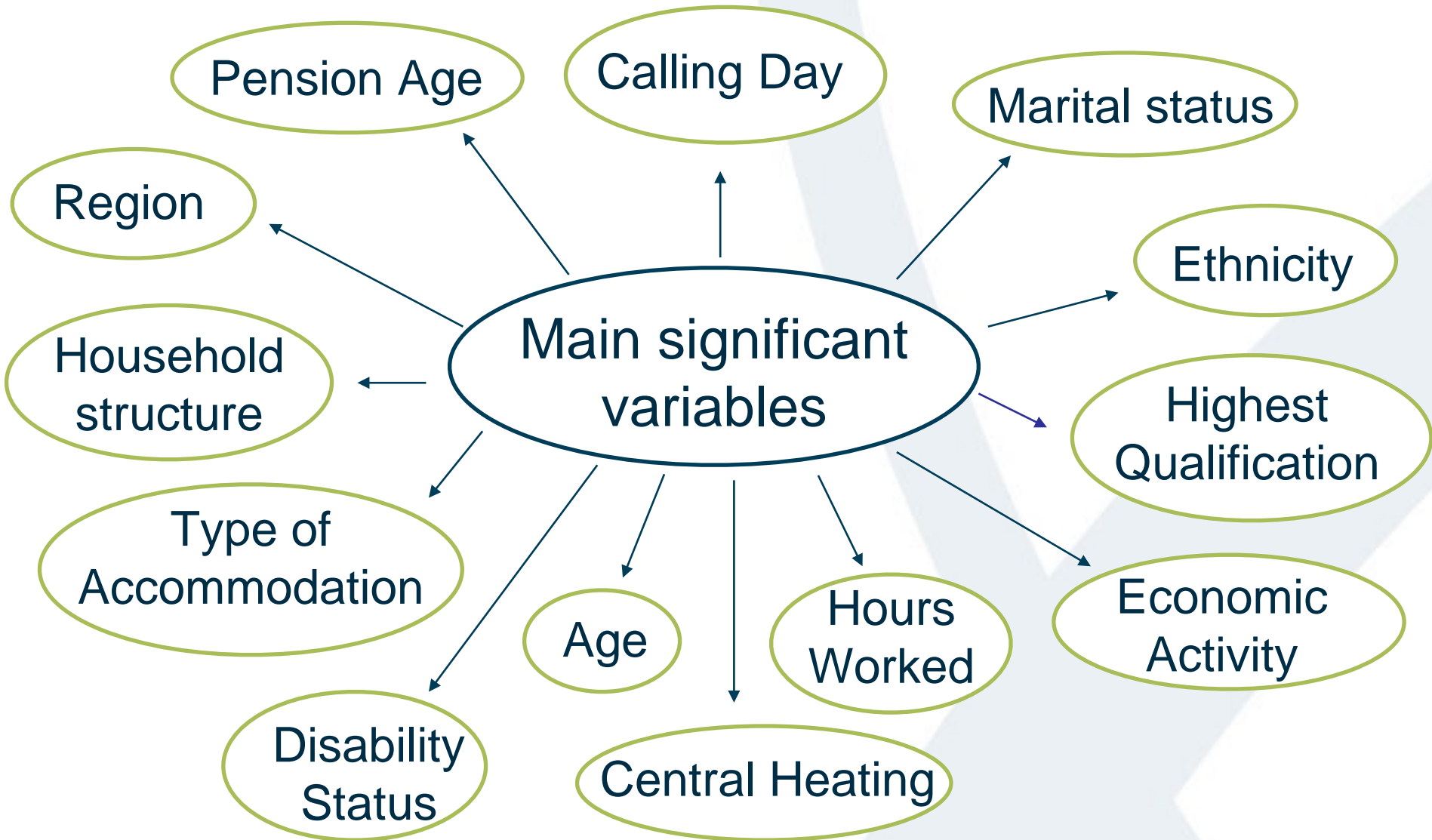
# Logistic Regression Findings

---



# Logistic Regression Findings

---



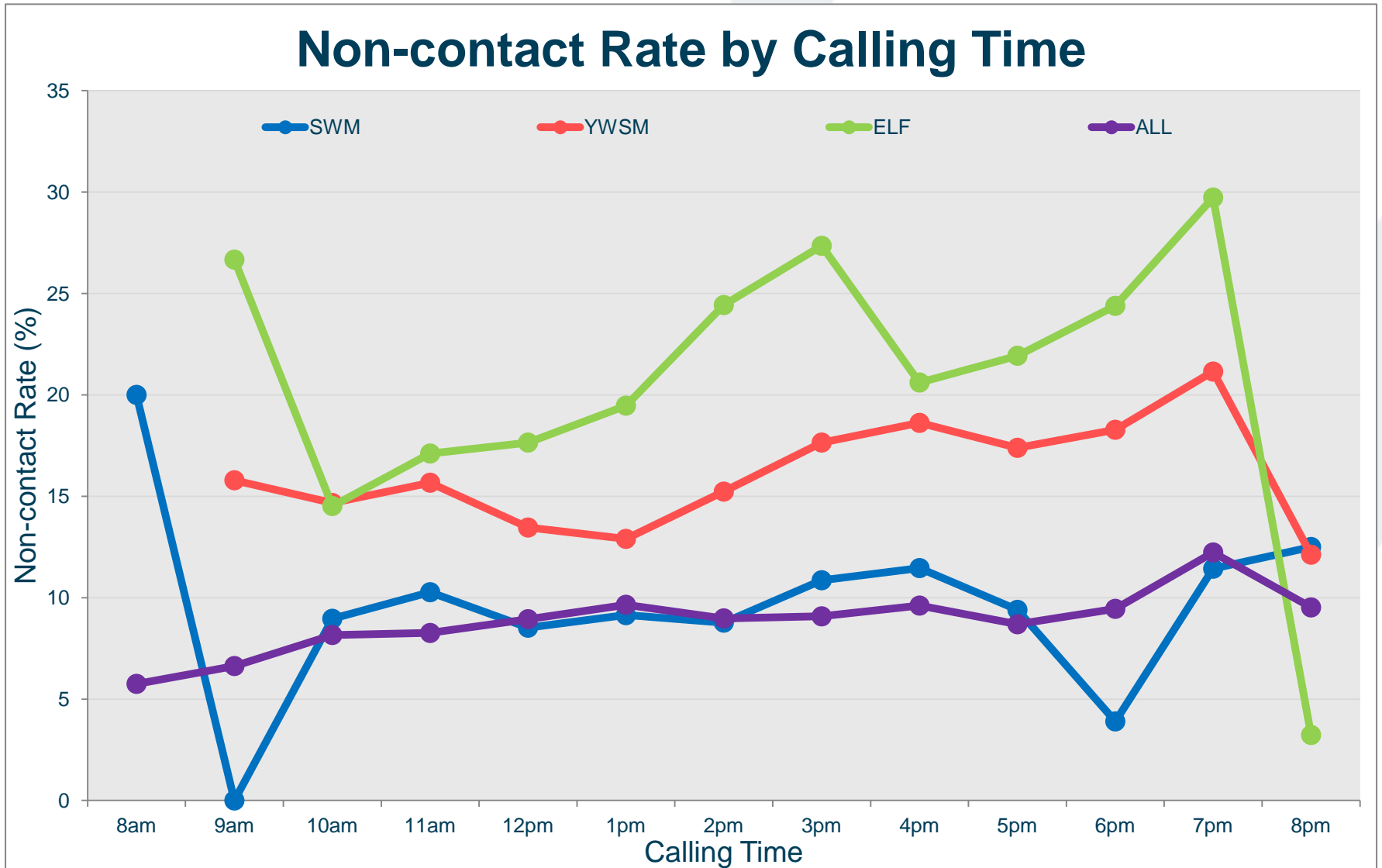


# Example of Profiles

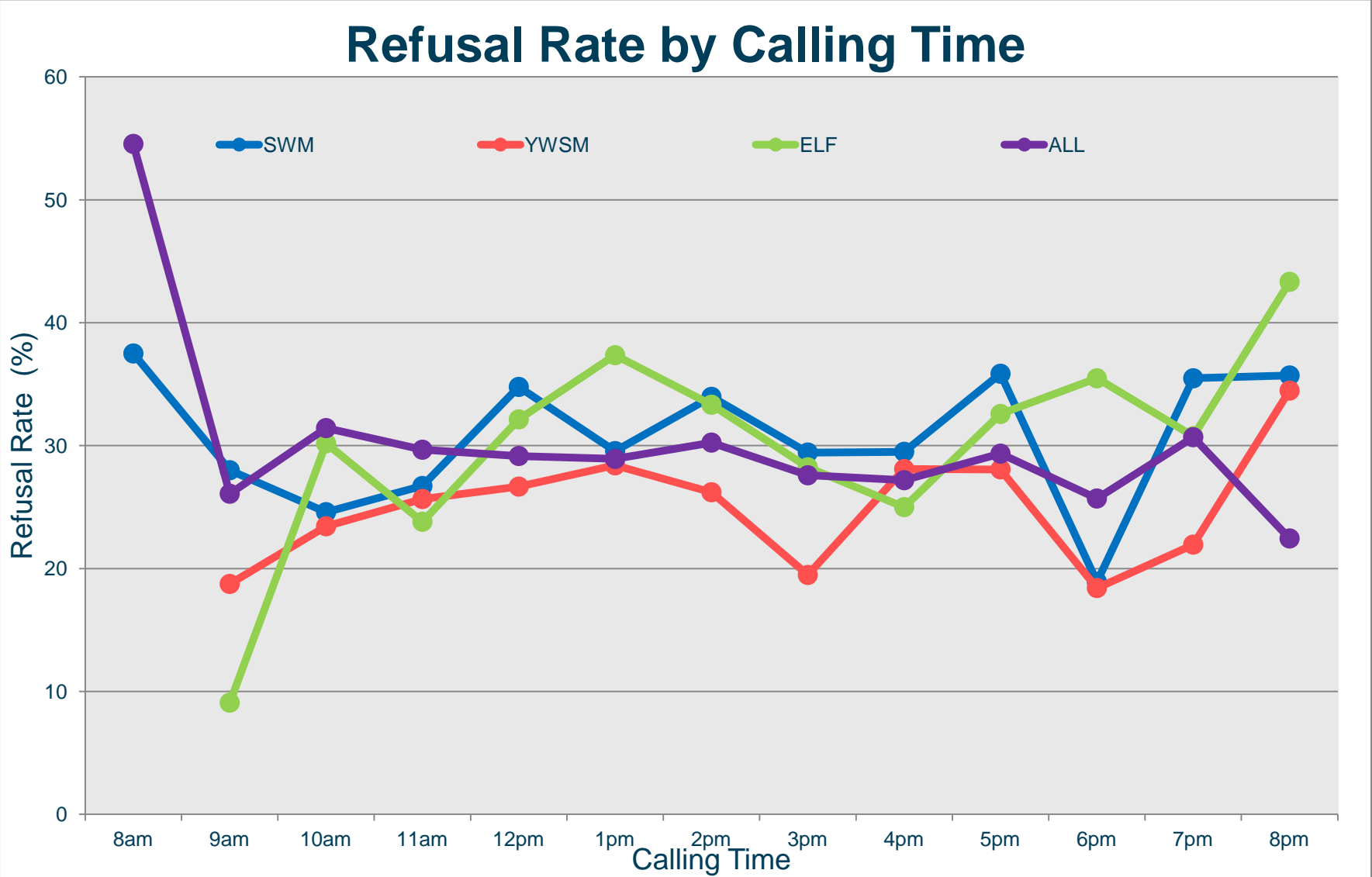
---

- Single Working Mothers (SWM)
  - Non-contact Rate = 13%
- Young Working Single Males (YWSM)
  - Non-contact Rate = 16%
- Employed in London Flats (ELT)
  - Non-contact Rate = 20%
- All
  - Non-contact Rate = 10%

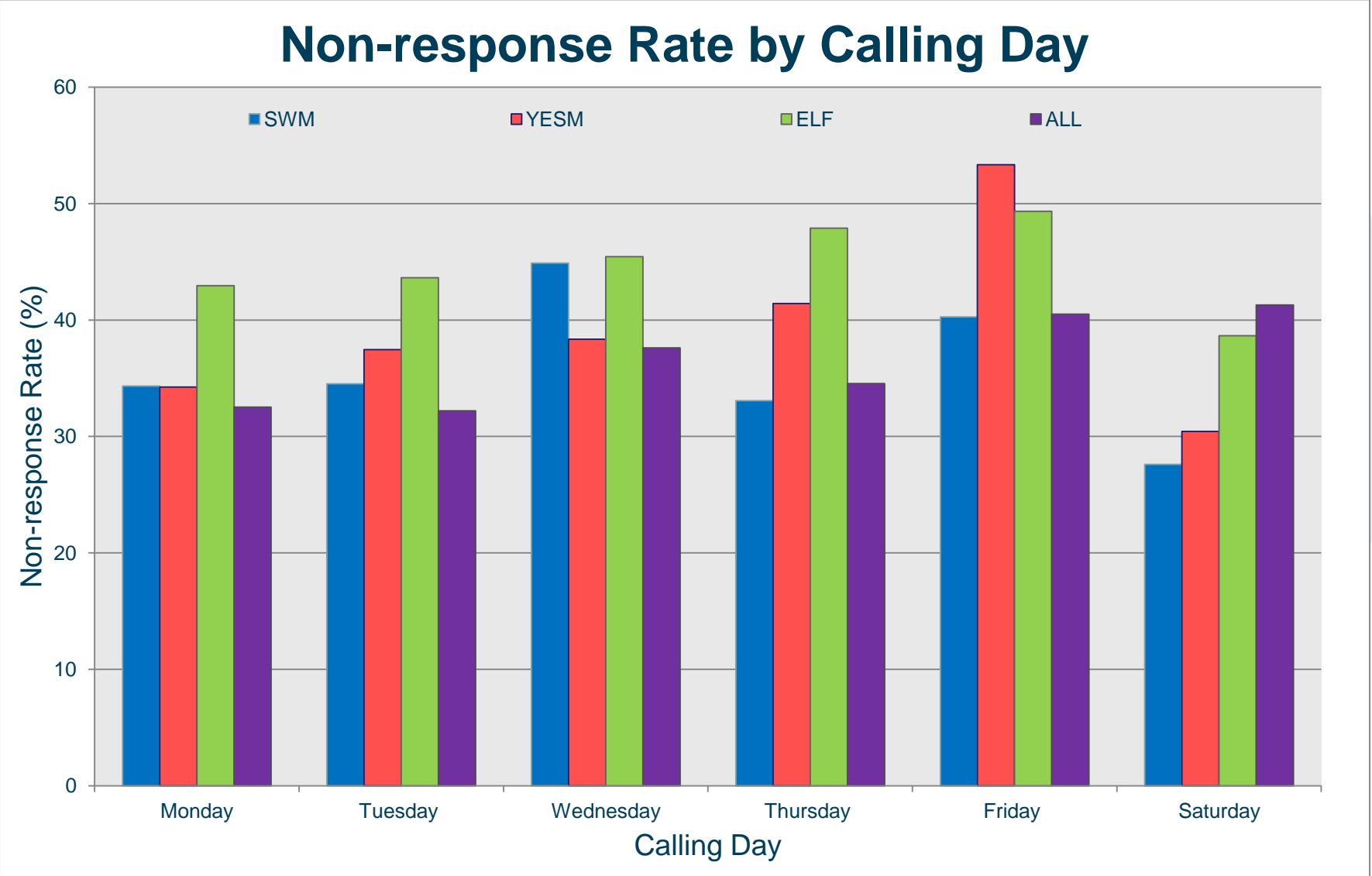
# Profiles - Calling Time



# Profiles - Calling Time



# Profiles - Calling Day



# Example of Profiles

---

- Contact rates vary when looking at different subgroups
- Higher contact rates possible when contact strategies take account of the different subgroups
- Still difficult to implement in the field! How do we target these people??
- Use of additional variables or admin data

# Conclusion

---

- CNRLS – a great but rare opportunity to assess response patterns and effectiveness of non-response weighting strategies
- Difficult to identify non-contacts from paradata alone
- Challenging to modify data collection strategy
- Need to look at other sources of information

---



**Thank you!**