

Retirement in Portugal: drivers and implications

Beatriz Alexandre Mendes Fernandes

Dissertation

Master in Modeling, Data Analysis and Decision Support Systems

Supervised by:

Patrício Costa

2026

Acknowledgments

The first person I want to thank is, without a doubt, my mother. For all the conversations that helped me get this far, for all the unconditional love and for all the hope you have placed in me, my deepest thank you.

To my whole family, for always being here to support me and for understanding the times that I couldn't be there.

To my supervisor, Professor Patrício Costa, for his time and for supporting me throughout this past year. I wouldn't be here if it weren't for you.

To all my friends, the ones who are near and the ones who are not, for all the moments we've shared over the last year, for all the collective desperation and for all the words of hope.

To those who are not here, for being my guardian angels.

And last, but certainly not least, to Tábua, O Encanto das Beiras and Baguim do Monte.

Abstract

The increase ageing of the population has intensified concerns regarding the sustainability of pension systems and the continuity of active lives across Europe. In Portugal, successive reforms in the pension system have increased both the legal retirement age and the flexibility retirement. However, individuals continue to follow heterogeneous paths out of the labour market.

This study analyses the timing of retirement, distinguishing between early, on-time, and late retirement, as well as situations where individuals remain employed, despite being legally eligible to retire. Using the legal retirement age in the year in each individual has retired as a reference, the study examines the determinants associated with these different trajectories.

Based on data from the Survey of Health, Ageing and Retirement in Europe (SHARE), Wave 9 (2021–2022), this paper uses statistical and machine learning models to explore the role of health, working conditions, financial incentives, social networks, education, and gender in each different path. The results contribute to the literature by providing evidence for the Portuguese context, emphasising the importance of analysing retirement as a process rather than a single event.

Keywords: Population Ageing, Pension Systems, Retirement Timing, Early Retirement, Late retirement, Portugal, Multiclass classification, Machine Learning models

Resumo

O envelhecimento cada vez mais acelerado da população tem intensificado as preocupações relativamente à sustentabilidade dos sistemas de pensões e à continuidade da vida ativa em toda a Europa. Em Portugal, as sucessivas reformas do sistema de pensões aumentaram tanto a idade legal de reforma como a flexibilidade na reforma. No entanto, os indivíduos continuam a seguir percursos heterogéneos em relação à saída do mercado de trabalho.

Este estudo analisa o momento da reforma, distinguindo entre reforma antecipada, na idade prevista e tardia, bem como situações em que os indivíduos permanecem empregados, apesar de serem legalmente elegíveis para se reformarem. Utilizando como referência a idade legal de reforma no ano em que cada indivíduo se reformou, o estudo examina os fatores determinantes associados a estas trajetórias diferentes.

Com base em dados do Survey of Health, Ageing and Retirement in Europe (SHARE), Wave 9 (2021–2022), este artigo utiliza modelos estatísticos e de machine learning para explorar o papel da saúde, das condições de trabalho, dos incentivos financeiros, das redes sociais, da educação e do género em cada um destes percursos. Os resultados contribuem para a literatura ao fornecerem evidências para o contexto português, salientando a importância de analisar a reforma como um processo e não como um evento isolado.

Palavras-chave: Envelhecimento Demográfico, Sistemas de Pensões, Momento da Reforma, Reforma Antecipada, Reforma Tardia, Portugal, Classificação Multiclasse, Modelos de Machine Learning

Glossary

ADL - Activities of Daily Living

AIC - Aitake Information Criteria

BIC - Bayesian Information Criteria

BMI - Body Mass Index

CART - Classification and Regression Trees

IADL - Instrumental Activities of Daily Living

ISCED - International Standard Classification of Education

MAR - Missing at Random

MCAR - Missing Completely at Random

MNAR - Missing not at Random

MICE - Multiple Imputation by Chained Equations

NRA - Normal Retirement Age

OECD - Organisation for Economic Co-operation and Development

SHAP - SHapley Additive exPlanations

SHARE - Survey of Health, Ageing and Retirement in Europe

SVM - Support Vector Machine

Contents

Abstract iii

Resumo iv

1 Introduction **1**

2 Literature Review **3**

2.1 Retirement as a Process 3

2.2 Determinants 6

2.3 Retirement in Portugal 9

2.3.1 Legal Framework 9

2.3.2 Legal Retirement Age 11

2.3.3 Early Retirement: Access Schemes, Penalties and Sustainability factor 11

2.3.4 Late Retirement: Bonuses and Incentives to Prolong Working Life . 13

2.3.5 Portugal in the European context 14

2.4 Methodological Foundations 16

2.4.1 Multiple Imputation 17

2.4.2 Multinomial Logistic Regression Model 18

2.4.3 Decision Trees 19

2.4.4 Random Forest 19

2.4.5 Support Vector Machine 20

3 Methodology **21**

3.1 Analytical Sample 21

3.1.1 Normal Retirement Age 22

3.1.2 Outcome Definition 23

3.2 Predictor selection and characterization 25

3.3 Exploratory Data analysis and Data Quality 27

3.4 Missing Data and Multiple Imputation 28

3.5 Modelling Framework 30

4 Results **35**

4.1 Outcome Distribution 35

4.2 Feature Analysis 36

4.3 Multiple Imputation Analysis 41

4.4 Multinomial Logistic Regression Results 43

4.5 Decision Tree Results 46

4.6 Random Forest 46

4.7 Support Vector Machine Results 47

4.8 Friedman test 47

4.9 Model Interpretation 49

5 Discussion **55**

5.1 Main findings 55

5.2 Literature comparison 56

5.3 Methodology Considerations 56

5.4 Study Limitations 57

5.5 Future research 58

6 Conclusions **59**

Bibliography **61**

List of Figures

1 Main legislative milestones of the Portuguese pension system 10

2 Construction of the outcome 25

3 Distribution of Retirement Timing 35

4 Box Plots - Continous Predictor Variables 40

5 Missing Values per Predictor Variable 41

6 MICE Trace plot (Network Satisfaction, Network Closeness and CASP . . . 42

7 MICE Trace plot (Max Grip Strenght, Immediate Recall, Delayed Recall . . 42

8 MICE distributions 43

9 Mean SHAP plot 51

10 SHAP distribution plot 52

11 SHAP Dependence plot. 53

List of Tables

1 Determinants of retirement timing by type of retirement and push/pull mechanisms 8

2 Fertility rate and remaining life expectancy at age 65 in Portugal and the OECD 15

3 Distribution of retirement timing categories across alternative retirement windows without hybrid classification 24

4 Operationalization of predictor variables 26

5 Imputations methods by variable 29

6 Categorical Predictors – Recodification 31

7 Multinomial Model Specification 32

8 Distribution of categorical predictors by retirement timing 37

9 Descriptive statistics for continuous variables 39

10 Model fit statistics for the multinomial models 43

11 Significant and borderline associations from the multinomial logistic regression 44

12 Predictive performance of the Decision Tree 46

13 Predictive performance of the Random Forest 46

14 Predictive performance of the Support Vector Machine 47

15 Comparison of machine learning pipelines 48

16 Stable variables by retirement timing category 50

Chapter 1

Introduction

Like in the rest of the European Union, the Portuguese population is ageing rapidly. This phenomenon is related to the combination of increased average life expectancy and declining fertility rates, which is increasing pressure on Portuguese pension systems. As a result, most European countries have been promoting alterations to their social security systems, increasing the legal retirement age, and creating incentives for extending active life, and postponing the exit from the labour market.

However, increasing the legal retirement age does not translate into a proportional increase in the effective age at which people leave the labour market. Individuals can follow different paths when they leave the workforce, and these highlights that retirement does not correspond to the moment when an individual reaches the legal age, but rather to a decision that can be postponed over time.

In this way, retirement is no longer a homogeneous and predictable event, but rather a process that can take multiple forms and occur at different times throughout the individual's life. Some individuals leave the workforce before the legal retirement age, while others remain working until that age or extend their presence in the labour force. It is important to also highlight the individuals who, despite having already met the legal conditions to retire, choose to continue working.

Despite the extensive literature on the retirement matter, there is not a consensual definition on early, on-time, and late retirement, being these categories usually operationalised depending on the objectives of the study. In addition, most studies tend to use binary approaches that distinguish between retired and non-retired individuals, which limits the analysis of the different paths of retirement.

These different behaviours show that retirement decisions do not depend only on legal eligibility, but also on a combination of incentives, constraints and individual preferences, which translate into different retirement trajectories. In this context, retirement decision is influenced by a combination of factors, being them individuals, work-related, economic, family and institutional. So, it becomes relevant to analyse the moment of retirement, considering not only individuals who retire before or after the legal age but also those who, despite being eligible for retirement, remain in the labour market.

Understanding the determinants that influence the timing of retirement is relevant from a social and economic perspective, as these decisions have direct implications for the financial sustainability of pension systems, job opportunities for older workers, and individual well-being. These choices were influenced by incentives and institutional rules, making it necessary to analyse them in the specific contexts in which they occur.

This project focuses on analysing the timing of retirement, distinguishing between the different transitions from the labour market, that is, early retirement, late retirement, on-time retirement, and the situations in which the individual chooses to work, despite being legally eligible to retire.

Using the legal retirement age as a reference, this study compares the determinants that lead to these different exit trajectories, highlighting similarities and differences between the paths. By doing so, it aims to contribute with empirical evidence to the Portuguese context.

The present study adopts an empirical approach, based on the analysis of data from the Survey of Health, Ageing and Retirement in Europe (SHARE). It will use information about Portugal from the 9th Wave of the survey, collected during the period 2021-2022. This allows for the observation of individual's labour status, the timing of the transition to retirement, and various demographic, socioeconomic, and health characteristics relevant to the analysis.

Multiple imputation is applied to address missing values, while different strategies are considered to mitigate the class imbalance between the four classes. By combining multinomial logistic regression with machine learning models, this study provides a complementary evaluation of model predictive performance and the determinants associated with different retirement paths.

Chapter 2

Literature Review

2.1 Retirement as a Process

Retirement generally refers to the withdrawal from the paid labour market, occurring, in most cases, among individuals who are close to the legal retirement age and who typically gain access to pensions or other social protection benefits (Dorn and Sousa-Poza (2010)).

Recent literature has emphasised that retirement should not be viewed as a single event, but rather as a process. The transition from work to retirement can occur gradually and take on multiple forms in the labour market, rather than a direct transition to retirement. Therefore, retirement trajectories may include partial retirement, temporary interruptions in professional activity, or continued work beyond the normal retirement age (Wang and Shultz (2010)).

The literature also shows that the definition of retirement varies according to the criteria used, and may be based on labour market status, receipt of pensions or self-identification as retired ((Dorn & Sousa-Poza, 2010)). These differences mainly reflect methodological choices and specific institutional contexts.

This study focuses on the timing of retirement, analysing the different types of exit from the labour force (early, on-time and late retirement), when compared with the statutory retirement age in Portugal.

Although not officially classified as retirees, this study also considers individuals who, despite having reached the legal retirement age and being eligible for retirement, remain active in the labour market. As these individuals have not yet made the transition out of paid employment, they cannot be classified as late retirees and were studied as a separate retirement group.

Early Retirement

There is no universally accepted definition of early retirement in the literature.

Some studies define early retirement as leaving the labour market before the legal retirement age (De Wind et al. (2014); Litwin and Tur-Sinai (2015); Reeuwijk et al. (2013)), while others define early retirement as leaving the workforce due to pensions and social protection benefits associated with leaving the labour market before the legal retirement age (Axelrad (2018); OECD (2025)). At the same time, some literature argues that early retirement is not based exclusively on objective criteria, such as age or eligibility, but also on subjective dimensions, such as psychological detachment from work and self-identification as a retiree (Topa, Depolo, and Alcover (2018)).

In this study, individuals are classified as early retirees if they were offer the option of early retirement or if they retired at least 6 months before the legal retirement age. This definition allows for a clear operationalisation of different retirement trajectories based on the moment of transition out of working life (Litwin and Tur-Sinai (2015)). Furthermore, this approach is supported by the literature, having already been used in other significant empirical studies that analysed the exit moment from the workforce in the European context (Axelrad (2018)).

Depending on the circumstances in which the exit from the labour market occurred, the literature distinguishes between different forms of early retirement. Specifically, Dorn and Sousa-Poza (2010) distinguishes between early exit from the labour market as voluntary, if the individual chooses to leave the labour market despite being able to continue working, and involuntary, if the exit results from unexpected circumstances that prevent the individual from continuing their professional activity. Although there is a decision to opt for involuntary early retirement, this choice is influenced by external factors, such as health problems or difficulty in finding employment, that makes it impossible to remain active in the labour market.

The literature has relied on the distinction between push and pull factors to explain early retirement decisions (Casas and Román (2024); Ebbinghaus and Hofäcker (2013); Gerke (2013b); Reeuwijk et al. (2013)). Push factors correspond to negative circumstances that drive individuals towards early retirement. These include, for example, poor health and unfavourable work conditions. Pull factors, in contrast, refer to aspects that make early retirement more attractive to workers. These include, for example, financial incentives from pension schemes, family responsibilities and personal motivations. It is important to mention that these mechanisms often operate simultaneously, as individuals may face both constraints that limit continued employment and incentives that encourage early exit from the labour market (Casas and Román (2024)).

On-Time Retirement

In the literature, on-time retirement corresponds to leaving the labour market when an individual reaches the legally defined retirement age, if there is one, or when they become eligible to receive full retirement pensions, whether they are public, private, or occupational (Banks (2006)). This path represents the standard transition from work to retirement, in contrast with early and late retirement.

Although on-time retirement is not always explicitly defined in the literature, it is frequently used as a reference category when analysing other retirement trajectories, since the legal retirement age or the eligibility for full pension benefits serve as the main temporal criterion for distinguishing between different types of retirement (Banks (2006); Wang and Shultz (2010)).

In this study, on-time retirement is define as the exit from the labour market between 6 months before and 6 months after the legal retirement age.

Late Retirement

Some workers continue their professional activity beyond the legal retirement age or after reaching the age of eligibility to receive full pension benefits. The literature refers to these cases as late retirement, and this trajectory represents a delay in exiting the workforce in relation to the normal transition between work and retirement (Axelrad (2018); Banks (2006)).

So, in this study, late retirement is define as the exit from the labour market after, at least, 6 months from the legal retirement age.

Working beyond Legal Retirement Age

As previously mentioned, this work also studies individuals who, despite already meeting the legal conditions to access the old-age pension, remain in the labour market and have not yet transitioned to retirement.

Despite not being officially categorised as retirees, these people are conceptually associated with late retirement trajectories due to their continuing employment after the legal retirement age. The inclusion of these cases allows for a more comprehensive understanding of the behaviours related to delaying exit from the labour market, contributing to a more complete analysis of the decisions associated with extending active life.

2.2 Determinants

The literature highlights that the timing of retirement results from a combination of individual, employment-related and socio-economic factors (De Preter, Van Looy, and Mortelmans (2013); De Wind et al. (2014); Reeuwijk et al. (2013); Scharn et al. (2018)). Comparative analyses suggests that the decision to retire, whether early, on-time or late, is strongly shaped by the rules and incentives of pension systems and the macroeconomic context, which generates systematic differences between countries (Axelrad (2018)). In this sense, early, on-time and late retirement are not just individual choices, but also responses to opportunities and constraints created by external factors, with different effects according to the social groups.

Health, functional limitations and wellbeing

The literature recognises health as one of the key determinants of early retirement. Evidence shows that health problems, functional limitations, and physical and emotional exhaustion act primarily as push factors, with a tendency to reduce the ability to work and increase the likelihood of early exit from the workforce (Börsch-Supan et al. (2013); De Wind et al. (2014); Koçak and Burdorf (2025); Litwin and Tur-Sinai (2015)). On the other hand, better health is associated with a higher probability of delaying the transition to retirement (Axelrad (2018)). When analysing empirical studies in a national context, data from Denmark indicate that better physical capacity, measured through indicators such as grip strength, is associated with lower probability of early retirement (Gerke (2013b)). In the case of Bulgaria, perceived health stands out as a key determinant of early retirement decisions, emphasising gender differences, with poor health being a more decisive factor for men than for women (Markova (2020)).

Working conditions and job quality

Working conditions appear to be push factors for early retirement. Workers exposed to high physical demands, organisational pressures, labour conflicts or lack of support in the workplace tend to retire earlier (Gerke (2013b); Reeuwijk et al. (2013)). Evidence from a qualitative study shows that negative working conditions reduce motivation and/or ability to remain in the labour market (Reeuwijk et al. (2013)).

In contrast, jobs that are less physically demanding, offer greater autonomy, have lower levels of stress and greater flexibility tend to make it easier to remain in the labour market at an old age (Wang and Shultz (2010)). Following this logic, the literature also suggests that people working in the primary or secondary sectors are more likely to retire early, than those working in the tertiary sector (Casas and Román (2024)).

Financial Incentives and "opportunity" to leave VS need to stay

Financial circumstances can influence the decision to retire in two ways.

The existence of financial possibilities or favourable retirement schemes emerges as a decisive element for the choice of early retirement (De Preter et al. (2013); De Wind et al. (2014); Litwin and Tur-Sinai (2015); Reeuwijk et al. (2013)).

On the contrary, the decision to extend one's working life is associate to financial difficulties or less favourable retirement benefits (Axelrad (2018); De Wind, Geuskens, Ybema, Bongers, and Van Der Beek (2015); Markova (2020); Scharn et al. (2018)).

Social Network and Family

The social network and family context appears in the literature as one of the most relevant determinants of early retirement, not only in terms of support, but also through the intensity of contact and the composition of the network. Empirical evidence shows that frequently contact with the principal confident and a larger social network is positively related to the probability of early exit from the workforce Litwin and Tur-Sinai (2015).

The role of the spouse is also relevant. Being married is associated with a more positive attitude towards early retirement (Axelrad (2018); De Preter et al. (2013); De Wind et al. (2015); Litwin and Tur-Sinai (2015)). This probability is also affected by the spouse's labour status, suggesting that the decision to retire is not only an individual one, but rather made within the family context (Litwin and Tur-Sinai (2015)). The literature shows that this effect is more significant among women, whose likelihood of retirement increases if their spouse is not working (Wang and Shultz (2010)). Additionally, the presence of grandchildren is associated with greater propensity for early retirement (Litwin and Tur-Sinai (2015)).

The available literature is limited when discussing the role of social network and family in late retirement. These factors are especially relevant to the decision to retire early, and not to remaining in the labour market at older ages.

Education and qualifications

The level of education emerges as a central factor in the literature when discussing retirement age (Wang and Shultz (2010)). Low levels of education tend to be associated with a higher probability of early retirement, while higher levels of education are related to longer periods in the workforce (Axelrad (2018)). Following this logic, higher-skilled workers less likely to leave the labour market at an early age (Casas and Román (2024)).

Gender

Gender appears in the literature as a relevant factor in defining the timing of retirement, reflecting differentiated professional and family trajectories between women and men.

In various studies, men are slightly more likely to retire late than women (Axelrad (2018); De Preter et al. (2013); Markova (2020)).

Labour market restrictions

Besides individual choices, the literature identifies that early retirement may come from limitations enforced by the labour market (Fasang (2010)). In these cases, early retirement is mostly involuntary (Wang and Shultz (2010)).

Table 1 shows an overview of these determinants, including the types of retirement they influence and whether they are considered Push/Pull, in the case of early retirement.

Table 1: Determinants of retirement timing by type of retirement and push/pull mechanisms

Determinants	Early Retirement	On-time Retirement	Late Retirement / Working past LRA	Push / Pull
Health, functional limitations and wellbeing	Poor health and functional limitations	–	Good health	Push (poor health) / Pull (good health)
Working conditions and job quality	Physically demanding or stressful jobs; jobs in primary and secondary sectors	–	Flexible, autonomous and low-stress jobs; jobs in the tertiary sector	Push / Pull
Financial incentives and pension generosity	Financial security	–	Low pension generosity and financial need	Pull (early) / Push (late)
Labour market restrictions	Unemployment, age discrimination and rigid labour markets	–	–	Push
Social network and family context	Spousal retirement, grandchildren and dense networks	–	Limited evidence for late retirement	Pull
Education and qualifications	Lower education	–	Higher education	Push / Pull
Gender	Women more influenced by family context; men by health shocks	–	Gendered patterns persist in delayed retirement	Context-dependent

2.3 Retirement in Portugal

2.3.1 Legal Framework

After the Revolution of 25th April 1974, Portugal underwent profound demographic, political and socio-economic changes Costa, Lima, Loureiro, and França (2025), which influenced the construction of the current Social Security System. The evolution of social rights, universal access to social security and the need to respond to the progressive ageing of the population led to growing pressure on the pension system, requiring successive legislative reforms. The Constitution of the Portuguese Republic, following the **VII Constitutional Revision of 2005**, establishes in **Artigo n.º 63** (Article 63) the right to social security, determining that the State must protect citizens in old age and that all working time must be taken into account in the calculation of pensions. This principle is reinforced by **Artigo n.º 72** (Article 72), which recognises the right of elderly people to economic security and participation in community life. Together, these provisions constitute the normative basis for the development of legal protection schemes for old age, including mechanisms allowing flexibility in the retirement age.

Although **Lei n.º 4/2007, de 16 de janeiro** (Law No. 4/2007) defines the general framework of the social security system, it is **Decreto-Lei n.º 187/2007, de 10 de maio** (Decree-Law No. 187/2007) that structures the old-age and disability pension scheme. This decree establishes the legal age for access to an old-age pension, initially set at 65 years (**Artigo n.º 20** (Article 20)), and introduces a flexible retirement age regime. This regime includes penalties associated with early retirement (**Artigo n.º 25 e 36** (Articles 25 and 36)) and bonuses in cases of postponed retirement (**Artigo n.º37 e 38** (Articles 37 and 38)).

Since 2012, in the context of the financial crisis and the Economic and Financial Assistance Programme, the flexible retirement age regime has been conditioned. Under **Decreto-Lei n.º 85-A/2012, de 5 de abril** (Decree-Law No. 85-A/2012), voluntary access to early retirement was restricted, except for long-term unemployed individuals. Subsequently, **Decreto-Lei n.º 167-E/2013, de 31 de dezembro** (Decree-Law No. 167-E/2013) indexed the legal retirement age to changes in average life expectancy.

In 2015, **Decreto-Lei n.º 8/2015, de 14 de janeiro** (Decree-Law No. 8/2015) revoked the suspension of early access to retirement pensions, introducing stricter eligibility conditions, namely a minimum age of 60 and at least 40 years of contributions. This regime was lately revised by **Decreto-Lei n.º 10/2016, de 8 de março** (Decree-Law No. 10/2016).

Decreto-Lei n.º 126-B/2017, de 6 de outubro (Decree-Law No. 126-B/2017) introduced

penalty-free early retirement for individuals with contributory careers of 48 years or more, or for those who began working at age 14 or younger and, at age 60, had completed at least 46 years of contributions.

In 2018, **Decreto-Lei n.º 73/2018, de 31 de agosto** (Decree-Law No. 73/2018) extended this scheme by increase the minimum age for the beginning of a contributory career from 14 to 16. At the same time, **Decreto-Lei n.º 119/2018, de 27 de dezembro** (Decree-Law No. 119/2018) introduced the concept of personal pension age and clarified that the flexible regime allows access to a pension below or above the reference age, while maintaining incentives for postponed retirement.

Finally, **Lei n.º 5/2022, de 11 de janeiro** (Law No. 5/2022) created a new early retirement scheme for old-age pensions due to disability, which was regulated by **Decreto-Lei n.º 18/2023, de 17 de março** (Decree-Law No. 18/2023).

Portuguese retirement legislation is characterised by a system based on flexibility in the age of access to pensions, which has been successively adjusted over time. Legislative changes have primarily focused on early retirement, while postponed retirement remains part of the general system.

Figure 1 shows a timeline of the most important legislative changes in Portugal on pension system.

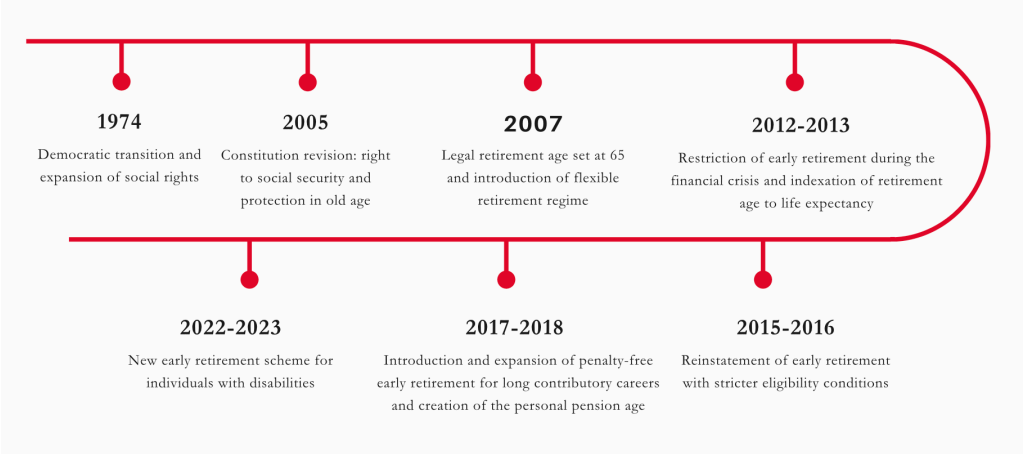


Figure 1: Main legislative milestones of the Portuguese pension system

The following sections analyse the early retirement schemes and their respective penalties, the incentives associated with late retirement, and the distinction between the normal age and personal age for accessing a pension.

2.3.2 Legal Retirement Age

The legal age of access to the old-age pension serves as a central reference for the pension allocation without penalties and for the calculation of early retirement situations.

Life Expectancy Adjustment Mechanism

Since 2014, the legal retirement age is no longer fixed and now varies according to changes in average life expectancy at age 65. This mechanism was introduced with the aim of protecting the financial sustainability of the social security system.

Evolution of the Legal Retirement Age

During the period covered by the data in this study, the legal retirement age was fixed in 66 years and 6 years, in 2021, and 66 years and 7 months, in 2022, the last one being the legal retirement age used in the analysis. In the following years, 2023 and 2024, there was an exceptional reduction to 66 years and 4 months, associated with the COVID-19 pandemic impact on life expectancy, and this was temporary. It is not part of the analysed period. In 2025, the normal age for accessing the old-age pension was again set at 66 years and 7 months, reflecting the increase longevity of the Portuguese population.

Personal Retirement Age

The concept of personal retirement age was first introduced by the **Decreto-Lei n.º 119/2018, de 27 de dezembro** (Decree-Law No. 119/2018) and corresponds to the age resulting from reducing the legal pension age by four months for each year of contributions exceeding 40 years. Under no circumstances it is allow access to a pension before the age of 60.

2.3.3 Early Retirement: Access Schemes, Penalties and Sustainability factor

In Portugal, access to early retirement is conditioned by a set of factors, including age, length of contribution history, and the beneficiary's employment status at the time of retirement.

The general Social Security System provides four different early retirement schemes that reflect the diversity of career paths and situations of vulnerability throughout the working life.

Early Retirement for Long Careers

The first scheme corresponds to Early Retirement for Long Careers, which allows beneficiaries with at least 40 years of contributions to access their pensions from the age of 60. However, whenever the pension is requested before the personal retirement age, the respective amount is subject to the general calculation rules applicable to early retirement. This penalty does not apply in cases of very long contributory careers, i.e. when the beneficiary has 48 or more years of contributions, or 46 years when the career began before the age of 17. In these cases, access to the pension at the age of 60 is allowed without any reduction.

Early Retirement for Long-Term Unemployment

The Early Retirement for Long-Term Unemployment scheme is another form of early access to a pension and applies to beneficiaries who have lost their jobs involuntarily and have exhausted their unemployment benefits. Depending on the age and contribution history, it is possible to grant a pension from the age of 57, in which case the amount is subject to penalties until the age of 62. Alternatively, if the person is 57 or older on the date of unemployment and has at least 15 years of contributions, they can retire at 62 without any reduction.

Early Retirement under Special Schemes

The third scheme covers Early Retirement under Special Schemes and is aimed at dangerous or physically demanding professions. It covers professional groups such as miners, maritime workers and fishermen, firefighters, dancers, embroiderers from Madeira, among others. Each one of these cases has its own eligibility rules, which are often more favourable and, in many cases, without penalties.

Early Retirement with Penalties

Finally, there is the Early Retirement with Penalties scheme, also known as voluntary retirement. This scheme allows workers to retire before the legal retirement age, as long as they meet the legal requirements, namely in terms of age and contribution history. In this case, early retirement before the personal or normal retirement age implies the application of legally established penalties

Penalties and Sustainability Factor

Early Retirement is generally associated with penalties on pension amounts, regardless of the access regime, although there are legally established exceptions. These penalties can take two main forms: monthly reductions and, in certain situations, the application of the sustainability factor. The monthly reduction for early retirement corresponds to a penalty of 0.5% for each month prior to the personal age of eligibility for an old-age pension. The number of months of anticipation is calculated by reference to the personal retirement age or, where applicable, the legal pension age in force. This reduction is permanent and applies to most early retirement situations.

In addition to this monthly penalty, the sustainability factor can also be applied. This mechanism was introduced with the aim of adjusting the pension amount to changes in average life expectancy. Its application depends on the specific early retirement scheme and the legal framework in force. However, the legislation provides for exceptions to its application, particularly in cases of very long contributory carers.

2.3.4 Late Retirement: Bonuses and Incentives to Prolong Working Life

In Portugal, late retirement is integrated into the general old-age pension scheme. Its legal framework is based on incentives for active ageing, promoting continued participation in the labour market beyond retirement age.

Late retirement and pension bonus

The legal system provides for the granting of a bonified pension if the beneficiary chooses to postpone retirement beyond the normal or personal retirement age. The bonus is then calculated based on the number of additional months of actual work after the reference age and based on the beneficiary's contribution history, subject to legally established maximum limits.

Bonus in situations of early access without penalty

In addition to this bonus, there is also a mechanism applicable to beneficiaries who, having met the conditions for early retirement without penalty, have chosen to continue working.

2.3.5 Portugal in the European context

Recent projections suggest that, although population ageing will accelerate significantly until mid-century, the growth of the dependency ratio is expected to slow down after 2050, and may even decline slightly in some countries, including Portugal (OECD (2025)).

The current situation of the Portuguese pensions system already reflects this demographic ageing. In 2023, the number of old-age pension beneficiaries in Portugal was 2 117 487 individuals, an increase of 61,6% compared to 1990.

The ageing of the population in developed countries is essentially driven by two demographic factors: declining fertility rates and increased average life expectancy.

Fertility Rates

According to the OECD report: Pensions at a Glance 2025 (OECD (2025)), OECD countries have been witnessed a continuous decline in total fertility rates over the last decades. This evolution translates into levels of fertility below the so-called replacement level, that is, the average number of children per woman needed to ensure long-term population replacement, estimated at 2.1 in developed countries. In 2024, the average fertility rate in the OECD was 1.46 (OECD (2025)).

In the Portuguese case, although the fertility rate in 2024 (1.51) was above the average of OECD, Portugal follows the same tendency observed in the developed countries. The fertility rate in Portugal decreases from 3.24 children per woman in 1964 to 1.51 in 2024, reflecting an accentuated reduction over the last six decades. The projection indicated that, in 2064, the fertility rate will be at 1.59, remaining below the replacement level.

Average Life Expectancy

At the same time, it has been observed a continuous increase in average life expectancy. Despite this being one of the biggest achievements in the last century, it brings serious challenges to the financial sustainability of the pensions systems.

According to the Pensions at a Glance 2025 report (OECD (2025)), in 2024, the average life expectancy at 65 years in OECD countries stood, in average, at 86.6 years for women and 83.5 years for men. In Portugal, these values are also high. In 2024, women aged 65 could expect to live approximately another 23 years, while men at the same age were expected to live more 19 years. In most OECD countries, remaining life expectancy at 65 years is predicted

to increase 3.7 years among women and 4.2 years among men by 2065. Thus, projections estimates that, in Portugal, the remaining life expectancy at 65 years will increase to 26.7 years for women and 23.2 years for men.

Table 2: Fertility rate and remaining life expectancy at age 65 in Portugal and the OECD

	Fertility Rate			**Average Life Expectancy at 65**			
	1964	2024	2064	2024		2065	
				Women	Men	Women	Men
Portugal	3.24	1.51	1.59	23.0	19.0	26.7	23.2
OECD	3.27	1.46	1.53	21.6	18.5	25.3	22.7

Normal Retirement Age

Due to this ageing of population, most OECD countries have been changing their pensions system, mainly through the increase in the normal retirement age.

The OECD (2025) definition of normal retirement age is the age at which workers with, at least, 22 years of contributions can access their pensions without penalty. According with this definition, the normal retirement age in Portugal, in 2024, was 65 years and 3 months, value above the average of OECD (64 years and 7 months). It is important to highlight that the normal retirement age reported by OECD does not corresponds to the legal retirement age in place in Portugal for the same year (66 years and 4 months). This happens because the OECD incorporates Portugal's existing long-career schemes into its calculations, which reduces the retirement age without penalties, compared to the legal retirement age.

Projections indicate that the tendency to increase the normal retirement age will continue in the next years (OECD (2025)). For individuals entering the labour market in 2024, the normal retirement age in Portugal, according to the OECD's harmonized methodology, will be, approximately, 68 years. Although this value is higher than the average normal retirement age in OECD countries, it is below the projections for other countries, such as Denmark or Estonia, where the normal retirement age could reach 74 and 71 years, respectively.

Early Retirement, Late Retirement and Financial Sustainability

Alongside raising the legal retirement age, most OECD countries have strengthened mechanisms to control early retirement, discouraging early exits from the labour market through permanent reductions in the value of benefits. Portugal is one of these countries, where the permanent reduction corresponds to a reduction of 0.5% for each month of early retirement, as defined in social security legislation.

On the other hand, Portugal stands out from other OECD countries due to incentives associated with postponing retirement. In average, bonuses rate for retiring one year after the normal retirement age is 4.8% across OECD countries. In Portugal, this value is 12% for those with more than 40 years of contributions.

The financial impact of these decisions must be well interpreted. According with the OECD (2025) report, the public expenditure on pensions in Portugal was 12.9% of GPD, which is above the OECD average (8.1%). Despite these changes, Portugal will continue to face significant pressure on public expenditure related to pensions over the coming decades, as shown in the European Commission. Directorate General for Economic and Financial Affairs. (2024) Report , which predicts that between 2022 and 2040 there will be an increase from 12.2 to 14.7, before it stabilizes in the following decades.

To sum up, Portugal is dealing with an intense demographic aging, driven by persistently low fertility rates and increasing life expectancy. The response has been to raise the legal retirement age, control early retirement, and create strong incentives for late retirement.

2.4 Methodological Foundations

The literature on the study of retirement decisions mostly uses traditional econometric models, generally focused on binary frameworks that distinguish between retired and non-retired individuals, or models that analyse the timing of exit from the labour market (Delattre and Moussa (2018); Leinonen et al. (2022)).

Although these approaches are relevant in the study of determinants, they are limited when the main objective is to study different retirement paths (Axelrad (2018); Dorn and Sousa-Poza (2010)).

Recently, there has been an increasing interest in the use of machine learning methods for applied economics, especially in contexts where the goal is the prediction and classification of individual behaviours. Mullainathan and Spiess (2017) describes machine learning as a complement to traditional econometric tools, capable of uncovering generalised patterns and complex structures. Independent of parameter estimation, machine learning models are able to fit data in highly complex ways without overfitting.

When an outcome involves categorising different classes, there is a classification problem (Alnuaimi and Albaldawi (2024); Grandini, Bagli, and Visani (2020)). This is a supervised learning technique, meaning that labelled data is used to train a predictive system, with the aim of this technique being to assign unknown patterns to known classes (Alnuaimi and Albaldawi

(2024)). When studying the retirement decisions, a multi-classification problem appears as the best approach, as it allows for the classification of different outcomes into more than two categories (Alnuaimi and Albaldawi (2024); Grandini et al. (2020)), going beyond the simplification of binary approaches.

2.4.1 Multiple Imputation

According to Dong and Peng (2013), the presence of missing values represents a frequent challenge in empirical research, as it can lead to a reduction in statistical power, an inflation of standard errors, and the loss of important information.

There are several methods for dealing with missing values. The most common are: complete cases, which removes all cases containing missing values; mean substitution, which imputes missing values using the mean of all available information; last observed carried forward, which replaces all missing values with the most recent observed value; and multiple imputation, that creates complete data sets, replacing each missing value with m plausible values (Rubin (1987); Zhu (2014)).

Before applying any imputation technique, it is necessary to identify the statistical mechanism that caused the data loss. The literature on missing values distinguishes between three possible mechanisms: Missing Completely at Random (MCAR), where the probability of a data point being missing does not depend on both observed and unobserved data; Missing at Random (MAR), where the probability can depend on the variables that were actually observed in the data set, but not on the missing variable itself; Missing not at Random (MNAR), where the probability of a value being missing depends on the value that would be observed (Dong and Peng (2013); Rubin (1987)).

As an alternative to Single Imputation, Multiple Imputation creates complete data sets, replacing each missing value with m plausible values. This method incorporates the uncertainty associated with estimating the gaps with missing values Zhang (2015).

One of the most commonly used implementations is the Multiple Imputation by Chained Equations (MICE). This method, also known as Fully Conditional Specification (Van Buuren, Brand, Groothuis-Oudshoorn, and Rubin (2006)) creates an individual predictive model for each variable containing missing values, using the remaining variables as predictors.

Different imputation methods can be used for different types of variables. This strategy allows each variable with missing values to be imputed using its own conditional algorithm. For numeric variables, predictive mean matching, Bayesian linear regression, linear regres-

sion (non-Bayesian), Unconditional mean imputation and two-level linear model can be used. For binary categorical variables, logistic regression can be used. For nominal categorical variables with more than two categories, the multinomial logit model can be applied. For ordinal variables with more than two categories, the odered logit model can be used (Buuren and Groothuis-Oudshoorn (2011)).

Each dataset is then analysed independently, with the results combined using Rubin's rule to ensure the correct integration of the uncertainty associated with the multiple imputation process. The final pooled point estimate ($\bar{Q}$) is calculated as the arithmetic mean of the estimates obtained from each of the m imputed datasets. The pooled variance (T) is expressed as the sum of within-imputation variance ($\bar{U}$), which represents conventional sampling uncertainty, and the between-imputation variance (B), which quantifies the statistical uncertainty generated by imputing missing data. This calculation includes an adjustment factor to compensate for the simulation error arising from the use of a finite number of imputations (Dong and Peng (2013); Rubin (1987)).

2.4.2 Multinomial Logistic Regression Model

The multinomial logistic regression model extends binary logistic regression to classification scenarios in which the outcome contains three or more categories (Hashimoto, Ortega, Cordeiro, Suzuki, and Kattan (2020)).

For an outcome with K categories, the model estimates K-1 log-odds equations relative to the reference category. Each set of coefficients represents the association between the predictor variable and the log-odds of belonging to a particular category relative to the reference category. Positive coefficients indicate higher probabilities of belonging to the corresponding category rather than to the reference category, while negative coefficients indicate lower relative probabilities (Hastie, Tibshirani, and Friedman (2001)).

Pooled standard error, p-values and 95% confidence intervals were derived from the combined estimations. Regression coefficients were reported as Odds-Ratio. Values above one indicates higher relative odds, and values below one indicates lower relative odds, of belonging to the respective category rather than to the reference category, conditional on the other variable sin the model.

2.4.3 Decision Trees

Decision Trees are a classic and transparent methodology for predictive modelling and supervised learning, widely used in classification tasks (Alnuaimi and Albaldawi (2024)). Graphically structured in a top-down way, the observations are classified by following a sequence of decision rules consisting of interconnected nodes and branches. At the top of the tree, the Root node is the starting point and corresponds to the best predictor (Alnuaimi and Albaldawi (2024)). The Internal nodes are decision points, where the value of a single variable determines which node is visited next. The Branches represent feature combinations that lead to decisions, and the Leaf nodes are the terminal nodes of the tree, which indicate the final classification (Alnuaimi and Albaldawi (2024); Mullainathan and Spiess (2017)).

The tree is constructed using recursive binary splitting, in which the algorithm repeatedly searches for splits that increase the homogeneity of the resulting nodes.

Decision trees can be estimated using the Classification and Regression Trees (CART) algorithm. This is a non-parametric supervised algorithm that repeatedly splits the attribute into rectangular sub-regions through binary splits, assigning a constant predictive value (the mean or the majority class) to each terminal leaf node of the tree (Breiman, Friedman, Olshen, and Stones (1984); Lewis (2000)).

Various impurity criteria can be used to evaluate impurity in the nodes, including the Gini index and entropy-based measures. The Gini Index measures the probability of incorrectly classifying a randomly selected element at the node (Breiman et al. (1984)).

Although Decision Trees are highly interpretable and capable of modelling non-linear relationships and interactions between predictors, the absence of an intrinsic stopping mechanism during splitting can lead to overfitting and high variance. Consequently, the complexity of the tree must be controlled (Alnuaimi and Albaldawi (2024)).

2.4.4 Random Forest

Radom forest, officially introduced by Breiman (2001), is a non-parametric supervisor ensemble learning method. Random Forests consist of a combination of decision trees, where each tree is based on a randomly selected vector sampled independently and from the same distribution for all trees in the forest, that will be used to determine the majority vote (Alnuaimi and Albaldawi (2024)).

To build a forest of X trees, the algorithm repeatly drawns n random samples with replace-

ment from the original training set (Breiman (2001)). When growing each tree, the algorithm does not evaluate all p available attributes at each split, but rather a random subset of m variables, and then selects the best one.

2.4.5 Support Vector Machine

The Support Vector Machine (SVM) was developed by Cortes and Vapnik (1995) has a supervised learning technique applicable to classification and prediction problems.

The method maps the relationship between the dependent variable and the independent variables, by identifying an optimal hyperplane that maximizes the margin of separation between classes, defined by the observations closest to the decision boundary, known as support vectors (Alnuaimi and Albaldawi (2024)).

In its original formulation, the SVM was designed specifically for binary classification problems. Its extension to multi-class problems is achieved using one of the following strategies: 'one-vs-rest', meaning to adjust a binary classification model for each class, relative to the other classes, or 'one-vs-one', meaning to adjust a binary classification model for each pair of classes (Alnuaimi and Albaldawi (2024)).

Chapter 3

Methodology

This chapter describes the methodological framework adopted to study the determinants and predictive patterns of retirement timings.

It begins by presenting the Survey of Health, Ageing and Retirement in Europe (SHARE) data set and define the analytical sample, the outcome variable and the predictor variables included in the study.

The sections describe the data preparation procedures, followed by the modelling strategy. Common procedures applied to the models are also described, particularly the class imbalance strategies, as well as others specifically applied to machine learning models, such as the division of the sample for training and model evaluation tools.

3.1 Analytical Sample

This analysis used data from the SHARE, a European longitudinal study that collects data on health, socioeconomic conditions, and social networks for individuals aged 50 and over, as well as their partners, in all participating European countries.

Portugal entered SHARE in 2011, in the Wave 4, and was part of all surveys to date, with the exception of the Wave 5 survey in 2013.

The population consists of Portuguese individuals who responded to the Wave 9 SHARE questionnaire, that corresponds to the year of 2022. To improve the data available for certain variables, information from Waves 4, 6 and 7, corresponding to the years of 2011, 2015 and 2017, respectively, were used, having been linked at the individual level using the SHARE

respondent identifier. Where the same variable appeared in several Waves, the information used corresponded to the latest valid information from the most recent Wave.

SHARE missing codes were recorded as Missing Values and treated prior to the construction of the sample.

The total number of portuguese individuals in Wave 9 is 1638. They were divided into two groups, based on the employment status information from the Employment & Pensions module: Retired (= 880) and Not Retired (= 739). Additionally, 19 individuals were considered as Missing Values.

For retired individuals, the year of retirement was obtained from the Wave containing the most recent valid response. The month of retirement, meanwhile, was obtained from the same Wave as the year of retirement. For instance, if an individual's year of retirement was obtained from Wave 7, but that Wave provided no information on the month of retirement, then that value is treated as a missing value.

The individual's year and month of birth were obtained from the Demographics module, following the same approach as for the year and month of retirement.

Age at Retirement was then calculated by comparing the year of birth with the year of retirement, and this value was recorded in months.

$$\text{Age at Retirement} = \text{Date of Retirement} - \text{Date of Birth}$$

For non-retired individuals, the year and month of birth were obtained following the same approach as the retired individuals, and the individual's age in December 2022 was calculated in months.

$$\text{Age} = (2022 - \text{Year of Birth}) * 12 + (12 - \text{Month of Birth})$$

3.1.1 Normal Retirement Age

Age and Age at Retirement were compared with the Normal Retirement Age (NRA).The normal retirement age corresponds to the age at which workers can access their pension without penalty.

For retired individuals, the difference between the age at retirement and the NRA was calcu-

lated in order to determine each individual's retirement timing, considering the year in which they retired. For example, if an individual retired in 2015, then their retirement timing will be calculated based on the 2015 NRA in Portugal.

$$\text{Retirement Timing} = \text{Age at Retirement (months)} - \text{Normal Retirement Age (months)}$$

If the value is negative, then they retired before the NRA; if positive, they retired after the NRA.

For non-retired individuals, their age in 2022 was compared with the NRA to determine whether or not they had already reached the retirement age. For context, the NRA in 2022 was 66 years and 7 months.

$$\text{Status relative to NRA} = \text{Age in 2022 (months)} - \text{Normal Retirement Age (months)}$$

If the value is negative, then they are Below NRA; if it's positive, they are Beyond NRA.

3.1.2 Outcome Definition

The dependent variable was constructed as a multi-class problem, in which individuals were classified into four mutually exclusive categories: On-time, Early, Late and Working beyond NRA.

To classify retired individuals, hybrid classification rules were used, so that the classification was not based solely on age. These rules were applied in the following order of priority:

1) If the individual indicated that they had become eligible to receive a state pension, a retirement pension, or a public pension, then they are classified as 'On-time'
2) If the individual indicated that they were offered the option of early retirement, they are classified as 'Early'
3) If none of the above applies, retirement age was used with a window of +/- 6 months
    3a) If they retired before the 6-month window, they are classified as 'Early'

3b) If they retired between the 6 months before and 6 months after the NRA, they are classified as 'On-time'

3c) If they retired after the 6-month window, they are classified as 'Late'

For the On-time category, the individual's reason for retiring always has priority over the difference between the retirement age and the NRA, because the individual may already have the conditions for claiming a pension without penalty, regardless of said distance.

The decision to add a window for classifying the outcome came from the necessity to adopt a tolerance system that would take into consideration small differences between the retirement age and the normal retirement age. So, a six-month threshold was adopted and used as a conservative temporal margin rather than as a parameter selected to optimise the distribution of the outcome categories.

Although a descriptive robustness check was carried out to analyse the distribution of retirement categories, a formal sensitivity analysis using these windows was not included in the main analytical design. This absence is recognised as a limitation, since alternative windows could affect the classification of individuals close to the NRA.

Table 4 displays the distribution of retirement categories across alternative windows, without being applied the hybrid rules for classification.

Table 3: Distribution of retirement timing categories across alternative retirement windows without hybrid classification

Window	On-time	Early	Late	NA
0 months	64 (48.64%)	664 (44.55%)	133 (6.14%)	19 (0.68%)
1 month	81 (49.43%)	659 (44.09%)	121 (5.80%)	19 (0.68%)
3 months	102 (49.89%)	648 (43.86%)	111 (5.57%)	19 (0.68%)
6 months	122 (50.91%)	638 (43.30%)	101 (5.11%)	19 (0.68%)
12 months	213 (55.11%)	596 (41.48%)	52 (2.73%)	19 (0.68%)

The 'Working beyond NRA' category consists of individuals who have already reached the NRA (Status relative to the NRA = Beyond NRA) and who have indicated that they remain employed or self-employed.

Individuals who do not meet any of these criteria do not have a valid analytical result and are, therefore, not included in the final sample.

Figure 2 summarizes the available information in each outcome.

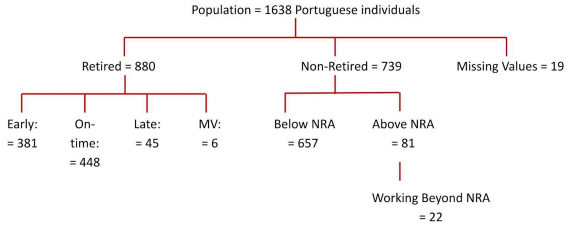


Figure 2: Construction of the outcome

3.2 Predictor selection and characterization

The variables were selected on the relevance found in the literature review and were divided into five different domains:

1. Demographic domain: gender, education, and marital status
2. Labour-market domain: employment type and industry sector
3. Financial domain: perceived financial difficulty of the household
4. Social network domain: network size, network satisfaction, and network closeness
5. Health and functioning domain: self-reported health, functional limitations, body mass index, quality of life, depressive symptoms, grip strength, chronic conditions, mobility and cognitive indicators

The Education variable was obtained from the Education modules in SHARE. For each individual, the most recent valid information from the Waves under analysis was selected. The reference international classification ISCED 2011 was used for Waves 6, 7 and 9, and, for Wave 4, the ISCED 1997. The education levels were then harmonized into Low (ISCED 0-2), Medium (ISCED 3-4) and High (ISCED 5+). Just like Axelrad (2018), this study assumes that education is acquired in the early years of life. Therefore, the level of education will be the same at the time of the interview and at the time of retirement.

Marital status at retirement was constructed for retired individuals by comparing the date of marriage, divorce and widowhood with the date of retirement. If no dated information is available for retired individuals, or the variable is being calculated for non-retired individuals, then the most recent observation is used.

Financial difficulty was measured at household level, with available responses being assigned to the other members of the same household. Responses that differed between households members were coded as missing.

Industry group was obtained from each individual's economic activity and was then grouped into three categories: Agriculture, Industry and Services.

The operationalization of the remaining variables follows the original SHARE structure. The final coding, measurement scale and interpretation of each variable are summarized in Table 4.

Table 4: Operationalization of predictor variables

Domain	**Variable**	**Description**	**Final coding/measurement**
Demographic	Gender	Self-reported gender obtained from the SHARE demographic module	0 = Female; 1 = Male
	Education level	Highest available education, harmonized using ISCED	Low; Medium; High
	Marital status	Marital status at retirement for retired individuals, or most recent observed marital status	Married/Partnership; Single; Divorced; Widowed
Labour-market	Employment type	Employment type associated with the current or previous job	Private-sector employee; Public-sector employee; Self-employed
	Industry sector	Economic activity of the current or previous occupation, grouped into broader sectors	Agriculture; Industry; Services
Financial	Financial difficulty	Self-reported difficulty in making ends meet at the household level	With great difficulty; With some difficulty; Fairly easily; Easily
Social Networks	Network size	Number of individuals included in the respondent's social network	0–7 persons
	Network satisfaction	Self-reported satisfaction with the respondent's social network	1–10
	Network closeness	Average perceived closeness to members of the respondent's social network	Continuous mean score

Domain	Variable	Description	Final coding/measurement
Health and functioning	General health	Respondent's self-reported general health status	Excellent; Very good; Good; Fair; Poor
	ADL limitations	Indicator derived from limitations in Activities of Daily Living	0 = No ADL limitations; 1 = 1+ ADL limitations
	IADL limitations	Indicator derived from limitations in Instrumental Activities of Daily Living	0 = No IADL limitations; 1 = 1+ IADL limitations
	Body Mass Index	Body mass index derived from SHARE health information	Underweight; Normal; Overweight; Obese
	CASP	Quality-of-life measure based on the CASP instrument	12–48
	EURO-D	Measure of depressive symptoms based on the EURO-D instrument	0 = Not a case; 1 = Depressive-symptom case
	Maximum grip strength	Maximum available hand-grip strength measurement	Continuous grip-strength measure
	Chronic conditions	Number of reported chronic health conditions	Count
	Mobility limitations	Indicator of reported mobility limitations	0 = No mobility limitations; 1 = 1+ mobility limitations
	Immediate recall	Cognitive-function measure test of memory	0–10 words
	Delayed recall	Cognitive-function measure test of memory after a delay time	0–10 words
	Orientation	Cognitive orientation measure	0–4

3.3 Exploratory Data analysis and Data Quality

Prior to the exploratory data analysis, a series of data quality checks were conducted to verify data leakage, the presence of implausible values and coding issues.

An exploratory data analysis was then performed to characterize and analyze the sample. Categorical variables were analyzed using absolute and relative frequencies, while continuous variables were analyzed using descriptive statistics.

3.4 Missing Data and Multiple Imputation

Missing values were analyzed for all variables, and it was concluded that the proportion of missing values varies substantially depending on the variable.

Of the 896 individuals in the sample, 538 (=58.9%) have complete data across all variables, while 41.1% have at least one missing value. If relying only on complete cases, a large proportion of the sample would be lost. For this reason, multiple imputation was adopted to retain the available information, while accounting for uncertainty associated with the missing values.

Missing values were handled using Multiple Imputation by Chained Equations (MICE). Twenty imputed datasets were generated, with 20 iterations for each imputation procedure and fixed random seed of 1234. In this study, the MICE model assumes MAR. As it was not possible to test statistically whether the probability of missing data depends on the missing value itself, its plausibility had to be substantially argued. Two decisions were taken to make the MAR assumption more acceptable. Firstly, the MICE matrix includes, for each variable with missing values, all variables in the dataset as potential predictors. Secondly, the dependent variable itself was included as a predictor in all imputation models, even though it was never imputed itself.

Following the Fully Conditional Specification method, different imputation methods were used for the different types of variables. Table 5. displays the different imputation methods by variable.

Table 5: Imputations methods by variable

Variable	Missing n	Missing %	Imputation method
Industry group	243	27.12 %	Multinomial logistic
Employment type	230	25.67%	Multinomial logistic
Network closeness	93	10.38 %	Predictive mean matching
Network satisfaction	84	9.38 %	Predictive mean matching
Maximum grip strength	39	4.35 %	Predictive mean matching
Education level	21	3.24 %	Proportional-odds logistic
EURO-D	17	1.90 %	Binary logistic
CASP	15	1.67 %	Predictive mean matching
Delayed word recall	14	1.56 %	Predictive mean matching
Marital Status	13	1.45 %	Multinomial logistic
Immediate word recall	13	1.45 %	Predictive mean matching
BMI	10	1.12 %	Proportional-odds logistic

The quality of the multiple imputation process was assessed through five complementary checks. Firstly, the events recorded by the algorithm were examined, identifying instances where a variable had to be excluded as a predictor in a specific iteration.

Secondly, the convergence of the imputation chains was assessed by visually inspecting the mean and standard deviation of the imputed values across the 20 iterations, for each of the 20 imputations.

Thirdly, the plausibility of the imputed values was assessed by comparing the distribution of the observed values with that of the imputed values, for the numerical variables with the highest proportion of missing data.

After that, it was confirmed that there were no remaining missing values after the imputation process; and, finally, it was verified whether the proportions of the categories for categorical variables remained stable between the observed data and the complete data.

3.5 Modelling Framework

Common modelling strategy

Due to the existence of class imbalance in the outcome, the four models were estimated using three strategies: baseline, class-weighted and oversampling. In addition, the three machine learning models followed a common modelling and evaluation framework, including the same training/test split, cross validation procedure, performance metrics, and class-specific permutation importance approach.

The baseline strategy preserved the original class distribution, serving as a benchmarking for assessing the impact of the other two strategies. The class-weighted strategy assigns greater importance to observation from less frequent outcome categories. The weights for each class were calculated using the formula N/((K * n_K)), where N = sample size, K = number of outcome categories and n_K = the number of observations in category K. The results of the weights were 0.588 for Early, 0.500 for On-time, 4.978 for Late and 10.182 for Working Beyond NRA. The oversampling strategy increases the representation of minority classes in the training dataset, using randomly oversampled with replacement until each outcome group matched the size of the larger class.

For the three machine learning models, the analytical sample was divided into a stratified train set, containing approximately 80% of the observations, and into an independent test set, containing the remaining 20%. The train dataset contains 718 observations, being 359 from the On-time category, 305 from Early, 36 from Late and 18 from Working Beyond NRA. The test sample contained 178 observations: 89 On-time, 76 Early, 9 Late and 4 Working Beyond NRA.

Hyperparameter selection was carried out exclusively using 5-fold cross-validation, repeated five times.

To robustly evaluate the classifiers, three overall measures were used: Accuracy, Balanced Accuracy, and Macro F1-score. Balanced Accuracy was selected as the primary metric for hyperparameter tuning, to ensure that minority and majority classes have equal influence during model selection. In addition, class-specific Recall and F1 scores were calculated to monitor the model's performance on under-represented categories (specifically Late and Working Beyond NRA). If a minority class was never predicted by the models, it's class-specific F1 score was strictly set to zero and included in the denominator of the macro F1 score, to prevent an artificial inflation of the overall average.

An exploratory comparative analysis was carried out for the nine pipelines using the Friedman's test to analyze the relative consistency of performance across the 20 imputed datasets.

A class-specific permutation importance was then applied to the pipeline with the most consistent overall predictive performance

Multinomial Logistic Regression

Considering that the outcome of this analysis contains four categories of retirement timing (Early, On-time, Late and Working Beyond NRA), a multinomial logistic regression model was used.

The 'On-time' category was defined as the reference category in the model.

Prior to modelling, categorical predictors with low or zero frequencies in the outcome were adjusted in order to mitigate problems of quasi-separation and numerical instability in the estimation of the likelihood.

Table 6 displays the variables that were recoded and why.

Table 6: Categorical Predictors – Recodification

Variable	Initial categories	Recodification	Reason
Marital Status	Married/Partnership; Single; Divorced; Widowed	Married/Partnership; Not married	Sparse cells in "Late" and "Working Beyond NRA"
Industry group	Agriculture; Industry; Services	Services; Non-services	Empty cells in "Working Beyond NRA"
General health	Excellent; Very good; Good; Fair; Poor	Good or better; Fair or poor	Empty cell in "Late"
BMI	Underweight; Normal; Overweight; Obese	Underweight/Normal; Overweight/Obese	Empty cells in "Working Beyond NRA"

Additionally, the outcome retirement timing and the variables Education Level, Financial Difficulty, General Health, and BMI were modeled as categorical factors (using dummy coding) rather than as ordered factors, to ensure the direct estimations of the log-odds ration without imposing ad-hoc polynomial contrasts.

The predictor variables were introduced in blocks corresponding to different domains so that associations with simpler specifications could be compared with those obtained as more domains were added. ML5 contains all the variables considered in the analysis.

Table 7: Multinomial Model Specification

Model	Predictors included
M1	Gender + education level + marital status
M2	M1 + employment type + industry sector
M3	M2 + financial difficulty
M4	M3 + network size + network satisfaction + network closeness
M5	M4 + general health + ADL2 + IADL2 + BMI + CASP + EURO-D + maximum grip strength + chronic conditions + mobility + immediate recall + delayed recall + orientation

The resulting parameters estimates were pooled using the Rubin's rule, and pooled regression coefficients, standard errors, p-values and 95% confidence intervals were obtained from the multinomial logistic regression models.

For easier interpretation, the coefficients were exponentiated and presented as odds ratios, using On-time as the reference outcome category.

The fit of the model was evaluated using the Aitake Information Criteria (AIC), the Bayesian Information Criteria (BIC) and the McFadden's Pseudo R^2. AIC and BIC evaluate the model while also penalizing its complexity. Smaller values indicate a more favorable balance between complex and fit. McFadden's Pseudo R^2 assesses the improvement of the model in study relative to the model without variables. Higher values indicate an increase in the improvement. The Variance Inflation Factor was used to analyze multicollinearity among the predictor variables. The higher this value, the higher the degree of multicollinearity.

Decision Trees

Following the Literature Review, Decision Trees were the second model to be implemented in this analysis.

Decision Trees can be affected by class imbalance, as splits and final predictions may become dominated by observations from the more frequent classes, reducing the model's ability to identify minority classes.

Decision Trees were estimated using the Classification and Regression Trees (CART) algorithm, which used the Gini criterion for classification splits.

Model complexity was controlled by the complexity parameter (cp), which determines the

extent to which additional splits are retained. Lower values allow for a deeper and more complex tree, while higher values prune them. A grid of 30 candidates cp values ranging from 0.0001 to 0.05 was evaluated, and the optimal value in each imputed dataset and modelling strategy was selected by maximizing cross-validated Balanced Accuracy.

Random Forest

Although Random Forest approach offers high stability and robustness against noise and outliers, this model is highly sensitive to class imbalance, meaning that, in scenarios of significant imbalance, the probability of successfully distinguish minority classes can be extremely low.

Random Forest models were estimated using the ranger implementation. Each fitted model contained 500 trees, and three model components were considered during tuning: mtry, which controls the number of candidate predictors considered in each split; minimum node size; and the splitting rule.

Five candidate mtry values (3, 5, 7, 10, 14) and three minimum node size values (1, 5, 10) were evaluated, resulting in 15 hyperparameters combinations, and the Gini impurity criterion was used as the splitting rule.

For each imputed dataset and imbalance strategy, the combination with the highest cross-validation Balanced Accuracy was selected.

Support Vector Machine (SVM)

Unlike the other algorithms, the SVM exhibits remarkable insensitivity to dataset composition. Despite this, imbalance correction strategies were still used in this model.

As SVM models are sensitive to differences in the scale of predictors, numerical predictors were centered and standardized prior to estimation. A radial kernel was used to capture any non-linear relationships between the predictors and the dependent variable. The multi-class problem strategy used in this study was 'one-vs-one'.

Two hyperparameters were tuned: C, which controls the penalization associated with classification errors, and sigma, which determines the width of the radial kernel.

The tuning grid contained five values of C (0.25, 0.5, 1, 2, 4) and four values of sigma (0.001, 0.005, 0,01, 0.05), which results in 20 candidate combinations. The selection of hyperparameters was performed using the same five-fold stratified cross-validation, with five repetitions, using 'Balanced Accuracy' as the optimization criterion.

Model Comparison

The Friedman test is a non-parametric statistical test designed to detect significant differences between multiple treatments repeated on a set of blocks.

The methodological pipeline of this study involves the tuning and comparison of 9 classification pipelines (three machine learning algorithms combined with three strategies of class imbalance), each evaluated across 20 imputations, that correspond to the 'treatments'.

If only the means of the performance metrics are compared, it is not possible to guarantee that a particular pipeline is statistically superior to another. The Friedman test can be used to overcome this limitation and confirm if the observed differences are genuinely significant or merely random.

For each of the evaluations, the pipelines were ranked, according to the metric under assessment, with the best-performing pipeline being assigned rank 1 and so on. The Friedman test compares the sum of these ranks to test the null hypothesis of equivalent performance across pipelines. When the overall p-value is less than 0.05, than the hypothesis is rejected, and a pair-wise comparison is carried out using the Wilcoxon test, with a Holm correction.

The Friedman test answers the question of which pipeline presents the most consistent predictive performance, but does not explain which characteristics most influence the distinction between each reform time. To find this out, the most balanced model was chosen for feature analysis.

Model Interpretation

A class-specific permutation importance analysis was implemented to the most balanced model to address the limitation of the class imbalance in the predictor identification. The importance of a predictor was operationally defined as the difference, in percentage points, in the Recall for a given class before and after the random permutation of its values.

Afterwards, the predictors were ranked independently for each outcome category. Lower classifications (closer to 1) robustly identify the most relevant determinants for characterizing each transition to retirement according to the model.

Lastly, a SHAP analysis was conducted to complement the class-specific permutation importance analysis, to assess the direction and magnitude of each predictor's contribution.

Chapter 4

Results

4.1 Outcome Distribution

The distribution of the retirement timing is imbalanced. The On-time category (=448) was the most frequent, followed by Early (=381). Approximately 92% of the observations are concentrated in these two categories. In contrast, Late (=45) and Working Beyond NRA (=22) are the smallest categories.

Figure 3 provides a visual representation of this distribution.

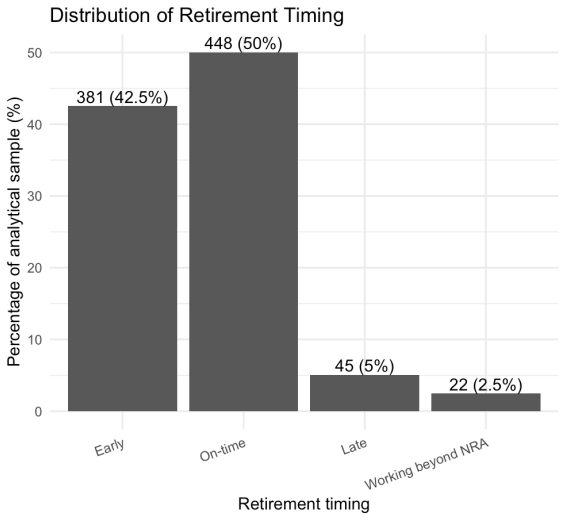


Figure 3: Distribution of Retirement Timing

All results in this section must be interpreted with caution, with regard to Late retirement and Working Beyond the NRA, due to the smaller number of observations in these categories.

4.2 Feature Analysis

First, it was confirmed the absence of data leakage, meaning that no variable used to construct the retirement timing was included as a predictor variable.

Univariate and bivariate analyses were conducted on the categorical variables. The univariate analysis describes the distribution of each categorical variable, while the bivariate analysis investigates the distribution of each variable according to the outcome.

Table 8: Distribution of categorical predictors by retirement timing

Variable	Category	n	Early	On-time	Late	Working Beyond NRA
Gender	Female	442 (49.3%)	184 (41.6%)	217 (49.1%)	33 (7.5%)	8 (1.8%)
	Male	454 (50.7%)	197 (43.4%)	231 (50.9%)	12 (2.6%)	14 (3.1%)
Education Level	Low	696 (77.7%)	310 (44.5%)	336 (48.3%)	37 (5.3%)	13 (1.9%)
	Medium	89 (9.9%)	37 (41.6%)	46 (51.7%)	4 (4.5%)	2 (2.2%)
	High	90 (10.0%)	27 (30.0%)	54 (60.0%)	2 (2.2%)	7 (7.8%)
Marital Status	Married/Partnership	721 (80.5%)	319 (44.2%)	349 (48.4%)	36 (5.0%)	17 (2.4%)
	Single	29 (3.2%)	15 (51.7%)	12 (41.4%)	2 (6.9%)	–
	Divorced	45 (5.0%)	20 (44.4%)	22 (48.9%)	–	3 (6.7%)
	Widowed	88 (9.8%)	27 (30.7%)	52 (59.1%)	7 (8.0%)	2 (2.3%)
Employment Type	Private-sector employee	404 (45.1%)	202 (50.0%)	175 (43.3%)	17 (4.2%)	10 (2.5%)
	Public-sector employee	162 (18.1%)	62 (38.3%)	86 (53.1%)	9 (5.6%)	5 (3.1%)
	Self-employed	100 (11.2%)	38 (38.0%)	46 (46.0%)	10 (10.0%)	6 (6.0%)
Industry Sector	Agriculture, forestry and fishing	57 (6.4%)	19 (33.3%)	31 (54.4%)	6 (10.5%)	1 (1.8%)
	Industry, construction, energy and water	204 (22.8%)	107 (52.5%)	91 (44.6%)	6 (2.9%)	–
	Services	392 (43.8%)	176 (44.9%)	186 (47.4%)	24 (6.1%)	6 (1.5%)
Financial Difficulty	With great difficulty	207 (23.1%)	110 (53.1%)	81 (39.1%)	13 (6.3%)	3 (1.4%)
	With some difficulty	299 (33.4%)	125 (41.8%)	151 (50.5%)	16 (5.4%)	7 (2.3%)
	Fairly easily	246 (27.5%)	79 (32.1%)	149 (60.6%)	12 (4.9%)	6 (2.4%)
	Easily	144 (16.1%)	67 (46.5%)	67 (46.5%)	4 (2.8%)	6 (4.2%)

Table 8 continued: Distribution of categorical predictors by retirement timing

Variable	Category	n	Early	On-time	Late	Working Beyond NRA
General Health	Excellent	20 (2.2%)	7 (35.0%)	11 (55.0%)	1 (5.0%)	1 (5.0%)
	Very good	30 (3.3%)	12 (40.0%)	14 (46.7%)	–	4 (13.3%)
	Good	215 (24.0%)	70 (32.6%)	131 (60.9%)	6 (2.8%)	8 (3.7%)
	Fair	378 (42.2%)	173 (45.8%)	176 (46.6%)	22 (5.8%)	7 (1.9%)
	Poor	253 (28.2%)	119 (47.0%)	116 (45.8%)	16 (6.3%)	2 (0.8%)
ADL	No ADL limitations	704 (78.6%)	302 (42.9%)	348 (49.4%)	34 (4.8%)	20 (2.8%)
	ADL limitations	192 (21.4%)	79 (41.1%)	100 (52.1%)	11 (5.7%)	2 (1.0%)
IADL	No IADL limitations	696 (77.7%)	304 (43.7%)	343 (49.3%)	29 (4.2%)	20 (2.9%)
	IADL limitations	200 (22.3%)	77 (38.5%)	105 (52.5%)	16 (8.0%)	2 (1.0%)
BMI	Underweight	15 (1.7%)	5 (33.3%)	9 (60.0%)	1 (6.7%)	–
	Normal weight	266 (29.7%)	100 (37.6%)	144 (54.1%)	14 (5.3%)	8 (3.0%)
	Overweight	404 (45.1%)	175 (43.3%)	197 (48.8%)	20 (5.0%)	12 (3.0%)
	Obese	201 (22.4%)	99 (49.3%)	90 (44.8%)	10 (5.0%)	2 (1.0%)
EURO-D	No depressive symptoms	511 (57.0%)	213 (41.7%)	264 (51.7%)	20 (3.9%)	14 (2.7%)
	Depressive symptoms	368 (41.1%)	162 (44.0%)	175 (47.6%)	25 (6.8%)	6 (1.6%)
Mobility	No mobility limitations	334 (37.3%)	132 (39.5%)	172 (51.5%)	12 (3.6%)	18 (5.4%)
	Mobility limitations	562 (62.7%)	249 (44.3%)	276 (49.1%)	33 (5.9%)	4 (0.7%)

Mobility in Working Beyond NRA is the strongest finding in the table: individuals without mobility limitations (=18) are observed almost eight times more often than those with limitations (=4).

Functional health affected Late and Working Beyond NRA in opposite directions: individuals with ADL, IADL limitations and depressive symptoms are more present in Late, while they are less present in Working Beyond NRA, when compared with individuals without these limitations.

Individuals with greater financial difficulties are more present in the early retirement outcome. Self-employed individuals stand out in Late and Working Beyond NRA categories. Women are more present in the Late category than men.

Lastly, this analysis also revealed empty cells, which were recoded, as mentioned before: for Working Beyond NRA, Marital status = Single, Industry group = Industry and BMI = Underweighted showed 0 cases; for Late, Marital status = Divorced and General health = 2 showed 0 cases.

Table 9 displays the descriptive analysis of continuous variables, with the box plots, shown in Figure 4, presenting information on the distributions of the continuous predictor variables across the different retirement timing categories.

Table 9: Descriptive statistics for continuous variables

Variable	N	Missing	Mean	SD	Median	Q1	Q3	Min	Max
network_size	896	0	2.80	1.77	3	1.00	4.00	0	7
network_satisfaction	812	84	9.23	1.25	10	9.00	10.00	0	10
network_closeness	803	93	3.04	0.61	3	2.67	3.50	1	4
casp	881	15	34.29	6.08	34	30.00	38.00	15	48
maxgrip	857	39	27.83	9.73	27	20.00	35.00	3	65
chronic	896	0	2.62	1.72	2	1.00	4.00	0	10
Immediate Reccal	883	13	3.87	1.84	4	3.00	5.00	0	9
Delayed Recall	882	14	2.41	1.96	2	0.00	4.00	0	9
orienti	896	0	3.33	1.24	4	3.00	4.00	0	4

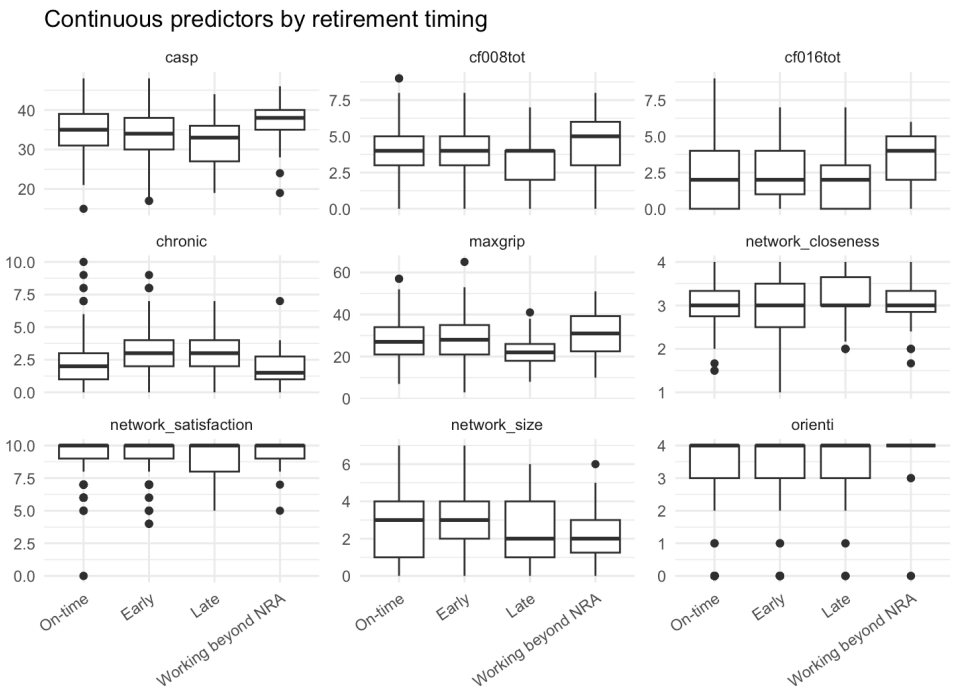


Figure 4: Box Plots - Continous Predictor Variables

The verification of outliers, based on the box plots in Figure 4, recorded no values outside of the possible ranges for each predictor variable. There was a higher concentration of outliers in Network satisfaction (= 66) and in temporal orientation (= 96), which is to be expected given that both exhibit a ceilling effect: the majority of individuals reported high levels of network satisfaction and high temporal orientation.

Several variables revealed noticeable differences between categories: individuals working beyond the NRA tend to have better results on cognitive tests (Immediate and Delayed recall), also presenting a higher CASP and maximum grip strength, while the category Late retirees had comparatively lower results for these last two variables. The number of chronic conditions appears to be higher in early and late retirees, and lower among those who are still working beyond the NRA.

Network closeness was found to be slightly higher in late retirees, while network size appeared to be smaller in the late retirement and Working Beyond NRA categories. Meanwhile, scores for network satisfaction and orientation were heavily concentrated at the upper ends of their respective scales, with little differentiation between categories.

Spearman's correlation analysis of the numerical predictors identified a single pair with potentially problematic redundancy. Immediate recall and Delayed recall showed a correlation of 0.747, which is higher than the reference threshold of 0.60. Both variables were still retained

in the model, as they represent conceptually distinct dimensions of cognitive function.

Missing values were also analyzed for all variables, and it was concluded that the proportion of missing values varies substantially depending on the variable. The highest proportion of missing values was found in Industry group (27.1%), followed by Employment type' (25.7%). Network closeness (10.4%) and network satisfaction (9.4%) also show a considerable proportion of missing values.

The percentage of missing values was less than 5% for all other incomplete variables (maximum grip strength, education level, EURO-D, CASP, Delayed word recall, Marital Status, Immediate word recall, BMI). The remaining variables (Gender, Financial difficulty, Network size, Health general, ADL2, IADL2, Chronic, Mobility and Orientation) present no missing values.

Figure 5 displays the distribution of missing values across variables.

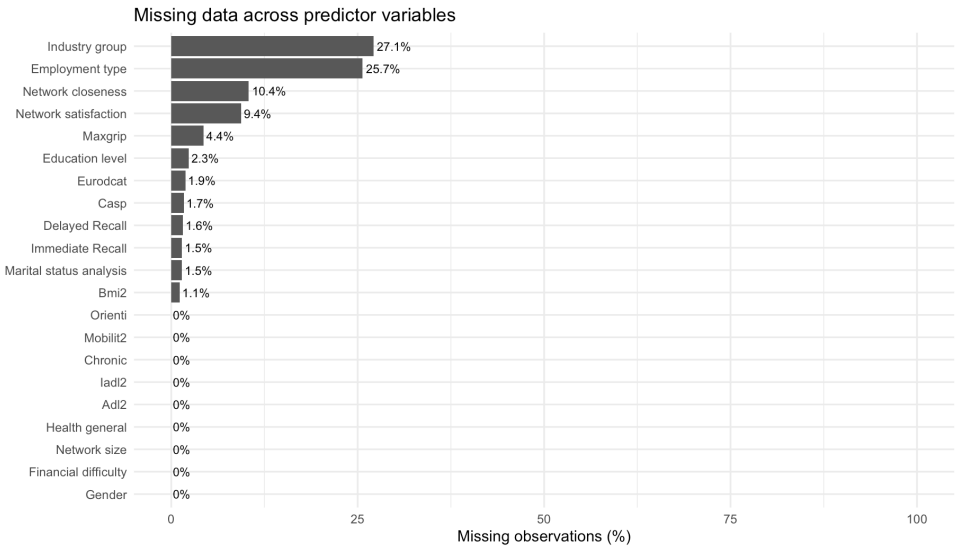


Figure 5: Missing Values per Predictor Variable

4.3 Multiple Imputation Analysis

Prior to imputation, 528 of the 896 observations (58.9%) were complete across the variables considered, while 368 (41.1%) contained at least one missing value.

Diagnostic checks of the MICE procedure did not identify any logged events, and no missing values remained in any of the completed datasets after imputation.

The following trace plots showed overlap between the imputation chains and no persistent

upward or downward trends.

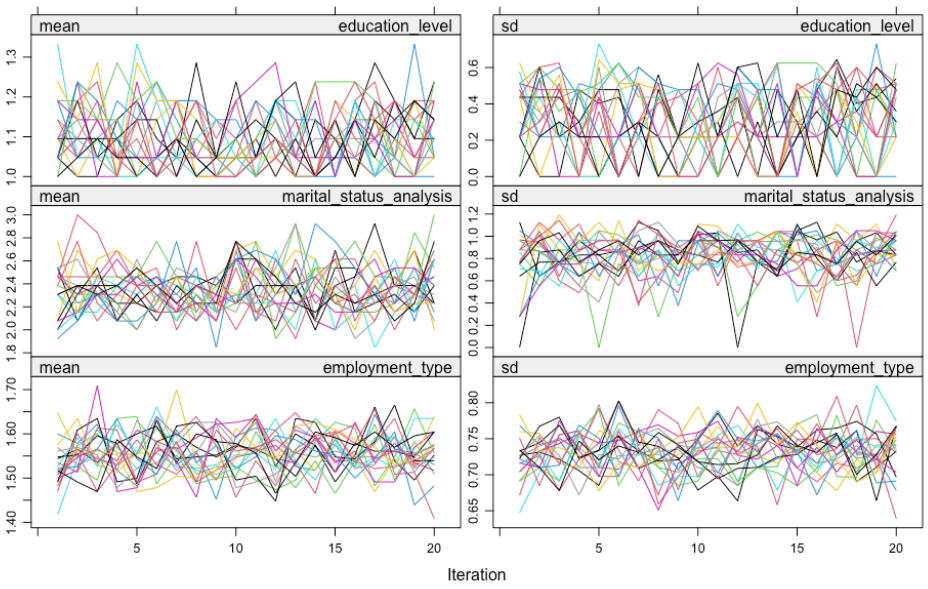


Figure 6: MICE Trace plot (Network Satisfaction, Network Closeness and CASP

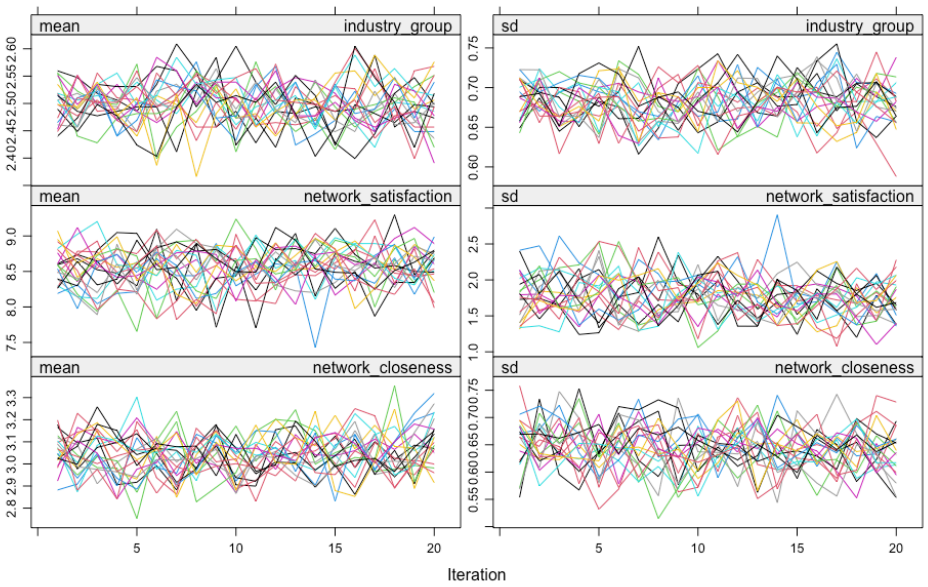


Figure 7: MICE Trace plot (Max Grip Strenght, Immediate Recall, Delayed Recall

Observed and imputed distributions were generally similar, with some discrepancies visible in network satisfaction and immediate and delayed recall.

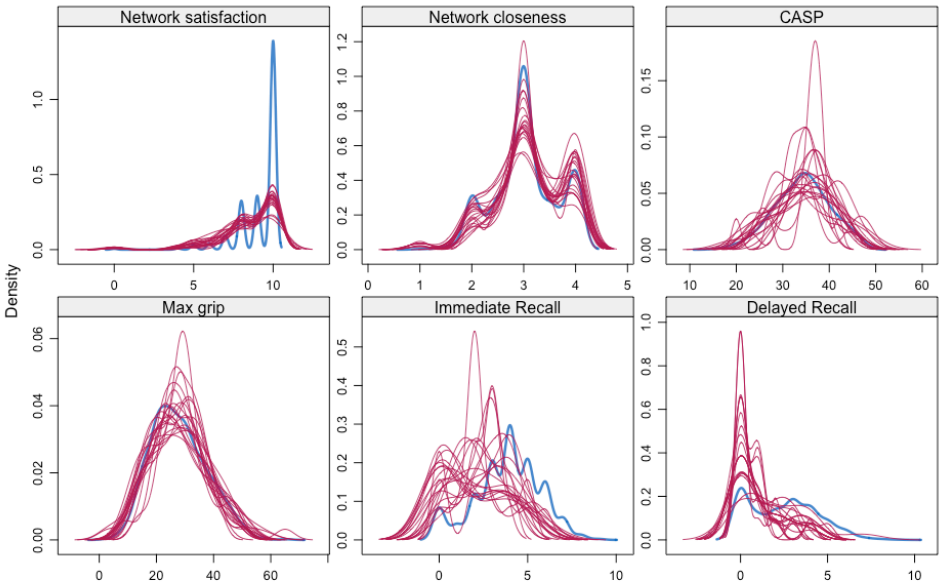


Figure 8: MICE distributions

For categorical variables, observed and completed proportions remained similar. For instance, private-sector employment changed from 60.7% to 60.3%, while services changed from 60.0% to 59.4%.

Overall, the diagnostic did not indicate problems introduced by the imputation procedure.

4.4 Multinomial Logistic Regression Results

Firstly, the results for AIC, BIC and Pseudo R^2 were analyzed across the M1, M2, M3, M4 and M5.

Table 10: Model fit statistics for the multinomial models

Model	AIC (mean)	BIC (mean)	Pseudo R^2 (mean)
M1	1704	1776	0.0183
M2	1703	1818	0.0296
M3	1698	1856	0.0428
M4	1706	1908	0.0487
M5	1714	2088	0.0863

Pseudo R^2 increased from 0.0183 in Model 1 to 0.0863 in Model 5, indicating that as more variables are added, the fit of the model improves. However, the AIC has the lowest value in

M3 (AIC = 1698), meaning that this model showed a more balanced performance in terms of both fit and complexity. At the same time, the BIC increased from 1776 in M1 to 2088 in M5, which indicates a preference for simpler variables.

The conclusion drawn from these results is that adding more variables to the models improves the fit, on one hand, but also increases its complexity.

The selection of the complete model (M5) as the main model comes from the theoretical nature rather than simply its predictive power. Each domain was introduced to test the contribution of the theoretically relevant variables to the retirement decision.

Finally, the multicollinearity analysis revealed no problematic evidence regarding the variables present in M5, with the maximum value being 2.67 for variable Immediate Recall.

To identify the following associations, the On-time category was used as the reference category.

Table11 displays coefficients organized by outcome. Only variables with $p < 0.10$ in at least one strategy are displayed.

Table 11: Significant and borderline associations from the multinomial logistic regression

Category	Term	OR	95% CI	p-value
Early	Financial difficulty (2 vs 1)	0.642	0.435–0.949	0.026
Early	Financial difficulty (3 vs 1)	0.469	0.303–0.725	< 0.001
Early	General health (Fair/Poor vs Good+)	1.569	1.080–2.280	0.018
Early	Chronic conditions	1.129	1.027–1.242	0.012
Early	Employment type (Public vs Private)	0.660	0.432–1.009	0.055
Late	Self-employed (vs Private)	2.155	0.901–5.158	0.085
Late	CASP	0.938	0.876–1.005	0.069
Working beyond NRA	Mobility limitation (vs none)	0.108	0.024–0.483	0.004
Working beyond NRA	Self-employed (vs Private)	3.448	0.987–12.051	0.052

Early Retirement

For the Early category, health and financial difficulties were the most significant variables found. The number of chronic conditions is positively associated with this outcome (OR = 1.13, p = 0.012), which indicates that each additional chronic condition is associated with an 13% increase in the odds of early retirement rather than On-time.

Poor self-rated health is also positively associated with early retirement. Individuals reporting fair/poor health have a higher odds ratio of retiring earlier (OR = 1.57, P = 0.018) than those in other health categories.

Financial difficulties also represent a significant factor in early retirement. When compared with the reference category for financial difficulty (1 - With great difficulty), individuals that responded With some difficulty and Fairly easily were associated with lower odds of early retirement.

The direction and magnitude of these associations generally remained stable across the class-weighted and oversampled strategies. However, some lost statistical significance when these strategies were implemented.

Late Retirement

For the Late category, there are not many statistically significant associations in the baseline strategy. Self-employment appears to be the most consistent association, with self-employment individuals having higher odds of being in the Late category than in the On-time category in the baseline model (OR = 2.16), while maintaining similar estimates in the class imbalance correction strategies.

In the class-weights and oversampling strategies, other associations were found. As they are sensitive to class imbalance, these results must be interpreted with caution. Male respondents have lower odds of late retirement than female respondents. Additionally, higher CASP values are also associated with lower odds of late retirement. In contrast, higher network closeness is associated with higher odds of late retirement.

Working Beyond NRA

Finally, for the Working Beyond NRA category, mobility limitations are the strongest association in the baseline strategy, with individuals with mobility limitations having lower odds of continuing to work beyond the NRA rather than retiring on-time (OR = 0.12, p = 0.004). Self-employment is also positively associated with Working Beyond NRA, with an OR = 3.45 in the baseline strategies and similar estimates in the other strategies.

Further associations emerged in the class-weights and oversampling strategies. Higher education was associated with higher odds of Working Beyond NRA while lower financial difficulties and greater network closeness are associated with reduced odds relative to their respective categories.

4.5 Decision Tree Results

Table 12 displays the predictive performance of the three class imbalance strategies of the Decision Tree model, plus the class-specific Recall.

Table 12: Predictive performance of the Decision Tree

Strategy	Accuracy	Balanced Accuracy	Macro F1	Recall On-time	Recall Early	Recall Late	Recall Working Beyond NRA
Baseline	0.509	0.521	0.261	0.611	0.476	0.000	0.000
Oversampling	0.213	0.569	0.169	0.120	0.268	0.483	0.650
Class weights	0.222	0.571	0.178	0.065	0.355	0.467	0.638

Class weights strategy had the highest Balanced Accuracy (= 0.571), while the Baseline strategy presented the highest Macro F1 (= 0.261) and the highest Accuracy (= 0.509).

When looking at Recall, the baseline strategy performed best for the 'On-time' category (=0.611) and 'Early' category (=0.476), although it was unable to identify the minority classes.

The oversampling strategy produced the most balanced Recall levels, with Recall Late = 0.483 and Recall Working Beyond NRA = 0.650, although this improvement was accompanied by lower Recall for On-time (= 0.120) and 'Early' (=0.268).

The class weights strategy reached a similar Recall for Late (=0.467) and Working Beyond NRA (=0.638), but substantially lower Recall for On-time (=0.065).

Overall, the increase in the detection of minority classes was accompanied by a decrease in the overall accuracy and in the recall for the majority classes.

4.6 Random Forest

Table 13 displays the predictive performance of the three class imbalance strategies of the Decision Tree model, plus the class-specific Recall.

Table 13: Predictive performance of the Random Forest

Strategy	Accuracy	Balanced Accuracy	Macro F1	Recall On-time	Recall Early	Recall Late	Recall Working Beyond NRA
Baseline	0.533	0.529	0.267	0.738	0.384	0.000	0.000
Oversampling	0.505	0.520	0.261	0.634	0.440	0.000	0.000
Class weights	0.476	0.512	0.256	0.556	0.457	0.011	0.000

The baseline strategy performance was very similar to the baseline performance in the Decision Tree model. This strategy achieved the highest value for Accuracy (= 0.533), Balanced Accuracy (=0.529), Macro F1 (=0.267) and Recall On-time (=0.738).

The baseline and the oversampling strategies have Recall = 0 for both the Late and Working Beyond NRA categories,

In this model, the use of class weights slightly improved Recall for the minority classes, with Recall Late = 0.011 and Recall Working Beyond NRA = 0.00, despite increasing the Recall for Early (=0.457).

4.7 Support Vector Machine Results

Table 14 displays the predictive performance of the three class imbalance strategies of the Decision Tree model, plus the class-specific Recall.

Table 14: Predictive performance of the Support Vector Machine

Strategy	Accuracy	Balanced Accuracy	Macro F1	Recall On-time	Recall Early	Recall Late	Recall Working Beyond NRA
Baseline	0.542	0.535	0.274	0.694	0.456	0.000	0.000
Oversampling	0.267	0.551	0.215	0.185	0.342	0.383	0.388
Class weights	0.312	0.539	0.238	0.266	0.371	0.306	0.250

Oversampling was the strategy with the highest Balanced Accuracy (=0.551), while the Baseline strategy achieved the highest Accuracy (=0.542) and Macro F1 (=0.274) values.

Regarding recall, the baseline strategy achieved the highest Recall for On-time (=0.694) and Early (=0.456), while achieving Recall = 0 for minority classes.

The use of Oversampling and Class-weights strategy improved the Recall in the minority classes, with oversampling providing the highest recall values for Working Beyond NRA (=0.388) and Late (=0.383), despite significantly reducing On-time recall (=0.185).

4.8 Friedman test

The three models show a very similar pattern when it comes to the Baseline strategy: they consistently achieve the highest Accuracy and Macro F1 scores and have highest Recall in the On-time and Early categories. However, the Late and Working Beyond NRA always have Recall equal to zero.

The class weights and oversampling strategies, in Decision Tree and Support Vector Machine models, increase recall for minority classes, although this was accompanied by a reduction in Recall for On-time. On the other hand, Random Forest remains essentially unable to identify minority classes.

Although this comparison revealed differences between the algorithms and class imbalance strategies, the differences alone do not determine whether a particular pipeline performs consistently better than the others across all imputations.

A Friedman test was then applied as an exploratory rank based-comparison of the nine pipelines across the 20 imputed datasets, being only considered as a robustness check of the relative pipelines ranking

The result revealed statistically significant differences between the pipelines, both for Balanced Accuracy (X^2=150.51, df = 8, p < 0.01) and for Macro F1 (X^2=148.24, df = 8, p < 0.01), suggesting that the models did not have equivalent performance across the 20 imputed datasets.

To complement these results, the nine pipelines were compared on the basis of their cross-validation performance. Balanced Accuracy and Macro F1 were retained as the main performance metrics, while the standard deviation of class-specific Recall was used as an additional indicator of the consistency of each pipeline's performance across the categories.

This analysis indicates a compromise between the overall predictive performance and the balance of class classification.

The mean ranks indicate how well each pipeline performed relative to the others across the different imputations, with lower ranks corresponding to better relative performance.

Table 15: Comparison of machine learning pipelines

Pipeline	Balanced Accuracy	Macro F1	Rank BA	Rank F1	Recall SD
SVM Oversampling	0.592	0.231	1.50	6.70	0.082
SVM Weights	0.590	0.258	1.50	5.55	0.045
DT Weights	0.558	0.162	3.40	8.65	0.208
DT Oversampling	0.557	0.168	3.65	8.15	0.202
RF Weights	0.542	0.297	5.25	1.00	0.253
RF Oversampling	0.539	0.286	6.45	2.00	0.277
RF Baseline	0.538	0.279	6.70	3.30	0.307
SVM Baseline	0.537	0.276	7.55	4.05	0.300
DT Baseline	0.520	0.260	9.00	5.60	0.276

The Oversample SVM achieved the highest Balanced Accuracy (= 0.592), followed by the SVM with weights (= 0.590), with both of them sharing the best average ranking for Balanced Accuracy (1.5). However, the SVM with weights has a higher Macro F1 and the lowest recall variance across all pipelines, indicating a more balanced performance across all classes.

The Random Forest with class weights achieved the highest Macro-F1 score, and the best corresponding average ranking (1), but it has a lower Balanced Accuracy and a higher Recall variance, which suggests that the performance was not uniform.

Overall, these results suggest that the SVM with weights provides the most balanced trade-off between performance and consistency in recall balance at class level. This was the reasons why it was selected for subsequent model interpretation.

4.9 Model Interpretation

Feature Importance

As mentioned in the previous section, the feature importance was conducted only for the SVM model with class weights.

Each feature was ranked independently for each retirement outcome, from most to least important in reducing the Recall of that category when permuted.

Table 16: Stable variables by retirement timing category

On-time		Early		Late		Working beyond NRA	
Variable	**Value**	**Variable**	**Value**	**Variable**	**Value**	**Variable**	**Value**
Financial difficulty	0.044 (20/20)	Chronic conditions	0.039 (20/20)	Gender	0.071 (19/20)	General health	0.142 (16/20)
Immediate Recall	0.024 (19/20)	Employment type	0.034 (19/20)	General health	0.041 (17/20)	Mobility	0.138 (19/20)
BMI	0.017 (19/20)	Mobility	0.025 (19/20)	Grip strength	0.011 (12/20)	Delayed memory	0.064 (15/20)
Orientation	0.015 (18/20)	IADL	0.014 (18/20)	Education	0.005 (12/20)	Employment type	0.061 (12/20)
ADL	0.008 (16/20)	BMI	0.014 (17/20)	–	–	–	–

Note: Values in parentheses indicate the number of imputations, out of 20, in which the variable was identified as stable.

Regarding Early Retirement, chronic conditions were the most significant and consistent factor (mean importance = 0.039, and identified as relevant in 20 out of 20 imputations), followed by type of employment (0.034, 19 out of 20) and mobility limitations (0.025, 19 out of 20).

In Late Retirement, gender appeared as the dominant characteristic (0.071, 19/20), followed by self-assessed health status (0.041, 17/20), while the remaining predictors proved to be less important.

Regarding On-time, financial difficulties were the strongest predictor (0.044, 20/20), followed by Immediate recall (0.024, 19/20), BMI (0.017, 19/20) and orientation (0.015, 18/20).

Finally, Working Beyond NRA showed the highest class specific importance values, with regard to self-rated health (0.142, 16/19) and mobility limitations (0.138, 19/20).

But this analysis still does not indicate the direction of the effect. For that, SHAP was used.

SHAP

SHAP analysis was used to further interpret the selected SVM pipeline by examining both the global importance of predictors and the direction of their contributions to each retirement timing category.

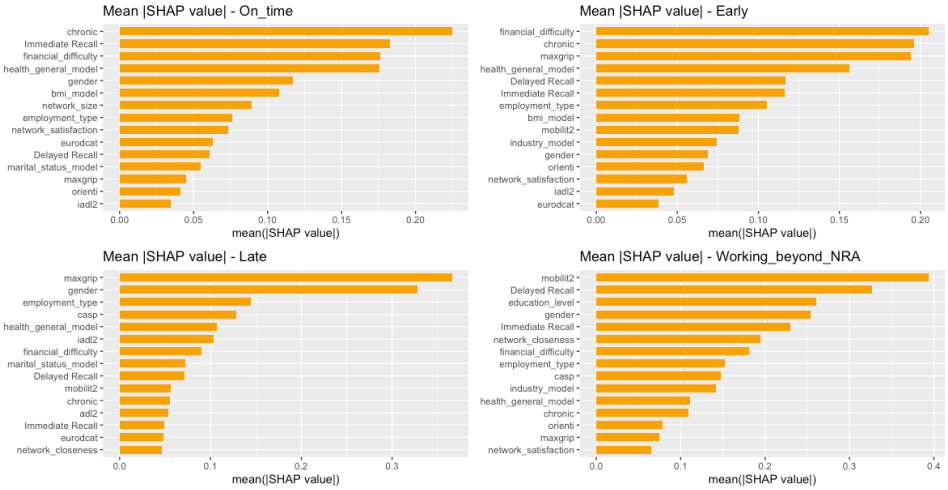


Figure 9: Mean SHAP plot

The global mean absolute SHAP values revealed distinct importance profiles across the four outcomes. For On-time retirement, the most influential predictors were chronic conditions, immediate recall, financial difficulty and general health, suggesting that the classification relied on a combination of health, cognitive and financial characteristics.

A similarly distributed pattern was observed for Early retirement, where financial difficulty, chronic conditions, maximum grip strength and general health showed the largest mean absolute SHAP values.

Late retirement was particularly influenced by maximum grip strength and gender, followed by employment type, while Working beyond NRA showed the strongest contribution from mobility limitations, followed by delayed recall, education level, gender and immediate recall.

The SHAP value distributions additionally provided information on the direction of these relationships.



Figure 10: SHAP distribution plot

For On-time retirement, lower numbers of chronic conditions were associated with positive SHAP contributions, whereas higher values shifted the prediction away from this category.

For Early retirement, chronic conditions showed the opposite pattern, with higher values contributing positively to the classification. Financial difficulty displayed a non-monotonic relationship with Early retirement, as levels 1 and 4 contributed positively while intermediate levels, particularly level 3, contributed negatively.

For Late retirement, lower maximum grip strength was associated with positive SHAP values, while higher grip strength progressively shifted predictions away from this category. Female gender also contributed positively to Late retirement predictions, whereas male gender contributed negatively. Self-employment and lower CASP values similarly tended to favour the Late classification.

For Working beyond NRA, mobility showed the clearest class-specific pattern. The absence of mobility limitations contributed strongly and positively to the prediction of Working beyond NRA, while the presence of limitations contributed negatively. Higher delayed-recall scores and higher educational attainment also tended to shift predictions towards this category.

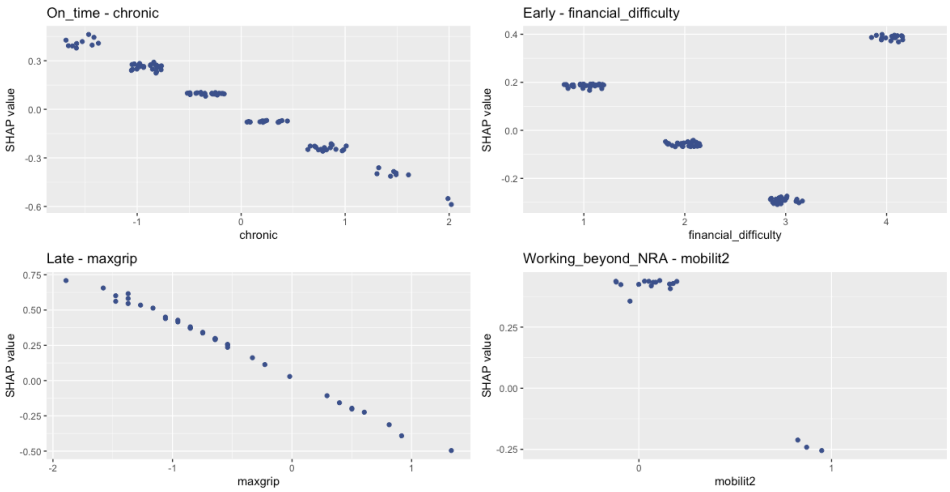


Figure 11: SHAP Dependence plot.

The dependence plots further highlighted the four strongest class-specific relationships. Chronic conditions showed a clear negative gradient for On-time retirement, while financial difficulty exhibited a non-linear pattern for Early retirement. Maximum grip strength showed a strong negative relationship with Late retirement, and mobility limitations clearly separated positive and negative contributions to Working Beyond NRA. Together, these results reinforce the existence of distinct predictive profiles across the four retirement timing trajectories.

SVM VS Multinomial Regression

The SHAP results were compared with the estimates from the multinomial logistic regression to assess whether the predictive patterns identified by the selected SVM pipeline were consistent with the associations obtained from the regression model.

Regarding early retirement, the two approaches showed strong agreement on health-related characteristics. In the multinomial model, each additional chronic condition was associated with higher odds of early retirement, while a self-assessed health status of Fairly/Poor was also associated with higher odds of early retirement. he SHAP analysis revealed the same directional pattern. Financial difficulties also showed strong consistency between the two approaches. In the multinomial model, 'With some difficulty' and 'Fairly easily' responses were associated with lower probabilities of early retirement compared with the reference category ('With great difficulty'), while the SHAP values for these categories were also negative.

With regard to late retirement, convergence was observed in terms of gender, type of employment and CASP. The multinomial model indicated lower probabilities of late retirement among men and higher probabilities among self-employed, while higher CASP scores were

associated with lower probabilities of late retirement in the class-imbalance-adjusted specifications. The SHAP analysis reflected the same directional relationships: female respondents, the self-employed and lower CASP scores positively contributed to the prediction of late retirement, while male gender and higher CASP scores moved the predictions away from this category. Additionally, the SHAP analysis identified that a lower maximum grip strength contributed to late retirement, a finding that proved to be less evident in the multinomial model.

Since the On time category was used as the reference outcome category in the multinomial regression, it does not have its own set of estimated coefficients. However, the SHAP patterns were generally consistent with the inverse of several associations identified for the non-reference categories. In particular, a lower number of chronic conditions and a self-assessed health status of Good/Better contributed positively to the prediction of retirement at the expected age.

With regard to Working Beyond NRA, mobility represented the most evident convergence between the two approaches. In the multinomial model, limitations on mobility were associated with substantially lower probabilities of Working Beyond NRA than of On-time. The SHAP analysis revealed that the absence of mobility constraints contributed strongly to the prediction of Working Beyond NRA. A higher level of Educational Level also showed a similar pattern across the approaches, with high Educational Level being associated with higher probabilities of continuing to work beyond NRA in the multinomial analysis and with strongly positive SHAP contributions in the SVM. The SHAP analysis also suggested positive contributions from higher scores on 'Immediate Recall' and from being male, although these patterns were not the strongestest results of the multinomial model.

Overall, the comparison indicates considerable convergence between the inferential and predictive approaches, particularly for chronic conditions and self-rated health in Early retirement, gender, self-employment and CASP in Late retirement, and mobility and education in Working beyond NRA. At the same time, SHAP revealed additional class-specific predictive patterns that were not equally prominent in the multinomial regression. Rather than being interpreted as contradictory, these differences reflect the complementary nature of the two approaches: multinomial regression identifies adjusted statistical associations, while SHAP captures how the selected machine learning model uses individual predictor values, including potentially non-linear and class-specific relationships, to generate predictions.

Chapter 5

Discussion

5.1 Main findings

To investigate the determinants of the three categories of retirement (Early, On-time, Late) and of Working Beyond NRA, a Multinomial Logistic Regression model was used, along with three machine learning models (Decision Tree, Random forest and Support Vector Machine). For these four models, 3 strategies were implemented to address class imbalance (baseline model, class-weights and oversampling).

Of the three machine learning models, the one that achieved the best compromise between Balanced Accuracy, Macro-F1 and Recall was the SVM class-weights.

When the results of the SVM and multinomial models were compared, the following conclusions were drawn: there is considerable alignment between the two approaches, particularly regarding to chronic conditions, self-assessed health and financial difficulties in the case of early retirement; gender, self-employment and CASP in the case of late retirement; and mobility and education in the case of Working Beyond NRA.

At the same time, SHAP revealed additional patterns that did not emerge in the regression. Rather than being interpreted as contradictory, these differences reflect the complementary nature of the two approaches.

Overall, the findings indicate that retirement timing cannot be characterised by a single set of determinants. Instead, the four trajectories appear to reflect different combinations of health, employment, financial, social and cognitive characteristics.

5.2 Literature comparison

Most of the findings are consistent with those described in the literature. Health and functional capacity proved to be particularly relevant, as the findings support the idea that poor health can contribute to a higher probability to exit the labour force earlier (Börsch-Supan et al. (2013); De Wind et al. (2014); Litwin and Tur-Sinai (2015)), while good functional capacity can support continuity of employment (Axelrad (2018)).

Financial difficulties also influenced the decision to retire. In this study, individuals in situations where the household is able to make ends meet with some difficulty or fairly easily are more likely to retire earlier than those who are experiencing great difficulty. In the literature, the impact of financial difficulty follows the same path (De Preter et al. (2013); De Wind et al. (2014); Litwin and Tur-Sinai (2015); Reeuwijk et al. (2013)), although it is also noted that financial difficulties are associated with late retirement (Axelrad (2018); De Wind et al. (2015); Markova (2020); Scharn et al. (2018)).

Higher education also contributed positively to Working Beyond NRA, consistent with the literature associating higher educational level with later retirement (Axelrad (2018)).

Being self-employed also appears as a positive factor to Late retirement and Working Beyond NRA, may reflect greater autonomy over work. However, this pattern should be interpreted with caution, as self-employment has not been consistently identified as a key determinant in the literature

Working conditions (Casas and Román (2024); Gerke (2013a); Reeuwijk et al. (2013)) and Social networks (Axelrad (2018); Litwin and Tur-Sinai (2015)) have been identified in the literature as relevant determinants of the timing of retirement. However, they did not emerge as significant determinants in this analysis. Although the results do not provide empirical support for these mechanisms, this should not be interpreted as evidence that they are irrelevant to retirement decisions.

5.3 Methodology Considerations

Multiple imputation introduced a methodological dimension. By analysing 20 imputed datasets, the study incorporated uncertainty associated with missing information rather than relying on a single completed dataset or complete-case analysis.

Class imbalance represented a major methodological challenge. Late retirement and Working Beyond NRA contained substantially fewer observations than Early and On-time retirement.

When these strategies were not implemented, the machine learning models obtained reasonable Accuracy, while failing to recognize minority outcomes. This was the reason why Balanced Accuracy, Macro F1 and class-specific Recall were considered. Despite this, these strategies should not be interpreted as equivalent to increasing the amount of independent information available for the minority classes. Results that were only important after weighting or oversampling were therefore interpreted more cautiously than associations consistently identified in the original-data specification.

The choice of combining multinomial logistic regression and machine learning was made to provide different but complementary forms of evidence. Although the AIC and BIC did not presented the best values for the M5 model, this was still the chosen for this analyze because it represented the theoretically complete model required to assess the different domains simultaneously, while assuming the cost of higher complexity.

5.4 Study Limitations

Several limitations should be considered when interpreting the findings.

First, the strongest limitation concerns the small number of observations in the Late and Working beyond NRA categories. With only a limited number of cases available for these categories, both models are less stable than for Early and On-time retirement.

Second, the outcome was constructed with a 6-month threshold, without a sensitivity analysis regarding the re-estimation of the models using other windows.

Third, the main analytical dataset was constructed using the most recent valid information available for each predictor. This strategy maximized sample coverage but does not guarantee that all predictor variables preceded the retirement transition. Consequently, temporal ordering is weaker for some variables, and the observed associations should not be interpreted causally.

Fourth, several predictors are based on self-reported SHARE information and may be affected by measurement error, recall error or subjective reporting differences.

Fifth, although the predictor set covers demographic, employment, financial, social-network, health and cognitive domains, retirement decisions may also depend on external factors not fully captured in the available data.

Finally, the analysis focuses on the Portuguese SHARE sample included in the analytical dataset. The findings should therefore not automatically be generalised to other national

contexts, particularly given differences in pension systems, labour-market institutions and statutory retirement arrangements across countries.

5.5 Future research

Future research could first examine the sensitivity of the present findings to the timing of the predictor variables being measured. Compare models with predictors measured before the retirement moment and with the most recent information would help determine whether the findings are sensitive to the temporal construction of the analytical dataset.

The analysis would also benefit from larger samples of individuals retiring Late or remaining in employment beyond the NRA. Combining additional SHARE waves or extending the analysis to other countries could increase the number of observations in these categories and allow more stable results.

Finally, longitudinal approaches that model retirement transitions over time could help to distinguish between predictors that precede retirement decisions and characteristics that change as a result of retirement itself.

Chapter 6

Conclusions

This study analysed the factors associated with the timing of retirement in Portugal, distinguishing between Early retirement, On-time retirement, Late retirement and Working Beyond the Normal Retirement Age. Based on data from the Survey of Health, Ageing and Retirement in Europe, statistical and machine learning models were implemented to assess how different characteristics were associated with these retirement paths.

The results suggest that the different retirement timings cannot be explained by a single group of determinants, given that each retirement timing presents different demographic, labour-market, financial, social and health characteristics.

The results of the combination between multinomial logistic regression model and support vector machine model provided complementary evidence on retirement timing. While the multinomial models quantified the associations, machine learning and SHAP analysis uncovered some class-specific patterns and relationships.

These findings should be interpreted with caution due to the small number of observations in the Late and Working Beyond NRA categories, which limits the stability and generalisability of the conclusions that were identified for these categories. The existence of missing values in some variables and the decisions that were made regarding on how to define the timing of retirement also limits the conclusions.

Even with these limitations, the study shows the value of analysing retirement timing using more than a simple retired versus non-retired distinction. Separating individuals according to when retirement occurred in relation to the NRA provides a more detailed view of retirement behaviour and also reveals how difficult some of these patterns are to predict. Future research could extend this analysis using larger samples and information collected closer to

the retirement transition.

Bibliography

Alnuaimi, A. F., & Albaldawi, T. H. (2024). An overview of machine learning classification techniques. *BIO Web of Conferences*, *97*, 00133. Retrieved from https://www.bio-conferences.org/10.1051/bioconf/20249700133 doi: 10.1051/bioconf/20249700133

Axelrad, H. (2018). Early Retirement and Late Retirement: Comparative Analysis of 20 European Countries. *International Journal of Sociology*, *48*(3), 231–250. Retrieved from https://www.tandfonline.com/doi/full/10.1080/00207659.2018.1483004 doi: 10.1080/00207659.2018.1483004

Banks, J. (2006). Retirement in the UK. *Oxford Review of Economic Policy*, *22*(1), 40–56. Retrieved from https://academic.oup.com/oxrep/article-lookup/doi/10.1093/oxrep/grj003 doi: 10.1093/oxrep/grj003

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. Retrieved from https://doi.org/10.1023/A:1010933404324 doi: 10.1023/A:1010933404324

Breiman, L., Friedman, J., Olshen, R., & Stones, C. J. (1984). *Classification and Regression Trees*. doi: 10.1201/9781315139470

Buuren, S. V., & Groothuis-Oudshoorn, K. (2011). **mice** : Multivariate Imputation by Chained Equations in *R*. *Journal of Statistical Software*, *45*(3). Retrieved from http://www.jstatsoft.org/v45/i03/ doi: 10.18637/jss.v045.i03

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., … Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, *42*(4), 992–1001. Retrieved from https://academic.oup.com/ije/article/42/4/992/657275 doi: 10.1093/ije/dyt088

Casas, P., & Román, C. (2024). The impact of artificial intelligence in the early retirement decision. *Empirica*, *51*(3), 583–618. Retrieved from https://link.springer.com/10.1007/s10663-024-09613-3 doi: 10.1007/s10663-024-09613-3

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. Retrieved from http://link.springer.com/10.1007/BF00994018 doi: 10.1007/BF00994018

Costa, J., Lima, M., Loureiro, H., & França, L. (2025). Tomada de decisão da reforma em portugal: a importância dos Preditores micro. *RIAGE - Revista Ibero-Americana de*

Gerontologia, *6*, 281–292. Retrieved from https://www.riagejournal.com/index.php/riage/article/view/280 doi: 10.61415/riage.280

Delattre, E., & Moussa, R. (2018). Early retirement decisions: Lessons from a dynamic structural modelling.

De Preter, H., Van Looy, D., & Mortelmans, D. (2013). Individual and institutional push and pull factors as predictors of retirement timing in Europe: A multilevel analysis. *Journal of Aging Studies*, *27*(4), 299–307. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S0890406513000273 doi: 10.1016/j.jaging.2013.06.003

De Wind, A., Geuskens, G. A., Ybema, J. F., Blatter, B. M., Burdorf, A., Bongers, P. M., & Van Der Beek, A. J. (2014). Health, job characteristics, skills, and social and financial factors in relation to early retirement – results from a longitudinal study in the Netherlands. *Scandinavian Journal of Work, Environment & Health*, *40*(2), 186–194. Retrieved from http://www.sjweh.fi/show_abstract.php?abstract_id=3393 doi: 10.5271/sjweh.3393

De Wind, A., Geuskens, G. A., Ybema, J. F., Bongers, P. M., & Van Der Beek, A. J. (2015). The role of ability, motivation, and opportunity to work in the transition from work to early retirement – testing and optimizing the Early Retirement Model. *Scandinavian Journal of Work, Environment & Health*, *41*(1), 24–35. Retrieved from http://www.sjweh.fi/show_abstract.php?abstract_id=3468 doi: 10.5271/sjweh.3468

Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*(1), 222. Retrieved from https://springerplus.springeropen.com/articles/10.1186/2193-1801-2-222 doi: 10.1186/2193-1801-2-222

Dorn, D., & Sousa-Poza, A. (2010). 'Voluntary' and 'involuntary' early retirement: an international analysis. *Applied Economics*, *42*(4), 427–438. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/00036840701663277 doi: 10.1080/00036840701663277

Ebbinghaus, B., & Hofäcker, D. (2013). Reversing Early Retirement in Advanced Welfare Economies: A Paradigm Shift to Overcome Push and Pull Factors. *Comparative Population Studies*, *38*(4). Retrieved from https://www.comparativepopulationstudies.de/index.php/CPoS/article/view/138 doi: 10.12765/CPoS-2013-24

European Commission. Directorate General for Economic and Financial Affairs. (2024). *2024 ageing report: economic & budgetary projections for the EU Member States (2022 2070)*. LU: Publications Office. Retrieved from https://data.europa.eu/doi/10.2765/022983 doi: 10.2765/022983

Fasang, A. E. (2010). Retirement: Institutional Pathways and Individual Trajectories in Britain and Germany. *Sociological Research Online*, *15*(2), 1–16. Retrieved from https://journals.sagepub.com/doi/10.5153/sro.2110 doi: 10.5153/sro.2110

Gerke, O. (2013a). Determinants of early retirement in Denmark. An empirical investigation using SHARE data.

Gerke, O. (2013b). Gerke, Oke & Lauridsen, Jørgen T., 2013. "Determinants of early retirement in Denmark. An empirical investigation using SHARE data," Discussion Papers on Economics 4/2013, University of Southern Denmark, Department of Economics.

Grandini, M., Bagli, E., & Visani, G. (2020). *Metrics for Multi-Class Classification: an Overview*. arXiv. Retrieved from http://arxiv.org/abs/2008.05756 doi: 10.48550/arXiv.2008.05756

Hashimoto, E. M., Ortega, E. M. M., Cordeiro, G. M., Suzuki, A. K., & Kattan, M. W. (2020). The multinomial logistic regression model for predicting the discharge status after liver transplantation: estimation and diagnostics analysis. *Journal of Applied Statistics*, *47*(12), 2159–2177. Retrieved from https://www.tandfonline.com/doi/full/10.1080/02664763.2019.1706725 doi: 10.1080/02664763.2019.1706725

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.)*. doi: 10.1007/978-0-387-84858-7

Koçak, C., & Burdorf, A. (2025). The Relationship of Limitations in Daily Living Activities and Other Health-Related Factors with Early Retirement. *Journal of Prevention*, *46*(3), 339–353. Retrieved from https://link.springer.com/10.1007/s10935-024-00818-7 doi: 10.1007/s10935-024-00818-7

Leinonen, T., Boets, I., Pletea, E., Vandenbroeck, S., Mehlum, I. S., Hasselhorn, H. M., & De Wind, A. (2022). A conceptual framework addressing the complex labour market dynamics of the work-to-retirement process. *European Journal of Ageing*, *19*(4), 1631–1637. Retrieved from https://link.springer.com/10.1007/s10433-022-00704-3 doi: 10.1007/s10433-022-00704-3

Lewis, R. J. (2000). An Introduction to Classification and Regression Tree (CART) Analysis.

Litwin, H., & Tur-Sinai, A. (2015). The Role of the Social Network in Early Retirement Among Older Europeans. *Work, Aging and Retirement*, *1*(4), 340–349. Retrieved from https://academic.oup.com/workar/article-lookup/doi/10.1093/workar/wav013 doi: 10.1093/workar/wav013

Markova, E. (2020). WHY TO GO FOR EARLY RETIREMENT? DETERMINANTS FOR EARLY EXIT FROM THE LABOUR MARKET: THE EVIDENCE FROM BULGARIA. *Balkan Social Science Review*, *16*, 299–314. Retrieved from https://js.ugd.edu.mk/index.php/BSSR/article/view/4010 doi: 10.46763/BSSR2016299m

Mullainathan, S., & Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, *31*(2), 87–106. Retrieved from https://pubs.aeaweb.org/doi/10.1257/jep.31.2.87 doi: 10.1257/jep.31.2.87

OECD. (2025). *Pensions at a Glance 2025: OECD and G20 Indicators*. OECD Publishing. Retrieved from https://www.oecd.org/en/publications/pensions-at-a-glance-2025_e40274c1-en.html doi: 10.1787/e40274c1-en

Reeuwijk, K. G., De Wind, A., Westerman, M. J., Ybema, J. F., Van Der Beek, A. J., & Geuskens, G. A. (2013). 'All those things together made me retire': qualita-

tive study on early retirement among Dutch employees. *BMC Public Health*, *13*(1), 516. Retrieved from https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-13-516 doi: 10.1186/1471-2458-13-516

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York Chichester Brisbane Toronto Singapore: John Wiley & Sons.

Scharn, M., Sewdas, R., Boot, C. R. L., Huisman, M., Lindeboom, M., & Van Der Beek, A. J. (2018). Domains and determinants of retirement timing: A systematic review of longitudinal studies. *BMC Public Health*, *18*(1), 1083. Retrieved from https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-018-5983-7 doi: 10.1186/s12889-018-5983-7

Topa, G., Depolo, M., & Alcover, C.-M. (2018). Early Retirement: A Meta-Analysis of Its Antecedent and Subsequent Correlates. *Frontiers in Psychology*, *8*, 2157. Retrieved from http://journal.frontiersin.org/article/10.3389/fpsyg.2017.02157/full doi: 10.3389/fpsyg.2017.02157

Van Buuren, S., Brand, J. P., Groothuis-Oudshoorn, C. G., & Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation*, *76*(12), 1049–1064. Retrieved from https://www.tandfonline.com/doi/full/10.1080/10629360600810434 doi: 10.1080/10629360600810434

Wang, M., & Shultz, K. S. (2010). Employee Retirement: A Review and Recommendations for Future Investigation. *Journal of Management*, *36*(1), 172–206. Retrieved from https://journals.sagepub.com/doi/10.1177/0149206309347957 doi: 10.1177/0149206309347957

Zhang, Z. (2015). Multiple imputation with multivariate imputation by chained equation (MICE) package. *Annals of Translational Medicine*, *4*(2), 30. Retrieved from http://www.atmjournal.org/article/view/8847/9618 doi: 10.3978/j.issn.2305-5839.2015.12.63

Zhu, X. (2014). Comparison of Four Methods for Handling Missing Data in Longitudinal Data Analysis through a Simulation Study. *Open Journal of Statistics*, *04*(11), 933–944. Retrieved from https://www.scirp.org/journal/paperinformation?paperid=52855 doi: 10.4236/ojs.2014.411088