Models of Deterministic and Stochastic Comparison

Two Studies in Applied Operations Research

by

Kiefer Joe Burgess

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Naoki Masuda Professor, Dept. of Mathematics, The State University of New York, University at Buffalo
Supervisors:	Stanko Dimitrov Professor, Dept. of Management Sciences, University of Waterloo Qi-Ming He Professor, Dept. of Management Sciences, University of Waterloo
Internal Members:	Jim Bookbinder Professor, Dept. of Management Sciences, University of Waterloo Ken McKay Professor, Dept. of Management Sciences, University of Waterloo

Internal-External Member: Ben Feng

Assistant Professor, Dept. of Statistics and Actuarial Science,

University of Waterloo

Other Member: Xuan Zhao

Professor, Lazaridis School of Business and Economics,

Wilfrid Laurier University

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This dissertation includes two essays on applications of management science methods to modelling service systems and developing novel improvements to sports team ranking systems. The first essay proposes a novel approach to modelling changes in business procedures that have neither explicitly positive nor explicitly negative effects on operational performance, but are changes to operating rules; we call these procedure changes Operational Protocol Modifications (OPMs). Our approach is to model these OPMs via distributional censoring. Using the scenario of a technical support employee at a SaaS firm, we model changes in OPMs as censoring effects on the distributions of both service quality and service time. We demonstrate the nonlinear effects OPMs can have on the optimal service contract and the employer's (principal's) expected utility in hiring the technical support employee (agent), under certain distributional assumptions. This modelling approach arms operations management analysts with a new tool to better capture the impact of OPMs and their non-linear impacts on operational performance.

The second essay proposes a number of additions to both static and dynamic network ranking models for professional soccer teams. We introduce ways to incorporate relevant home/away game status and goal difference information. Further, we introduce a collection of methods to measure the competitive similarity between teams, which we then integrate into the ranking systems. We demonstrate, using a large collection of data on five of the top European professional soccer leagues, that our methods produce superior empirical performance when compared to comparable approaches. Importantly, our work is the first to integrate the competitive similarity notion directly into network ranking models, providing the first direct link between two related bodies of literature.

Acknowledgements

First and foremost, I would like to thank Professors Stanko Dimitrov and Qi-Ming He for their dedicated supervision throughout my PhD program. They demonstrated superb patience in dealing with my transition into the management sciences and taught me how to both research and write for academic purposes. I am grateful for the insight they shared with me in crafting this thesis and for the example they set as scholars, which was crucial in my development.

Second, I am grateful to Professor Naoki Masuda from University at Buffalo of The State University of New York for insightfully evaluating my dissertation as an external examiner. I would like to express my gratitude to my advisory committee members Professors Jim Bookbinder, Ken McKay, Xuan Zhao and Assistant Professor Ben Feng for their comments and encouragement throughout the process.

I am eternally grateful for the love and support I receive from my parents, Ken and Linda, whose many trips to Waterloo always brightened my spirits and emboldened me to press on. I am grateful for receiving plenty of love, support, and prayers from my grandmother, Catherine, who encourages me every day. I am grateful for my brother, Roman, for the love and support he shows me, and who always provides me with inspiration to surpass my limits and reach my potential.

Finally, I'd be remiss if I did not give thanks to Matt and Phil, my two best friends. Our many chats about student and research life were pivotal in bringing camaraderie and levity to this endeavour.

Dedication

To my wonderful parents, ***Ken*** & ***Linda***,
my lovely grandmother, ***Catherine***,
and my incredible brother, ***Roman***.

Table of Contents

Examining Committee **ii**

Author's Declaration **iv**

Abstract **v**

Acknowledgements **vi**

Dedication **vii**

List of Figures **xv**

List of Tables **xxi**

1 Introduction **1**

2 Distributional Censoring and Optimal Service Contracts **4**

2.1 Introduction 4

2.2 Theoretical Background 6

2.2.1 Game Theory 6

2.2.2 Principal-Agent Modelling and Mechanism Design 8

2.2.3 Distributional Censoring and Stochastic Dominance 12

2.3 Literature Review 13

2.4 The Model of Interest 16

2.4.1 Dynamics 16

2.4.2 Parameters and Distributions 18

2.4.3 Wage and Utility Functions 20

2.4.4 Mathematical Programs 24

2.5 Optimization Analysis 28

2.6 Business Process Modifications 33

2.6.1 Operational Protocol Modifications (OPMs) 34

2.6.2 Training and Knowledge Enhancements 42

2.6.3 Changing Business Context 48

2.7 Practical Considerations 51

2.8 Conclusion 53

3 Approaches for Incorporating Additional Information into Network Ranking Models 55

3.1 Introduction 55

3.2 Theoretical Background 58
3.2.1 Ranking Primer 58
3.2.2 League Points Approach 59
3.2.3 Elo Approach 59
3.2.4 Paired Comparison Approach 61
3.2.5 Network Model Approach 63
3.3 Literature Review 75
3.4 Data Description 78
3.5 New Network Ranking Methods 80
3.5.1 Home/Away Effects 81
3.5.2 Goal/Difference Effects 83
3.5.3 Ranking Crossings as Competition 84
3.5.4 Direct Similarity Approach 86
3.5.5 Mean-Based Direct Similarity Approach 90
3.5.6 Matched Set Similarity Approach 92
3.5.7 Mean-Based Matched Set Similarity Approach 96
3.5.8 Unmatched Set Similarity Approach 98
3.5.9 Mean-Based Unmatched Set Similarity Approach 101
3.5.10 Combination Methods 102
3.5.11 Similarity Approach Comparison 103

3.6 Performance Results and Analysis 104
3.6.1 Performance Metric 105
3.6.2 Ranking Method Performance Comparison 107
3.6.3 Rank Correlation Analysis 120
3.6.4 Sensitivity Analysis 124
3.6.5 Betting Odds Comparison 128
3.7 Discussion 130
3.7.1 Methodological Considerations 130
3.7.2 Managerial Implications 132
3.8 Conclusion 133

4 Conclusion **135**

References **137**

APPENDICES **144**

A Contract Theory Proofs **145**
A.1 Solutions of the Optimization Problem (POPI) 145
A.2 Candidate Comparison Conditions 156
A.3 Proofs of Candidate Comparison Conditions 157
A.4 Operational Protocol Modification Proofs 161
A.4.1 $\Delta_{\hat{S}}(x)$ 161

A.4.2 $\Delta_{\hat{T}}(x)$ 162

A.4.3 Property Derivations 163

B Soccer Ranking Appendices **176**

B.1 Notation 176

B.2 Combination Method Explanations 178

B.2.1 Home/Away + Goal Difference 179

B.2.2 Static Home/Away + Direct Similarity 179

B.2.3 Dynamic Home/Away + Direct Similarity 179

B.2.4 Static Home/Away + Matched Set Similarity 179

B.2.5 Dynamic Home/Away + Matched Set Similarity 180

B.2.6 Static Home/Away + Unmatched Set Similarity 180

B.2.7 Dynamic Home/Away + Unmatched Set Similarity 180

B.2.8 Static Home/Away + Mean-Based Direct Similarity 180

B.2.9 Dynamic Home/Away + Mean-Based Direct Similarity 181

B.2.10 Static Home/Away + Mean-Based Matched Set Similarity 181

B.2.11 Dynamic Home/Away + Mean-Based Matched Set Similarity . . . 181

B.2.12 Static Home/Away + Mean-Based Unmatched Set Similarity 181

B.2.13 Dynamic Home/Away + Mean-Based Unmatched Set Similarity . . 182

B.2.14 Static Goal Difference + Direct Similarity 182

B.2.15 Dynamic Goal Difference + Direct Similarity 182

B.2.16 Static Goal Difference + Matched Set Similarity 182
B.2.17 Dynamic Goal Difference + Matched Set Similarity 183
B.2.18 Static Goal Difference + Unmatched Set Similarity 183
B.2.19 Dynamic Goal Difference + Unmatched Set Similarity 183
B.2.20 Static Goal Difference + Mean-Based Direct Similarity 183
B.2.21 Dynamic Goal Difference + Mean-Based Direct Similarity 184
B.2.22 Static Goal Difference + Mean-Based Matched Set Similarity 184
B.2.23 Dynamic Goal Difference + Mean-Based Matched Set Similarity . . 184
B.2.24 Static Goal Difference + Mean-Based Unmatched Set Similarity . . 184
B.2.25 Dynamic Goal Difference + Mean-Based Unmatched Set Similarity 185
B.2.26 Static Home/Away + Goal Difference + Direct Similarity 185
B.2.27 Dynamic Home/Away + Goal Difference + Direct Similarity 185
B.2.28 Static Home/Away + Goal Difference + Matched Set Similarity . . 186
B.2.29 Dynamic Home/Away + Goal Difference + Matched Set Similarity 186
B.2.30 Static Home/Away + Goal Difference + Unmatched Set Similarity 186
B.2.31 Dynamic Home/Away + Goal Difference + Unmatched Set Similarity 186
B.2.32 Static Home/Away + Goal Difference + Mean-Based Direct Similarity 187
B.2.33 Dynamic Home/Away + Goal Difference + Mean-Based Direct Similarity 187
B.2.34 Static Home/Away + Goal Difference + Mean-Based Matched Set Similarity 187

B.2.35 Dynamic Home/Away + Goal Difference + Mean-Based Matched Set Similarity . . . 188
B.2.36 Static Home/Away + Goal Difference + Mean-Based Unmatched Set Similarity . . . 188
B.2.37 Dynamic Home/Away + Goal Difference + Mean-Based Unmatched Set Similarity . . . 188
B.3 Additional Performance Figures . . . 188
B.3.1 Static Approaches . . . 189
B.3.2 Dynamic Approaches . . . 213
B.4 Attempts to Beat Betting Odds . . . 238
B.4.1 Tie Prediction Within k Ranks . . . 239
B.4.2 Tie Prediction Within k Total Score Standard Deviations . . . 239
B.4.3 Randomized Result Prediction Within k Total Score Standard Deviations . . . 240
B.4.4 Similarity-Based Tie Prediction . . . 240
B.4.5 Machine Learning Approaches to Predicting Ties, Upsets, and Non-Upsets . . . 241

List of Figures

2.1 Timeline of Events 18

2.2 U^P for various values of τ_L and τ_H 36

2.3 U^P for various values of τ_L and τ_H 39

2.4 Optimal U^P as a function of k 49

2.5 Optimal x as a function of k 50

3.1 Basic Network Model Structure 64

3.2 Average Predictive Accuracy as a function of k_A and k_D 125

3.3 Average Predictive Accuracy as a function of k_A and k_D; $\alpha = 0.2$, $\beta = 0.1$. 126

3.4 Average Predictive Accuracy as a function of α and β; $k_A = 2$ and $k_D = 1.6$ 127

3.5 Per Season Betting Odds Predictive Accuracy 129

B.1 English Premier League 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 189

B.2 English Premier League 4-Year Rolling Predictive Accuracy: Static Similarity Approaches 190

B.3 English Premier League 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches 191

B.4 English Premier League 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches 192

B.5 English Premier League 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches 193

B.6 Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 194

B.7 Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Similarity Approaches 195

B.8 Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches 196

B.9 Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches 197

B.10 Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches 198

B.11 German Bundesliga 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 199

B.12 German Bundesliga 4-Year Rolling Predictive Accuracy: Static Similarity Approaches 200

B.13 German Bundesliga 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches 201

B.14 German Bundesliga 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches 202

B.15 German Bundesliga 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches 203

B.16 Italian Serie A 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 204

B.17 Italian Serie A 4-Year Rolling Predictive Accuracy: Static Similarity Approaches 205

B.18 Italian Serie A 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches 206

B.19 Italian Serie A 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches 207

B.20 Italian Serie A 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches 208

B.21 French Ligue 1 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 209

B.22 French Ligue 1 4-Year Rolling Predictive Accuracy: Static Similarity Approaches 210

B.23 French Ligue 1 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches 211

B.24 French Ligue 1 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches 212

B.25 French Ligue 1 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches 213

B.26 English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 214

B.27 English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches 215

B.28 English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches 216

B.29 English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches 217

B.30 English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches 218

B.31 Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 219

B.32 Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches 220

B.33 Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches 221

B.34 Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches 222

B.35 Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches 223

B.36 German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 224

B.37 German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches 225

B.38 German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches 226

B.39 German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches 227

B.40 German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches 228

B.41 Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 229

B.42 Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches 230

B.43 Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches 231

B.44 Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches 232

B.45 Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches 233

B.46 French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches 234

B.47 French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches 235

B.48 French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches 236

B.49 French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches 237

B.50 French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches 238

List of Tables

2.1 Wage contract decision variables 21

2.2 Example Distributions 32

2.3 Example Parameters 32

2.4 Example Candidate Values 33

2.5 Post-operational-change optimal solutions for different τ_L 37

2.6 Post-operational-change optimal solutions for different τ_H 37

2.7 Post-operational-change optimal solutions for different Υ_L 40

2.8 Post-operational-change optimal solutions for different Υ_H 40

2.9 Post-Training Service Quality Distribution 44

2.10 Post-training optimal solutions for different service quality $\tilde{S}$ 45

2.11 Post-Training Service Time Distribution 46

2.12 Post-training optimal solutions for different service time $\tilde{T}$ 47

3.1 Per-League Total Number of Games and Goals 79

3.2 Per-League Game and Goal Percentages 79

3.3 Existing Approaches, Home/Away, and Goal Difference Predictive Accuracy Average of 4-season Rolling Windows 108

3.4 Original Static Network vs Similarity Approaches Predictive Accuracy Average of 4-season Rolling Windows 109

3.5 Original Static Network + Home/Away vs Similarity Approaches + Home/Away Predictive Accuracy Average of 4-season Rolling Windows 111

3.6 Original Static Network + Goal Difference vs Similarity Approaches + Goal Difference Predictive Accuracy Average of 4-season Rolling Windows . . . 112

3.7 Original Static Network + Home/Away + Goal Difference vs Similarity Approaches + Home/Away + Goal Difference Predictive Accuracy Average of 4-season Rolling Windows 113

3.8 Existing Approaches, Home/Away, and Goal Difference Predictive Accuracy Average of 4-season Rolling Windows 115

3.9 Original Dynamic Network vs Similarity Approaches Predictive Accuracy Average of 4-season Rolling Windows 116

3.10 Original Dynamic Network + Home/Away vs Similarity Approaches + Home/Away Predictive Accuracy Average of 4-season Rolling Windows 117

3.11 Original Dynamic Network + Goal Difference vs Similarity Approaches + Goal Difference Predictive Accuracy Average of 4-season Rolling Windows 118

3.12 Original Dynamic Network + Home/Away + Goal Difference vs Similarity Approaches + Home/Away + Goal Difference Predictive Accuracy Average of 4-season Rolling Windows 119

3.13 Original Dynamic Network vs Similarity Approaches Spearman Rank Correlation Average of 4-season Rolling Windows 121

3.14 Original Dynamic Network + Home/Away vs Similarity Approaches + Home/Away Spearman Rank Correlation Average of 4-season Rolling Windows 122

3.15 Original Dynamic Network + Goal Difference vs Similarity Approaches + Goal Difference Spearman Rank Correlation Average of 4-season Rolling Windows 122

3.16 Original Dynamic Network + Home/Away + Goal Difference vs Similarity Approaches + Home/Away + Goal Difference Spearman Rank Correlation Average of 4-season Rolling Windows 123

3.17 Average Betting Odds Predictive Accuracy Per Country 129

B.1 Notation for Network Ranking Models 177

Chapter 1

Introduction

Underpinning the management science discipline is an appreciation for the rigourous application of mathematical models and methods that suit the problem at hand. The problem domains addressed by management scientists are incredibly varied, ranging from highly technical mathematical work to deploying real-time analytics systems in a variety of operations contexts. As such, the methods deployed by a management scientist may be diverse. This dissertation demonstrates this methodological diversity by using tools from game theory/mechanism design, optimization, probability theory, statistics, and network science in its two essays.

The focus of both essays is the introduction of novel methodologies. The first essay, which is independent from the second, introduces distributional censoring as an approach to modelling certain changes in business operation procedures and explores its use in a mechanism design context. The second essay presents a collection of techniques for incorporating additional information into network ranking models for professional soccer teams. We present an overview of each essay:

- **Essay One**: *Distributional Censoring and Optimal Service Contracts*

  We consider a customer service scenario wherein an employer seeks to hire an employee to complete a customer service request. After establishing the game dynamics and solving for the optimal set of contracts, we consider how these optimal contracts change under various changes in operating conditions. Our main contribution is in modelling changes in business operating rules or procedures, which we call Operational Protocol Modifications (OPMs), as instances of distributional censoring on the distributions of service quality and service time. We round out our analysis by modelling training and knowledge enhancement initiatives as first-order stochastic dominant shifts in the distributions of service quality and service time.

- **Essay Two**: *Approaches for Incorporating Additional Information into Network Ranking Models*

  Network ranking models are an alternative to statistically-driven approaches that use the notion of direct and indirect wins and losses to rank sports players and teams. We introduce a collection of methods to incorporate additional information into these models. As a smaller contribution, we introduce ways to incorporate both home/away status and goal difference into network ranking models. Our main contribution is the introduction of techniques that incorporate information from the time series of rankings to develop similarity metrics between teams, which we use to tune the ranking models. We test our approaches on a large sample of data from five of the top European soccer leagues and demonstrate their superior empirical performance.

While the two essays of this dissertation cover a variety of domains and methodologies, at their core, they are related to something fundamental to management science: comparison. A key component of management science is to compare two or more objects, as this is the basis for implementing operational improvement decisions and policies. For example, we compare the rewards earned by making various decisions whenever we find the optimal solution to an optimization problem. Each essay of this dissertation explores comparison in a different way. The first essay uses distributional censoring and stochastic dominance to compare how the service system performs under different OPMs and training levels. The second essay is entirely concerned with comparing team performance to better assess which team has the highest capability. Both works highlight the importance methodology plays in making comparisons and the roles these comparisons play in arriving at optimal system performance.

The remainder of this dissertation proceeds as follows. Essay One is found in Chapter 2. Essay Two is found in Chapter 3. We conclude the dissertation in Chapter 4.

Chapter 2

Distributional Censoring and Optimal Service Contracts

2.1 Introduction

Given the numerous upheavals brought about due to the COVID-19 pandemic, the old adage of change being the only constant may resonate now more than ever. Change comes in many forms, and it may be experienced passively or implemented by management. The latter is the focus of this work. As management scientists, we focus on devising ways to improve businesses, be these improvements through operations, personnel, or technology. We propose adjustments to standard operating procedures, often with the goal of maximizing some objective subject to constrained resources. However, these proposals may have non-obvious impacts via the organization's employment structures. Specifically, there is an important interplay between operating procedures and a firm's incentive contracts.

Before proceeding, we clarify our terminology around changes to business procedures.

We use the term Operational Protocol Modifications (OPMs) to refer to any proposed or implemented change in standard operating procedures or rules at a firm. Given that these operational changes are rule-based, it is non-obvious if they will benefit or harm firm performance. This is in contrast to employee training and knowledge enhancement which, *ceteris paribus*, is almost universally assumed to be a performance enhancing initiative.

The motivating example of our work stems from an industry partnership with a Software-as-a-Service (SaaS) firm. We use insights gained from studying their operations as the basis for structuring our mathematical model and mechanism design problem. Namely, we consider a contracting game in which an employer (principal) seeks to hire an employee (agent) to complete the technical support requests of the firm's clients. We assume there is a degree of asymmetric information in this game; namely, the agent observes the difficulty of the support request while the principal can only discover this information after a costly audit. We propose an optimal contract that maximizes the principal's utility with respect to the agent's generated service quality and service time; the contract incentivizes the agent to exert the optimal effort level for both the agent and the principal. This mechanism design problem then serves as the foundation for our modelling and methodological contributions.

The contributions of our work are twofold. Firstly, we introduce distributional censoring as a rigourous and intuitive way to model OPMs. Censoring naturally captures how these rule changes restrict the range of outcomes on service outputs, like how providing a service agent with a checklist or script can impose a new minimum on service time. Secondly, we connect OPMs and censoring to mechanism design via our aforementioned contracting game. We show that, under certain distributional assumptions, the censoring brought about by OPMs yields nonlinear responses in the principal's optimal utility. These results imply that managers should think carefully about the expected impact of their OPMs; the incentive structures of the organization will ultimately be a key determinant in their efficacy.

To round out our analysis, we explore the effects of training and knowledge enhancement initiatives and conduct a brief sensitivity analysis of our model's key parameters. We model training and knowledge enhancement as instances of first-order stochastic dominant shifts in the distributions of service quality and service time. We show that these initiatives, when modelled this way, yield monotonic improvements in the principal's optimal utility. Our subsequent sensitivity analysis highlights the importance of the agent's cost of effort and suggests that the principal may benefit from any efforts that, over time, reduce this cost.

The rest of the paper proceeds as follows. Section 2.3 discusses related work and specifies our contributions to the literature. Section 2.4 details the service contract game, our model parameters, wage and utility functions, and the mathematical program we use to find the optimal contract. Section 2.5 presents our optimal solution and a numerical example highlighting an optimal contract. Section 2.6 presents our analysis on the impacts of OPMs, training and knowledge enhancements, and our sensitivity analysis. Section 2.7 discusses factors that may arise in implementing our contracts and OPMs in practice. Section 2.8 concludes.

2.2 Theoretical Background

2.2.1 Game Theory

The content in this section and Section 2.2.2 roughly follow explanations from Rasmusen (2006), which contains greater detail. Game theory is the study of interactions amongst strategic agents (players). We concern ourselves with game theory that uses mathematical models to describe and analyse these interactions. The key difference between decision theory and game theory is that game-theoretic models treat all players as strategic; each chooses a strategy based on the available information and the anticipated responses by other players. This is typically framed as the various agents each solving an optimization or optimal control problem which inherently depends on the strategies of the other agents in some way.

We now formalize these notions and introduce some necessary terminology. The key elements to any game under consideration are the players, actions, payoffs, and information. Collectively, these elements form the rules of the game. We describe each in turn:

- **Players**: Players of the game are the strategic agents interacting in the game.
- **Actions**: These are the sets of actions each agent can take as they proceed through the game. Note: the collection of actions an agent can take is called the *action set*. Action sets may differ across players and across time. Actions are distinct from strategies. Strategies are information-contingent sets of actions that dictate which action to take based on the observed information. For example, an action may be something like "take an umbrella to work", while a strategy is "if the weather forecast calls for rain, take an umbrella to work, otherwise do not take an umbrella". Note: a strategy must cover all states of information, hence the use of the "otherwise" in the above example. A strategy profile is a vector where each component contains a strategy for the corresponding player. For example, in an n-player game, we can write the strategy vector $\mathbf{s}$, where component s_i is the strategy for player i.
- **Payoffs**: Payoffs are the results of the game. Once all players choose their strategies and the game proceeds, the results are realized and each player receives their payoff. Payoffs can be positive or negative, where negative payoffs represent a loss (typically

in terms of utility, though other explicit units can be used). It is important to mention that payoffs are tied to strategies, and a key component of game-theoretic modelling is to determine expected payoffs from a strategy. This is where the optimization or control comes into play, wherein a player attempts to maximize their expected payoff.

- **Information**: Information is what each player knows about the game as well as the state of the world, including the actions of other players and payoffs. Crucially, each player has their own information set which may be different from other players' information sets. Further, these information sets may change during the course of the game, and information may or may not be relevant to the decision at hand.

Beyond the players of interest in a game, we sometimes define a pseudo-player whose actions serve a mechanical purpose in a game. Most notably, we often declare Nature as a pseudo-player who takes random actions at certain points in a game.

Often, the goal of a game-theoretic analysis is to investigate some form of equilibrium of the game. While there are a variety of equilibrium concepts, in general, an equilibrium consists of a strategy profile $\mathbf{s}^*$, where each component is the optimal strategy of the corresponding player. The game theorist is interested in determining what equilibria will arise based on the rules of the game. If we have modelled our game after a real-world scenario, these equilibria can then provide insights into that scenario.

2.2.2 Principal-Agent Modelling and Mechanism Design

Principal-agent modelling is a tool used to model game-theoretic scenarios. Specifically, it is often used to pose mechanism design problems. Before discussing the principal agent model, we briefly introduce mechanism design. Mechanism design is perhaps best thought of as the inverse of classical game theory. In game theory, we treat the rules of the game as fixed and determine and analyze the equilibria that arise from those rules. In mechanism design, we treat the rules of the game as variable, wherein we propose rules to optimize some sort of mechanism-designer-relevant objective.

To elaborate, it helps if we consider a specific scenario. Consider an employer who seeks to hire an employee to complete a task for the business. This employer drafts up and offers a contract (or menu of contracts) to the potential employee. If the potential employee deems the contract beneficial, they accept the contract and perform the task; otherwise the potential employee rejects the contract and the game ends. In this scenario, the employer is the principal and the potential employee is the agent. In general, we call the individual proposing the contract and offering compensation via the contract the principal; the individual who is in the position to accept the contract and get paid for their performance is the agent. The contract is the mechanism.

Information plays a crucial role in mechanism design. The following classifications of information are useful:

- **Perfect**: Information is perfect if each information set is a singleton, meaning that the player knows which state of the world they are in and what moves the other players have taken before they take their action.

- **Certain**: Information is certain when Nature does not move after any player moves. In essence, this means that the consequences of a player's actions are not subject to randomness if information is certain.

- **Symmetric**: Information is symmetric when no player has information that differs from other players when they move. As one may imagine, scenarios of asymmetric information are common and often the focus of mechanism design problems.

- **Complete**: Information is complete when Nature does not move first, or Nature's initial move is observed by every player.

Mechanism design is often concerned with games wherein information either is or becomes asymmetric at some point. Two common scenarios are those of moral hazard and adverse selection:

- **Moral Hazard**: These games feature symmetric information at the time a contract is offered and accepted, but the agent then takes a hidden action after the contract is accepted. Information is complete.

- **Adverse Selection**: These games feature incomplete information as Nature chooses the agent's type (for example, a high or low ability worker) as the first move of the game, which the principal cannot observe. The principal then offers a contract.

As one can see by the definitions of moral hazard and adverse selection, these scenarios are not uncommon in the workplace and other contract-driven environments. Using mechanism design, we attempt to elicit the agent's private information (their hidden actions or type) through the contract's features; this is typically called auditing. It is important to note that we may never be able to observe certain components of the game. For example, it may be impossible to observe the effort an agent exerts in the game. We may only be able to observe, say, an output of the agent's effort that is correlated with effort, but the mapping from effort space to output space may not be a bijection.

With some terminology introduced, we now introduce the typical mathematical form these models take on. Note: this is by no means a comprehensive treatise on all forms such models can take, but it parallels the form we use later. Consider the general scenario of a principal hiring an agent to complete some task. Nature is a pseudo-player in this game.

Let S be a random variable for the state of the world, the probability distribution of which is known by both the principal and the agent. Suppose only the agent gets to view the state of the world after the contract is accepted. Let $w(\rho)$ be a wage contract, which is a function of the output of the agent completing the task, $\rho(x)$, where ρ is a function of the agent's effort x. The principal offers the wage contract to the agent. If the agent accepts the contract, the agent proceeds to complete the task by exerting effort x; if the agent rejects the contract, the game ends. Let the principal's utility function be $U(\rho, w(\rho))$; we see that it is a function of the agent's output and the wage paid to the agent. Typically, the principal's utility increases as a function of ρ and decreases as a function of w. Let the agent's utility function be $V(w(\rho), x)$. As one may expect, the agent's utility typically increases in terms of w and decreases in terms of x. Given that neither the principal nor the agent get to view the state of the world prior to agreeing to the contract, we use expected utilities in the decision-making process. Assume both the principal and the agent are risk-neutral (this means they want to maximize the expected utility). To participate in the contract, the agent most likely has a reserve utility $\bar{V}$, which the expected utility of the contract must meet or exceed.

We now present the standard optimization model facing the principal offering the contract, which emulates the formulation in Chapter 7 of Rasmusen (2006):

$$\max_{w} \mathrm{E}[U(\rho(x^*), w(\rho(x^*)))] \tag{2.1}$$

$$\text{subject to: } x^* = \operatorname*{argmax}_{x} \mathbf{E}[V(w(\rho), x)]; \tag{2.2}$$

$$\mathbf{E}[V(w(\rho(x^*)), x)] \geq \bar{V}; \tag{2.3}$$

$$0 \leq x \leq 1. \tag{2.4}$$

Examining (2.1)–(2.4), we see the following. The principal seeks to maximize their expected utility by choosing the contract w. In (2.2), we see that one of the constraints of the principal's problem is that the agent maximizes their own expected utility by exerting effort, given the contract the principal selects. Constraint (2.3) indicates that the agent's expected utility must meet or exceed their reserve utility $\bar{V}$. Finally, (2.4) indicates that the agent's effort is bounded between 0 and 1. Note: a more specific model will likely have sign restrictions on the coefficients of terms included in the wage contract. In solving the optimization problem (if a solution exists) the principal finds the optimal contract or menu of contracts that maximize their expected utility.

2.2.3 Distributional Censoring and Stochastic Dominance

Underpinning our contributions is the notion of censoring a probability distribution. Consider a random variable X with distribution F with infinite support on the real line $\mathbb{R}$. This random variable, upon realization, can take on any value in $\mathbb{R}$. Now, suppose we define a new random variable $\hat{X}$ as follows:

$$\hat{X} = \max\{0, X\}. \tag{2.5}$$

In (2.5), we see that if X realizes a value above or equal to zero, $\hat{X} = X$. However, if X realizes a value less than zero, $\hat{X} > X$ as $\hat{X}$ is forced to take the value 0. In this way, some potential realizations of X are not available to $\hat{X}$; we call this censoring. It is possible to censor the other tail of the distribution as well, so the upperbound becomes constrained. Indeed, the most general case is that both tails of the distribution are censored.

The other probability-related notion we use in our work is the notion of first-order stochastic dominance. Consider X again, and we now consider a random variable $\tilde{X}$ such that $X \leq^{\text{st}} \tilde{X}$, where the stochastic order $\leq^{\text{st}}$ is defined as:

$$\mathrm{P}\{X > z\} \leq \mathrm{P}\{\tilde{X} > z\}, \ z \in (-\infty, \infty). \tag{2.6}$$

Examining (2.6), we see that, for any value z, the probability to the right of z under the graph of the distribution of X will be less than or equal to that of $\tilde{X}$. An equivalent condition, which helps convey the intuitive usefulness of first-order stochastic dominance, is:

$$\mathrm{E}f(X) \leq \mathrm{E}f(\tilde{X}), \tag{2.7}$$

for all increasing functions f such that the expectations exist.

2.3 Literature Review

Our work most closely relates to existing work on principal-agent modeling, auditing, linear contracts, and stochastic comparison.

Our work contributes to the extensive literature on principal-agent modelling; specifically, we augment the contract theory literature concerning moral hazard with hidden actions. Due to the size of this body of literature, we refer to Bolton and Dewatripont (2004) and Laffont and Martimort (2009) for comprehensive discussions. Our model formulation is relatively standard insofar that we establish a bilevel program that the principal solves to find the optimal contract. In our solution, we make use of the first-order approach discussed in Rogerson (1985) to simplify the agent's embedded optimization problem, which allows us to solve an equivalent but more tractable problem. Similar to Bénabou and Tirole

(2016) and Helm and Wirl (2020), we model the agent as having a quadratic disutility of effort; intuitively, factors such as fatigue and stress contribute to this growing disutility.

Typically, the principal uses an audit action or observation of monitors to gain access to information that is, otherwise, solely known to the agent. The auditing and monitoring literature is extensive, considering scenarios including government regulation of firms (Baron and Besanko (1984)), games of tax reporting and evasion (Reinganum and Wilde (1985), Reinganum and Wilde (1986), Graetz et al. (1986), Border and Sobel (1987)), and audits to investigate managerial actions (Kofman and Lawarree (1993)). The type of auditing mechanism is also the subject of study. Classic work by Townsend (1979) focuses on deterministic auditing commitments, while subsequent work (Reinganum and Wilde (1985), Mookherjee and Png (1989)) examine threshold policies and random audits. The type of auditing commitment is another area of consideration. Khalil (1997) considers when the principal cannot commit to an audit policy. Khalil and Lawarree (2001) show how the principal can benefit from choosing what performance metrics to monitor ex-post. H.-C. Chen and Liu (2008) examine optimal incentive contracts under imperfect auditing under both no-commitment and commitment schemes; they find certain classic contracts (those of Baron and Besanko (1984)) fail under the commitment case under imperfect auditing schemes. More recently, Barbos (2019) studies dynamic contracting under moral hazard in an infinitely repeated game where contracts are implemented with random monitoring technology. M. Chen et al. (2020) investigate a setting where a principal induces effort from an agent to reduce the arrival rate of a Poisson process of adverse events; this setting models effort as unobservable unless the principal engages in costly monitoring. Hoffmann et al. (2021) investigate how to design incentives for an agent who engages in activities that produce a time-delayed signal that can be observed and contracted on.

We deploy a classic random auditing scheme wherein the principal precommits to the probability of auditing the agent's work. Somewhat differently from existing work, we focus the audit on task difficulty; we assume that the principal cannot observe it without auditing. We also introduce two terms to our contract design corresponding to this auditing element. They account for the reported task difficulty and the need to compensate or punish the agent depending on the result of an audit. We find that by establishing a condition relating these terms, we can remove any incentive to misreport (which is in line with the Revelation Principle).

We propose an additively separable wage contract with mostly linear terms, so our work relates closely to existing findings about linear contracts. Holmstrom and Milgrom (1987) find that a linear function is the optimal incentive scheme over time for an agent with constant absolute risk aversion. Diamond (1998) examines contracts in which outcomes depend on managers' choices as well as efforts and finds that, if the control space of the agent has full dimensionality, the optimal contract converges to a linear payoff as the cost of effort shrinks. More recent work on linear contracts has incorporated a focus on robustness, which is a key feature of Carroll (2015), Yu and Kong (2020), Garrett (2021). While our proposed contract does have a mostly linear form, we do incorporate implicit nonlinearity. Specifically, (as can be seen in the agent's expected utility function (2.16)), we have two terms that compensate the agent based on the expected values of service quality and service time. These expectations are modelled as functions of the agent's exerted effort level. As such, it is easy to incorporate nonlinear expectations depending on the distributional choice.

In Section 2.6 we consider how business process modifications impact the optimal contract design. We model training and knowledge enhancements as first-order stochastic dominant shifts on the distributions of service quality and service time. For a thorough discussion on stochastic dominance viewed through the lens of majorization, see Marshall et al. (2011). Stole (1993) provides an useful tutorial on standard approaches to incorporating first-order stochastic dominance in principal-agent models. Rogerson (1985) identifies sufficient conditions (the monotone likelihood ratio condition and the convexity of the distribution function condition) for the first-order approach to be valid in solving principal-agent problems; in the work he highlights the equivalence of those conditions with first-order stochastic dominance.

Moving beyond standard first-order stochastic dominance considerations, we explore the effects of OPMs. To capture these non-trivial effects, while also corresponding well to real-world operating environments, we model these operational changes as instances of censoring our service quality and service time distributions. To our knowledge, our work is the first instance of modelling operational changes in this way.

2.4 The Model of Interest

In this Section, we outline the various components of our model. In Section 2.4.1, we describe how the principal, agent, and Nature interact. In Section 2.4.2, we define our distributions for service time, service quality, and task difficulty. In Section 2.4.3, we define the agent's wage function and the principal and agent's profit functions; we also declare certain simplifying assumptions. In Section 2.4.4, we state the final version of our mathematical program.

2.4.1 Dynamics

In the game we examine, the principal is the employer, and the agent is the employee. The principal wants to hire an agent to handle technical support requests. As such, the principal offers a contract to the agent that pays them a wage after they complete the support request. The agent can accept or reject the contract; if the agent rejects the contract, the game concludes. If the agent accepts the contract, the game continues, and Nature, a pseudo-player, chooses the difficulty of the support request and delivers the request to the agent (e.g., via email). The agent observes the difficulty of the request and chooses an effort level to exert to resolve it. Once the request is resolved, it generates two outputs that all players observe: service quality and service time.

This observation is possible in practice via modern Customer Relationship Management (CRM) technologies. According to Deloitte, CRM tools allow full observation of a customer interaction, from the initial point of contact through the final resolution and followup customer satisfaction surveys (Micallef (2018)). Such observation even encompasses different modes of communication, allowing firms to be contacted via one medium and seamlessly address the customer via another, with all of these interactions being tracked and logged in an electronic system. After completing the support request, the agent records the difficulty of said request. Notably, the agent can strategically misrepresent task difficulty and record a value that is different from the one Nature assigned. The principal may, at a cost to himself, audit the agent's recorded difficulty. We assume that the principal has perfect accuracy in assessing difficulty via the audit.

The game concludes once the principal either decides not to audit or decides to and completes an audit. Note: we model the audit as a pre-committed probability of auditing the agent's reported difficulty. We choose to model auditing this way as we make the implicit assumption that the principal is busy with their own work and can only afford to allocate a fraction of their time to auditing the agent's work; the pre-committed probability corresponds nicely with this notion.

Figure 2.1 displays the precise timeline of events described above. We note that the focus of our research is not the game but, rather, the design of the principal's optimal contract. Section 2.4.2 defines our model parameters and distributions.

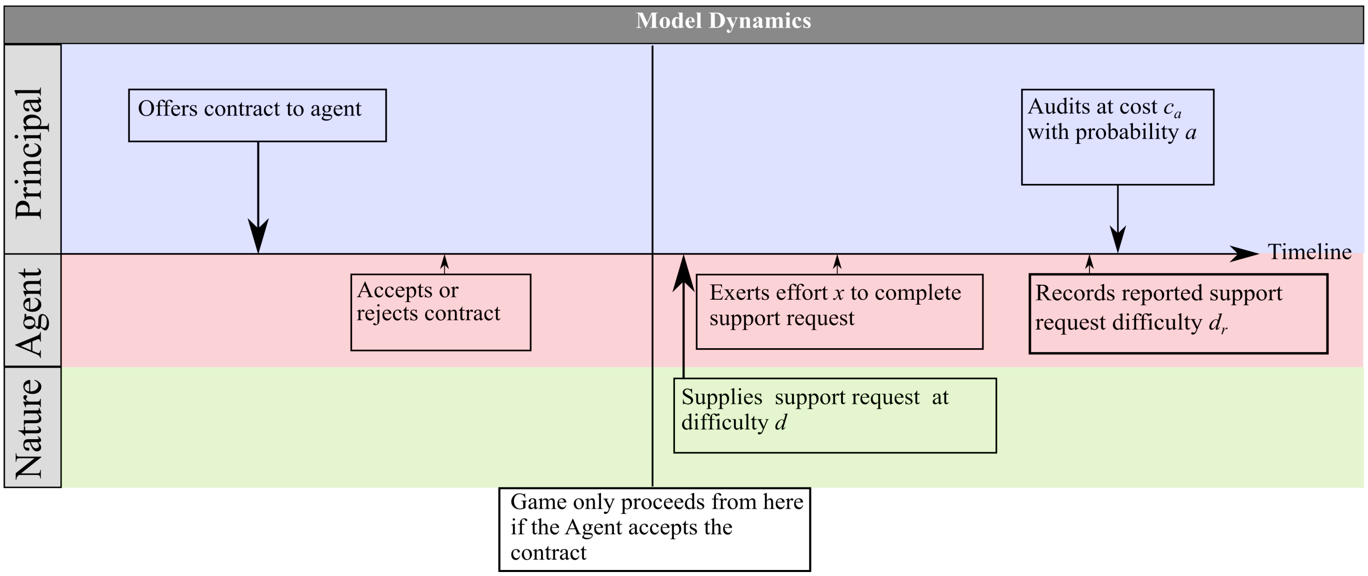

Figure 2.1: Timeline of Events

2.4.2 Parameters and Distributions

With the general dynamics explained, we now precisely define our model parameters and distributions. We introduce, in turn, the difficulty level of support requests, the agent's action choices, the auditing parameters, service quality, and service time.

1. *Difficulty Level of a Support Request* Let D be the difficulty level of a support request, which is a random variable with $\text{Prob}\{D = 1\} = 1 - \text{Prob}\{D = 0\} = p$, where 0 denotes a non-difficult support request, and 1 denotes a difficult support request. To exclude the uninteresting case where all support requests are non-difficult, we assume

that $p > 0$. We assume that both the principal and the agent know Nature's prior distribution of D.

2. *Agent Action Choices* The agent completes the support request by exerting effort x, where $x \in [0, 1]$. The agent records $d_r \in \{0, 1\}$ after completing the support request. Both x and d_r are the agent's decision variables.

3. *Auditing Parameters* The principal precommits to auditing with probability a. If the principal audits, a cost of $c_a > 0$ is incurred.

With the initial set of parameters defined, we next define service quality and service time. However, we need to clarify the units of both service metrics before proceeding. We assume that both service quality and service time are generated stochastically and influenced by the agent's effort x and task difficulty D. Notably, in real-world operations, service quality will likely be on a different scale and of different units than service time. For example, service quality can be evaluated on a rating from 0 to 100 percent, while service time is evaluated in minutes. To ensure service quality and service time achieve the same units in the utility functions, we assume that the distributions of service quality and service time map via bijection to utility values the principal receives from realized values of service quality and service time. As such, the utility values of service quality and service time are distributed stochastically, but in utility space. For convenience, we will refer to these utilities as service quality and service time for the remainder of the paper.

4. *Service Quality* Let service quality S be a random variable. The distribution of S depends on the agent's effort level x and the support request difficulty D. We assume that $0 \leq S \leq M$, as this corresponds to a bounded interval of service quality. Note that higher levels of service quality are superior, from the principal's perspective.

Further, we assume that if effort $x = 0$, service quality automatically realizes a value of zero.

5. *Service Time* Let the random variable T denote service time. Similar to service quality S, the distribution of T depends on the agent's effort level x and the support request difficulty D. We assume $0 \leq T \leq M$, where M is a constant. We assume that the upperbound on the real-world service time process is finite. This corresponds to a service time threshold; namely, once service time reaches a pre-ordained maximum tolerable level, an intervention occurs. The request is removed from the agent's purview to be expedited by another employee. Further, we assume that if effort $x = 0$, service time automatically realizes a value of M.

Before concluding this section, we briefly clarify the observability of various elements of the game, from the perspective of the principal and the agent. Both the principal and the agent observe the realizations of S and T. The agent observes support request difficulty D; the principal only observes D if an audit occurs. The agent observes their effort x; the principal does not observe x.

2.4.3 Wage and Utility Functions

We assume that the wage contract is proportional to service quality s, service time t, recorded support request difficulty d_r, and the realized difficulty d. As such, the proposed wage contract w has the following form:

$$w(s, t, d_r, d, a) = \alpha s - \beta t + \gamma d_r - \delta a(d_r - d), \tag{2.8}$$

where coefficients $\{\alpha, \beta, \gamma, \delta\}$ are in fact the principal's decision variables that are explained in Table 2.1.

Variable	Purpose
α	A reward for service quality
β	A penalty for service completion time
γ	A reward for reported difficulty
δ	A misreporting penalty

Table 2.1: Wage contract decision variables

The objective of this paper - contract design - is to choose the best $\{\alpha, \beta, \gamma, \delta\}$ for the principal. We note that the linear form of (2.8) is well-supported in the literature (see Carroll (2015)).

Note: in the following, we take expectations with respect to D in both the principal and the agent's expected utility functions. We do this as these functions are considered from the perspective of the principal. Recall that the principal can only view the task difficulty D with certainty after an *ex-post* audit. As such, the principal anticipates the agent's choice of effort based on the expectation of the difficulty state, D. The principal designs the contract (selects $\{\alpha, \beta, \gamma, \delta\}$) based on this expectation.

We consider risk-neutral agents. We assume that the utility of the agent is the wage less the costs associated with the effort x. For chosen x, the agent's expected utility is:

$$U^A(x) = \mathrm{E}_D\big[w(S, T, d_r, D, a) - kx^2\big]; \tag{2.9}$$

$$= \mathrm{E}_D\big[\alpha S - \beta T + \gamma d_r - \delta a(d_r - d) - kx^2\big]; \tag{2.10}$$

$$= \alpha \Delta_S(x) + \beta \Delta_T(x) + \mathrm{E}_D\big[\gamma d_r - \delta a(d_r - d) - kx^2\big], \tag{2.11}$$

where

$$\Delta_S(x) = E_D\big[E[S \mid D]\big]; \tag{2.12}$$

$$\Delta_T(x) = -E_D\big[E[T \mid D]\big] \tag{2.13}$$

are the expected service quality and expected service time for chosen effort x, respectively, kx^2 is the disutility of effort, and k is a constant. The choice of kx^2 is supported in the literature, as seen in Bénabou and Tirole (2016) and Helm and Wirl (2020). This form implies that an agent generates more disutility per unit of effort the closer their exertion is to their maximal effort capacity. Note: $\Delta_T(x)$ is negative since the length of service time contributes negatively to the utility of both the agent and the principal. We make this choice with regards to $\Delta_T(x)$ so we can present the contract in an additive form later.

We assume that the utility of the principal is proportional to service quality and service time less the disutility of expected auditing cost and expected wage payment. Thus, the expected utility of the principal is given by

$$U^P(x) = c_1\Delta_S(x) + c_2\Delta_T(x) - ac_a - w(\Delta_S(x), \Delta_T(x), d_r, d, a), \tag{2.14}$$

where c_1 and c_2 are nonnegative and real-valued scalar parameters that sum to 1. We have c_1 and c_2 sum to 1 to reflect the priority the principal assigns to service quality and service time (i.e. as one grows in importance, the other naturally declines in importance).

Using the explicit expression of the wage equation given in (2.8), the utility functions of the principal and the agent are given by

$$U^P = (c_1 - \alpha)\Delta_S(x) + (c_2 - \beta)\Delta_T(x) - ac_a - \delta ap - (\gamma - \delta a)d_r; \tag{2.15}$$

$$U^A = \alpha\Delta_S(x) + \beta\Delta_T(x) + \delta ap - kx^2 + (\gamma - \delta a)d_r, \tag{2.16}$$

respectively, as functions of the exerted effort x.

It is important to note that despite the linear utility form, both the principal and agent's expected utility functions are nonlinear in the agent's effort level x, and this effort level drives the service outputs S and T. Given that we are working in units of utility, not wealth, explicitly, the linear form does correspond well to work using additively separable utilities like Grossman and Hart (1983).

In Section 2.4.4, we use (2.15) and (2.16) in a bilevel mathematical program to design a contract that maximizes the principal's expected utility. To solve our optimization problem, we make the following tractability assumption:

Based on practical, real-world considerations (similar to the agent's quadratic disutility of effort), exerting higher levels of effort elicits higher incidence of factors that interfere with productive outputs (e.g., stress, strain, fatigue). These factors become more acute as effort level increases, so it is natural to consider outputs (e.g., service quality and the additive inverse of service time) increasing with effort, but at a decreasing rate. Therefore, we make the following assumption about expected service quality and service time:

Assumption 2.4.0.1. *We assume that $\Delta_S(x)$ and $\Delta_T(x)$ are monotone increasing and concave in x. Let $\Delta_S^{(1)}(x)$ and $\Delta_T^{(1)}(x)$ denote the first derivatives of $\Delta_S(x)$ and $\Delta_T(x)$ respectively with respect to x. We assume in the following that $\Delta_S^{(1)}(0) > 0$ and $\Delta_T^{(1)}(0) > 0$.*

We note that $\Delta_S^{(1)}(x), \Delta_T^{(1)}(x) \geq 0$ due to Assumption (2.4.0.1). The positivity assumptions on $\Delta_S^{(1)}(0)$ and $\Delta_T^{(1)}(0)$ imply that any effort makes a difference in the system at the effort boundary (i.e., $x = 0$). Assumption 2.4.0.1 is well-justified from an economic perspective. With an effort level of $x = 0$, service quality is automatically 0, and service time is M. As such, any chosen effort greater than $x = 0$ has performance-enhancing implications.

2.4.4 Mathematical Programs

The principal faces the problem of designing a contract by choosing $\{\alpha, \beta, \gamma, \delta\}$. This problem can be modelled as the following optimization problem, which is often called a Pareto-Optimization Problem (POP0):

$$\max_{\alpha,\beta,\delta,\gamma} U^P; \tag{2.17a}$$

$$\text{s.t. } x, d_r \in \operatorname*{argmax}_{x,d_r} U^A; \tag{2.17b}$$

$$U^A \geq 0; \tag{2.17c}$$

$$\alpha, \beta, \delta, \gamma \geq 0; \quad 0 \leq x \leq 1; \quad d_r \in \{0, 1\}. \tag{2.17d}$$

This formulation is standard, and is similar to Holmstrom (1979), Rogerson (1985), and other subsequent work. We see in (2.17a) that the principal maximizes U^P by choosing contract decision variables α, β, γ, and δ. In (2.17b), the agent chooses effort x and reports difficulty d_r to maximize their own utility, given the contract the principal has chosen. The agent has a reserve utility of 0, as shown in (2.17c), so the agent's expected utility must meet or exceed this for the agent to be willing to participate in the principal's contract. We see in (2.17d) that all decision variables have bounds according to their purpose in the contract or conventional interpretation.

The constraint set (2.17b) is the element that makes this optimization problem challenging. The following observation simplifies the above optimization problem by removing one decision variable. Specifically, it provides a condition that, given our contract form, ensures we have an incentive compatible contract. We know, via the Revelation Principle, that we need only focus on direct mechanisms (incentive compatible contracts), but the motivating managerial scenario implies we must show how the employer can arrive at such a contract.

The observation is:

Observation 2.4.1. The principal can induce the agent to tell the true difficulty level by setting $\gamma = \delta a$.

We justify this Observation as follows.

If $\gamma > \delta a$ in (2.16), the agent gains expected utility $\gamma - \delta a$ by reporting $d_r = 1$; if $\gamma < \delta a$, the agent avoids losing expected utility $\gamma - \delta a$ by reporting $d_r = 0$. By setting $\gamma = \delta a$, the agent has no incentive to report one difficulty level over the other. In the meantime, this equality has no negative effect on the expected utility of the principal $U^P(x)$. Consequently, the principal sets $\gamma = \delta a$ and induces the agent to tell the true difficulty level.

We now have an incentive compatible contract, insofar as task difficulty is concerned. However, simply setting $\gamma = \delta a$ does not guarantee the agent will exert an effort level that maximizes the principal's expected utility. To do that, we must find the optimal set of other contract decision variables. As such, we proceed with solving the bilevel program stated in (2.17a)-(2.17d).

Substituting $\gamma = \delta a$ into (2.15) and (2.16), the updated U^P and U^A are:

$$U^P = (c_1 - \alpha)\Delta_S(x) + (c_2 - \beta)\Delta_T(x) - ac_a - \delta ap, \tag{2.18}$$

$$U^A = \alpha\Delta_S(x) + \beta\Delta_T(x) + \delta ap - kx^2. \tag{2.19}$$

Using the updated values of U^P and U^A, the original optimization problem (POP0) is rewritten as (POPI); mathematical program (2.20):

$$\max_{\alpha,\beta,\delta} \; U^P = (c_1 - \alpha)\Delta_S(x) + (c_2 - \beta)\Delta_T(x) - ac_a - \delta ap; \tag{2.20a}$$

$$\text{s.t. } \; x \in \underset{x}{\operatorname{argmax}} \; U^A = \alpha\Delta_S(x) + \beta\Delta_T(x) + \delta ap - kx^2; \tag{2.20b}$$

$$\alpha\Delta_S(x) + \beta\Delta_T(x) + \delta ap - kx^2 \geq 0; \tag{2.20c}$$

$$\alpha, \beta, \delta \geq 0; \quad 0 \leq x \leq 1. \tag{2.20d}$$

Note that we have reduced our set of decision variables from $\{\alpha, \beta, \delta, x, \gamma, d_r\}$, to $\{\alpha, \beta, \delta, x\}$.

To solve the optimization problem, we consider three cases according to the value of x: i) $x = 0$; ii) $x = 1$; and iii) $0 < x < 1$. For cases i) and ii), the value of x is fixed and (POPI) becomes a standard optimization problem. For case iii), since $U^A(x)$ is assumed to be differentiable in x, then (POPI) becomes (POPII); mathematical program (2.21):

$$\max_{\alpha,\beta,\delta,x} U^P = (c_1 - \alpha)\Delta_S(x) + (c_2 - \beta)\Delta_T(x) - ac_a - \delta ap; \tag{2.21a}$$

$$\text{s.t. } \; \frac{\partial U^A(x)}{\partial x} = \alpha\Delta_S^{(1)}(x) + \beta\Delta_T^{(1)}(x) - 2kx = 0; \tag{2.21b}$$

$$\alpha\Delta_S(x) + \beta\Delta_T(x) + \delta ap - kx^2 \geq 0; \tag{2.21c}$$

$$\alpha, \beta, \delta \geq 0; \quad 0 < x < 1. \tag{2.21d}$$

Without conditions, (POPII) is not equivalent to (POPI). Under Assumption 2.4.0.1, the agent's expected utility function is concave; this condition ensures equivalence between (POPI) and (POPII). Solving (POPII) generates two candidate solutions for (POPI); if an optimal solution to (POPI) exists with $x \in (0, 1)$, it will be one of the two candidates generated from (POPII). We present the candidate solutions generated by cases i), ii), and iii) in Section 2.5.

It is worth providing a construction of probability density functions and distributions that satisfy Assumption 2.4.0.1, given its importance. Let $\Phi(s, x, d)$ and $\phi(s, x, d)$ denote the probability distribution and the probability density function of S, respectively, parameterized by x and d (where d is a realization of D). Let $\Psi(t, x, d)$ and $\psi(t, x, d)$ denote the probability distribution and the probability density function of T, respectively, parameterized by x and d. Let ϕ and ψ be twice continuously differentiable in x.

As is standard in the literature (e.g. Rogerson (1985), Stole (1993), and Marshall et al. (2011)), we make some assumptions on how S and T are influenced by changes in effort x and difficulty d. Let service quality S stochastically increase when x increases; this increase takes the form $\Phi_x(s, x, d) < 0$, where $\Phi_x(\cdot)$ is the partial derivative with respect to x. Let service time T stochastically decrease when x increases; this decrease takes the form $\Psi_x(t, x, d) > 0$, where $\Psi_x(\cdot)$ is the partial derivative respect to x In other words, the agent's effort produces a first-order stochastic dominant shift on the support of S and T. Increases in d produce a stochastic dominant shift in the form $\Phi(s, x, 0) \leq \Phi(s, x, 1)$ on the support of s and x and a stochastic dominant shift in the form $\Psi(t, x, 0) \geq \Psi(t, x, 1)$ on the support of t and x. Finally, a strong condition that guarantees Assumption 2.4.0.1 holds is if ϕ'', $\psi'' \leq 0$, where ϕ'', and ψ'' are the second partial derivatives with respect to x. Such a construction guarantees the subsequent optimization analysis will hold, though we do not claim it is the only such construction; it is merely a representative example.

2.5 Optimization Analysis

Consider the cases i), ii), and iii) identified at the end of Subsection 2.4.4. We find the optimal solution to (POPI), which leads to the following Theorem:

Theorem 2.5.1. *Consider the four solutions* $\{(\alpha^{(1)} = 0, \beta^{(1)}, \delta^{(1)}, x^{(1)}), (\alpha^{(2)} = 0, \beta^{(2)}, \delta^{(2)}, x^{(2)}),$ $(\alpha^{(3)}, \beta^{(3)} = 0, \delta^{(3)}, x^{(3)})$, $(\alpha^{(4)}, \beta^{(4)}, \delta^{(4)}, x^{(4)})\}$*, given in equations* (2.22)*,* (2.23)*,* (2.24)*, and* (2.25)*. The one that maximizes* U^P *is the optimal solution of (PIPO).*

In Appendix A.1, we prove Theorem 2.5.1. Here, we display our four candidate solutions with brief descriptions of their salient features. We put the details of our optimization approach in Appendix A.1.

Candidate 1: Case with $x = 0$.

$$\begin{aligned}
\alpha^* &= \alpha^{(1)} = 0;\\
\beta^* &= \beta^{(1)} = \min_{0<x\leq 1}\left\{\frac{kx^2}{\Delta_T(x) - \Delta_T(0)}\right\};\\
\delta^* &= \delta^{(1)} = -\frac{\beta^{(1)}\Delta_T(0)}{ap};\\
x^* &= x^{(1)} = 0;\\
U^{A*} &= U^{A(1)} = 0;\\
U^{P*} &= U^{P(1)} = c_1\Delta_S(0) + c_2\Delta_T(0) - ac_a.
\end{aligned} \tag{2.22}$$

This candidate is one of our boundary cases where $x = 0$.

Candidate 2: Case with $x = 1$.

$$\begin{aligned}
\alpha^* &= \alpha^{(2)} = 0; \\
\beta^* &= \beta^{(2)} \geq \max_{0 \leq x < 1} \left\{ \frac{k(1-x^2)}{\Delta_T(1) - \Delta_T(x)} \right\}; \\
\delta^* &= \delta^{(2)} = \frac{k - \beta^{(1)}\Delta_T(1)}{ap}; \\
x^* &= x^{(2)} = 1; \\
U^{A*} &= U^{A(2)} = 0; \\
U^{P*} &= U^{P(2)} = c_1\Delta_S(1) + c_2\Delta_T(1) - ac_a - k.
\end{aligned} \tag{2.23}$$

Another boundary case, here $x = 1$. We note that this candidate defines a lowerbound for $\beta^{(1)}$ but no upperbound; $\delta^{(1)}$ scales off $\beta^{(1)}$ to accommodate this, ensuring the agent's participation constraint is satisfied.

Candidate 3: Case with $0 < x < 1$, $\delta = 0$.

$$
\begin{aligned}
\alpha^* &= \alpha^{(3)} = \underline{\xi}(x^{(3)}) \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}; \\
\beta^* &= \beta^{(3)} = \left(1 - \underline{\xi}(x^{(3)})\right) \frac{2kx^{(3)}}{\Delta_T^{(1)}(x^{(3)})}; \\
\delta^* &= \delta^{(3)} = 0; \\
x^* &= x^{(3)} = \underset{0<x<1:\ 2\Delta_S(x) \geq x\Delta_S^{(1)}(x)}{\operatorname{argmax}} F_3(x); \\
F_3(x) &= \left(c_1 - \frac{\underline{\xi}(x) 2kx}{\Delta_S^{(1)}(x)}\right) \Delta_S(x) + \left(c_2 - \left(1 - \underline{\xi}(x)\right) \frac{2kx}{\Delta_T^{(1)}(x)}\right) \Delta_T(x); \qquad (2.24) \\
U^{A*} &= U^{A(3)} = \alpha^{(3)} \Delta_S(x^{(3)}) + \beta^{(3)} \Delta_T(x^{(3)}) - k(x^{(3)})^2; \\
U^{P*} &= U^{P(3)} = \left(c_1 - \alpha^{(3)}\right) \Delta_S(x^{(3)}) + \left(c_2 - \beta^{(3)}\right) \Delta_T(x^{(3)}) - ac_a; \\
\underline{\xi}(x) &= \frac{x - 2\frac{\Delta_T(x)}{\Delta_T^{(1)}(x)}}{2\left(\frac{\Delta_S(x)}{\Delta_S^{(1)}(x)} - \frac{\Delta_T(x)}{\Delta_T^{(1)}(x)}\right)}; \\
0 &\leq \underline{\xi}(x) \leq 1.
\end{aligned}
$$

This candidate is more elaborate than the previous two. While the full derivation is reserved for Appendix A.1, we briefly describe why it has this structure. This candidate assumes that $\delta^{(3)} = 0$. Further, this candidate results from applying the first-order approach, wherein (2.21b) must be satisfied. Examining this constraint, we note that setting either α or β to zero establish endpoints for a line segment of points that satisfy this condition. However, not all of these points satisfy the participation constraint (2.21c). As such, we find that the optimal solution is at a convex combination of the two end points; that weighting is determined by $\underline{\xi}(x)$.

Candidate 4: Case with $0 < x < 1$, $\delta > 0$.

$$\begin{aligned}
\alpha^* &= \alpha^{(4)} = 0; \\
\beta^* &= \beta^{(4)} = \frac{2kx^{(4)}}{\Delta_T^{(1)}(x^{(4)})}; \\
\delta^* &= \delta^{(4)} = \frac{1}{ap}\left(k\big(x^{(4)}\big)^2 - \frac{2kx^{(4)}}{\Delta_T^{(1)}(x^{(4)})}\Delta_T(x^{(4)})\right); \\
x^* &= x^{(4)} = \underset{0<x<1}{\operatorname{argmax}}\{c_1\Delta_S(x) + c_2\Delta_T(x) - ac_a - kx^2\}; \\
U^{A*} &= U^{A(4)} = 0; \\
U^{P*} &= U^{P(4)} = c_1\Delta_S(x^{(4)}) + c_2\Delta_T(x^{(4)}) - ac_a - k(x^{(4)})^2.
\end{aligned} \tag{2.25}$$

This candidate results from setting δ to satisfy the participation constraint while the first order condition is also satisfied. Given how δ is set, we find that the principal's utility is invariant to the choice of α and β, provided we choose a feasible solution. As such, we choose the endpoint of the feasible set to represent this candidate.

The four candidates for the optimal solution indicate that there is an optimal solution with either $\alpha = 0$ or $\beta = 0$. This is intuitive since the utility functions for both the principal and the agent are linear in α and β. In practice, this implies that the principal will design a system that either promotes the quality of service or reduces the service time, but not both.

Further, we note that α, β, and δ are not symmetric. When $\alpha > 0$, in Candidate 3, the agent is rewarded for service quality directly. Candidates 1, 2, and 4 set $\alpha = 0$ and have $\beta > 0$; given that β is a punishment term, we see that δ must, simultaneously, be positive to ensure the agent receives nonnegative utility and participates in the contract. In Appendix 2.5, we provide some useful conditions for comparing the various candidate optimal solutions.

We now present a straightforward numerical example to demonstrate an optimal solution.

Variable	$(d = 0)$	$(d = 1)$
S	$5 \cdot \text{Beta}(3x,\, 2)$	$5 \cdot \text{Beta}(2.5x,\, 2)$
T	Trunc. Expon.$(1 + x)$ on $[0, 5]$	Trunc. Expon.$(1 + 0.5x)$ on $[0, 5]$

Table 2.2: Example Distributions

We purposefully use both the Beta distribution and the Truncated Exponential distribution in this and subsequent examples. We use the Beta distribution for service quality as one often thinks of service quality as having a bound of excellence, typically from 0 to 100 percent, which is reflected nicely in the Beta distribution's bounded support. Normally, we would default to the classic Exponential distribution for modelling time. However, in this service scenario, we posit that it is extremely unrealistic to allow for unlimited service time. As such, we believe a truncated exponential distribution better reflects the real operating scenario.

a	c_a	p	c_1	k
0.3	1	0.3	0.5	1.5

Table 2.3: Example Parameters

Using our formulae from our candidate solutions, we arrive at the following:

Variable or Utility	Candidate 1	Candidate 2	Candidate 3	Candidate 4
α	0	0	0.4512	0
β	0.022857	12.5512	0.8026	3.6340
δ	0.0229	93.2307	0	32.4969
x	0	1	0.482	0.482
U^A	0	0	0	0
U^P	-0.7818	-0.6076	0.0138	0.0138

Table 2.4: Example Candidate Values

We see from Table 2.4 that both Candidate 3 and Candidate 4 are optimal for this example. This pattern holds with other examples, implying that the optimal contract is not unique.

2.6 Business Process Modifications

We now examine several business process modifications which influence the parameters of the game. In Section 2.6.1, we investigate the impact on our candidate optimal contracts after OPMs induce various censoring effects on S and T. In Section 2.6.2, we consider how our candidate optimal contracts are impacted by various types of training and employee knowledge enhancements eliciting first-order stochastic dominant shifts in the distributions of S and T. In Section 2.6.3, we investigate how changes in the business context, both inside and outside the firm, can influence the other model parameters; this investigation is in the form of a sensitivity analysis.

2.6.1 Operational Protocol Modifications (OPMs)

To reiterate, by OPMs, we mean modifications in the standard operating procedures of the firm. Below are two examples of such modifications:

- Implementing a service script: Agents can be provided with a script or checklist for what to cover in a service call. Such a script can include components like a reminder for a friendly introductory statement, confirming the software version and account number, and a flowchart for common issue resolution.
- Deploying chatbots: Prior to the customer interacting with an agent, they could be forced to first supply problem information to a chatbot designed to elicit salient problem details from them.

The mixed efficacy of these two examples is well-documented in the literature. Durgin et al. (2014) note that checklists and job aids (like a service script) are generally seen as positive, but Kaufman (2015) notes that if service is too standardized, customers can be left feeling "cold" from the interaction. Similarly, Schanke et al. (2022) note that chatbots need to be carefully designed to ensure customers have both effective and enjoyable service experiences; Castillo et al. (2021) find that customers report feeling annoyed and frustrated when the co-created service of using a chatbot fails to meet their expectations. In sum, these operational protocol changes, in comparison to the training and knowledge resource enhancements of Section 2.6.2, are less likely to lead to universal improvements in service quality and service time.

2.6.1.1 Service Quality

We now explore the impact of OPMs on service quality S through a censoring perspective. Suppose that the principal has instituted some or all of the OPMs mentioned in Section

2.6.1. Let $\hat{S}$ be the post-protocol-change random variable for service quality. We assume that the protocol changes have produced a censoring effect on service quality such that:

$$\hat{S} = \max\{\tau_L, \min\{\tau_H, S\}\}; \tag{2.26}$$

for $\tau_L, \tau_H \in [0, M]$, where $\tau_L \leq \tau_H$. The interpretation of (2.26) is such that if a realization of S falls outside the interval $[\tau_L, \tau_H]$, it automatically takes the value of τ_L or τ_H (whichever is closer). If S generates a realization within $[\tau_L, \tau_H]$, then the realization occurs as in the non-censored case.

Note that one-directional censoring (where $\tau_L = 0$ or $\tau_H = M$ but not both) is a special case of (2.26). We assume that the values of τ_L and τ_H are not dependent on x nor d. This follows from the real-world inspiration for such protocol modifications; given the structural nature, they are defined exogenous to the model.

To elaborate on the economic meaning of these parameters, τ_L and τ_H are best thought of as restrictions imposed on service quality imposed by some operational rule-change. For example, if the manager provides an employee with a plan for conflict resolution that covers various potential issues, the lowerbound of service quality may increase (τ_L), but if the plan is adhered to in a robotic nature or if it stifles creativity, the upperbound of service quality may simultaneously decline (τ_H). Depending on the proposed operational change, the effects may be stronger on one bound, or it may leave one bound unaffected.

We reconsider the example from Section 2.5. We first present the general case where τ_L and τ_H change simultaneously. Note that service time is modelled as a Beta distributed random variable that is multiplied by M. As such, we simply vary τ_L and τ_H between 0 and 1, though they are multiplied in the computation by M. To best-represent this simultaneous change, we refer to Figure 2.2.

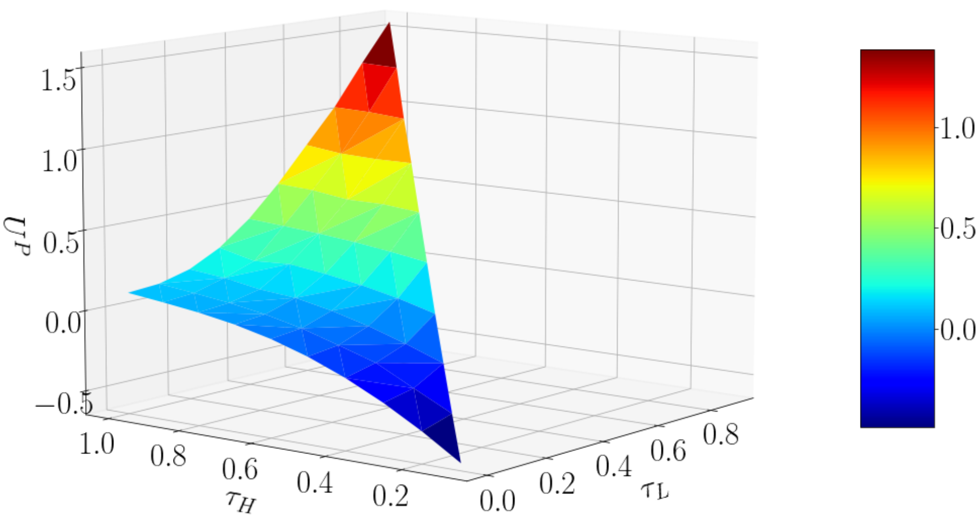

Figure 2.2: U^P for various values of τ_L and τ_H

We make the following observations.

Observation 2.6.1. For fixed τ_L, U^P increases as a concave function of τ_H.

Observation 2.6.2. For fixed τ_H, U^P increases as a convex function of τ_L.

Next, we present a special case demonstrating how our solutions change as the value of τ_L changes from 0 to 1 while τ_H is fixed at 1:

τ_L	α^*	β^*	δ^*	x^*	U^A	U^P
0	0	3.6571	32.6787	0.484	0	0.1070
0.1	0	3.4633	31.1466	0.467	0	0.1226
0.2	0	3.0394	27.7419	0.428	0	0.1746
0.5	0	1.3198	12.9795	0.233	0	0.6096
0.9	0	0.5533	5.6953	0.113	0	1.5504

Table 2.5: Post-operational-change optimal solutions for different τ_L

Observation 2.6.3. U^P improves monotonically as τ_L tends to 1 from the left.

Observation 2.6.4. As τ_L increases, β^*, δ^*, and x^* all decrease.

We note that there exist corresponding contracts to Candidate 3 for the optimal contracts in Table 2.5.

Here, we present how our solutions change as τ_H changes from 1 to 0 while τ_L remains fixed at 0:

τ_H	α^*	β^*	δ^*	x^*	U^A	U^P
1.0	0	3.6571	32.6787	0.484	0	0.1070
0.9	0	3.5311	31.6840	0.473	0	0.0801
0.4	0	2.4605	22.9606	0.370	0	-0.2104
0.2	0.5614	0.3623	0	0.282	0	-0.04546
0.1	0.6962	0.2197	0	0.326	0	-0.6046

Table 2.6: Post-operational-change optimal solutions for different τ_H

Observation 2.6.5. U^P decreases monotonically as τ_H decreases.

Observation 2.6.6. As τ_H decreases, β^*, δ^*, and x^* all decrease.

Observation 2.6.7. As τ_H decreases, for contracts with α, α increases.

We note that in the above examples, Candidate 4 is always optimal (though Candidate 3 is optimal simultaneously). As such, we use the formula of the principal's optimal utility in Candidate 4 to derive the following property, which provides insight into how τ_L and τ_H affect the princpal's utility (the proof of which is in Appendix A.4):

Property 2.6.7.1. *The principal's optimal utility increases convexly as τ_L increases and increases concavely as τ_H increases.*

The proof of Property 2.6.7.1 is in Appendix A.4.3, along with analytical formulae of the associated derivatives.

2.6.1.2 Service Time

We now explore the impact of OPMs on service time T. Let $\hat{T}$ be the post-protocol-change random variable for service time. We assume that the protocol changes have produced a censoring effect on service time such that:

$$\hat{T} = \max\{\Upsilon_L, \min\{\Upsilon_H, T\}\}; \tag{2.27}$$

for $\Upsilon_L, \Upsilon_H \in [0, M]$, where $\Upsilon_L \leq \Upsilon_H$. Equation (2.27) has a similar interpretation to (2.26). If a realization of T falls outside the interval $[\Upsilon_L, \Upsilon_H]$, it automatically takes the value of Υ_L or Υ_H (whichever is closer). If T generates a realization within $[\Upsilon_L, \Upsilon_H]$, then the realization occurs as in the non-censored case.

Note that one-directional censoring (where $\Upsilon_L = 0$ or $\Upsilon_H = M$ but not both) is a special case of (2.27). We assume that the values of Υ_L and Υ_H are not dependent on x nor d. This makes intuitive sense; for example, if a service script takes two minutes to complete, minimum, this should not change depending on the effort or task difficulty.

We reconsider the example from Section 2.5. We first present how our solutions change as both Υ_L and Υ_H change simultaneously. To best-represent this simultaneous change, we refer to Figure 2.3.

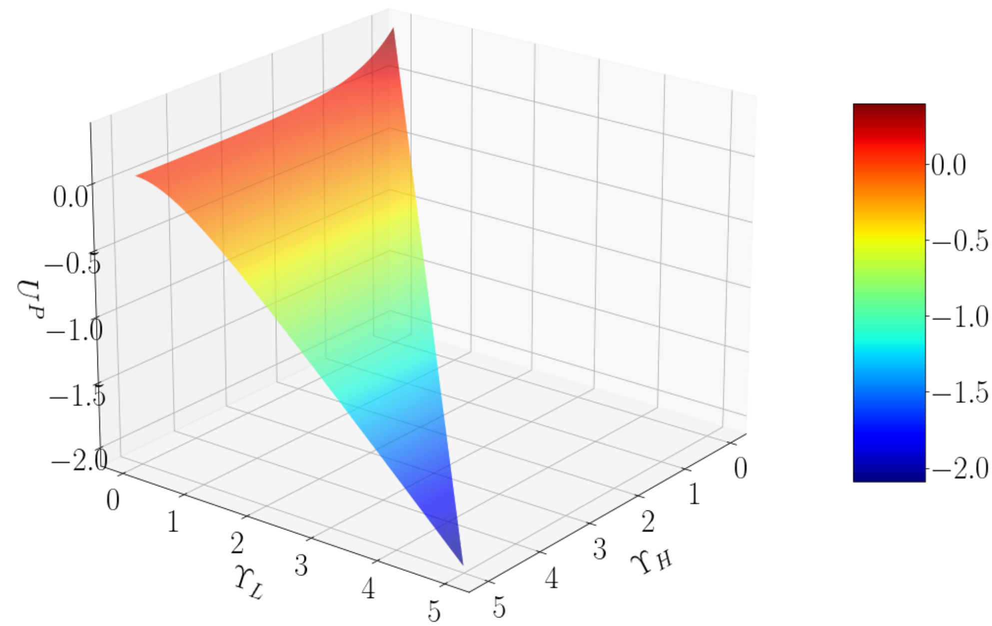

Figure 2.3: U^P for various values of τ_L and τ_H

Observation 2.6.8. When holding Υ_H fixed, U^P as a function of Υ_L is concave decreasing.

Observation 2.6.9. When holding Υ_L fixed, U^P as a function of Υ_H is convex increasing.

In Table 2.7 we present a special case of how our solutions change as Υ_L goes from 0 to 4.9 while Υ_H remains fixed:

Υ_L	α^*	β^*	δ^*	x^*	U^A	U^P
0	0	3.6687	32.7697	0.485	0	0.0465
0.1	0	3.6929	33.2365	0.484	0	0.0432
1.1	0	6.6187	95.7649	0.467	0	-0.2232
2.1	0	18.0494	432.0483	0.454	0	-0.6645
3.1	0	67.5353	2335.4606	0.449	0	-1.1498
4.1	0	475.7598	21681.9626	0.448	0	-1.6463
4.9	0	53215.0222	2.8972×10^6	0.447	0	-2.0458

Table 2.7: Post-operational-change optimal solutions for different Υ_L

Observation 2.6.10. Increasing Υ_L causes U^P to decrease monotonically.

Observation 2.6.11. As Υ_L increases, β^* and δ^* increase.

Observation 2.6.12. As Υ_L increases, x^* decreases but not markedly.

In Table 2.8 we present how our solutions change as Υ_H goes from 1 to 0:

Υ_H	α^*	β^*	δ^*	x^*	U^A	U^P
5	0	3.6687	32.7697	0.485	0	0.0465
4.9	0	3.6689	32.7711	0.485	0	0.0465
3.9	0	3.6955	32.9208	0.484	0	0.0473
2.9	0	3.8855	34.0413	0.483	0	0.0516
1.9	0	4.6112	37.7731	0.476	0	0.0702
0.9	0	9.0729	55.2891	0.462	0	0.147
0.1	0	349.2698	365.7991	0.448	0	0.3575

Table 2.8: Post-operational-change optimal solutions for different Υ_H

Observation 2.6.13. Decreasing Υ_H causes U^P to improve monotonically, in a noticeably convex fashion.

Observation 2.6.14. Decreasing Υ_H causes β^* and δ^* to increase.

Observation 2.6.15. Decreasing Υ_H causes x to decrease monotonically.

We note that in the above examples, Candidate 4 is always optimal (though Candidate 3 is optimal simultaneously). As such, we use the formula of the principal's optimal utility in Candidate 4 to derive the following property, which provides insight into how Υ_L and Υ_H affect the princpal's utility (the proof of which is in Appendix A.4):

Property 2.6.15.1. *The principal's optimal utility decreases concavely as* Υ_L *increases and decreases convexly as* Υ_H *increases.*

The proof of Property 2.6.15.1 is in Appendix A.4.3, along with analytical formulae of the associated derivatives.

2.6.1.3 Managerial Implications

The main takeaway from this work is that managers now have a tool (distributional censoring) to use that can explicitly connect OPMs and the incentive contracts within (and potentially without) the firm. Accurately modelling the connection between operational change and incentives is important. With a game-theoretic mindset, managers can better anticipate how their OPMs not only modify or restrict the behaviour of their employees, but managers can also use this information to adjust the incentive mechanisms in light of the anticipated modifications. As we noted above, under censoring effects, the values of the principal's decision variables in the optimal contract can change quite markedly. Without an adequate modelling tool, the principal may misjudge how to adjust the agent's contract, leading to rent extraction or the agent rejecting the contract.

Further, adopting a contract-focused view may enhance ideation phase of the operational improvement workflow. Specifically, an operations management team that considers the way contracts influence internal and external response to OPMs will be able to better forecast if the proposed modifications will help the firm reach its goals. These improved forecasts can then be used to discard ideas earlier before wasting excessive time and resources exploring their feasibility.

Realizing that OPMs can be viewed through this censoring lens can enhance an operations analyst's efforts. Consider an analyst generating a set of potential OPMs from which a manager will choose to implement. An analyst who can properly forecast how τ_L, τ_H, Υ_L, and Υ_H impact operations will be able to augment and refine existing cost-benefit analysis approaches. This, in turn, will help the manager make a more-informed, and likely more impactful, decision.

2.6.2 Training and Knowledge Enhancements

In this Section, we provide a contrast to OPMs by considering training and knowledge base enhancements. Employee training and knowledge base enhancements are, perhaps, the most obvious business process modifications. Training is a natural component of the onboarding process for junior to midlevel positions, provided the employee has not moved laterally into a similar role as a previous employment engagement (in which case, training may not be necessary). Butcher et al. (2009) note that the literature on employee training indicates that firm size is highly correlated with the investment in training and willingness to train employees. As firms move away from the survival phase of small operations to the growth phase, staffing needs increase, as does the focus on human resource management. Also, as firms feel more financially secure, they tend to have more bandwidth for such training investments.

Below are a few examples of training:

- Cordiality training: By cordiality, we mean that how pleasant the employee is when interacting with customers.

- Resiliency training: By resiliency, we mean how resilient the employee is to provocation or other belligerent behaviour on the part of the customer.
- Service speed training: This type of training could be conducted via timed, simulated service requests after which the agent is evaluated.
- Language training: A somewhat nonobvious form of training is that of language fluency and dialect, at least after the agent is hired. Subsidized language courses and vocal coaching are two examples of how this service component could be enhanced.

Aside from training, a firm can enhance internal knowledge resources available to agents. These resources go beyond the software manual for the software product that the firm sells. As an example, service employees can collaboratively build a shared knowledge resource that details solutions for common customer requests.

2.6.2.1 Service Quality

We now explore the impact of training and knowledge enhancement on service quality S through the lens of stochastic comparison. Suppose that the principal has instituted some or all of the training and knowledge system modifications mentioned in Section 2.6.2. Let $\tilde{S}$ be the post-training random variable for service quality. We assume that the training has produced a first-order stochastic dominant shift on service quality such that $S \leq^{\text{st}} \tilde{S}$, where the stochastic order $\leq^{\text{st}}$ is defined as:

$$P\{S > z\} \leq P\{\tilde{S} > z\}, \quad 0 \leq z \leq 1. \tag{2.28}$$

We reconsider the example from Section 2.5. We model post-training service quality changes by introducing a new parameter θ. We use θ to define the distribution of $\tilde{S}$ in the following way:

$\tilde{S}$ $(d=0)$	$\tilde{S}$ $(d=1)$
$5 \cdot \text{Beta}(3\theta x,\, 2)$	$5 \cdot \text{Beta}(2.5\theta x,\, 2)$

Table 2.9: Post-Training Service Quality Distribution

We can think of θ as an agent-specific performance metric, like innate ability or productive capacity, which is enhanced by training (better training yields larger values of θ).

We now precisely define first-order stochastic dominant shifts in the Beta distribution. Suppose that X and Y are random variables. If Y stochastically dominates X in the first-order sense, this is equivalent to the following condition from Marshall et al. (2011):

$$Ef(X) \leq Ef(Y), \tag{2.29}$$

for all increasing functions f such that the expectations exist. Next, we recall (see N. L. Johnson et al. (1995)) that the expectation of a Beta(a, b)-distributed random variable Z is:

$$E[Z] = \frac{a}{a+b}. \tag{2.30}$$

Substituting $\tilde{S}$ (and its parameters) for Z, we have:

$$E[\tilde{S}] = (1-p)\frac{3\theta x}{3\theta x + 2} + p\frac{2.5\theta x}{2.5\theta x + 2}. \tag{2.31}$$

We see that (2.31) is an increasing function of θ. As such, increasing θ yields a first-order stochastic dominant shift in $\tilde{S}$. All other model parameters are unchanged from Section 2.5.

We see the results of varying θ in Table 2.10:

θ	α^*	β^*	δ^*	x^*	U^A	U^P
1	0	3.6340	32.4969	0.482	0	0.0138
2.5	0	3.5538	31.8639	0.475	0	0.5748
5	0	3.0605	27.9128	0.430	0	0.9400
7.5	0	2.7225	25.1446	0.397	0	1.1125
10	0	2.4890	23.1997	0.373	0	1.2160

Table 2.10: Post-training optimal solutions for different service quality $\tilde{S}$

Examining Table 2.10, we make the following observations.

Observation 2.6.16. U^P improves as θ increases.

Observation 2.6.17. Optimal effort x^* and contract variables β^* and δ^* decrease as θ increases.

2.6.2.2 Service Time

We now explore the impact of training and knowledge enhancement on service quality T through the lens of stochastic comparison. Let $\tilde{T}$ be the post-training random variable for service quality. We assume that the training has produced a first-order stochastic dominant shift on service quality such that:

$$\tilde{T} \leq^{\text{st}} T. \tag{2.32}$$

We consider the impact of improved service time on the setting described in the example from Section 2.5. We model post-training service time changes by introducing a new parameter θ. We use θ to define the distribution of $\tilde{T}$ in the following way:

$\tilde{T}$ $(d = 0)$	$\tilde{T}$ $(d = 1)$
Trunc. Expon.$(1 + \theta x)$ on $[0, 5]$	Trunc. Expon.$(1 + 0.5\theta x)$ on $[0, 5]$

Table 2.11: Post-Training Service Time Distribution

Similar to with service quality, θ reflects innate ability or productive capacity, which grows through training.

We now define first-order stochastic dominant shifts on the truncated exponential distribution. First, we once again concentrate on the expectation condition (2.29) from Marshall et al. (2011). Next, we recall (see Al-Athari (2008)) that the expectation of a Truncated Exponential(λ, b) distributed random variable Z, where λ is the usual Exponential distribution parameter and b is the upperbound truncation point, is:

$$E[Z] = \frac{1}{\lambda} - \frac{b}{e^{\lambda b} - 1}. \tag{2.33}$$

Considering $\tilde{T}$ as defined in Table 2.11, we note that our expectation is the probability-weighted average of the expectations in each difficulty state. In particular, when $d = 0$, $\lambda = 1 + \theta x$ and when $d = 1$, $\lambda = 1 + 0.5\theta x$. Further, $b = M$ in both states. Substituting $\tilde{T}$ (and its parameters) for Z, λ, and b in (2.33), we have:

$$E[\tilde{T}] = (1 - p)\left[\frac{1}{1 + \theta x} - \frac{M}{e^{(1+\theta x)M} - 1}\right] + p\left[\frac{1}{1 + 0.5\theta x} - \frac{M}{e^{(1+0.5\theta x)M} - 1}\right]. \tag{2.34}$$

We see that (2.34) is a decreasing function of θ. As such, increasing θ yields a first-order stochastic dominant shift in $\tilde{T}$. All other model parameters are unchanged.

We see the results of varying θ in Table 2.12:

θ	α^*	β^*	δ^*	x^*	U^A	U^P
1	0	3.6340	32.4969	0.482	0	0.0138
2.5	0	13.5687	55.7461	0.454	0	0.1983
5	0	46.0975	98.5695	0.447	0	0.2790
7.5	0	98.1281	141.7190	0.445	0	0.3085
10	0	170.0277	185.3524	0.445	0	0.3238

Table 2.12: Post-training optimal solutions for different service time $\tilde{T}$

Examining Table 2.12, we make the following observations.

Observation 2.6.18. U^P improves as θ increases.

Observation 2.6.19. β^* and δ^* increase as θ increases.

Observation 2.6.20. x^* decreases as θ increases.

2.6.2.3 Managerial Implications

The benefits of employee training and knowledge enhancements are somewhat obvious, though still important. However, it is important to note that firm willingness to invest in training, especially for small firms, is mixed. Thomson and Gray (1999) find that growth-oriented small firms tend to have positive views on training activity. Vinten (2000), similarly, finds that firms that undertake more training investments view training as beneficial and important to success. Conversely, S. Johnson (2002) finds that, especially for training activities, small business owners see business improvement measures as an act of faith. Patton and Marlow (2002) document a chronic fear of staff mobility in small firms leads to owners and managers viewing these activities as wasteful.

What the above indicates is that attitudes towards training tend to carry some momentum; namely, firms appear initially skeptical in such process modifications. However, once investments are made, there is a documented tendency to believe in the efficacy of such process improvements. Our analysis in Sections 2.6.2.1 and 2.6.2.2 demonstrates potential positive results of training, which may encourage firms to reconsider their choice to avoid training investments.

2.6.3 Changing Business Context

Any firm can experience changes that modify the context of that business. Some relevant examples for our setting include:

- Market position: The firm's value proposition may change; namely, the firm may choose to emphasize more service quality or service time. As such, service priorities could change, which may induce a change in c_1 and c_2.
- Managerial role: The principal's responsibilities at the firm could change. This could affect their ability to audit the agent, impacting a.
- Customer mix: The SaaS firm could experience a shift in customer mix (e.g., they develop a new software product to sell to a different type of firm). This could impact the probability p of receiving a difficult support request.
- Agent factors: The agent could experience changes either within themselves or outside the work environment that impact the disutility they get from exerting effort (e.g., if they have to suddenly start acting as a caregiver at home, they may want to save effort exerted at work). These changes could impact k.

We conducted sensitivity analysis on each of the above-mentioned parameters. Changing c_1, c_2, a, and p all yield linear changes in the principal's optimal utility. Specifically, increasing c_1 (thus, decreasing c_2) yields a linear increase in the principal's optimal utility, which makes some intuitive sense as the principal gains more from service quality, the positive service output. Increasing a and p each yield decreases in the principal's optimal utility; this makes sense as auditing more often increases expected auditing costs, and increasing the probability of a difficult support request increases the probability of worse service outcomes. The parameter k is the only one that yields nonlinear utility responses; as such, we focus on it.

We display how U^P changes with respect to k in Figure 2.4.

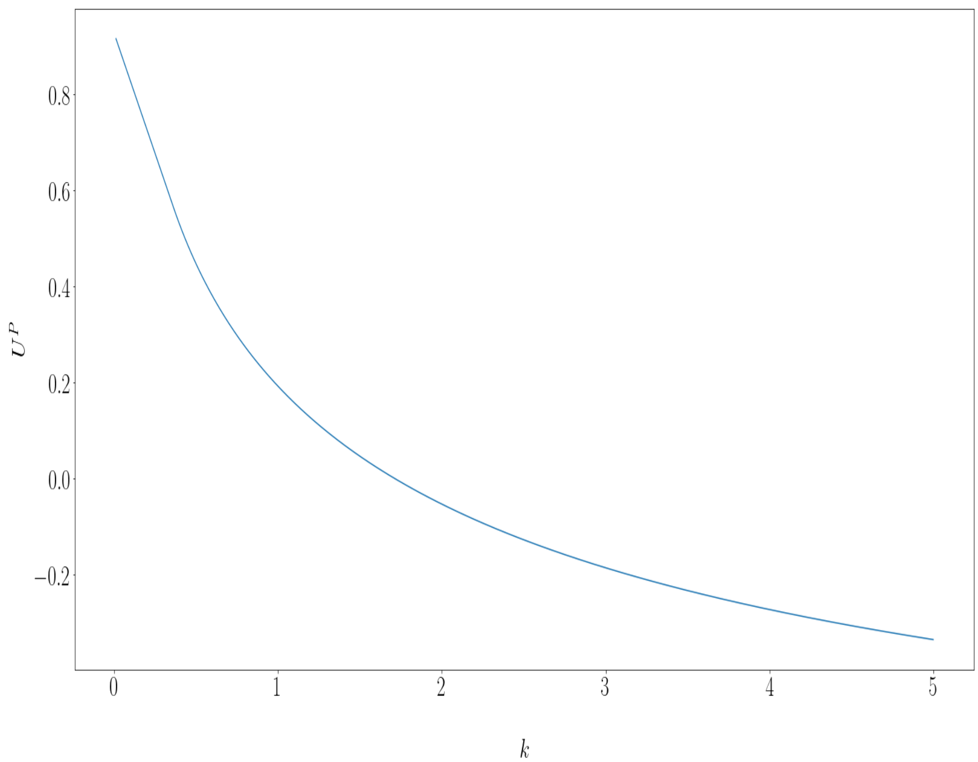

Figure 2.4: Optimal U^P as a function of k

Observation 2.6.21. Increasing k is associated with monotone, convex decreases in U^P.

We also show how optimal x changes in response to changes to k in Figure 2.5.

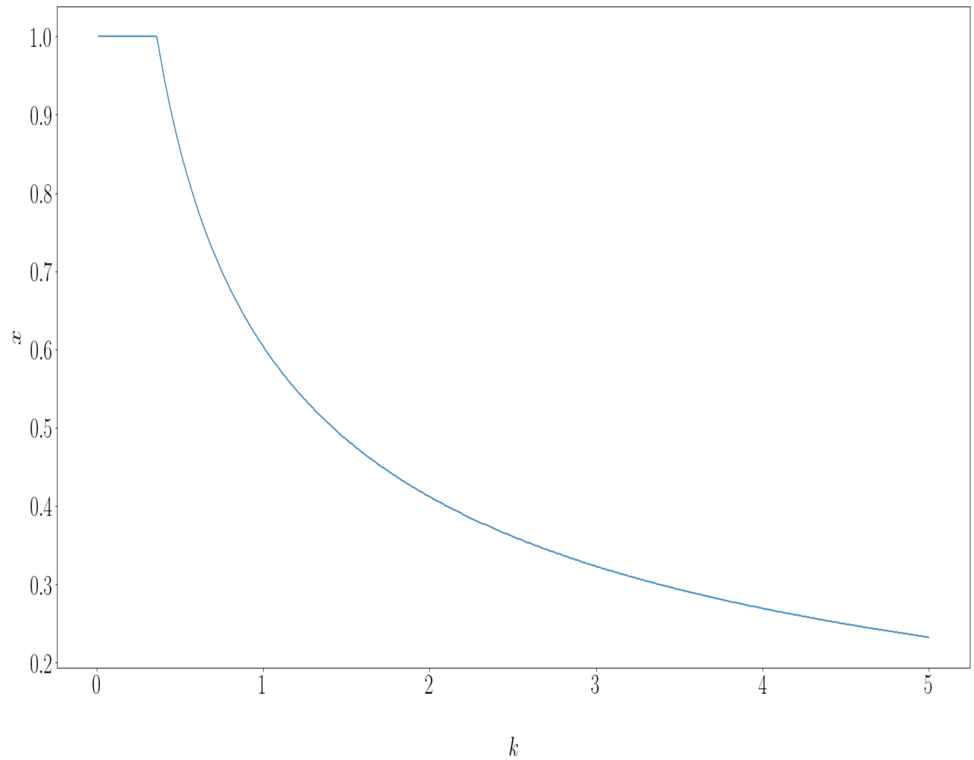

Figure 2.5: Optimal x as a function of k

Observation 2.6.22. Initially, x is unchanged as k increases, only for small values of k. Thereafter, x decreases monotonically with k

This makes intuitive sense; if the cost of effort is extremely low for an agent, exerting maximal effort is optimal.

2.6.3.1 Managerial Implications

Of the changing elements of the business context, it is worth noting that k is the most interesting and produces the most noticeable effects on the principal's utility. Further, the contract structure changes for small values of k as we shift from Candidate 2 to Candidate 3 and Candidate 4 being optimal.

We see in Figure 2.4 that improving the agent's outside-of-work circumstances may lead to utility benefits for the principal, given a sufficiently large change in k. However, modifying the agent's disutility of exerting effort is not necessarily an easy task. Earlier, we gave the example that an employee may have a high disutility of effort because they need to reserve energy for a caregiving role off the job. This may serve to explain why employers provide things like childcare benefits and paid time off for parental leave. Further, the potency of k in impacting the principal's utility implies that spending time getting to know one's employees may be materially beneficial. Specifically, with better knowledge of an employee's outside-of-work circumstances, a principal may be better able to provide support that materially benefits both parties.

2.7 Practical Considerations

As with any theoretical model inspired by a real-world scenario, it is important to clearly lay out the limitations and assumptions and how these may be affected in a real-world scenario. In particular, we point out scenarios in which our model may need adjustment.

We note that our model and the dynamics therein were constructed with a specific service scenario in mind. In particular, we model a service employee working for a SaaS firm. This type of service can typically be done remotely and is supported by electronic documentation tools, like a support ticketing system. In this scenario, auditing is relatively seamless in that it can be completed by accessing relevant service information via the electronic system, though it still costs time. Further, in our model, we have assumed perfect auditing accuracy, as this best-reflects the level of expertise at our industry partner.

Not all service environments operate in this fashion. Firstly, many service operations are not completed virtually; some examples include pest control, security system maintenance, and automotive repair. As such, the ability to audit completed work may be more costly, and potentially more difficult. In some cases, auditing may even be destructive to the original work; an example of this could be examining wiring done by an electrician after the walls have already been resealed. Such auditing environments may impose additional costs on the principal or force them to use costly real-time monitoring methods to avoid having work redone ex-post.

Beyond just the difficulty of conducting an audit, it may be the case that the principal does not have the requisite expertise to actually validate or invalidate the agent's work. A homeowner, for example, may not be able to accurately scrutinize the work done by a contractor during a renovation, aside from determining if they like the aesthetics or not. From a modelling perspective, this would require a re-interpretation of the auditing component of our contracts. The easiest addition would be to add a probability of correct assessment to the model, which would come into play if an audit occurred; this would end up increasing the effective cost of the audit.

Another limitation of our model is that it only considers a single period. Some service scenarios are multi-period or involve additional decision points. For example, a mechanic may inspect a vehicle that was initially brought in for a single issue and find multiple issues, opening up the option for the principal to negotiate a price after information is revealed. These scenarios would likely need to be framed as a multi-period, sequential game. Notably, these scenarios wherein information is revealed by the agent who has more expertise (e.g., the mechanic) than the principal would likely benefit from including reputation in some fashion, to gauge the trustworthiness of the information.

While not relevant to our target service scenario, it is worth mentioning other ways in which compensation can be computed in practice. Many services operate on a fixed-price model, though there is also the "cost +" model that can be common in general contracting. The fixed price model is reasonably amenable to a static model. The "cost +" model, in which price is based on the cost plus some margin, can be problematic in practice and would likely require a dynamic model. In particular, this pricing approach can often be leveraged by the agent to collect an informational rent, especially when information is only revealed partway through the task. In a home renovation, for example, information may be revealed during the renovation after some walls are demolished. This information could be that more work needs to be done on the house than expected. With a destroyed wall, many homeowners feel compelled to do the extra work and pay for it, which may or may not be truly necessary. This also gives the agent more bargaining power, which adds an additional difficulty into the modelling.

2.8 Conclusion

Change and turmoil are constant features of dynamic business environments, and these disruptive forces can often spur management to implement a variety of adjustments to the business process. One likely candidate for such adjustments is the introduction of new operating rules and procedures, the Operational Protocol Modifications (OPMs) that are the main subject of this work. Specifically, we have introduced distributional censoring as a method for modelling these OPMs. We frame the use of this new tool in the context of a mechanism design scenario wherein an employer seeks to hire an employee to handle customer service requests. By connecting OPMs to mechanism design via distributional censoring, we aim to provide managers and operations analysts an intuitive yet rigourous approach to forecast the impact of OPMs on service performance and the efficacy of current compensation contracts.

To round out our operational modification analysis, we model instances of training as a form of first-order stochastic dominant shifts on the distributions of service quality and service time. We also highlight how the disutility the agent receives from exerting effort has a noticeable effect on the principal's optimal utility; namely, the more disutility the agent generates, the less utility the principal can expect.

This work has a few noteworthy limitations. First, we only consider one utility function and contract form: an additively separable form. Secondly, we only pursue a numerical investigation of the first-order stochastic dominance effects. Finally, one could argue that we should compare multiple forms of stochastic dominance in the modelling of training effects.

We aim to pursue a few vectors of inquiry for future work in this area; some of these stem directly from the aforementioned limitations. Generalizing our model by relaxing our assumptions will entail a different optimization approach. So far, we have begun investigating the effects of relaxing the monotonicity of expected service quality and service time with respect to effort x. Another extension worth mentioning is changing the style of censoring. So far, we have only discussed the impact of censoring the tails of the distributions. It could be that some operational process modifications induce censoring across non-tail intervals of the distributional support. At first consideration, these censoring effects likely would arise from structural elements of the service process, at least with regards to service time.

Chapter 3

Approaches for Incorporating Additional Information into Network Ranking Models

3.1 Introduction

Professional sports present a captivating duality. For the casual spectator, they yield hours of wholesome entertainment. For the management scientist, they offer up a myriad of fascinating problems to explore and research. Perhaps the most fundamental of these problems is to evaluate the teams of a given sports league and rank them. However, what appears to be a simple exercise is anything but; the approaches to this problem are numerous.

The focus of our work is one subset of these team ranking methodologies: network models. These models treat teams as nodes in a network and games between teams as vertices, and we consider these models in the context of professional European soccer teams. Unlike other methods, network ranking models eschew probabilistic assumptions in favour of an intuitive and mathematically straightforward concept. This concept is that of an indirect win or loss. Put simply, a team earns an indirect win if a team they have beaten defeats another team. An indirect loss is earned similarly if a team that was lost to in the past loses to another team. The network ranking model appeals to a common conversation among sports fans, who typically use these indirect wins and losses to justify why their favourite team is better or more likely to win in a head-to-head confrontation.

Broadly, network models can be classified into two types: static and dynamic. What separates the two is that dynamic models incorporate discount factors, so older games have less influence on current rankings. Additionally, the computational approach of static models relies on convergence, so the model parameters have to lie in a narrower range. This issue is not present in dynamic models, at least not the type we examine in our work.

The main contribution of our work is twofold. Firstly, we provide methods to incorporate additional match information into both static and dynamic network ranking models. The information we focus on is home/away and goal difference information. Secondly, we introduce multiple ways to use the time series of team rankings to produce similarity metrics between teams. We then incorporate this similarity information into static and dynamic network models. This last point is especially important as our work connects two streams of literature that have long-acknowledged each other but have yet to find a methodological intersection.

To elaborate on this last point somewhat, there exist two heretofore related but unconnected streams of literature: network ranking models and analysis of competitivity graphs. The former seek to use network models to rank various entities. The latter form graphs (called competitivity graphs) where ranked entities are nodes and weighted edges are constructed between the nodes based on how many times the entities swap relative positions in sequential rankings. The weight of the edges corresponds to how many times the entities swap positions, where each position swap is called a competition. Prior to our work, the information generated from these competitivity graphs was used largely in a descriptive sense; it was used to describe how competitive sports leagues are. However, our work actually takes this information and uses it to refine the way network ranking models rank teams.

After introducing our methodological additions, we evaluate the models' performance using data from five of the most popular professional men's soccer leagues: the English Premier League, the Spanish La Liga, the German Bundesliga, the Italian Serie A, and the French Ligue 1. After demonstrating our models' superior empirical performance, we note that the type of model that performs best is league dependent. We then conduct sensitivity analysis on a subset of the methods and provide guidance on optimal parameter choice.

The rest of this chapter proceeds as follows. Section 3.2 explains the relevant theoretical background to support the rest of the chapter; this includes explanations of extant ranking methods. Section 3.3 explains our work's connection to other work in the ranking methods literature. Section 3.4 describes the data used in our analysis. Section 3.5 explains our aforementioned additions to both static and dynamic ranking models. Section 3.6 reports our results and analysis of the performance of our empirical testing. Section 3.7 contains our discussion of the results. Section 3.8 concludes.

3.2 Theoretical Background

This section serves to give the reader the requisite theoretical background and familiarity with sports ranking systems in order to make this work self-contained. This will also make the explanations of our additions to the network ranking models more efficient. Section 3.2.1 provides a very brief explanation ranking and ranking systems. Section 3.2.2 describes the standard way teams are ranked in professional soccer leagues, which we call the league point approach. Section 3.2.3 explains the Elo ranking system, which was formulated originally by Arpad Elo to rank chess players but has since been adopted by a wide variety of competitions and sports. Section 3.2.4 introduces the paired comparison approach, with particular attention paid to the Bradley-Terry method. Section 3.2.5 arms the reader with the requisite background on both the static Park and Newman (2005) and dynamic Motegi and Masuda (2012) network ranking models.

3.2.1 Ranking Primer

The goal of a ranking system is to compare a collection of items and be able to sort them according to some system-designer-specified metric. In the case of sports ranking systems, this metric is usually, broadly defined, ability or capability. We cannot directly view the ability of any team, so ranking systems use estimates of ability. These estimates are then used to rank teams, with more capable teams ranking higher. Note: in a ranking system, a lower numerical rank is deemed better. As such, in a 20-team soccer league, the best rank is 1 and the worst rank is 20.

The main factor that distinguishes different ranking systems is how they compute estimates of team capability. As we show in subsequent sections, all of the systems are based off of the number of wins, losses, and ties teams accumulate throughout a season. Some systems incorporate additional information, like which teams were played when these wins, losses, and ties were accumulated. We now proceed to describe the systems used in this work, either as bases of comparison or theoretical foundations for our contributions.

3.2.2 League Points Approach

The league points approach to ranking is straightforward. In professional soccer leagues, there are three potential outcomes to any game: win, loss, or tie. To generate the table of team performance that decides who wins the league, leagues assign point values to each of these results. A win provides a team 3 points, a loss provides 0 points, and a tie provides 1 point.

To rank teams via a league point system, we simply total the league points each team in the league has accrued by time t and rank them according to their total league points, with more points being better.

3.2.3 Elo Approach

This approach was proposed by Arpad Elo and implemented by the United States Chess Federation in 1960, and Elo described his work in detail in Elo (1978). The Elo approach assigns each team a value in points; the higher a team's point total, the higher their rank. To accomplish this, the method initially assigns each team the same point value; it then updates these values using the result of each game. After each game, only the point totals of the two teams involved in the game are adjusted. Notably, these point adjustments take into account the pre-match difference in point totals between the two competing teams. A team with a large point total will not earn many points by defeating a team with a significantly lower total, but the larger-totaled team will lose many points if they lose such a game.

The Elo model is relatively simple, insofar as the mathematics of its construction. We reproduce a description of the standard Elo model adapted from Wunderlich and Memmert (2018). Note that this model considers home and away effects, with the home team denoted by superscript H and the away team denoted by superscript A.

The original Elo approach estimates the win probability according to the following equations, where A_t denotes the point total of the away team at the start of ranking period t and H_t denotes the point total of the home team at the start of ranking period t:

$$\eta^H = \frac{1}{1 + c^{(A_t - H_t - \omega)/d}} \tag{3.1}$$

$$\eta^A = 1 - \eta^H, \tag{3.2}$$

where ω is a measure of the home advantage (in Elo-points), while c and d are freely selectable parameters that influence the scale of the rating. A common parameter choice is $\omega = 80$, $c = 100$, and $d = 400$; these come from the original application in the United States Chess Federation. These win probabilities factor into the point total adjustments, as shown below.

After the match, the actual result a^H for the home team is observed; $a^H = 1$ if the home team wins, $a^H = 0.5$ in case of a draw and $a^H = 0$ if the home team loses. As such, the result for the away team is $a^A = 1 - a^H$ and the ratings for both teams are adjusted as follows:

$$H_{t+1} = H_t + k(a^H - E^H) \tag{3.3}$$

$$A_{t+1} = A_t + k(a^A - E^A), \tag{3.4}$$

where k is an adjustment factor that is chosen via calibration.

To perform an Elo ranking over an entire season, we simply compute adjustments as outlined in (3.1)−(3.4) for each match in the season, in sequence. We then rank the teams based on their final Elo point totals, ranking from largest to smallest.

For our testing, we use the following parameter values (following Wunderlich and Memmert (2018)):

- $\omega = 80$;
- $c = 100$;
- $d = 400$;
- $k = 0.5$;
- initialized point value = 50.

3.2.4 Paired Comparison Approach

Paired comparison models are typically called Bradley-Terry models after the seminal work Bradley and Terry (1952). As the name suggests, a paired comparison model focuses on interactions between pairs of entities; in our case, matches between pairs of soccer teams. The model assumes that each team i has an unobservable ability parameter γ_i that we attempt to estimate from the data on match outcomes. These ability parameters, once estimated, form the basis for our team ranking. Namely, we rank the teams in descending order, with larger ability parameters being superior.

Proceeding to specifics, we reproduce the representation from Hunter (2004). Let $\gamma_i > 0$ be the ability of team i. The Bradley-Terry model without ties assumes the following:

$$\text{P(individual } i \text{ beats individual } j) = \frac{\gamma_i}{\gamma_i + \gamma_j}. \tag{3.5}$$

As we see from (3.5), this model makes a probabilistic assumption of how team abilities are related. In this way, it is somewhat similar to the Elo model. A key assumption of this model is the following:

Assumption 3.2.0.1. *In every possible partition of the individuals into two non-overlapping, nonempty subsets, some individual in the second set beats some individual in the first set at least once.*

For a model with ties, we use the following to denote such a result probability:

$$\text{P}(i \text{ ties } j) = \frac{\theta\sqrt{\gamma_i\gamma_j}}{\gamma_i + \gamma_j + \theta\sqrt{\gamma_i\gamma_j}}, \tag{3.6}$$

where $\theta > 0$ is the constant of proportionality if the probability of a tie is proportional to the geometric mean of the probabilities of a win by either individual. This construction yields the following ratios for the probabilities:

$$\text{P}(i \text{ beats } j) : \text{P}(j \text{ beats } i) : \text{P}(i \text{ ties } j) = \gamma_i : \gamma_j : \theta\sqrt{\gamma_i\gamma_j}. \tag{3.7}$$

Examining (3.5) and (3.6) suggests that we use maximum-likelihood estimation to estimate all γ_i and θ.

The likelihood function, which we seek to maximize via our choice of γ_is and θ, for the model with ties is:

$$\mathcal{L}(\boldsymbol{\gamma}) = \prod_{i=1}^{m}\prod_{j=1}^{m}\left(\frac{\gamma_i}{\gamma_i+\gamma_j+\theta\sqrt{\gamma_i\gamma_j}}\right)^{w_{ij}} \cdot \prod_{i=1}^{m}\prod_{j=1,j<i}^{m}\left(\frac{\theta\sqrt{\gamma_i\gamma_j}}{\gamma_i+\gamma_j+\theta\sqrt{\gamma_i\gamma_j}}\right)^{d_{ij}}, \tag{3.8}$$

where w_{ij} denotes the number of times individual i has beaten individual j (we assume $w_{ii} = 0$ by convention) and d_{ij} is the number of times team i and team j draw in the measurement period. Taking the logarithm yields:

$$\begin{aligned}\ell(\boldsymbol{\gamma}) &= \sum_{i=1}^{m}\sum_{j=1}^{m}[w_{ij}\ln(\gamma_i) - w_{ij}\ln(\gamma_i+\gamma_j+\theta\sqrt{\gamma_i\gamma_j})] \\ &+ \sum_{i=1}^{m}\sum_{j=1,j<i}^{m}[d_{ij}\ln(\theta\sqrt{\gamma_i\gamma_j}) - d_{ij}\ln(\gamma_i+\gamma_j+\theta\sqrt{\gamma_i\gamma_j})].\end{aligned} \tag{3.9}$$

Simplifying further yields:

$$\begin{aligned}\ell(\boldsymbol{\gamma}) &= \sum_{i=1}^{m}\sum_{j=1}^{m}[w_{ij}\ln(\gamma_i) - w_{ij}\ln(\gamma_i+\gamma_j+\theta\sqrt{\gamma_i\gamma_j})] \\ &+ \sum_{i=1}^{m}\sum_{j=1,j<i}^{m}\left[d_{ij}\ln(\theta) + \frac{d_{ij}}{2}\ln(\gamma_i) + \frac{d_{ij}}{2}\ln(\gamma_j) - d_{ij}\ln(\gamma_i+\gamma_j+\theta\sqrt{\gamma_i\gamma_j})\right].\end{aligned} \tag{3.10}$$

We impose the constraint on the abilities such that $\sum_{i=1}^{m}\gamma_i = 1$. We then use any standard maximum-likelihood estimation solver to compute estimates for the γ_is and θ. We then rank the teams according to the estimated γ_is, with larger γ_is being superior.

3.2.5 Network Model Approach

This section provides the necessary background on both static and dynamic network ranking models. Section 3.2.5.1 describes the static network ranking model introduced in Park and Newman (2005). Section 3.2.5.2 explains the dynamic network ranking model introduced in Motegi and Masuda (2012). Section 3.2.5.3 provides a brief summary of the key modelling and implementation benefits provided by network ranking models as a whole.

3.2.5.1 Park and Newman (2005) Static Model

Network ranking models take a different approach to ranking teams than the league point and Elo models. As explained in Section 3.1, network models do not make assumptions about teams' relative abilities and how these abilities influence the probability of victory. Instead, network models are inspired by a sport spectator's natural intuition regarding direct and indirect wins.

The classic network sports ranking model is described in Park and Newman (2005). Before presenting the mathematics, we briefly describe how the method works. The network is visualized with nodes being teams and edges pointing from winning teams to losing teams in any given match.

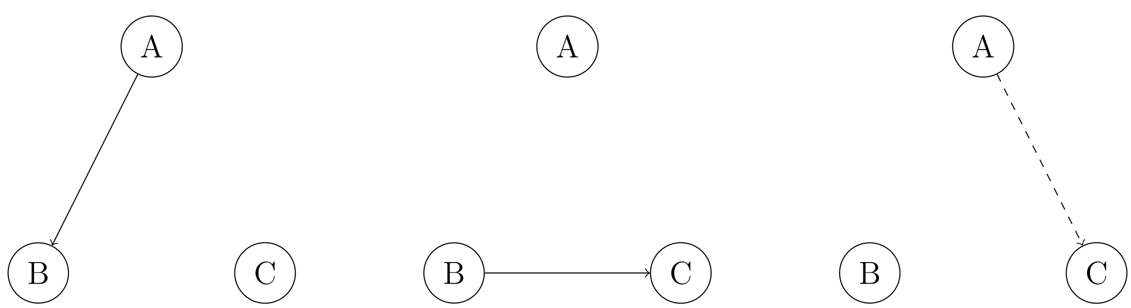

Figure 3.1: Basic Network Model Structure

In Figure 3.1, we display a brief example. In this league, there are three teams: A, B, and C. Let each set of nodes, moving from left to right, represent a gameweek. In gameweek 1, A defeats B. In gameweek 2, B defeats C. We now would be curious if A or C would win the next gameweek's match. Given that A has defeated B and that B has defeated C, a network model would assume that A has an advantage over C in the upcoming match, since A defeated B who defeated C.

Using the match history of a league, the method computes both a win score and loss score per team in the league. These win scores and loss scores are computed by tallying the number of direct and discounted indirect wins and losses, respectively, that each team has earned throughout the season. The indirect wins and losses typically range from 0.1 to 0.3 of a direct win or loss, with the discount compounding per degree of indirection. Once all teams have win and loss scores computed, we compute the difference between these scores, which is called the total score. These total scores then form the basis for ranking, with higher total scores being superior.

We proceed with the mathematics, starting with adjacency matrix A, which is an $n \times n$ real matrix, where n is the number of teams, with element A_{ij} equal to the number of times team j has beaten team i (usually 0 or 1, but occasionally 2). The number of direct wins for a team can be written as:

$$\text{direct wins for team } i = \sum_{j=1}^{n} A_{ji}. \tag{3.11}$$

We see that (3.11) makes sense as each column i holds the wins against the teams in each row j. The number of indirect wins at distance 2 (A beats B beats C) can be written as:

$$\text{indirect wins at distance 2 for team } i = \sum_{j=1}^{n} \sum_{k=1}^{n} A_{kj} A_{ji}. \tag{3.12}$$

Note: in the above, we can sum from 1 through n inclusive because the diagonal of the matrix will be zeros (since no team plays itself). We explain (3.12) as follows. In computing indirect wins by team i, we first select opponent j. If no wins were scored against this team, we have a multiplicative factor of 0, if one win, 1, and so on. We then take this multiplicative factor and multiply it by the direct win score of team j, who plays against opponents indexed by k. Note, we don't skip any indices; we include wins scored against i, since these will be balanced out by a loss score.

The type of computation illustrated in (3.12) is extended to all directed path lengths available in the network. We discount indirect wins over direct ones by a constant factor α for every level of indirection, so that an indirect win two steps removed is discounted by α, an indirect win three steps removed by α^2, and so forth. The parameter α is the only free parameter in the ranking scheme.

The total **win score** w_i of a team i is the sum of direct and indirect wins at all distances, with discounting. It is expressed as follows:

$$w_i = \sum_{j=1}^{n} A_{ji} + \alpha \sum_{j=1}^{n} \sum_{k=1}^{n} A_{kj} A_{ji} + \alpha^2 \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{h=1}^{n} A_{hk} A_{kj} A_{ji} + \cdots . \tag{3.13}$$

We factor by combining all the sums across j, yielding:

$$= \sum_{j=1}^{n} \left(1 + \alpha \sum_{k=1}^{n} A_{kj} + \alpha^2 \sum_{k=1}^{n} \sum_{h=1}^{n} A_{hk} A_{kj} + \cdots \right) A_{ji}. \tag{3.14}$$

We notice that the second and subsequent terms in the inner sum are the components of w_j multiplied by α, which leads to a more compact expression:

$$= \sum_{j=1}^{n} (1 + \alpha w_j) A_{ji} \tag{3.15}$$

If we let k_i^{out} denote the **out-degree** of vertex i in the network (the number of edges leading away from vertex i), we can restate (3.15) as follows:

$$= k_i^{\text{out}} + \alpha \sum_{j=1}^{n} A_{ij}^T w_j. \tag{3.16}$$

Using a similar logic, we can construct the **loss score** l_i:

$$l_i = \sum_{j=1}^{n} A_{ij} + \alpha \sum_{j=1}^{n} \sum_{k=1}^{n} A_{ij} A_{jk} + \alpha^2 \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{h=1}^{n} A_{ij} A_{jk} A_{kh} + \cdots . \tag{3.17}$$

We combine the various sums over j, yielding:

$$= \sum_{j=1}^{n} A_{ij} \left(1 + \alpha \sum_{k=1}^{n} A_{jk} + \alpha^2 \sum_{k=1}^{n} \sum_{h=1}^{n} A_{jk} A_{kh} + \cdots \right). \tag{3.18}$$

We note that we can rewrite this in a more compact form using l_j:

$$= \sum_{j=1}^{n} A_{ij} (1 + \alpha l_j). \tag{3.19}$$

Let k_i^{in} be the **in-degree** of vertex i (the number of edges leading to vertex i). Then we can rewrite (3.19) as follows:

$$= k_i^{\text{in}} + \alpha \sum_{j=1}^{n} A_{ij} l_j. \tag{3.20}$$

The total score for a team is the difference $s_i = w_i - l_i$. Teams are ranked on the basis of their total score. Notably, a win against a strong team rewards a team heavily, and a loss against a weak team punishes a team heavily.

We can rearrange (3.16) and (3.20) using vector notation. We set the following:

$$\mathbf{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}, \tag{3.21}$$

$$\mathbf{l} = \begin{bmatrix} l_1 & l_2 & \cdots & l_n \end{bmatrix}, \tag{3.22}$$

$$\mathbf{k}^{\text{out}} = \begin{bmatrix} k_1^{\text{out}} & k_2^{\text{out}} & \cdots & k_n^{\text{out}} \end{bmatrix}, \tag{3.23}$$

$$\mathbf{k}^{\text{in}} = \begin{bmatrix} k_1^{\text{in}} & k_2^{\text{in}} & \cdots & k_n^{\text{in}} \end{bmatrix}. \tag{3.24}$$

We can now write (3.16) and (3.20) as

$$\mathbf{w} = \mathbf{k}^{\text{out}} + \alpha \mathbf{A}^T \cdot \mathbf{w} \tag{3.25}$$

$$\mathbf{l} = \mathbf{k}^{\text{in}} + \alpha \mathbf{A} \cdot \mathbf{l}. \tag{3.26}$$

We rearrange (3.25) and (3.26) as follows. First, we subtract $\alpha \mathbf{A}^T \cdot \mathbf{w}$ from both sides of (3.25) and $\alpha \mathbf{A} \cdot \mathbf{l}$ from both sides of (3.26), yielding:

$$\mathbf{w} - \alpha \mathbf{A}^T \cdot \mathbf{w} = \mathbf{k}^{\text{out}} \tag{3.27}$$

$$\mathbf{l} - \alpha \mathbf{A} \cdot \mathbf{l} = \mathbf{k}^{\text{in}}. \tag{3.28}$$

Factoring out the common vector on each lefthand side, we have:

$$(I - \alpha \mathbf{A}^T)\mathbf{w} = \mathbf{k}^{\text{out}} \tag{3.29}$$

$$(I - \alpha \mathbf{A})\mathbf{l} = \mathbf{k}^{\text{in}}. \tag{3.30}$$

We then take the matrix inverse of the matrices on each lefthand side and premultiply both sides of both equations by them, yielding:

$$\mathbf{w} = (I - \alpha \mathbf{A}^T)^{-1} \cdot \mathbf{k}^{\text{out}} \tag{3.31}$$

$$\mathbf{l} = (I - \alpha \mathbf{A})^{-1} \cdot \mathbf{k}^{\text{in}}. \tag{3.32}$$

Now we deal with the parameter α. Firstly, we note that larger values of α place more weight on indirect wins relative to direct wins, while smaller values place more weight on direct wins.

Park and Newman (2005) note there are, in general, limits on the values α can take, claiming that it is straightforward to show that the series in (3.16) and (3.20) converge only if $\alpha < \lambda_{max}^{-1}$, where λ_{max} is the largest eigenvalue of the adjacency matrix $\mathbf{A}$. If the network is *acyclic*, then the largest eigenvalue is zero, which imparts no limit on the value of α. The authors find this has never happened in the data they observe, with the normal upperbound value on α being between 0.2 and 0.3.

The authors find empirical performance works well setting $\alpha = 0.8\lambda_{max}^{-1}$, but they note that this is for retrodictive performance, since it requires knowing the full schedule.

We re-compute α based on the adjacency matrix at each epoch of the ranking period to best-reflect the information available. We use the following to determine α for our testing:

$$\alpha = \begin{cases} 0.2 & \lambda_{max} = 0 \\ 0.8\lambda_{max}^{-1} & \text{otherwise.} \end{cases} \tag{3.33}$$

This is the same choice of α as used originally in Park and Newman (2005) and will ensure convergence and good baseline performance.

3.2.5.2 Motegi and Masuda (2012) Dynamic Model

The dynamic network ranking model was introduced by Motegi and Masuda (2012). What follows is an explanation of this model. Specifically, this model extends the win-lose score of Park and Newman (2005) dynamically.

The logic behind the modification is as follows. The original, static network model does not account for changes in team strength throughout the ranking period. For example, if team i defeated team j early in the ranking period when team j was weak, they are subsequently entitled to a portion of all future wins of team j. However, if team j subsequently became rather strong, the win that team i gained over team j, and the indirect wins they accrue because of this win, do not properly reflect the value of the win, since team i won when team j was weak.

To proceed, two assumptions are made.

Assumption 3.2.0.2. *The increment of the win score of team i through i's win against team j depends on j's win score at that moment. It does not explicitly depend on j's score in the past or future.*

Assumption 3.2.0.3. *Each team's win and lose scores decay exponentially in time.*

Let A_t be our win-lose adjacency matrix for the game that occurs at time t ($1 \leq t \leq t_{max}$). The original work sets t to a resolution of one day, though that work examined tennis, which revolves around tournament play over a short number of days. For our application in professional soccer, we consider t with a resolution of one match week.

If team j wins against team i at time t, we set the (i, j) element of the matrix A_t to be 1 (note: total wins in a gameweek can then be found by summing the rows). All other elements of A_t are set to 0. We continue to populate the matrix A_t for all other games played in the period t before doing any subsequent computations.

We define the dynamic win score at time t in vector form, denoted by $\mathbf{w}_t$ as follows:

$$\mathbf{w}_t = W_t^\top \mathbf{1}, \tag{3.34}$$

where $W_t^\top$ is defined as follows:

$$\begin{aligned} W_t = A_t &+ e^{-\beta(t-(t-1))} \sum_{m_n \in \{0,1\}} \alpha^{m_n} A_{(t-1)} A_t^{m_n} \\ &+ e^{-\beta(t-(t-2))} \sum_{m_{n-1}, m_n \in \{0,1\}} \alpha^{m_{n-1}+m_n} A_{(t-2)} A_{(t-1)}^{m_{n-1}} A_t^{m_n} \\ &+ \cdots + e^{-\beta(t-1)} \sum_{m_2,\ldots,m_n \in \{0,1\}} \alpha^{\sum_{i=2}^{n} m_i} A_1 A_{t_2}^{m_2} \cdots A_t^{m_n}. \end{aligned} \tag{3.35}$$

We now provide an interpretation to (3.35). Firstly, α is the same as the original term in Park and Newman (2005); it is the weight of an indirect win. However, it should be noted that, unlike in that static method, α is fixed at the start of the ranking procedure (as opposed to being scaled off the eigenvalues of adjacency matrix A_t). While not stated in explicitly in Motegi and Masuda (2012), α will be too large if scaled as in the classic, static method, since A_t is reset each period (leading to α values around 1.5, at least in our testing). In line with Motegi and Masuda (2012), $\alpha \in (0, 0.2]$ is chosen, and our precise parameter choice is indicated whenever we present specific numerical examples.

Next, β, where β is a nonnegative real number, is the decay rate of the score. The first term in (3.35), A_t is the effect of the direct win at time t. The second term, being a sum, consists of two contributions to the win score. Firstly, for $m_n = 0$, the quantity inside the summation represents a direct win at time $(t-1)$, which is decayed by $e^{-\beta(t-(t-1))}$. For

$m_n = 1$, the quantity represents the indirect wins. The (i, j) element of $A_{t_{n-1}} A_t$ is positive if and only if player j wins against a player k at time t and k wins against i at time $(t-1)$. Player j gains $e^{-\beta(t-t_{n-1})} \cdot \alpha$ win score from this, assuming only one win was earned by k over i at time $(t-1)$.

Now we consider the third term, which covers four different cases. For $m_{n-1} = m_n = 0$, the quantity inside the summation represents the direct win at t_{n-2}, which is decayed by $e^{-\beta(t-(t-2))}$. For $m_{n-1} = 0$ and $m_n = 1$, the quantity inside the summation represents the indirect win based on the games at $(t-2)$ and t, resulting in additional $e^{-\beta(t-t_{n-2})} \cdot \alpha$ decayed indirect win. For $m_{n-1} = 1$ and $m_n = 0$, the quantity inside the summation represents the indirect win based on games played at $(t-2)$ and $(t-1)$ which becomes an additional $e^{-\beta(t-t_{n-2})} \cdot \alpha$ decayed win. The final case is $m_{n-1} = 1$ and $m_n = 1$. This represents the second-degree indirect wins in period $(t-2)$ resulting from indirect wins in period $(t-1)$ due to wins in period t. These indirect wins are decayed by the usual factor of $e^{-\beta(t-t_{n-2})}$ and then weighted by α^2 since they are second-degree. The j column of the third term (the result of the sum of the four cases) accounts for the effect of j's direct and indirect wins at time $(t-2)$.

More simply, what the approach does is allow teams to accrue discounted indirect wins and losses from teams they win or lose to, but only those wins and losses that happened before the win or loss result to that team. For example, consider teams A, B, and C. In the first period, A defeats B. In the second period, B defeats C. In the third period, C defeats A. In the static system, the win scores will all be the same since A gets credit for B's subsequent win over C, and B gets credit for C's subsequent win over A. In the dynamic model, A no longer gets any credit for B's win over C, since that did not occur by the time A defeated B, nor does B get credit for C's win over A, for the same reason.

Now we reconsider (3.35) with the aim of finding an update equation. Namely, we seek to represent W_t in terms of $W_{(t-1)}$. With some factoring, we find the following:

$$
\begin{aligned}
W_t =& A_t + e^{-\beta(t-t_{n-1})}\Bigg[A_{(t-1)}+ \\
& e^{-\beta((t-1)-(t-2))}\sum_{m_{n-1}\in\{0,1\}} \alpha^{m_{n-1}} A_{(t-2)} A_{(t-1)}^{m_{n-1}} + \cdots \\
& + e^{-\beta((t-1)-(t-2))}\sum_{m_2,\ldots,m_{n-1}} \alpha^{\sum_{i=2}^{n-1} m_i} A_1 A_{t_2}^{m_2} \cdots A_{(t-1)}^{m_{n-1}}\Bigg] \qquad (3.36) \\
& \cdot \sum_{m_n\in\{0,1\}} \alpha^{m_n} A_t^{m_n} \\
& = A_t + e^{-\beta(t-t_{n-1})} W_{(t-1)}(I+\alpha A_t). \qquad (3.37)
\end{aligned}
$$

From (3.34) and (3.37) we get the following update equation for the dynamic win score:

$$
\mathbf{w}_t = \begin{cases} A_1^\top \mathbf{1}, & n = 1 \\ A_t^\top \mathbf{1} + e^{-\beta(t-(t-1))}\Big(I + \alpha A_{tn}^\top\Big)\mathbf{w}_{(t-1)}, & n > 1. \end{cases} \qquad (3.38)
$$

Noting that our matrices of win-lose scores contain the necessary information we need to compute lose scores by simply transposing the matrices in (3.38), we obtain the update equation for our loss score $\mathbf{l}_t$ as follows:

$$
\mathbf{l}_t = \begin{cases} A_1 \mathbf{1}, & n = 1 \\ A_t \mathbf{1} + e^{-\beta(t-(t-1))}\Big(I + \alpha A_{tn}\Big)\mathbf{l}_{(t-1)}, & n > 1. \end{cases} \qquad (3.39)
$$

The dynamic win-lose score at time t, denoted by $\mathbf{s}_t$ is given by

$$\mathbf{s}_t = \mathbf{w}_t - \mathbf{l}_t. \tag{3.40}$$

3.2.5.3 Benefits of Network Models

Before proceeding, we highlight a pair of benefits of network ranking models. The first, and most obvious, is that they are relatively simple to compute, relying largely on linear algebraic tools and making no probabilistic assumptions. The second, and ostensibly more important, is that these methods allow us to easily incorporate multiple leagues into a ranking model, provided at least one team from each league plays each other. In professional soccer, this is an attractive feature, as every year there are large, multi-league, European club soccer championships. Specifically, both the UEFA Champions League and Europa League are annual spectacles that allow European soccer fans to see how their favourite teams fare against the best teams from other countries' leagues.

As a concrete example, it is possible for, say, Arsenal F.C. of the English Premier League to play Bayern Munich of the German Bundesliga. With such a game having occurred, we could then rank all of the English Premier League teams against the German Bundesliga Teams in a seamless fashion. These ranking results could then be used for a variety of purposes, including participating in betting markets, improving player transfer pricing models, and enhancing projections of club season performance for promotion and relegation concerns.

3.3 Literature Review

Our work most-closely relates to existing work on network ranking models and competitivity graphs, and we use both Elo models and paired comparison models as comparable models for our empirical tests.

Our work contributes to the literature on network ranking models for sports. Specifically, we introduce numerous approaches to including additional information into these models. Such models were first proposed for use in ranking sports teams by Park and Newman (2005), who deploy a static network model on collegiate football. In this application, the authors found it performed comparable to a composite measure involving both statistical approaches and expert opinions. Radicchi (2011) introduces an alternative form of a static network model based on the PageRank algorithm called the "prestige score" and applies it to professional tennis data, where it is effective in matching the ATP tennis rankings. Motegi and Masuda (2012) introduce a dynamic version of the Park and Newman (2005) model and also introduce a dynamic version of the Radicchi (2011) prestige score. Both dynamic methods outperform their static counterparts when applied to professional tennis data. More recently, Abernethy (2018) examines both static and dynamic network ranking models and compares their predictive performance to FIFA's own system in international soccer. A key feature of the work is the addition of an adjustment for the importance of games based on the game type, with, for example, World Cup game wins being worth 4 wins in friendly matches. Similar to Abernethy (2018), we introduce ways to incorporate home/away and goal difference information individually and in combination to both static and dynamic network ranking models. As an example of another sport application of these types of models, Kim and Jeon (2019) apply a PageRank-based network ranking algorithm for ranking taekwondo athletes as an alternative to the current Olympic system. The applicability of PageRank-based network ranking algorithms was called into question by Zhou et al. (2020), who use a variety of sports leagues to show that the base PageRank method does not perform well when there is highly random data. However, they introduce adjustments akin to the prestige score in Radicchi (2011), which improves the results significantly; the main benefit they cite is the separation of the team's ability into win and loss scores. Other application areas include ranking baseball teams (Hyo-jun et al. (2021)), ranking sports team managers (Erkol and Radicchi (2021)), and analyzing social media traffic (Ahmad et al. (2021)).

While the static and dynamic network models form the basis for our approaches, the inspiration for our main methodological contribution comes from the literature on competitivity graphs. Criado et al. (2013) introduce the notion of a competitivity graph and teams "competing" across rankings by interchanging relative positions across consecutive rankings. See Section 3.5.3 for a formal definition. In brief, teams are defined as competing if they swap relative rank positions in sequential gameweeks. A competitivity graph is a graph drawn by connecting teams, the nodes, if they compete, with the weight of the edge connecting them being determined by the number of times they compete. We use this competition notion to define several similarity metrics for teams, which we then use to scale our network ranking approaches. Criado et al. (2014) then extend this notion of competition to what they call 'eventual' competition, wherein teams eventually compete if they can be connected by a sequence of competitions (for example, if team i competes with team j and team j competes with team k, we say that team i and team k eventually compete). Pedroche et al. (2014) extend this work on competitivity graphs to incorporate rankings with ties, where ties defined via a threshold of closeness between teams and define a multiplex network (a multilayered competitivity graph) using different types of competition. Criado et al. (2016) extend the PageRank algorithm to these multiplex networks and apply them to analyze subway networks. Pedroche and Conejero (2020) extend the framework of competitivity graphs to handle and compare incomplete rankings of teams. Tuesta et al. (2020) combine both competitivity network techniques and Data Envelopment Analysis to develop benchmarks for university ranking indicators; the goal of the work was to identify a set of universities that compete with any selected from the incumbent ranking providers.

To our knowledge, ours is the first work to directly integrate the competition concept from the competitivity graph literature into network ranking models. While these two literature streams have cited each other, it was only due to both using network techniques. Our work provides the first explicit connection and integration of both techniques.

While our work does not directly modify Elo ranking approaches, we do use a version of the Elo model as a comparable method in our empirical performance tests. As such, we briefly summarize some work in this literature. Originally formulated by Arpad Elo in Elo (1978), the system was originally used for ranking chess players. For a comprehensive coverage of how Elo is applied in sports like football, rugby, and soccer, Stefani and Pollard (2007) is a comprehensive reference. Elo methods are used as components in other analytics systems. For example, Hvattum and Arntzen (2010) use Elo models to develop metrics that are then incorporated in multinomial logistic regression models for ranking professional soccer teams. As another application, Yang et al. (2014) use Elo models as part of a pricing model for digital goods in online games. The Elo approach has been compared to betting odds, in terms of predictive accuracy, as shown in Wunderlich and Memmert (2018), where the authors then integrate betting odds information into the Elo model to improve forecasting accuracy. We feature a comparison of our proposed methods to betting odds, in terms of predictive accuracy, and we document some (unsuccessful) attempts to beat the odds.

The final class of methods we mention here are paired comparison models, like those of the Bradley-Terry model originally introduced in Bradley and Terry (1952). This model was expanded to include ties in Davidson (1970) and this is the approach we deploy in our comparisons. Many additions have been made to these models. Dynamic paired comparison models for tennis were explored in Glickman (1999), which includes an algorithm that allows for variability in the parameter estimates as a function of time. Further building on these models, Knorr-Held (2000) introduces an extension that has close connections to nonparametric smoothing methods; this model is applied to German Bundesliga data from 1996-1997 and American National Basketball Association data from 1996-1997. Different temporal dependence structures are explored in Cattelan et al. (2013), who use exponentially weighted moving average processes to model the dependence of team abilities on historical home and away results. An alternative to the dynamic Bradley-Terry model is proposed in Baker and McHale (2014) in which the parameter variation is deterministic, not stochastic. Tutz and Schauberger (2014) propose a general paired comparison model that allows for the use of additional information; as an application, they are able to show their model uses the budget of the various teams in the German Bundesliga to enhance ranking performance. While not the focus of our work, the literature on paired comparison models does highlight that dynamic models are often used in sports applications. This reinforces our choice to explore both static and dynamic approaches.

3.4 Data Description

This section describes salient features of our data. All datasets discussed below were acquired from worldfootball.net, a comprehensive resource for international and professional soccer data ("worldfoolball.net" (2021)). The datasets on the English Premier League, German Bundesliga, and French Ligue 1 contain data for the 2000/2001−2017/2018 seasons. The dataset on the Spanish La Liga contains data for the 2000/2001−2015/2016 seasons. The dataset on the Italian Serie A contains data for the 2005/2006−2017/2018 seasons. The omitted seasons are due to data quality issues (missing data and duplicates). We only include seasons for which we have all the games.

In Table 3.1, we list the number of games of each type as well as the number of goals scored in each league for the whole dataset. In Table 3.2 we include the same figures in percentage terms, to facilitate easier comparison.

League	Games	Home Wins	Draws	Away Wins	Total Goals	Home Goals	Away Goals	Goals per Game
England	6,840	3,176	1,751	1,913	18,179	10,448	7,731	2.66
Spain	6,080	2,926	1,494	1,660	16,378	9,557	6,821	2.69
Germany	5,508	2,577	1,358	1,573	15,803	9,055	6,748	2.88
Italy	4,940	2,287	1,303	1,350	13,018	7,441	5,577	2.64
France	6,692	3,092	1,909	1,691	18,013	10,395	7,618	2.69

Table 3.1: Per-League Total Number of Games and Goals

League	Home Wins	Draws	Away Wins	Home Goals	Away Goals
England	46.43%	25.60%	27.97%	57.47%	42.53%
Spain	48.13%	24.57%	27.30%	58.35%	41.65%
Germany	46.79%	24.66%	28.56%	57.30%	42.70%
Italy	46.30%	26.38%	27.33%	57.16%	42.84%
France	46.20%	28.53%	25.27%	57.71%	42.29%

Table 3.2: Per-League Game and Goal Percentages

Examining the summarized data in Tables 3.1 and 3.2 yields some interesting observations. First, we should note that the German league only has 18 teams, hence the lower overall number of games and goals. The French Ligue 1 had 18 teams in the 2000/2001 and 2001/2002 seasons, which is why it has fewer games than the English Premier League.

The Spanish La Liga has the highest percentage of home wins and home goals. Interestingly, German Bundesliga football fans are also treated to the most average goals per game. Both the Italian Serie A and the French Ligue 1 have markedly more draws in percentage terms (and in absolute terms for France) than the other leagues. This higher incidence of ties will influence the performance of the ranking methods considered, as most, if not all the methods, either struggle to predict ties (network models) or over-weight ties (Bradley-Terry). This is discussed more thoroughly in Section 3.6.5.

3.5 New Network Ranking Methods

This section describes all of the additions we make to both the Park and Newman (2005) static network model and the Motegi and Masuda (2012) dynamic network model. Section 3.5.1 explains how we incorporate home/away information into the network ranking models. Section 3.5.2 describes our approach to including goal difference information into the network ranking models. Section 3.5.3 formally introduces the ranking position interchange form of competition from Criado et al. (2013), which we use in the ensuing sections. Section 3.5.4 explains our first approach using the competition concept from Criado et al. (2013), which we call the Direct Similarity Approach. Section 3.5.5 presents an alternative formulation to the Direct Similarity Approach, which we call the Mean-Based Direct Similarity Approach, that modifies the approach based on an average. Section 3.5.6 explains our second approach using the competition concept from Criado et al. (2013), which we call the Matched Set Similarity Approach. Section 3.5.7 presents an alternative formulation to the Matched Set Similarity Approach, which we call the Mean-Based Matched Set Similarity Approach, that modifies the approach based on an average. Section 3.5.8 explains our third and final approach using the competition concept from Criado et al.

(2013), which we call the Unmatched Set Similarity Approach. Section 3.5.9 presents an alternative formulation to the Unmatched Set Similarity Approach, which we call the Mean-Based Unmatched Set Similarity Approach, that modifies the approach based on an average. Section 3.5.10 lists the various combinations of features we examine, the explanations of which are relegated to Appendix B.2. Section 3.5.11 briefly compares the various similarity measures and highlights what they emphasize.

3.5.1 Home/Away Effects

We note that the network ranking models we discuss were originally developed for sports other than soccer. In particular, the Park and Newman (2005) model was initially tested on collegiate football in the U.S.A., and the Motegi and Masuda (2012) model was initially tested on professional tennis data. Tennis does not have the same notion of a home or away game as soccer, since games are played in a series of tournaments in various hosting countries and venues; a player does not have a home stadium. It should be noted that tennis players do receive additional encouragement from their countrymen, but this is a different phenomenon. However, in sports that do have this distinction, this home/away status is valuable information. It is common wisdom to assume that there is an inherent advantage to being the home team in a professional soccer game, which is supported by the data presented in Section 3.4. As such, home wins and losses should be treated differently than away wins and losses. Here, we introduce a way to incorporate this information in network models.

To better elucidate our approach, we deploy a running example. Consider a soccer league with four teams: A, B, C, and D. Suppose that in the first gameweek team A plays team B at A's home stadium and team C plays team D at team D's home stadium. As such, both teams B and team C are playing away. Now, suppose that team B wins against team A and team C ties team D.

In the Park and Newman (2005) static model, our adjacency matrix for this gameweek is the following (where the rows and columns correspond to the teams in alphabetical order):

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0 \end{bmatrix}. \tag{3.41}$$

For the Motegi and Masuda (2012) model, we simply index the adjacency matrix by time forming A_1, in which case $A_1 = A$ from (3.41), at least in the first period (we recall that the static model uses a single adjacency matrix constructed iteratively each gameweek).

The most straightforward approach to accounting for differences in home and away wins is to construct our adjacency matrix A as in (3.41), but instead of adding 1 for an away win and 0.5 for an away tie, we introduce a parameter $k_A > 1$ and add k_A and $0.5k_A$ instead. To demonstrate, our modified adjacency matrix is as follows:

$$A = \begin{bmatrix} 0 & k_A & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5k_A & 0 \end{bmatrix}. \tag{3.42}$$

For the dynamic model, $A_t = A$ from (3.42). This approach implies that home losses and home ties are worse than away losses and away ties, which also holds with conventional wisdom.

For our base case testing, we use a value of $k_A = 1.3$ (we explore more parameter values for a subset of methods in Section 3.6.4).

3.5.2 Goal/Difference Effects

Another element to consider in a professional soccer game is the goal difference. Specifically, while any given game has three outcome possibilities (i.e. home win, draw, away win), there is nuance in home and away wins. For example, if team i defeats team j with a final scoreline of $1-0$, we may believe that the game is a relatively close contest. However, if team i defeats team j with a final scoreline of $4-0$, we may believe that team i handily defeated team j. Given that we are ranking teams, the difference in team ability implied by the scoreline is valuable information to include in our ranking model.

We return to our running example with four teams: A, B, C, and D. Suppose that in the first gameweek team A plays team B and team C plays team D. Now, suppose that team B wins against team A with a scoreline of $4-2$ and team C ties team D. For the game between teams A and B, the goal difference is 2.

We can incorporate this information into the adjacency matrix in a similar fashion as our home-away information. We introduce a parameter, $k_D > 1$; whenever team i scores a win against team j we add

$$(k_D)^{d-1}$$

to the jith element of A, where d is the goal difference. As such, instead of the adjacency matrix in (3.41), we use the following:

$$A = \begin{bmatrix} 0 & k_D^1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0 \end{bmatrix}. \tag{3.43}$$

For the dynamic model, $A_t = A$ in (3.43). The only difference in the static and dynamic approaches, as far as adjacency matrices are concerned, is that the static model constructs a cumulative adjacency matrix over the gameweeks, where the dynamic model constructs a new adjacency matrix per gameweek (which are subsequently discounted as the ranking period proceeds). Notice that this does not affect the treatment of ties. This has the effect of making non-close losses more penalizing as well, which is ideal, since such games indicate a more noticeable ability gap between the two teams.

For our testing, we use a value of $k_D = 1.3$. We consider other values for this parameter in Section 3.6.4.

We note that it is is possible to combine both home/away and goal difference information. We explain how to do this in Appendix B.2.

3.5.3 Ranking Crossings as Competition

So far, we consider two adjustments to the basic static and dynamic network ranking models, and both adjustments focus on the type of match, be it location or result. Next, we turn our attention to α. The base model we have considered thus far assumes that indirect wins and losses are accrued because an indirect win or loss would reflect a similar result if the indirectly connected teams faced off in direct confrontation. Reflecting the strength of this assumption, both the static and dynamic network approach use a common α for all teams, which treats all indirect wins and losses the same, in terms of attribution. However, this assumption is predicated on the teams being closely matched or competitive with one another. If the teams are not similar or competitive, then the assumption is called into question; it may be the case that the direct confrontation is less likely to go as predicted by the indirect confrontations.

To weaken this assumption, we propose scaling α according to how similar the two teams under consideration are. As indirect wins and losses pass through the network, each team acts as a filter, where the attribution of indirect wins and losses will increase or decrease according to how similar the teams are along each edge in the network. We posit that this will better-reflect the relationship between direct and indirect confrontations and will yield superior accuracy.

Consulting the literature, Criado et al. (2013) provide a convenient and intuitive metric. To understand their metric, first consider two vectors of team rankings, one for game week t, and one for game week $t+1$, which we denote by $\mathbf{r}_t$ and $\mathbf{r}_{t+1}$ respectively, where the row number is the rank of the team in that row. Criado et al. (2013) define two teams as **competing** if they exchange relative positions from ranking $\mathbf{r}_t$ to $\mathbf{r}_{t+1}$. As a specific example, if team A was ranked better than team B in ranking $\mathbf{r}_t$ and then team B was ranked better than team A in ranking $\mathbf{r}_{t+1}$, we say A and B competed one time, and they are thereby mutual competitors. Note: the metric requires that competition happen between consecutive rankings, so if the above example occurred between rankings $\mathbf{r}_t$ and $\mathbf{r}_{t+2}$, it would not count as an instance of competition. This notion is intuitively appealing, since we would expect teams that are closer in terms of capability to oscillate above and below one another in a ranking, while teams that are very different would settle in different regions of the ranking list.

It is important to note that to measure competition in this fashion, we must have a minimum of two rankings. If we have only one ranking, there is no opportunity to have teams cross each other. As such, for the first two gameweeks, our proposed approaches reduce to either the static Park and Newman (2005) model or the dynamic Motegi and Masuda (2012) model, if they are static or dynamic models respectively. This allows our approaches to be applied to leagues that, for example, have only two gameweeks. We call these first two ranking periods our calibration period, which allows us to generate potential instances of competition. Starting in the third ranking period, we implement our new approaches.

For use in later sections, let C^t be a matrix where the ijth entry is the number of times teams i and j have crossed in the rankings from periods 1 through $t-1$ inclusive. This matrix is, therefore, computed before we rank the teams based on results in period t.

3.5.4 Direct Similarity Approach

Our first similarity approach is a relatively simple one. Specifically, we consider the normalized number of competitions that occur between every pair of teams. We call this approach the Direct Similarity Approach to contrast it with our subsequent methods, which consider more complicated, indirect measures of similarity.

Common to both the static and dynamic model variants of the Direct Similarity Approach is the creation of a matrix that stores the pairwise similarity metrics. Recall that, at time t, the greatest potential number of rank crossings (competition instances) is $t-1$. We form a matrix D^t where each component d_{ij}^t is equal to:

$$d_{ij}^t = \frac{C_{ij}^t}{t-1}. \tag{3.44}$$

We now need to scale this matrix, since, as pointed out in Park and Newman (2005), the sums composing the win and loss scores do not converge in the static model if the discount factor is larger than α. As such, we must scale D^t appropriately. To preserve the relative strength of the similarity metric, we first divide all elements of D^t by the maximum element of this matrix. Given that D^t is a nonnegative matrix, this division normalizes the entries between 0 and 1. We then multiply the matrix by α, which then scales α according to the relative strength of the similarity metric.

Formally, we denote the maximum element of D^t as $D^t_{\max}$. We then let $\gamma D^t_{\max} = \alpha$, so $\gamma = \frac{\alpha}{D^t_{\max}}$ is the desired factor. Let $\tilde{D}^t = \gamma D^t$.

This matrix is used in both the static model of Section 3.5.4.1 and the dynamic model of Section 3.5.4.2

3.5.4.1 Static Model

We use the matrix $\tilde{D}^t$ to adjust the static model of Park and Newman (2005). We compute the adjusted win score as:

$$\begin{aligned} w_i^{t+1} = &\sum_{j=1}^{n} A_{ji} + \sum_{j=1}^{n}\sum_{k=1}^{n} A_{kj}\tilde{D}^t_{ji}A_{ji} \\ &+ \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{h=1}^{n} A_{hk}\tilde{D}^t_{kj}A_{kj}\tilde{D}^t_{ji}A_{ji} + \cdots . \end{aligned} \tag{3.45}$$

Note: we have removed α, as we use multiple degrees of similarity as a discount factor.

We see from the form of (3.45), we scale indirect wins according to how competitive all the teams in the degree of indirection are.

We propose a similar scheme for losses:

$$
\begin{aligned}
l_i^{t+1} = \sum_{j=1}^{n} A_{ij} + \sum_{j=1}^{n}\sum_{k=1}^{n} A_{ij}\tilde{D}_{i,j}^{t} A_{jk} \\
+ \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{h=1}^{n} A_{ij}\tilde{D}^{ij} A_{jk}\tilde{D}_{j,k}^{t} A_{kh} + \cdots .
\end{aligned}
\tag{3.46}
$$

In the same fashion as the Park and Newman (2005) static network ranking model, we then compute the score via:

$$
s_i^t = w_i^t - l_i^t. \tag{3.47}
$$

We then rank the teams according to the scores.

3.5.4.2 Dynamic Model

We use the matrix $\tilde{D}^t$ to adjust the dynamic model of Motegi and Masuda (2012). In the following, note that $\circ$ denotes the Hadamard product, which is the element-wise matrix product. We define the dynamic win score at time t in vector form, denoted by $\mathbf{w}_t$ as follows:

$$
\mathbf{w}_t = W_t^{\top}\mathbf{1}, \tag{3.48}
$$

where $W_t^\top$ is defined as follows:

$$\begin{aligned} W_t = A_t &+ e^{-\beta(t-(t-1))} \sum_{m_n \in \{0,1\}} A_{(t-1)} (\tilde{D}^t \circ A_t)^{m_n} \\ &+ e^{-\beta(t-(t-2))} \sum_{m_{n-1}, m_n \in \{0,1\}} A_{(t-2)} (\tilde{D}^{t-1} \circ A_{(t-1)})^{m_{n-1}} (\tilde{S}^t \circ A_t)^{m_n} \\ &+ \cdots + e^{-\beta(t-1)} \sum_{m_2, \ldots, m_n \in \{0,1\}} A_1 (\tilde{D}^2 \circ A_2)^{m_2} \cdots (\tilde{D}^t \circ A_t)^{m_n}. \end{aligned} \tag{3.49}$$

The lose score is defined analagously:

$$\mathbf{l}_t = L_t \mathbf{1}, \tag{3.50}$$

where L_t is defined as follows:

$$\begin{aligned} L_t = A_t^\top &+ e^{-\beta(t-(t-1))} \sum_{m_n \in \{0,1\}} A_{(t-1)}^\top (\tilde{D}^t \circ A_t^\top)^{m_n} \\ &+ e^{-\beta(t-(t-2))} \sum_{m_{n-1}, m_n \in \{0,1\}} A_{(t-2)}^\top (\tilde{D}^{t-1} \circ A_{(t-1)}^\top)^{m_{n-1}} (\tilde{D}^t \circ A_t^\top)^{m_n} \\ &+ \cdots + e^{-\beta(t-1)} \sum_{m_2, \ldots, m_n \in \{0,1\}} A_1^\top (\tilde{D}^2 \circ A_2^\top)^{m_2} \cdots (\tilde{D}^t \circ A_t^\top)^{m_n}. \end{aligned} \tag{3.51}$$

Our total score is defined as:

$$\mathbf{s}_t = \mathbf{w}_t - \mathbf{l}_t. \tag{3.52}$$

3.5.4.2.1 Implementation Details To implement this algorithm, we could attempt to compute the W_t and L_t matrices at each iteration using the above formulae. However, when implementing this approach, the number of terms to consider grows at an exponential rate, leading to computational issues. As such, it is useful to form an update equation, as the original dynamic model used.

Note: for implementation purposes, we would be unable to compute D^t when $t < 3$, as we need at least two ranking periods to measure competitions. However, if we set $D^2 = I$ (meaning the similarity effects are not used for the first two periods, since they cannot be), where I is the identity matrix, then the following update equations hold:

$$\mathbf{w}_t = \begin{cases} A_1^\top \mathbf{1}, & t = 1 \\ A_t^\top \mathbf{1} + e^{-\beta}\Big(I + (\tilde{D}^t \circ A_t^\top)\Big)\mathbf{w}_{(t-1)}, & t > 1. \end{cases} \tag{3.53}$$

$$\mathbf{l}_t = \begin{cases} A_1 \mathbf{1}, & t = 1 \\ A_t \mathbf{1} + e^{-\beta}\Big(I + (\tilde{D}^t \circ A_t)\Big)\mathbf{l}_{(t-1)}, & t > 1. \end{cases} \tag{3.54}$$

Using (3.53) and (3.54) allows for much faster computation.

3.5.5 Mean-Based Direct Similarity Approach

We propose an alternative formulation of the Direct Similarity Approach. Note that the denominator of (3.44) grows as t increases. If a league reaches a state of relative stability early in the ranking period, it could be that minimal competitions occur over the remaining gameweeks. This may cause the average Direct Similarity to decrease over time. To investigate what happens when the denominator scales with the average number of competitions, we change the denominator.

Given that C^t is symmetric, all of the direct competition information is contained in either the the upper or lower triangular portion of the matrix. As such, we define the following as a metric of average pairwise competition at epoch t:

$$\bar{d}_t = \frac{\sum_{i=1}^{n} \sum_{j \geq i}^{n} C_{ij}^t}{(\sum_{k=1}^{n} k) - n} \tag{3.55}$$

where the numerator is the sum of the elements in the upper triangular portion of the matrix C^t and the denominator is the number of elements in the upper triangular portion minus the number of diagonal elements. We can simplify this to:

$$\bar{d}_t = \frac{\sum_{i=1}^{n-1} \sum_{j \geq i+1}^{n} C_{ij}^t}{\sum_{k=1}^{n-1} k}. \tag{3.56}$$

Next, we create our similarity metric matrix $\bar{D}^t$, which is defined as:

$$\bar{D}^t = \frac{1}{\bar{d}_t} C^t. \tag{3.57}$$

In Section 3.5.5.1, we demonstrate how we use the matrix $\bar{D}^t$ to modify the static Park and Newman (2005) model. In Section 3.5.5.2, we use $\bar{D}^t$ to modify the dynamic Motegi and Masuda (2012) model.

3.5.5.1 Static Model

Here we use $\bar{D}^t$ to modify the static Park and Newman (2005) model. We can use this matrix exactly the same as we do our original Direct Similarity Approach matrix, but like that approach, we have to normalize the matrix according to the value of α, otherwise the win and lose scores will not converge. We take the maximum element of $\bar{D}^t$, which we denote as $\bar{D}^t_{\max}$. We then let $\gamma \bar{D}^t_{\max} = \alpha$, so $\gamma = \frac{\alpha}{\bar{D}^t_{\max}}$ is the desired factor. Let $\tilde{\bar{D}}^t = \gamma \bar{D}^t$. The computations are the same as in Section 3.5.4.1.

3.5.5.2 Dynamic Model

Here we use $\bar{D}^t$ to modify the dynamic Motegi and Masuda (2012) model. Unlike the static Mean-Based Direct Similarity Approach, the dynamic model for Mean-Based Direct Similarity will no longer feature normalization. We used normalization for the static models largely out of computational necessity. The dynamic models have finitely many terms, so convergence is not an issue. We choose to not use normalization so we can emphasize any benefit or detriment of the adaptive denominator. As such, we define the following matrix, which we use analogously to the matrix $\tilde{D}^t$ in Section 3.5.4.2:

$$\bar{D}^t_\alpha = \alpha \bar{D}^t. \tag{3.58}$$

3.5.6 Matched Set Similarity Approach

So far, we consider direct similarity, as measured by the number of times teams compete across rankings. However, it is possible to define metrics of indirect similarity. In particular, if we wish to determine how similar two teams are to one another, one natural basis of comparison to consider is the set of teams both teams compete with. Here, we take inspiration from the notion proposed in Criado et al. (2014) of **eventual competition**, where two teams i and j eventually compete if they can be connected by a sequence of competing teams (e.g. i competes with x, x competes with y, and y competes with j).

To use this idea in a similarity metric, we take the number of common teams they compete with and enumerate them. Let $I^t_{ij}(k)$ be the set of teams that both teams i and j compete with k times by epoch t. We call this the "matched" approach as both team i and team j compete with each team in the set k times. The cardinality of this set, denoted by $|I^t_{ij}(k)|$ is, therefore, the number of teams that both i and j compete with precisely k

times by epoch t. Note that we treat k strictly; a team that i and j both compete with k times by epoch t is only used once in this computation; it is not counted again when we consider, say, $k-1$ competitions.

Let n be the number of teams we are ranking. We define a similarity index between teams based on this notion of competition as follows:

$$m_{ij}^{t} = \frac{\sum_{k=1}^{t-1} k \cdot |I_{ij}^{t}(k)|}{(t-1)\cdot(n-1)}, \tag{3.59}$$

where the numerator computes the number of mutual competitors per number of competitions k and the denominator denotes the maximum possible number of competitions multiplied by the maximum number of competitors. We can collect all of these scores into a matrix, which we denote as M^t. As noted earlier, convergence can be an issue. As such, we can scale the entries of M^t with respect to α. Specifically, at every epoch, we recompute α using our adjacency matrix, then we compute M^t. We find the maximum entry of M^t, which we denote $M_{\max}^t$. We then find $\gamma_t = \alpha / M_{\max}^t$. We define our adjusted similarity score matrix as:

$$\tilde{M}^t = \gamma_t M^t. \tag{3.60}$$

In Section 3.5.6.1, we show how we use this matrix to adjust the Park and Newman (2005) static ranking model. In Section 3.5.6.2, we show how we use this matrix to adjust the Motegi and Masuda (2012) dynamic ranking model.

3.5.6.1 Static Model

In a similar fashion to our Direct Similarity Approach, we now use the $\tilde{M}^t$ matrix to adjust the Park and Newman (2005) static ranking model. We compute our win score as follows:

$$\begin{aligned} w_i^{t+1} &= \sum_{j=1}^{n} A_{ji} + \sum_{j=1}^{n}\sum_{k=1}^{n} A_{kj}\tilde{M}_{ji}^{t}A_{ji} \\ &+ \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{h=1}^{n} A_{hk}\tilde{M}_{k,j}^{t}A_{kj}\tilde{M}_{j,i}^{t}A_{ji} + \cdots . \end{aligned} \tag{3.61}$$

We define a lose score similarly:

$$\begin{aligned} l_i^{t+1} &= \sum_{j=1}^{n} A_{ij} + \sum_{j=1}^{n}\sum_{k=1}^{n} A_{ij}\tilde{M}_{ij}^{t}A_{jk} \\ &+ \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{h=1}^{n} A_{ij}\tilde{M}_{ij}^{t}A_{jk}\tilde{M}_{jk}^{t}A_{kh} + \cdots . \end{aligned} \tag{3.62}$$

We then compute the total score as follows:

$$s_i^{t+1} = w_i^{t+1} - l_i^{t+1}. \tag{3.63}$$

We then use this total score to rank the teams.

3.5.6.2 Dynamic Model

We now introduce a dynamic approach to our Matched Set Similarity approach of Section 3.5.6.1. We define the win score as follows:

$$\mathbf{w}_t = W_t^\top \mathbf{1}, \tag{3.64}$$

where $W_t^\top$ is defined as follows:

$$\begin{aligned} W_t = A_t &+ e^{-\beta(t-(t-1))} \sum_{m_n \in \{0,1\}} A_{(t-1)} (\tilde{M}^t \circ A_t)^{m_n} \\ &+ e^{-\beta(t-(t-2))} \sum_{m_{n-1}, m_n \in \{0,1\}} A_{(t-2)} (\tilde{M}^{t-1} \circ A_{(t-1)})^{m_{n-1}} (\tilde{M}^t \circ A_t)^{m_n} \\ &+ \cdots + e^{-\beta(t-1)} \sum_{m_2, \ldots, m_n \in \{0,1\}} A_1 (\tilde{M}^2 \circ A_2)^{m_2} \cdots (\tilde{M}^t \circ A_t)^{m_n}. \end{aligned} \tag{3.65}$$

The lose score is defined analagously:

$$\mathbf{l}_t = L_t \mathbf{1}, \tag{3.66}$$

where L_t is defined as follows:

$$\begin{aligned} L_t = A_t^\top &+ e^{-\beta(t-(t-1))} \sum_{m_n \in \{0,1\}} A_{(t-1)}^\top (\tilde{M}^t \circ A_t^\top)^{m_n} \\ &+ e^{-\beta(t-(t-2))} \sum_{m_{n-1}, m_n \in \{0,1\}} A_{(t-2)}^\top (\tilde{M}^{t-1} \circ A_{(t-1)}^\top)^{m_{n-1}} (\tilde{M}^t \circ A_t^\top)^{m_n} \\ &+ \cdots + e^{-\beta(t-1)} \sum_{m_2, \ldots, m_n \in \{0,1\}} A_1^\top (\tilde{M}^2 \circ A_2^\top)^{m_2} \cdots (\tilde{M}^t \circ A_t^\top)^{m_n}. \end{aligned} \tag{3.67}$$

Our total score is defined as:

$$\mathbf{s}_t = \mathbf{w}_t - \mathbf{l}_t. \tag{3.68}$$

3.5.6.2.1 Implementation Details Similar to the Direct Similarity Approach, we introduce update equations to allow for faster computation. We allow $M^2 = I$, where I is the identity matrix; the following update equations then hold:

$$\mathbf{w}_t = \begin{cases} A_1^\top \mathbf{1}, & t = 1 \\ A_t^\top \mathbf{1} + e^{-\beta}\Big(I + (\tilde{M}^t \circ A_t^\top)\Big)\mathbf{w}_{(t-1)}, & t > 1. \end{cases} \tag{3.69}$$

$$\mathbf{l}_t = \begin{cases} A_1 \mathbf{1}, & t = 1 \\ A_t \mathbf{1} + e^{-\beta}\Big(I + (\tilde{M}^t \circ A_t)\Big)\mathbf{l}_{(t-1)}, & t > 1. \end{cases} \tag{3.70}$$

Using (3.69) and (3.70) allows for much faster computation.

3.5.7 Mean-Based Matched Set Similarity Approach

Similar to Section 3.5.5, we consider an alternative formulation to our Matched Set Similarity approach. Specifically, we introduce a new denominator that is not explicitly tied to the gameweek t. The numerator is the same as in Section 3.5.5.

Let n be the number of teams we are ranking. We can compute an average of matched set pairwise competitions via:

$$\bar{m}^t = \frac{\sum_{i=1}^{n-1} \sum_{j \geq i+1}^{n} \sum_{k=1}^{t-1} k \cdot |I_{ij}^t(k)|}{\sum_{i=1}^{n-1} i}. \tag{3.71}$$

The numerator in (3.71) computes the sum of the number teams each pair of teams both compete with k times, multiplied by k, where k goes from 1 through $t-1$ inclusive. The denominator in (3.71) is the number of pairs of teams. We then define the matrix $\bar{M}^t$, where each entry m_{ij} is defined as:

$$m_{ij} = \frac{\sum_{k=1}^{t-1} k \cdot |I_{ij}^t(k)|}{\bar{m}^t}. \tag{3.72}$$

In Section 3.5.7.1, we briefly explain how this alternative matrix $\bar{M}^t$ is used in the static model. In Section 3.5.7.2, we describe the dynamic model.

3.5.7.1 Static Model

As noted before, the static model requires normalization to ensure convergence. We take the maximum element of $\bar{M}^t$, which we denote as $\bar{M}^t_{\max}$. We then let $\gamma \bar{M}^t_{\max} = \alpha$, so $\gamma = \frac{\alpha}{\bar{M}^t_{\max}}$ is the desired factor. Let $\tilde{\bar{M}}^t = \gamma \bar{M}^t$. The computations are the same as in Section 3.5.6.1.

3.5.7.2 Dynamic Model

Similar to the Mean-Based Direct Similarity Approach, convergence is no longer a concern, since the method relies on a finite number of terms. As such, we define the following matrix, which we use analogously to the matrix $\tilde{M}^t$ in Section 3.5.6.2:

$$\bar{M}^t_\alpha = \alpha \bar{M}^t. \tag{3.73}$$

3.5.8 Unmatched Set Similarity Approach

We now consider a relaxed version of the Matched Set Similarity Approach of Section 3.5.6. In particular, we note that we can relax the requirement that both teams i and j have competed with the common team k times in the computation. Instead, we sum the number of times i and j compete with each team in the set of common competitors, then normalize this sum. Let I_{ij} be an index set denoting the teams that both i and j compete with. We define our unmatched similarity metric (where "unmatched" is used since, in this case, we do not care if teams i and j have competed with a team the same number of times) as:

$$u_{ij}^t = \frac{\sum_{k \in I_{ij}} C_{ki} + \sum_{k \in I_{ij}} C_{kj}}{2|I_{ij}|(t-1)} \text{ if } |I_{ij}| > 0, \tag{3.74}$$

$$= 0 \text{ if } |I_{ij}| = 0. \tag{3.75}$$

We collect all of these scores into a matrix, which we denote as U^t. To handle convergence concerns, we scale the entries of U^t with respect to α. Specifically, at every epoch, we recompute α using our adjacency matrix, then we compute U^t. We find the maximum entry of U^t, which we denote U^t_{max}. We then find $\gamma_t = \alpha / U^t_{\text{max}}$. We define our adjusted similarity score matrix as:

$$\tilde{U}^t = \gamma_t U^t. \tag{3.76}$$

In Section 3.5.8.1 we show how $\tilde{U}^t$ is used in the static model. In Section 3.5.8.2 we show how $\tilde{U}^t$ is used in the dynamic model.

3.5.8.1 Static Model

Using $\tilde{U}^t$, we compute our win score as follows:

$$\begin{aligned} w_i^{t+1} &= \sum_{j=1}^{n} A_{ji} + \sum_{j=1}^{n}\sum_{k=1}^{n} A_{kj}\tilde{U}_{ji}^{t} A_{ji} \\ &+ \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{h=1}^{n} A_{hk}\tilde{U}_{kj}^{t} A_{kj}\tilde{U}_{ji}^{t} A_{ji} + \cdots . \end{aligned} \tag{3.77}$$

We define a lose score similarly:

$$\begin{aligned} l_i^{t+1} &= \sum_{j=1}^{n} A_{ij} + \sum_{j=1}^{n}\sum_{k=1}^{n} A_{ij}\tilde{U}_{ij}^{t} A_{jk} \\ &+ \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{h=1}^{n} A_{ij}\tilde{U}_{ij}^{t} A_{jk}\tilde{U}_{jk}^{t} A_{kh} + \cdots . \end{aligned} \tag{3.78}$$

We then compute the total score as follows:

$$s_i^{t+1} = w_i^{t+1} - l_i^{t+1}. \tag{3.79}$$

We then use this total score to rank the teams.

3.5.8.2 Dynamic Model

Here we introduce a dynamic version of our Unmatched Set Similarity approach of Section 3.5.8.1. Our win score is as follows:

$$\mathbf{w}_t = W_t^{\top}\mathbf{1}, \tag{3.80}$$

where $W_t^\top$ is defined as follows:

$$
\begin{aligned}
W_t = A_t &+ e^{-\beta(t-(t-1))} \sum_{m_n \in \{0,1\}} A_{(t-1)} (\tilde{U}^t \circ A_t)^{m_n} \\
&+ e^{-\beta(t-(t-2))} \sum_{m_{n-1}, m_n \in \{0,1\}} A_{(t-2)} (\tilde{U}^{t-1} \circ A_{(t-1)})^{m_{n-1}} (\tilde{U}^t \circ A_t)^{m_n} \\
&+ \cdots + e^{-\beta(t-1)} \sum_{m_2, \ldots, m_n \in \{0,1\}} A_1 (\tilde{U}^2 \circ A_2)^{m_2} \cdots (\tilde{U}^t \circ A_t)^{m_n}.
\end{aligned}
\tag{3.81}
$$

The lose score is defined analogously:

$$
\mathbf{l}_t = L_t \mathbf{1}, \tag{3.82}
$$

where L_t is defined as follows:

$$
\begin{aligned}
L_t = A_t^\top &+ e^{-\beta(t-(t-1))} \sum_{m_n \in \{0,1\}} A_{(t-1)}^\top (\tilde{U}^t \circ A_t^\top)^{m_n} \\
&+ e^{-\beta(t-(t-2))} \sum_{m_{n-1}, m_n \in \{0,1\}} A_{(t-2)}^\top (\tilde{U}^{t-1} \circ A_{(t-1)}^\top)^{m_{n-1}} (\tilde{U}^t \circ A_t^\top)^{m_n} \\
&+ \cdots + e^{-\beta(t-1)} \sum_{m_2, \ldots, m_n \in \{0,1\}} A_1^\top (\tilde{U}^2 \circ A_2^\top)^{m_2} \cdots (\tilde{U}^t \circ A_t^\top)^{m_n}.
\end{aligned}
\tag{3.83}
$$

Our total score is defined as:

$$
\mathbf{s}_t = \mathbf{w}_t - \mathbf{l}_t. \tag{3.84}
$$

3.5.8.2.1 Implementation Details Once again, if we allow $U^2 = I$, where I is the identity matrix, then we can use the following update equations:

$$\mathbf{w}_t = \begin{cases} A_1^\top \mathbf{1}, & t = 1 \\ A_t^\top \mathbf{1} + e^{-\beta}\Big(I + (\tilde{U}^t \circ A_t^\top)\Big)\mathbf{w}_{(t-1)}, & t > 1. \end{cases} \tag{3.85}$$

$$\mathbf{l}_t = \begin{cases} A_1 \mathbf{1}, & t = 1 \\ A_t \mathbf{1} + e^{-\beta}\Big(I + (\tilde{U}^t \circ A_t)\Big)\mathbf{l}_{(t-1)}, & t > 1. \end{cases} \tag{3.86}$$

Using (3.85) and (3.86) allows for much faster computation.

3.5.9 Mean-Based Unmatched Set Similarity Approach

Similar to the Mean-Based Direct Similarity and Mean-Based Matched Set Similarity methods, we also provide an alternative formulation for the Unmatched Set Similarity Approach.

We can compute an average of unmatched set pairwise competitions via:

$$\bar{u}^t = \frac{\sum_{i=1}^{n-1} \sum_{j \geq i+1}^{n} (\sum_{k \in I_{ij}} C_{ki} + \sum_{k \in I_{ij}} C_{kj})}{\sum_{i=1}^{n-1} i}. \tag{3.87}$$

The numerator in (3.87) computes the sum of competitions each pair of teams have with teams in their set of mutual competitors I_{ij}. The denominator in (3.87) is the number of pairs of teams.

We then define the matrix $\bar{U}^t$, where each entry u_{ij} is defined as:

$$u_{ij} = \frac{\sum_{k \in I_{ij}} C_{ki} + \sum_{k \in I_{ij}} C_{kj}}{\bar{u}^t}. \tag{3.88}$$

In Section 3.5.9.1 we describe how $\bar{U}^t$ is used in the static model. In Section 3.5.9.2 we explain how $\bar{U}^t$ is used in the dynamic model.

3.5.9.1 Static Model

To handle convergence concerns, we scale the entries of $\bar{U}^t$ with respect to α. Specifically, at every epoch, we recompute α using our adjacency matrix, then we compute $\bar{U}^t$. We find the maximum entry of $\bar{U}^t$, which we denote $\bar{U}^t_{\max}$. We then find $\gamma_t = \alpha/\bar{U}^t_{\max}$. We define our adjusted similarity score matrix as:

$$\tilde{\bar{U}}^t = \gamma_t \bar{U}^t. \tag{3.89}$$

The computations are the same as in Section 3.5.8.1.

3.5.9.2 Dynamic Model

Convergence is no longer a concern, since the method relies on a finite number of terms. As such, we define the following matrix, which we use analogously to the matrix $\tilde{U}^t$ in Section 3.5.8.2:

$$\bar{U}^t_\alpha = \alpha \bar{U}^t. \tag{3.90}$$

3.5.10 Combination Methods

We consider the above approaches in isolation, but we also consider them in combination. Specifically, we consider the following combination methods in both static and dynamic form:

1. Home/Away + Goal Difference
2. Home/Away + Direct Similarity
3. Home/Away + Matched Set Similarity
4. Home/Away + Unmatched Set Similarity
5. Home/Away + Mean-Based Direct Similarity
6. Home/Away + Mean-Based Matched Set Similarity
7. Home/Away + Mean-Based Unmatched Set Similarity
8. Goal Difference + Direct Similarity
9. Goal Difference + Matched Set Similarity
10. Goal Difference + Unmatched Set Similarity
11. Goal Difference + Mean-Based Direct Similarity
12. Goal Difference + Mean-Based Matched Set Similarity
13. Goal Difference + Mean-Based Unmatched Set Similarity
14. Home/Away + Goal Difference + Direct Similarity
15. Home/Away + Goal Difference + Matched Set Similarity
16. Home/Away + Goal Difference + Unmatched Set Similarity
17. Home/Away + Goal Difference + Mean-Based Direct Similarity
18. Home/Away + Goal Difference + Mean-Based Matched Set Similarity
19. Home/Away + Goal Difference + Mean-Based Unmatched Set Similarity.

The details of these approaches are included in Appendix B.2.

3.5.11 Similarity Approach Comparison

With all of the formulae presented, it is useful to briefly compare what each similarity approach emphasizes. Note: the following also holds for the Mean-Based versions. Direct

Similarity is the most straightforward; it simply tracks the number of times teams cross in successive ranking periods, scaling indirect wins and losses based on this. Both Matched and Unmatched Set Similarity are concerned, instead, with how many crossings teams tally with a set of mutual competitors. We examine a normalized similarity score between teams that share common opponents.

On this last statement is where the Matched and Unmatched Set Similarity differ somewhat. The Matched Set Similarity can be thought of as our most strict similarity approach; in a league where teams play each other, albeit at different times, only teams that have played another team the same number of times up to a specific point in the season will count. However, in the Umatched Set Similarity, we allow for teams that have not played the same number of times to determine the similarity score.

3.6 Performance Results and Analysis

Section 3.6.1 explains our chosen performance metric for evaluating our ranking methods. Section 3.6.2 presents and discusses the performance of the various ranking methods, both static and dynamic. Section 3.6.3 contains our rank correlation analysis, which we use to see if performance discrepancies for our dynamic models in particular leagues can be explained. Section 3.6.4 contains our sensitivity analysis, in which we examine the impact of our model parameters on a subset of our ranking models. Section 3.6.5 uses composite betting odds as a final basis of comparison for our methods.

3.6.1 Performance Metric

We use a simple predictive accuracy metric for evaluating the performance of our ranking methods. For a given dataset under consideration, we step forward iteratively gameweek-by-gameweek generating new rankings after accounting for the matches played in the gameweek. We then use these rankings to predict the results of the matches in the subsequent gameweek.

For a match between team i and team j, if the rank of team i is lower than team j (where a lower rank is superior, as the best ranked team is ranked 1), then the system predicts that team i will win. If the ranks are tied, the system predicts a tie. If the rank of team i is higher than team j, then the system predicts team j will win. To compute predictive accuracy, we use the following formula:

$$\text{predictive accuracy} = \frac{\text{number of correct predictions}}{\text{number of predictions made}} \times 100\%. \tag{3.91}$$

It should be noted that there exist specific edge cases when a prediction is non-standard. In all the leagues under consideration, there exists a promotion/relegation system. Namely, at the start of a new season, some teams leave the league and teams enter the league to replace them. In the English Premier League, for example, the bottom three teams are relegated and leave the league after the season, joining the English Football League (the league below the Premier League). Three teams then join the Premier League from the English Football League. Such promotion and relegation can cause the ranking system to make a prediction about a team it has yet to see against a team it has ranked before, or two unranked teams.

If the ranking system encounters a team it has not yet ranked, the following occurs:

- If both teams are unranked, the prediction is always a tie.
- If only one team is unranked, the ranked team is predicted to win.

One may wonder why we choose predictive accuracy. Predictive accuracy, when calculated as shown in (3.91), can be interpreted as showing how well a given ranking model captures the information from match performance. In theory, if a ranking model more readily captures information and translates this information to rankings, the ranking-based prediction should better reflect the real ability of the teams, which, in turn, should lead to better predictions about future matches.

It should also be noted that this metric, or similar analogues of it, have been used in both Motegi and Masuda (2012) and Abernethy (2018). However, unlike those works, we do not remove tied games from our dataset. We choose to do so as this allows for direct comparison with betting odds in Section 3.6.5, as betting odds make predictions on all three potential match results.

For the remainder of this section, we compute average predictive accuracy from models deployed on four-season chunks of data. For example, one chunk is the 2000/2001 season through (and including) the 2003/2004 season of the English Premier League. The next chunk is the 2001/2002 season through (and including) the 2004/2005 season. We create chunks for all available data in each of our leagues. For each of these chunks, we compute the predictive accuracy of each method. To compute the averages, we simply average the predictive accuracy over all chunks. The total number of chunks differs per league, based on the data availability mentioned in Section 3.4.

3.6.2 Ranking Method Performance Comparison

In Section 3.6.2.1 we examine our static models. In Section 3.6.2.2 we consider our dynamic models. Note: in this section we focus on average performance for concision, but in Appendix B.3 we include the line plots demonstrating per-data-chunk performance for each league. Those figures, in general, support the conclusions demonstrated below.

3.6.2.1 Static Models

To begin, Table 3.3 presents both preexisting ranking methods and our home/away and goal difference network models of Sections 3.5.1 and 3.5.2 along with associated estimated standard errors. The standard error is computed as the sample standard deviation divided by the square root of the sample size; in this case, the sample size is the number of starting years we use for rolling data windows.

The Static Network model is the Park and Newman (2005) model. Note that the Static Network ($\alpha = 0$) column is the same as the Park and Newman (2005) static ranking model, but we fix $\alpha = 0$, meaning that the method does not use any indirect wins or losses. Further, the HA + GD column is a combination of the home/away and goal difference approaches.

League	League Points	Elo	Bradley-Terry	Static Network	Static Network ($\alpha = 0$)	Home /Away (HA)	Goal Difference (GD)	HA + GD
England	49.11% (0.41%)	39.68% (0.77%)	38.18% (0.60%)	48.62% (0.45%)	25.40% (0.35%)	49.08% (0.46%)	49.16% (0.45%)	49.25% (0.44%)
Spain	46.49% (0.70%)	40.24% (0.68%)	34.67% (1.00%)	46.20% (0.67%)	28.87% (0.37%)	46.51% (0.73%)	46.84% (0.68%)	46.83% (0.69%)
Germany	46.92% (0.32%)	39.73% (0.41%)	34.92% (0.67%)	46.43% (0.29%)	28.06% (0.24%)	46.93% (0.24%)	46.78% (0.29%)	46.87% (0.31%)
Italy	47.68% (0.73%)	38.25% (1.19%)	36.20% (0.77%)	47.27% (0.68%)	26.05% (0.61%)	47.54% (0.64%)	47.56% (0.63%)	47.54% (0.64%)
France	43.50% (0.59%)	36.62% (0.38%)	33.38% (0.55%)	43.05% (0.52%)	27.56% (0.26%)	43.49% (0.51%)	43.66% (0.57%)	43.77% (0.54%)

Table 3.3: Existing Approaches, Home/Away, and Goal Difference Predictive Accuracy Average of 4-season Rolling Windows

Examining Table 3.3, we immediately see the role α plays in the network models; when $\alpha = 0$, the model performs rather poorly. Further, this validates the usefulness of the indirect wins and losses concept, as the Static Network model performs markedly better when α is allowed to be as intended. We note that both the Elo and Bradley-Terry models only marginally outperform random guessing; as such, we will not consider them in the remaining tables and figures. Turning our attention to our new methods, we see that the Home/Away, Goal Difference, and HA + GD combination all outperform the Static Network model. This follows our intuition; including additional, useful information should yield better results.

Next, we consider the League Points approach, which outperforms the Static Network approach. This, unto itself, is somewhat interesting, as it shows that the relatively simple system that is in place already does a reasonably good job of capturing the strengths of the teams. We note, however, that four of the five leagues considered (all except Italy) has at least one network approach (either Home/Away, Goal Difference, or HA + GD) that outperforms the League Points approach. This is important, as the League Points approach is locally valid, but it doesn't provide a well-defined way to incorporate teams from multiple leagues into the system. In contrast, the network model approach allows for multiple leagues to be added seamlessly, provided there is at least match between teams from the included leagues.

Next, in Table 3.4, we compare the Park and Newman (2005) model with our similarity approaches from Sections 3.5.4.1, 3.5.6.1, 3.5.8.1, 3.5.5.1, 3.5.7.1, and 3.5.9.1.

League	Static Network	Direct Similarity (DS)	Mean-Based Direct Similarity (MBDS)	Matched Set Similarity (MSS)	Mean-Based Matched Set Similarity (MBMSS)	Unmatched Set Similarity (USS)	Mean-Based Unmatched Set Similarity (MBUSS)
England	48.62% (0.45%)	48.94% (0.44%)	48.94% (0.44%)	48.77% (0.47%)	48.77% (0.47%)	48.83% (0.46%)	48.80% (0.46%)
Spain	46.20% (0.67%)	46.38% (0.67%)	46.38% (0.67%)	46.31% (0.66%)	46.31% (0.66%)	46.40% (0.69%)	46.37% (0.68%)
Germany	46.43% (0.29%)	46.67% (0.26%)	46.67% (0.26%)	46.84% (0.25%)	46.84% (0.25%)	46.72% (0.28%)	46.60% (0.27%)
Italy	47.27% (0.68%)	47.27% (0.65%)	47.27% (0.65%)	47.19% (0.69%)	47.19% (0.69%)	47.25% (0.69%)	47.12% (0.62%)
France	43.05% (0.52%)	43.40% (0.51%)	43.38% (0.51%)	43.51% (0.49%)	43.51% (0.49%)	43.27% (0.50%)	43.38% (0.51%)

Table 3.4: Original Static Network vs Similarity Approaches Predictive Accuracy Average of 4-season Rolling Windows

Examining Table 3.4, we see that all of our similarity approaches outperform the Park and Newman (2005) Static Network model except for in the Italian Serie A, of which only the Direct Similarity (DS) and Mean-Based Direct Similarity (MBDS) match the performance. As one might expect, many of the mean-based methods produce very similar results to their non-mean-based counterparts. This is partially due to the normalization we introduce, which is required for the algorithms to converge, and diminishes the difference between the two classes of algorithms. In Section 3.6.2.2 covering our dynamic model results, we see that the dynamic versions, which are non-normalized, are more distinct in terms of performance.

Interestingly, the best-performing similarity approach changes per league. For the English Premier League and Italian Serie A, Direct Similarity and Mean-Based Direct Similarity were the most accurate. For the Spanish La Liga, Unmatched Set Similarity (USS) was the most accurate. For the German Bundesliga and the French Ligue 1, the Matched Set Similarity (MSS) and Mean-Based Matched Set Similarity (MBMSS) were the best-performing approaches. This suggests that for practical use, the similarity metric should vary depending on the league of interest.

Next, in Table 3.5 we consider a set of combination approaches incorporating home/away and similarity information, with our home/away approach of Section 3.5.1 serving as our basis of comparison.

League	HA	HA + DS	HA + MBDS	HA + MSS	HA + MBMSS	HA + USS	HA + MBUSS
England	49.08% (0.46%)	48.94% (0.45%)	48.94% (0.45%)	48.96% (0.45%)	48.96% (0.45%)	49.00% (0.45%)	48.76% (0.41%)
Spain	46.51% (0.73%)	46.50% (0.73%)	46.50% (0.73%)	46.55% (0.73%)	46.55% (0.73%)	46.50% (0.72%)	46.60% (0.67%)
Germany	46.93% (0.24%)	46.69% (0.28%)	46.69% (0.28%)	46.87% (0.26%)	46.87% (0.26%)	46.75% (0.25%)	46.63% (0.26%)
Italy	47.54% (0.64%)	47.34% (0.70%)	47.34% (0.70%)	47.38% (0.67%)	47.38% (0.67%)	47.47% (0.63%)	47.23% (0.65%)
France	43.39% (0.51%)	43.50% (0.53%)	43.50% (0.53%)	43.62% (0.50%)	43.62% (0.50%)	43.56% (0.48%)	43.54% (0.52%)

Table 3.5: Original Static Network + Home/Away vs Similarity Approaches + Home/Away Predictive Accuracy Average of 4-season Rolling Windows

A cursory examination of Table 3.5 indicates that there are somewhat mixed results to combining home/away and similarity approaches. For the English Premier League, the German Bundesliga, and the Italian Serie A, performance is superior when no similarity information is incorporated. For the Spanish La Liga, three approaches, the Matched Set Similarity, Mean-Based Matched Set Similarity, and Mean-Based Unmatched Set Similarity all demonstrate superior performance. The French Ligue 1 is the apparent outlier for this comparison, as all similarity approaches outperform the regular static network model with home/away information. These observations further support our conclusion that the network model choice in implementation should be league-dependent.

Next, in Table 3.6, we compare the Park and Newman (2005) model with goal difference information of Section 3.5.2 with combination approaches that use both similarity approaches and goal difference information.

League	GD	GD + DS	GD + MBDS	GD + MSS	GD + MBMSS	GD + USS	GD + MBUSS
England	49.16% (0.45%)	49.09% (0.46%)	49.09% (0.46%)	49.20% (0.44%)	49.20% (0.44%)	49.16% (0.43%)	48.96% (0.42%)
Spain	46.84% (0.68%)	46.76% (0.67%)	46.76% (0.67%)	46.76% (0.62%)	46.76% (0.62%)	46.79% (0.71%)	46.56% (0.68%)
Germany	46.78% (0.29%)	46.82% (0.33%)	46.82% (0.33%)	47.01% (0.27%)	47.01% (0.27%)	46.93% (0.29%)	46.69% (0.31%)
Italy	47.56% (0.63%)	47.42% (0.65%)	47.24% (0.65%)	47.61% (0.61%)	47.61% (0.61%)	47.49% (0.59%)	47.23% (0.57%)
France	43.66% (0.57%)	43.79% (0.52%)	43.79% (0.52%)	43.68% (0.53%)	43.68% (0.53%)	43.64% (0.53%)	43.57% (0.49%)

Table 3.6: Original Static Network + Goal Difference vs Similarity Approaches + Goal Difference Predictive Accuracy Average of 4-season Rolling Windows

In Table 3.6, we see that all leagues except the Spanish La Liga have a similarity-based approach that exhibits superior performance to the comparable non-similarity-based approach. In particular, for England, Germany, and Italy, the Matched Set Similarity and Mean-Based Matched Set Similarity Approaches demonstrated the strongest performance. For France's Ligue 1, the Direct Similarity and Mean-Based Direct Similarity were the best-performing approaches. Interestingly, for all leagues, the Mean-Based Unmatched Set Similarity was the worst performing approach; this method may be muting the effect of goal difference information in some fashion.

Our final performance table for our static model results, Table 3.7 presents the Park and Newman (2005) model with both home/away and goal difference information as the basis of comparison with the similarity approaches combined with home/away and goal difference information.

League	HA + GD	HA + GD + DS	HA + GD + MBDS	HA + GD + MSS	HA + GD + MBMSS	HA + GD + USS	HA + GD + MBUSS
England	49.25% (0.44%)	49.14% (0.42%)	49.14% (0.42%)	49.13% (0.44%)	49.13% (0.44%)	49.24% (0.45%)	48.99% (0.39%)
Spain	46.83% (0.69%)	46.77% (0.67%)	46.67% (0.67%)	46.77% (0.67%)	46.77% (0.67%)	46.77% (0.73%)	46.56% (0.71%)
Germany	46.87% (0.31%)	46.97% (0.30%)	46.97% (0.30%)	46.87% (0.29%)	46.87% (0.29%)	47.05% (0.29%)	46.80% (0.29%)
Italy	47.54% (0.64%)	47.55% (0.59%)	47.55% (0.59%)	47.63% (0.63%)	47.63% (0.63%)	47.51% (0.62%)	47.28% (0.59%)
France	43.77% (0.54%)	43.76% (0.54%)	43.76% (0.54%)	43.89% (0.54%)	43.89% (0.54%)	43.67% (0.54%)	43.68% (0.50%)

Table 3.7: Original Static Network + Home/Away + Goal Difference vs Similarity Approaches + Home/Away + Goal Difference Predictive Accuracy Average of 4-season Rolling Windows

Table 3.7 further supports the notion that the league choice and method are interlinked. For both the English Premier League and Spanish La Liga, simply including home/away and goal difference information into the Park and Newman (2005) model generates the best performance. For the German Bundesliga, the Unmatched Set Similarity with home/away and goal difference information is the clear winner. For both the Italian Serie A and French Ligue 1, the Matched Set Similarity and Mean-Based Set Similarity approaches yielded the best results in this comparison.

Across all the static network methods tested on the data used in this study, these are the best methods per league and the corresponding predictive accuracy achieved:

- England: (49.25%) HA + GD
- Spain: (46.84%) GD
- Germany: (47.05%) USS + HA + GD
- Italy: (47.63%) MSS + HA + GD and MBMSS + HA + GD

- France: (43.89%) MSS + HA + GD and MBMSS + HA + GD.

The key result to note from this is that the best method is different per league. It should be noted that these methods are not guaranteed to be the best for future performance. While we demonstrate it more thoroughly in Section 3.6.3, we note that the additional information methods (any with home/away and goal difference effects), have the effect of increasing the total score difference between teams. As shown by the above list, all leagues benefit from the inclusion of one or both types of this information.

We also note that the similarity models, in general, reduce the average total score difference amongst the teams, since they scale the effect of indirect wins and losses downwards. We note that different similarity approaches are superior for predicting different leagues. This may imply that certain underlying relationships are more important. For example, in both Italy and France, the Matched Set Similarity and Mean-Based Matched Set Similarity approaches performed best. This suggests that, given the strict nature of these approaches, only a few pairwise relationships are most important for driving the attribution of indirect wins and losses. Such differences are good motivation for future work.

3.6.2.2 Dynamic Models

To begin, Table 3.8 presents both preexisting ranking methods and our home/away and goal difference network models of Sections 3.5.1 and 3.5.2. Specifically, the Dynamic Network model is the Motegi and Masuda (2012) model.

League	League Points	Elo	Bradley-Terry	Dynamic Network	HA	GD	HA + GD
England	49.11% (0.41%)	39.68% (0.77%)	38.18% (0.60%)	48.42% (0.37%)	48.44% (0.38%)	48.98% (0.40%)	48.78% (0.43%)
Spain	46.49% (0.70%)	40.24% (0.68%)	34.67% (1.00%)	45.95% (0.79%)	46.10% (0.82%)	46.40% (0.83%)	46.43% (0.84%)
Germany	46.92% (0.32%)	39.73% (0.41%)	34.92% (0.67%)	46.23% (0.31%)	46.41% (0.27%)	46.74% (0.26%)	46.84% (0.26%)
Italy	47.68% (0.73%)	38.25% (1.19%)	36.20% (0.77%)	46.62% (0.51%)	46.84% (0.59%)	46.94% (0.57%)	47.09% (0.54%)
France	43.50% (0.59%)	36.62% (0.38%)	33.38% (0.55%)	41.92% (0.65%)	42.21% (0.60%)	42.65% (0.52%)	42.88% (0.55%)

Table 3.8: Existing Approaches, Home/Away, and Goal Difference Predictive Accuracy Average of 4-season Rolling Windows

First, we note that including home/away and goal difference information immediately improves the dynamic model noticeably in all leagues. Secondly, unlike the static model cases, none of the dynamic network models outperform the League Points approach. However, it should be noted that the dynamic network model, when incorporating both home/away and goal/difference information, has four parameters (k_A, k_D, α, and β). In our results display, we simply used comparable parameters to the static models. However, in Section 3.6.4.2 we examine a subset of the methods and explore how predictive accuracy changes when these parameters are modified.

Next, in Table 3.9, we compare the Motegi and Masuda (2012) model with our similarity approaches from Sections 3.5.4.2, 3.5.6.2, 3.5.8.2, 3.5.5.2, 3.5.7.2, and 3.5.9.2.

League	Dynamic Network	Direct Similarity (DS)	Mean-Based Direct Similarity (MBDS)	Matched Set Similarity (MSS)	Mean-Based Matched Set Similarity (MBMSS)	Unmatched Set Similarity (USS)	Mean-Based Unmatched Set Similarity (MBUSS)
England	48.42% (0.37%)	48.50% (0.36%)	46.67% (0.43%)	48.46% (0.36%)	48.45% (0.39%)	48.58% (0.38%)	48.53% (0.37%)
Spain	45.95% (0.79%)	45.43% (0.79%)	44.37% (0.59%)	45.57% (0.79%)	45.51% (0.82%)	45.57% (0.79%)	45.58% (0.78%)
Germany	46.23% (0.31%)	46.37% (0.28%)	45.19% (0.33%)	46.43% (0.28%)	46.32% (0.29%)	46.42% (0.28%)	46.49% (0.26%)
Italy	46.62% (0.51%)	47.16% (0.48%)	44.90% (0.59%)	47.22% (0.50%)	47.13% (0.53%)	47.13% (0.47%)	47.29% (0.50%)
France	41.92% (0.65%)	41.83% (0.60%)	41.27% (0.55%)	41.90% (0.59%)	42.05% (0.59%)	41.88% (0.59%)	41.81% (0.62%)

Table 3.9: Original Dynamic Network vs Similarity Approaches Predictive Accuracy Average of 4-season Rolling Windows

The first thing we note from Table 3.9 is that the Mean-Based Direct Similarity noticeably underperforms the other approaches here. Note that unlike the static network models, for the mean-based dynamic approaches we do not normalize the similarity matrices, so performance differences are more pronounced. This suggests that the scaling in the Mean-Based Direct Similarity approach is potentially over-weighting some relationships and under-weighting others. For the remainder of this section, we omit this method and its variants for concision.

Continuing our examination of Table 3.9, we see that all leagues except the Spanish La Liga have at least one similarity approach that exhibits superior performance to the standard dynamic model. For the English Premier League, German Bundesliga, and Italian Serie A, all similarity approaches except the Mean-Based Direct Similarity outperform the standard dynamic model. In particular, the Italian Serie A exhibts the largest performance gains.

Next, in Table 3.10 we consider a set of combination approaches incorporating home/away and similarity information, with our home/away approach of Section 3.5.1 serving as our basis of comparison.

League	HA	HA + DS	HA + MSS	HA + MBMSS	HA + USS	HA + MBUSS
England	48.44% (0.38%)	48.74% (0.37%)	48.80% (0.36%)	48.71% (0.36%)	48.82% (0.38%)	48.81% (0.35%)
Spain	46.10% (0.82%)	45.75% (0.78%)	45.88% (0.77%)	45.87% (0.84%)	45.85% (0.82%)	45.83% (0.78%)
Germany	46.41% (0.27%)	46.64% (0.29%)	46.68% (0.29%)	46.63% (0.30%)	46.69% (0.29%)	46.72% (0.27%)
Italy	46.84% (0.59%)	47.48% (0.49%)	47.56% (0.50%)	47.25% (0.47%)	47.54% (0.47%)	47.64% (0.48%)
France	42.21% (0.60%)	42.28% (0.54%)	42.27% (0.57%)	42.31% (0.55%)	42.18% (0.57%)	42.35% (0.57%)

Table 3.10: Original Dynamic Network + Home/Away vs Similarity Approaches + Home/Away Predictive Accuracy Average of 4-season Rolling Windows

Examining Table 3.10, we first notice that, for the Spanish La Liga, the Motegi and Masuda (2012) model with home/away information is superior to the similarity approaches incorporating home/away information. However, for all the other leagues considered, the similarity methods outperform, sometimes markedly, except for the French Ligue 1 and the Unmatched Set Similarity approach. It is worth noting again that the Italian Serie A seems to respond the most, in the dynamic model context, to the addition of similarity information.

Next, in Table 3.11, we compare the Motegi and Masuda (2012) model with goal difference information of Section 3.5.2 with combination approaches that use both similarity approaches and goal difference information.

League	GD	GD + DS	GD + MSS	GD + MBMSS	GD + USS	GD + MBUSS
England	48.98% (0.40%)	49.05% (0.31%)	48.99% (0.32%)	48.92% (0.34%)	49.03% (0.32%)	49.03% (0.32%)
Spain	46.40% (0.83%)	46.21% (0.78%)	46.30% (0.80%)	46.27% (0.79%)	46.30% (0.80%)	46.33% (0.79%)
Germany	46.74% (0.26%)	46.84% (0.25%)	46.90% (0.23%)	47.04% (0.24%)	46.93% (0.24%)	46.83% (0.24%)
Italy	46.94% (0.57%)	47.63% (0.61%)	47.55% (0.62%)	47.72% (0.61%)	47.59% (0.57%)	47.63% (0.60%)
France	42.65% (0.52%)	42.63% (0.57%)	42.65% (0.58%)	42.55% (0.57%)	42.64% (0.56%)	42.63% (0.58%)

Table 3.11: Original Dynamic Network + Goal Difference vs Similarity Approaches + Goal Difference Predictive Accuracy Average of 4-season Rolling Windows

In Table 3.11, we see that the Spanish La Liga continues to exhibit underperformance in similarity approaches, when compared to the comparable dynamic network model without similarity information. In light of the consistent underperformance, we examine rank correlations in Section 3.6.3. Conversely, all similarity approaches outperform in the German Bundesliga and Italian Serie A under similar comparison. The English Premier League and French Ligue 1 demonstrate mixed performance, with the former exhibiting some approaches outperforming and underperfoming while the latter only manages to match the Motegi and Masuda (2012) model with goal difference information in one instance, underperfoming in the others.

Our final performance table for our dynamic model results, Table 3.12 presents the Motegi and Masuda (2012) model with both home/away and goal difference information as the basis of comparison with the similarity approaches combined with home/away and goal difference information.

League	HA + GD	HA + GD + DS	HA + GD + MSS	HA + GD + MBMSS	HA + GD + USS	HA + GD + MBUSS
England	48.78% (0.43%)	49.09% (0.32%)	49.18% (0.32%)	49.12% (0.33%)	49.20% (0.33%)	49.12% (0.33%)
Spain	46.43% (0.84%)	46.58% (0.80%)	46.60% (0.80%)	46.55% (0.82%)	46.64% (0.81%)	46.55% (0.82%)
Germany	46.84% (0.26%)	47.06% (0.25%)	47.13% (0.22%)	47.02% (0.25%)	47.12% (0.21%)	47.02% (0.25%)
Italy	47.09% (0.54%)	47.88% (0.59%)	47.84% (0.61%)	47.75% (0.64%)	47.94% (0.58%)	47.75% (0.64%)
France	42.88% (0.55%)	43.01% (0.59%)	43.03% (0.59%)	42.90% (0.60%)	42.90% (0.60%)	42.90% (0.60%)

Table 3.12: Original Dynamic Network + Home/Away + Goal Difference vs Similarity Approaches + Home/Away + Goal Difference Predictive Accuracy Average of 4-season Rolling Windows

Table 3.12 reveals that when compared to the Motegi and Masuda (2012) method with home/away and goal difference information, all the remaining similarity approaches outperform. Perhaps most-notably, while the individual home/away and goal difference comparisons showed that similarity approaches were underperfoming for the Spanish La Liga, here they all outperform. The Italian Serie A demonstrates the largest response to adding similarity information in the dynamic models, and this is consistent across all the dynamic models we have examined.

Across all the dynamic network methods tested on the data used in this study, these are the best methods per league and the corresponding predictive accuracy achieved:

- England: (49.20%) USS + HA + GD
- Spain: (46.64%) USS + HA + GD
- Germany: (47.13%) MSS + HA + GD
- Italy: (47.94%) USS + HA + GD

- France: (43.03%) MSS + HA + GD.

As noted with the static models, these models are not guaranteed to be the best-performing on future data.

3.6.3 Rank Correlation Analysis

In this section, we examine rank correlation of our dynamic network models to see if any noteworthy patterns emerge that may help explain the performance on the Spanish La Liga data. We focus on the various similarity metric approaches, as those are where the performance discrepancies arose. Section 3.6.3.1 describes the method we use for computing rank correlation. Section 3.6.3.2 contains our correlation data analysis.

3.6.3.1 Spearman Rank Correlation

Our chosen rank correlation metric is the Spearman rank correlation. This metric computes correlation between two vectors of rankings. In our application, we consider two vectors of rankings, corresponding to two sequential ranking periods, $\mathbf{r}_t$ and $\mathbf{r}_{t+1}$. Let these vectors be of length n. Let the ith component of each vector be the rank of team i. To compute the Spearman rank correlation, we use the following formula:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \tag{3.92}$$

where $d_i^2 = (\mathbf{r}_{t+1,i} - \mathbf{r}_{t,i})^2$. To compute the average Spearman rank correlation for our testing, for each 4-year data chunk, per ranking method, we compute the sequential rank correlations, gameweek to gameweek. After we have all these correlations for a data chunk, we average them. We then average over all chunks to get our overall average rank correlation, per method.

Note that higher rank correlation implies that ranks change less often, on average, between gameweeks. A higher rank correlation average is an indicator that the league settles into its final standing sooner, which is an indication that the league is less competitive.

3.6.3.2 Data Analysis

In Table 3.13, we present the average rank correlation of the Motegi and Masuda (2012) dynamic network approach and our basic similarity metric approaches.

League	Dynamic Network	Direct Similarity (DS)	Mean-Based Direct Similarity (MBDS)	Matched Set Similarity (MSS)	Mean-Based Matched Set Similarity (MBMSS)	Unmatched Set Similarity (USS)	Mean-Based Unmatched Set Similarity (MBUSS)
England	93.86%	92.88%	93.80%	92.85%	92.83%	92.88%	92.90%
Spain	92.42%	90.65%	92.96%	90.74%	90.77%	90.67%	90.75%
Germany	92.72%	91.31%	93.06%	91.30%	91.28%	91.27%	91.28%
Italy	93.84%	92.42%	93.57%	92.46%	92.42%	92.43%	92.45%
France	92.41%	90.73%	92.87%	90.74%	90.71%	90.77%	90.74%

Table 3.13: Original Dynamic Network vs Similarity Approaches Spearman Rank Correlation Average of 4-season Rolling Windows

Examining Table 3.13, we see that Spain, for most of the approaches, features noticeably lower average rank correlation than the other leagues. This suggests that, at least for these methods, the teams continue to move throughout the rankings more so than the other leagues.

In Table 3.14, we consider the average rank correlation when home/away information is added.

League	HA	HA + DS	HA + MSS	HA + MBMSS	HA + USS	HA + MBUSS
England	94.23%	93.87%	93.01%	92.98%	93.01%	93.02%
Spain	92.70%	90.91%	90.94%	90.93%	90.91%	90.93%
Germany	92.85%	91.43%	91.42%	91.51%	91.44%	91.38%
Italy	94.04%	92.59%	92.63%	92.54%	92.64%	92.60%
France	92.81%	91.00%	90.98%	90.97%	91.08%	91.01%

Table 3.14: Original Dynamic Network + Home/Away vs Similarity Approaches + Home/Away Spearman Rank Correlation Average of 4-season Rolling Windows

A similar pattern emerges in Table 3.14, with Spain having a noticeably lower rank correlation. Interestingly, we should note that the lower rank correlation is not necessarily associated with worse performance. Specifically, Italy, of which all the similarity approaches yield noticeable performance improvements, demonstrates lower rank correlation for the similarity approaches. This suggests that what is occurring with Spain could be league-specific.

In Table 3.15, we examine the average rank correlation of the dynamic network approaches with goal difference information.

League	GD	GD + DS	GD + MSS	GD + MBMSS	GD + USS	GD + MBUSS
England	94.42%	93.39%	93.38%	93.38%	93.39%	93.36%
Spain	93.01%	91.29%	91.30%	91.37%	91.32%	91.30%
Germany	93.16%	91.83%	91.82%	91.91%	91.84%	91.81%
Italy	94.21%	92.80%	92.91%	92.88%	92.89%	92.85%
France	92.64%	90.98%	90.84%	91.04%	91.05%	90.96%

Table 3.15: Original Dynamic Network + Goal Difference vs Similarity Approaches + Goal Difference Spearman Rank Correlation Average of 4-season Rolling Windows

Examining Table 3.15, we see that it is no longer evident that Spain has a markedly lower average rank correlation than the other leagues. Conversely, it is apparent that England has the highest rank correlation consistently. This could partially be explained by the influence of the large clubs in the English Premier League that have had historically strong multi-season performances (Manchester United is the canonical example).

In Table 3.16, we examine the dynamic network ranking models that include both home/away and goal difference information.

League	HA + GD	HA + GD + DS	HA + GD + MSS	HA + GD + MBMSS	HA + GD + USS	HA + GD + MBUSS
England	94.47%	93.54%	93.51%	93.53%	93.54%	93.53%
Spain	92.98%	91.52%	91.57%	91.55%	91.47%	91.55%
Germany	93.12%	92.08%	92.08%	92.11%	92.03%	92.11%
Italy	94.21%	93.05%	93.04%	93.01%	93.08%	93.01%
France	92.81%	91.38%	91.33%	91.37%	91.41%	91.37%

Table 3.16: Original Dynamic Network + Home/Away + Goal Difference vs Similarity Approaches + Home/Away + Goal Difference Spearman Rank Correlation Average of 4-season Rolling Windows

In Table 3.16, we again see that performance and rank correlation are not necessarily associated. Further, after examining all the tables in this section, it is worth noting that, on average, incorporating similarity information decreases rank correlation. This is somewhat intriguing as the inclusion of additional information related to home/away and goal difference information seems to increase rank correlation.

There are explanations for both phenomena. The home/away and goal difference information, in essence, cause teams to have greater differences in total score. This means that any differences in rank are effectively larger. As such, it is more difficult for those ranks to change per gameweek, leading to higher rank correlation. The similarity information has the opposite effect. Specifically, the similarity information will reduce the influence of teams that do not exchange positions with other teams as often; this is likely to occur to both the best and worst teams. Additionally, the influence of teams that exchange positions more often, likely those in the middle, will grow. As such, the teams are all closer together, in terms of total score, which leads to easier rank changes and more rank oscillations. This drives average rank correlation downwards. Overall, both of these effects combine and their influence on the rank correlation is league-dependent.

3.6.4 Sensitivity Analysis

In Section 3.6.4.1 we examine the sensitivity of our top performing static model to parameter changes. In Section 3.6.4.2 we examine the sensitivity of our top performing dynamic model to parameter changes.

3.6.4.1 Static Model

Recall that our top performing static model was the Park and Newman (2005) static network ranking model using both home/away and goal difference information using data from the English Premier League. In Figure 3.2, we plot average predictive accuracy as a function of k_A and k_D, our parameters for home/away and goal difference information respectively.

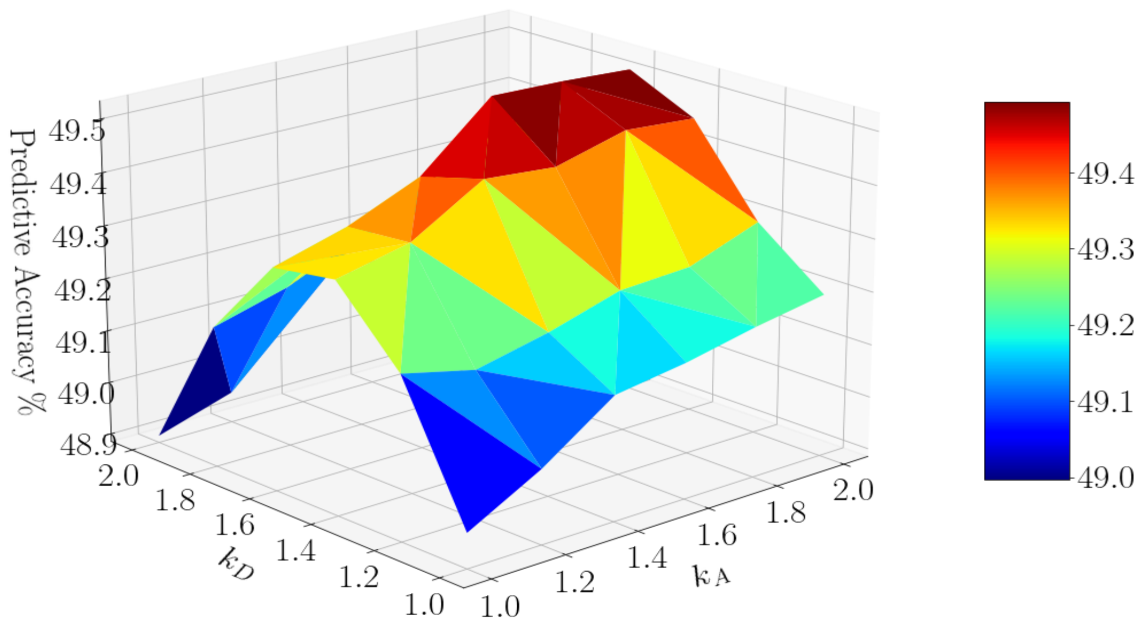

Figure 3.2: Average Predictive Accuracy as a function of k_A and k_D

Examining Figure 3.2, we can clearly see a parabolic shape induced by changes in the goal difference parameter k_D. It appears that an ideal value for this parameter is approximately 1.6. The home/away parameter k_A, conversely, appears to produce a near-monotonic increase in average predictive accuracy as it grows. From this particular sensitivity analysis, the combination of parameters that yielded the highest average predictive accuracy is $k_A = 1.6$, $k_D = 1.6$ with an average predictive accuracy of 49.52%.

3.6.4.2 Dynamic Model

Recall that our top-performing dynamic model was the Unmatched Set Similarity Approach with home/away and goal difference information using data from the English Premier league. In Figure 3.3, we present average predictive accuracy of this method as a function of k_A and k_D, keeping $\alpha = 0.2$ and $\beta = 0.1$ from our overall empirical testing.

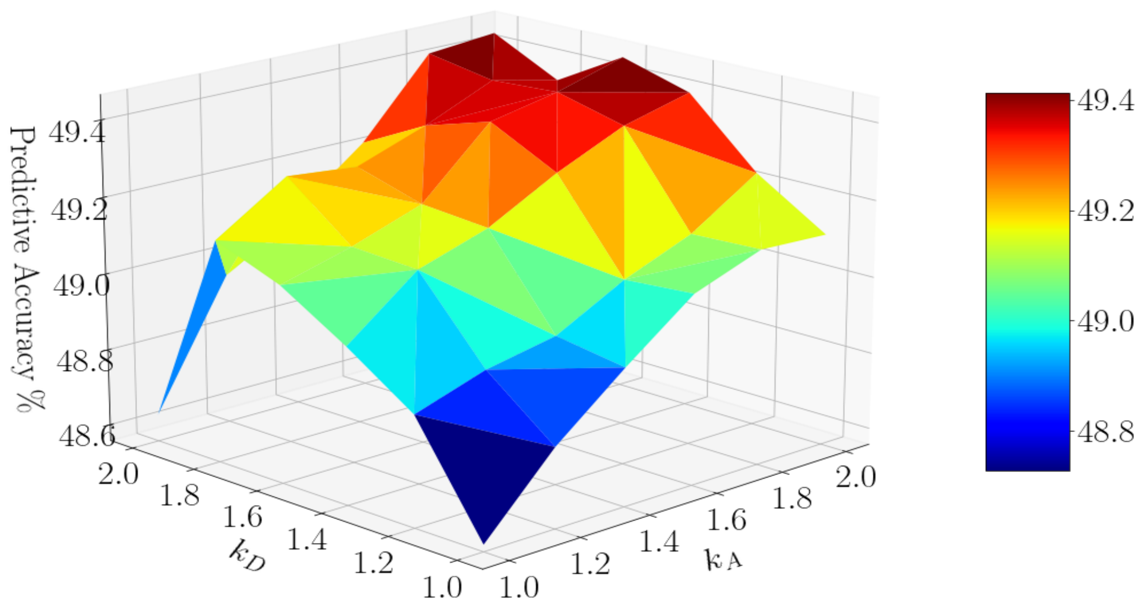

Figure 3.3: Average Predictive Accuracy as a function of k_A and k_D; $\alpha = 0.2$, $\beta = 0.1$

Figure 3.3 suggests that k_D induces a similar parabolic shape to the static case, though for higher values of k_A, higher values of k_D are favourable. For this particular sensitivity analysis, the optimal parameter combination is $k_A = 2.0$ and $k_D = 1.6$, with an average predictive accuracy of 49.45%.

In Figure 3.4, we use the highest-performing parameterization from the previous sensitivity analysis ($k_A = 2.0$ and $k_D = 1.6$) and plot average predictive accuracy as a function of α and β.

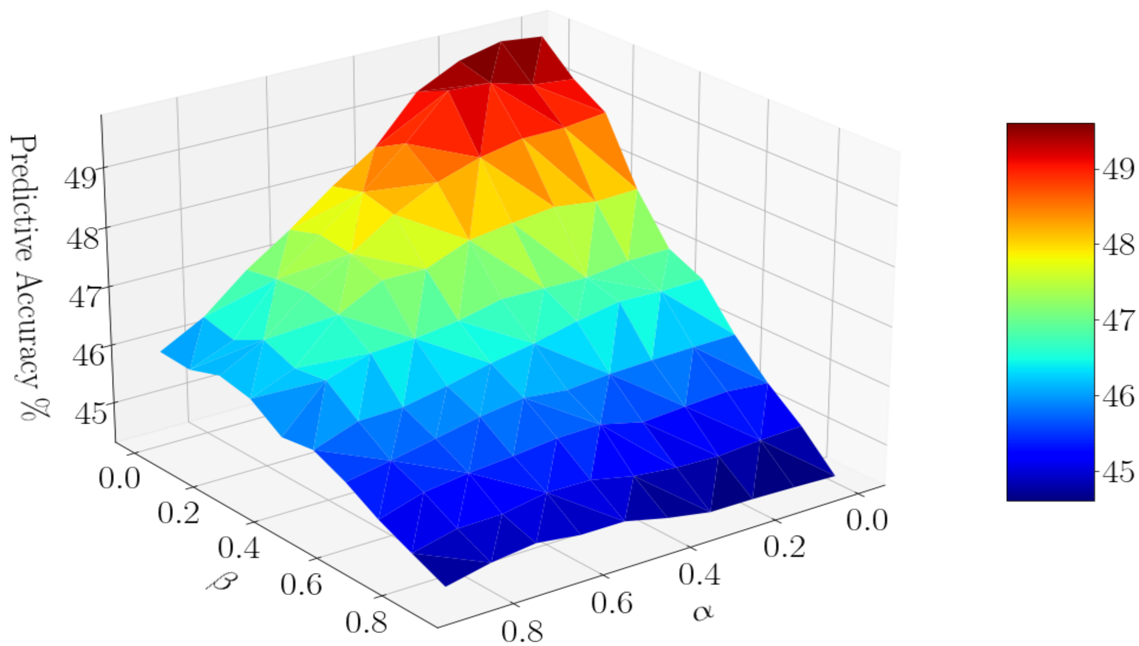

Figure 3.4: Average Predictive Accuracy as a function of α and β; $k_A = 2$ and $k_D = 1.6$

Note that, for easier viewing, the axes in Figure 3.4 increase towards the centre of the image. This figure clearly shows that increasing α and β, at least for these values of k_A and k_D, produces notable decreases in average predictive accuracy. For this particular sensitivity analysis, the optimal parameter combination is $\alpha = 0.1$ and $\beta = 0$, with an average predictive accuracy of 49.76%.

3.6.5 Betting Odds Comparison

As a final basis of comparison, we acquire data from website Football-Data.co.uk, which contains comprehensive betting odds data for all the leagues considered in our study ("Football-Data.co.uk" (2023)). Betting odds can be considered the approach to beat, since if one can consistently outperform the betting odds, a profit can be made. We provide the comparison to show that there still exists a gap, in terms of performance, between our approaches and the betting odds. However, we do so while acknowledging that betting odds are set through a complicated process; not only to bookmakers use complicated ensemble models to predict game outcomes, but they incorporate the current betting pool into their evaluations.

To compute the predictive accuracy of the betting odds, we simply assume that the betting odds predict the result with the most likely odds (for example, in a decimal betting odds system, the outcome with the lowest odds is the one with the largest estimated probability of occurring). We then compare this predicted result to the actual result; if the predicted result matches the actual result, a correct prediction is recorded; otherwise, an erroneous prediction is recorded. The predictive accuracy is then computed as

$$\text{Betting Odds Predictive Accuracy} = \frac{\text{correct predictions}}{\text{total number of matches}} \times 100\%. \tag{3.93}$$

The per year predictive accuracy is shown in Figure 3.5.

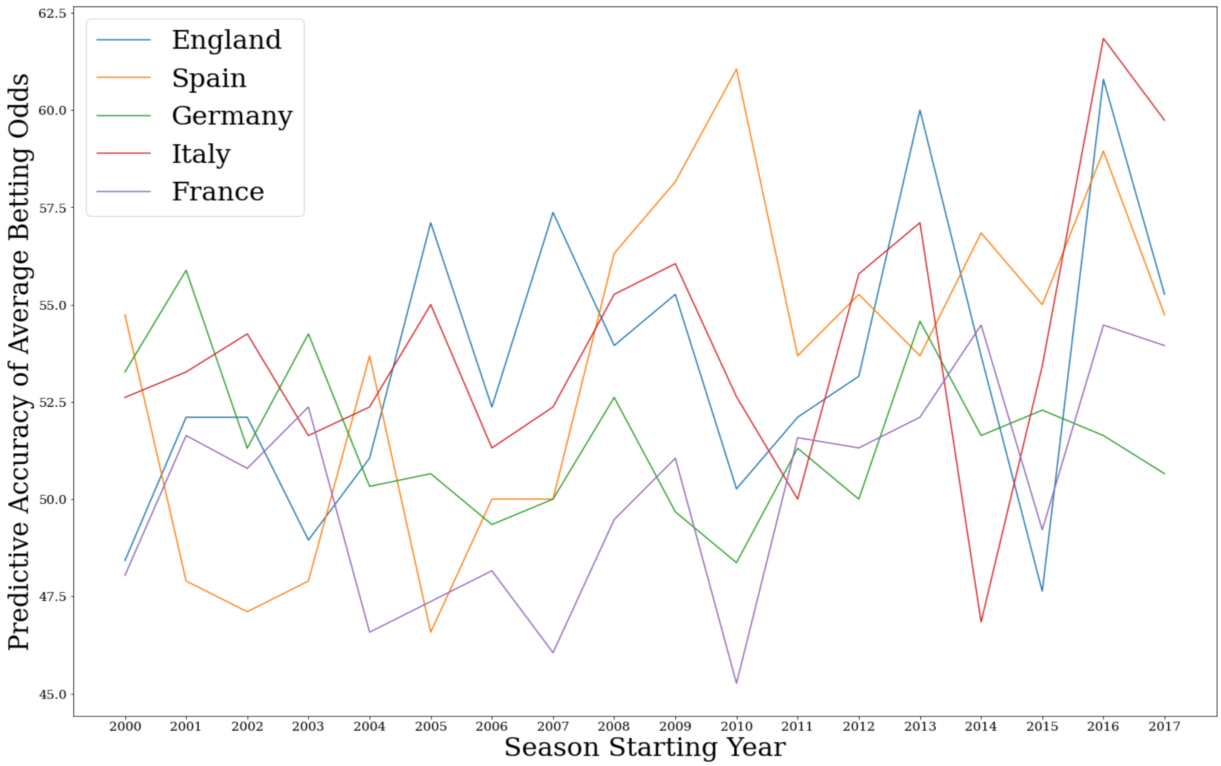

Figure 3.5: Per Season Betting Odds Predictive Accuracy

The average per-league predictive accuracy is shown in Table 3.17.

England	Spain	Germany	Italy	France
53.42%	53.42%	51.54%	53.97%	50.22%

Table 3.17: Average Betting Odds Predictive Accuracy Per Country

Examining Figure 3.5, we can see that the betting odds predictive accuracy can oscillate rather widely year-over-year. However, the average predictive accuracies in Table 3.17 do indicate that their performance, on average, outpaces our proposed methods. As alluded to earlier, the chief performance superiority of the betting odds is in predicting ties.

Consider, for example, our highest-performing static network approach: the Park and

Newman (2005) model featuring both home/away and goal difference information used in ranking the English Premier League. The total predictive accuracy for this model is 49.25%. However, the method predicts 63.76% of home wins and 70.21% of away wins correctly. Unfortunately, the method predicts 0.07% of draws correctly. The reason for this is largely due to the chosen predictive scheme. Namely, we only predict a tie if teams have the same rank. However, teams only have the same rank if they have the same total score, which is an extremely rare occurrence partially driven by the fact that the method relies on convergence rather than a closed form.

We have attempted several approaches to improve our methods' predictive accuracy, all of which are detailed in Appendix B.4. So far, there appears to be a challenging tradeoff wherein any gains made in predicting ties are overshadowed by losses in accuracy in predicting home and away wins. However, this challenge is something we plan to address in future research, which we make reference to in Section 3.8.

3.7 Discussion

In Section 3.7.1 we discuss the main takeaways from our methodology. In Section 3.7.2 we consider the managerial implications of our results.

3.7.1 Methodological Considerations

As demonstrated in Section 3.6, we see that our additions to both static and dynamic network ranking models show superior performance in terms of predictive accuracy than the approaches they are based on. As mentioned earlier, the methodology choice does matter for optimal results, as one similarity metric is not optimal for all leagues. Further, our sensitivity analyses reveals that parameter choice matters, as both the home/away parameter k_A and the goal difference parameter k_D produce noticeable changes in the performance of methods using these parameters.

One key takeaway from our work is that there is valuable, performance-enhancing information contained within the time-series of rankings. Specifically, competition information, using the Criado et al. (2013) definition, is useful for scaling the influence of the α parameter in network ranking models, in terms of enhancing prediction accuracy. Further, by incorporating this information, we are able to treat teams heterogeneously, which opens up future inquiries. For example, we note that Set Similarity methods do particularly well in the Italian Serie A; this points to a potential hidden structure to uncover in terms of how the teams are similar or dissimilar with each other.

It remains to be be seen whether our results here generalize to wider samples of teams. To our knowledge, our work has considered the largest sample of professional soccer games in the ranking literature to date. We leave it for future work to investigate whether these results hold for larger networks. However, using the network ranking models developed here, we can add more leagues without much difficulty, provided we include cross-league competitions like the UEFA Champions League.

It should also be noted that the best methods as indicated from our testing are not guaranteed to be the best for future performance. When actually implementing these models for predictive purposes, it would likely be best to combine this model with a sort of meta-model that predicts which network model will do best for the upcoming gameweek. This is beyond the scope of this work, but potentially a subject for future study.

3.7.2 Managerial Implications

Beyond methodological considerations, how can firms use our results? Firstly, the methods we introduce yield greater predictive accuracy than existing, comparable approaches. As such, team managers, sports analysts, and betting odds bookmakers who use ranking models as part of their analytics suite will likely benefit from incorporating our approaches into their systems.

Secondly, as briefly mentioned earlier, our similarity metrics provide a useful basis through which to examine team interrelationships in professional soccer. The type of similarity that leads to the best performance in terms of predictive accuracy could be an indication of a particular relationship, in terms of performance, among the teams of a given league. For example, these types of similarity could help uncover incidents of match fixing if, say, the rankings appear to be fluctuating more or less than expected and similarity values diverge from projections.

Finally, our methods introduced here can be modified for other network model domains. One such domain that has become relevant in recent years is the modelling of epidemics via networks. Infection networks are typically arranged where nodes are individuals and edges are drawn between nodes pointing in the direction of infection. Applying network ranking models to these networks allows us to rank those individuals who are the most prevalent spreaders of the infection. Depending on the particular question being studied, the modeller may wish to encode additional information into the model; our approaches to home/away and goal/difference information can be modified for this purpose. Further, studying rank crossings via similarity metrics like those we introduce in this work could be used to categorize various types of individuals in the network and help predict which individuals have the highest potential of becoming highly infectious individuals.

3.8 Conclusion

Ultimately, sports enthusiasts watch and analyze sports as part of a grand spectacle of comparison. The base component of these comparisons is an assessment of the strength of the athletes and teams in the competition, and this is most commonly estimated via ranking methodologies. We introduce augmentations to both the static Park and Newman (2005) and dynamic Motegi and Masuda (2012) network ranking models that incorporate additional information. Further, our work provides the first integration of the rank competition concept of Criado et al. (2013) directly into network ranking models. We demonstrate that these additions yield superior empirical performance by using data on five of the most popular European men's professional soccer teams.

Our results highlight that the best-performing model differs between leagues. As such, model choice should be carefully considered for practical implementation. Further, our sensitivity analyses indicate that parameter choices for incorporating home/away and goal difference information have a material effect on performance.

As with any work, ours has limitations. Firstly, our performance conclusions are based off of a specific set of data, though for the leagues covered, this does represent the full population of data for the years used. It remains to be seen if our performance improvements hold for leagues in other countries or continents. A second limitation is that our work concerns only one sport; we cannot extrapolate these results to, say, professional basketball without testing. Thirdly, our approaches only consider one form of similarity and fixed parameters throughout the testing procedure; it could be the case that varying these components each period could lead to superior performance not demonstrated here.

This work presents many future vectors of inquiry, some of which we mention here briefly. One of the most natural follow-up studies would be to work on developing a method to consistently beat the betting odds, using our ranking approaches as a basis. Our initial thinking is that some sort of ensemble approach could be useful, though balancing efficacy with model parsimony would be a meaningful part of the challenge. Another avenue to pursue would be using the efficacy of the additional information we introduce here as a basis to pursue further econometric analyses. For example, if we find home/away information to be useful, especially in certain years, we could use this as a basis to investigate factors that may have influenced the importance of this information. One potential idea that comes to mind is that some countries may experience marked weather disparities when comparing their northernmost and southernmost points; home/away information might be indicative of more noteworthy weather differentials (as teams play in weather they are less-accustomed to).

Chapter 4

Conclusion

Management science provides organizations of all sizes with approaches to enhance operational efficiency and efficacy. Chief among these approaches is the modelling of systems using mathematics to capture salient features, from which we can derive insights otherwise unobtainable. Models will differ significantly depending on the problem at hand, as they should, since the model must be well-suited to the problem. However, a common goal from the modelling exercise is to incorporate new features or information to better-represent the system than existing approaches.

In this dissertation, both essays focus on introducing new approaches to modelling their problems of interest. The first essay uses distributional censoring to capture the effects of Operational Protocol Modifications (OPMs) on the service process and the associated employment contracts. The second essay proposes ways to incorporate useful soccer-relevant information like home/away status and goal difference into network ranking models. Further, and more significantly, it ties two streams of literature together by incorporating similarity metrics based on rank crossings directly into these network ranking models. Notably, while both essays deploy different techniques to address distinct problem domains, they are unified through the overarching concept of comparison. Be it between different system states or competing entities, the notion of comparison is fundamental to management science.

While generating a model is interesting, as management scientists, we are most-concerned with how our models impact our understanding of management and operational practice. Both of our essays contribute insights as follows:

- The first essay provides operations management analysts a useful tool – distributional censoring – to use in analyzing service systems. As demonstrated by this essay, the effects on expected employer utility are nonlinear in nature, so any proposed changes to OPMs should consider both the current system state as well as the state imposed by the changes.

- The second essay provides any user of network ranking models with a collection of tools to enhance these models. Team managers, league policymakers, and sports betting bookmakers can all use these tools to better-understand the relative positioning of teams in their league(s) of interest. Further, our approaches are somewhat general, insofar as similar approaches could be used to incorporate other useful information.

As demonstrated by this dissertation, management science is well-situated to continue providing insights to managers and stakeholders of various organizations. As it relates to our work here, there are many potential future vectors of inquiry. We plan on further generalizing our modelling in the first essay to incorporate more general contract forms and alternative forms of censoring. As for the second essay, the most-evident extension is to more-seriously explore the problem of predicting ties and upsets in professional soccer, using our network ranking models as a base input.

References

Abernethy, S. (2018). Dynamic network $3-0$ fifa rankings: Replacing an inaccurate, biased, and exploitable ranking system.

Ahmad, W., Porter, M. A., & Beguerisse-Diaz, M. (2021). Tie-decay networks in continuous time and eigenvector-based centralities. *Information Sciences*, *8*(2).

Al-Athari, F. M. (2008). Estimation of the mean of truncated exponential distribution. *Journal of Mathematics and Statistics*, *4*(4), 284–288.

Baker, R. D., & McHale, I. G. (2014). A dynamic paired comparisons model: Who is the greatest tennis player? *European Journal of Operational Research*, *236*(2), 677–684.

Barbos, A. (2019). Dynamic contracts with random monitoring. *Journal of Mathematical Economics*, *85*, 1–16.

Baron, D. P., & Besanko, D. (1984). Asymmetric Information, and Auditing. *The RAND Journal of Economics*, *15*(4), 447–470.

Bénabou, R., & Tirole, J. (2016). Bonus Culture: Competitive Pay, Screening, and Multitasking. *Journal of Political Economy*, *124*(2), 305–370.

Bolton, P., & Dewatripont, M. (2004). *Contract Theory*. The MIT Press.

Border, K. C., & Sobel, J. (1987). Samurai Accountant: A Theory of Auditing and Plunder. *The Review of Economic Studies*, *15*(4), 447–470.

Bradley, R. A., & Terry, M. E. (1952). The rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, *39*, 324–345.

Butcher, K., Sparks, B., & McColl-Kennedy, J. (2009). Predictors of customer service training in hospitality firms. *International Journal of Hospitality Management*, *28*, 389–396.

Carroll, G. (2015). Robustness and Linear Contracts. *American Economics Review*, *105*(2), 536–563.

Castillo, D., Canhoto, A. I., & Said, E. (2021). The dark side of AI-powered service interactions: exploring the process of co-destruction from the customer perspective. *The Service Industries Journal*, *41*(13-14), 900–925.

Cattelan, M., Varin, C., & Firth, D. (2013). Dynamic Bradley-Terry modelling of sports tournaments. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *62*(1), 135–150.

Chen, H.-C., & Liu, S.-M. (2008). Incentive Contracts Under Imperfect Auditing. *The Manchester School*, *76*(2), 131–159.

Chen, M., Sun, P., & Xiao, Y. (2020). Optimal Monitoring Schedule in Dynamic Contracts. *Operations Research*, *68*(5), 1285–1314.

Criado, R., Garcia, E., Pedroche, F., & Romance, M. (2014). Comparing rankings by means of competitivity graphs: Structural properties and computation.

Criado, R., Garcia, E., Pedroche, F., & Romance, M. (2016). On graphs associated to sets of rankings. *Journal of Computational and Applied Mathematics*, *291*, 497–508.

Criado, R., García, E., Pedroche, F., & Romance, M. (2013). A new method for comparing rankings through complex networks: Model and analysis of competitiveness of major european soccer leagues. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *23*(4), 043114. https://doi.org/10.1063/1.4826446

Davidson, R. R. (1970). On extending the bradley-terry model to accommodate ties in paired comparison experiments. *Journal of the American Statistical Association*, *65*(329), 317–328.

Diamond, P. (1998). Managerial incentives: On the near linearity of optimal compensation. *Journal of Political Economy*, *106*(5), 931–957.

Durgin, A., Mahoney, A., Cox, C., Weetjens, B. J., & Poling, A. (2014). Using Task Clarification and Feedback Training to Improve Staff Performance in an East African Nongovernmental Organization. *Journal of Organizational Behavior Management*, *34*, 122–143.

Elo, A. (1978). *The Rating of Chessplayers, Past and Present*. Arco Publishing Inc.

Erkol, Ş., & Radicchi, F. (2021). Who is the best coach of all time? a network-based assessment of the career performance of professional sports coaches (J. Méndez, Ed.). *Journal of Complex Networks*, *9*(1). https://doi.org/10.1093/comnet/cnab012

Football-data.co.uk [Accessed: 2023-05-01]. (2023).

Garrett, D. F. (2021). Payoff implications of incentive contracting. *Theoretical Economics*, *16*, 1281–1312.

Glickman, M. E. (1999). Parameter estimation in large dynamic paired comparison experiments. *Applied Statistics*, *48*(3), 377–394.

Graetz, M. J., Reinganum, J. F., & Wilde, L. L. (1986). The Tax Compliance Game: Toward and Interactive Theory of Law Enforcement. *Journal of Law Economics and Organization*, *2*(1), 1–32.

Grossman, S. J., & Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, *51*(1), 7–45.

Helm, C., & Wirl, F. (2020). Multitasking: Incentivizing agents differing either in their work ethic or intrinsic motivation. *Journal of Economics*.

Hoffmann, F., Inderst, R., & Opp, M. (2021). Only Time Will Tell: A Theory of Deferred Compensation. *Review of Economic Studies*, *88*, 1253–1278.

Holmstrom, B. (1979). Moral Hazard and Observability. *The Bell Journal of Economics*, *10*(1), 74–91.

Holmstrom, B., & Milgrom, P. (1987). Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica*, *55*(2), 303–328.

Hunter, D. R. (2004). Mm algorithms for generalized bradley-terry models. *The Annals of Statistics*, *32*(1), 384–406.

Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, *26*, 460–470.

Hyo-jun, Y., Jae-Hyeon, P., Jiwun, Y., & Minsoo, J. (2021). Evaluation on the team perforamnce of mlb applying pagerank algorithm. *Annals of Applied Sport Science*, *9*(4).

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous Univariate Distributions: Volume 2 (Second Edition)*. John Wiley & Sons Inc.

Johnson, S. (2002). Lifelong learning and SMEs: issues for research and policy. *Journal of Small Business and Enterprise Development*, *9*(3), 285–295.

Kaufman, R. (2015). Why Your Customer Service Training Won't Lead to Happy Customers (of Inspired Employees). *The Journal for Quality and Participation*, *37*(4), 33–38.

Khalil, F. (1997). Auditing without Commitment. *RAND Journal of Economics*, *28*(4), 629–640.

Khalil, F., & Lawarree, J. (2001). Catching the agent on the wrong foot: ex post choice of monitoring. *Journal of Public Ecnomics*, *82*(4), 327–347.

Kim, E., & Jeon, M. (2019). Proposal for implementation of a ranking model for olympic taekwondo competitions using pagerank. *International Journal of Performance Analysis in Sport*, *19*(2), 227–235.

Knorr-Held, L. (2000). Dynamic Rating of Sports Teams. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *49*(2), 261–276.

Kofman, F., & Lawarree, J. (1993). Collusion in Hierarchical Agency. *Econometrica*, *61*(3), 629–656.

Laffont, J.-J., & Martimort, D. (2009). *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Marshall, A. W., Olkin, I., & Arnold, B. C. (2011). *Inequalities: Theory of Majorization and its Applications*. Springer Science+Business Media, LLC.

Micallef, L. (2018). CRM 101: What is CRM [data retrieved from Deloitte official website, https://www2.deloitte.com/mt/en/pages/technology/articles/mt-salesforce-crm-101.html].

Mookherjee, D., & Png, I. (1989). Optimal Auditing, Insurance, and Redistribution. *The Quarterly Journal of Economics*, *104*(2), 399–415.

Motegi, S., & Masuda, N. (2012). A network-based dynamical ranking system for competitive sports. *Scientific Reports*, *2*(1). https://doi.org/10.1038/srep00904

Myers, R. H., Montgomery, D. C., Vining, G. G., & Robinson, T. J. (2010). *Generalized Linear Models with Applications in Engineering and the Sciences*. Wiley.

Park, J., & Newman, M. E. J. (2005). A network-based ranking system for us college football. *Journal of Statistical Mechanics: Theory and Experiment*.

Patton, D., & Marlow, S. (2002). The determinants of management training within smaller firms in the UK. What role does strategy play? *Journal of Small Business and Enterprise Development*, *9*(3), 260–270.

Pedroche, F., Criado, R., Garcia, E., Romance, M., & Sanchez, V. E. (2014). Comparing series of rankings with ties by using complex networks: An analysis of the spanish stock market (ibex-35 index).

Pedroche, F., & Conejero, J. A. (2020). Corrected evolutive kendall's τ coefficients for incomplete rankings with ties: Application to case of spotify lists. *Mathematics*, *8*, 1828.

Radicchi, F. (2011). Who is the best player ever? a complex network analysis of the history of professional tennis. *PLoS One*.

Rasmusen, E. (2006). *Games and Information: An Introduction to Game Theory*. Blackwell Publishing.

Reinganum, J. F., & Wilde, L. L. (1985). Income Tax Compliance in a Principal-Agent Framework. *Journal of Public Economics*, *26*(1), 1–18.

Reinganum, J. F., & Wilde, L. L. (1986). Equilibrium Verification and Reporting Policies in a Model of Tax Compliance. *International Economic Review*, *27*(3), 739–760.

Rogerson, W. P. (1985). The First-Order Approach to Principal-Agent Problems. *Econometrica*, *53*(6), 1357–1367.

Schanke, S., Burtch, G., & Ray, G. (2022). Estimating the Impact of "Humanizing" Customer Service Chatbots. *Information Systems Research*, *32*(3), 736–751.

Stefani, R., & Pollard, R. (2007). Football rating systems for top-level competition: A critical survey. *Journal of Quantitative Analysis in Sports*, *3*(3), 1–22.

Stole, L. A. (1993). Lectures on the theory of contracts.

Thomson, A., & Gray, C. (1999). Determinants of management development in small businesses. *Journal of Small Business and Enterprise Development*, *6*(2), 113–127.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, *21*(2), 265–293.

Tuesta, E. F., Bolanos-Pizarro, M., Neves, D. P., Fernandez, G., & Axel-Berg, J. (2020). Complex networks for benchmarking in global universities rankings. *Scientometrics*, *125*, 405–425.

Tutz, G., & Schauberger, G. (2014). Extended Ordered Paired Comparison Models with Application to Football Data from German Bundesliga [Technical Report Number 151. Department of Statistics. University of Munich].

Vinten, G. (2000). Training in small- and medium-sized enterprises. *Industrial and Commercial Training*, *32*(1), 9–14.

Worldfoolball.net [Accessed: 2021-09-30]. (2021).

Wunderlich, F., & Memmert, D. (2018). The Betting Odds Rating System: Using soccer forecasts to forecast soccer. *PLoS ONE*, *13*(6), e0198668.

Yang, L., Dimitrov, S., & Mantin, B. (2014). Forecasting sales of new virtual goods with the elo rating system. *Journal of Revenue and Pricing Management*, *13*(6), 457–469.

Yu, Y., & Kong, X. (2020). Robust Contract Designs: Linear Contracts and Moral Hazard. *Operations Research*, *68*(5), 1457–1473.

Zhou, Y., Wang, R., Zhang, Y.-C., Zeng, A., & Medo, M. (2020). Limits of pagerank-based ranking methods in sports data.

APPENDICES

Appendix A

Contract Theory Proofs

A.1 Solutions of the Optimization Problem (POPI)

We prove the solutions given in Subsection 2.5. We note that the solutions are obtained under Assumption 2.4.0.1 (i.e., $\Delta_S(x)$ and $\Delta_T(x)$ are monotone increasing and concave in x, for $0 \leq x \leq 1$). First, we state a property that will be used repeatedly in the solution process.

Proposition A.1.0.1. *We have* $U^P + U^A = c_1\Delta_S(x) + c_2\Delta_T(x) - ac_a - kx^2$, *for* $0 \leq x \leq 1$.

According to Subsection 2.5, we consider three cases: i) $x = 0$; ii) $x = 1$; and iii) $0 < x < 1$ to find the optimal solution for (POPI). We find solutions that maximize the objective function U^P for each case. Those optimal solutions are called candidates and one of them is the optimal solution for (POPI).

First, we consider $x = 0$. The best solution for this case can be found by solving

$$\begin{aligned} &\max_{\alpha,\beta,\delta} \ \{(c_1 - \alpha)\Delta_S(0) + (c_2 - \beta)\Delta_T(0) - ac_a - \delta ap\} \\ &\text{s.t. } \max_{0 \le x \le 1} \left\{\alpha\Delta_S(x) + \beta\Delta_T(x) + \delta ap - kx^2\right\} \le \alpha\Delta_S(0) + \beta\Delta_T(0) + \delta ap; \\ &\qquad \alpha\Delta_S(0) + \beta\Delta_T(0) + \delta ap \ge 0; \\ &\qquad \alpha, \beta, \delta \ge 0. \end{aligned} \tag{A.1}$$

By Proposition A.1.0.1, for this case, maximizing U^P is equivalent to finding $\{\alpha, \beta, \delta\}$ to minimize U^A. The optimization problem (A.1) is equivalent to

$$\begin{aligned} \min_{\alpha,\beta,\delta} \left\{U^A = \alpha\Delta_S(0) + \beta\Delta_T(0) + \delta ap\right\} \\ \text{s.t. } \alpha(\Delta_S(x) - \Delta_S(0)) + \beta(\Delta_T(x) - \Delta_T(0)) \le kx^2, \quad 0 \le x \le 1; \\ U^A = \alpha\Delta_S(0) + \beta\Delta_T(0) + \delta ap \ge 0; \\ \alpha, \beta, \delta \ge 0. \end{aligned} \tag{A.2}$$

It is easy to see that δ should be small, as long as $U^A \ge 0$. Since we can always choose δ to ensure $U^A \ge 0$, we further consider optimization problem

$$\begin{aligned} \min_{\alpha,\beta} \{\alpha\Delta_S(0) + \beta\Delta_T(0)\} \\ \text{s.t. } \alpha(\Delta_S(x) - \Delta_S(0)) + \beta(\Delta_T(x) - \Delta_T(0)) \le kx^2, \quad 0 \le x \le 1; \\ \alpha, \beta \ge 0. \end{aligned} \tag{A.3}$$

Proposition A.1.0.2. *If* $(\alpha^{(1)}, \beta^{(1)})$ *is an optimal solution of* (A.3)*, then* $(\alpha^{(1)}, \beta^{(1)}, \delta^{(1)})$ *is an optimal solution of* (A.2)*, where* $\delta^{(1)} = \max\left\{0, -\frac{\alpha^{(1)}\Delta_S(0) + \beta^{(1)}\Delta_T(0)}{ap}\right\}$.

Proof. If $(\alpha^{(1)}, \beta^{(1)})$ is an optimal solution of (A.3) and $\alpha^{(1)}\Delta_S(0) + \beta^{(1)}\Delta_T(0) \le 0$, then we must have $\alpha^{(1)}\Delta_S(0) + \beta^{(1)}\Delta_T(0) + \delta^{(1)}ap = 0$, which implies that $(\alpha^{(1)}, \beta^{(1)}, \delta^{(1)})$ is an optimal solution of (A.2). If $\alpha^{(1)}\Delta_S(0) + \beta^{(1)}\Delta_T(0) > 0$, we choose $\delta^{(1)} = 0$. Then $(\alpha^{(1)}, \beta^{(1)}, 0)$ must be an optimal solution to (A.2). Otherwise, there exists $\{\hat{\alpha}, \hat{\beta}, \hat{\delta}\}$ that is a better solution to (A.2). As a result, we must have

$$\hat{\alpha}\Delta_S(0) + \hat{\beta}\Delta_T(0) + \hat{\delta}ap < \alpha^{(1)}\Delta_S(0) + \beta^{(1)}\Delta_T(0) + \delta^{(1)}ap, \tag{A.4}$$

which leads to $\hat{\alpha}\Delta_S(0) + \hat{\beta}\Delta_T(0) \leq \hat{\alpha}\Delta_S(0) + \hat{\beta}\Delta_T(0) + \hat{\delta}ap < \alpha^{(1)}\Delta_S(0) + \beta^{(1)}\Delta_T(0)$. Since $\{\hat{\alpha}, \hat{\beta}\}$ is also a feasible solution to (A.3), the last inequality contradicts the fact that $\{\alpha^{(1)}, \beta^{(1)}\}$ is the optimal solution to (A.3). This completes the proof. $\square$

For (A.3), there are always nonnegative (α, β) such that the constraint holds for any x. On the other hand, the objective function is minimized if α is as small as possible, since $\Delta_S(0) \geq 0$, and β is as large as possible, since $\Delta_T(0) \leq 0$. The first constraint in (A.3) becomes

$$\beta \leq \frac{kx^2 - \alpha(\Delta_S(x) - \Delta_S(0))}{\Delta_T(x) - \Delta_T(0)} \leq \frac{kx^2}{\Delta_T(x) - \Delta_T(0)}, \quad \text{for } 0 < x \leq 1. \tag{A.5}$$

The above equation holds and its right-hand-side is positive since both $\Delta_S(x)$ and $\Delta_T(x)$ are strictly increasing. Thus, there is an optimal solution to the above optimization problem: $\alpha^{(1)} = 0$ and $\beta^{(1)} = \min_{0<x\leq 1}\left\{\frac{kx^2}{\Delta_T(x)-\Delta_T(0)}\right\}$. Then solution $(\alpha^{(1)}, \beta^{(1)}, \delta^{(1)}, x^{(1)})$ is a

candidate for the optimal solution of the original problem (P):

$$\begin{aligned}
\alpha^{(1)} &= 0;\\
\beta^{(1)} &= \min_{0<x\leq 1}\left\{\frac{kx^2}{\Delta_T(x)-\Delta_T(0)}\right\};\\
\delta^{(1)} &= -\frac{\beta^{(1)}\Delta_T(0)}{ap};\\
x^{(1)} &= 0;\\
U^{A(1)} &= 0;\\
U^{P(1)} &= c_1\Delta_S(0)+c_2\Delta_T(0)-ac_a.
\end{aligned} \tag{A.6}$$

Note that the optimal solution to (A.1) may be not unique. For example, if $\Delta_S(0)=0$ and $\Delta_T(0)=-M$, then we have another optimal solution: $\alpha=\min_{0<x\leq 1}\{kx^2/(\Delta_S(x)-\Delta_S(0))\}$ and $\beta=\delta=0$.

Second, we consider $x=1$. The best solution for this case can be found by solving

$$\begin{aligned}
&\max_{\alpha,\beta,\delta} \quad \{(c_1-\alpha)\Delta_S(1)+(c_2-\beta)\Delta_T(1)-ac_a-\delta ap\}\\
&\text{s.t.} \ \max_{0\leq x\leq 1}\left\{\alpha\Delta_S(x)+\beta\Delta_T(x)+\delta ap-kx^2\right\}\leq \alpha\Delta_S(1)+\beta\Delta_T(1)+\delta ap-k;\\
&\qquad \alpha\Delta_S(1)+\beta\Delta_T(1)+\delta ap-k\geq 0;\\
&\qquad \alpha,\beta,\delta\geq 0.
\end{aligned} \tag{A.7}$$

By Proposition A.1.0.1, (A.7) is equivalent to

$$\begin{aligned} &\min_{\alpha,\beta,\delta} \{\alpha\Delta_S(1) + \beta\Delta_T(1) + \delta ap - k\} \\ \text{s.t. } &\alpha(\Delta_S(x) - \Delta_S(1)) + \beta(\Delta_T(x) - \Delta_T(1)) \leq kx^2 - k, \quad 0 \leq x \leq 1; \\ &\alpha\Delta_S(1) + \beta\Delta_T(1) + \delta ap - k \geq 0; \\ &\alpha, \beta, \delta \geq 0. \end{aligned} \tag{A.8}$$

Similar to the case with $x = 0$, we introduce a new optimization problem

$$\begin{aligned} &\min_{\alpha,\beta} \{\alpha\Delta_S(1) + \beta\Delta_T(1)\} \\ \text{s.t. } &\alpha(\Delta_S(1) - \Delta_S(x)) + \beta(\Delta_T(1) - \Delta_T(x)) \geq k - kx^2, \quad 0 \leq x \leq 1; \\ &\alpha, \beta \geq 0. \end{aligned} \tag{A.9}$$

Proposition A.1.0.3. *If* $(\alpha^{(2)}, \beta^{(2)})$ *is an optimal solution of* (A.9)*, then* $(\alpha^{(2)}, \beta^{(2)}, \delta^{(2)})$ *is an optimal solution of* (A.2)*, where* $\delta^{(2)} = \max\left\{0, -\frac{\alpha^{(2)}\Delta_S(1)+\beta^{(2)}\Delta_T(1)-k}{ap}\right\}$.

For a given x, there are always $\{\alpha, \beta\}$ such that the constraint in (A.9) holds. On the other hand, the objective function is minimized if α is as small as possible, since $\Delta_S(1) > 0$, and β is as large as possible, since $\Delta_T(1) < 0$. Thus, there is an optimal solution to the above optimization problem (A.9): $\alpha^{(2)} = 0$, $\beta^{(2)} \geq \max_{0 \leq x < 1}\left\{\frac{k(1-x^2)}{\Delta_T(1)-\Delta_T(x)}\right\}$, $x^{(2)} = 1$, and $\delta^{(2)} = \frac{k-\beta^{(2)}\Delta_T(1)}{ap} > 0$. The solution $(\alpha^{(2)}, \beta^{(2)}, \delta^{(2)}, x^{(2)})$ is a candidate for the optimal solution of the original problem (P). The original objective function corresponding to the solution is $c_1\Delta_S(1) + c_2\Delta_T(1) - ac_a$. Consequently, a candidate for the optimal solution of

(P) is

$$
\begin{aligned}
\alpha^{(2)} &= 0; \\
\beta^{(2)} &\geq \max_{0 \leq x < 1} \left\{ \frac{k(1-x^2)}{\Delta_T(1) - \Delta_T(x)} \right\}; \\
\delta^{(2)} &= \frac{k - \beta^{(2)} \Delta_T(1)}{ap}; \\
x^{(2)} &= 1; \\
U^{A(2)} &= 0; \\
U^{P(2)} &= c_1 \Delta_S(1) + c_2 \Delta_T(1) - ac_a - k.
\end{aligned}
\tag{A.10}
$$

Now, we consider the case $0 < x < 1$. By applying the *first order condition*, the above optimization problem can be equivalent to

$$
\begin{aligned}
&\max_{\alpha,\beta,\delta,x} \quad \{(c_1 - \alpha)\Delta_S(x) + (c_2 - \beta)\Delta_T(x) - ac_a - \delta ap\} \\
&\text{s.t. } \alpha \Delta_S^{(1)}(x) + \beta \Delta_T^{(1)}(x) = 2kx; \\
&\qquad \alpha \Delta_S(x) + \beta \Delta_T(x) + \delta ap - kx^2 \geq 0; \\
&\qquad \alpha, \beta, \delta \geq 0, 0 < x < 1.
\end{aligned}
\tag{A.11}
$$

We choose $\delta = \max_{0 \leq x \leq 1} \left\{ 0, \frac{kx^2 - \alpha \Delta_S(x) - \beta \Delta_T(x)}{ap} \right\}$ to ensure that the second constraint holds and the objective function is maximized. Then, (A.11) becomes, for $\delta = 0$,

$$\begin{aligned} \max_{\alpha,\beta,x} \quad & \{(c_1 - \alpha)\Delta_S(x) + (c_2 - \beta)\Delta_T(x)\} \\ \text{s.t. } \; & \alpha\Delta_S^{(1)}(x) + \beta\Delta_T^{(1)}(x) = 2kx; \\ & \alpha\Delta_S(x) + \beta\Delta_T(x) - kx^2 \geq 0; \\ & \alpha, \beta \geq 0, 0 < x < 1, \end{aligned} \tag{A.12}$$

or, for $\delta > 0$,

$$\begin{aligned} \max_{\alpha,\beta,x} \quad & \{c_1\Delta_S(x) + c_2\Delta_T(x) - kx^2\} \\ \text{s.t. } \; & \alpha\Delta_S^{(1)}(x) + \beta\Delta_T^{(1)}(x) = 2kx; \\ & \alpha\Delta_S(x) + \beta\Delta_T(x) - kx^2 < 0; \\ & \alpha, \beta \geq 0, 0 < x < 1. \end{aligned} \tag{A.13}$$

For (A.12), given $0 < x < 1$, the constraint $\alpha\Delta_S^{(1)}(x) + \beta\Delta_T^{(1)}(x) = 2kx$ implies that the feasible (α, β) is contained in the line segment

$$L_x = \left\{(\alpha, \beta) = \xi\left(\frac{2kx}{\Delta_S^{(1)}(x)}, 0\right) + (1 - \xi)\left(0, \frac{2kx}{\Delta_T^{(1)}(x)}\right) : \; 0 \leq \xi \leq 1\right\}. \tag{A.14}$$

Note that both $\Delta_S^{(1)}(x)$ and $\Delta_T^{(1)}(x)$ are positive due to Assumption 2.4.0.1. However, not all (α, β) in L_x are feasible solutions. Using the constraint $\alpha\Delta_S(x) + \beta\Delta_T(x) - kx^2 \geq 0$, the set of feasible (α, β) is obtained as, if $2\Delta_S(x) \geq x\Delta_S^{(1)}(x)$,

$$L_{x0} = \left\{ \left(\xi \frac{2kx}{\Delta_S^{(1)}(x)},\ (1-\xi)\frac{2kx}{\Delta_T^{(1)}(x)} \right) :\ \underline{\xi}(x) = \frac{x - 2\frac{\Delta_T(x)}{\Delta_T^{(1)}(x)}}{2\left(\frac{\Delta_S(x)}{\Delta_S^{(1)}(x)} - \frac{\Delta_T(x)}{\Delta_T^{(1)}(x)}\right)} \leq \xi \leq 1 \right\}. \quad \text{(A.15)}$$

We note that the condition $2\Delta_S(x) \geq x\Delta_S^{(1)}(x)$ arises from setting $\beta = 0$; if the condition is satisfied it implies that the endpoint of the line segment where $\alpha \geq 0$, $\beta = 0$ is in the feasible set. This naturally raises the question: what happens if $2\Delta_S(x) < x\Delta_S^{(1)}(x)$? We show that this contradicts our participation constraint. Suppose $2\Delta_S(x) < x\Delta_S^{(1)}(x)$, and recall that we have already assumed that $\delta = 0$. We consider our agent's participation constraint:

$$\alpha\Delta_S(x) + \beta\Delta_T(x) \geq kx^2. \quad \text{(A.16)}$$

Next, we substitute in the maximum feasible value of α, $\frac{2kx}{\Delta_S^{(1)}(x)}$ while setting $\beta = 0$:

$$\frac{2kx}{\Delta_S^{(1)}(x)}\Delta_S(x) \geq kx^2. \quad \text{(A.17)}$$

Next, divide both sides by kx and multiply both sides by $\Delta_S^{(1)}(x)$:

$$2\Delta_S(x) \geq x\Delta_S^{(1)}(x). \quad \text{(A.18)}$$

Thus, given that our supposition directly contradicts our participation constraint, we see that (A.18) must hold for feasibility. This means that the maximum feasible value of α must be in the feasible set.

With the reasoning behind $2\Delta_S(x) \geq x\Delta_S^{(1)}(x)$ explained, we now proceed to explain the lowerbound weighting $\underline{\xi}(x)$. First, we restate the agent's participation constraint's left-hand side as a weighting between the maximum feasible values of α and β:

$$\xi \frac{2kx}{\Delta_S^{(1)}(x)} \Delta_S(x) + (1-\xi) \frac{2kx}{\Delta_T^{(1)}(x)} \Delta_T(x) \geq kx^2. \tag{A.19}$$

We distribute $\frac{2kx}{\Delta_T^{(1)}(x)} \Delta_T(x)$ on the left hand side, yielding:

$$\xi \frac{2kx}{\Delta_S^{(1)}(x)} \Delta_S(x) + \frac{2kx}{\Delta_T^{(1)}(x)} \Delta_T(x) - \xi \frac{2kx}{\Delta_T^{(1)}(x)} \Delta_T(x) \geq kx^2. \tag{A.20}$$

We isolate for ξ. We subtract $\frac{2kx}{\Delta_T^{(1)}(x)} \Delta_T(x)$ from both sides of (A.20):

$$\xi \left[\frac{2kx}{\Delta_S^{(1)}(x)} \Delta_S(x) - \frac{2kx}{\Delta_T^{(1)}(x)} \Delta_T(x) \right] \geq kx^2 - \frac{2kx}{\Delta_T^{(1)}(x)} \Delta_T(x). \tag{A.21}$$

Dividing both sides by $\frac{2kx}{\Delta_S^{(1)}(x)} \Delta_S(x) - \frac{2kx}{\Delta_T^{(1)}(x)} \Delta_T(x)$ yields:

$$\xi \geq \frac{x - \frac{2\Delta_T(x)}{\Delta_T^{(1)}(x)}}{2\left[\frac{\Delta_S(x)}{\Delta_S^{(1)}(x)} - \frac{\Delta_T(x)}{\Delta_T^{(1)}(x)} \right]} = \underline{\xi}. \tag{A.22}$$

The mapping $(c_1 - \alpha, c_2 - \beta)$ transforms L_{x0} into the following line segment:

$$L_{Px0} = \left\{ (c_1 - \alpha, c_2 - \beta) = \left(c_1 - \frac{\xi 2kx}{\Delta_S^{(1)}(x)}, c_2 - \frac{(1-\xi)2kx}{\Delta_T^{(1)}(x)} \right),\ \underline{\xi}(x) \leq \xi \leq 1 \right\}. \tag{A.23}$$

Then the maximum of the linear objective function $(c_1 - \alpha)\Delta_S(x) + (c_2 - \beta)\Delta_T(x)$ should be achieved at the end point of the line segment L_{Px0} for which β is as large as possible, i.e., $\alpha = \underline{\xi}(x)2kx/\Delta_S^{(1)}(x)$ and $\beta = (1-\underline{\xi}(x))2kx/\Delta_T^{(1)}(x)$. Consequently, we obtain

$$\begin{aligned}
&\max_{\alpha,\beta,x} \{(c_1-\alpha)\Delta_S(x) + (c_2-\beta)\Delta_T(x)\} \\
&= \max_x \left\{ \left(c_1 - \underline{\xi}(x)\frac{2kx}{\Delta_S^{(1)}(x)}\right)\Delta_S(x) + \left(c_2 - (1-\underline{\xi}(x))\frac{2kx}{\Delta_T^{(1)}(x)}\right)\Delta_T(x)\right\}.
\end{aligned} \tag{A.24}$$

Let $x^{(3)}$ be a point that the above function is maximized. Let $\alpha^{(3)} = \underline{\xi}(x^{(3)})\frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}$, $\beta^{(3)} = (1-\underline{\xi}(x^{(3)}))\frac{2kx^{(3)}}{\Delta_T^{(1)}(x^{(3)})}$, $\delta^{(3)} = 0$. If $2\Delta_S(x^{(3)}) \geq x^{(3)}\Delta_S^{(1)}(x^{(3)})$, the following solution is optimal to (A.12).

$$\begin{aligned}
\alpha^{(3)} &= \underline{\xi}(x^{(3)})\frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}; \\
\beta^{(3)} &= \left(1-\underline{\xi}(x^{(3)})\right)\frac{2kx^{(3)}}{\Delta_T^{(1)}(x^{(3)})}; \\
\delta^{(3)} &= 0; \\
x^{(3)} &= \underset{0<x<1:\ 2\Delta_S(x)\geq x\Delta_S^{(1)}(x)}{\operatorname{argmax}} F_3(x); \\
F_3(x) &= \left(c_1 - \frac{\underline{\xi}(x)2kx}{\Delta_S^{(1)}(x)}\right)\Delta_S(x) + \left(c_2 - \left(1-\underline{\xi}(x)\right)\frac{2kx}{\Delta_T^{(1)}(x)}\right)\Delta_T(x); \\
U^{A(3)} &= \alpha^{(3)}\Delta_S(x^{(3)}) + \beta^{(3)}\Delta_T(x^{(3)}) - k(x^{(3)})^2; \\
U^{P(3)} &= \left(c_1-\alpha^{(3)}\right)\Delta_S(x^{(3)}) + \left(c_2-\beta^{(3)}\right)\Delta_T(x^{(3)}) - ac_a.
\end{aligned} \tag{A.25}$$

Proposition A.1.0.4. *Solution* (A.25)*, if it exists, is optimal to optimization problem* (A.12)*.*

Now we consider when solution (A.25) exists, but is not an optimal solution to optimization problem (A.12). Suppose that $\{\hat{\alpha}, \hat{\beta}, \hat{x}\}$ is the optimal soluiton to (A.12). Once

$x = \hat{x}$ is given, then we can follow the above steps to show that $\hat{x}$ maximizes function $F_2(x)$ under the two constraints in (A.12). Consequently, we must have $\hat{x} = x^{(3)}$. Then $\hat{\alpha} = \alpha^{(3)}$ and $\hat{\beta} = \beta^{(3)}$. This completes the proof.

Next, we consider the optimal solution to (A.13). For this case, for any given x, the set of feasible solutions contains at least one pair of (α, β): $\alpha = 0$ and $\beta = 2kx/\Delta_T^{(1)}(x)$. In fact, the following set of (α, β) are all optimal solutions:

$$L_{x+} = \left\{ \left(\xi \frac{2kx}{\Delta_S^{(1)}(x)},\ (1-\xi)\frac{2kx}{\Delta_T^{(1)}(x)} \right) :\quad 0 \le \xi \le \hat{\xi}(x) = \max\left\{ 1,\ \frac{x - 2\frac{\Delta_T(x)}{\Delta_T^{(1)}(x)}}{2\left(\frac{\Delta_S(x)}{\Delta_S^{(1)}(x)} - \frac{\Delta_T(x)}{\Delta_T^{(1)}(x)}\right)} \right\} \right\}. \tag{A.26}$$

Let $x^{(4)}$ be maximizing the objective function of (A.13): $c_1\Delta_S(x) + c_2\Delta_T(x) - ac_a - kx^2$. Then the following solution is optimal to (A.13).

$$\begin{aligned}
\alpha^{(4)} &= 0; \\
\beta^{(4)} &= \frac{2kx^{(4)}}{\Delta_T^{(1)}(x^{(4)})}; \\
\delta^{(4)} &= \frac{1}{ap}\left(k\big(x^{(4)}\big)^2 - \frac{2kx^{(4)}}{\Delta_T^{(1)}(x^{(4)})}\Delta_T(x^{(4)}) \right); \\
x^{(4)} &= \underset{0<x<1}{\operatorname{argmax}}\{c_1\Delta_S(x) + c_2\Delta_T(x) - ac_a - kx^2\}; \\
U^{A(4)} &= 0; \\
U^{P(4)} &= c_1\Delta_S(x^{(4)}) + c_2\Delta_T(x^{(4)}) - ac_a - k(x^{(4)})^2.
\end{aligned} \tag{A.27}$$

Proposition A.1.0.5. *Solution* (A.27) *is optimal to optimization problem* (A.13).

Finally, we summarize the solutions to find the optimal solution to (PIPO).

Theorem A.1.1. *Consider the four solutions* $\{(\alpha^{(1)} = 0, \beta^{(1)}, \delta^{(1)}, x^{(1)}), (\alpha^{(2)} = 0, \beta^{(2)}, \delta^{(2)}, x^{(2)}), (\alpha^{(3)}, \beta^{(3)} = 0, \delta^{(3)}, x^{(3)}), (\alpha^{(4)}, \beta^{(4)}, \delta^{(4)}, x^{(4)})\}$*, given in equations* (A.6)*,* (A.10)*,* (A.25)*, and* (A.27)*. The one that maximizes* U^P *is the optimal solution of (PIPO).*

A.2 Candidate Comparison Conditions

With our solution candidates displayed, we now present conditions that compare Candidate 3 with the other solutions candidates. We focus on Candidate 3 as it is the only candidate wherein $\alpha > 0$. The proofs of these conditions are presented in Appendix A.3.

Proposition A.2.0.1. *If* (A.28) *does not hold, then Candidate 3 is not an optimal solution:*

$$\frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})} \geq \frac{k(x^{(3)})^2}{\Delta_S(x^{(3)})}. \tag{A.28}$$

If (A.29) *holds, then Candidate 3 a superior solution to Candidate 1:*

$$c_2 M \geq \Big(\frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})} - c_1\Big)\Delta_S(x^{(3)}) - c_2\Delta_T(x^{(3)}) \tag{A.29}$$

If (A.30) *holds, then Candidate 3 a superior solution to Candidate 2:*

$$k \geq c_1\Big(\Delta_S(1) - \Delta_S(x^{(3)})\Big) + c_2\Big(\Delta_T(1) - \Delta_T(x^{(3)})\Big) + \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Delta_S(x^{(3)}) \geq 0. \tag{A.30}$$

If (A.31) *holds, then Candidate 3 a superior solution to Candidate 4:*

$$\Big(c_1 - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Big)\Delta_S(x^{(3)}) + c_2\Delta_T(x^{(3)}) \geq c_1\Delta_S(x^{(4)}) + c_2\Delta_T(x^{(4)}) - k(x^{(4)})^2. \quad \text{(A.31)}$$

Of these conditions, (A.28) is perhaps the most useful. Namely, if one evaluates both sides of the inequality for values of x in $(0, 1)$, it provides a useful approximation of which x could be optimal, and which need not be considered. This could lead to a more efficient evaluation of $F_2(x)$, if a smaller range of x can be considered. It could also outright remove Candidate 3 from consideration prior to computing each candidate solution.

Conditions (A.29) and (A.30) can be computed similarly to condition (A.28), before any optimization takes place. Namely, as the value of x changes from 0 to 1, we can see which of conditions (A.29) and (A.30) are satisfied. If condition (A.29) is not satisfied, it indicates that the more complicated structure of Candidate 3 is too expensive for the given effort level x, so the zero effort level is preferred. If condition (A.29) is not satisfied, it implies that the principal can more-profitably extract more effort from the agent, but they must use a more complicated contract.

Condition (A.31) requires a more expensive computation to check (since we have to check all pairs of x). However, if one candidate uniformly dominates the other (either Candidate 3 or 4), then one can be eliminated from consideration.

A.3 Proofs of Candidate Comparison Conditions

Below we restate the various components of Proposition A.2.0.1 and prove each in turn.

Proposition A.2.0.1 first condition: *If* (A.32) *does not hold, then Candidate 3 is* *not an optimal solution:*

$$\frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})} \geq \frac{k(x^{(3)})^2}{\Delta_S(x^{(3)})}. \tag{A.32}$$

Proof. To begin, we know that based on arguments in Section A.1, this candidate is only valid if $k(x^{(3)})^2 - \alpha\Delta_S(x^{(3)}) - \beta\Delta_T(x^{(3)}) \leq 0$. In (2.24), $\beta^* = 0$. As such, the following holds:

$$k(x^{(3)})^2 - \alpha\Delta_S(x^{(3)}) \leq 0. \tag{A.33}$$

Adding $\alpha\Delta_S(x^{(3)})$ to both sides of (A.33) and dividing both sides by $\Delta_S(x^{(3)})$ (which, by assumption, is positive), yields:

$$\alpha \geq \frac{k(x^{(3)})^2}{\Delta_S(x^{(3)})}. \tag{A.34}$$

Further, we know that, via the line-segment argument in A.1, the candidate optimal α takes on the value:

$$\alpha = \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}. \tag{A.35}$$

Substituting (A.35) in (A.34) yields:

$$\frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})} \geq \frac{k(x^{(3)})^2}{\Delta_S(x^{(3)})}. \tag{A.36}$$

Thus, we arrive at our result. □

Proposition A.2.0.1 second condition: *If (A.37) holds, then Candidate 3 a superior*

solution to Candidate 1:

$$c_2 M \geq \Big(\frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})} - c_1\Big)\Delta_S(x^{(3)}) - c_2\Delta_T(x^{(3)}). \tag{A.37}$$

Proof. We consider how Candidate 3 compares with Candidate 1. Suppose (2.24) is optimal. This implies the following:

$$\Big(c_1 - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Big)\Delta_S(x^{(3)}) + c_2\Delta_T(x^{(3)}) - ac_a \geq c_1\Delta_S(0) + c_2\Delta_T(0) - ac_a. \tag{A.38}$$

Recalling that $\Delta_T(0) = -M$ and $\Delta_S(0) = 0$ we see the following:

$$\Big(c_1 - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Big)\Delta_S(x^{(3)}) + c_2\Delta_T(x^{(3)}) - ac_a \geq -c_2 M - ac_a. \tag{A.39}$$

We subtract $-c_2M - ac_a$ from both sides of (A.39), yielding:

$$\Big(c_1 - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Big)\Delta_S(x^{(3)}) + c_2\Delta_T(x^{(3)}) + c_2 M \geq 0. \tag{A.40}$$

Subtracting all terms except c_2M from both sides of (A.40) yields:

$$c_2 M \geq \Big(\frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})} - c_1\Big)\Delta_S(x^{(3)}) - c_2\Delta_T(x^{(3)}). \tag{A.41}$$

This concludes the proof. □

Proposition A.2.0.1 third condition: *If* (A.42) *holds, then Candidate 3 a superior*

solution to Candidate 2:

$$k \geq c_1\Big(\Delta_S(1) - \Delta_S(x^{(3)})\Big) + c_2\Big(\Delta_T(1) - \Delta_T(x^{(3)})\Big) + \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Delta_S(x^{(3)}) \geq 0. \quad \text{(A.42)}$$

Proof. We consider how Candidate 3 compares with the Candidate 2. Suppose (2.24) is optimal. This implies the following:

$$\Big(c_1 - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Big)\Delta_S(x^{(3)}) + c_2\Delta_T(x^{(3)}) - ac_a \geq c_1\Delta_S(1) + c_2\Delta_T(1) - ac_a - k. \quad \text{(A.43)}$$

If we subtract the right hand side of (A.43) from both sides it yields:

$$c_1\Big(\Delta_S(x^{(3)}) - \Delta_S(1)\Big) - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Delta_S(x^{(3)}) + c_2\Big(\Delta_T(x^{(3)}) - \Delta_T(1)\Big) + k \geq 0. \quad \text{(A.44)}$$

We have assumed that both $\Delta_S(x)$ and $\Delta_T(x)$ are concave in x and that both are improving in x monotonically. As such, the following must hold:

$$\Delta_S(x^{(3)}) - \Delta_S(1) \leq 0; \quad \text{(A.45)}$$

$$\Delta_T(x^{(3)}) - \Delta_T(1) \leq 0. \quad \text{(A.46)}$$

With (A.45), we see that the only positive term in (A.44) is 1. Thus, we have the following necessary condition for a solution with $\alpha > 0$ to be optimal:

$$k \geq c_1\Big(\Delta_S(1) - \Delta_S(x^{(3)})\Big) + c_2\Big(\Delta_T(1) - \Delta_T(x^{(3)})\Big) + \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Delta_S(x^{(3)}) \geq 0. \quad \text{(A.47)}$$

This concludes the proof. □

Proposition A.2.0.1 fourth condition: *If* (A.48) *holds, then Candidate 3 a superior*

solution to Candidate 4:

$$\Big(c_1 - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Big)\Delta_S(x^{(3)}) + c_2\Delta_T(x^{(3)}) \geq c_1\Delta_S(x^{(4)}) + c_2\Delta_T(x^{(4)}) - k(x^{(4)})^2. \tag{A.48}$$

Proof. We consider how Candidate 3 compares with Candidate 4. Suppose (2.24) is optimal. This implies the following:

$$\Big(c_1 - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Big)\Delta_S(x^{(3)}) + c_2\Delta_T(x^{(3)}) - ac_a \geq c_1\Delta_S(x^{(4)}) + c_2\Delta_T(x^{(4)}) - ac_a - k(x^{(4)})^2. \tag{A.49}$$

Cancelling ac_a from both sides of (A.49) leaves us with:

$$\Big(c_1 - \frac{2kx^{(3)}}{\Delta_S^{(1)}(x^{(3)})}\Big)\Delta_S(x^{(3)}) + c_2\Delta_T(x^{(3)}) \geq c_1\Delta_S(x^{(4)}) + c_2\Delta_T(x^{(4)}) - k(x^{(4)})^2. \tag{A.50}$$

This concludes the proof. □

A.4 Operational Protocol Modification Proofs

In proving Properties 2.6.7.1 and 2.6.15.1, we first establish some necessary quantities in Appendices A.4.1 and A.4.2. We then proceed to derive the Properties in Appendix A.4.3.

A.4.1 $\Delta_{\hat{S}}(x)$

Let $\hat{S}$ be a random variable denoting service quality. Let $\hat{S}$ be distributed according to a censored beta distribution with parameters $A_{x,d}$ and B_d, with censoring points $0 \leq \tau_L \leq$

$\tau_H \leq M$. Note that $A_{x,d}$ is dependent on x, the agent's effort decision variable. Let the difficulty states be $d = 0$ for the non-difficult state and $d = 1$ for the difficult state. Let the parameters $A_{x,0}$ and B_0 be for the non-difficult state, and $A_{x,1}$ and B_1 be for the difficult state.

We can express the expectation $\mathrm{E}[\hat{S}] = \Delta_{\hat{S}}(x)$ as:

$$\Delta_{\hat{S}}(x) = \begin{aligned} &= p\Bigg[M\tau_L \int_0^{\tau_L} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}-1}(1-s)^{B_1-1}ds \\ &+ M\int_{\tau_L}^{\tau_H} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}}(1-s)^{B_1-1}ds \\ &+ M\tau_H \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}-1}(1-s)^{B-1}ds\Bigg] \\ &+ (1-p)\Bigg[M\tau_L \int_0^{\tau_L} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}-1}(1-s)^{B_0-1}ds \\ &+ M\int_{\tau_L}^{\tau_H} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}}(1-s)^{B_0-1}ds \\ &+ M\tau_H \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}-1}(1-s)^{B_0-1}ds\Bigg]. \end{aligned} \tag{A.51}$$

A.4.2 $\Delta_{\hat{T}}(x)$

Let $\hat{T}$ be a random variable denoting service time. Let $\hat{T}$ be distributed according to a censored, truncated exponential distribution with parameter $\lambda_{x,d}$, with censoring points $0 \leq \Upsilon_L \leq \Upsilon_H \leq M$. Let the difficulty states be $d = 0$ for the non-difficult state and $d = 1$ for the difficult state. Let the parameter $\lambda_{x,0}$ be for the non-difficult state, and $\lambda_{x,1}$ be for the difficult state.

We can express the expectation $\mathrm{E}[\hat{T}]$ as

$$\mathrm{E}[\hat{T}] = \begin{array}{r} p \cdot \dfrac{1}{1-e^{-\lambda_{x,1}M}}\left[\Upsilon_L - \Upsilon_H e^{-\lambda_{x,1}M} + \dfrac{e^{-\lambda_{x,1}\Upsilon_L} - e^{-\lambda_{x,1}\Upsilon_H}}{\lambda_{x,1}}\right] \\ +(1-p) \cdot \dfrac{1}{1-e^{-\lambda_{x,0}M}}\left[\Upsilon_L - \Upsilon_H e^{-\lambda_{x,0}M} + \dfrac{e^{-\lambda_{x,0}\Upsilon_L} - e^{-\lambda_{x,0}\Upsilon_H}}{\lambda_{x,0}}\right]. \end{array} \tag{A.52}$$

Let λ_x be a function of x (effort impacts the rate). Let $\Delta_{\hat{T}}(x) = -\mathrm{E}[\hat{T}]$. Now we write $\Delta_{\hat{T}}(x)$ as follows:

$$\Delta_{\hat{T}}(x) = \begin{array}{r} -p \cdot \dfrac{1}{1-e^{-\lambda_{x,1}M}}\left[\Upsilon_L - \Upsilon_H e^{-\lambda_{x,1}M} + \dfrac{e^{-\lambda_{x,1}\Upsilon_L} - e^{-\lambda_{x,1}\Upsilon_H}}{\lambda_{x,1}}\right] \\ -(1-p) \cdot \dfrac{1}{1-e^{-\lambda_{x,0}M}}\left[\Upsilon_L - \Upsilon_H e^{-\lambda_{x,0}M} + \dfrac{e^{-\lambda_{x,0}\Upsilon_L} - e^{-\lambda_{x,0}\Upsilon_H}}{\lambda_{x,0}}\right]. \end{array}. \tag{A.53}$$

A.4.3 Property Derivations

We begin by restating Candidate 4:

$$\begin{aligned} \alpha^{(4)} &= 0; \\ \beta^{(4)} &= \frac{2kx^{(4)}}{\Delta_T^{(1)}(x^{(4)})}; \\ \delta^{(4)} &= \frac{1}{ap}\left(k\big(x^{(4)}\big)^2 - \frac{2kx^{(4)}}{\Delta_T^{(1)}(x^{(4)})}\Delta_T(x^{(4)})\right); \\ x^{(4)} &= \operatorname*{argmax}_{0<x<1}\{c_1\Delta_S(x) + c_2\Delta_T(x) - ac_a - kx^2\}; \\ U^{A(3)} &= 0; \\ U^{P(3)} &= c_1\Delta_S(x^{(4)}) + c_2\Delta_T(x^{(4)}) - ac_a - k(x^{(4)})^2. \end{aligned} \tag{A.54}$$

A.4.3.1 $\hat{S}$ Censoring

In the solution, x is determined first. As such, we reconsider the function x is chosen to maximize, which is the principal's utility function:

$$U^P = c_1\Delta_{\hat{S}}(x) + c_2\Delta_T(x) - ac_a - kx^2. \tag{A.55}$$

Now we consider this with the expression for $\Delta_{\hat{S}}(x)$:

$$\begin{aligned}
&= c_1 p\Bigg[M\tau_L \int_0^{\tau_L} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}-1}(1-s)^{B_1-1}dx \\
&+ M\int_{\tau_L}^{\tau_H} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}}(1-s)^{B_1-1}dx \\
&+ M\tau_H \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}-1}(1-s)^{B-1}dx\Bigg] \\
&+ c_1(1-p)\Bigg[M\tau_L \int_0^{\tau_L} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}-1}(1-s)^{B_0-1}dx \\
&+ M\int_{\tau_L}^{\tau_H} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}}(1-s)^{B_0-1}dx \\
&+ M\tau_H \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}-1}(1-s)^{B_0-1}dx\Bigg] \\
&+ c_2\Delta_T(x) - ac_a - kx^2.
\end{aligned} \tag{A.56}$$

We investigate derivatives of (A.56) with respect to the censoring parameters τ_L and τ_H.

Derivatives with respect to τ_L

We state this first derivative $\partial U^P(x^*)/\partial\tau_L$ as:

$$\begin{aligned}\frac{\partial U^P(x^*)}{\partial\tau_L} &= \frac{\partial}{\partial\tau_L}c_1 p\Bigg[M\tau_L\int_0^{\tau_L}\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}s^{A_{x,1}-1}(1-s)^{B_1-1}ds \\ &+ M\int_{\tau_L}^{\tau_H}\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}s^{A_{x,1}}(1-s)^{B_1-1}ds \\ &+ M\tau_H\int_{\tau_H}^{1}\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}s^{A_{x,1}-1}(1-s)^{B-1}ds\Bigg] \\ &+ \frac{\partial}{\partial\tau_L}c_1(1-p)\Bigg[M\tau_L\int_0^{\tau_L}\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}s^{A_{x,0}-1}(1-s)^{B_0-1}ds \\ &+ M\int_{\tau_L}^{\tau_H}\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}s^{A_{x,0}}(1-s)^{B_0-1}ds \\ &+ M\tau_H\int_{\tau_H}^{1}\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}s^{A_{x,0}-1}(1-s)^{B_0-1}ds\Bigg].\end{aligned} \tag{A.57}$$

We note that we can drop any terms in (A.57) that do not contain τ_L:

$$\begin{aligned}\frac{\partial U^P(x^*)}{\partial\tau_L} &= \frac{\partial}{\partial\tau_L}c_1 p\Bigg[M\tau_L\int_0^{\tau_L}\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}s^{A_{x,1}-1}(1-s)^{B_1-1}ds \\ &+ M\int_{\tau_L}^{\tau_H}\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}s^{A_{x,1}}(1-s)^{B_1-1}ds\Bigg] \\ &+ \frac{\partial}{\partial\tau_L}c_1(1-p)\Bigg[M\tau_L\int_0^{\tau_L}\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}s^{A_{x,0}-1}(1-s)^{B_0-1}ds \\ &+ M\int_{\tau_L}^{\tau_H}\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}s^{A_{x,0}}(1-s)^{B_0-1}ds\Bigg].\end{aligned} \tag{A.58}$$

Taking the derivatives (using the second Fundamental Theorem of Calculus and the product rule for differentiation), we obtain the following:

$$
\begin{aligned}
\frac{\partial U^P(x^*)}{\partial \tau_L} &= c_1 p\Bigg[M\int_0^{\tau_L}\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}s^{A_{x,1}-1}(1-s)^{B_1-1}ds \\
&+ M\tau_L\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}\tau_L^{A_{x,1}-1}(1-\tau_L)^{B_1-1} \\
&- M\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}\tau_L^{A_{x,1}}(1-\tau_L)^{B_1-1}\Bigg] \\
&+ c_1(1-p)\Bigg[M\int_0^{\tau_L}\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}s^{A_{x,0}-1}(1-s)^{B_0-1}ds \\
&+ M\tau_L\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}\tau_L^{A_{x,0}-1}(1-\tau_L)^{B_0-1} \\
&- M\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}\tau_L^{A_{x,0}}(1-\tau_L)^{B_0-1}\Bigg].
\end{aligned}
\tag{A.59}
$$

In simplifying, we note that the second and third terms in each pair of brackets cancel, leading to the following:

$$
\begin{aligned}
\frac{\partial U^P(x^*)}{\partial \tau_L} &= c_1 pM\int_0^{\tau_L}\frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt\int_0^\infty t^{B_1-1}e^{-t}dt}s^{A_{x,1}-1}(1-s)^{B_1-1}ds \\
&+ c_1(1-p)M\int_0^{\tau_L}\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt\int_0^\infty t^{B_0-1}e^{-t}dt}s^{A_{x,0}-1}(1-s)^{B_0-1}ds.
\end{aligned}
\tag{A.60}
$$

Now we determine the sign of (A.60). We note that the coefficients c_1pM and $c_1(1-p)M$ are nonnegative by definition (c_1 is the only term that is nonnegative, not positive). Next, we consider the fractions resulting from the gamma functions:

$$\frac{\Gamma(A_{x,1}+B_1)}{\Gamma(A_{x,1})\Gamma(B_1)} = \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt}; \tag{A.61}$$

$$\frac{\Gamma(A_{x,0}+B_0)}{\Gamma(A_{x,0})\Gamma(B_0)} = \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt}. \tag{A.62}$$

Given that the integrations in (A.61) and (A.62) are over nonnegative ranges, the numerators and denominators of both fractions are positive. As such, (A.61) and (A.62) are positve. We now consider the remaining portions of (A.60):

$$\int_0^{\tau_L} s^{A_{x,1}-1}(1-s)^{B_1-1}ds \tag{A.63}$$

$$\int_0^{\tau_L} s^{A_{x,0}-1}(1-s)^{B_0-1}ds. \tag{A.64}$$

We see that (A.63) and (A.64) are positive based on the integration ranges.

We observe that all the portions of (A.60) are positive or nonnegative, so (A.60) is nonnegative. As such, increasing τ_L will increase the principal's expected utility, which aligns with our intuitive understanding of the distribution of service quality.

Now we consider the second derivative with respect to τ_L. We begin by restating the derivative of the principal's utility function with respect to τ_L.

$$\begin{aligned}\frac{\partial U^P(x^*)}{\partial \tau_L} &= c_1 p M \int_0^{\tau_L} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}-1}(1-s)^{B_1-1}ds \\ &+ c_1(1-p)M \int_0^{\tau_L} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}-1}(1-s)^{B_0-1}ds.\end{aligned} \tag{A.65}$$

Taking the derivative with respect to τ_L (once again using the second Fundamental

Theorem) yields:

$$\begin{aligned}\frac{\partial^2 U^P(x^*)}{\partial \tau_L^2} &= c_1 p M \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt}\tau_L^{A_{x,1}-1}(1-\tau_L)^{B_1-1} \\ &\quad + c_1(1-p)M\frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt}\tau_L^{A_{x,0}-1}(1-\tau_L)^{B_0-1}.\end{aligned} \tag{A.66}$$

By similar arguments to our discussion of the sign of $\partial U^P(x^*)/\partial \tau_L$, we can see that (A.66) is nonnegative. As such, increasing τ_L yields a convex, increasing response in $U^P(x^*)$.

Derivatives with respect to τ_H We state this first derivative $\partial U^P(x^*)/\partial \tau_H$ as:

$$\begin{aligned}\frac{\partial U^P(x^*)}{\partial \tau_H} &= \frac{\partial}{\partial \tau_H} c_1 p \Bigg[M\tau_L \int_0^{\tau_L} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}-1}(1-s)^{B_1-1}ds \\ &\quad + M\int_{\tau_L}^{\tau_H} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}}(1-s)^{B_1-1}ds \\ &\quad + M\tau_H \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,1}+B_1-1}e^{-t}dt}{\int_0^\infty t^{A_{x,1}-1}e^{-t}dt \int_0^\infty t^{B_1-1}e^{-t}dt} s^{A_{x,1}-1}(1-s)^{B-1}ds \Bigg] \\ &\quad + \frac{\partial}{\partial \tau_H} c_1(1-p) \Bigg[M\tau_L \int_0^{\tau_L} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}-1}(1-s)^{B_0-1}ds \\ &\quad + M\int_{\tau_L}^{\tau_H} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}}(1-s)^{B_0-1}ds \\ &\quad + M\tau_H \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,0}+B_0-1}e^{-t}dt}{\int_0^\infty t^{A_{x,0}-1}e^{-t}dt \int_0^\infty t^{B_0-1}e^{-t}dt} s^{A_{x,0}-1}(1-s)^{B_0-1}ds \Bigg].\end{aligned} \tag{A.67}$$

We note that we can drop any terms in (A.67) that do not contain τ_H:

$$\begin{aligned}\frac{\partial U^P(x^*)}{\partial \tau_H} &= \frac{\partial}{\partial \tau_H} c_1 p \Bigg[M \int_{\tau_L}^{\tau_H} \frac{\int_0^\infty t^{A_{x,1}+B_1-1} e^{-t} dt}{\int_0^\infty t^{A_{x,1}-1} e^{-t} dt \int_0^\infty t^{B_1-1} e^{-t} dt} s^{A_{x,1}} (1-s)^{B_1-1} ds \\ &\quad + M\tau_H \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,1}+B_1-1} e^{-t} dt}{\int_0^\infty t^{A_{x,1}-1} e^{-t} dt \int_0^\infty t^{B_1-1} e^{-t} dt} s^{A_{x,1}-1} (1-s)^{B-1} ds \Bigg] \\ &\quad + \frac{\partial}{\partial \tau_H} c_1 (1-p) \Bigg[M \int_{\tau_L}^{\tau_H} \frac{\int_0^\infty t^{A_{x,0}+B_0-1} e^{-t} dt}{\int_0^\infty t^{A_{x,0}-1} e^{-t} dt \int_0^\infty t^{B_0-1} e^{-t} dt} s^{A_{x,0}} (1-s)^{B_0-1} ds \\ &\quad + M\tau_H \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,0}+B_0-1} e^{-t} dt}{\int_0^\infty t^{A_{x,0}-1} e^{-t} dt \int_0^\infty t^{B_0-1} e^{-t} dt} s^{A_{x,0}-1} (1-s)^{B_0-1} ds \Bigg].\end{aligned} \tag{A.68}$$

Invoking the product rule and the second Fundamental Theorem yields:

$$\begin{aligned}\frac{\partial U^P(x^*)}{\partial \tau_H} &= c_1 p \Bigg[M \frac{\int_0^\infty t^{A_{x,1}+B_1-1} e^{-t} dt}{\int_0^\infty t^{A_{x,1}-1} e^{-t} dt \int_0^\infty t^{B_1-1} e^{-t} dt} \tau_H^{A_{x,1}} (1-\tau_H)^{B_1-1} \\ &\quad + M \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,1}+B_1-1} e^{-t} dt}{\int_0^\infty t^{A_{x,1}-1} e^{-t} dt \int_0^\infty t^{B_1-1} e^{-t} dt} s^{A_{x,1}-1} (1-s)^{B-1} ds \\ &\quad - M\tau_H \frac{\int_0^\infty t^{A_{x,1}+B_1-1} e^{-t} dt}{\int_0^\infty t^{A_{x,1}-1} e^{-t} dt \int_0^\infty t^{B_1-1} e^{-t} dt} \tau_H^{A_{x,1}-1} (1-\tau_H)^{B_1-1} \Bigg] \\ &\quad + c_1 (1-p) \Bigg[M \frac{\int_0^\infty t^{A_{x,0}+B_0-1} e^{-t} dt}{\int_0^\infty t^{A_{x,0}-1} e^{-t} dt \int_0^\infty t^{B_0-1} e^{-t} dt} \tau_H^{A_{x,0}} (1-\tau_H)^{B_0-1} \\ &\quad + M \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,0}+B_0-1} e^{-t} dt}{\int_0^\infty t^{A_{x,0}-1} e^{-t} dt \int_0^\infty t^{B_0-1} e^{-t} dt} s^{A_{x,0}-1} (1-s)^{B_0-1} ds \\ &\quad - M\tau_H \frac{\int_0^\infty t^{A_{x,0}+B_0-1} e^{-t} dt}{\int_0^\infty t^{A_{x,0}-1} e^{-t} dt \int_0^\infty t^{B_0-1} e^{-t} dt} \tau_H^{A_{x,0}-1} (1-\tau_H)^{B_0-1} \Bigg].\end{aligned} \tag{A.69}$$

Cancelling terms in (A.69) yields:

$$\frac{\partial U^P(x^*)}{\partial \tau_H} = c_1 p M \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,1}+B_1-1} e^{-t} dt}{\int_0^\infty t^{A_{x,1}-1} e^{-t} dt \int_0^\infty t^{B_1-1} e^{-t} dt} s^{A_{x,1}-1} (1-s)^{B-1} ds \\ + c_1(1-p) M \int_{\tau_H}^{1} \frac{\int_0^\infty t^{A_{x,0}+B_0-1} e^{-t} dt}{\int_0^\infty t^{A_{x,0}-1} e^{-t} dt \int_0^\infty t^{B_0-1} e^{-t} dt} s^{A_{x,0}-1} (1-s)^{B_0-1} ds. \tag{A.70}$$

By similar arguments to τ_L, we see that (A.70) is nonnegative. As such, increasing the upperbound τ_H on service quality improves the principal's expected utility, which confirms our intuition.

We now compute the second derivative of $U^P(x^*)$ with respect fo τ_H. Once again invoking the second Fundamental Theorem yields:

$$\frac{\partial^2 U^P(x^*)}{\partial \tau_H^2} = -c_1 p M \frac{\int_0^\infty t^{A_{x,1}+B_1-1} e^{-t} dt}{\int_0^\infty t^{A_{x,1}-1} e^{-t} dt \int_0^\infty t^{B_1-1} e^{-t} dt} \tau_H^{A_{x,1}-1} (1-\tau_H)^{B-1} \\ - c_1(1-p) M \frac{\int_0^\infty t^{A_{x,0}+B_0-1} e^{-t} dt}{\int_0^\infty t^{A_{x,0}-1} e^{-t} dt \int_0^\infty t^{B_0-1} e^{-t} dt} \tau_H^{A_{x,0}-1} (1-\tau_H)^{B_0-1}. \tag{A.71}$$

We note that both terms in (A.71) are nonpositive based on previous arguments, so increasing τ_H yields concave increases in $U^P(x^*)$ and confirms our numerical results.

A.4.3.2 $\hat{T}$ Censoring

We again reconsider the principal's expected utility; this time where service time $\hat{T}$ is censored. We restate the principal's expected utility as:

$$U^P = c_1 \Delta_S(x) + c_2 \Delta_{\hat{T}}(x) - a c_a - k x^2. \tag{A.72}$$

Now we consider this with the expression for $\Delta_{\hat{T}}(x)$:

$$\begin{aligned} & c_1\Delta_S(x) - p \cdot \frac{c_2}{1-e^{-\lambda_{x,1}M}}\left[\Upsilon_L - \Upsilon_H e^{-\lambda_{x,1}M} + \frac{e^{-\lambda_{x,1}\Upsilon_L} - e^{-\lambda_{x,1}\Upsilon_H}}{\lambda_{x,1}}\right] \\ & - (1-p)\cdot \frac{c_2}{1-e^{-\lambda_{x,0}M}}\left[\Upsilon_L - \Upsilon_H e^{-\lambda_{x,0}M} + \frac{e^{-\lambda_{x,0}\Upsilon_L} - e^{-\lambda_{x,0}\Upsilon_H}}{\lambda_{x,0}}\right] - ac_a - kx^2. \end{aligned} \tag{A.73}$$

We investigate the derivatives of (A.73) with respect to Υ_L and Υ_H.

Derivatives with respect to Υ_L

First, we consider derivatives with respect to Υ_L:

$$\begin{aligned} \frac{\partial U^P(x^*)}{\partial \Upsilon_L} = & \frac{-pc_2}{1-e^{-\lambda_{x,1}M}} \\ & + \frac{\begin{array}{c}(pc_2\lambda_{x,1}e^{-\lambda_{x,1}\Upsilon_L}+0)(\lambda_{x,1}(1-e^{-\lambda_{x,1}M})) \\ -(-pc_2e^{-\lambda_{x,1}\Upsilon_L}+pc_2e^{-\lambda_{x,1}\Upsilon_H})(0)\end{array}}{(\lambda_{x,1}(1-e^{-\lambda_{x,1}M}))^2} \\ & + \frac{(p-1)c_2}{1-e^{-\lambda_{x,0}M}} \\ & + \frac{\begin{array}{c}((1-p)c_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_L}+0)(\lambda_{x,0}(1-e^{-\lambda_{x,0}M})) \\ -((p-1)c_2e^{-\lambda_{x,0}\Upsilon_L}+(1-p)c_2e^{-\lambda_{x,0}\Upsilon_H})(0)\end{array}}{(\lambda_{x,0}(1-e^{-\lambda_{x,0}M}))^2}. \end{aligned} \tag{A.74}$$

Simplifying yields:

$$\begin{aligned} &= \frac{-pc_2}{1-e^{-\lambda_{x,1}M}} + \frac{pc_2e^{-\lambda_{x,1}\Upsilon_L}}{1-e^{-\lambda_{x,1}M}} \\ &+ \frac{(p-1)c_2}{1-e^{-\lambda_{x,0}M}} + \frac{(1-p)c_2e^{-\lambda_{x,0}\Upsilon_L}}{1-e^{-\lambda_{x,0}M}}. \end{aligned} \tag{A.75}$$

We place the two components over common denominators:

$$\begin{aligned} &= \frac{-pc_2 + pc_2 e^{-\lambda_{x,1}\Upsilon_L}}{1 - e^{-\lambda_{x,1}M}} \\ &+ \frac{-(1-p)c_2 + (1-p)c_2 e^{-\lambda_{x,0}\Upsilon_L}}{1 - e^{-\lambda_{x,0}M}}. \end{aligned} \tag{A.76}$$

Further simplifying yields:

$$\frac{\partial U^P(x^*)}{\partial \Upsilon_L} = -pc_2 \frac{1 - e^{-\lambda_{x,1}\Upsilon_L}}{1 - e^{-\lambda_{x,1}M}} - (1-p)c_2 \frac{1 - e^{-\lambda_{x,0}\Upsilon_L}}{1 - e^{-\lambda_{x,0}M}}. \tag{A.77}$$

Consider the first fraction:

$$\frac{1 - e^{-\lambda_{x,1}\Upsilon_L}}{1 - e^{-\lambda_{x,1}M}}. \tag{A.78}$$

The pre-multiplication by $-pc_2$ implies that this term is nonpositive in (A.77). A similar argument implies the second fraction is similar in (A.77). Expanding the second term, we have:

$$-c_2 \frac{1 - e^{-\lambda_{x,0}\Upsilon_L}}{1 - e^{-\lambda_{x,0}M}} + pc_2 \frac{1 - e^{-\lambda_{x,0}\Upsilon_L}}{1 - e^{-\lambda_{x,0}M}}. \tag{A.79}$$

As $0 < p \leq 1$ by assumption, we see that this term is nonpositive.

Therefore, (A.77) is nonpositive.

We now consider the second derivative with respect to Υ_L:

$$\frac{\partial^2 U^P(x^*)}{\partial \Upsilon_L^2} = \frac{(-pc_2\lambda_{x,1}e^{-\lambda_{x,1}\Upsilon_L})(1-e^{-\lambda_{x,1}M}) - (-pc_2 + pc_2 e^{-\lambda_{x,1}\Upsilon_L})(0)}{(1-e^{-\lambda_{x,1}M})^2} + \frac{\begin{array}{c}(-c_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_L} + pc_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_L})(1-e^{-\lambda_{x,0}M}) \\ -(-c_2 + pc_2 + c_2 e^{-\lambda_{x,0}\Upsilon_L} - pc_2 e^{-\lambda_{x,0}\Upsilon_L})(0)\end{array}}{(1-e^{-\lambda_{x,0}M})^2} \tag{A.80}$$

Simplifying yields:

$$= \frac{-pc_2\lambda_{x,1}e^{-\lambda_{x,1}\Upsilon_L}}{1-e^{-\lambda_{x,1}M}} + \frac{-c_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_L} + pc_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_L}}{1-e^{-\lambda_{x,0}M}}. \tag{A.81}$$

Inspecting (A.81) we see that the first fraction is nonpositive. We also see that the numerator of the second fraction is nonpositive as it is $(p-1)c_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_L}$, and by assumption $0 < p \leq 1$. As such, the whole expression is nonpositive. This means that increases in Υ_L yield concave decreases in the principal's expected utility.

Derivatives with respect to Υ_H

We first restate the principal's optimal utility function:

$$c_1\Delta_S(x) - p \cdot \frac{c_2}{1-e^{-\lambda_{x,1}M}}\left[\Upsilon_L - \Upsilon_H e^{-\lambda_{x,1}M} + \frac{e^{-\lambda_{x,1}\Upsilon_L} - e^{-\lambda_{x,1}\Upsilon_H}}{\lambda_{x,1}}\right] - (1-p) \cdot \frac{c_2}{1-e^{-\lambda_{x,0}M}}\left[\Upsilon_L - \Upsilon_H e^{-\lambda_{x,0}M} + \frac{e^{-\lambda_{x,0}\Upsilon_L} - e^{-\lambda_{x,0}\Upsilon_H}}{\lambda_{x,0}}\right] - ac_a - kx^2. \tag{A.82}$$

When we take the derivative of (A.82) with respect to Υ_H, we only need to consider certain terms. As such, we restate the derivative as follows:

$$\begin{aligned}\frac{\partial U^P(x^*)}{\partial \Upsilon_H} &= \frac{\partial}{\partial \Upsilon_H}\frac{pc_2\Upsilon_H e^{-\lambda_{x,1}M}}{1-e^{-\lambda_{x,1}M}} + \frac{\partial}{\partial \Upsilon_H}\frac{pc_2 e^{-\lambda_{x,1}\Upsilon_H}}{\lambda_{x,1}(1-e^{-\lambda_{x,1}M})} \\ &+ \frac{\partial}{\partial \Upsilon_H}\frac{(1-p)c_2\Upsilon_H e^{-\lambda_{x,0}M}}{1-e^{-\lambda_{x,0}M}} + \frac{\partial}{\partial \Upsilon_H}\frac{(1-p)c_2 e^{-\lambda_{x,0}\Upsilon_H}}{\lambda_{x,0}(1-e^{-\lambda_{x,0}M})}\end{aligned} \tag{A.83}$$

We now compute the derivative:

$$\begin{aligned}&= \frac{pc_2 e^{-\lambda_{x,1}M}(1-e^{-\lambda_{x,1}M})}{(1-e^{-\lambda_{x,1}M})^2} + \frac{-pc_2\lambda_{x,1}e^{-\lambda_{x,1}\Upsilon_H}(\lambda_{x,1}(1-e^{-\lambda_{x,1}M})}{(\lambda_{x,1}(1-e^{-\lambda_{x,1}M}))^2} \\ &+ \frac{(1-p)c_2 e^{-\lambda_{x,0}M}(1-e^{-\lambda_{x,0}M})}{(1-e^{-\lambda_{x,0}M})^2} + \frac{-(1-p)c_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_H}(\lambda_{x,0}(1-e^{-\lambda_{x,0}M}))}{(\lambda_{x,0}(1-e^{-\lambda_{x,0}M}))^2}.\end{aligned} \tag{A.84}$$

Simplifying yields:

$$= \frac{pc_2(e^{-\lambda_{x,1}M} - e^{-\lambda_{x,1}\Upsilon_H})}{1-e^{-\lambda_{x,1}M}} + \frac{(1-p)c_2(e^{-\lambda_{x,0}M} - e^{-\lambda_{x,0}\Upsilon_H})}{1-e^{-\lambda_{x,0}M}}. \tag{A.85}$$

We note that $\lambda_{x,1}M > \lambda_{x,1}\Upsilon_H$ and $\lambda_{x,0}M > \lambda_{x,0}\Upsilon_H$. As such, the numerators of both fractions are nonpositive, so the derivative is nonpositive. This makes sense, as increasing the upperbound of service time should, intuitively, worsen the principal's expected utility.

Now we pursue the second derivative. We distribute the coefficients across the numerator of each fraction in (A.85). We then isolate the terms that contain Υ_H.

$$\begin{aligned}\frac{\partial U^P(x^*)}{\partial \Upsilon_H} &= \frac{pc_2 e^{-\lambda_{x,1}M}}{1-e^{-\lambda_{x,1}M}} - \frac{pc_2 e^{-\lambda_{x,1}\Upsilon_H}}{1-e^{-\lambda_{x,1}M}} \\ &+ \frac{(1-p)c_2 e^{-\lambda_{x,0}M}}{1-e^{-\lambda_{x,0}M}} - \frac{(1-p)c_2 e^{-\lambda_{x,0}\Upsilon_H}}{1-e^{-\lambda_{x,0}M}}.\end{aligned} \tag{A.86}$$

We now take the first partial derivative of (A.86) with respect to Υ_H:

$$\frac{\partial^2 U^P(x^*)}{\partial \Upsilon_H^2} = \frac{(pc_2\lambda_{x,1}e^{-\lambda_{x,1}\Upsilon_H})(1-e^{-\lambda_{x,1}M})}{(1-e^{-\lambda_{x,1}M})^2} + \frac{((1-p)c_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_H})(1-e^{-\lambda_{x,0}M})}{(1-e^{-\lambda_{x,0}M})^2}. \tag{A.87}$$

Simplifying (A.87) yields:

$$\frac{\partial^2 U^P(x^*)}{\partial \Upsilon_H^2} = \frac{pc_2\lambda_{x,1}e^{-\lambda_{x,1}\Upsilon_H}}{1-e^{-\lambda_{x,1}M}} + \frac{(1-p)c_2\lambda_{x,0}e^{-\lambda_{x,0}\Upsilon_H}}{1-e^{-\lambda_{x,0}M}}. \tag{A.88}$$

By inspection, we see that both fractions in (A.88) are nonegative. This implies that increasing Υ_H yields a convex decreasing response in U^P.

Appendix B

Soccer Ranking Appendices

B.1 Notation

In this Appendix, we present a table of our notation.

Table B.1: Notation for Network Ranking Models

A	Cumulative adjacency matrix for static model
A_t	Adjacency matrix for gameweek t for dynamic model
k_A	Adjustment coefficient for home/away status
k_D	Adjustment coefficient for goal difference status
$\tilde{D}^t$	Normalized Direct Similarity Approach matrix
$\bar{\tilde{D}}^t$	Normalized Mean-Based Direct Similarity Approach matrix for static model
$\bar{D}^t_\alpha$	α-scaled Mean-Based Direct Similarity Approach matrix for dynamic model
$\tilde{M}^t$	Normalized Matched Set Similarity Approach matrix
$\bar{\tilde{M}}^t$	Normalized Mean-Based Matched Set Similarity Approach matrix for static model
$\bar{M}^t_\alpha$	α-scaled Mean-Based Matched Set Similarity Approach matrix for dynamic model
$\tilde{U}^t$	Normalized Unmatched Set Similarity Approach matrix
$\bar{\tilde{U}}^t$	Normalized Mean-Based Unmatched Set Similarity Approach matrix for static model
$\bar{U}^t_\alpha$	α-scaled Mean-Based Unmatched Set Similarity Approach matrix for dynamic model

B.2 Combination Method Explanations

As mentioned in Section 3.5.10, we consider the following combination methods in both static and dynamic form:

1. Home/Away + Goal Difference
2. Home/Away + Direct Similarity
3. Home/Away + Matched Set Similarity
4. Home/Away + Unmatched Set Similarity
5. Home/Away + Mean-Based Direct Similarity
6. Home/Away + Mean-Based Matched Set Similarity
7. Home/Away + Mean-Based Unmatched Set Similarity
8. Goal Difference + Direct Similarity
9. Goal Difference + Matched Set Similarity
10. Goal Difference + Unmatched Set Similarity
11. Goal Difference + Mean-Based Direct Similarity
12. Goal Difference + Mean-Based Matched Set Similarity
13. Goal Difference + Mean-Based Unmatched Set Similarity
14. Home/Away + Goal Difference + Direct Similarity
15. Home/Away + Goal Difference + Matched Set Similarity
16. Home/Away + Goal Difference + Unmatched Set Similarity
17. Home/Away + Goal Difference + Mean-Based Direct Similarity
18. Home/Away + Goal Difference + Mean-Based Matched Set Similarity
19. Home/Away + Goal Difference + Mean-Based Unmatched Set Similarity.

We now explain each in turn.

B.2.1 Home/Away + Goal Difference

This approach is simply a combination of the approaches in Sections 3.5.1 and 3.5.2. The implementation is as follows:

- For away wins by team i over team j, we add $k_A(k_D^{d-1})$ to A_{ji}.
- For away ties by team i over team j, we add $0.5k_A$ to A_{ji}.
- For home wins by by team i over team j, we add (k_D^{d-1}) to A_{ji}.
- For home ties by team i over team j, we add 0.5 to A_{ji}.

For our testing, we set $k_A = k_D = 1.3$.

B.2.2 Static Home/Away + Direct Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.4.1. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our static Direct Similarity Approach.

B.2.3 Dynamic Home/Away + Direct Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.4.2. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our dynamic Direct Similarity Approach.

B.2.4 Static Home/Away + Matched Set Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.6.1. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our static Matched Set Similarity Approach.

B.2.5 Dynamic Home/Away + Matched Set Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.6.2. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our dynamic Matched Set Similarity Approach.

B.2.6 Static Home/Away + Unmatched Set Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.8.1. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our static Unmatched Set Similarity Approach.

B.2.7 Dynamic Home/Away + Unmatched Set Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.8.2. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our dynamic Unmatched Set Similarity Approach.

B.2.8 Static Home/Away + Mean-Based Direct Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.5.1. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our static Mean-Based Direct Similarity Approach.

B.2.9 Dynamic Home/Away + Mean-Based Direct Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.5.2. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our dynamic Mean-Based Direct Similarity Approach.

B.2.10 Static Home/Away + Mean-Based Matched Set Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.7.1. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our static Mean-Based Matched Set Similarity Approach.

B.2.11 Dynamic Home/Away + Mean-Based Matched Set Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.7.2. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our dynamic Mean-Based Matched Set Similarity Approach.

B.2.12 Static Home/Away + Mean-Based Unmatched Set Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.9.1. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our static Mean-Based Unmatched Set Similarity Approach.

B.2.13 Dynamic Home/Away + Mean-Based Unmatched Set Similarity

This approach is a combination of the approaches in Sections 3.5.1 and 3.5.9.2. Specifically, we prepare the adjacency matrix with the home/away modifications and then use this matrix in our dynamic Mean-Based Unmatched Set Similarity Approach.

B.2.14 Static Goal Difference + Direct Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.4.1. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our static Direct Similarity Approach.

B.2.15 Dynamic Goal Difference + Direct Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.4.2. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our dynamic Direct Similarity Approach.

B.2.16 Static Goal Difference + Matched Set Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.6.1. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our static Matched Set Similarity Approach.

B.2.17 Dynamic Goal Difference + Matched Set Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.6.2. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our dynamic Matched Set Similarity Approach.

B.2.18 Static Goal Difference + Unmatched Set Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.8.1. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our static Unmatched Set Similarity Approach.

B.2.19 Dynamic Goal Difference + Unmatched Set Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.8.2. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our dynamic Unmatched Set Similarity Approach.

B.2.20 Static Goal Difference + Mean-Based Direct Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.5.1. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our static Mean-Based Direct Similarity Approach.

B.2.21 Dynamic Goal Difference + Mean-Based Direct Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.5.2. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our dynamic Mean-Based Direct Similarity Approach.

B.2.22 Static Goal Difference + Mean-Based Matched Set Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.7.1. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our static Mean-Based Matched Set Similarity Approach.

B.2.23 Dynamic Goal Difference + Mean-Based Matched Set Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.7.2. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our dynamic Mean-Based Matched Set Similarity Approach.

B.2.24 Static Goal Difference + Mean-Based Unmatched Set Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.9.1. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our static Mean-Based Unmatched Set Similarity Approach.

B.2.25 Dynamic Goal Difference + Mean-Based Unmatched Set Similarity

This approach is a combination of the approaches in Sections 3.5.2 and 3.5.9.2. Specifically, we prepare the adjacency matrix with the goal difference modifications and then use this matrix in our dynamic Mean-Based Unmatched Set Similarity Approach.

B.2.26 Static Home/Away + Goal Difference + Direct Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.4.1. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our static Direct Similarity Approach.

B.2.27 Dynamic Home/Away + Goal Difference + Direct Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.4.2. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our dynamic Direct Similarity Approach.

B.2.28 Static Home/Away + Goal Difference + Matched Set Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.6.1. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our static Matched Set Similarity Approach.

B.2.29 Dynamic Home/Away + Goal Difference + Matched Set Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.6.2. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our dynamic Matched Set Similarity Approach.

B.2.30 Static Home/Away + Goal Difference + Unmatched Set Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.8.1. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our static Unmatched Set Similarity Approach.

B.2.31 Dynamic Home/Away + Goal Difference + Unmatched Set Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.8.2. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our dynamic Unmatched Set Similarity Approach.

B.2.32 Static Home/Away + Goal Difference + Mean-Based Direct Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.5.1. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our static Mean-Based Direct Similarity Approach.

B.2.33 Dynamic Home/Away + Goal Difference + Mean-Based Direct Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.5.2. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our dynamic Mean-Based Direct Similarity Approach.

B.2.34 Static Home/Away + Goal Difference + Mean-Based Matched Set Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.7.1. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our static Mean-Based Matched Set Similarity Approach.

B.2.35 Dynamic Home/Away + Goal Difference + Mean-Based Matched Set Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.7.2. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our dynamic Mean-Based Matched Set Similarity Approach.

B.2.36 Static Home/Away + Goal Difference + Mean-Based Unmatched Set Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.9.1. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our static Mean-Based Unmatched Set Similarity Approach.

B.2.37 Dynamic Home/Away + Goal Difference + Mean-Based Unmatched Set Similarity

This approach is a combination of the approaches in Sections B.2.1 and 3.5.9.2. Specifically, we prepare the adjacency matrix with the home/away and goal difference modifications and then use this matrix in our dynamic Mean-Based Unmatched Set Similarity Approach.

B.3 Additional Performance Figures

This section contains additional figures related to our ranking model performance. We separate them by approach (static and dynamic) and league.

B.3.1 Static Approaches

This section contains the line plots for the static network approaches for 4 year rolling windows starting with the number indicated on the x-axis.

B.3.1.1 England

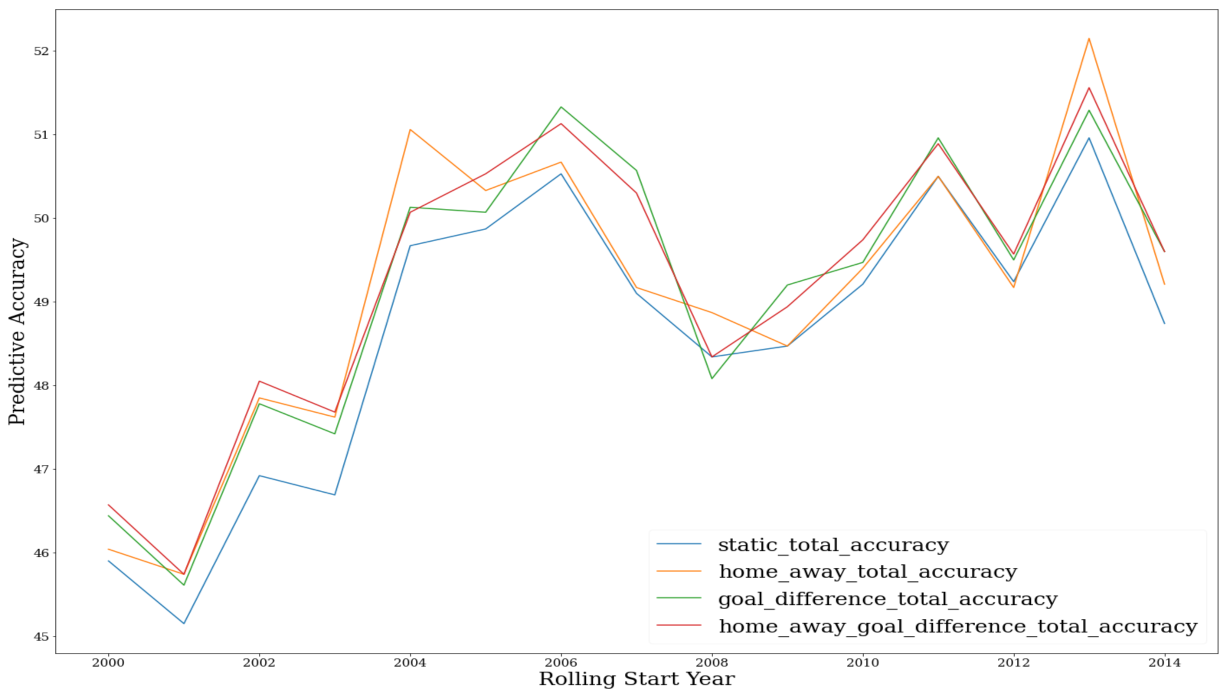

Figure B.1: English Premier League 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

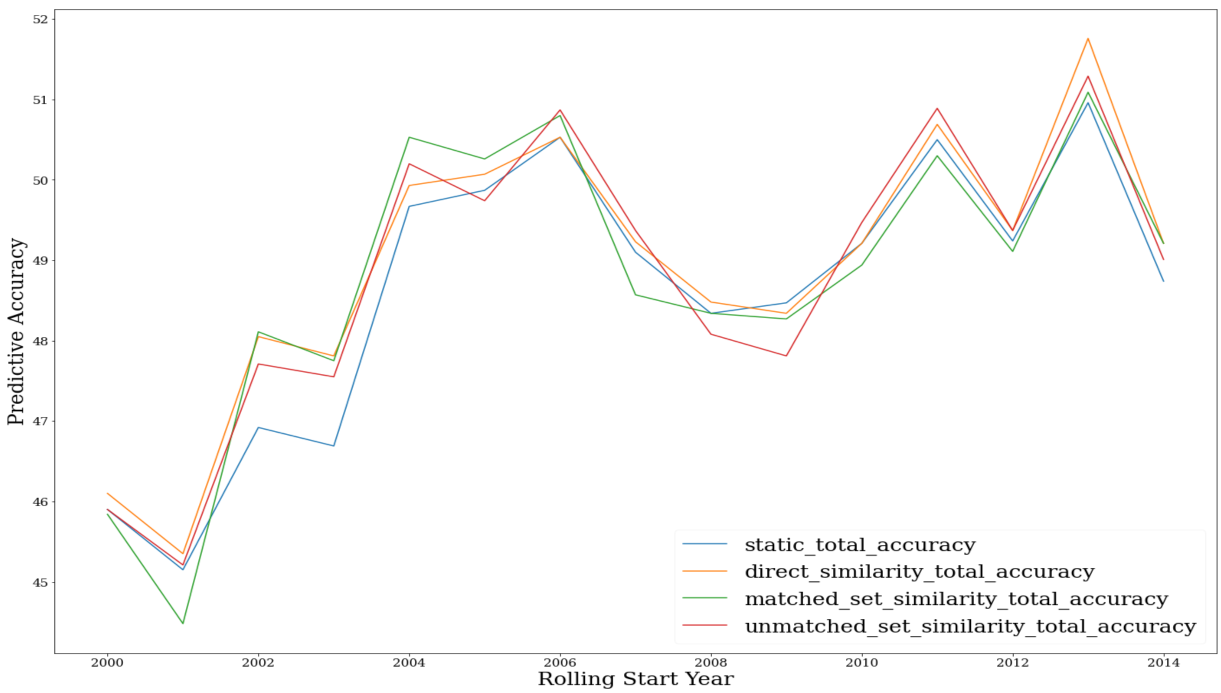

Figure B.2: English Premier League 4-Year Rolling Predictive Accuracy: Static Similarity Approaches

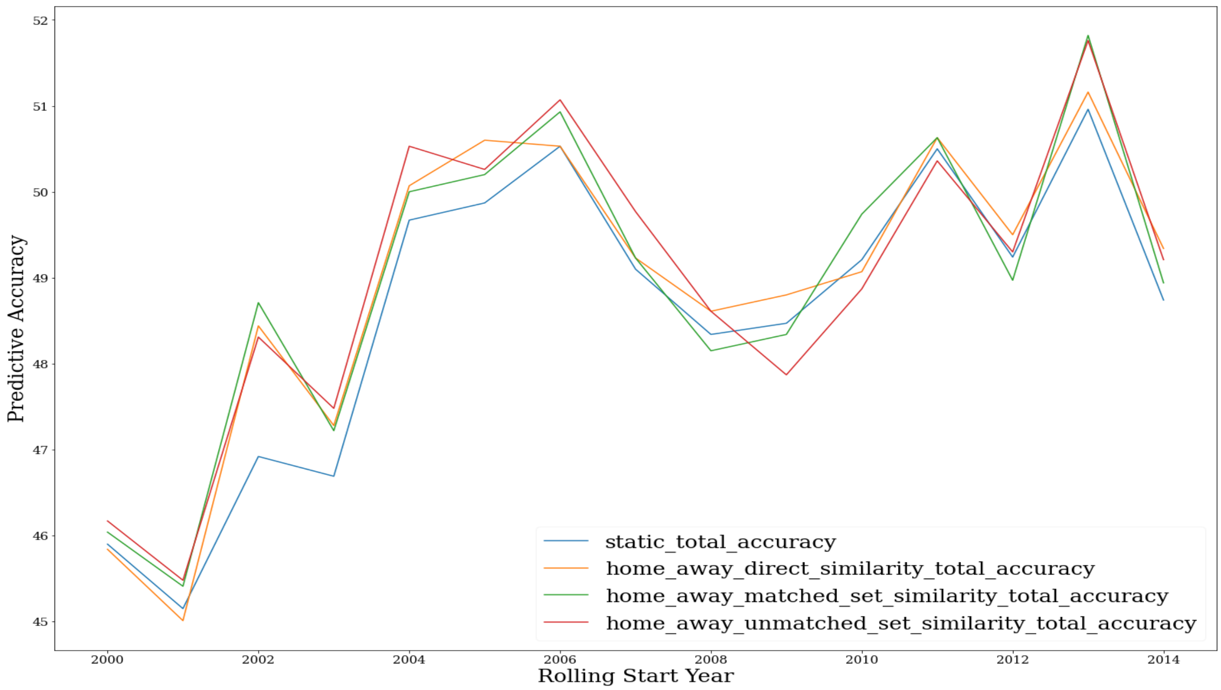

Figure B.3: English Premier League 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches

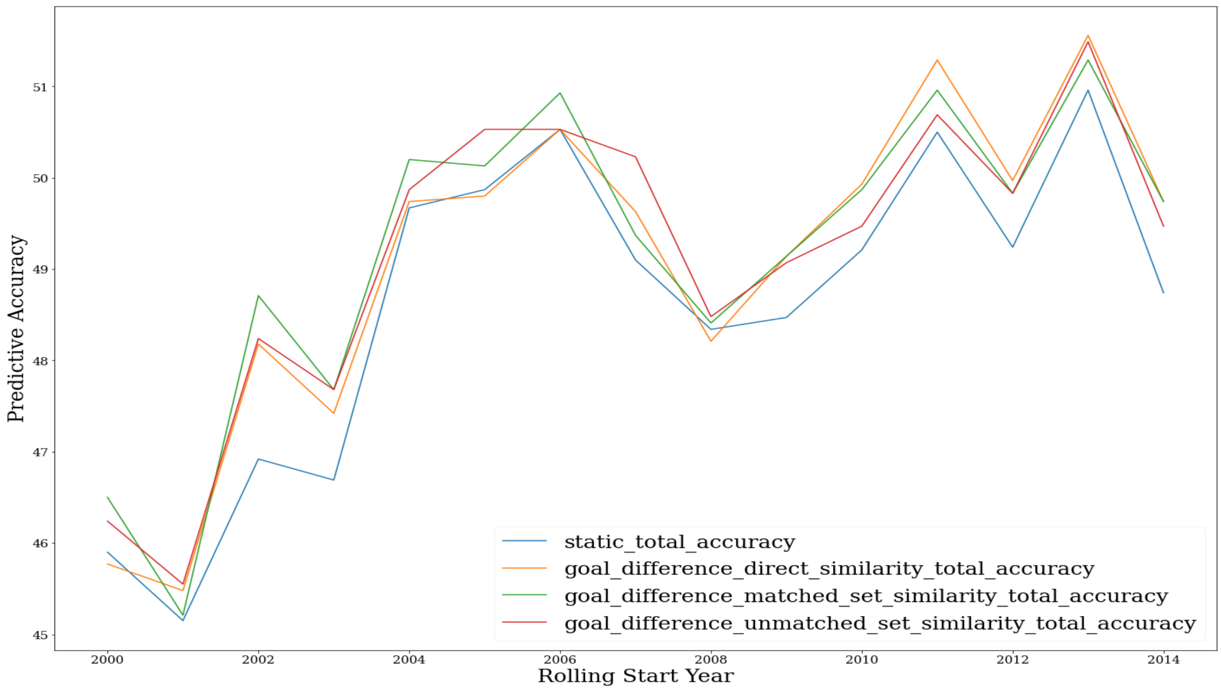

Figure B.4: English Premier League 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches

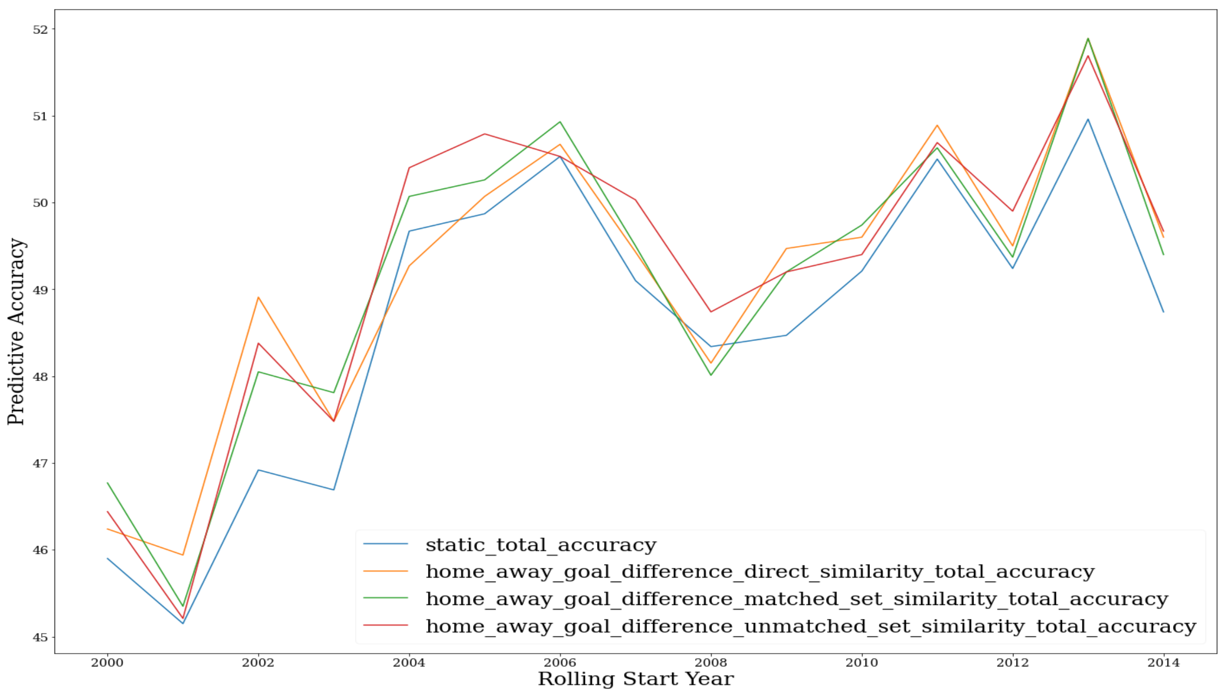

Figure B.5: English Premier League 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches

B.3.1.2 Spain

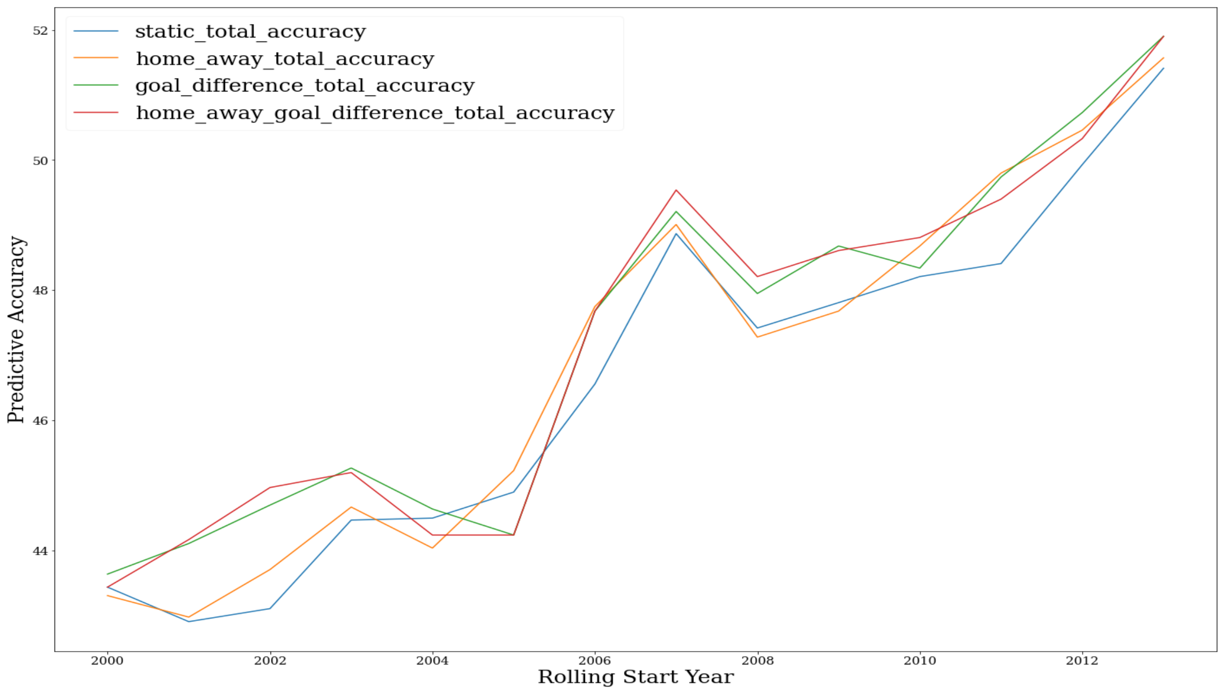

Figure B.6: Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

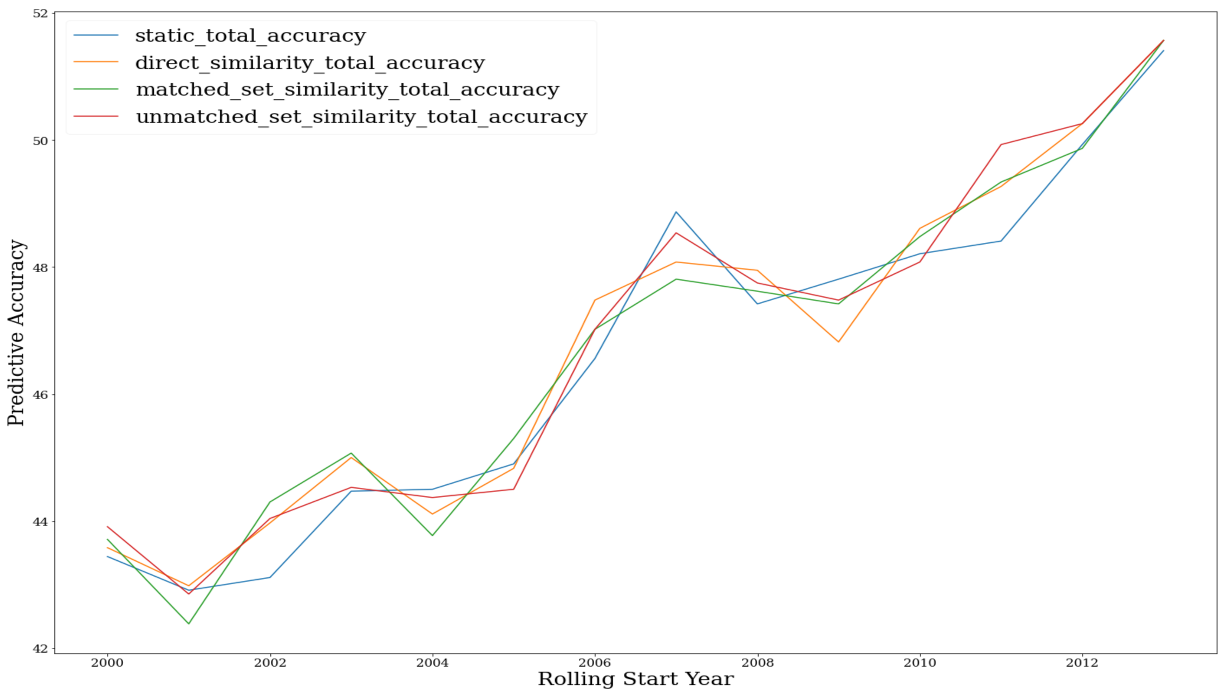

Figure B.7: Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Similarity Approaches

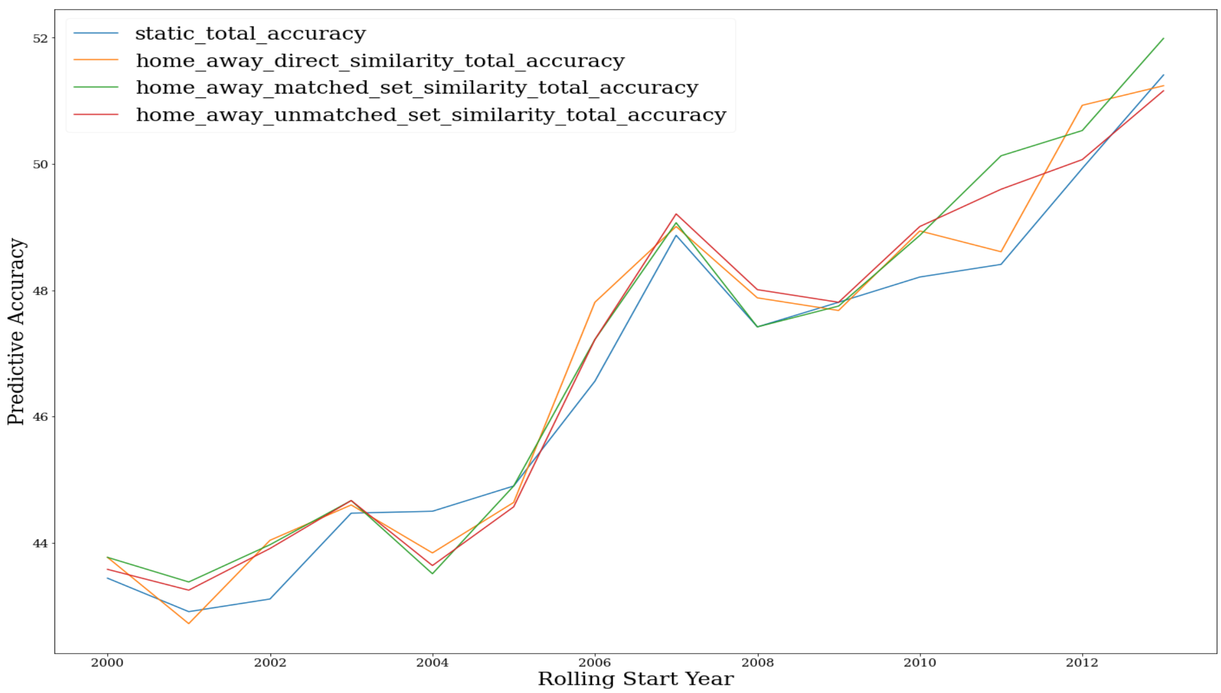

Figure B.8: Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches

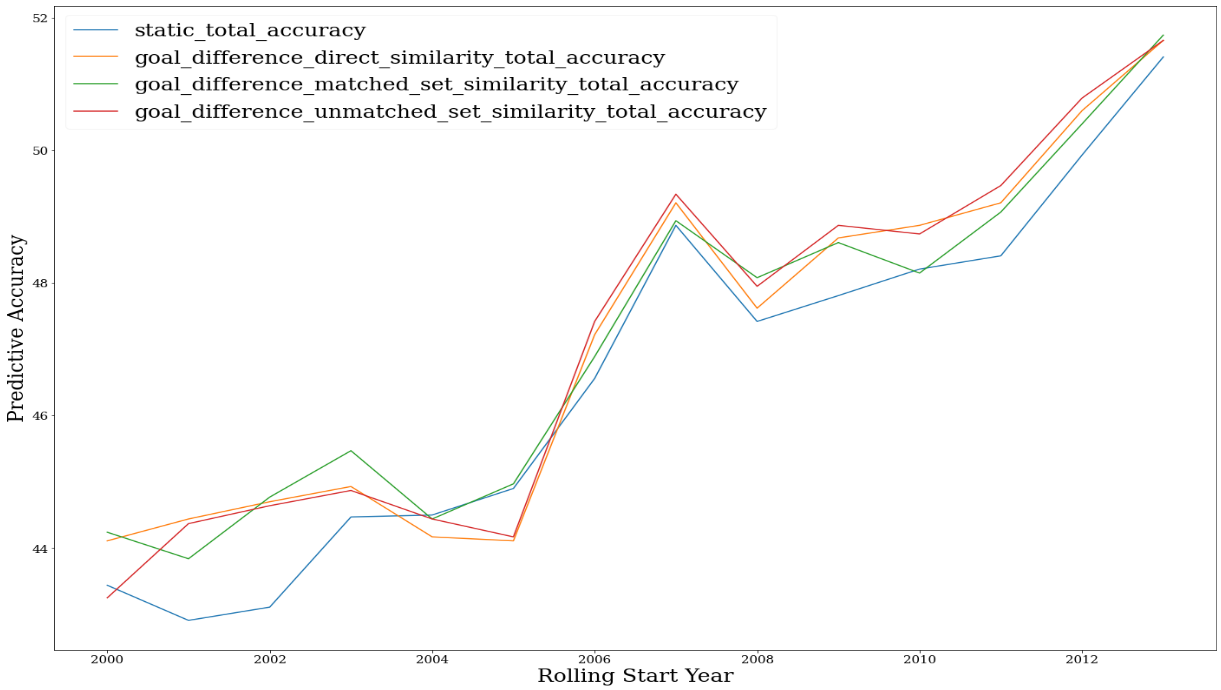

Figure B.9: Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches

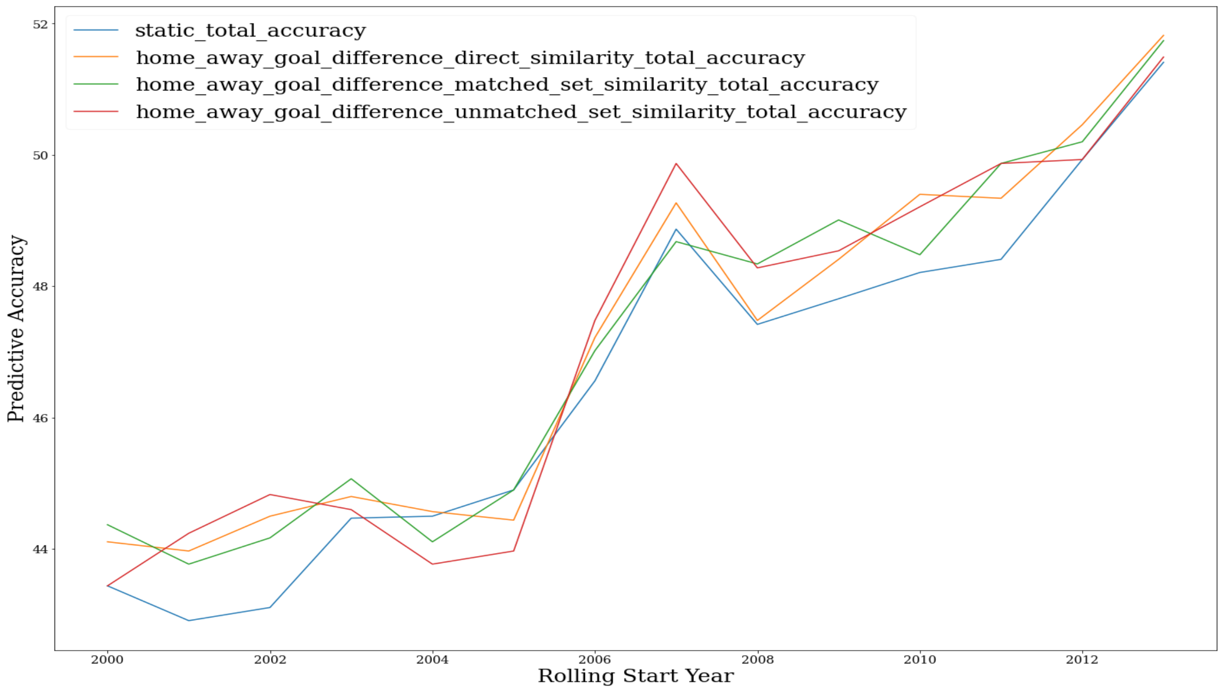

Figure B.10: Spanish La Liga 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches

B.3.1.3 Germany

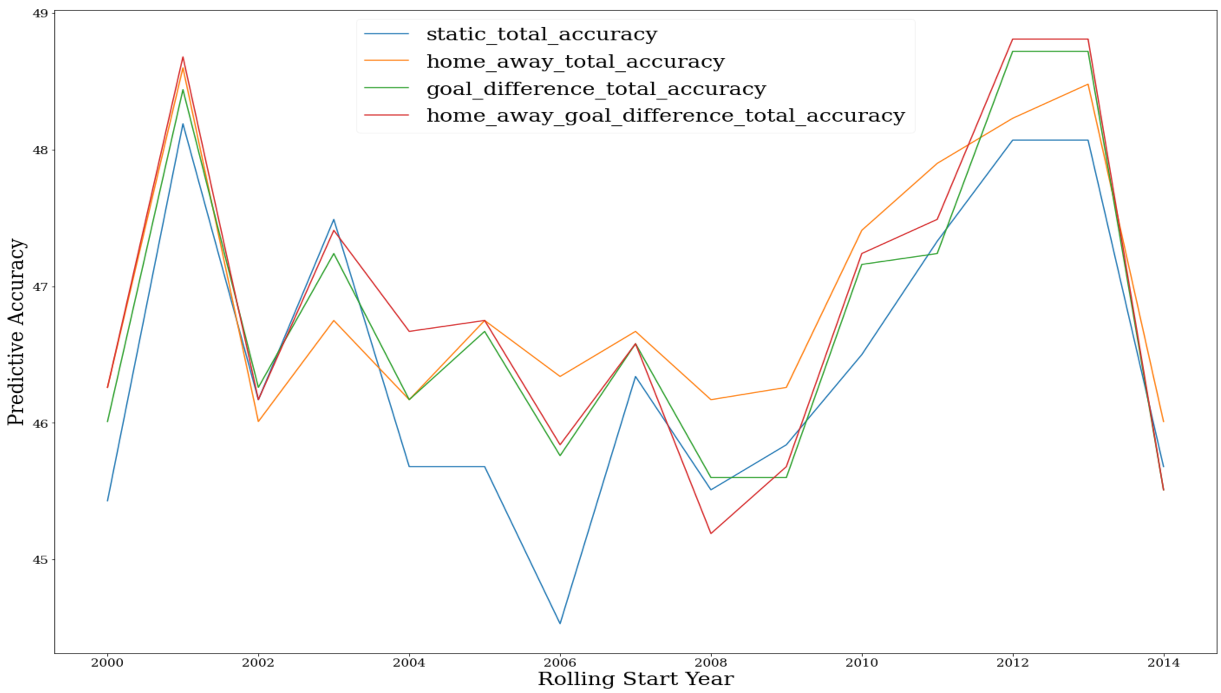

Figure B.11: German Bundesliga 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

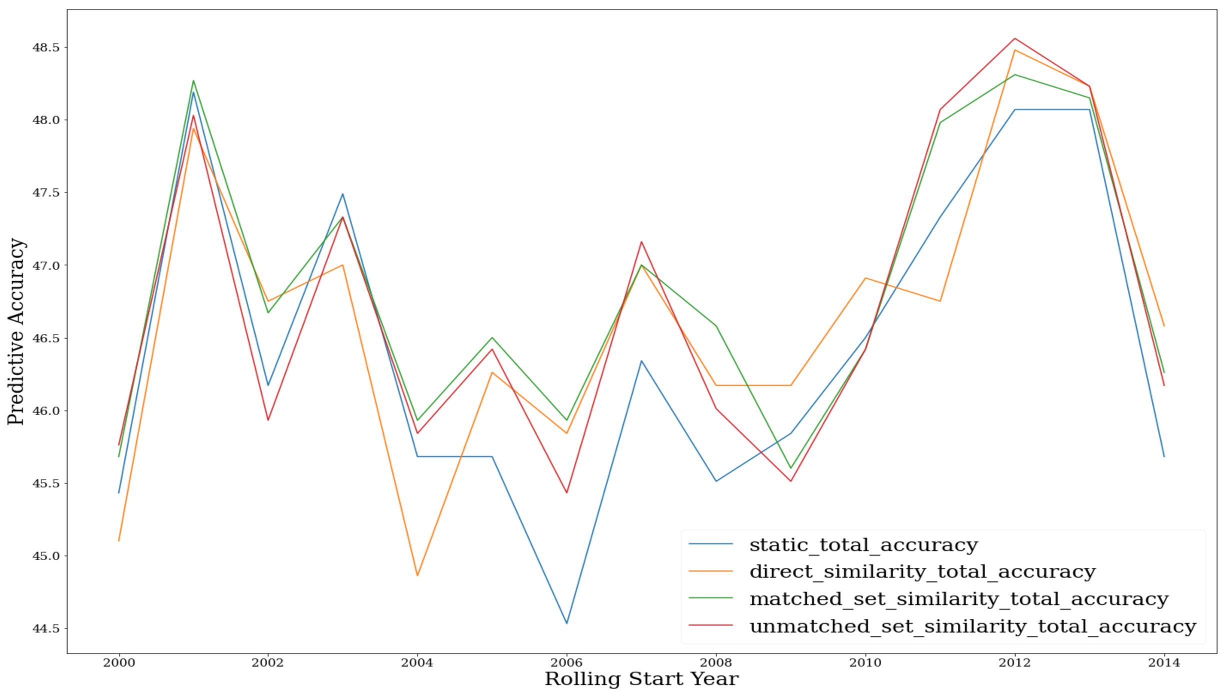

Figure B.12: German Bundesliga 4-Year Rolling Predictive Accuracy: Static Similarity Approaches

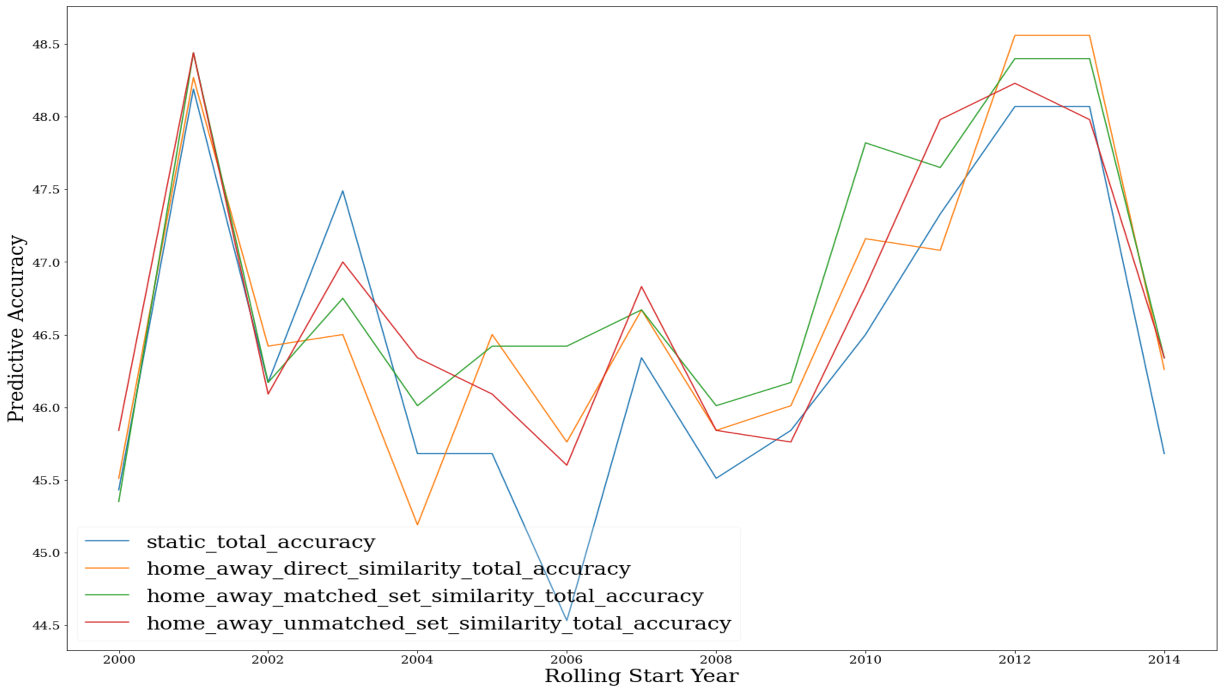

Figure B.13: German Bundesliga 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches

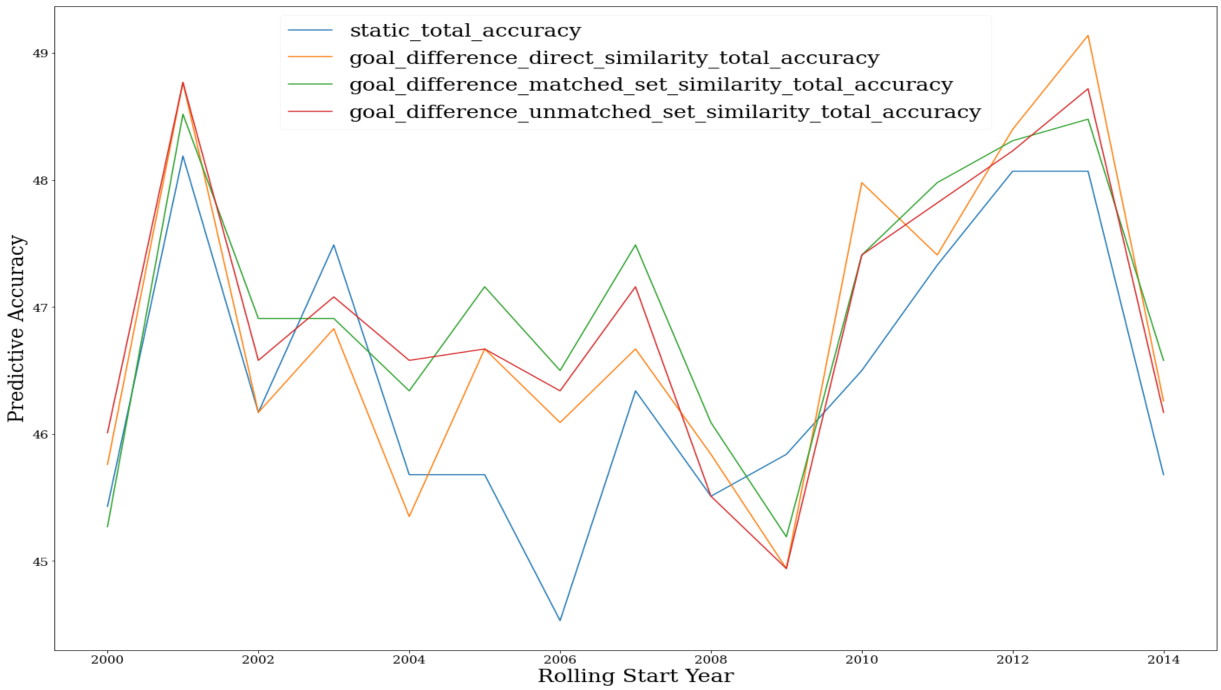

Figure B.14: German Bundesliga 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches

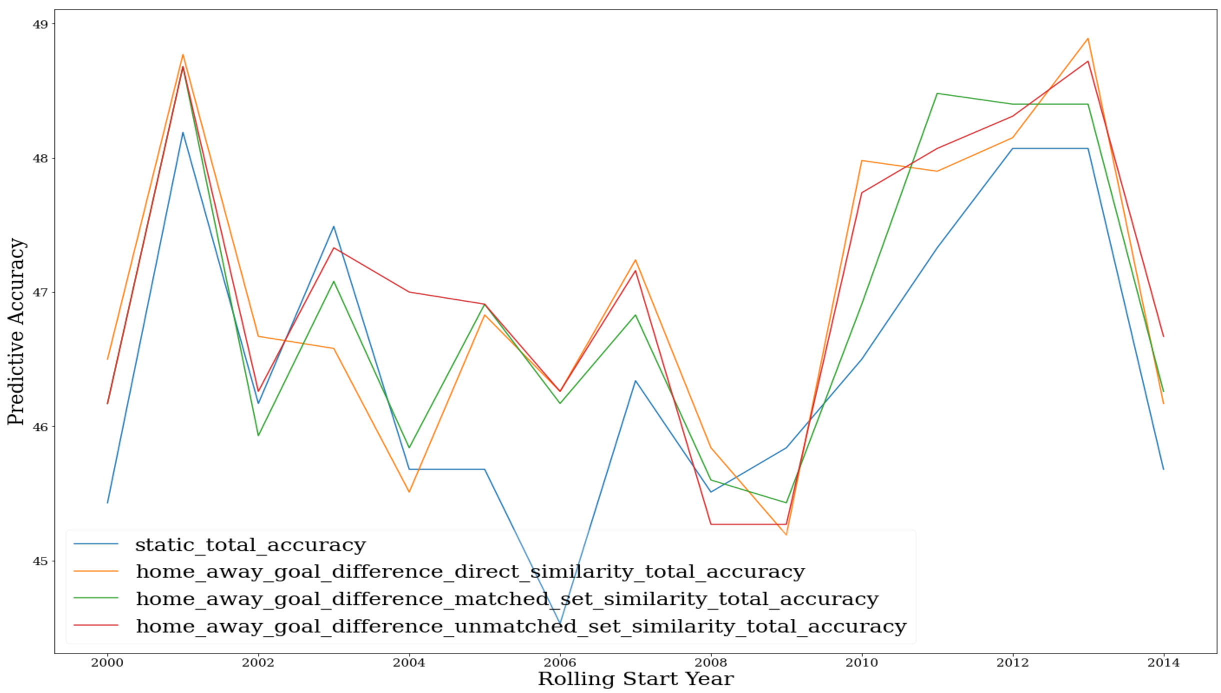

Figure B.15: German Bundesliga 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches

B.3.1.4 Italy

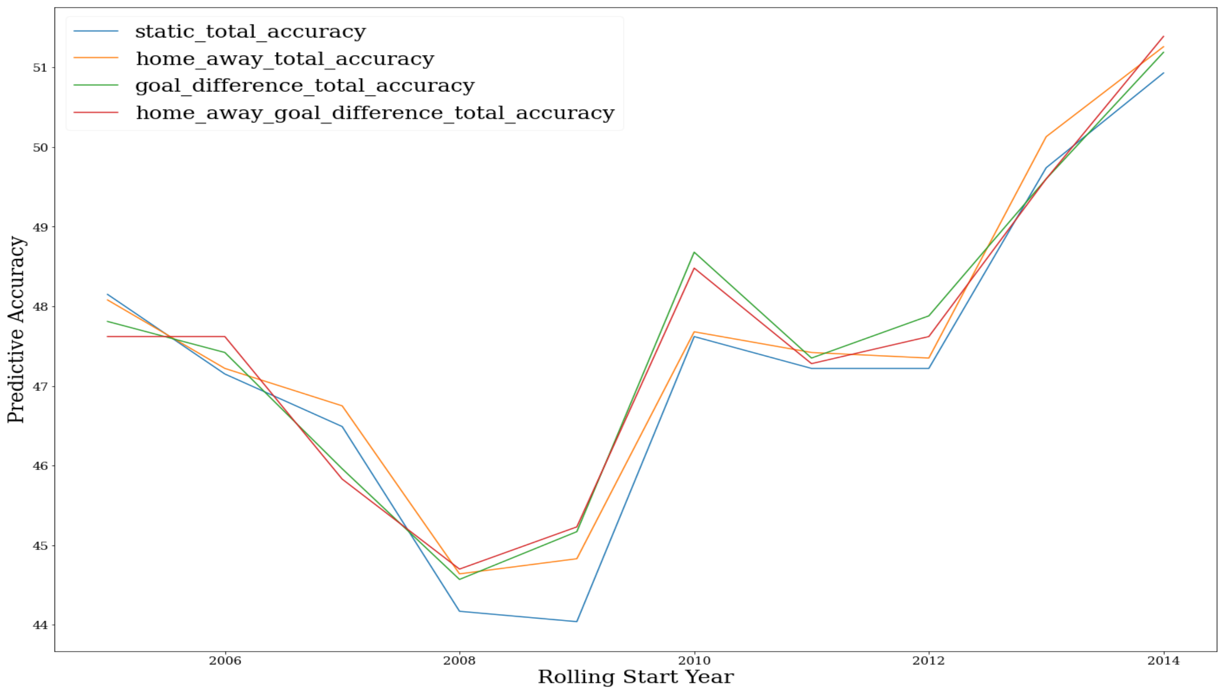

Figure B.16: Italian Serie A 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

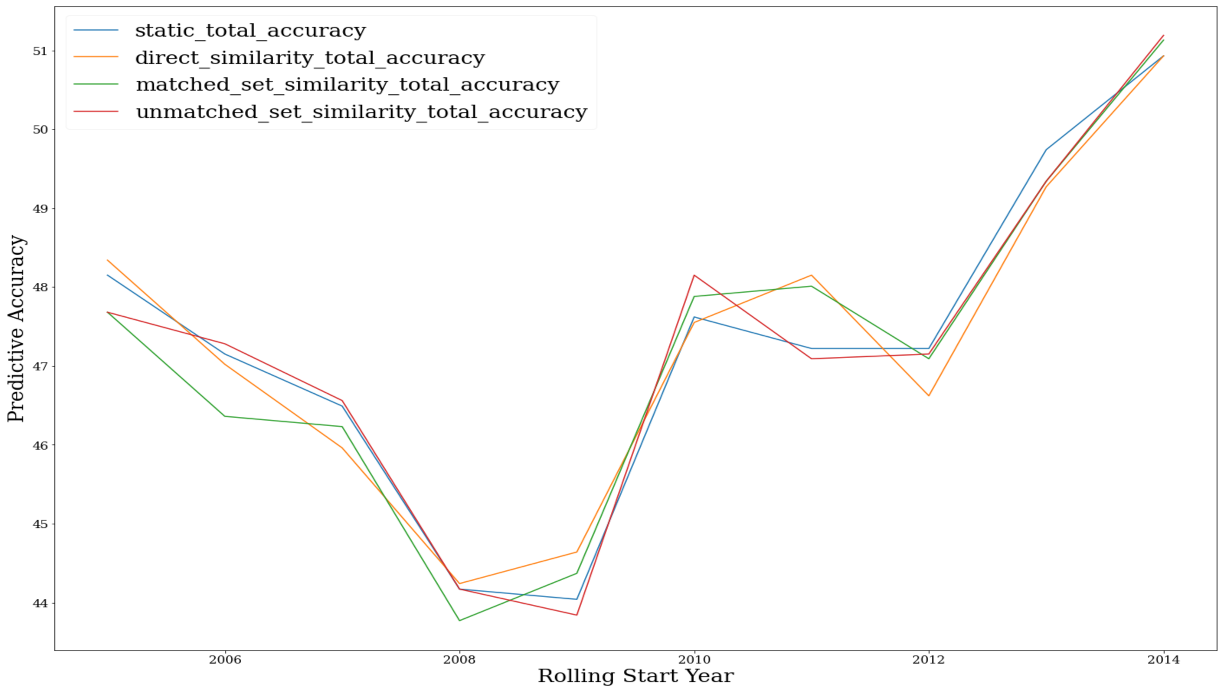

Figure B.17: Italian Serie A 4-Year Rolling Predictive Accuracy: Static Similarity Approaches

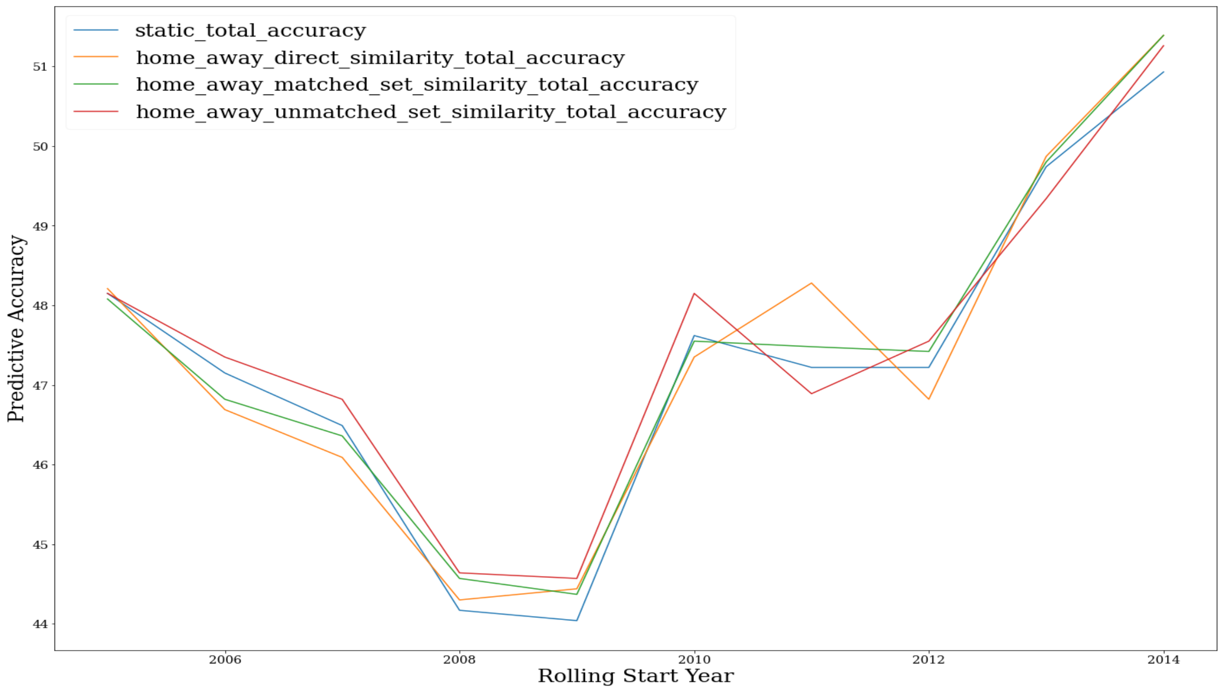

Figure B.18: Italian Serie A 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches

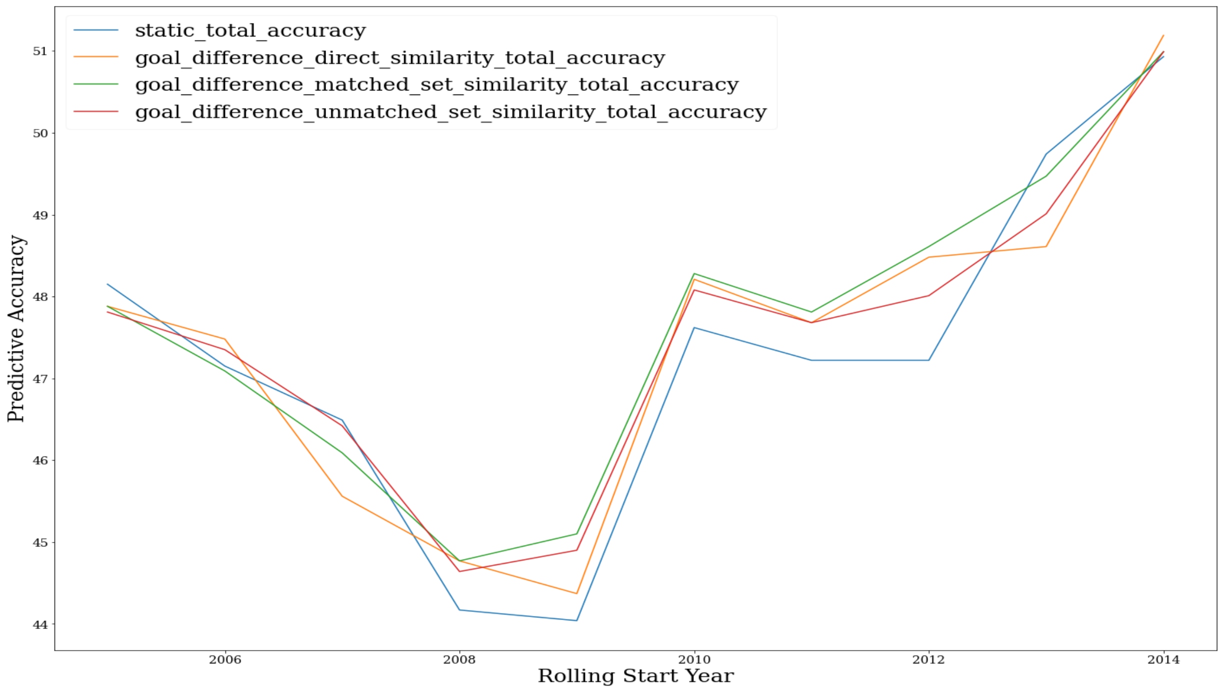

Figure B.19: Italian Serie A 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches

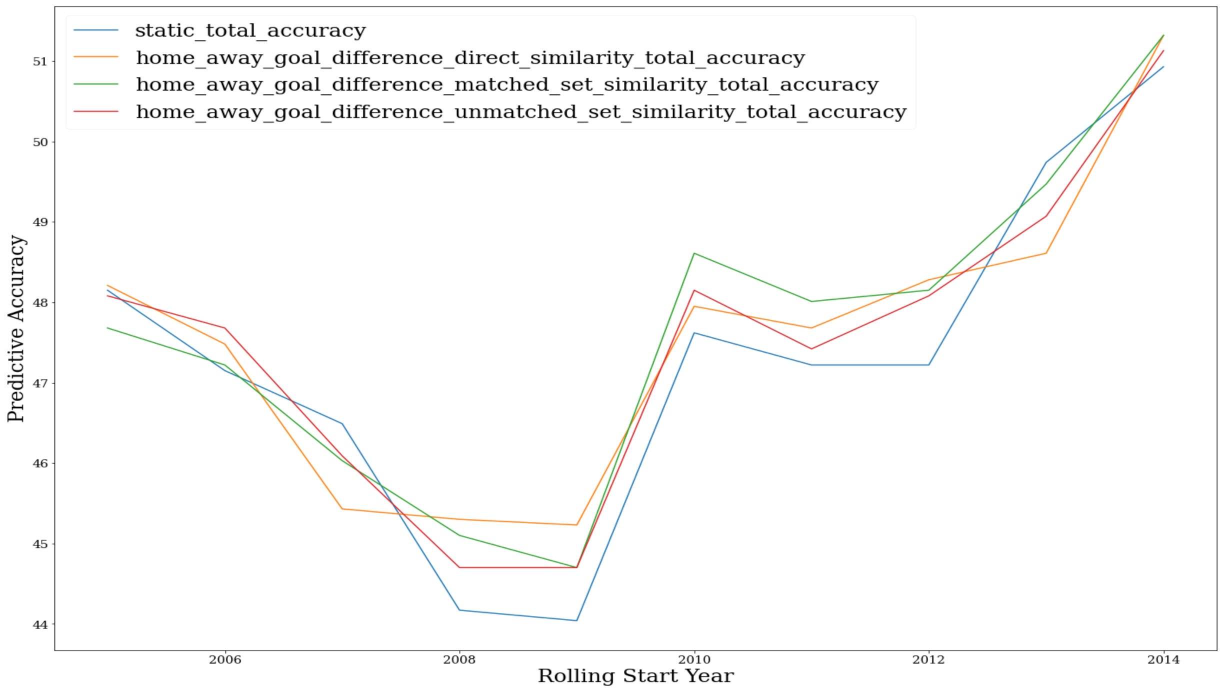

Figure B.20: Italian Serie A 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches

B.3.1.5 France

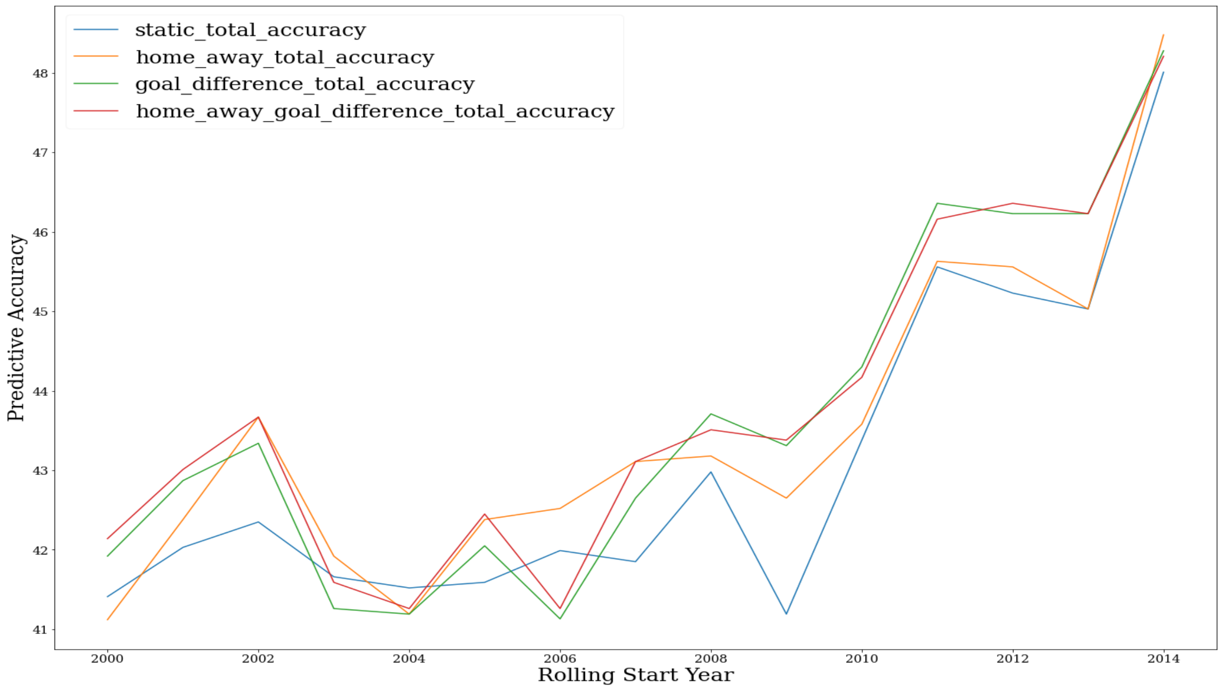

Figure B.21: French Ligue 1 4-Year Rolling Predictive Accuracy: Static Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

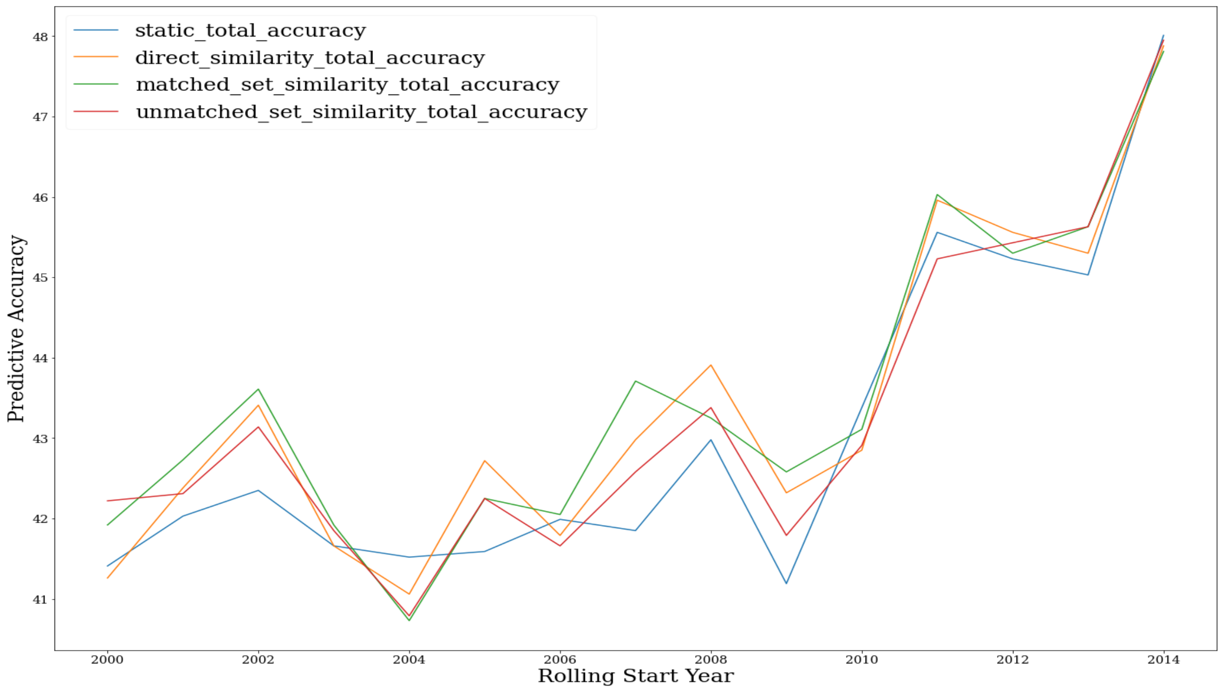

Figure B.22: French Ligue 1 4-Year Rolling Predictive Accuracy: Static Similarity Approaches

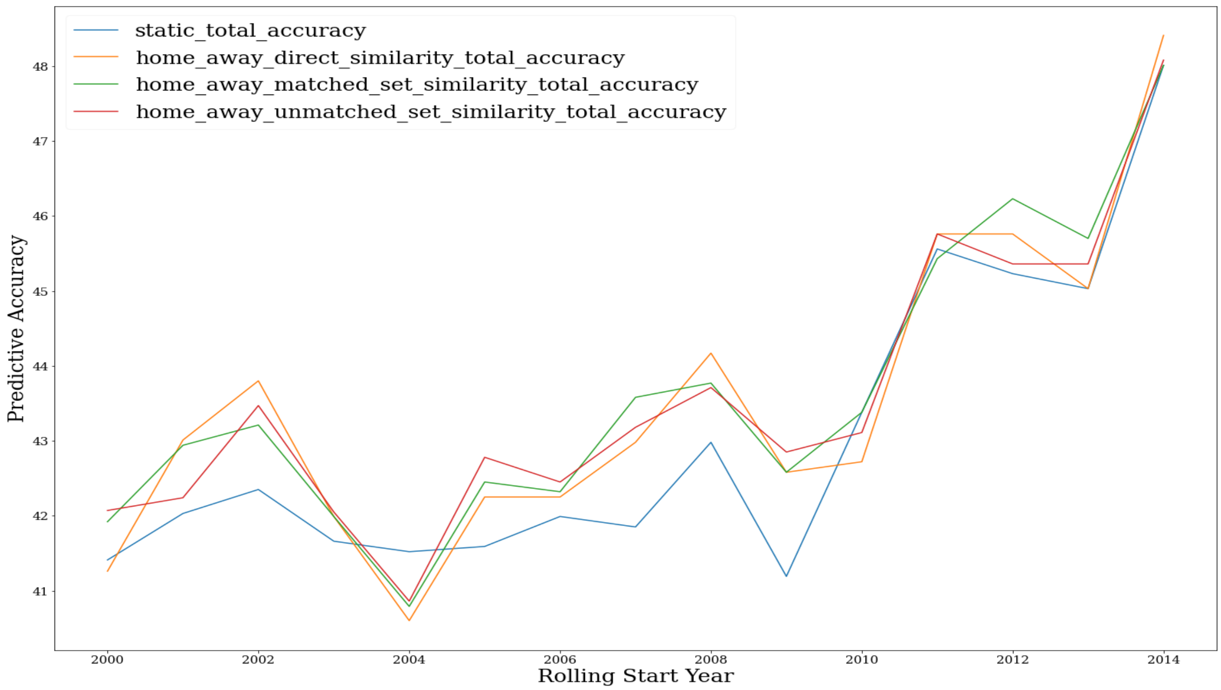

Figure B.23: French Ligue 1 4-Year Rolling Predictive Accuracy: Static Home/Away Similarity Approaches

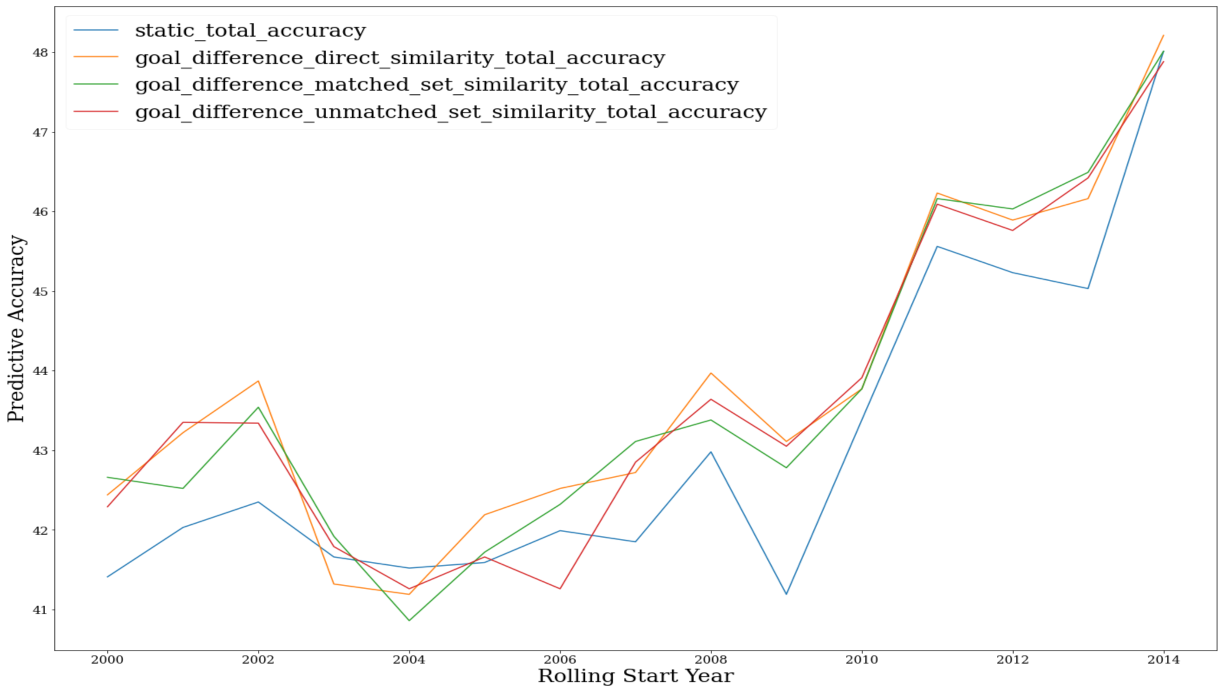

Figure B.24: French Ligue 1 4-Year Rolling Predictive Accuracy: Static Goal Difference Similarity Approaches

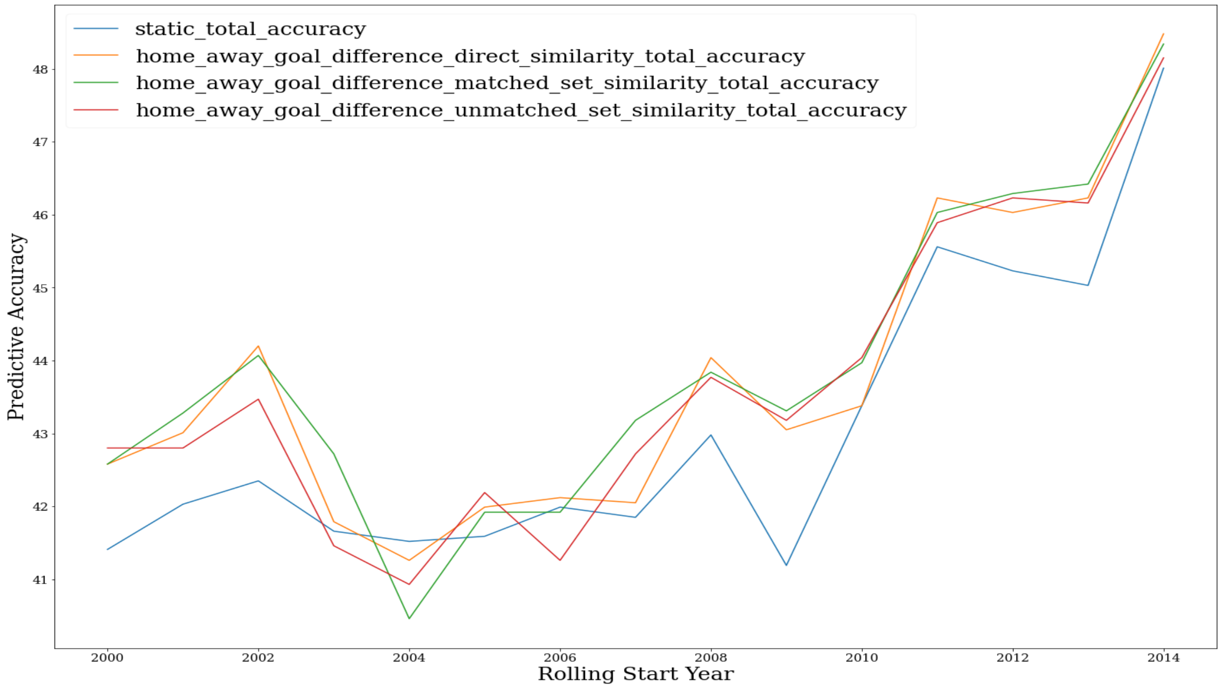

Figure B.25: French Ligue 1 4-Year Rolling Predictive Accuracy: Static Home/Away Goal Difference Similarity Approaches

B.3.2 Dynamic Approaches

This section contains the line plots for the dynamic network approaches for 4 year rolling windows starting with the number indicated on the x-axis.

B.3.2.1 England

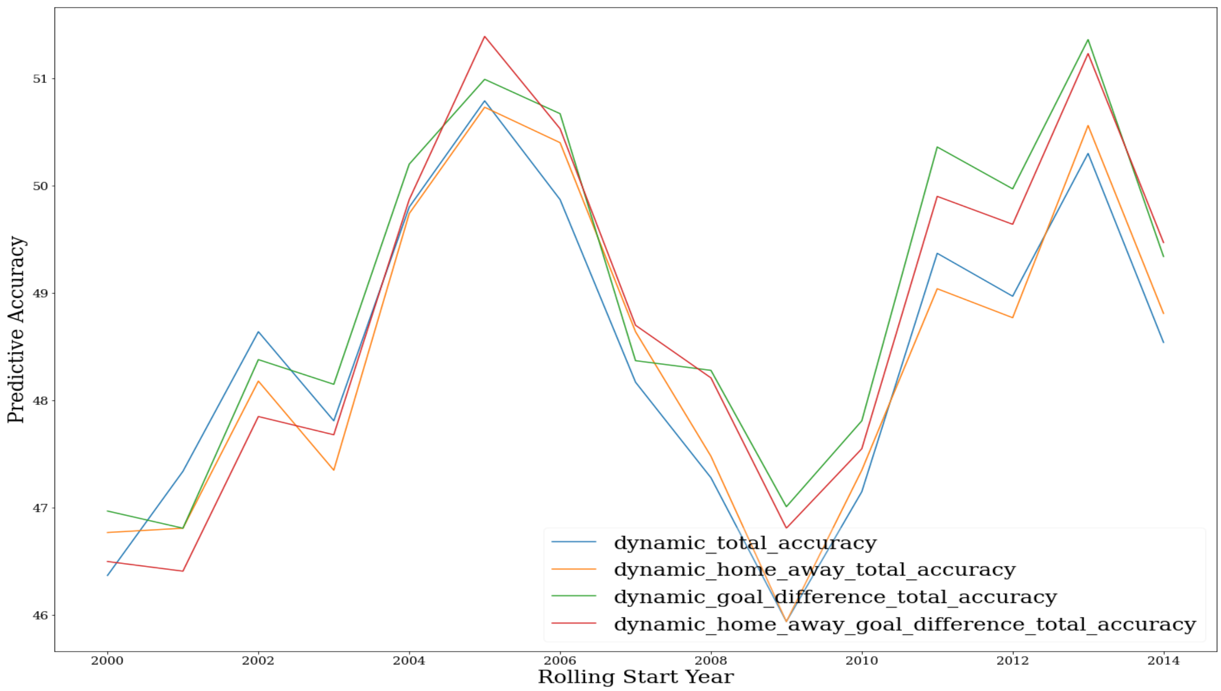

Figure B.26: English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

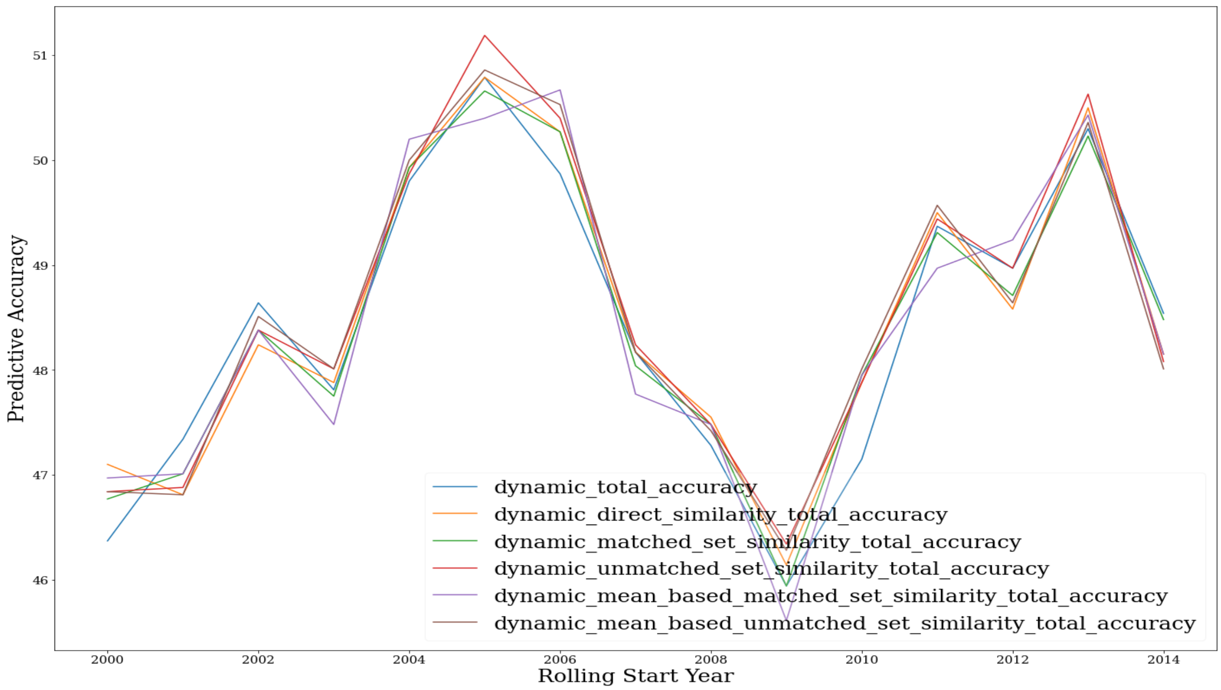

Figure B.27: English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches

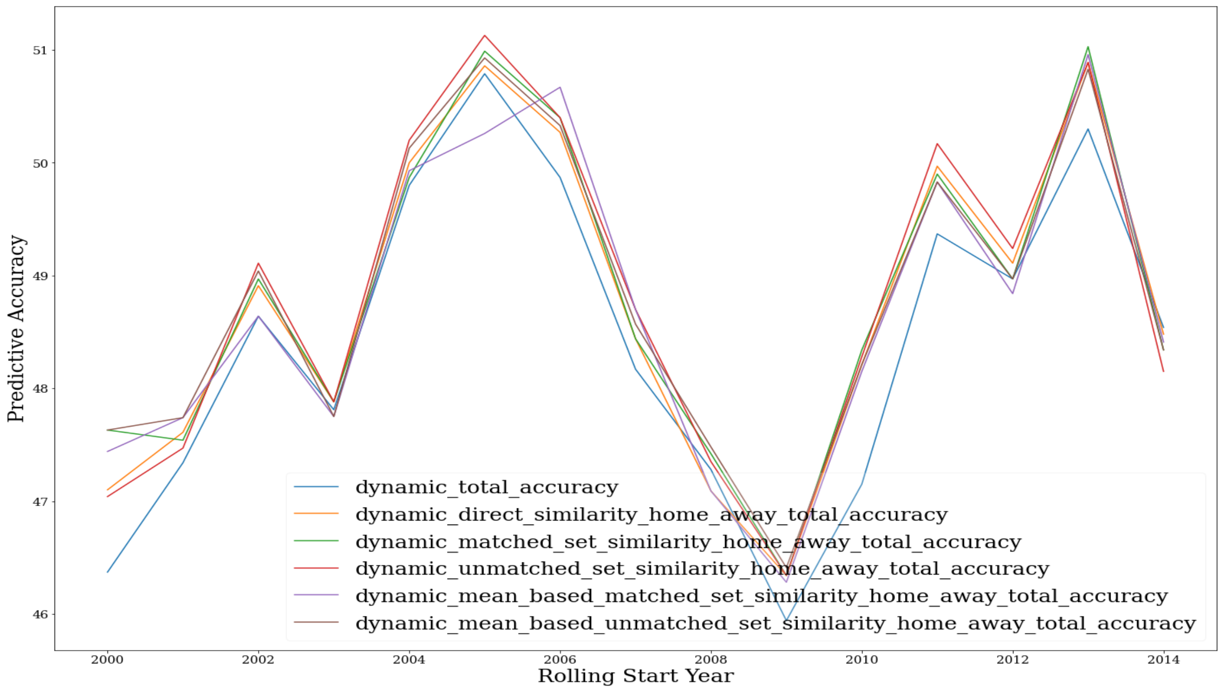

Figure B.28: English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches

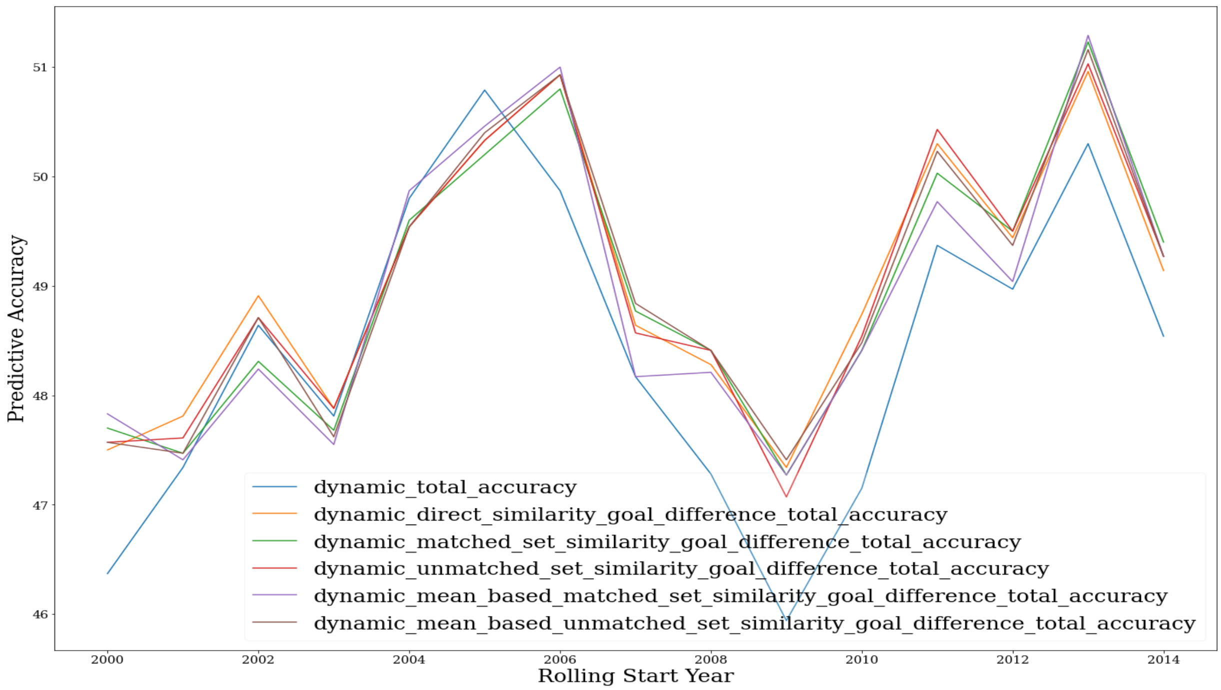

Figure B.29: English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches

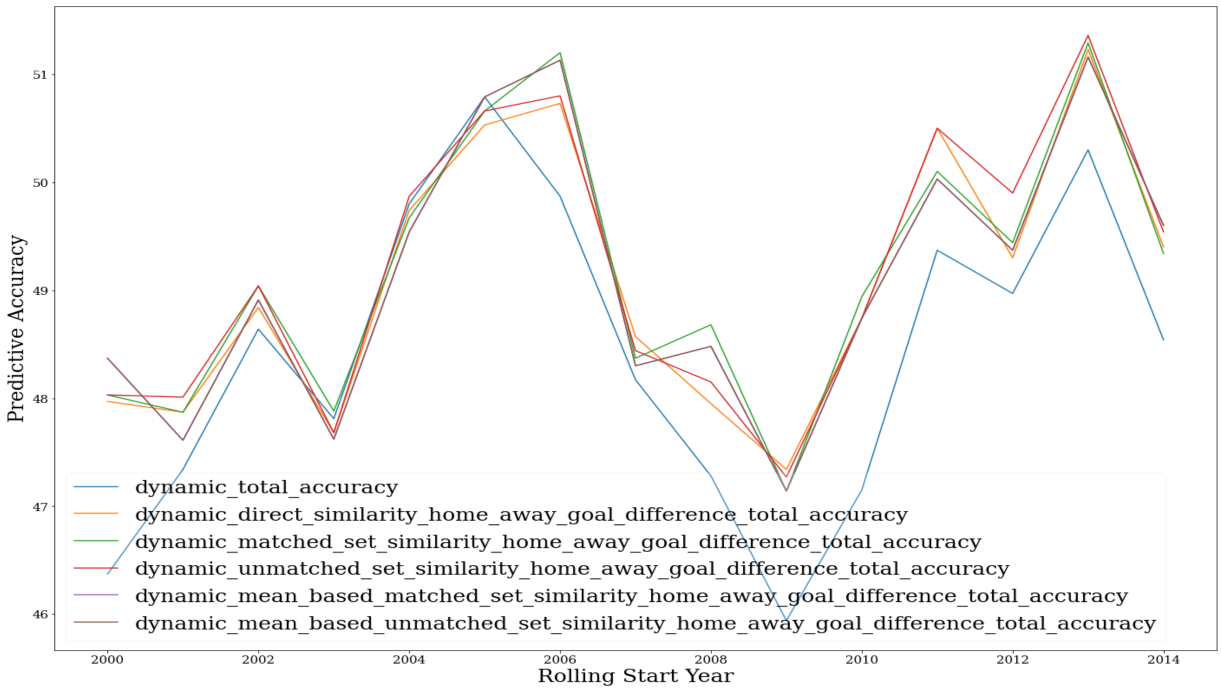

Figure B.30: English Premier League 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches

B.3.2.2 Spain

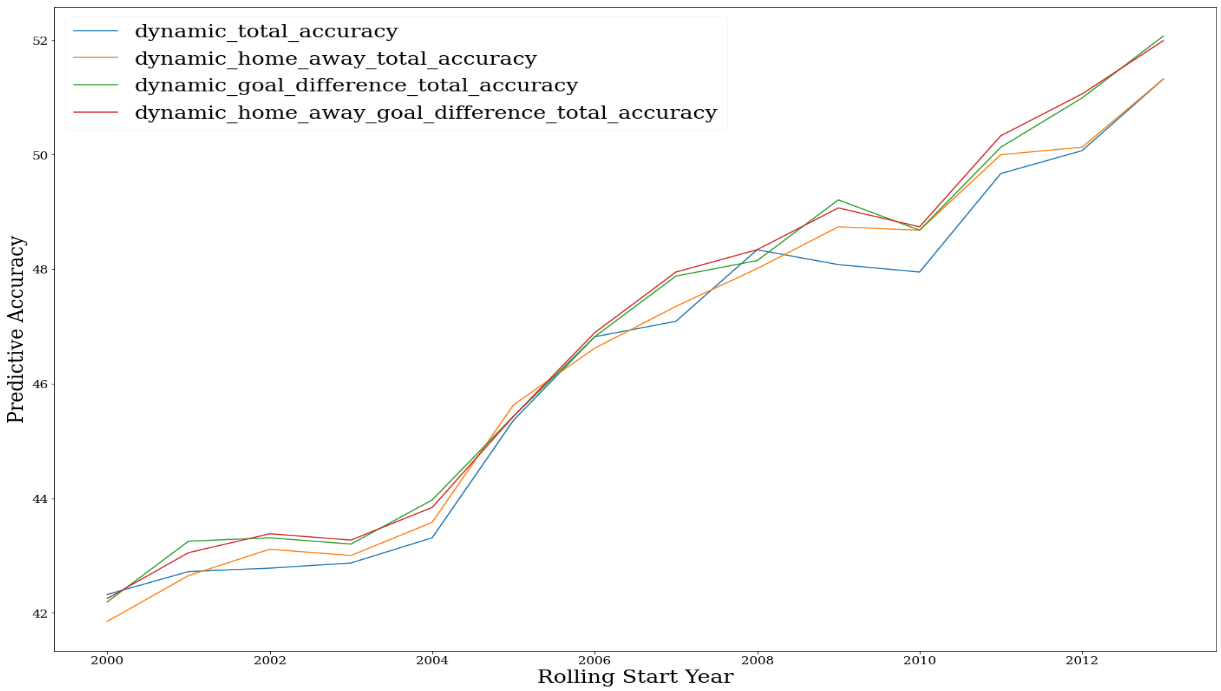

Figure B.31: Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

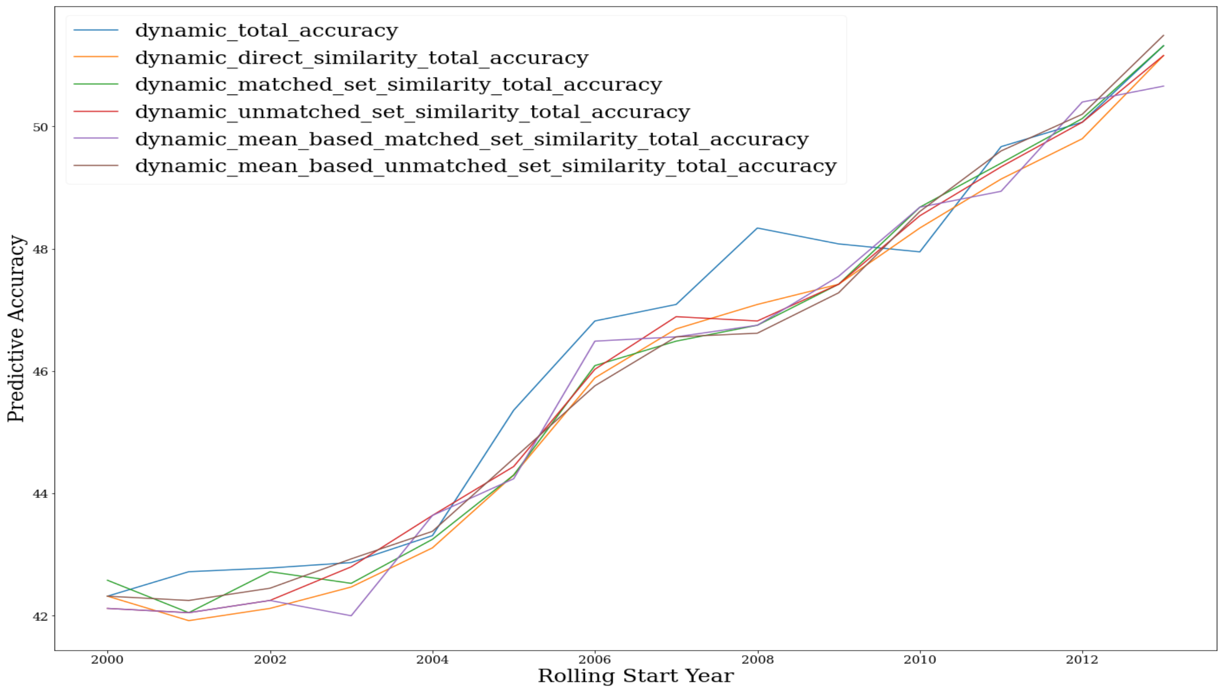

Figure B.32: Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches

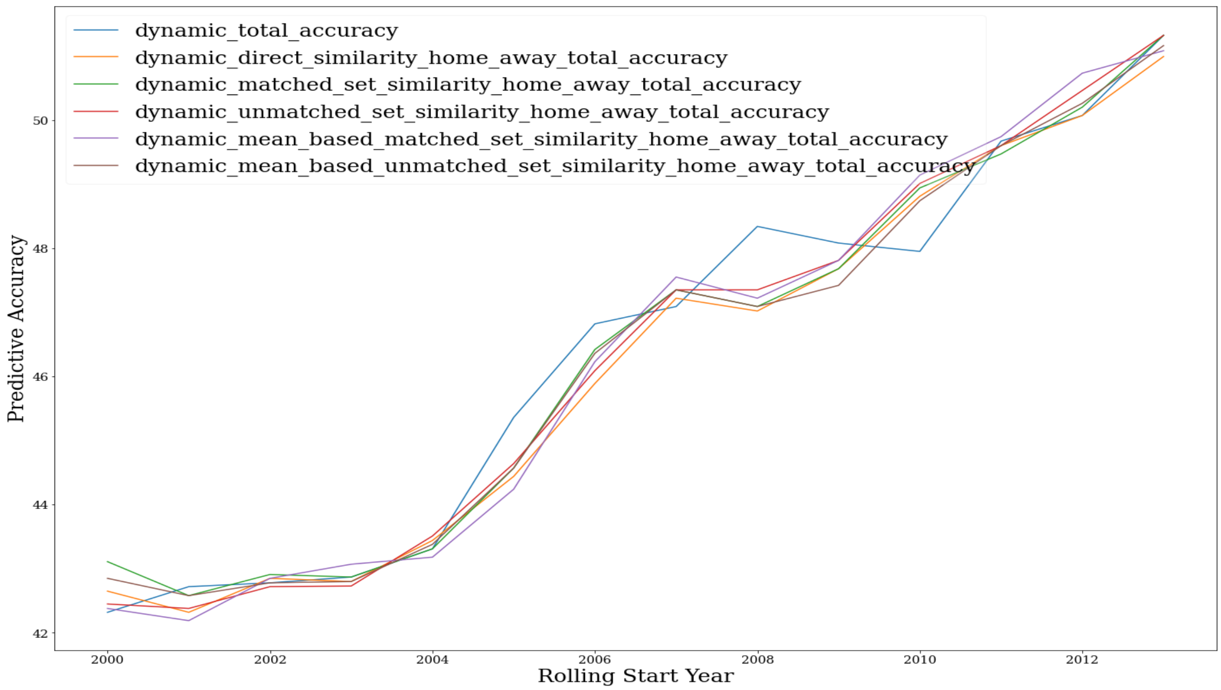

Figure B.33: Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches

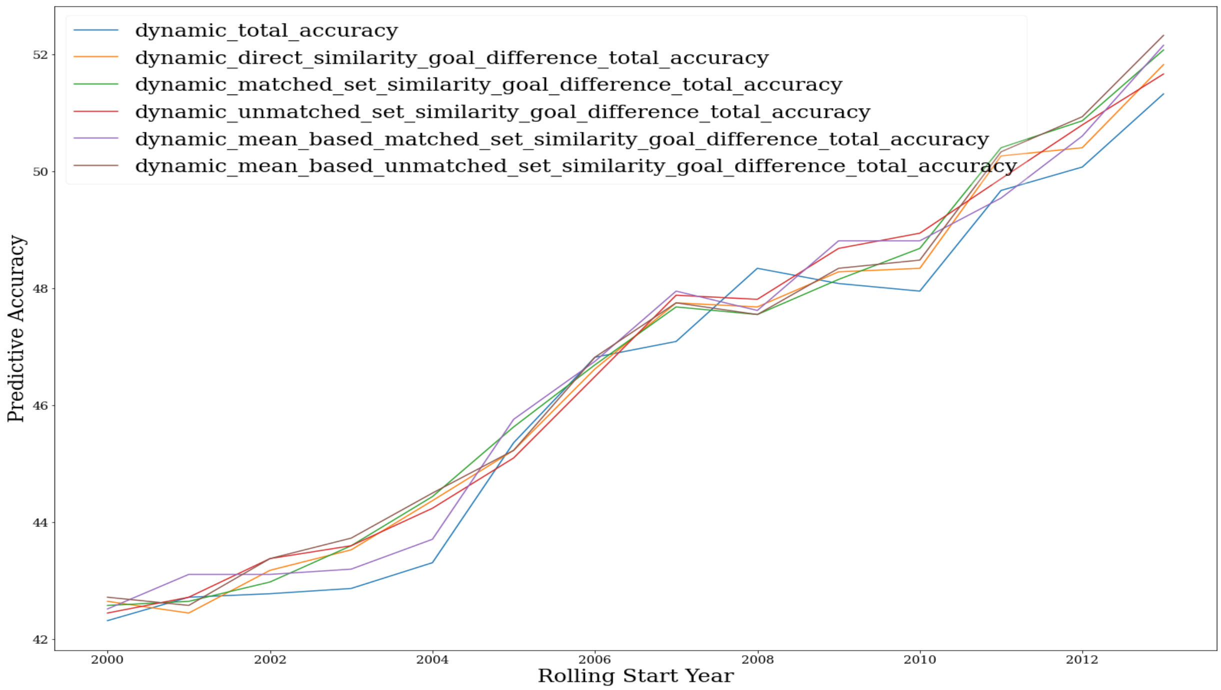

Figure B.34: Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches

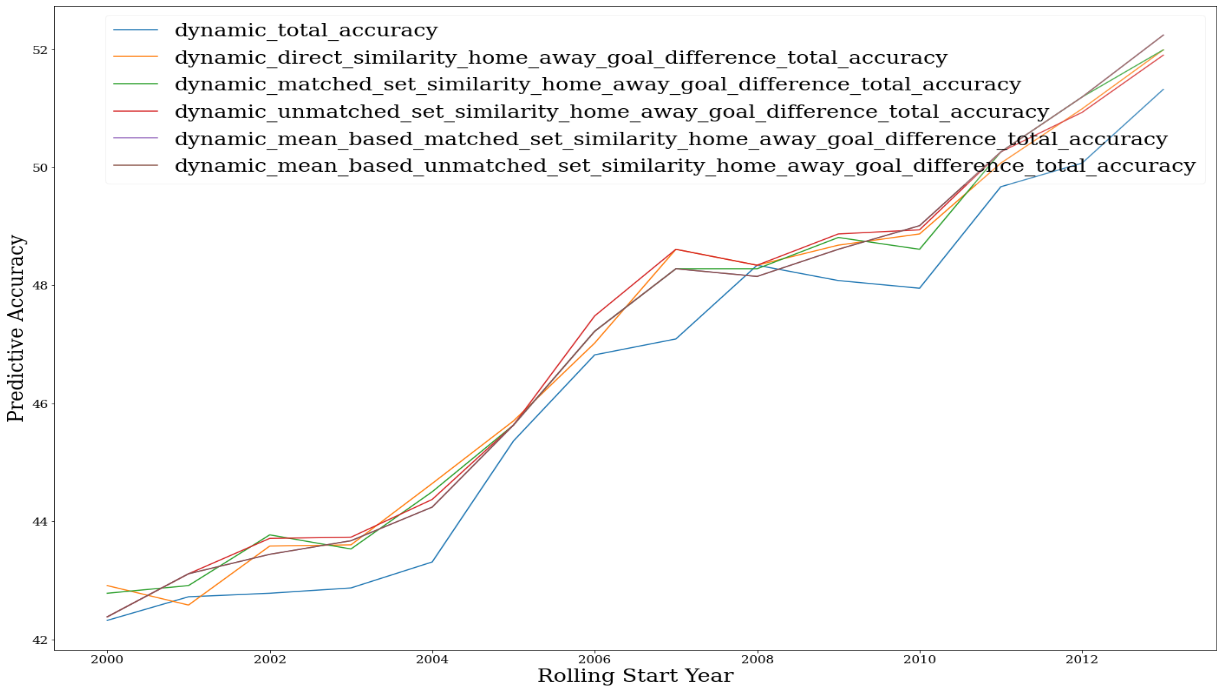

Figure B.35: Spanish La Liga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches

B.3.2.3 Germany

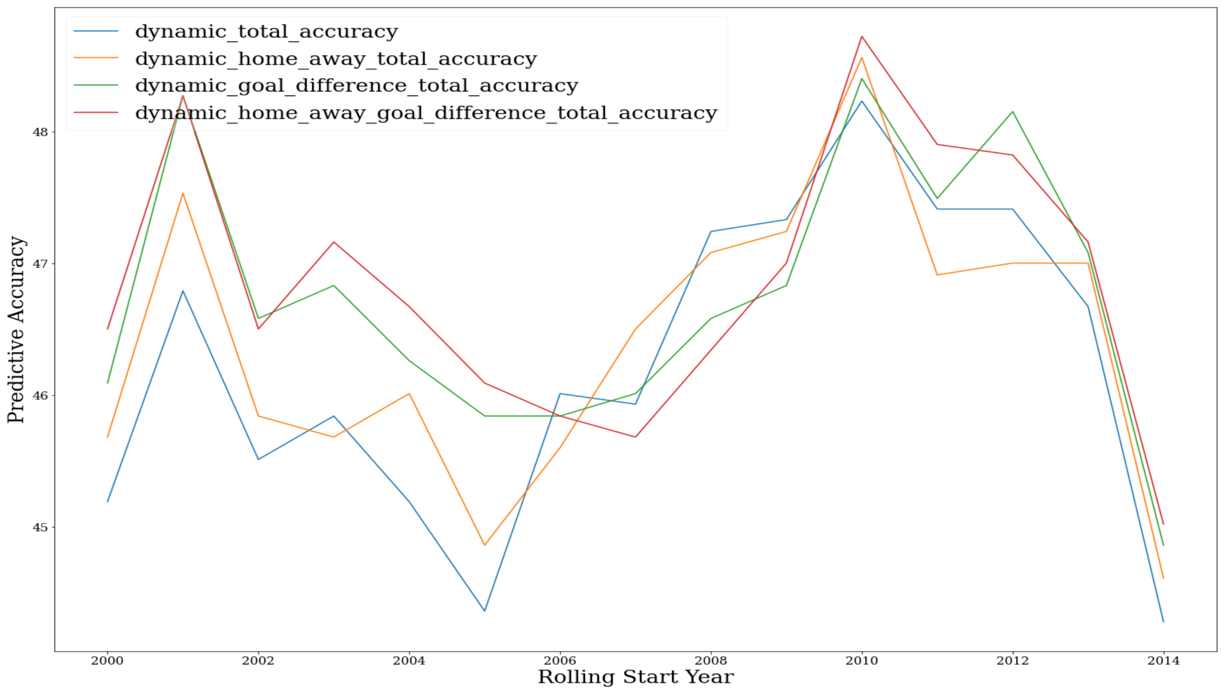

Figure B.36: German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

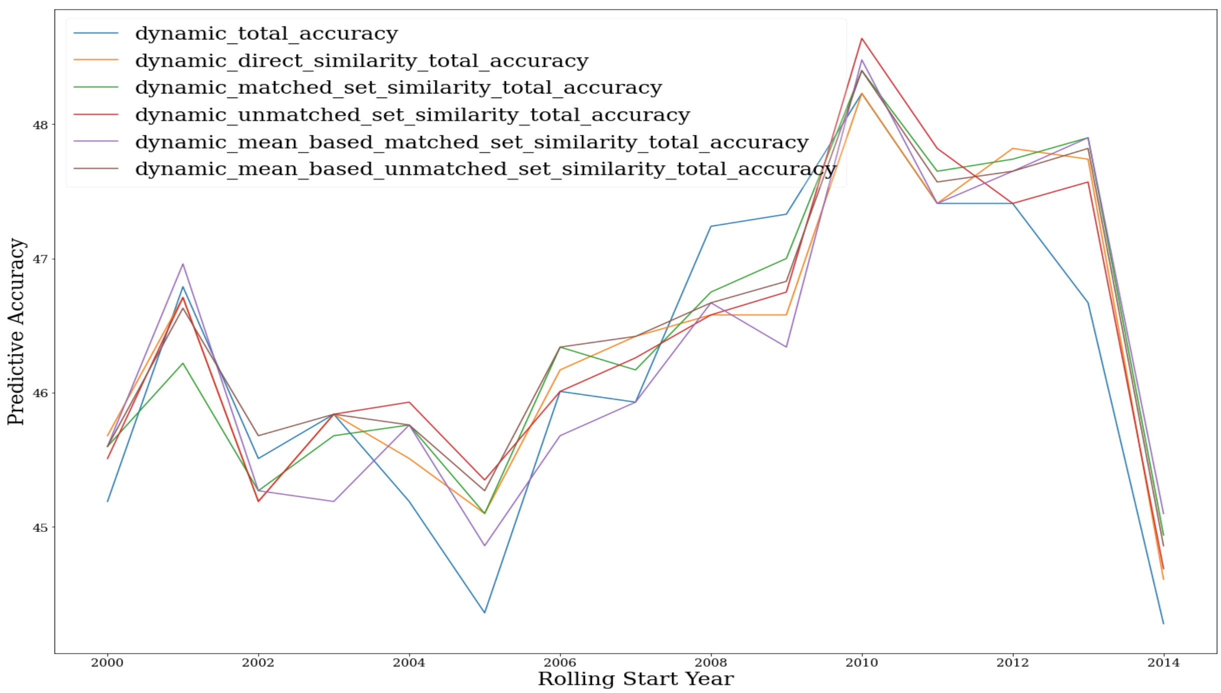

Figure B.37: German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches

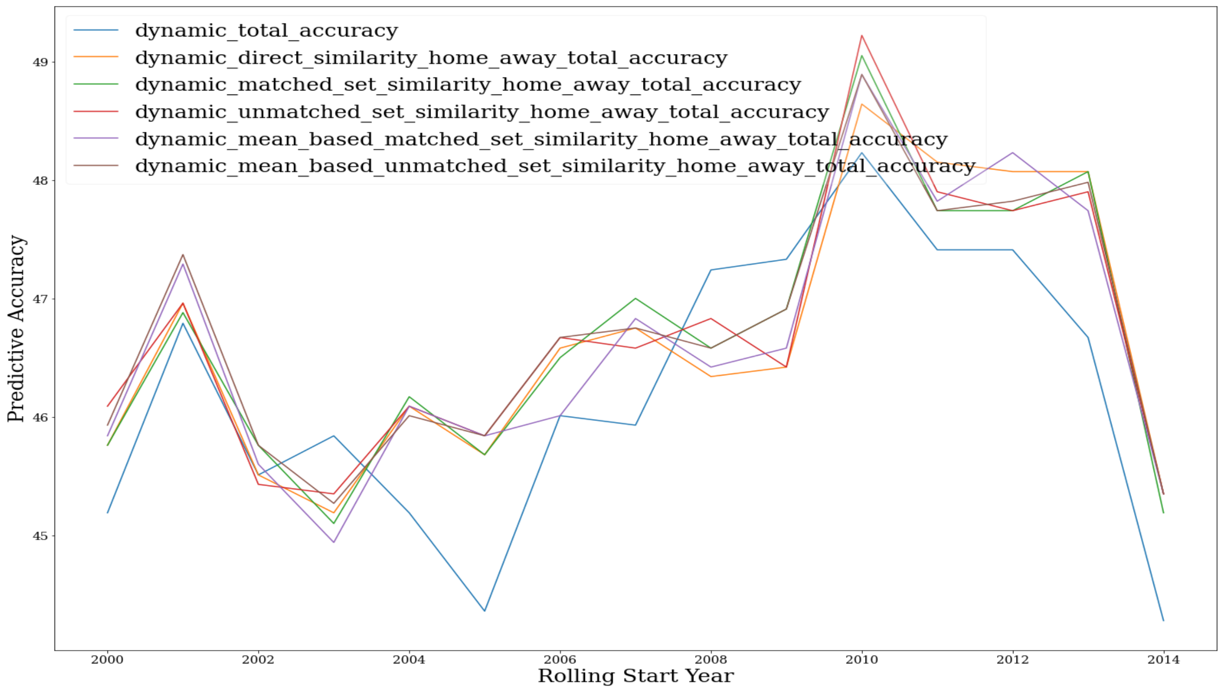

Figure B.38: German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches

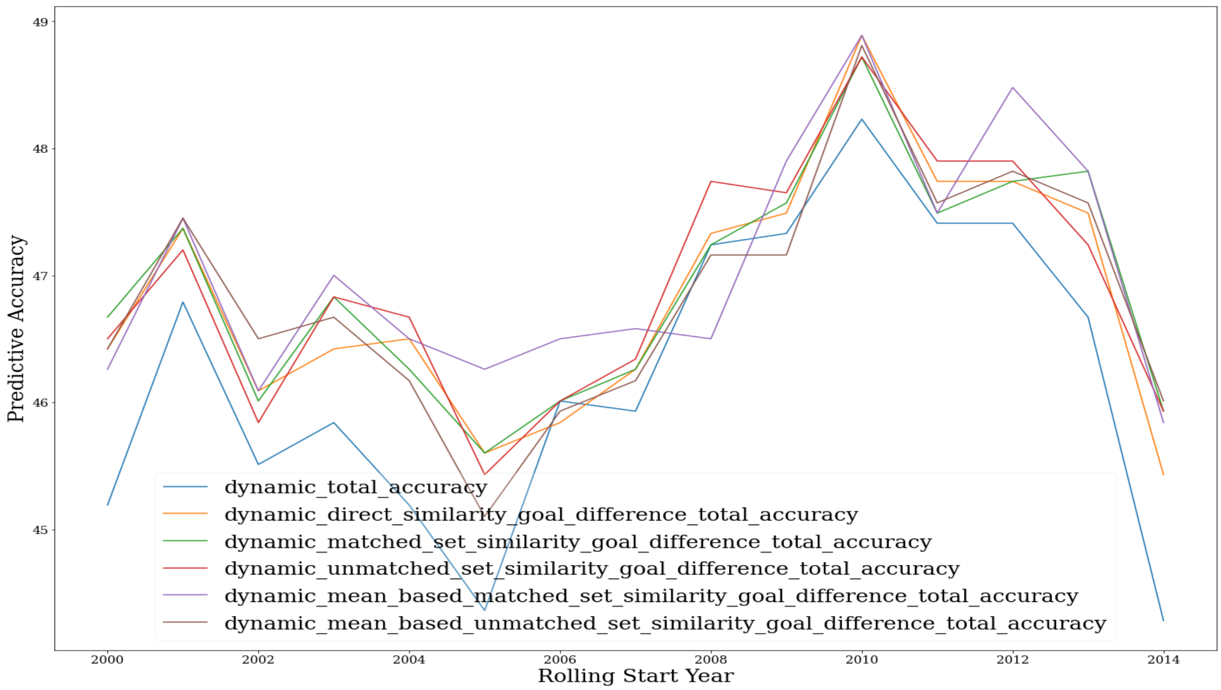

Figure B.39: German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches

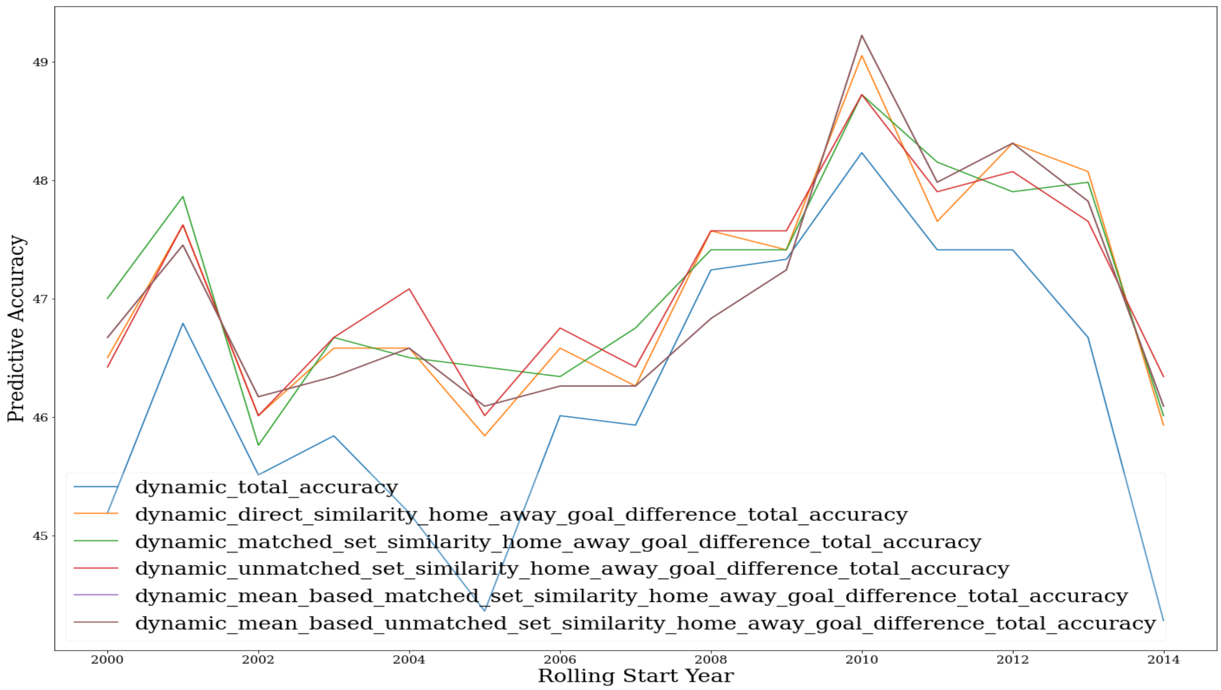

Figure B.40: German Bundesliga 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches

B.3.2.4 Italy

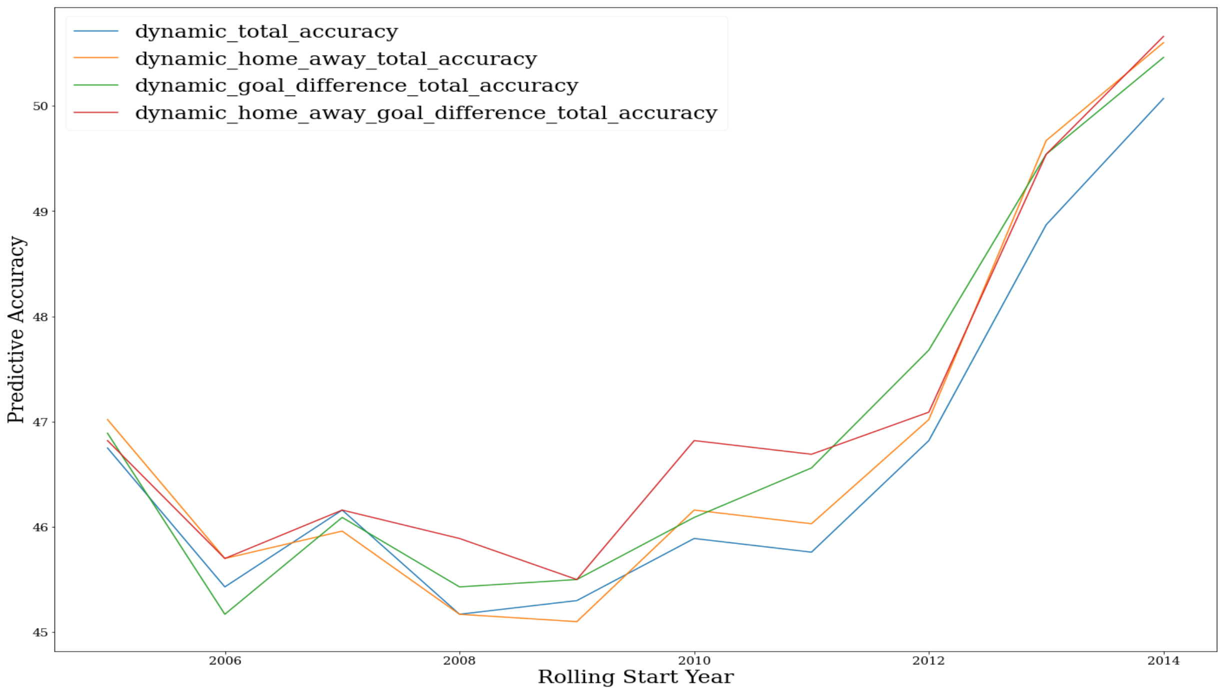

Figure B.41: Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

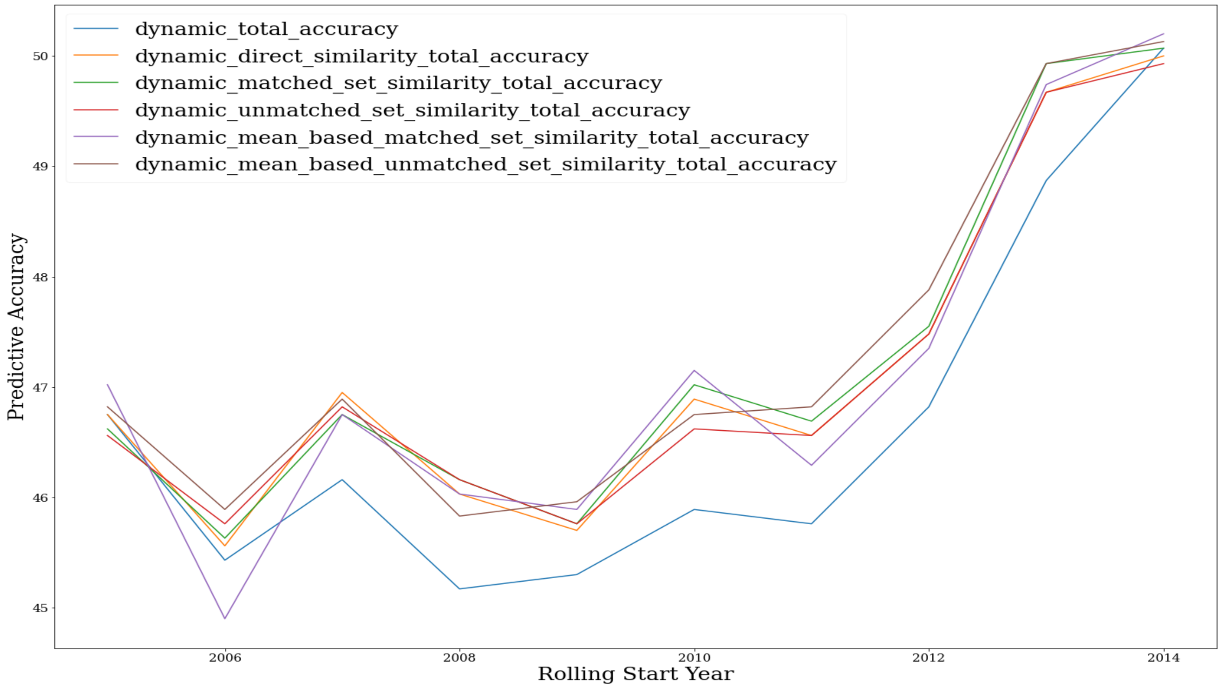

Figure B.42: Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches

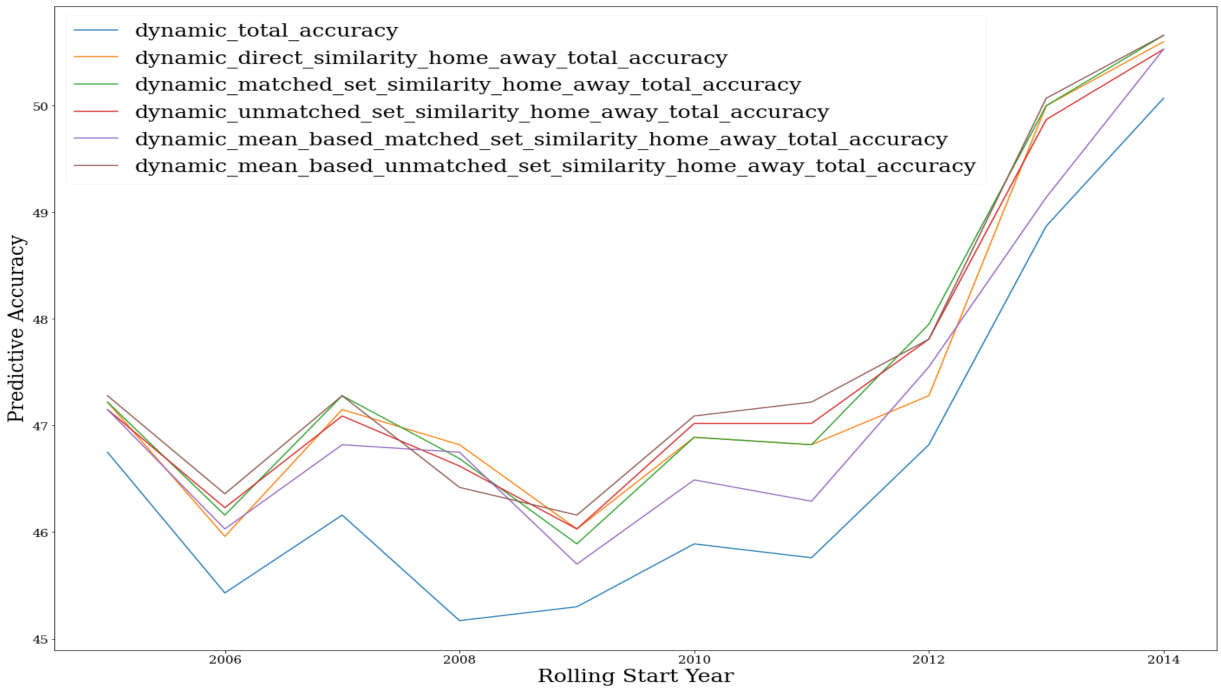

Figure B.43: Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches

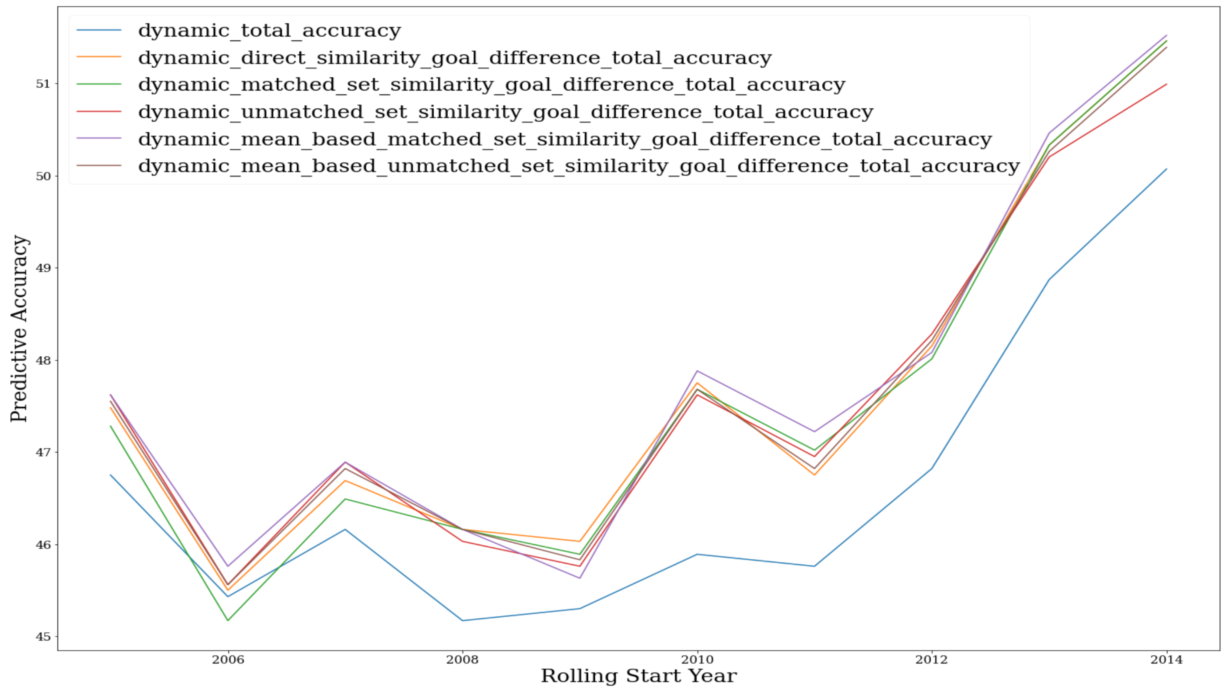

Figure B.44: Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches

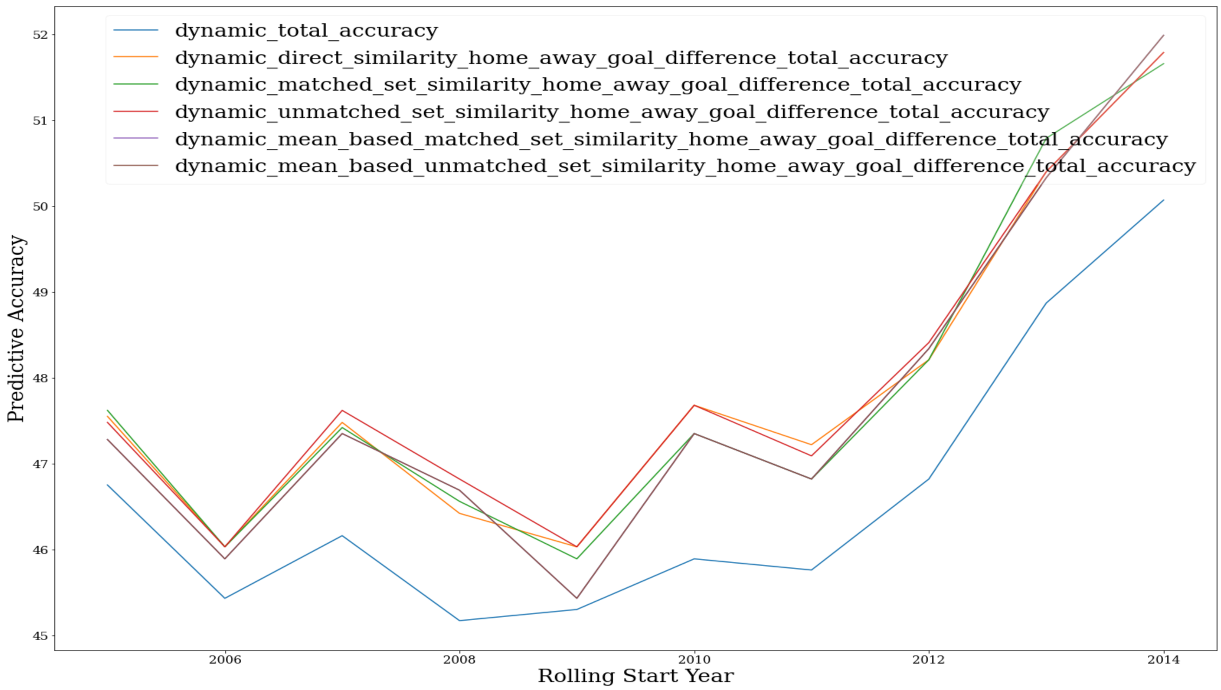

Figure B.45: Italian Serie A 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches

B.3.2.5 France

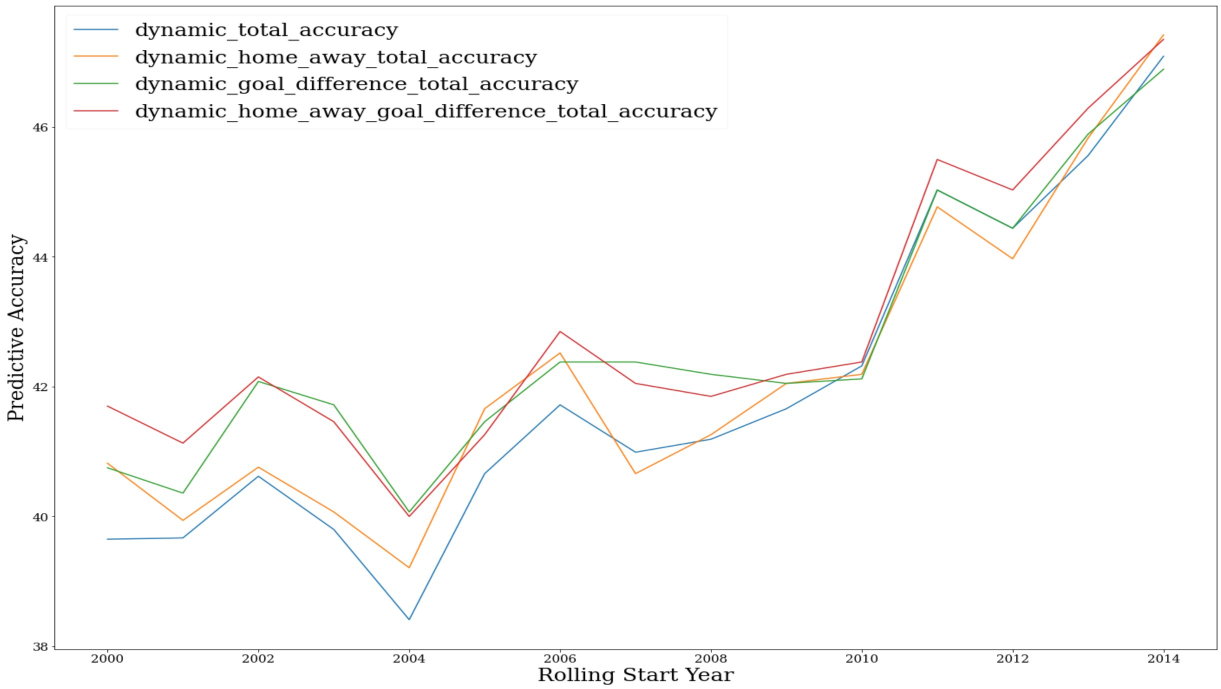

Figure B.46: French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Home/Away, Goal/Difference, and Home/Away + Goal Difference Approaches

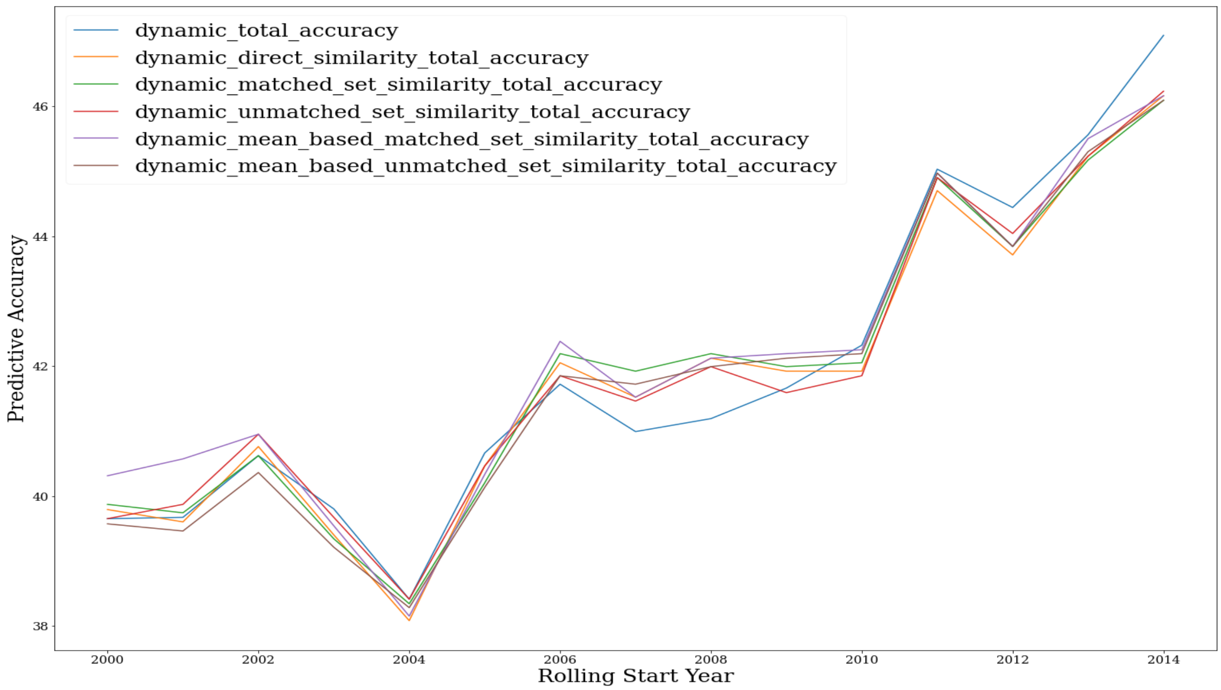

Figure B.47: French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Similarity Approaches

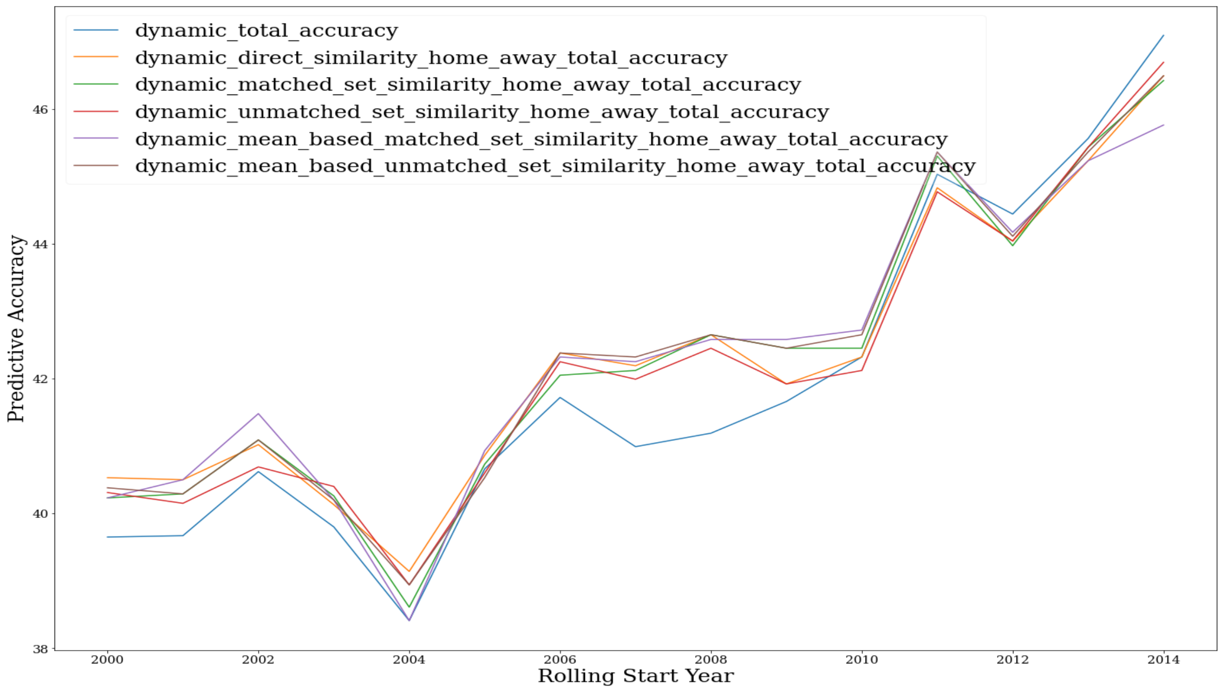

Figure B.48: French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Similarity Approaches

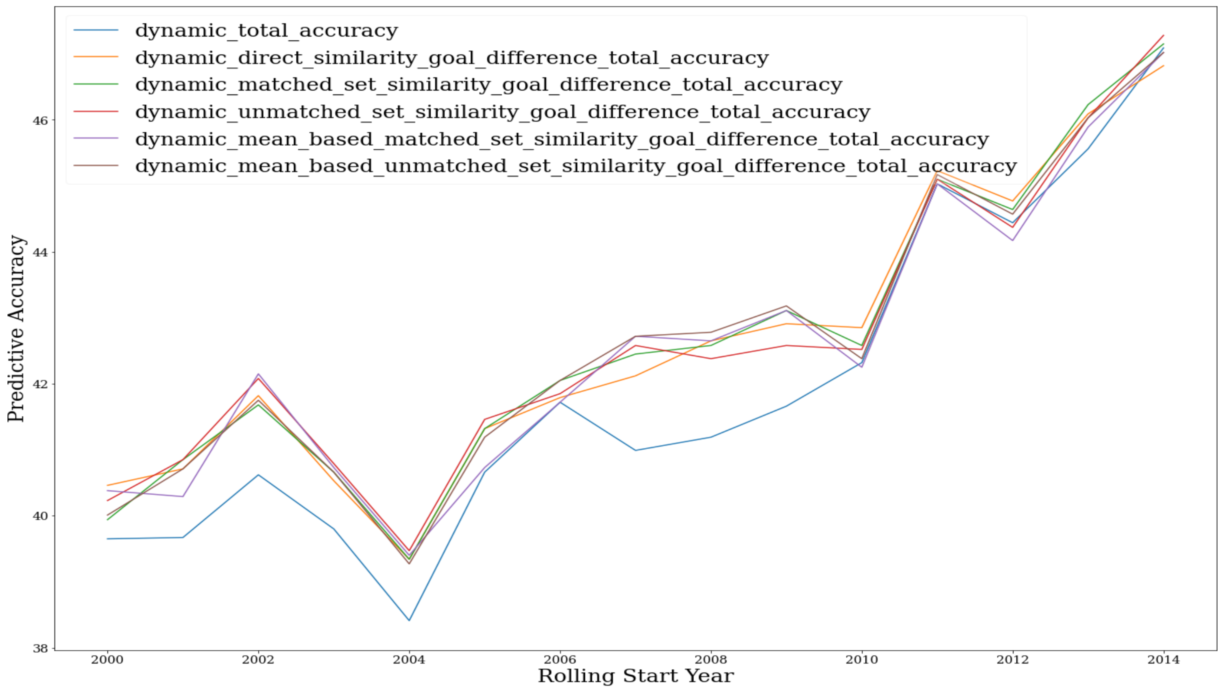

Figure B.49: French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Goal Difference Similarity Approaches

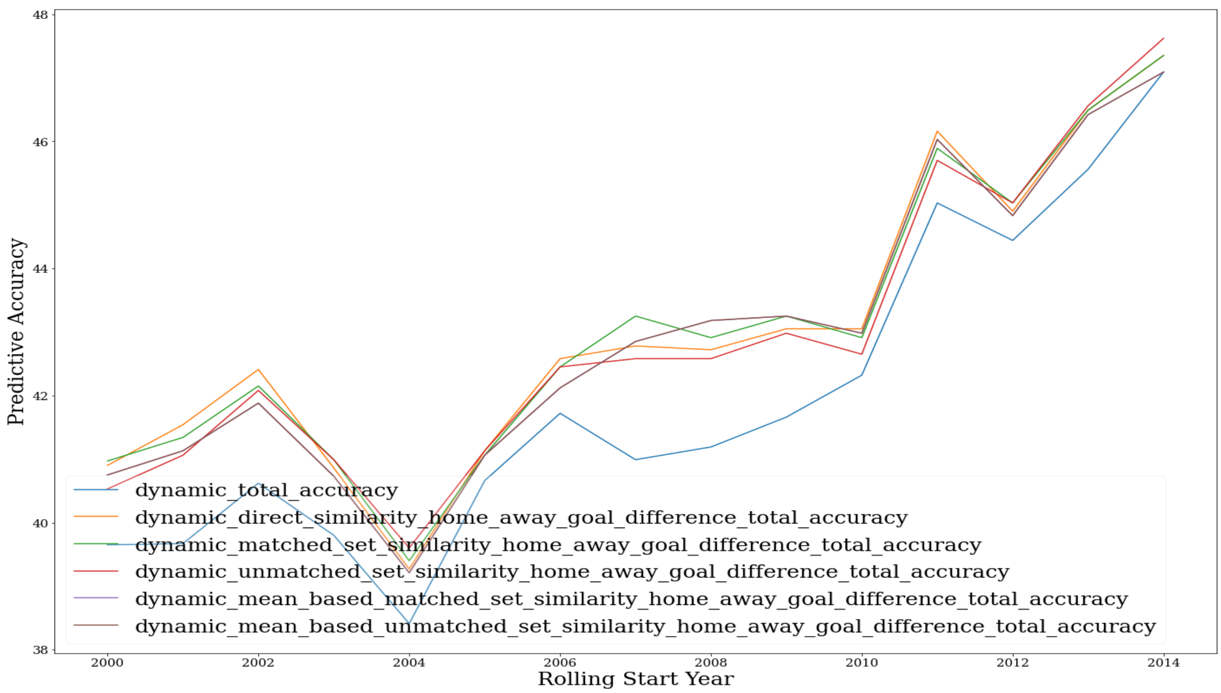

Figure B.50: French Ligue 1 4-Year Rolling Predictive Accuracy: Dynamic Home/Away Goal Difference Similarity Approaches

B.4 Attempts to Beat Betting Odds

This appendix explains the attempts we made to modify the prediction algorithm in order to better-capture ties and, in one case, upsets. Section B.4.1 describes an approach that treats match result predictions differently if teams are within k ranks of each other. Section B.4.2 explains an approach that treats match result predictions differently if teams are within k standard deviations in terms of total score (the number from which the ranks are derived). Section B.4.3 introduces an approach that predicts a random result in accordance with the empirical probabilities of home wins, draws, and away wins for teams that are within k standard deviations in terms of total score. Section B.4.4 details an approach that predicts ties for teams that have similarity metric values beyond a threshold. Section B.4.5 catalogues two attempts at using machine learning to build classification systems for the match results; Section B.4.5.3 uses a logistic regression model and B.4.5.4 uses a support vector machine.

B.4.1 Tie Prediction Within k Ranks

This approach deploys a simple heuristic in an attempt to generate tie predictions. At each gameweek, if two teams have ranks i and j, if $|i - j| > k$, where k is a user-chosen parameter, we predict that the team with the lower, better rank will win, regardless of home/away status. However, if the ranks are such that $|i - j| \leq k$, we use the following logic:

- If the home team is the better-ranked team, we predict a home win.
- If the away team is the better-ranked team, we predict a tie.

Essentially, what this heuristic supposes is that part of the benefit of being the better-ranked team is mitigated by home-field advantage, which softens the prediction of an away team win to a draw.

B.4.2 Tie Prediction Within k Total Score Standard Deviations

This approach is similar to that of Section B.4.1. However, instead of using ranks, we use the total score. Given that the total scores will grow as we proceed through any ranking window, we use the standard deviation of the total scores at each gameweek as a metric. Specifically, we use k standard deviations, where k is a user-defined parameter; let $\hat{d}_s$ be the standard deviation of the total scores in the given gameweek. Let s_i and s_j be the total scores of teams i and j respectively in our given gameweek. The heuristic operates as follows. If $|s_i - s_j| > k\hat{d}_s$, we predict that the better-ranked team will win. If $|s_i - s_j| \leq k\hat{d}_s$, then our prediction uses the following logic:

- If the home team is the better-ranked team, we predict a home win.
- If the away team is the better-ranked team, we predict a tie.

B.4.3 Randomized Result Prediction Within k Total Score Standard Deviations

This approach is similar to B.4.2 and uses the same standard deviation threshold. However, the prediction logic changes when $|s_i - s_j| \leq k\hat{d}_s$. Namely, we predict the outcome using a random draw. Specifically, we use the following probabilities:

- home win: 0.50
- draw: 0.25
- away win: 0.25.

B.4.4 Similarity-Based Tie Prediction

For this approach, we compute and use one of our similarity metrics (Direct Similarity, Matched Set Similarity, Unmatched Set Similarity, Mean-Based Direct Similarity, Mean-Based Matched Similarity, or Mean-Based Unmatched Set Similarity) for the ranking window up to the gameweek in question. We then compute both the mean similarity $\bar{x}$ and standard deviation of the similarities $\hat{d}_s$. Let the similarity of teams i and j be x_{ij}.

We then proceed to do the following for match result predictions. If $x_{ij} < \bar{x} + k\hat{d}_s$, we predict that the better-ranked team will win. If $x_{ij} \geq \bar{x} + k\hat{d}_s$, we predict that the teams will draw. The logic behind this approach is that teams with higher similarity should, on average, be closer to each other in terms of ability, so they are more likely to tie.

B.4.5 Machine Learning Approaches to Predicting Ties, Upsets, and Non-Upsets

B.4.5.1 Training and Prediction Approach

Before discussing the variables in Section B.4.5.2 and the models in sections B.4.5.3 and B.4.5.4, we first discuss how the machine learning model is trained and used for prediction. It does not make sense to train the machine learning models when there is insufficient data. As such, we use our regular predictive accuracy for the first N gameweeks, where N is the number of features we use in our machine learning models. We use this as a heuristic because this means we have roughly 10 games in the training dataset per feature, as we average 10 games per gameweek.

To predict game outcomes in gameweek $t + 1$, we train our machine learning model on all game data from time period 1 through t inclusive. We know that ties and upsets are relatively infrequent phenomena, so we use all the available data for training so as to give ourselves the best chance of finding useful patterns in our explanatory variables. Once our models are trained, we use our explanatory variable values for the upcoming games of gameweek $t + 1$ in our trained model to make predictions. We then compare these predictions with the realizations to compute our accuracy.

B.4.5.2 Model Variables

Both models in this section use the same set of dependent and independent variables. As such, we discuss those here. First, we mention our dependent variable: match outcome, which we denote as y_i for observation i. Match outcome has three categories: ties, upsets, and non-upsets. We define upsets as matches where the worse-ranked team wins. Non-upsets are games where the better-ranked team wins. For our models, we use the following to denote each of these outcomes:

- tie: 0
- upset: 1
- non-upset: 2.

We choose to use these classes because ranking systems, on average, tend to be good at predicting non-upsets, but they struggle with ties and upsets, since these defy the ordering proposed by the ranking system.

The dependent variables used for training and testing are explained below. Note that some variables are indexed by the time of the observation, which we denote as t_i, while some are indexed by the latest time period in the training dataset, which we denote as t.

- $rank\ difference$: a variable that equals home rank − away rank at the time t_i; note that the sign of this variable indicates whether the home team is better-ranked.
- $total\ score\ difference$: a variable that equals home total score − away total score at time t_i.
- $(total\ score\ difference)^2$: squared $total\ score\ difference$ at time t.
- $mean - based\ direct\ similarity$: the Mean-Based Direct Similarity of the teams in the match at time t; this means we use our latest similarity metric for each pair of teams. Note that we do this because the latest similarity value best-reflects how

competitive the teams are, but this value only gets revealed over many gameweeks. Further, for clarification, this does not introduce look-ahead bias because we retrain our machine learning models before each gameweek.

- $mean - based\ matched\ set\ similarity$: The Mean-Based Matched Set Similarity of the teams in the match at time t.
- $mean-based\ unmatched\ set\ similarity$: The Mean-Based Unmatched Set Similarity of the teams in the match at time t.
- $quantile_{ij}$: dummy variables indicating which quantile the home team, i and the away team, j are located in the rankings at time t_i. If we have q quantiles, we leave out the $quantile_{qq}$ dummy variable (both teams being in the worst quantile, rank-wise). This variable is included as we might suspect that there are certain structural properties to ties and upsets, where it might be more likely for, say, an upset to occur when teams play and are separated by more than one quantile.
- *team name*: we arrange our set of teams alphabetically and create dummy variables for each except the first; this is to capture if certain teams have greater or lesser tendencies to tie or have upsets.
- $\Delta total\ score$: For both the home and away team, we include proportional changes in total score, the number of which is determined by parameter p. For example, if we set $p = 3$, we include the proportional changes in total score generated in going from period t_{i-3} to t_{i-2}, from period t_{i-2} to t_{i-1}, and from period t_{i-1} to t_i.

B.4.5.3 Logistic Regression

Using Myers et al. (2010), we briefly outline our approach to the logistic regression model. We have three categories for prediction: ties, upsets, and non-upsets. These categories are changed to numerical values, where $y_i = 0$ for ties, $y_i = 1$ for upsets, and $y_i = 2$ for non-upsets. We collect all of the variables in Section B.4.5.2 into a matrix $\mathbf{X}$, where row $\mathbf{X}_i$ represents the collection of features associated with observation i. We then use the following equations to define our model:

$$\mathbf{P}(y_i = 0) = \frac{1}{1 + \sum_{j=1}^{2} \exp\left[\mathbf{X}_i \boldsymbol{\beta}^{(j)}\right]}; \tag{B.1}$$

$$\mathbf{P}(y_i = 1) = \frac{\exp\left[\mathbf{X}_i \boldsymbol{\beta}^{(1)}\right]}{1 + \sum_{j=1}^{2} \exp\left[\mathbf{X}_i \boldsymbol{\beta}^{(j)}\right]}; \tag{B.2}$$

$$\mathbf{P}(y_i = 2) = \frac{\exp\left[\mathbf{X}_i \boldsymbol{\beta}^{(2)}\right]}{1 + \sum_{j=1}^{2} \exp\left[\mathbf{X}_i \boldsymbol{\beta}^{(j)}\right]}, \tag{B.3}$$

where $\boldsymbol{\beta}^{(j)}$ denotes the parameter vector for class j. Note that we have only two category vectors, since we only need the probability of two classes to compute the probability of the third class. We compare each response category to the baseline, producing logits:

$$\ln \frac{\mathrm{P}(y_i = 1)}{\mathrm{P}(y_i = 0)} = \mathbf{X}_i \boldsymbol{\beta}^{(1)}; \tag{B.4}$$

$$\ln \frac{\mathrm{P}(y_i = 2)}{\mathrm{P}(y_i = 0)} = \mathbf{X}_i \boldsymbol{\beta}^{(2)}. \tag{B.5}$$

We use conventional maximum likelihood estimation to estimate the parameters.

B.4.5.4 Support Vector Machine

One approach to using support vector machines in the context of multiple class prediction is to build binary classifiers between all classes, then the classification of a new data point is done via a winner-takes-all strategy. Namely, the class that obtains the highest score via an output function is the one to which the new point is assigned. We briefly describe the linear support vector machine framework.

Let $(\mathbf{x_1}, y_1), \ldots, (\mathbf{x_n}, y_n)$ be our data points, where the $\mathbf{x_i}$ are the feature vectors and y_i are our target variables. For the sake of this explanation, y_i can take on two values, one per class: -1 or 1. The goal of a support vector machine is to find the maximum-margin hyperplane that best divides the $\mathbf{x_i}$ where y_i is one class from the $\mathbf{x_i}$ where y_i is the other class. A hyperplane can be written as the set of points $\mathbf{x}$ satisfying:

$$\mathbf{w}^T\mathbf{x} - b = 0. \tag{B.6}$$

where $\mathbf{w}$ is the normal vector to the hyperplane and the parameter $\frac{b}{\|\mathbf{w}\|}$ determines the offset of the hyperplane from the origin along the normal vector $\mathbf{w}$.

If the classes are linearly separable, we can cast this as the following optimization problem:

$$\min_{\mathbf{w},b} \|\mathbf{w}\|^2; \tag{B.7}$$

$$\text{subject to } y_i(\mathbf{w}^t\mathbf{x}_i - b) \geq 1 \; \forall i \in \{1, \ldots, n\}. \tag{B.8}$$

We solve this optimization problem for $\mathbf{w}$ and b.

If the training data is not linearly separable, we use what is called a soft-margin approach, which includes a loss penalty in case points lie on the incorrect side of the hyperplane. One example of this loss is the hinge loss function:

$$\max\big(0, 1 - y_i(\mathbf{w}^t\mathbf{x}_i - b)\big). \tag{B.9}$$

Through some adjustments, our optimization problem then becomes:

$$\min_{\mathbf{w},b,\gamma} \|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\gamma_i; \tag{B.10}$$

$$\text{subject to } y_i(\mathbf{w}^t\mathbf{x}_i - b) \geq 1 - \gamma_i, \gamma_i \geq 0 \ \forall i \in \{1, \ldots, n\}, \tag{B.11}$$

where C is a parameter for tuning the penalty imposed by the loss function.