

DOCTOR OF PHILOSOPHY IN MANAGEMENT

ESSAYS ON WARRANTY CLAIMS

By

AKSHITA SRIVASTAVA



भारतीय प्रबंध संस्थान बेंगलूर
INDIAN INSTITUTE OF MANAGEMENT
BANGALORE

2022

DOCTOR OF PHILOSOPHY IN MANAGEMENT

ESSAYS ON WARRANTY CLAIMS

By

Akshita Srivastava

A Dissertation submitted in Partial Fulfillment of the Requirements for the
degree of Doctor of Philosophy at

INDIAN INSTITUTE OF MANAGEMENT BANGALORE

2022

Prof. Ananth Krishnamurthy
Chairperson
Doctoral Programme

Prof. U Dinesh Kumar
Chairperson
Dissertation Advisory Committee

Members of the Dissertation Advisory Committee

- | | |
|----------------------------|-------------|
| 1. Prof. U Dinesh Kumar | Chairperson |
| 2. Prof. Srinivas Prakhya | Member |
| 3. Prof. Rajluxmi V Murthy | Member |

Copyright © 2022 by Akshita Srivastava
All rights reserved.

Dedication

I dedicate my thesis work to my husband Mr. Rajiv Mukul and my son Hemang, who fully supported me throughout my PhD journey.

This work is also dedicated to the memory of my beloved mother who always prayed for my success.

Last but not the least, I would like to thank my elders and the almighty God whose blessings guided me in every step of my journey.

ACKNOWLEDGEMENTS

Words cannot express my gratitude to the professor and chair of my committee, Prof. U Dinesh Kumar for his patience and invaluable feedback. He provided me with constant guidance and was always willing to assist me in every stage of my thesis.

I also could not have undertaken this journey without my defense committee members, Prof. Srinivas Prakhya and Prof. Rajluxmi V Murthy, who generously provided their comments and helped me through my thesis in every possible way.

I would also like to show my gratitude to my external review committee members, Prof. Ramanathan and Dr. Vimala Rani for their time and patience in reviewing my thesis.

I would also like to thank the Data Center and Analytics team of IIMB, who provided me with the data used in the study. During the engaging discussion with the DCAL members, I was able to clarify my doubts regarding the data.

Lastly, I want to thank my family for believing in me and supporting me.

Table of Contents

ABSTRACT	1
1. INTRODUCTION	1
1.1 BACKGROUND	1
1.2 MOTIVATION AND SIGNIFICANCE	4
1.3 PROBLEM STATEMENT AND CONTRIBUTION	5
1.4 FRAMEWORK OF THE STUDY.....	8
1.4.1 <i>Essay 1: Modelling variation in claim counts data</i>	9
1.4.2 <i>Essay 2: Modelling variation in lifetimes</i>	11
1.4.3 <i>Essay 3: Modelling variation in time between failures</i>	13
1.5 NEED FOR THE HYBRID APPROACH	14
1.6 STRUCTURAL FRAMEWORK OF THE THREE ESSAYS	15
1.7 CONTRIBUTION OF THE THREE ESSAYS	16
1.8 DATA SOURCE	17
1.9 INTRODUCTION TO THE UNDERLYING THEORY	18
1.9.1 <i>Survival analysis</i>	18
1.9.1.1 <i>Survivor function</i>	18
1.9.2 <i>Hazard function and hazard ratio</i>	18
1.9.3 <i>Cox proportional hazards model</i>	20
1.9.4 <i>Kendall's Tau correlation</i>	22
1.9.5 <i>Latent class</i> :	22
1.9.6 <i>Description of dependent and independent variables used in each</i>	22
1.9.6.1 <i>Variables used in essay 1</i>	24
1.9.6.1.1 Dependent variable in essay 1	24
1.9.6.1.2 Independent variables in essay 1	24
1.9.6.1.3 Variables used in clustering	27
1.9.6.2 <i>Variables used in essay 2</i>	29
1.9.6.2.1 Dependent variable in essay 2	29
1.9.6.2.2 Independent variables in essay 2	29
1.9.6.3 <i>Variables used in essay 3</i>	31
1.9.6.3.1 Dependent variable in essay 3	31
1.9.6.3.2 Independent variables in essay 3	32
2. LITERATURE REVIEW	35
2.1 HETEROGENEITY IN RELIABILITY	37
2.1.1 <i>Importance of addressing frailty effects</i>	38
2.2 HETEROGENEITY IN WARRANTY	39
2.2.1 <i>Modelling of lifetime data in reliability and warranty</i>	39

2.2.2	<i>Literature review on dependence in reliability and warranty</i>	41
2.2.3	<i>Literature review on dealership studies</i>	43
2.3	RESEARCH GAPS	44
2.3.1	<i>Critical literature review: Research gap 1</i>	45
2.3.2	<i>Critical literature review: Research gap 2</i>	46
2.3.3	<i>Critical literature review: Research gap 3</i>	48
2.4	CONTRIBUTIONS	49
2.4.1	<i>Contribution 1</i>	49
2.4.1.1	Contribution of thesis in using latent class Poisson regression methodology for variability in claim counts	50
2.4.1.2	Comparison with the previous models	51
2.4.2	<i>Contribution 2</i>	52
2.4.2.1	Development of a hybrid modelling approach	53
2.4.2.2	Comparison with the previous models	56
2.4.3	<i>Contribution 3</i>	56
2.4.3.1	Comparison with the previous models	58
2.5	BASIC CONCEPTS FOR MODELS USED TO ADDRESS THE CONTRIBUTIONS	58
2.5.1	<i>Econometric models</i>	58
2.5.1.1	Poisson model	58
2.5.1.2	Poisson regression model	59
2.5.1.3	Finite mixture model	60
2.5.1.4	Weibull mixture model and Weibull proportional hazards model	62
2.5.1.5	Frailty model	63
2.5.1.6	Shared frailty model	64
2.5.1.7	Recurrent event models	65
2.5.1.8	Empirical Bayes analysis	66
2.5.2	<i>Machine learning models</i>	67
2.5.2.1	Cluster analysis	67
2.6	LITERATURE REVIEW FOR MODELS USED IN ESSAY 1	69
2.6.1	<i>Cluster analysis in manufacturing</i>	69
2.6.1.1	Cluster analysis studies in warranty	70
2.6.1.2	Clustering of survival data	71
2.6.2	<i>Literature review for Poisson model</i>	72
2.6.3	<i>Literature review for Poisson regression</i>	73
2.6.4	<i>Literature review for finite mixture model</i>	74
2.7	NEED FOR THE MODELS USED IN ESSAY 1	74
2.7.1	<i>Justification for cluster analysis</i>	74
2.7.2	<i>Justification for choosing Poisson distribution</i>	75
2.7.3	<i>Justification for choosing Poisson regression</i>	75
2.7.4	<i>Justification for choosing latent class Poisson regression</i>	76
2.8	LITERATURE REVIEW FOR MODELS USED IN ESSAY 2	76

2.8.1	LITERATURE REVIEW FOR DEPENDENCE IN SURVIVAL.....	76
2.8.1.1	Observed covariates.....	77
2.8.1.2	Unobserved covariates.....	77
2.8.2	<i>Literature review for Weibull proportional hazards model (WPHM)</i>	77
2.8.3	<i>Literature review for frailty models</i>	78
2.8.4	<i>Literature review for shared frailty models</i>	79
2.9	NEED FOR THE MODELS USED IN ESSAY 2.....	80
2.9.1	<i>Need for Weibull mixture model and Weibull proportional hazards model</i>	81
2.9.2	<i>Need for the shared frailty model</i>	82
2.10	LITERATURE REVIEW FOR MODELS USED IN ESSAY 3.....	83
2.10.1	<i>Literature review on the Right first time approach</i>	83
2.10.2	<i>Literature review for recurrent event models</i>	84
2.10.3	<i>Literature review for empirical Bayes analysis</i>	85
2.11	NEED FOR THE MODELS USED IN ESSAY 3.....	86
2.11.1	<i>Justification for using recurrent event models</i>	86
2.11.2	<i>Justification for using empirical Bayes analysis</i>	87
ESSAY 1: MODELLING VARIATION IN THE WARRANTY CLAIM COUNTS DATA		89
3.1	RESEARCH QUESTION 1.....	89
3.2	RESEARCH QUESTION 2.....	90
3.3	PROPOSED METHODOLOGY FOR INCORPORATING VARIABILITY.....	91
3.4	CLUSTER ANALYSIS.....	93
3.4.1	<i>Data and variables for cluster analysis</i>	93
3.4.1.1	<i>Steps in data pre-processing for cluster analysis</i>	93
3.4.2	<i>Methodology for cluster analysis</i>	95
3.4.3	<i>Findings of cluster analysis</i>	96
3.5	LATENT CLASS POISSON REGRESSION METHODOLOGY FOR VARIABILITY IN CLAIM COUNTS.....	98
3.5.1	<i>Estimation</i>	101
3.6	DATA AND VARIABLES FOR LATENT POISSON REGRESSION METHODOLOGY.....	101
3.6.1	<i>Model implementation</i>	102
3.7	MODEL COMPARISON WITH PREVIOUS MODELS.....	103
3.8	DISCUSSION OF RESULTS.....	104
3.8.1	<i>Model application in Section 3.6 is the same as model development in section 3.5</i>	111
3.8.2	<i>Validity of FMM results</i>	113
3.9	CONCLUSIONS.....	115
ESSAY 2: MODELLING VARIATION IN LIFETIMES: A SHARED FRAILTY APPROACH		117
4.1	INTRODUCTION.....	117
4.2	RESEARCH QUESTION.....	117
4.3	CONSEQUENCES AND CAUSES OF DEPENDENCE.....	118

4.4	RESEARCH METHODOLOGY	119
4.4.1	<i>Weibull mixtures and Weibull proportional hazard Model</i>	120
4.4.2	<i>Frailty description</i>	122
4.4.3	<i>Frailty model</i>	122
4.4.4	<i>Gamma distribution for frailty</i>	122
4.4.5	<i>Clustered failures: the shared frailty approach</i>	123
4.4.6	<i>Estimation</i>	123
4.5	MODEL IMPLEMENTATION	123
4.5.1	<i>Data and variables for the shared frailty model</i>	123
4.5.2	<i>Clustering via mixture of Weibull proportional hazards model</i>	124
	Table 4.3: Cluster characteristics	Error! Bookmark not defined.
4.5.3	<i>Shared frailty model implementation</i>	128
4.5.3.1	<i>Steps for data pre-processing</i>	128
4.5.3.2	<i>Implementation</i>	128
4.5.3.3	<i>Comparison with previous models</i>	132
4.6	CONCLUSIONS	132

ESSAY 3: MODELLING VARIABILITY IN TIME BETWEEN CLAIMS: EXPLORING DEALER'S PERSPECTIVE IN ACHIEVING THE RIGHT FIRST TIME OBJECTIVE..... 134

5.1	INTRODUCTION	134
5.2	IMPORTANCE OF FIXING IT RIGHT FIRST TIME	136
5.3	ROLE OF DEALERS IN AUTOMOTIVE PROCESS	137
5.3.1	<i>Understanding the manufacturer-dealer relationship</i>	137
5.3.2	<i>Dealers and repeat failures: RFT perspective</i>	137
5.4	RESEARCH QUESTIONS	139
5.5	METHODOLOGY	140
5.5.1	<i>Risk intervals</i>	141
5.5.2	<i>Risk sets</i>	141
	143
5.5.3	<i>Common and event-specific parameter estimates</i>	143
5.5.4	RECURRENT EVENT MODELS	143
5.5.4.1	<i>Andersen-Gill (A-G) model</i>	144
5.5.4.2	<i>Prentice, Williams and Peterson model</i>	145
5.5.4.3	<i>Wei, Lin, and Weissfeld model</i>	145
5.6	EMPIRICAL BAYES ANALYSIS	146
5.7	DATA AND VARIABLES	148
5.7.1	<i>Steps in data pre-processing</i>	148
5.7.2	<i>Importance of variables included in the analysis</i>	149
5.7.3	<i>Results</i>	149
5.7.3.1	<i>Model comparison</i>	149
5.7.4	<i>Interpretation of the results</i>	153

5.7.5 Empirical Bayes estimates	156
5.8 CONCLUSIONS	159
6. MANAGERIAL IMPLICATIONS	161
APPENDIX A.....	171
DETAILS OF EM ALGORITHM (SECTION 3.5.1, CHAPTER 3).....	171
APPENDIX A.1	174
APPENDIX A.2	175
APPENDIX A.3	176
APPENDIX B.....	177
B.1 DETAILS OF EM ALGORITHM (SECTION 4.4.1, CHAPTER 4).....	177
EM estimation of parameters of WPHM	177
APPENDIX B.2	179
APPENDIX C.....	181
TABLE C.2: SHARED FRAILTY MODEL SUMMARY WITH AND WITHOUT COVARIATES	182
APPENDIX D.....	183
D.1 THE LAPLACE TRANSFORM (CHAPTER 5).....	183
D.2 LAPLACE TRANSFORM OF FRAILTY DISTRIBUTION OF SURVIVORS	183
D.3 DERIVATION OF EMPIRICAL BAYES FRAILTY ESTIMATES BASED ON LAPLACETRANSFORM	185
REFERENCES	187

LIST OF TABLES

3.1 Detailed variable description used in cluster analysis	94
3.2 Cluster centroids	98
3.3 Variable description (chapter 3)	102
3.4 Components with their AIC, BIC and log-likelihood values.....	104
3.5 Model comparisons (chapter 3)	104
3.6 Component-wise results.....	105
4.1 Variable description (chapter 4).....	125
4.2 BIC over different number of components.....	125
4.3 Cluster characteristics.....	127
4.4 Correlation for different basis of frailty.....	129
4.5 Model fit.....	129
4.6 Frailty model	129
5.1 Quality costs.....	136
5.2 Differences in the three risk intervals.....	141
5.3 Differences in the three types of risk sets	142
5.4 Data description	150

LIST OF TABLES

5.5 Cox proportional hazards (1), Andersen Gill (2) and Prentice Williams and Peterson (3) and Wei, Lin, and Weissfeld model (WLW) (4) models built on the dataset.....	151
5.6 Model comparison (essay 3)	152
5.7 Dealer-wise estimated frailty values	158

LIST OF FIGURES

Figure 1.1(a): Claims paid per quarter (in millions of dollars, 2016-2021) by U.S. based auto manufacturers.....	5
Figure 1.1(b): Warranty expenses (in millions of rupees, 2018-2020) by Indian-based auto manufacturers (TVS Motors, Bajaj Auto, Hero Motors and Eicher Motors)	6
Figure 1.2: Framework of the thesis.....	9
Figure 2.1: Flow of the literature review.....	36
Figure 2.2: Proposed hybrid approach	54
Figure 3.1: Proposed methodology for essay one.....	92
Figure 3.2: Stopping rule: percentage change in heterogeneity.....	96
Figure 3.3: Silhouette width.....	97
Figure 4.1: Cluster profile.....	127
Figure 5.1: Product selling structure.....	138
Figure 5.2: Different types of risk intervals.....	142
Figure 5.3: Empirical Bayes estimates of frailties.....	159

ACRONYMS

$i=1,2,\dots,N$	Index for the individuals
$l=1,2,\dots,L$	Index for the explanatory variables
$s=1,2,\dots,S$	Index for the latent classes
n_i	Number of claims for individual i in time period t_i
x_{il}	Value of the l^{th} explanatory variable for the i^{th} individual
b	Scale parameter
γ	Shape parameter
p	Failure mode
K	Number of latent segments
t_p	Duration time on category p
π_k	Mixing proportions corresponding to K latent segments
β	Regression parameters
X	Vector of covariates
h_{ij}	Hazard of the j^{th} individual from cluster i
Y	Matrix of the indicator variables
λ	Mean event rate

Abstract

Warranty management is a critical decision for consumer durables, automobiles and capital equipment manufacturers. Original equipment manufacturers (OEMs) must decide on warranty duration, exclusions on warranty and extended warranty policy. These decisions can have a significant impact on profitability. Many companies encounter an unexplained gap between warranty accruals and actual warranty costs. One of the reasons behind this is the ‘unexplained variation’, which if accounted for, will provide a valuable understanding of warranty claims and improve warranty prediction.

The unexplained variability in warranty claims can occur due to various factors. One of the most important causes of the variation are the usage and the operating conditions. There is extensive literature on known causes of variability such as usage and risk attitudes (i.e., whether a customer is risk averse, risk neutral or risk loving). Factors such as the operating conditions are difficult to incorporate in the model building process since they are unknown. Such hidden causes either need to be discovered or modelled appropriately for better warranty management. Therefore, through this thesis, we have first tried to uncover the latent factors that impact warranty claims and the survival of the components. Second, we have proposed wholistic modelling approaches to incorporate those factors into the model-building process.

We have divided our work into three different essays. The first essay deals with modelling variability in warranty claim counts itself. Although neglected in the literature, the variability in warranty claim counts can provide useful information about the underlying hidden components or the subpopulations. This essay presents a hybrid approach consisting of factor identification and model development and employs an integrated framework to better model warranty claims. In this essay, we group the claim population into clusters or components, which behave similarly with respect to location, climate and urbanization. Then, a latent class Poisson regression methodology is proposed. The cluster characteristics thus obtained are entered into the modelling framework using the finite mixture modelling approach, which gives a better fit than the existing models, thereby improving prediction.

ABSTRACT

In the second essay, we try to uncover the latent segments in the population that share dependence effects, owing to the unobserved risk factors. This would help the manufacturer identify segments in the population that are more likely to result in claims or are more ‘frail’ than the others. This is achieved through a two-step process. First, the hidden components have been obtained via clustering of survival data using the Weibull mixture model approach. Such an approach will not only serve as representative for unobservable covariates via cluster membership but also provide a way by which the product units can be classified into different subpopulations prior to statistical lifetime modelling. The second step of the approach incorporates cluster membership into the shared frailty model framework. Thus, this essay proposes a novel approach to obtain those segments in the population which share similar risks owing to the unobserved factors and are highly correlated.

The third essay studies the impact of dealer characteristics on time between claims and tries to identify the role of dealers in achieving the ‘right first time’ objective. In addition, this essay also provides a way to rank the dealers based on their unobserved characteristics using empirical Bayes methods and consequently enables best practices to be followed across the dealers. To the best of my knowledge, this is the first study examining dealer’s role on time between claims and provides a method to compare their relative performance.

The contribution of this thesis is to provide the manufacturer with a new perspective to look into different causes of variability apart from previously considered causes, i.e., usage and risk attitudes. Further, it provides a means to incorporate those causes in the model-building process for a better warranty management.

Keywords: Finite mixture modelling, sub-populations, warranty claims, variation, heterogeneity.