NAVAL POSTGRADUATE SCHOOL
Monterey, California

DISSERTATION

AGE REPLACEMENT POLICIES IN MULTIPLE TIME SCALES

by

Scott G. Frickenstein

June 2000

Dissertation Supervisor: Lyn R. Whitaker

Approved for public release; distribution is unlimited.

AGE REPLACEMENT POLICIES IN MULTIPLE TIME SCALES

Scott G. Frickenstein – Captain, United States Air Force
B.S., United States Air Force Academy, 1990
M.S., Florida State University, 1991
Advisor: Lyn R. Whitaker, Department of Operations Research
Committee Members: Robert R. Read, Department of Operations Research
Gerald G. Brown, Department of Operations Research
Samuel E. Buttrey, Department of Operations Research
Robert A. Koyak, Department of Operations Research
Craig W. Rasmussen, Department of Mathematics

We develop and estimate optimal age replacement policies for devices whose age is measured in multiple time scales. For example, the age of a jet engine can be measured in chronological time, the number of flight hours, and the number of landings. Under a single-scale age replacement policy, a device is replaced at age τ or upon failure, whichever occurs first. We show that a natural generalization to $k \geq 2$ scales is to replace non-failed devices when their usage path crosses the boundary of a k-dimensional region M, where M is a lower set with respect to the matrix partial order. For lifetimes measured in two scales, we consider two contexts. In the first, devices age along linear usage paths. For this case, we generalize the single-scale long-run average cost and estimate optimal two-scale policies. We show these policies are strongly consistent estimators of the true optimal policies under mild conditions, and study small-sample behavior using simulation. For the second context, in which device usage paths are unknown, we use two-dimensional renewal theory to derive the long-run average cost of a policy. We give examples in both settings and note that these ideas generalize to more than two scales.

DoD KEY TECHNOLOGY AREA: Other (Reliability)

KEYWORDS: Age replacement, Multiple time scales, Renewal theory

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	28.Jun.00	DISSERTATION

4. TITLE AND SUBTITLE
AGE REPLACEMENT POLICIES IN MULTIPLE TIME SCALES

5. FUNDING NUMBERS

6. AUTHOR(S)
CAPT FRICKENSTEIN SCOTT G

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NAVAL POSTGRADUATE SCHOOL

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
FY00-220

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

14. SUBJECT TERMS

15. NUMBER OF PAGES
125

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

Standard Form 298 (Rev. 2-89) (EG)
Prescribed by ANSI Std. 239.18
Designed using Perform Pro, WHS/DIOR, Oct 94

AGE REPLACEMENT POLICIES IN MULTIPLE TIME SCALES

Scott G. Frickenstein – Captain, United States Air Force
B.S., United States Air Force Academy, 1990
M.S., Florida State University, 1991
Advisor: Lyn R. Whitaker, Department of Operations Research
Committee Members: Robert R. Read, Department of Operations Research
Gerald G. Brown, Department of Operations Research
Samuel E. Buttrey, Department of Operations Research
Robert A. Koyak, Department of Operations Research
Craig W. Rasmussen, Department of Mathematics

We develop and estimate optimal age replacement policies for devices whose age is measured in multiple time scales. For example, the age of a jet engine can be measured in chronological time, the number of flight hours, and the number of landings. Under a single-scale age replacement policy, a device is replaced at age τ or upon failure, whichever occurs first. We show that a natural generalization to $k \geq 2$ scales is to replace non-failed devices when their usage path crosses the boundary of a k-dimensional region M, where M is a lower set with respect to the matrix partial order. For lifetimes measured in two scales, we consider two contexts. In the first, devices age along linear usage paths. For this case, we generalize the single-scale long-run average cost and estimate optimal two-scale policies. We show these policies are strongly consistent estimators of the true optimal policies under mild conditions, and study small-sample behavior using simulation. For the second context, in which device usage paths are unknown, we use two-dimensional renewal theory to derive the long-run average cost of a policy. We give examples in both settings and note that these ideas generalize to more than two scales.

DoD KEY TECHNOLOGY AREA: Other (Reliability)

KEYWORDS: Age replacement, Multiple time scales, Renewal theory

NAVAL POSTGRADUATE SCHOOL
Monterey, California

DISSERTATION

AGE REPLACEMENT POLICIES
IN MULTIPLE TIME SCALES

by

Scott G. Frickenstein

June 2000

Dissertation Supervisor: Lyn R. Whitaker

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2000	**3. REPORT TYPE AND DATES COVERED** Doctoral Dissertation	
4. TITLE AND SUBTITLE: Title (Mix case letters) Age Replacement Policies in Multiple Time Scales			**5. FUNDING NUMBERS**
6. AUTHOR(S) Frickenstein, Scott G.			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			**12b. DISTRIBUTION CODE**

13. ABSTRACT *(maximum 200 words)* We develop and estimate optimal age replacement policies for devices whose age is measured in multiple time scales. For example, the age of a jet engine can be measured in chronological time, the number of flight hours, and the number of landings. Under a single-scale age replacement policy, a device is replaced at age τ or upon failure, whichever occurs first. We show that a natural generalization to $k \geq 2$ scales is to replace non-failed devices when their usage path crosses the boundary of a k-dimensional region M, where M is a lower set with respect to the matrix partial order. For lifetimes measured in two scales, we consider two contexts. In the first, devices age along linear usage paths. For this case, we generalize the single-scale long-run average cost and estimate optimal two-scale policies. We show these policies are strongly consistent estimators of the true optimal policies under mild conditions, and study small-sample behavior using simulation. For the second context, in which device usage paths are unknown, we use two-dimensional renewal theory to derive the long-run average cost of a policy. We give examples in both settings and note that these ideas generalize to more than two scales.

14. SUBJECT TERMS Age replacement, multiple time scales, renewal theory			**15. NUMBER OF PAGES** 125
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

AGE REPLACEMENT POLICIES IN MULTIPLE TIME SCALES

Scott G. Frickenstein
Captain, United States Air Force
B.S., United States Air Force Academy, 1990
M.S., Florida State University, 1991

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2000

Author:
Scott Frickenstein

Approved by:

Lyn R. Whitaker
Associate Professor of
Operations Research
Dissertation Supervisor

Robert R. Read
Professor of Operations Research
Dissertation Committee Chair

Gerald G. Brown
Professor of Operations Research

Samuel E. Buttrey
Assistant Professor of
Operations Research

Robert A. Koyak
Assistant Professor of
Operations Research

Craig W. Rasmussen
Associate Professor of Mathematics

Approved by:
Richard E. Rosenthal, Chair, Department of Operations Research

Approved by:
Anthony P. Ciavarelli, Associate Provost for Instruction

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

We develop and estimate optimal age replacement policies for devices whose age is measured in multiple time scales. For example, the age of a jet engine can be measured in chronological time, the number of flight hours, and the number of landings. Under a single-scale age replacement policy, a device is replaced at age τ or upon failure, whichever occurs first. We show that a natural generalization to $k \geq 2$ scales is to replace non-failed devices when their usage path crosses the boundary of a k-dimensional region M, where M is a lower set with respect to the matrix partial order. For lifetimes measured in two scales, we consider two contexts. In the first, devices age along linear usage paths. For this case, we generalize the single-scale long-run average cost and estimate optimal two-scale policies. We show these policies are strongly consistent estimators of the true optimal policies under mild conditions, and study small-sample behavior using simulation. For the second context, in which device usage paths are unknown, we use two-dimensional renewal theory to derive the long-run average cost of a policy. We give examples in both settings and note that these ideas generalize to more than two scales.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

"Give thanks to the LORD, call on His name; make known among the nations what He has done." (2 Chronicles 16:8) Father God, I echo the words of King David and thank you for your faithfulness to me during this time. You have truly enabled me, provided for me, and protected me as you promised you would. "With your help I can advance against a troop; with my God I can scale a wall." (Psalm 18:29)

Kristi, Alex, and Jared, I appreciate your tremendous support and understanding. I thank God for you and love each of you with all my heart. Alex Filuk, thank you for your confidence in me and for taking a few moments to encourage me just hours before our Heavenly Father called you home. David Bretz, thank you for sharing your life with me as we covered the miles together.

Solomon, deemed to be the wisest man in the world, wrote "Of making many books there is no end, and much study wearies the body." (Ecclesiastes 12:12) In full knowledge of this fact, my mentors, Bob Reehm and Chaplain Phillips, provided much-needed "staying power," fervent prayers, gentle rebukes, and constant encouragement.

I am not sure where to begin, and I doubt I can adequately acknowledge all Professor Lyn Whitaker has done to make this happen. Were it not for her availability, empathy, honesty, and insight, this dissertation would still be a collection of jumbled thoughts. I am thankful for Professor Bob Read, who took me under his wing during my first 18 months and helped me wade through "out-of-hide" courses. I appreciate the many ways in which the faculty of the Operations Research department, under the leadership of Professor Rick Rosenthal, has invested in me over the last three years.

Professors Ilya Gertsbakh of Ben Gurion University of the Negev and Thierry Duchesne of the University of Toronto have also been extremely helpful. I greatly appreciate their gracious and immediate responses for clarification on their writings and for helping me understand concepts related to multiple time scales.

I am thankful for the support of my fellow "strugglers," especially Lieutenant Colonel Kirk Yost, my "point man" in what has often been a jungle. I also acknowledge the many ways I have been encouraged by Eitan Isreali, Mehmet Ayik, Gustavo Martins, Izumi Kobayashi, and Tom Cioppa during my tour of duty in Glasgow Hall. May our paths cross again soon.

I am also grateful for the constant logistical support I have had from Alan Jones, Alan Pires, Ronald del Presto, Tom Halwachs, and Debbi Kreider. Terry Bilodeau helped make this a document worth reading—her skills and graciousness are amazing. The staff of the Dudley Knox Library has been a great help too; keep up the good work!

Finally, I would like to thank Colonel Daniel Litwhiler for this opportunity to better prepare myself to serve future generations of Air Force leaders.

I. INTRODUCTION

In practice, the age of a device is often measured in more than one time scale. For example, automobiles age in the "parallel" scales of calendar time since purchase and number of miles driven. As such, routine engine maintenance depends on both of these scales: an oil change is recommended every three months or 3,000 miles, whichever comes first. For some devices, the scale most relevant for maintenance is clear. For example, Kordonsky and Gertsbakh (1993) note that for a jet engine turbine, the duration of the warmup period is the most relevant (of several possible scales) but for the undercarriage of an aircraft, the number of landings is most relevant. For other devices, however, the most relevant scale for maintenance is difficult to determine. For example, the joint between an aircraft wing and the fuselage is subjected simultaneously to corrosion (thus the scale "calendar time" is relevant), landing stresses (thus number of landings is relevant), and level flight stresses (thus total flight time is relevant), as noted by Kordonsky and Gertsbakh (1993). In any case, a maintenance policy should take into account the parallel scales in which an item operates. In a military setting, attempts are made to model the effect of chronological or operational time on the failure characteristics of a military device during the developmental testing phase. During this phase, however, it may be difficult or impossible to accurately model the effect of usage on the device resulting from military missions. Thus, classical failure models are used to develop single-scale maintenance policies, even though it is well known that the device will operate in the parallel scales of chronological (or operational) time and number of missions. Lifetime data including the total number of missions (e.g., landings) accrued at the time of device failure may become available later in the acquisition cycle, such as during operational testing or upon initial fielding. Military maintenance costs should be reduced by using policies that directly account for aging in multiple scales. In this dissertation we focus on developing, optimizing, and estimating maintenance policies, in particular age replacement policies, based on multiple time scales.

A. SINGLE-SCALE AGE REPLACEMENT POLICIES

The vast majority of methods for developing maintenance policies are based on a single time scale; see McCall (1965), Pierskalla and Voelker (1976), and Valdez-Flores and Feldman (1989) for comprehensive reviews. Among the most useful and most studied are age replacement policies, under which a device is replaced (or overhauled) at failure or at a predetermined age $\tau > 0$, whichever occurs first. Let X be a positive random variable (r.v.) representing the lifetime of a device, i.e., the time when the device fails. Let X have distribution function F; following Bather (1977) it will be convenient to define $F(x) = P(X < x)$ and the survivor function as $S(x) = P(X \geq x)$. Thus, under an age replacement policy, a device is replaced with a new one at time $\min\{X, \tau\}$. Let the cost for replacement be $K > 0$ if the device is replaced due to age (i.e., preventively, since $X \geq \tau$) and $K + C$ if it is replaced due to failure (i.e., $X < \tau$), where the additional cost of replacement at failure is $C > 0$. If devices have independent lifetimes, then replacement times occur according to a renewal process. From the Renewal Reward Theorem (e.g., Ross, 1997), the long-run average cost per unit of time that the device is in use is

$$C(\tau) = \frac{K + C\,F(\tau)}{\int_0^\tau S(u)du}, \ \tau > 0. \tag{1.1}$$

A complete derivation of this expression can be found in Appendix A. If F is absolutely continuous and has an increasing failure rate (IFR), then $C(\tau)$ has at most one minimum. In addition, if the failure rate is continuous and strictly increasing to ∞, there exists a unique and finite value τ^* minimizing $C(\tau)$ (e.g., Barlow & Proschan, 1965). Bergman (1982) shows that a unique, finite τ^* is attained under slightly less restrictive conditions.

When F is completely specified, τ^* can be found explicitly, but is more often found with numerical methods. Glasser (1967) uses numerical methods to obtain charts which can be used to find τ^* when F is truncated normal, gamma, or Weibull. When F is unknown, there are numerous approaches available for estimating τ^* based upon lifetime data. In most of these approaches, F in equation (1.1) is replaced with an estimator $\hat{F}$ based upon the data. This results in an estimator $\hat{C}(\tau)$ of the cost function $C(\tau)$; τ^* is then estimated by minimizing $\hat{C}(\tau)$. For example, given a simple random sample $X_1, \ldots, X_n$, of lifetimes from F, non-parametric estimators of $C(\tau)$ and τ^* can be found using the empirical survivor function

$$\hat{S}(\tau) = \sum_{j=1}^{n} I(X_j \geq \tau)/n, \tag{1.2}$$

where $I(X_j \geq \tau) = 1$ if $X_j \geq \tau$ and 0 otherwise. It follows that the estimator of $C(\tau)$ is

$$\hat{C}(\tau) = \frac{(K + C) - C\,\hat{S}(\tau)}{\int_0^\tau \hat{S}(u)du}, \ \tau > 0. \tag{1.3}$$

From the definition of $\hat{S}(\tau)$, it is seen that $\hat{C}(\tau)$ is lower semi-continuous with denominator strictly increasing on $(0,\infty)$ and numerator a lower semi-continuous step function constant between observations. As a result, local minima of $\hat{C}(\tau)$ are found at the observations and we define $\hat{\tau} = \text{argmin}\, \hat{C}(X_j)$. Also, $\hat{\tau}$ is not necessarily unique. Arunkumar (1972) proves that $\hat{C}(\tau)$ and $\hat{\tau}$ are strongly consistent estimators of $C(\tau)$ and τ^*, respectively. Ingram and Scheaffer (1976) address estimation using the non-parametric maximum likelihood estimator (MLE) of F under the restriction of F having an increasing failure rate. The optimal policy τ^* can also be estimated under other sampling schemes; for example, Kumar and Westberg (1997) estimate τ^* under right-censoring, and Bather (1977), Frees and Ruppert (1985), and Aras and Whitaker (1992) address sequential estimation of τ^*. Graphical approaches can also be used to minimize (1.1) and (1.3). Bergman (1977) uses the total time on test (TTT) plotting method of Barlow and Campo (1975) to estimate τ^*. This method is insightful since one can deduce ranges of the ratio K/C for which a particular τ^* is optimal. Two comprehensive treatments of this approach are contained in Bergman and Klefsjö (1982) and Klefsjö (1986).

B. FAILURE MODELING IN MULTIPLE TIME SCALES

Extending this theory so it can be used for maintenance of a device whose age is measured in multiple scales requires more than generalizing a univariate lifetime X to a multivariate lifetime, say (X, Y). This is not always immediately apparent. Confusion arises because data used to estimate multiple-scale policies often appear to be of the form $(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)$. Nonetheless, the actual implementation of an age replacement policy requires that a device be tracked continuously through time. Even in a single scale, a policy cannot be implemented by observing the age at failure; the device is monitored through time so that it can be replaced at failure or time τ, whichever comes first. The implementation of such a policy in more than one scale requires knowledge of the usage path, or "history" of the device; this notion is central to the literature of multiple time scales (e.g., Duchesne and Lawless (2000)). Let $x \geq 0$ denote the chronological time since introduction of a device into service, and let $y(x)$ represent usage accumulated by the device up to age x (e.g., the total number of miles an automobile has been driven up to age x). The *usage path* of a device up to chronological time x is defined to be $Z(x) = \{(u,y(u)): 0 \leq u \leq x\}$. In addition, if the random variable X represents the chronological age of the device at failure and $Y = y(X)$, then (X,Y) represents the time and cumulative usage at failure. In some cases a vector $\mathbf{y}(x)$ of various measures of usage is available (e.g., $y_1(x)$ could be the number of flight hours accrued as of chronological time x, and $y_2(x)$ could be the number of landings accrued as of chronological time x, etc.). Then, the usage path is $Z(x) = \{(u,\mathbf{y}(u)): 0 \leq u \leq x\}$. In most of what follows, however, we assume only a single measure of usage is available in order to simplify the presentation. Typically, a measure of usage is required to be both non-decreasing in x and an external covariate. The latter requirement (see Kalbfleisch and Prentice, 1980, Section 5.3) ensures the usage path Z is determined independently of the time to failure X.

Modeling the lifetime of a device whose failure depends upon the parallel effects of time and usage has received a great deal of attention in the past decade. Three main approaches are found in the literature. The first approach is to use a conditional model. Lawless et al (1995) model automobile warranty data by considering separately the distribution of X along each path Z and the distribution of the paths. The second is to use a joint model for failure times. This approach is taken by Singpurwalla and Wilson (1998), Murthy et al (1995), and Kordonsky and Gertsbakh (1994). Models built using this approach do not rely explicitly on the notion of a usage path. Due to the inherent complexity of explicitly modeling lifetimes and paths in multiple scales, much of the recent work in this area focuses on a third approach, that of finding appropriate methods for combining scales to form a single scale. When such a combined scale can be found, standard univariate reliability tools (including age replacement theory) can be brought to bear. Duchesne and Lawless (2000) unify and formalize all previous work in combining scales.

C. MAINTENANCE IN MULTIPLE TIME SCALES

Much less attention has been given to maintenance policies based on multiple scales. In the earliest work in this area, Nakagawa (1985) derives policies for devices that fail by either age or usage. He derives the expected cost rate $C(\tau,N)$ of the policy under which a device is replaced at failure, at chronological age τ, or at a discrete number N uses, whichever occurs first. In our setting, however, it is rarely evident whether failure occurred due to age or usage. In addition, since it is common to have both age and usage continuous (e.g., scales might be chronological time since production and total flight time), we need models that allow usage to be continuous as well as discrete. Unlike Nakagawa (1985), most recent work focuses on finding an appropriate combined scale to be used for preventive maintenance. With this approach, the cost of age replacement can then be computed in the combined scale, and, under appropriate conditions, an optimal replacement age can also be found in that scale. The major work in this area is done by Kordonsky and Gertsbakh (1994) and along slightly different lines by Kordonsky and Gertsbakh (1993, 1995, 1997). They restrict attention to linear combined scales $t(a) = (1-a)x + ay(x)$, where $a \in [0,1]$. Under an age replacement policy in such a scale, a device is replaced at age τ (in the combined scale) or upon failure at age $T(a) = (1-a)X + aY$, whichever occurs first. Most recently, Duchesne and Lawless (2000) propose an "ideal" time scale which generalizes some of the work of Kordonsky and Gertsbakh. Although not motivated specifically with preventive maintenance in mind, they suggest that their scale might be used for such purposes. The ideal scale is developed in order to capture chronological age and usage in such a way that, under appropriate conditions, the lifetime distribution of a device in this scale is independent of the path. Thus, in principle, an age replacement policy based on an ideal time scale could be used for devices regardless of their usage path.

Because using combined scales reduces the problem of maintenance in multiple scales to that of maintenance in a single scale, it has the advantage of being tractable and easily understood. Combined scales, however, do not completely address the problems of maintenance in multiple scales. Absent from the literature is discussion of the translation of policies developed in combined scales to policies in the original scales. Upon performing such a translation, it is clear that policies based on linear scales correspond to replacing devices if their joint failure time (X,Y) falls in the region $M = \{(x,y(x)): (1-a)x + ay(x) < \tau\}$ or when their usage curve crosses the boundary of this region, whichever occurs first. Similarly, policies based on an ideal time scale correspond to regions in the positive quadrant whose upper boundaries follow the contours of the ideal time scale. Considering such regions in the original scales suggests a more general class of policies that should be considered when searching for the optimal policy. Also absent from the literature are methods for comparing the cost of policies based on combined scales of different forms. The approach of Kordonsky and Gertsbakh (1994) does provide a means for comparing costs in the special case of the family of linear scales. As such, the need arises for a means to compare the cost of policies from a larger class of alternatives.

In this dissertation we directly attack the problem of estimating optimal age replacement policies for devices with age measured in multiple scales in two different settings. In both, our focus is to search over a large class of sensible policies to minimize estimated long-run costs. To do so, we first define a class of multiple-scale policies which generalize policies found in previous works. In Chapter II, we use several real data sets to help develop insight into our choice of this class of potential policies. Because this class of policies is related to policies produced by combined scales, in Chapter III we review and discuss in detail how multiple-scale policies are obtained using the scale-combining approaches found in the literature. In this chapter we also discuss how these policies fit into the framework established in Chapter II. In so doing, we raise significant concerns that reveal the need for new methods. Since usage paths are often well-approximated by straight lines, in Chapter IV we develop estimators of the cost function and optimal policy for the case in which devices age along linear usage paths. In Chapter V we discuss the large- and small-sample properties of these estimators and compare their performances with policies based on a common scale-combining approach. In Chapter VI we develop a cost function for policies under a joint model for (X,Y) and present numerical results obtained from solving the corresponding optimization problem for rectangular-shaped policies. In Chapter VII we highlight our contributions and present opportunities for further research.

THIS PAGE INTENTIONALLY LEFT BLANK

II. EXTENDING AGE REPLACEMENT THEORY TO MULTIPLE TIME SCALES

We seek to generalize the classical age replacement policy, under which a device is replaced at age τ or failure (whichever occurs first), to a policy based on age measured in multiple scales. The cost function used to define an optimal policy is based on the mechanism generating the failures. However, the general form of a sensible multiple-scale age replacement policy applies equally to many failure models. In this chapter, we introduce three data sets to help develop insight into an appropriate form for a multiple-scale age replacement policy. The data sets are chosen to represent situations for which either the conditional modeling approach or the joint modeling approach may be appropriate. In the first and third data sets, it is apparent that failures occur along fixed linear usage paths. In such a situation, an appropriate model is one which generates failures conditioned on the usage path and then utilizes a mixing distribution over the paths. However, in the second data set, there are no clear usage paths and the data are better modeled by a joint distribution. After considering the three data sets, we generalize the form of an age replacement policy to incorporate multiple time scales.

A. INTRODUCTORY CASE STUDIES

Under a single-scale policy with replacement time τ, a device is replaced if it fails in the interval $(0,\tau)$ or if its time in use (the one-dimensional equivalent of a usage path) crosses the right-most boundary of $(0,\tau)$. As we generalize to the case of multiple scales, it will be convenient to identify a policy by the multiple-scale equivalent of the failure replacement interval $(0, \tau)$. This leads to consideration of policies defined by regions M. Here, a device is replaced if (X,Y) is in M (i.e., upon failure) or when its usage path crosses the boundary of M, whichever occurs first. For now, we consider how such policies might be constructed based on observed bivariate failure times $(x_1,y_1),\ldots,(x_n,y_n)$. In what follows we use the notation $R(x,y)$ to denote the rectangle $(0,x) \times (0,y)$.

Case Study 1

Consider policy $M_X = R(\hat{\tau},\infty)$, where $\hat{\tau}$ minimizes the empirical cost function (1.3) based on the first components $x_1, \ldots, x_n$. Under this policy, we replace the device when its age reaches $\hat{\tau}$ or fails, whichever occurs first, regardless of the usage accrued. Although constructed in a rather naïve manner, such a policy may be adequate in some cases.

For example, consider the locomotive traction motor failure data in Singpurwalla and Wilson (1998). The data (see Appendix B) consists of the time since inception of service and mileage at failure of forty locomotive traction motors. Figure 2.1 shows a scatterplot of the failure data in the time scales number of days and number of miles and the regression fit through the origin. The coefficient of determination exceeds 99%. For these data, knowing the number of days at failure is almost equivalent to knowing the number of miles at failure since all exemplars have virtually identical usage rates (i.e., number of miles per day). Hence a "naïve" policy based solely on chronological age suffices. Similarly, we could consider a mileage-based policy $M_Y = R(\infty, \hat{\nu})$ where $\hat{\nu}$

minimizes (1.3) based on $y_1, \ldots, y_n$. In fact, for ratios $K/C \geq 0.25$, the two regions $M_X = R(\infty,1200)$ and $M_Y = R(57304,\infty)$ are based on the same observation, namely (1200,57304).

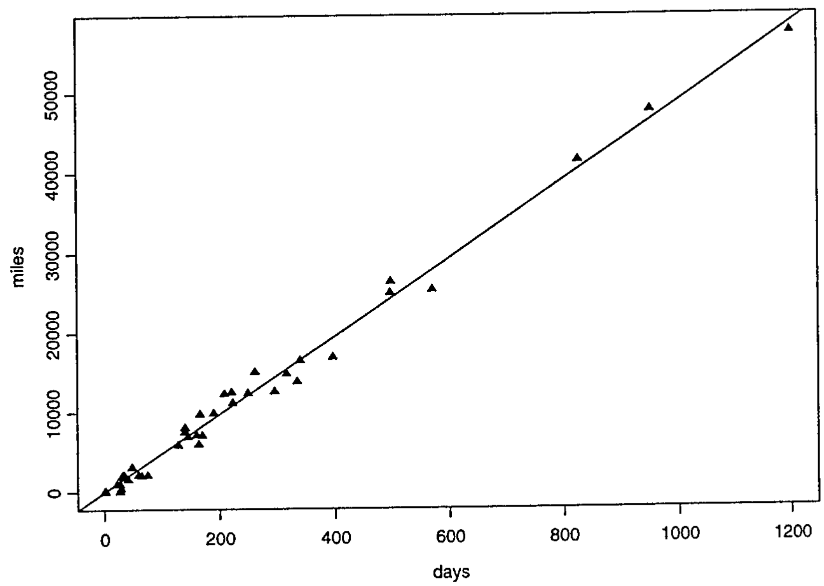

Figure 2.1: Traction Motor Data with Regression Line. Triangles represent the number of days and miles until a failure occurred.

Case Study 2

A policy based on a single scale may not be satisfactory for lifetime data arising from devices having differing usage paths. Figure 2.2 shows a scatterplot of failure times of jet engines, discussed by Gertsbakh and Kordonsky (1998). This data set (see Appendix B) contains the flight hours and number of landings at failure of 21 Aeroflot jet engines. Unlike the first data set, the failures have occurred along several usage paths, and these paths are not provided or evident from Figure 2.2. Thus, knowing the number of hours at failure is not equivalent to knowing the number of landings at failure. Hence, a policy based on only the flight hours at failure or only the number of landings at failure is likely to ignore information that could potentially reduce maintenance costs. In fact, for $K/C = 0.5$, $M_X = \mathrm{R}(4932,\infty)$ and $M_Y = \mathrm{R}(\infty,1152)$; these two policies (with boundaries delimited by the overlaid dashed lines in Figure 2.2) are based on the vastly "different" observations (4932, 1960) and (3227,1152), respectively.

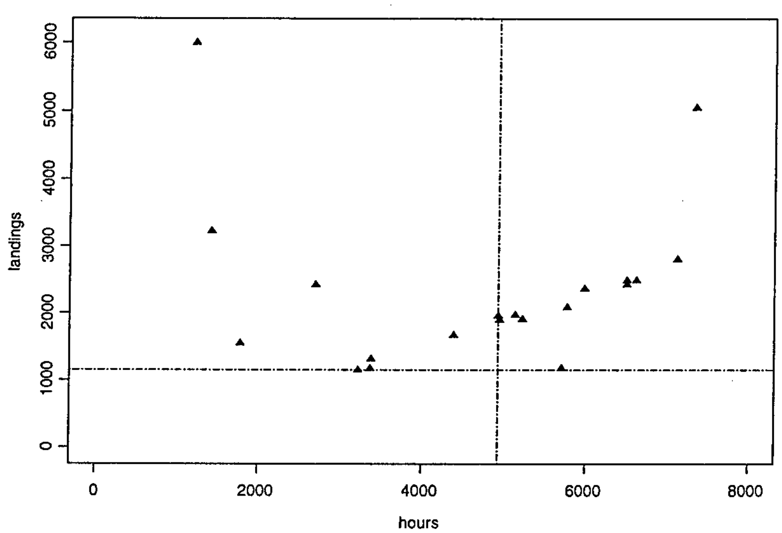

Figure 2.2: Jet Engine Data.
Triangles represent the number of flight hours and landings until a failure occurred. The vertical and horizontal lines represent the boundaries of, respectively, a policy triggered solely by the number of flight hours at failure and the policy triggered solely by the number of landings at failure.

Such policies, however, are often used. Gertsbakh and Kordonsky (1997) note that a single distribution is often fit to lifetime data arising from devices operating in heterogeneous environments. An "optimal" policy is estimated from this distribution and applied to the entire population. Policies of this form ignore the bivariate nature of the failure data. For example, under policy M_X, devices with lifetimes (x,y) and $(x,2y)$ are treated in the same manner, even though the latter device is "older" in some sense than the former. A policy which somehow incorporates the additional information contained in the paired failure times seems "better" than M_X. Consider the policy $M_{XY} = \mathrm{R}(\hat{\tau}, \hat{\nu})$, formed by combining M_X and M_Y. Under policy M_{XY} we replace a non-failed device when it accrues either age $\hat{\tau}$ or usage $\hat{\nu}$, whichever occurs first; $\hat{\tau}$ and $\hat{\nu}$ are estimates of the optimal replacement times in the two single-scale age replacement problems. Policy M_{XY} seems to be an improvement over both M_X and M_Y, since it is based on all the data and since in some cases (for example) devices with lifetimes (x,y) and $(x,2y)$ are treated differently. Nevertheless, the separate computation of $\hat{\tau}$ and $\hat{\nu}$ ignores the dependency between the failure times in the two scales. Policy M_{XY} is based only on estimates of the marginal distributions of failure time in the two scales, and thus does not fully account for the joint effect of age and usage on failure. A bivariate policy should somehow account for this dependence. Kordonsky and Gertsbakh (1995) explain, "Each particular time scale reflects indirectly a most relevant process of damage accumulation, but fails to reflect the joint, interactive action of these processes. For an aircraft ... 'time in the air' and 'number of flights' both reflect fatigue damage accumulation, but each scale separately is not able to reflect 'total' fatigue damage." In Chapter VI, we develop a cost function that can be used to find the "best" policy of the form M_{XY}.

Case Study 3

Consider failures due to metal fatigue, (see Appendix B) discussed in Kordonsky and Gertsbakh (1993). The metal fatigue data plotted in Figure 2.3 consists of 30 observations, five on each of six distinct paths. Specimens on a particular path are subjected to bending through a repetitive pattern of a fixed number of small-amplitude (low-load) cycles followed by a fixed number of large-amplitude cycles (high-load) until failure. In Figure 2.3, the scale along the horizontal axis is the number of low-load cycles and the scale along the vertical axis is the number of high-load cycles. By design, the observations fall almost perfectly on lines of slopes $\theta_1 = 0.053$, $\theta_2 = 0.250$, $\theta_3 = 0.667$, $\theta_4 = 1.5$, $\theta_5 = 4$, and $\theta_6 = 19$. The dashed lines in Figure 2.3 represent these approximate linear usage paths.

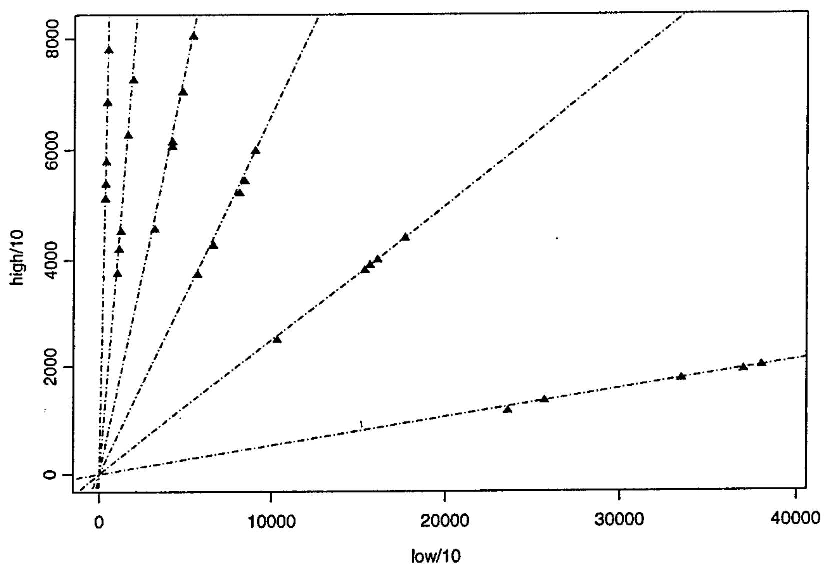

Figure 2.3: Metal Data with Approximate Linear Usage Paths. Each triangle represents the number of low-load and high-load cycles until a failure occurred, scaled by a factor of 1/10.

In data sets of this form, each device ages along a linear path of slope θ_i, $i = 1, \dots, m$, where $0 < \theta_1 < \theta_2 < \dots < \theta_m < \infty$. As such, the data set can be naturally partitioned into m samples, each consisting of failure data along a linear path. As with the traction motors, a policy can be specified for devices along a given usage path solely in terms of chronological time, since at any time $x > 0$ the position of a device along its usage path is known. To construct such a policy, consider the sample along each usage path separately. That is, use the n_i chronological ages at failure along the i^{th} path to estimate F_i, the conditional lifetime distribution of $X \mid \theta = \theta_i$. Then, use the empirical cost function (1.3) to estimate the optimal age replacement policy $\tilde{\tau}_i$ (which applies only to devices on the i^{th} path). The resulting policy, with replacement times summarized in vector $(\tilde{\tau}_1, \tilde{\tau}_2, \dots, \tilde{\tau}_m)$, takes the following form: replace a non-failed device on path i when its chronological age reaches $\tilde{\tau}_i$, $i = 1, \dots, m$.

For the metal data, suppose each F_i is estimated with the empirical distribution, placing mass $1/n_i = 0.2$ on each observation on the i^{th} path. Upon doing so, for $K/C = 0.5$ we obtain the following estimates: $\tilde{\tau}_1 = 23580$, $\tilde{\tau}_2 = 10300$, $\tilde{\tau}_3 = 5700$, $\tilde{\tau}_4 = 3200$, $\tilde{\tau}_5 = 1000$, $\tilde{\tau}_6 = 275$. Hence, the "composite" policy is as follows: replace non-failed devices on path 1 at age 23580; … ; replace non-failed devices on path 6 at age 275. The region corresponding to this policy is depicted on the left side of Figure 2.4.

At first glance, the proposed composite policy seems reasonable; however, the implementation of the policy is problematic. Consider two devices, say A and B. Suppose A has usage path 5, namely $\{(x, 4x), x > 0\}$ and B has usage path 4, namely

$\{(x, 1.5x), x > 0\}$. Under the composite policy, if device A is still operating we would replace it preventively when its age reaches $x = 1000$; at this time, it has usage $y(x) = 4000$. However, if device B is still operating at $x = 3000$, we would not replace it; at this time, its usage is $y(x) = 4500$. The metal fatigue experiment was designed so that the accumulation of low-load cycles and the accumulation of high-load cycles are the only factors leading to device failure. As such, this composite policy does not seem sensible, because device B is older than device A in every respect.

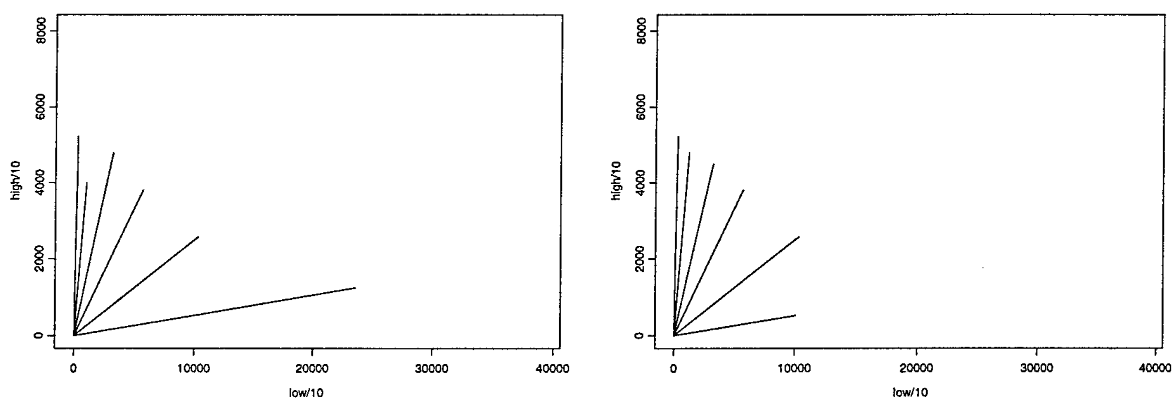

Figure 2.4: Composite Policies for the Metal Data.
The solid lines on left side of the figure represent the failure replacement region for the policy with replacement time vector (23580, 10300, 5700, 3200, 1000, 275). The right side of the figure depicts the failure replacement region for the policy with replacement time vector (10000, 10300, 5700, 3200, 1200, 275).

However, this is not the only problem we could encounter using this approach. Consider the same data, and suppose that instead of the policy suggested above, we obtain policy (10000, 10300, 5700, 3200, 1200, 275). The region corresponding to this policy is depicted on the right side of Figure 2.4. Suppose device A is on path 1 and device B is on path 2. Under this policy, if device A is still operating at age $x = 10000$, we would replace it preventively; at this time its cumulative usage is $y(x) = 526$. However, if device B is still operating at age $x = 10000$, we would not replace it (as it has not yet reached age $x = 10300$); at age $x = 10000$ its cumulative usage is $y(x) = 2500$. Device B is older than A in every respect; this composite policy does not seem sensible either. We now investigate the notion of a "sensible" policy in more detail.

B. DESCRIPTION OF A CLASS OF MULTIPLE-SCALE POLICIES

In this section we describe a class of multiple-scale policies which generalizes the class of single-scale policies $\{(0,\tau): \tau > 0\}$. We assume devices under consideration may differ only in their age in chronological time and in the amount of usage accumulated. As such, we implicitly assume there are no "hidden" covariates (e.g., better environmental conditions for certain devices, or additional measures of usage) affecting the process leading to eventual device failure. One example of a policy which generalizes the policy $(0,\tau)$ is $M = (0,u) \times (0,v)$, where $u > 0$ and $v > 0$, as considered in Case Study 2. Under this policy, a device is replaced if it fails at a time (X,Y) where $X < u$ and $Y < v$ or when its usage path crosses the boundary $x = u$ or $y = v$, whichever occurs first. Kordonsky and Gertsbakh (1994) devise policies based on lifetimes in two scales by projecting failure times onto a single time scale of the form $t = (1\text{-}a)x + ay(x)$, in which they define a replacement age τ_a. This policy replaces at age $t = \tau_a$ or upon failure, whichever occurs first. In the original two scales, this policy corresponds to the region $M = \{(x,y(x)): (1\text{-}a)x + ay(x) < \tau_a\}$. In fact, for $a = 0$, $M = M_X$, as in Case 1;

similarly, for $a = 1$, $M = M_Y$; when $0 < a < 1$, M is a right triangle with right angle at the origin.

On the other hand, consider the policy M depicted in Figure 2.5. From a preventive maintenance standpoint, this policy is not sensible since the device with $(x,y(x)) = (50,25)$ would be replaced preventively, but a non-failed device with $(x,y(x)) = (55,90)$ would not be replaced, even though it is older than the first device in both time scales. In order to be sensible under the assumptions described above, a policy prescribing preventive replacement of a device should prescribe preventive replacement of any "older" device. On the other hand, if a policy stipulates that a device should not be replaced preventively, then any "younger" device should not be replaced preventively either. To describe this more formally, we need a means of ordering two-dimensional failure times.

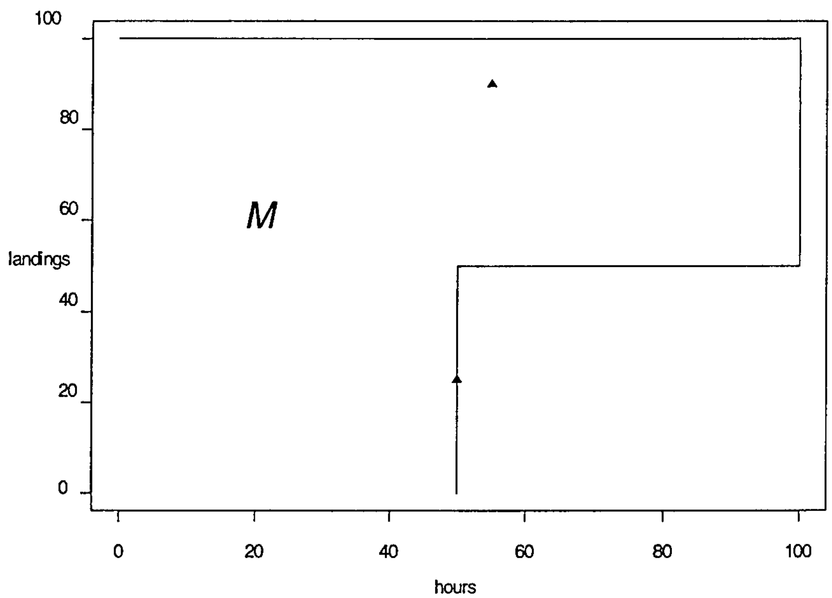

Figure 2.5: Undesirable Policy.
Under this policy, for example, a non-failed aircraft component with $x = 50$ flight hours and $y(x) = 25$ landings would be replaced, but one with $x = 55$ flight hours and $y(x) = 90$ landings would not be replaced.

A binary relation $\prec$ on a set $\mathcal{X}$ is a *simple order* on $\mathcal{X}$ if it is reflexive, transitive, anti-symmetric, and the members of every pair of elements of $\mathcal{X}$ are comparable. The relation $\prec$ is a *partial order* on a set $\mathcal{X}$ if it is reflexive, transitive, and anti-symmetric (thus, simple orders are partial orders; however, for partial orders, certain elements of $\mathcal{X}$ may be non-comparable). In addition, $L \subset \mathcal{X}$ is a *lower set* with respect to a partial order $\prec$ if $\boldsymbol{u} \in L$, $\boldsymbol{v} \in \mathcal{X}$ and $\boldsymbol{v} \prec \boldsymbol{u}$ imply $\boldsymbol{v} \in L$ (e.g., Robertson, Wright, and Dykstra, 1988); a lower set contains all "predecessors" of each of its members. For failure times $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$ in $\mathcal{X} \subseteq (0,\infty)^2$, we take $\prec$ to be the *matrix partial order* where $\boldsymbol{u} \prec \boldsymbol{v}$ if and only if $u_1 \leq v_1$ and $u_2 \leq v_2$. Note that $\mathcal{X}$ may be a proper subset of $(0,\infty)^2$, as in Case Study 3, where all failure times lie along one of six linear usage paths.

Using these definitions, we now characterize a class of policies for the multiple-scale age replacement problem. For ease of exposition, they are described in the plane. Let $\mathcal{X}$ denote the support of (X,Y), and $\mathcal{M}_{\mathcal{X}}$ denote the class of all open lower sets with respect to the matrix partial order on $\mathcal{X}$. Observe that for $\mathcal{X} = (0,\infty)$, the class of single-scale policies $\{(0,\tau)\colon \tau > 0\}$ is the class of open lower sets with respect to the simple order $\leq$ on $(0,\infty)$. Thus, $\mathcal{M}_{\mathcal{X}}$ is a natural generalization of the class of single-scale policies. In addition, members of $\mathcal{M}_{\mathcal{X}}$ are "sensible" policies from the standpoint of implementation when failure characteristics are captured by the two time scales. In the literature, Murthy et al (1995) use rectangular, triangular, other planar regions as warranty policies; every region they consider is a lower set with respect to the matrix partial order on $(0,\infty)^2$. Similarly, the policies developed in Case Studies 1 and 2 above are members of $\mathcal{M}_{\mathcal{X}}$, but the policies described in Case Study 3 are not. For ages measured in $k > 2$ scales, the notation is easily extended so that $\mathcal{M}_{\mathcal{X}}$ is the class of open lower sets in $\mathcal{X} \subseteq (0,\infty)^k$ with respect to the matrix partial order generalized to $(0,\infty)^k$.

C. NESTED POLICIES

Let $r = K/C$ denote the ratio of the preventive replacement cost and the additional cost to replace a device due to failure. As r decreases, it becomes proportionally more costly to replace at failure, so the replacement age based on a single scale should be more conservative. To show this, we make explicit the dependence on cost ratio r and define

$$D(\tau;r) = C(\tau)/C$$

$$= \frac{r+F(\tau)}{\int_0^\tau S(u)du}, \ \tau > 0, \tag{2.1}$$

where $C(\tau)$ is the cost function in (1.1). Let $\nu = \inf\{x: S(x) = 0\}$; $\nu \leq \infty$. Then, for cost ratio $s < r$,

$$D(\tau;r) - D(\tau;s) = \frac{r-s}{\int_0^\tau S(u)du} \tag{2.2}$$

is a positive, continuous, and strictly-decreasing function of τ on $(0,\nu)$. Suppose $D(\tau;s)$, and hence $C(\tau)$ with cost ratio s, attains a minimum at $\tau^*(s)$; there may be several minima. It can be shown that $\tau^*(s) \leq \nu$. For $\tau < \tau^*(s)$,

$$D(\tau;r) - D(\tau^*(s);r) = [D(\tau;r) - D(\tau;s)] + [D(\tau;s) - D(\tau^*(s);s)] + [D(\tau^*(s);s) - D(\tau^*(s);r)]. \tag{2.3}$$

Since $\tau^*(s)$ minimizes $D(\tau;s)$, the second term on the right-hand side of (2.3) is non-negative; in addition, because (2.2) is strictly decreasing on $(0,\nu)$, the sum of the first and third terms is positive. Thus, $D(\tau;r) > D(\tau^*(s);r) \ \forall \tau < \tau^*(s)$, and it follows that $C(\tau)$ with cost ratio r can only attain a minimum for $\tau \geq \tau^*(s)$.

Hence, for a decreasing sequence of cost ratios $r_1, r_2, \ldots$, the corresponding single-scale policies are nested. That is, if the corresponding optimal replacement times are, respectively, $\tau_1, \tau_2, \ldots$, then we know $\tau_1 \geq \tau_2 \geq \ldots$, so that the policies $(0,\tau_1) \supset (0,\tau_2) \supset \ldots$ form a sequence of nested lower sets.

Multiple-scale age replacement policies should also be more conservative as r decreases; in particular, policies for smaller r should be subsets of those for larger r. Let $\mathcal{X} = (0,\infty)^2$. Consider the policies based on region $M_1 = \{(x,y(x)): x + y(x) < 6, x > 0\}$ for $r_1 = 1$ and $M_2 = \{(x,y(x)): 5x + y(x) < 15, x > 0\}$ for $r_2 = 0.5$. Note both M_1 and M_2 are in $\mathcal{M}_{\mathcal{X}}$. Now, consider a device with linear usage path $y(x) = 5x$; the policies and usage path are depicted in Figure 2.6. This example illustrates that non-nested multiple-scale policies can prescribe replacement times that are not sensible. With $r_1 = 1$, the additional cost to replace a device due to failure is equal to the preventive replacement cost, while $r_2 = 0.5$ means the additional cost to replace a device due to failure is twice the preventive replacement cost. Thus, it seems policy M_2 should hedge against this higher failure replacement cost and suggest replacement at an earlier time than the time suggested by policy M_1.

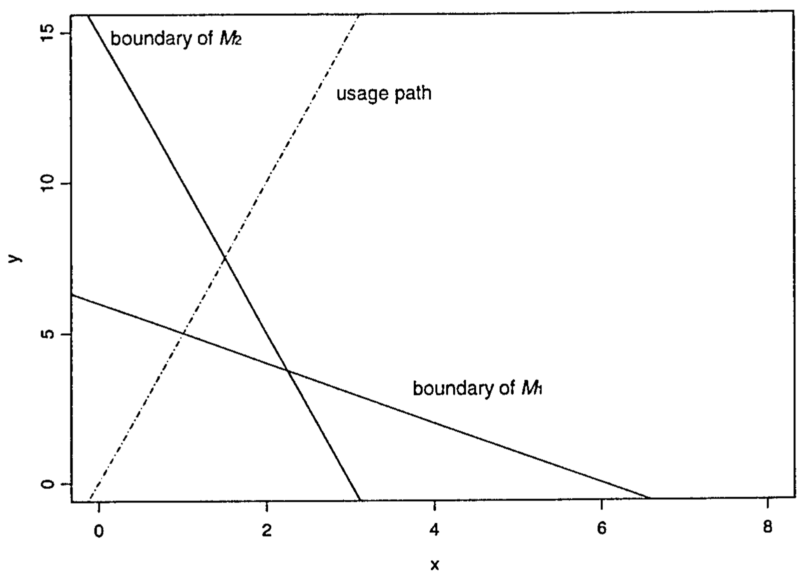

Figure 2.6: Non-nested Policies.
Solid lines represent boundaries of policies M_1 and M_2 and the dashed line represents a linear usage path of slope 5. Under policy M_1, non-failed devices on this path are replaced when $x = 1$; under policy M_2, non-failed devices on this path are replaced when $x = 1.5$.

THIS PAGE INTENTIONALLY LEFT BLANK

III. POLICIES BASED ON COMBINED SCALES

Due to the complexity of modeling lifetimes in multiple scales, much of the recent work in this area focuses on finding appropriate methods for combining scales to form a single time scale. Once such a combined scale is found, reliability tools such as age replacement theory can be brought to bear. We begin with a general discussion of combined scales. We then consider in detail three combined time scales in the literature that seem best suited for age replacement policies given failure data in two scales, age and usage. The first, and in a sense closest in spirit to our efforts, is the work of Kordonsky and Gertsbakh (1994) in which a combined scale is found for age replacement. The next two scales discussed are the "minimum CV" scale of Kordonsky and Gertsbakh (1993, 1995, 1997) and the "ideal" time scale of Duchesne and Lawless (2000). Both of these time scales are based solely on the underlying failure models and are developed independently of the age replacement problem. However, Gertsbakh and Kordonsky (1997) do suggest a context in which their min CV scale is "optimal" for preventive maintenance and Duchesne (1999) suggests his scales might be useful for maintenance planning.

A. COMBINED TIME SCALES

A formal definition of "time scale" is given by Duchesne and Lawless (2000). Let the set of all device usage paths $Z(x)$ be $\mathbf{Z}(x)$. For a particular device, let the "whole" usage path be $Z = Z(\infty)$; let the set of all such paths be $\mathbf{Z} = \mathbf{Z}(\infty)$. A *time scale* $\Phi(x,Z(x))$

is a non-negative real-valued functional of x and the path Z up to age x; it is required to be non-decreasing in x for all Z in $\mathbf{Z}$. Hence, a time scale is a function of chronological time and external covariates. Recent research efforts focus on finding a time scale Φ for which $t_z(x) = \Phi(x, Z(x))$ suffices for the calculation of probabilities for failures modeled in two scales. Oakes (1995) introduces the notion of the "collapsibility" of two time scales into one time scale which is "fully informative" in the sense that the probability of survival to a specified point (in the plane) depends only on the location of the point, not on the path taken to get to the point. Specifically, following Duchesne and Lawless (2000), the distribution of $X|Z$ is "collapsible in $y(x)$" if the survival probability at time x depends only on the path Z up to x only through its endpoint $(x, y(x))$. Thus, a time scale for a collapsible model can be written as $t_z(x) = \Phi(x, y(x))$. Collapsible models are common in the literature since in many cases X and $Y = y(X)$ are observable but the history $Z(X)$ is unknown. If the usage path is approximated by a straight line, the resulting models are collapsible since, $y(x) = \theta x$ and hence the path Z is known by its value $y(x)$ at any time x.

To illustrate the consequences of combining time scales in a collapsible model, consider the time scale $t = x + gy(x)$ for some $g > 0$. Note t induces a family of contours $\{y = (t - x)/g,\ t \in (0,\infty)\}$, as depicted in Figure 3.1 (Duchesne, 1999). The points where the usage paths intersect a given dotted contour line all have the same age (in the combined scale).

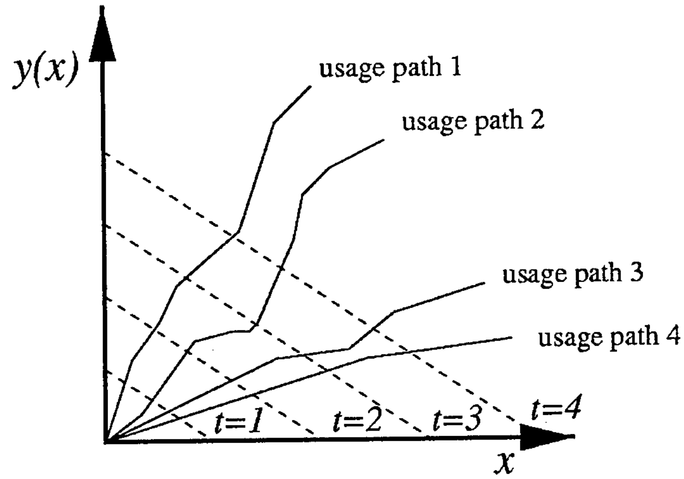

Figure 3.1: Contours of Linear Scale in a Collapsible Model. Jagged lines represent device usage paths and dashed lines represent contours of a linear time scale. Reproduced from Duchesne (1999).

This family of contours provides a means to compare points on different usage paths that may be non-comparable with respect to matrix partial order. Consider the points of intersection of contour $t = 4$ with the four usage paths in Figure 3.1. The matrix partial order does not enable us to determine the relative "age" of devices having age and usage represented by these points. On the other hand, the four points have the same age in scale t. Thus, the combined scale t induces an ordering (by age in this scale) of a set of points $(x_1,y(x_1)), (x_2,y(x_2)), \ldots, (x_n,y(x_n))$. In addition, as illustrated by the contours, the scale provides a means of specifying the relative age of one device in relation to another.

Different time scales order and "space" a given set of lifetimes differently. To illustrate this, consider Figure 3.2. Figure 3.2 contains a scatterplot of labeled points $(x_1,y_1), (x_2,y_2), \ldots, (x_{10},y_{10})$, randomly generated from the unit square; lines of slope -1 correspond to contours of scale $t = x + y(x)$ and lines of slope -10 correspond to contours of scale $s = x + 0.1y(x)$. Table 3.1 lists the coordinates of the points, their "age" in the two scales, and their ranks $r(t)$ and $r(s)$ in the two scales t and s.

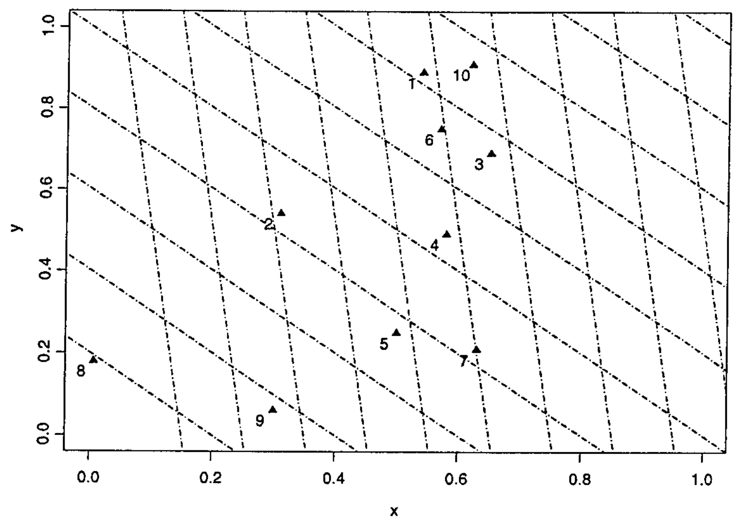

Figure 3.2: Contours of Linear Scales t and s.
Labeled points are randomly generated from the unit square. Lines of slope – 10 and –1 are contours of linear time scales s and t, respectively.

i	x	y	t	s	$r(t)$	$r(s)$
1	0.54	0.89	1.43	0.63	9	5
2	0.31	0.54	0.85	0.36	5	3
3	0.65	0.69	1.34	0.72	8	10
4	0.58	0.49	1.07	0.63	6	6
5	0.50	0.25	0.75	0.53	3	4
6	0.57	0.75	1.32	0.65	7	7
7	0.63	0.21	0.84	0.65	4	8
8	0.01	0.18	0.19	0.02	1	1
9	0.30	0.06	0.36	0.31	2	2
10	0.62	0.91	1.53	0.71	10	9

Table 3.1: The "Action" of Two Different Time Scales.
This table summarizes some of the information in Figure 3.2. Row 7 indicates (x_7, y_7) has age 0.84 in scale t, age 0.65 in scale s, is the fourth "youngest" point in scale t, and is the eighth "youngest" in scale s.

Using the combined scale, an age replacement policy can be expressed as $(0,\tau)$. In this form, a policy may have limited utility to the practitioner. On the other hand, a graphical depiction of this policy in terms of the original scales age and usage is very useful. In the original scales, the t-scale policy $(0,\tau)$ is equivalent to $M = \{(x,y(x)): \Phi(x,y(x)) < \tau\}$. For example, policy (0,0.4) in scale t above "translates" to the policy $\{(x,y(x)): x + y(x) < 0.4\}$ in Figure 3.2. In fact, the policy M_X discussed in Chapter II is a special case of such a "translation"; in this case, the combined scale is simply x. For most combined scales found in practice, an age replacement policy in the combined scale corresponds to a lower set in the original scales. This is only the case, however, when the combined time scale Φ is such that for $(x_1,y_1(x_1))$ and $(x_2,y_2(x_2))$ where $x_1 \le x_2$ and $y_1(x_1) \le y_2(x_2)$ we have $\Phi(x_1,y_1(x_1)) \le \Phi(x_2,y_2(x_2))$. In other words, since time scales are by definition required only to be increasing in x for any Z, it is possible to display combined scales for which the policy in the original scales is not a lower set.

B. A COMBINED SCALE FOR AGE REPLACEMENT

Kordonsky and Gertsbakh (1994) find the "best" scale for age replacement among the family of scales that are convex combinations of the two scales of age and usage. They consider the family of scales $\{t(a) = (1-a)x + ay(x),\ a \in [0,1]\}$; in scale $t(a)$ the lifetime is $T(a) = (1-a)X + aY$. The geometric interpretation of times in scale $t(a)$ is insightful. Time $t(a) = (1-a)x + ay(x)$ is proportional to the length of the orthogonal projection of the point $(x, y(x))$ onto vector $(1-a,a)$; the search for the "best" scale is essentially a search for the "best" such vector onto which to project the data.

For a fixed a, let $F_a(t) = P(T(a) < t)$, and define

$$C_a(\tau) = \frac{K + CF_a(\tau)}{\int_0^{\tau}(1 - F_a(u))du}, \ \tau > 0. \tag{3.1}$$

Thus, $C_a(\tau)$ is identical to the long-run average cost function (1.1). To find the "best" a, it seems reasonable to find, for a given a, the optimal replacement time in this scale (say τ_a), and then search [0,1] for the a yielding minimal $C_a(\tau_a)$. However, $C_a(\tau_a)$ has dimension cost per unit of time in the scale $t(a)$. Thus, values of $C_a(\tau_a)$ must be "converted" to make them comparable. To this end, Kordonsky and Gertsbakh convert (3.1) into a cost function with dimension cost per unit of chronological time in the following way. Because the average lifetime in scale $t(a)$ is $E[T(a)]$ and the average lifetime in chronological time x is $E[X]$, then from a damage accumulation perspective one unit of "$t(a)$-time" is equivalent to $E[X] / E[T(a)]$ units of x-time. Hence, the "converted" cost function is

$$D_a(\tau) = C_a(\tau)E[T(a)]/E[X], \qquad \tau > 0. \tag{3.2}$$

Let $\tau_a = argmin\ D_a(\tau)$. By definition, the "best" scale corresponds to the a^* which yields the minimum value of $D_a(\tau_a)$.

Kordonsky and Gertsbakh estimate a^* nonparametrically based on a simple random sample $(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)$. Care needs to be taken in applying their method, however. Consider the auto data set, taken from Wilson (1993), and which can be found in Appendix B. The boundaries of the policies for cost ratios $r = 0.5$, 0.25, and 0.125 are depicted in Figure 3.3. The policies are lower sets, but they exhibit the non-nested behavior exhibited in Figure 2.6. This suggests that the "best" scale is a function of the cost ratio. For the metal data, however, the policies are nested for $\{r: 0 < r < 1\}$. We suspect the non-nestedness of the policies derived from the auto data may be caused, in part, by the lack of sufficient spread in the distribution of usage paths. As such, it can be argued that non-nestedness is exhibited here since most observations in the auto data set fall roughly along a single regression line fit through the origin (unlike the metal data).

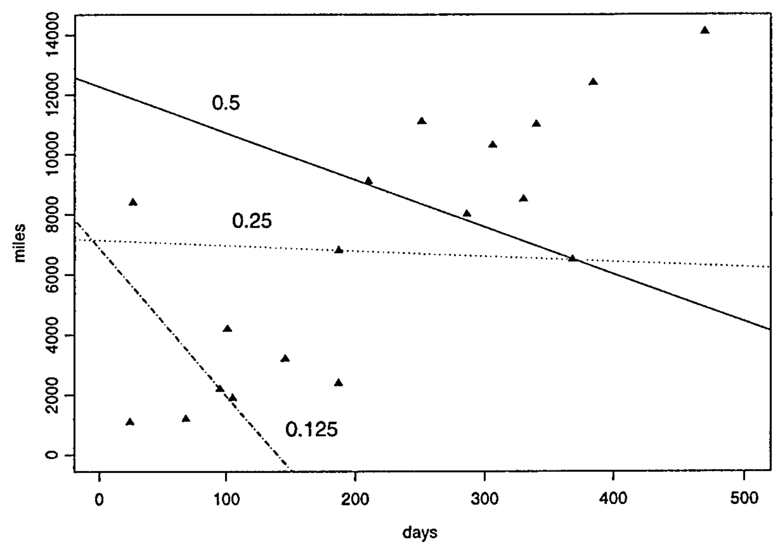

Figure 3.3: Non-nested Policies for Auto Data Based on "Best Scale" Method. Triangles represent the number of days and miles until a failure occurred. Labeled lines are policy boundaries for cost ratios r = 0.5, 0.25, and 0.125.

C. POLICIES BASED ON MINIMUM CV SCALE

We now examine another combined scale on which policies can be based. Consider again the family of linear scales $T_a = \{t(a) = (1-a)x + ay(x),\ a \in [0,1]\}$. Let $CV[T(a)]$ denote the coefficient of variation of the lifetime in scale $t(a)$; Kordonsky and

Gertsbakh (1993, 1995, 1997) identify the scale having a^* minimizing CV[$T(a)$]. They prove the (unrestricted) minimizer of CV^2[$T(a)$] has $a^* = g^*/(1+g^*)$, where

$$g^* = \frac{E[Y]Var[X] - E[X]Cov(X,Y)}{E[X]Var[Y] - E[Y]Cov(X,Y)}. \tag{3.3}$$

Since the family of scales specifies $a \in [0,1]$, it is important to describe the cases leading to $a^* \notin [0,1]$. In fact, from (3.3) we can show that $a^* \notin [0,1]$ iff either Case A or Case B holds in (3.4):

$$\begin{aligned} &Case\ A: CV^2(X) < \frac{Cov(X,Y)}{E[X]E[Y]} < CV^2(Y); \\ &Case\ B: CV^2(Y) < \frac{Cov(X,Y)}{E[X]E[Y]} < CV^2(X). \end{aligned} \tag{3.4}$$

In practice, an estimate $\hat{a}^*$ of a^* is obtained by replacing each of the terms in (3.3) with its sample estimate; Cases A and B are modified accordingly. Duchesne and Lawless (2000) note that when Case A holds, the minimizer of CV^2[$T(a)$] in T_a has $a^* = 0$, so that $t = x$ is the min CV scale. When Case B holds, $a^* = 1$, $t = y(x)$ is the min CV scale.

Consider using the min CV scale to construct a multiple-scale age replacement policy based on a simple random sample $(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)$. If the sample version of Case A holds, the policy is M_X (as in Case Study 1 of Chapter II). This means that if we use "min CV" as the criterion for time scale selection, it suffices to base the policy solely on the distribution of chronological time at failure. Similarly, if Case B holds, it suffices to base the policy solely on the distribution of cumulative usage at failure. Gertsbakh and Kordonsky (1997) note that if Cov(X,Y) < 0, then $a^* \in [0,1]$, so neither Case A nor Case B can occur. In this "more interesting" situation, we often find

$0 < a^* < 1$ (we note it is possible for a^* to be 0 or 1 if $\mathrm{Cov}(X,Y) < 0$). In this case, policies for a decreasing sequence of ratios r form a sequence of "nested" right triangles. For example, consider the metal data set. From the sample version of (3.3) we find $\hat{a}^* = 0.871$, so the min CV scale is $t = 0.129x + 0.871y(x)$. Using (1.3) in this scale we find the replacement time for $0.7 \le r \le 1$ is 3984; for $0.594 \le r < 0.7$ the replacement time is 3801; and for $r < 0.594$ the replacement time is 3396. These replacement times induce the set of nested right triangles depicted in Figure 3.4.

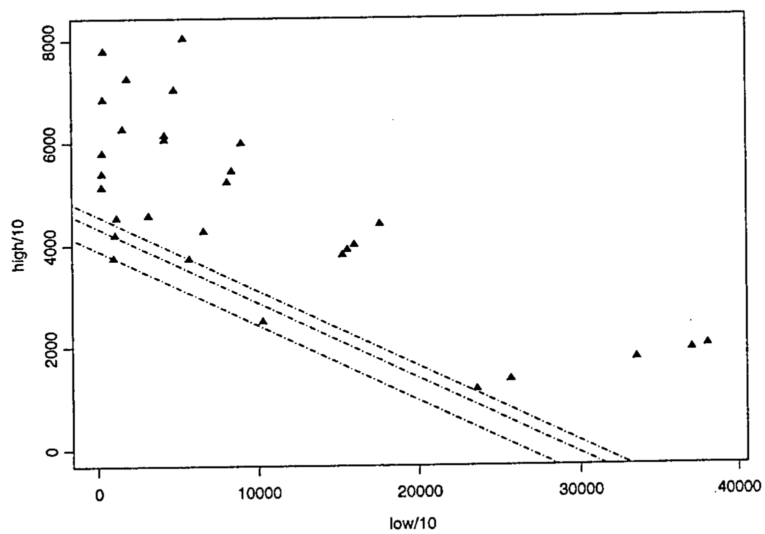

Figure 3.4: Nested Policies for Metal Data.
Dashed lines represent policy boundaries, based on the min CV scale. The policy for $r < 0.594$ is nested within the policy for $0.594 \le r < 0.7$, which is in turn nested within the policy for $0.7 \le r \le 1$.

D. POLICIES BASED ON IDEAL TIME SCALE

The long-run average cost $C(\tau)$ of a single-scale age replacement policy $(0,\tau)$ is given in (1.1); τ^* minimizes this expression. Using the transformation $p = F(\tau)$, with $F^{-1}(p)=\sup\{x\colon F(x) \le p\}$, equation (1.1) can be rewritten as

$$C(p) = \frac{K + C\,p}{\int_0^{F^{-1}(p)} S(u)du},\ 0 \le p \le 1. \tag{3.5}$$

Solving for p^* to minimize $C(p)$ in (3.5) and for τ^* in (1.1) are identical problems; the total time on test approach to solving the age replacement problem is based on this transformation. Thus, τ^* is the p^*-quantile of the lifetime distribution F. This latter formulation of the age replacement problem is insightful since it indicates that, under the policy, a device has probability p^* of failure before replacement. Thus, a "natural" generalization of policy $(0,\tau)$ is a multiple-scale policy for which the probability of failure before replacement is the same (say p) regardless of the path. With broader applications in mind, Duchesne and Lawless (2000) introduce an "ideal" time scale (ITS) which might be used to find such a policy.

Duchesne and Lawless (2000) motivate their definition of an ITS as follows. If a single-scale $t_z(x) = \Phi(x,Z(x))$ suffices for the calculation of failure probabilities, then the distribution of $T = \Phi(X,Z(X))$ along each Z should be independent of Z. That is, $P[T > t \mid Z] = P[T > t] = G(t)$, and $G(\cdot)$ does not depend upon Z. In addition, $t_z(x)$ must change whenever the conditional survivor functions $S_o(x, Z(x)) = P[X > x \mid Z]$ change. Duchesne and Lawless define $t_z(x) = \Phi(x,Z(x))$ to be an *ideal time scale* if it is a one-to-one function of $S_o(x, Z(x))$. In this case, $P[X > x \mid Z] = G[t_z(x)] = P[T > t_z(x)]$. Duchesne (1999) explains, "an ITS is a time scale in which we can directly compare the lifetimes of all the devices under study, no matter what their usage patterns are … it is 'ideal' in the sense that the age in the ITS is the only information needed to compute $P[X > x \mid Z]$, so it is 'sufficient' for computing the age of the units."

In fact, Duchesne (1999) mentions maintenance and inspection policies as potential applications of his ITS concept, and gives the following example. Suppose we want to inspect devices when their probability of failure is 0.25, regardless of the path. Suppose $t = x + 5y(x)$ is an ITS; let T denote the lifetime of a device in scale t and $t_{.25}$ denote the 25th percentile of that lifetime distribution. If $t_{.25} = 100$, devices should be inspected whenever $x + 5y(x) = 100$. Duchesne (1999) notes that ITSs are, by definition, unique up to one-to-one transformations. Hence, if t defines an ITS and ψ is a strictly increasing continuous function with $\psi(0) = 0$ and $\psi(\infty) = \infty$, then, $u = \psi(t)$ is also an ITS. Thus, for example, $u = t^2 = (x + 5y(x))^2$ is also an ITS; let U denote lifetime in this scale. Since $\Pr(U \leq 100^2) = \Pr(T \leq 100) = 0.25$, we have $u_{.25} = 100^2$. Thus, devices should be replaced whenever $(x + 5y(x))^2 = 100^2$, which is identical to the policy based upon scale t as defined above. This is a simple consequence of the monotone transformation.

Similarly, it seems we should be able to obtain a path-independent age replacement policy by finding the policy in any ITS and transforming this interval to a region in the positive quadrant (as described in section A above). There is a problem, however, stemming from the non-uniqueness of the ITS. Suppose T has an exponential distribution. It is well known that the optimal replacement time is infinite, so the policy in this scale would be to replace only at failure. The t-scale policy $(0,\infty)$ translates to the entire positive quadrant. Now, consider the policy based on scale $u = t^{1/2}$: U would then have a Weibull distribution, and the policy in scale u would be $(0,\nu)$ for some $\nu < \infty$. Translating to the plane results in the region $\{(x,y(x)): (x + 5y(x))^{1/2} < \nu\}$ which differs from the policy based on scale t.

To illustrate this, consider the metal fatigue data discussed in Case Study 3 of Chapter II. Duchesne and Lawless (2000) show that scale $t = x + 6.7y(x)$ is a reasonable approximation to the true, unknown ITS. Let T denote the lifetime in this scale; we first "reduce" each pair $(x, y(x))$ to scale t. Then, upon estimating $F_T(t) = P(T < t)$ with the empirical distribution, we find that for $r = 0.5$, the minimizer of (1.3) is $\hat{\tau} = 26125$. The ITS interval (0,26125) corresponds to the region $M_T = \{(x,y(x)): x + 6.7\, y(x) < 26125\}$. The boundary of this policy is the solid line in Figure 3.5. Under this policy, we replace the device upon failure or when the sum of its accumulated low cycles and 6.7 times its accumulated high cycles reaches 26125.

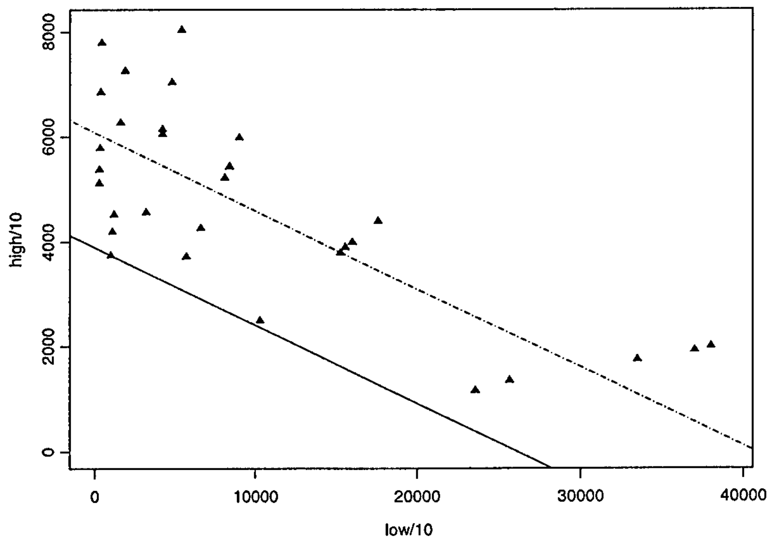

Figure 3.5: Policies Based on Ideal Scales *t* and *u*.
The solid line represents the policy boundary for $r = 0.5$ based on scale t and the dashed line represents the policy boundary for $r = 0.5$ based on scale u.

We now construct the age replacement policy for this data using another ITS. If $t = x + 6.7y(x)$ is ideal for the metal data, then the monotone transformation $u = t^2$ is also ideal. Proceeding as above, upon calculating the failure times U we find the minimizer of equation (1.3) is $\hat{V} = 40760^2$. In the plane, the ITS interval $(0, 40760^2)$ corresponds to the region $M_U = \{(x, y(x)): x + 6.7\, y(x) < 40760\}$. The boundary of this region is the dashed line in Figure 3.5. Observe M_U is not the same as M_T, the region derived from the first ideal scale.

In summary, path-independent, fixed-probability-of-failure inspection policies can be based on an ITS, but basing an age replacement policy on an ITS can pose significant problems. The reason ideal scales pose problems for age replacement but not fixed-probability inspection policies relates to our discussion of the ordering and "spacing" action of combined scales. An ITS Φ, like other combined scales, orders and induces spacings between the failure times. A monotone function ψ of Φ maintains the ordering of the times given by ITS Φ, but the spacings change. This fundamentally changes the nature of the failure distribution on which the optimal age replacement policy depends. (An obvious exception is when ψ is linear; see Lemma A.1 in Appendix A.) More specifically, let T and U denote the lifetimes in scales Φ and $\psi(\Phi)$, respectively; let τ^* and ν^* denote optimal replacement times in these scales. The observation above is that although $U = \psi(T)$, it is not necessarily true that $\nu^* = \psi(\tau^*)$. This is due to the fact that in transforming the cost function (1.1) from scale t to scale u, the numerator remains constant but the denominator changes.

E. DISCUSSION AND SUMMARY

In this chapter we have discussed how a multiple-scale age replacement policy might be obtained if scales age and usage are combined in various ways. One method of Kordonsky and Gertsbakh (1994) is motivated from the standpoint of cost. For a fixed $r > 0$, this method finds the "best" vector $(1-a, a)$ on which to project the data based on a "converted" cost function; the resulting policy M_r is triangular (or possibly of form M_X or M_Y). However, we note for $s < r$ the method is not guaranteed to have $M_s \subset M_r$; this is because the "best" scale depends on the cost ratio. For a fixed $r > 0$, policies based on the min CV scale are triangular (or possibly M_X or M_Y) and since minimizing CV results in a vector $(1-a,a)$ independent of r, the policies for a decreasing sequence of cost ratios are nested. Finally, we note that if, based on the failure data, a reasonable estimate of the ITS can be found, a policy in this scale has the property of fixed probability of failure before replacement, regardless of the path. While this property is attractive, we note monotone transformations of the ITS are also ideal, but do not necessarily result in the same policy as in the original ITS.

Combining scales is convenient in that it allows analysis to proceed along one scale. There is a drawback to the combining of scales, however. Kordonsky and Gertsbakh (1995) explain how damages in the different time scales can interact: in aviation, corrosion (as reflected by the time scale "calendar time") affects both fatigue damages due to the amount of time in level flight (as reflected by the time scale "flight hours") and the high-amplitude stresses incurred during the takeoff and landing cycle (as reflected by the time scale "number of landings"). As such, they observe "No single time scale is sufficient for a complete description of all wear and damage accumulation leading to failure in one of the aircraft parts." Thus, useful information may be lost even if the "best" single time scale is used (i.e., the one which best accounts for the damage accumulation processes and their interaction); for this reason, we proceed to the introduction of new methods which do not combine the scales.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. POLICIES GIVEN DATA ALONG SEVERAL LINEAR PATHS

In this chapter we generalize the single-scale failure replacement interval $(0,\tau)$ to the multiple-scale setting in which failure data fall along several linear paths. Such situations often arise in modeling real-world observational lifetime data in multiple scales (e.g., Gertsbakh and Kordonsky, 1998 and Lawless et al., 1995). In many cases X and Y are known but the usage curve Z is unknown and is approximated by a straight line. Linear usage paths may also arise by cyclic usage in fatigue life experiments (as exemplified in the metal data). The development is as follows. First, we establish notation to be used throughout the chapter. In so doing, we describe the cost function used to define an "optimal" policy in this setting. Next, we explain how to estimate the optimal policy for given failure data, and present an example. We then compare our approach to the methods found in the literature, and summarize.

A. "COMPOSITE" POLICIES

Consider a population of devices differing only in their rates of use, which remains constant throughout their lifetimes. Thus, suppose that upon entering service, a device is assigned a linear path Z_i (characterized by its slope θ_i) with probability p_i, $i = 1, \ldots, m$. Suppose also that $0 < \theta_1 < \theta_2 < \ldots < \theta_m < \infty$. Let F_i be the distribution of lifetime X (in chronological time) given $\theta = \theta_i$, $i = 1, \ldots, m$; as in Chapter I, $F_i(x) = P(X < x \mid \theta = \theta_i)$. From (1.1) the long-run average cost per unit time for a device operating with $\theta = \theta_i$ under policy $(0, \tau_i)$ is

$$C_i(\tau_i) = \frac{K + CF_i(\tau_i)}{\int_0^{\tau_i} S_i(u)du}, \ \tau_i > 0, \ i = 1, \ldots, m. \tag{4.1}$$

Let τ_i^* be an optimal age replacement time for devices on path i, $i = 1, \ldots, m$; that is, $\tau_i^* = \text{argmin } C_i(\tau_i)$. To form a composite policy from the path-specific policies $(0,\tau_i^*)$ for $i = 1, \ldots, m$, let $M_{\tau^*} = \{(x,\theta_i x): \ 0 < x < \tau_i^*, \ i = 1, \ldots, m\}$. This composite policy has replacement times summarized by the vector $(\tau_1^*, \tau_2^*, \ldots, \tau_m^*)$, meaning devices on path Z_i are replaced upon failure or when their age reaches τ_i^* (whichever occurs first), $i = 1, \ldots, m$. As in Case Study 1 of Chapter II, since at any given chronological time $x > 0$, the position of a device along its usage path is known, we can specify the replacement times solely in terms of chronological time.

In Case Study 3 of Chapter II, for the metal data, estimation of replacement times for such a composite policy did not result in a sensible policy. More specifically, with $\Theta = \{0.053, 0.250, 0.667, 1.5, 4, 19\}$ and $\mathcal{X} = \{(x,\theta_i x): \ 0 < x, \ \theta_i \text{ in } \Theta, \ i = 1, \ldots, 6\}$, the composite policy with replacement times (23580, 10300, 5700, 3200, 1000, 275) does not correspond to a region which is a lower set in $\mathcal{M}_{\mathcal{X}}$. We now give conditions on a replacement time vector $(\tau_1, \tau_2, \ldots, \tau_m)$ that ensure M_τ is a lower set.

Proposition 4.1. A composite policy $M_\tau = \{(x,\theta_i x): \ 0 < x < \tau_i, \ i = 1, \ldots, m\}$ for devices on linear usage paths where $0 < \theta_1 < \theta_2 < \ldots < \theta_m$ is a lower set with respect to the matrix partial order on $\mathcal{X} = \{(x,\theta_i x): \ 0 < x, \ i = 1, \ldots, m\}$ if and only if both $\tau_{i+1} \leq \tau_i$ and $\theta_{i+1}\tau_{i+1} \geq \theta_i\tau_i$, $i = 1, \ldots, m-1$.

Proof: Starting with the reverse statement, let $\boldsymbol{x} \in M_\tau$ and let $\boldsymbol{y} \in \mathcal{X}$ such that $\boldsymbol{y} \prec \boldsymbol{x}$. To show M_τ is a lower set with respect to the matrix partial order on $\mathcal{X}$, it suffices to show $\boldsymbol{y} \in M_\tau$. Because $\boldsymbol{x} \in M_\tau$, the age $\boldsymbol{x} = (t, \theta_j t)$ for some $0 < t < \tau_j$ and some θ_j in Θ. Similarly, because $\boldsymbol{y} \in \mathcal{X}$, $\boldsymbol{y} = (s, \theta_k s)$ for some $s > 0$ and some θ_k in Θ. Because $\boldsymbol{y} \prec \boldsymbol{x}$, it follows that $s \leq t$ and $\theta_k s \leq \theta_j t$. It suffices to show $0 < s < \tau_k$. First, treat the case $k \leq j$. Because $s \leq t$ and $\tau_j \leq \tau_k$, we have $0 < s \leq t < \tau_j \leq \tau_k$. On the other hand, if $k > j$, then because $\theta_k s \leq \theta_j t$ and $\theta_j \tau_j \leq \theta_k \tau_k$, we have $0 < s \leq (\theta_j/\theta_k)t < (\theta_j/\theta_k)\tau_j \leq \tau_k$. Thus, the policy is a lower set.

Turning to the direct statement, suppose M_τ is a lower set; let $i \in \{1, \ldots, m-1\}$. Suppose further that $\tau_{i+1} > \tau_i$. Let $x = (\tau_{i+1} + \tau_i)/2$; consider $\boldsymbol{u} = (x, \theta_{i+1} x) \in M_\tau$ and $\boldsymbol{v} = (x, \theta_i x) \in \mathcal{X}$. Note that $\boldsymbol{v} \prec \boldsymbol{u}$, but because $x > \tau_i$, $\boldsymbol{v} \notin M_\tau$. This contradicts the fact that M_τ is a lower set. Thus, $\tau_{i+1} \leq \tau_i$. Similarly, suppose $\theta_{i+1}\tau_{i+1} < \theta_i \tau_i$. Let $y = (\theta_{i+1}\tau_{i+1} + \theta_i \tau_i)/2$, $x = y/\theta_i$ and $z = y/\theta_{i+1}$. Consider $\boldsymbol{u} = (x,y) \in M_\tau$ and $\boldsymbol{v} = (z, y) \in \mathcal{X}$. Note that $\boldsymbol{v} \prec \boldsymbol{u}$, but because $z > \tau_{i+1}$, $\boldsymbol{v} \notin M_\tau$, contradicting the fact that M_τ is a lower set. Thus $\theta_{i+1}\tau_{i+1} \geq \theta_i \tau_i$.

This proposition reveals the problems encountered in Case Study 3 of Chapter II. The policy with $\tau = (23580, 10300, 5700, 3200, 1000, 275)$ has $\theta_5\tau_5 < \theta_4\tau_4$; in order for M_τ to be in $\mathcal{M}_\mathcal{X}$ we need $\theta_5\tau_5 \geq \theta_4\tau_4$ (all other requirements of the proposition are satisfied). Similarly, the policy with $\tau = (10000, 10300, 5700, 3200, 1200, 275)$ has $\tau_2 > \tau_1$; for M_τ to be in $\mathcal{M}_\mathcal{X}$ we need $\tau_2 \leq \tau_1$. Thus, for the metal data, the hypothetical policies we considered in Case Study 3 are not lower sets. Several ad hoc methods can be used to transform these policies into members of $\mathcal{M}_{\mathcal{X}}$. For one, a linear interpolation can be used to "smooth" sequential members of τ which violate either of the conditions $\tau_{i+1} \leq \tau_i$ or $\theta_{i+1}\tau_{i+1} \geq \theta_i\tau_i$. Another alternative is to use a pooling scheme (as is done in isotonic regression, ref. Robertson, Wright, and Dykstra 1988) to transform the policy. However, neither of these schemes takes into account the cost of implementing the resulting policy. Since it is desirable to obtain a sensible policy which is optimal with respect to some cost function, we now introduce such a cost function.

B. THE COST OF A COMPOSITE POLICY

The first policy of Case Study 3 of Chapter II is "optimal" in the sense that it minimizes the (estimated) long-run average cost per unit of time in use for devices on each path $i = 1,\ldots,m$. Unfortunately, the policy is not sensible from the standpoint of implementation. We need a means of obtaining a policy that is "optimal" in a sense which accounts for costs along each path, but is simultaneously "sensible." An equitable method of calculating the cost of policy M_τ with corresponding replacement time vector $\tau = (\tau_1, \tau_2, \ldots, \tau_m)$ is to form the average, weighted by the assigned probabilities, of the costs of the path-specific policies: let

$$C(\tau) = \sum_{i=1}^{m} p_i C_i(\tau_i),\ \tau_i > 0,\ i = 1, \ldots, m. \qquad (4.2)$$

A cost function of this form is studied by Gertsbakh and Kordonsky (1997) as they address the "optimal" time scale for maintenance in heterogeneous environments. Here

$C(\tau)$ represents the expected long-run average cost per unit of time in use of maintaining a device under a policy corresponding to its operating conditions. The dimension of $C(\tau)$ is in units of cost per unit of (chronological) time in use. If it is more meaningful to the decision maker, equation (4.2) can be easily transformed so it has dimension units of cost per unit of time in use in the second scale.

From Proposition 4.1, we note that in order for a policy M_τ with replacement time vector $\tau = (\tau_1, \tau_2, \ldots, \tau_m)$ to be in $\mathcal{M}_\mathcal{T}$, τ must lie in the set A, defined by

$$A = \{\tau \in (0,\infty)^m: \infty \geq \tau_1 \geq \tau_2 \geq \ldots \geq \tau_m > 0,\ \theta_1\tau_1 \leq \theta_2\tau_2 \leq \ldots \leq \theta_m\tau_m\}. \tag{4.3}$$

Thus, to find the optimal "sensible" policy for a given $r > 0$, one must minimize (4.2) subject to the restriction that τ is in A. Let τ^* denote this minimizer.

For a given $r > 0$, if a collection of conditional distributions $\{F_i\}$ has $(\tau_1^*, \tau_2^*, \ldots, \tau_m^*) \in A$, then by the optimality of each τ_i^* it follows that $\tau^* = (\tau_1^*, \tau_2^*, \ldots, \tau_m^*)$, regardless of the mixing probabilities. Collections of distributions with this property often arise from models common in the literature. Lawless, et al (1995) study failure data from automobile brake pads using a form of accelerated failure time model in which they form a time scale $u = x^{1-\eta}\, y(x)^\eta$, $\eta \in [0,1]$. They assume linear usage paths $y(x) = \theta x$, so that $u = x\theta^\eta$, and they fit a two-parameter Weibull distribution to failure times in scale u. Although their work does not pertain directly to age replacement theory, the resulting collection of distributions of $X|\theta$ has this property. Duchesne and Lawless (2000), Gertsbakh and Kordonsky (1998), and Oakes (1995) study linear time scales $t = x + gy(x)$, $g \geq 0$. Under a linear path assumption, time scale t takes form $x(1 + g\theta)$. When a parametric distribution including a scale parameter is fit to failure times in scale t, the resulting collection of distributions of $X|\theta$ has this property. In certain cases, proportional hazard models can also produce collections of conditional distributions with this property.

C. ESTIMATING THE OPTIMAL COMPOSITE POLICY

We now turn to estimation under constraints (4.3). Assume $\{F_i\}$ is a collection of distributions with $(\tau_1^*, \tau_2^*, \ldots, \tau_m^*) \in A$. Following (1.3), let $\hat{S}_i$ denote the empirical survivor function based on the ordered sample chronological lifetimes $x_{i,(1)} \leq x_{i,(2)} \leq \ldots \leq x_{i,(n_i)}$ from path i, where n_i is the number of observations on path i, and let

$$\hat{C}_i(\tau_i) = \frac{(K + C) - C\hat{S}_i(\tau_i)}{\int_0^{\tau_i} \hat{S}_i(u)du}, \quad \tau_i > 0,\ i = 1, \ldots, m. \tag{4.4}$$

Thus, $\hat{C}_i(\tau_i)$ estimates $C_i(\tau_i)$. The following is the analog of (1.3) for the multiple-path scenario:

$$\hat{C}(\tau) = \sum_{i=1}^{m} p_i \hat{C}_i(\tau_i), \quad \tau_i > 0,\ i = 1, \ldots, m. \tag{4.5}$$

In the univariate problem, the fact that the empirical cost function (1.3) is a piecewise decreasing function reduces the search for the minimum to a finite number of "strategic" points. Similar principles apply to searching for a minimizer of $\hat{C}(\tau)$; let $\hat{\tau}$

be such a minimizer, i.e., $\hat{C}(\hat{\tau}) \leq \hat{C}(\tau)$ for all τ in A. We now describe $\hat{\tau}$ and prove that it is globally optimal.

For convenience, suppose that along each path no two failure times are equal, so that $x_{i,(1)} < x_{i,(2)} < \ldots < x_{i,(n_i)}$; also let $x_{i,(0)} = 0$ and $x_{i,(n_i+1)} = \infty$, $i = 1, \ldots, m$. Form an m-dimensional grid

$$\Gamma = \bigtimes_{i=1}^{m} \{x_{i,(1)}, x_{i,(2)}, \ldots, x_{i,(n_i)}\} \tag{4.6}$$

based on the observations along each path. In each m-dimensional hypercube of the form

$$H = \bigtimes_{i=1}^{m} (x_{i,(j_i)}, x_{i,(j_i+1)}], \text{ where } j_i \in \{0, \ldots, n_i\}, i = 1, \ldots, m, \tag{4.7}$$

$\hat{C}(\tau)$ is decreasing in each argument; it follows that the minimum of $\hat{C}(\tau)$ in H occurs at the vertex $(x_{1,(j_1+1)}, x_{2,(j_2+1)}, \ldots, x_{m,(j_m+1)})$. Note this vertex dominates all other points in H with respect to the matrix partial order on $(0,\infty)^m$; that is, $\tau \prec (x_{1,(j_1+1)}, x_{2,(j_2+1)}, \ldots, x_{m,(j_m+1)})$ $\forall$ $\tau \in H$. Thus, to find the global minimum of $\hat{C}(\tau)$ in the absence of constraints, we evaluate $\hat{C}(\tau)$ at all such non-dominated vertices and select the one yielding the smallest cost. In the presence of constraints (4.3) defining set A, it seems reasonable to limit our search for $\hat{\tau}$ to the set of these vertices which lie in A, but it can be shown that checking only such vertices will not necessarily produce the global minimum. Such a procedure, though, will yield a point corresponding to an upper bound for the optimal cost.

Let $\boldsymbol{H}$ denote the set of all hypercubes H as in (4.7) for which $H \cap A \neq \varnothing$. For some H in this set, the non-dominated vertex $(x_{1,(j_1+1)}, x_{2,(j_2+1)}, \dots, x_{m,(j_m+1)})$ lies in A; for others, this vertex lies outside of A. In the latter case, the non-dominated point in $H \cap A$ (i.e., the point that simultaneously maximizes the value of each coordinate) yields the smallest value of $\hat{C}(\tau)$. To find $\hat{\tau}$, an enumeration procedure is utilized to find the non-dominated point, say $u(H)$, in $H \cap A$ for all H in $\boldsymbol{H}$. Then, $\hat{\tau} = \operatorname{argmin} \hat{C}(u(H))$ among all $H \in \boldsymbol{H}$. For each $H \in \boldsymbol{H}$, the non-dominated point $u(H)$ is constructed explicitly based on the following results.

Proposition 4.2. For any $\boldsymbol{x} = (x_1, x_2, \dots, x_m)$ in $(0,\infty)^m$, let $B_{\boldsymbol{x}} = \{\tau \in (0,\infty)^m : \tau \prec \boldsymbol{x}\}$. Define $\boldsymbol{u}(\boldsymbol{x})$ as follows: $\boldsymbol{u}(\boldsymbol{x}) = (u_1(\boldsymbol{x}), u_2(\boldsymbol{x}), \dots, u_m(\boldsymbol{x}))$ where

$$u_1(\boldsymbol{x}) = \min\left\{x_1, \frac{\theta_2}{\theta_1}x_2, \dots, \frac{\theta_m}{\theta_1}x_m\right\},$$
$$u_2(\boldsymbol{x}) = \min\left\{x_1, x_2, \frac{\theta_3}{\theta_2}x_3, \dots, \frac{\theta_m}{\theta_2}x_m\right\},$$
$$\vdots$$
$$u_i(\boldsymbol{x}) = \min\left\{x_1, x_2, \dots, x_i, \frac{\theta_{i+1}}{\theta_i}x_{i+1}, \dots, \frac{\theta_m}{\theta_i}x_m\right\},$$
$$\vdots$$
$$u_m(\boldsymbol{x}) = \min\{x_1, x_2, \dots, x_m\}.$$

Then, (1) $\boldsymbol{u}(\boldsymbol{x}) \in A \cap B_{\boldsymbol{x}}$ and (2) $\boldsymbol{y} \prec \boldsymbol{u}(\boldsymbol{x})\ \forall \boldsymbol{y} \in A \cap B_{\boldsymbol{x}}$.

Proof: First, $\boldsymbol{u}(\boldsymbol{x}) \in B_{\boldsymbol{x}}$ since $u_i(\boldsymbol{x}) \le x_i$, $i = 1, \ldots, m$; that is, $\boldsymbol{u}(\boldsymbol{x}) \prec \boldsymbol{x}$. To show $\boldsymbol{u}(\boldsymbol{x}) \in A$ it suffices to show $\theta_i u_i(\boldsymbol{x}) \le \theta_{i+1} u_{i+1}(\boldsymbol{x})$ and $u_i(\boldsymbol{x}) \ge u_{i+1}(\boldsymbol{x})$ for $i = 1, \ldots, m-1$. Let $i \in \{1, \ldots, m-1\}$. Since $\theta_{i+1} > \theta_i$,

$$\begin{aligned}
\theta_i u_i(\boldsymbol{x}) &= \theta_i \min\left(x_1, x_2, \ldots, x_i, \frac{\theta_{i+1}}{\theta_i} x_{i+1}, \ldots, \frac{\theta_m}{\theta_i} x_m\right) \\
&= \min(\theta_i x_1, \theta_i x_2, \ldots, \theta_i x_i, \theta_{i+1} x_{i+1}, \ldots, \theta_m x_m) \\
&\le \min(\theta_{i+1} x_1, \theta_{i+1} x_2, \ldots, \theta_{i+1} x_i, \theta_{i+1} x_{i+1}, \ldots, \theta_m x_m) \\
&= \theta_{i+1} \min\left(x_1, x_2, \ldots, x_i, x_{i+1}, \frac{\theta_{i+2}}{\theta_{i+1}} x_{i+2}, \ldots, \frac{\theta_m}{\theta_{i+1}} x_m\right) \\
&= \theta_{i+1} u_{i+1}(\boldsymbol{x});
\end{aligned}$$

also

$$\begin{aligned}
u_i(\boldsymbol{x}) &= \min\left(x_1, x_2, \ldots, x_i, \frac{\theta_{i+1}}{\theta_i} x_{i+1}, \frac{\theta_{i+2}}{\theta_i} x_{i+2}, \ldots, \frac{\theta_m}{\theta_i} x_m\right) \\
&\ge \min\left(x_1, x_2, \ldots, x_i, x_{i+1}, \frac{\theta_{i+2}}{\theta_{i+1}} x_{i+2}, \ldots, \frac{\theta_m}{\theta_{i+1}} x_m\right) \\
&= u_{i+1}(\boldsymbol{x}).
\end{aligned}$$

Thus, $\boldsymbol{u}(\boldsymbol{x}) \in A \cap B_{\boldsymbol{x}}$, proving (1). To show (2), let $\boldsymbol{y} \in A \cap B_{\boldsymbol{x}}$, and let $i \in \{1, \ldots, m\}$. Since $\boldsymbol{y} \in A$, $y_1 \ge y_2 \ge \ldots \ge y_i$ and $\theta_m y_m \ge \ldots \ge \theta_{i+2} y_{i+2} \ge \theta_{i+1} y_{i+1} \ge \theta_i y_i$, then $(\theta_m/\theta_i) y_m \ge \ldots \ge (\theta_{i+2}/\theta_i) y_{i+2} \ge (\theta_{i+1}/\theta_i) y_{i+1} \ge y_i$, so by definition of $\boldsymbol{u}(\boldsymbol{y})$, $u_i(\boldsymbol{y}) = y_i$. It follows that $\boldsymbol{u}(\boldsymbol{y}) = \boldsymbol{y}$. Since $\boldsymbol{y} \in B_{\boldsymbol{x}}$, it follows that $y_i \le x_i$, $i = 1, \ldots, m$. Since each $u_i(z)$ is non-decreasing in each argument of $z \in (0,\infty)^m$, we have $\boldsymbol{y} = \boldsymbol{u}(\boldsymbol{y}) \prec \boldsymbol{u}(\boldsymbol{x})$, as required.

We now use this result to find $u(H)$ for $H \in \boldsymbol{H}$.

Proposition 4.3. Let H as in (4.7) be a member of $\boldsymbol{H}$; let $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$ denote the vertex $(x_{1,(j_1+1)}, x_{2,(j_2+1)}, \ldots, x_{m,(j_m+1)})$ and let $\boldsymbol{z} = (z_1, z_2, \ldots, z_m)$ denote the vertex $(x_{1,(j_1)}, x_{2,(j_2)}, \ldots, x_{m,(j_m)})$. Let $u(H) = u(\boldsymbol{x})$ as in Proposition 4.2. Then (1) $\boldsymbol{y} \prec u(H)$ $\forall \boldsymbol{y} \in A \cap H$, and (2) $u(H) \in A \cap H$.

Proof: Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_m) = u(H)$. Let $\boldsymbol{y} \in A \cap H$ (such a $\boldsymbol{y}$ exists, since $A \cap H \neq \varnothing$). Since $H \subset B_x$ it follows that $\boldsymbol{y} \in A \cap B_x$; from Proposition 4.2 we know that $\boldsymbol{y} \prec \boldsymbol{u}$, thus proving (1). By (1), we have $y_i \leq u_i$, $i = 1, \ldots, m$. Since $\boldsymbol{y} \in H$, we know that $z_i < y_i \leq x_i$, $i = 1, \ldots, m$. Because $\boldsymbol{u} \prec \boldsymbol{x}$, we know $u_i \leq x_i$, $i = 1, \ldots, m$. From these inequalities it follows that $z_i < u_i \leq x_i$, $i = 1, \ldots, m$, so that $u(H) \in H$. By Proposition 4.2 we know $u(H) \in A$; thus, we have shown (2).

We now show that our procedure returns the global minimum of $\hat{C}(\tau)$.

Theorem 4.1: $\hat{C}(\hat{\tau}) \leq \hat{C}(\tau)\ \forall \tau \in A$.

Proof: Let $\tau \in A$. Because the grid Γ defines a partition of the positive orthant, $\tau \in H$ for some $H \in \boldsymbol{H}$. Form $u(H)$ as described above. By definition, $\hat{C}(\hat{\tau}) \leq \hat{C}(u(H))$, so it remains to show $\hat{C}(u(H)) \leq \hat{C}(\tau)$. By construction of $u(H)$ we have $\tau \prec u(H)$; in H, $\hat{C}(\tau)$ is decreasing in each argument, so it follows that $\hat{C}(u(H)) \leq \hat{C}(\tau)$, as required.

D. EXAMPLE

Returning to the metal fatigue data from Case Study 3 of Chapter II, Table 4.1 contains the policy vector $\hat{\tau} = (\hat{\tau}_1, \hat{\tau}_2, \ldots, \hat{\tau}_m)$ for $r = 0.5$ along with the values $\theta_i \hat{\tau}_i$ to amplify the fact that $M_{\hat{\tau}}$ is a member of $\mathcal{M}_{\mathcal{X}}$. In the policy for $r = 0.75$, $\hat{\tau}_2$=15200, so that $\theta_2 \hat{\tau}_2$=3800. All other components are identical to the policy for $r = 0.5$. The policy for $r = 1$ is identical to the policy for $r = 0.75$. Thus, for this data the procedure produces nested policies for these values of the cost ratio. Figure 4.1 contains a scatterplot of the data overlaid with line segments representing paths curtailed by their corresponding replacement times for $r = 0.5$.

i	Slope θ_i	$\hat{\tau}_i$	$\theta_i \hat{\tau}_i$
1	0.053	23580	1241
2	0.250	10300	2575
3	0.667	5700	3800
4	1.500	2666.67	4000
5	4.000	1000	4000
6	19.00	275	5225

Table 4.1: Composite Policy for the Metal Data, $r = 0.5$.
For example, row 5 indicates that non-failed devices on a linear usage path of slope 4 are replaced when the number of low-load cycles accrued reaches 1000. At this time, the number of high-load cycles accrued is 4000.

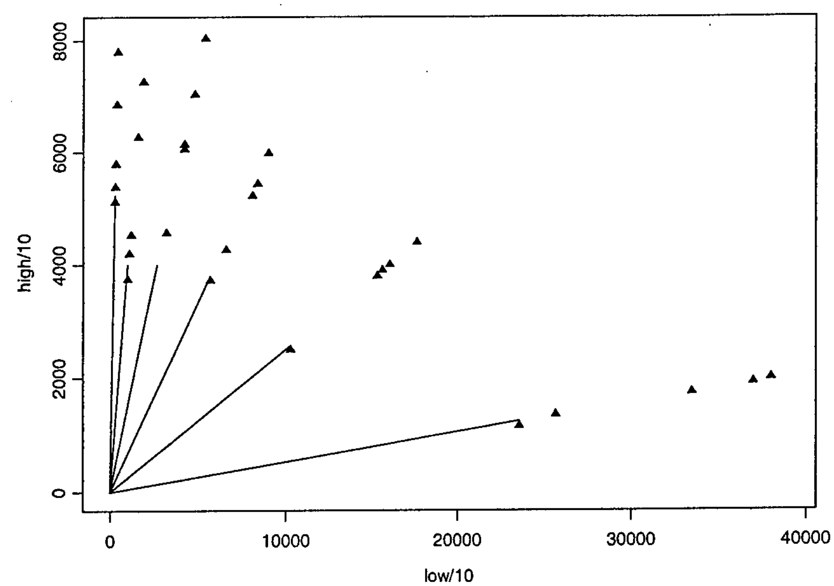

Figure 4.1: Metal Data with Policy for $r = 0.5$.
The solid lines represent the failure replacement region for the policy with replacement time vector (23580, 10300, 5700, 2666.67, 1000, 275).

This example also sheds light on ways to reduce the computational burden of finding $\hat{\tau}$: it is often unnecessary to compute $\hat{C}$ at the non-dominated point in every $H \in \boldsymbol{H}$. We recommend first finding the unrestricted minimizer $\tilde{\tau}$. A basic optimization principle is that if the solution of a relaxation happens to satisfy a restriction, then it solves the restriction. This principle implies that if $\tilde{\tau} \in A$, then $\hat{\tau} = \tilde{\tau}$. Thus, if the unrestricted minimizer lies in the set A, no further computation is necessary. Computing $\tilde{\tau}$ can save computation even if $\tilde{\tau} \notin A$. In some cases, $\tilde{\tau}$ may violate only one constraint defining the set A; restricting the coordinates causing the violation (while leaving the others relaxed) may lead to an optimal solution. More specifically, suppose that

$\tilde{\tau} = (\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_m)$ is such that for some k in $\{1,..,m-1\}$, either $\tilde{\tau}_k < \tilde{\tau}_{k+1}$ or $\theta_k \tilde{\tau}_k > \theta_{k+1} \tilde{\tau}_{k+1}$. Let $\tilde{\tau}'_k$ and $\tilde{\tau}'_{k+1}$ minimize

$$C'_k(\tau_k, \tau_{k+1}) = \hat{C}_k(\tau_k) p_k + \hat{C}_{k+1}(\tau_{k+1}) p_{k+1},$$

subject to

$$A_k = \{(\tau_k, \tau_{k+1}) \in (0,\infty)^2 \colon \tau_k \geq \tau_{k+1},\ \theta_k \tau_k \leq \theta_{k+1} \tau_{k+1}\}.$$

Let $\tilde{\tau}'$ denote the vector formed by replacing $\tilde{\tau}_k$ and $\tilde{\tau}_{k+1}$ in $\tilde{\tau}$ with $\tilde{\tau}'_k$ and $\tilde{\tau}'_{k+1}$, respectively. It can be shown that if $\tilde{\tau}' \in A$, then $\hat{\tau} = \tilde{\tau}'$. This approach works for the metal data for $r = 0.5$; recall from Case Study 3 that $\tilde{\tau}$ violates one constraint defining set A. This approach applies sequentially on the metal data for $r = 0.75$; in this case $\tilde{\tau}$ violates two constraints.

E. COMPARISON WITH SCALE-COMBINING APPROACHES

The scale-combining methods discussed in Chapter III differ fundamentally from our estimation procedure in their motivation, but in some cases produce sensible policies. The "best scale" method seeks the linear time scale $t(a) = (1-a)x + ay(x)$, $a \in [0,1]$, with corresponding $t(a)$-scale replacement time τ_a, that yields the lowest long-run average cost (per unit of chronological time, after "conversion"). The min CV method seeks the linear scale corresponding to the smallest lifetime CV. Both of these procedures use the data to produce a linear time scale and hence a policy of the form M_X, M_Y, or a triangular set $M_a = \{(x,y(x)) \colon t(a) < \tau_a\}$. In contrast, the policies produced by our procedure are required only to be lower sets. This is a broader class of policies than those resulting from a linear scale.

Ideal time scale methods seek the scale t such that $P[T > t_0 \mid Z]$ does not depend on the path Z; hence, a policy based on an ITS has the property that the probability of failure before replacement in this scale is the same, regardless of the path. Most of the focus of Duchesne (1999) is on inference procedures for the parameters of ITS models which are either linear (i.e., $t = x + gy(x)$, $g \geq 0$) or multiplicative (i.e., $u = x^{1-\eta}\, y(x)^{\eta}$, $0 \leq \eta \leq 1$). In the case of linear paths with slopes $\theta \in \{\theta_1,\ldots,\theta_m\}$, these scales always result in sensible policies. To demonstrate this, suppose the data are reasonably described by a linear ITS model $t = x + gy(x)$. The "best" scale for age replacement and min CV scale can be re-parameterized to this form. The policy takes the following form: replace non-failed devices when $x + gy(x) = \hat{\tau}$. It follows that the replacement time vector is $(\hat{\tau}/(1 + g\theta_1),\ldots,\hat{\tau}/(1 + g\theta_m)) \in A$. Restricting attention to preventive maintenance policies formed by ITS models may be appropriate in some cases; however, we have noted in Chapter III that the non-uniqueness of an ITS can cause problems for estimation of age replacement policies even when the ITS has a simple parametric form. Unfortunately, given a set of lifetime data (along linear paths or otherwise) it is rarely clear which (if any) parametric form the ITS should take. Duchesne (1999) suggests a non-parametric procedure for estimating the true, unknown ITS; this procedure links the quantiles along the paths. Policies based on the resulting scale can be constructed which are not lower sets.

In Chapter II we have noted that in the single-scale problem, policies corresponding to a sequence of decreasing cost ratios are "nested." We have also observed that this quality is desirable for multiple-scale policies because non-nested, multiple-scale policies prescribe replacement times for devices on some paths that are inconsistent with respect to the corresponding cost ratios. We have also observed in Chapter III that policies based on either the min CV scale or on an ITS are nested, but policies based on the "best" scale for age replacement method are not guaranteed to be nested. Due to the nature of the single-scale cost function (1.3) and in turn (4.5), the policies produced by our procedure are not necessarily nested. However, we show in Chapter V that in practice, our procedure tends to produce nested policies even with small samples. In such cases, our procedure forms a time scale based on the cost ratio r. The points along each path corresponding to the replacement time for a given r have the same age in this scale. Also, in a manner analogous to the cost sensitivity analyses conducted with the aid of TTT plots, we find there are ranges of r over which the same composite policy is valid. Combined scales, on the other hand, essentially order the observations based on their lifetimes in the combined scales; points along contours of these scales are the same "age" in these scales, indicating they have, in a sense, accumulated the same level of damage.

F. DISCUSSION AND SUMMARY

In this chapter we developed a method of estimating the optimal "sensible" policy given lifetimes from a population of devices which age along linear paths. Under such a policy, non-failed devices on path Z_i are replaced when their chronological age reaches τ_i, $i = 1, \ldots, m$. As such, this composite policy technically applies only to devices on these paths. Policies based on combined scales of the form considered in Chapter III have this same form when applied to data on linear paths. The assumption that devices age exactly along linear paths is usually an approximation of reality; thus, it is worthwhile to consider ways to extend these policies to ones that apply to devices on any path. The policy $(0,\tau)$ in a combined scale t extends in a natural way to the region $\{(x,y(x)): t < \tau\}$ in the positive quadrant, as exemplified in Figure 3.4 and Figure 3.5.

The key consideration for extending the policy produced by our estimation procedure is to ensure that the resulting policy is a lower set with respect to the matrix partial order on $(0,\infty)^2$. Consider, for example, a population of devices aging along lines of slope $\theta_1 = 0.5$, $\theta_2 = 2$, or $\theta_3 = 8$. Suppose that for some $r > 0$ the replacement times are $\tau_1 = 20$, $\tau_2 = 10$, and $\tau_3 = 5$, respectively. The solid lines segments in Figure 4.2 represent the failure replacement region for this policy. To extend this policy to the positive quadrant, we need a non-increasing function on $(0,\infty)$ that is contained within the rectangular regions delimited by the dashed lines in Figure 4.2. This function induces a boundary of the failure replacement region; non-failed devices are replaced when their usage curve crosses this boundary. A "conservative" extension is to choose a step function coincident with the lower boundaries of the boxes; a more "aggressive" extension is to choose a step function coincident with the upper boundaries of the boxes (in this case there is no usage limit for devices with $x < 5$). Between the two extremes, we arbitrarily choose a smooth curve through the policy points $\{(20,10), (10,20), (5,40)\}$, as depicted in Figure 4.2. We address the problem of determining the cost of implementing such policies in Chapter VI.

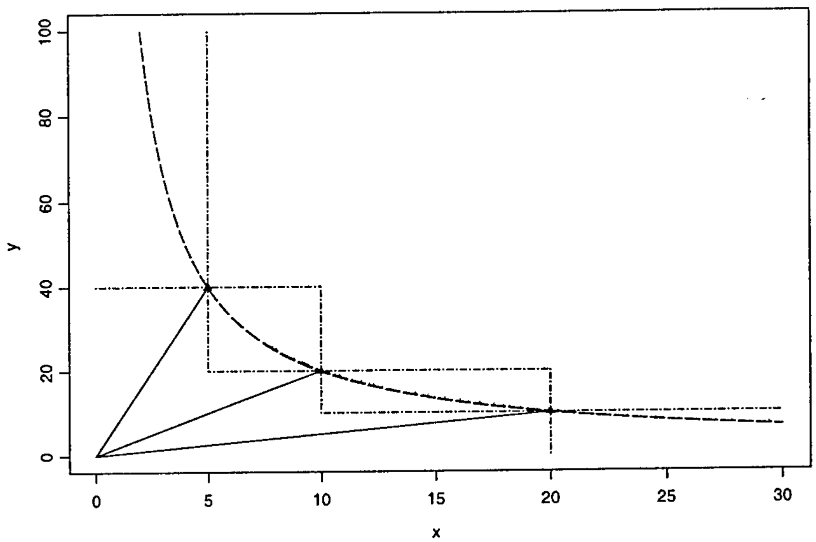

Figure 4.2: Extension of Estimated Optimal Policy.
The solid lines represent the failure replacement region for the policy with replacement time vector (20, 10, 5). The dashed lines represent bounds for a non-increasing function serving as a policy boundary under the lower set restriction. The smooth curve represents the boundary of one possible extension of the policy based on the linear paths of slope 0.5, 2, and 8.

Additionally, we note that our focus in this chapter has been on completely non-parametric estimation of the optimal policy. We acknowledge it is also possible to estimate the F_i under the restriction that the estimates be IFR. Ingram and Scheaffer (1976), however, find little value added from the increased computational burden over empirical estimation. We remark that if parametric (or other nonparametric) distributions are fit to each F_i and a $\tilde{\tau}_i$ estimating τ_i* is found for a given $r > 0$, the vector $(\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_m)$ is not necessarily in A. It is possible, however, to estimate parameters of certain collections $\{F_i\}$ under the restriction that $(\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_m)$ be in A. Gertsbakh and Kordonsky (1998) consider an example of such a collection. They discuss estimation in the Weibull family under which the shape parameter is constant for all paths but the scale parameters are allowed to vary. Geurts (1983) acknowledges optimal age replacement times in the Weibull family are relatively insensitive to the shape parameter, so in our setting this seems to be a reasonable approach. In such a case, it can be shown that if the scale parameters satisfy conditions akin to (4.3), the resulting composite policy $(\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_m)$ is in A. General conditions under which $(\tilde{\tau}_1, \tilde{\tau}_2, ..., \tilde{\tau}_m)$ is in A need further study.

Finally, in this chapter we focus on linear paths in two scales. The concepts developed here can be generalized to more than two scales. For example, m linear paths in k+1 scales can be represented by $(x, y_1(x), \ldots, y_k(x))$ where $y_j(x) = \theta_{ij}x$, $i = 1, \ldots, m$, $j = 1, \ldots, k$. For m such paths, as in two scales, an age replacement policy need only specify replacement ages $(\tau_1, \tau_2, \ldots, \tau_m)$ in chronological time. In addition, the cost function (4.2) remains the same. The difference comes in specifying constraints (4.3) so that the policy $(\tau_1, \tau_2, \ldots, \tau_m)$ is indeed sensible in the original scales. We speculate that the constraints and the estimator will have form similar to those developed in this chapter. We do note, however, that it is difficult to imagine practical applications of extending these results to more than three scales.

V. PROPERTIES OF THE ESTIMATED OPTIMAL COMPOSITE POLICY

In this chapter we address the properties of the policy $\hat{\tau}$. We begin with a discussion of its large-sample properties, and then investigate its small-sample behavior through simulation. We conclude with simulation results aimed at comparing the performance of the policies produced by our procedure with those based on the min CV method.

A. LARGE-SAMPLE PROPERTIES

Let $\hat{S}_i$ be a uniformly strongly consistent estimator of S_i, $i = 1,\ldots,m$. For example, if lifetimes along path i are from a simple random sample, then taking $\hat{S}_i$ to be the empirical survivor function (1.2) gives a non-parametric estimator of S_i, which by the Glivenko-Cantelli lemma converges uniformly to S_i with probability 1. On the other hand, should lifetimes along path i be right-censored, depending on the censoring mechanism, the Kaplan-Meier estimator is an appropriate choice for $\hat{S}_i$. With such an estimator and the assumption that $\tau_i^* < \infty$ exists and is unique (e.g., if F_i is IFR with failure rate strictly increasing to ∞) then it is well known (e.g., Arunkumar, 1972) that $\tilde{\tau}_i$ minimizing $\hat{C}_i(\tau_i)$ is a strongly consistent estimator of τ_i^*. From this it follows, for the composite policy with replacement time vector $\tilde{\tau} = (\tilde{\tau}_1,\ldots,\tilde{\tau}_m)$, that

$$\max_{1\le i\le m}\left|\tilde{\tau}_i - \tau_i^*\right| \to 0$$

with probability 1 as $n_i \to \infty$, for all $i = 1, \dots, m$. This result, of course, does not require the estimated policy $\tilde{\tau}$ to be in A even if $(\tau_1^*, \dots, \tau_m^*)$ is in A.

The showing of the strong consistency of the estimator $\hat{\tau}$, which is required to be in A, takes a bit more care. With the individual $\tau_i^* < \infty$ and unique, and $\hat{S}_i$ a uniformly strongly consistent estimator of S_i, then a small modification of Ingram and Scheaffer's (1976) argument shows that $\hat{C}_i$ converges uniformly to C_i in an interval bounded away from zero with probability 1. In particular, Ingram and Scheaffer (1976) show that $\int_0^\infty S_i(u)du < \infty$ by appealing to the condition that F_i be IFR. However, this is also true if $\tau_i^* < \infty$ and unique because for $t > \tau_i^*$, $0 < C_i(\tau_i^*) < C_i(t)$ and hence $\int_0^\infty S_i(u)du = (K+C)\lim_{t\to\infty}(1/C_i(t)) < \infty$. For the multiple-scale functions $C(\tau)$ and $\hat{C}(\tau)$, we have

$$\left|\hat{C}(\tau) - C(\tau)\right| = \left|\sum_{i=1}^m p_i \hat{C}_i(\tau_i) - \sum_{i=1}^m p_i C_i(\tau_i)\right| \le \sum_{i=1}^m p_i \left|\hat{C}_i(\tau_i) - C_i(\tau_i)\right|.$$

Thus, for $a > 0$ we see that $\hat{C}(\tau)$ converges uniformly to $C(\tau)$ in the m-dimensional region $[a,\infty)^m$. Suppose $\tau \notin [a,\infty)^m$, so that $\tau_j < a$ for some $j = 1,\dots,m$. Then $C(\tau)$, and similarly $\hat{C}(\tau)$, are bounded below as follows:

$$\begin{aligned} C(\tau) &\ge p_j C_j(\tau_j) \\ &\ge p_j \frac{(K+C)-C}{\int_0^a S_j(u)du} \\ &\ge \frac{K}{a} \min_{1\le i\le m} p_i. \end{aligned}$$

An application of the multivariate analog of Theorem 1 of Arunkumar (1972, p. 252) then gives strong consistency of $\hat{\tau}$, as an estimate of τ^*, as stated in the following theorem.

Theorem 5.1. Let $r > 0$, $(\tau_1^*, \ldots, \tau_m^*) \in A$ be unique, where A is defined in (4.3), $\tau_i^* < \infty$, and $\hat{S}_i$ be a uniformly strongly consistent estimator of S_i, $i = 1, \ldots, m$. Then

$$\max_{1 \le i \le m} \left| \hat{\tau}_i - \tau_i^* \right| \to 0$$

with probability 1 as each $n_i \to \infty$, $i = 1, \ldots, m$.

We note that the proof of Theorem 5.1 does not require $\hat{\tau}$ to be unique. Indeed with $\hat{S}_i$ as the empirical survivor function, uniqueness of $\hat{\tau}$ is not guaranteed. In addition, although $(\tau_1^*, \ldots, \tau_m^*) \in A$ for most practical cases, this is not a strict requirement. What is required in the proof of Theorem 5.1 is the existence of a unique τ^* minimizing $C(\tau)$ among $\tau \in A$ and that τ^* has finite elements. Weak convergence of $\hat{\tau}$ is not studied here. Arunkumar (1972) does develop the asymptotic distribution of the minimizer of (1.3) in the one-dimensional case. Perhaps Arunkumar's approach can be used to establish weak convergence for the multi-dimensional, restricted estimator $\hat{\tau}$.

Furthermore, for large samples, the estimators of the optimal policies are nested. Suppose $s < r$, and let $\tau_i^*(s)$ and $\tau_i^*(r)$ minimize $C_i(\tau_i)$ with respective cost ratios s and r, $i = 1, \ldots, m$. If $(\tau_1^*(s), \ldots, \tau_m^*(s)) \prec (\tau_1^*(r), \ldots, \tau_m^*(r))$, the corresponding failure replacement regions are nested. Suppose both $(\tau_1^*(s), \ldots, \tau_m^*(s))$ and

$(\tau_1{*}(r), \ldots, \tau_m{*}(r))$ are in A and $\tau_i{*}(s) < \tau_i{*}(r)$, $i = 1, \ldots, m$. Then, it follows from Theorem 5.1 that with probability 1, for all $n_1, n_2, \ldots, n_m$ large enough, the estimated policies $\hat{\tau}(s) \prec \hat{\tau}(r)$, and thus their corresponding failure replacement regions will be nested.

B. SMALL-SAMPLE BEHAVIOR

1. General Simulation Results

We use simulation to gain insight into the behavior of the estimated cost function and policy for small sample sizes. In this simulation, devices have "low," "medium," or "high" rates of use, corresponding to usage paths of slope $\theta_1 = 1$, $\theta_2 = 2$ or $\theta_3 = 5$. For each path, lifetimes arise from the Weibull distribution, with density

$$f(t;\beta,\varphi) = \frac{\beta}{\varphi}\left(\frac{t}{\varphi}\right)^{\beta-1} \exp\left(-\left(\frac{t}{\varphi}\right)^{\beta}\right), t > 0. \tag{5.2}$$

As in the simulations of Ingram and Scheaffer (1976) we fix the shape parameter $\beta = 2$ for each path. Gertsbakh and Kordonsky (1998) also assume the Weibull shape parameter is constant over paths. The scale parameter φ is varied for the three paths so that $\varphi_1 = 40/21$, $\varphi_2 = 10/7$, $\varphi_3 = 1$ for paths 1, 2, and 3, respectively. These scale parameters ensure $(\tau_1{*}, \tau_2{*}, \tau_3{*})$ lies in A for any $r > 0$.

Four groups of simulations are performed to investigate the small-sample behavior of $\hat{C}(\tau)$ and $\hat{\tau}$ as sample sizes along paths $\boldsymbol{n} = (n_1, n_2, n_3)$, mixing probabilities $\boldsymbol{p} = (p_1, p_2, p_3)$, and cost ratio r vary. Each group corresponds to realistic settings for $\boldsymbol{n}$

and $\boldsymbol{p}$. There are three runs within each group, to investigate the effects of varying r. Table 5.1 depicts the settings used in each run.

	Run	n	p	r
Group 1	1	(5,5,5)	(1/3,1/3,1/3)	1.0
	2	(5,5,5)	(1/3,1/3,1/3)	0.5
	3	(5,5,5)	(1/3,1/3,1/3)	0.1
Group 2	4	(5,5,5)	(0.1,0.8,0.1)	1.0
	5	(5,5,5)	(0.1,0.8,0.1)	0.5
	6	(5,5,5)	(0.1,0.8,0.1)	0.1
Group 3	7	(10,10,10)	(1/3,1/3,1/3)	1.0
	8	(10,10,10)	(1/3,1/3,1/3)	0.5
	9	(10,10,10)	(1/3,1/3,1/3)	0.1
Group 4	10	(10,10,10)	(0.1,0.8,0.1)	1.0
	11	(10,10,10)	(0.1,0.8,0.1)	0.5
	12	(10,10,10)	(0.1,0.8,0.1)	0.1

Table 5.1: Settings for General Simulation Runs

Sample sizes of 5 and 10 are common, particularly in observational data or experiments designed to study the lifetime of high-cost prototypic devices. Mixing probabilities (1/3, 1/3, 1/3) represent populations for which devices are evenly spread across several usage rates; and mixing probabilities (0.1, 0.8, 0.1) represent populations for which a large majority of the devices have a "medium" rate of use (e.g., automobiles). Table 5.1 contains runs for which the relative frequencies of the sample sizes along paths differ from the mixing probability vector since it is not uncommon for the mixture of test assets to differ from the mixture in the actual population. Finally, the cost ratios 1, 0.5, and 0.1 are common in the literature.

Each run of the simulation consists of 200 replications. In replication j, we generate a data set consisting of n_i Weibull$(2,\varphi_i)$ lifetimes, $i = 1,2,3$ and for this data set we find $\hat{\tau}^{(j)}$ corresponding to the given p and r using the procedure described in Chapter IV. The random number seed is set in advance for replicability. For each run, we compute several quantities to gain insight into the small-sample performance of $\hat{\tau}$ as an estimator of τ^*. Table 5.2 contains τ^*, the minimizer of the true cost function $C(\tau)$, found numerically. It also lists $Av(\hat{\tau}) = (1/200)\sum_{j=1}^{200} \hat{\tau}^{(j)}$, an estimate of the expected value of $\hat{\tau}$ and the difference $Av(\hat{\tau}) - \tau^*$, an estimate of the bias of $\hat{\tau}$. Finally, it includes $p(\hat{\tau})$, the proportion of the replications for which $\hat{\tau} = (\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3)$; this quantity reveals how often $\tilde{\tau} \in A$ and hence we find $\hat{\tau}$ "automatically," with minimal computation.

	τ^*			$Av(\hat{\tau})$			$Av(\hat{\tau}) - \tau^*$			$p(\hat{\tau})$
1	2.078	1.558	1.091	2.005	1.440	0.988	-0.073	-0.118	-0.103	0.225
2	1.406	1.054	0.738	1.471	1.048	0.713	0.066	-0.007	-0.025	0.260
3	0.607	0.456	0.319	0.866	0.607	0.407	0.258	0.151	0.088	0.210
4	2.078	1.558	1.091	2.180	1.464	0.973	0.102	-0.094	-0.118	0.225
5	1.406	1.054	0.738	1.607	1.078	0.706	0.201	0.024	-0.032	0.260
6	0.607	0.456	0.319	0.923	0.634	0.424	0.316	0.179	0.105	0.210
7	2.078	1.558	1.091	2.121	1.545	1.035	0.044	-0.013	-0.056	0.250
8	1.406	1.054	0.738	1.466	1.064	0.747	0.061	0.010	0.009	0.330
9	0.607	0.456	0.319	0.753	0.538	0.357	0.146	0.082	0.038	0.225
10	2.078	1.558	1.091	2.301	1.601	1.031	0.223	0.043	-0.060	0.250
11	1.406	1.054	0.738	1.514	1.054	0.725	0.108	-0.001	-0.013	0.330
12	0.607	0.456	0.319	0.768	0.551	0.364	0.160	0.095	0.045	0.225

Table 5.2: Small-Sample Performance of $\hat{\tau}$

First, by comparing rows 1-6 with rows 7-12 in Table 5.2, we note that increasing sample sizes generally results in an increase in the (estimated) accuracy of $\hat{\tau}$. As expected, increasing sample sizes increases the proportion of replications for which $\hat{\tau} = (\tilde{\tau}_1, \tilde{\tau}_2, \tilde{\tau}_3)$. To investigate the effect of a non-uniform $\boldsymbol{p}$ on $\hat{\tau}$ in small-sample situations, compare rows 1-3 with rows 4-6 and rows 7-9 with rows 10-12. In general, the accuracy of $\hat{\tau}$ decreases slightly, but this effect is reduced as the sample sizes increase. By examining columns 4-6 of the rows within each group, we note the "average" policies are nested.

We proceed as follows to determine if the policies produced in each individual replication of a given run are nested. By setting the random seed, we generate the same lifetimes for each run in the first two groups and in the last two groups. Hence, for example, the estimated policies for replication j of runs 1, 2, and 3 are based on the same random numbers. For a fixed group, let $\hat{\tau}^{(j)}(r)$ denote the estimated policy for cost ratio r given the data for replication j. It can be shown that these policies are nested if $\hat{\tau}^{(j)}(0.1) \prec \hat{\tau}^{(j)}(0.5) \prec \hat{\tau}^{(j)}(1)$. For each of the four groups, we find that nesting occurs in each of the 200 replications.

For each run, we also compute several quantities to gain insight into the small-sample performance of $\hat{C}(\tau)$ as an estimator of the true cost $C(\tau)$. First, we compute $C(\tau^*)$, the exact cost of the true optimal policy, from (4.2). Next, we compute $Av[\hat{C}(\hat{\tau})] = (1/200)\sum_{j=1}^{200} \hat{C}^{(j)}\left(\hat{\tau}^{(j)}\right)$, an estimate of the expected estimated minimum cost of age replacement, and then the sample standard deviation of the $\hat{C}(\hat{\tau})$. We also compute $Av[C(\hat{\tau})] = (1/200)\sum_{j=1}^{200} C\left(\hat{\tau}^{(j)}\right)$, an estimate of the expected true cost at the optimal policy. Finally, we compute $b[\hat{C}(\hat{\tau})] = Av[\hat{C}(\hat{\tau})] - Av[C(\hat{\tau})]$ and

$MSE\left[\hat{C}(\hat{\tau})\right] = (1/200)\sum_{j=1}^{200}\left(\hat{C}^{(j)}\left(\hat{\tau}^{(j)}\right) - C\left(\hat{\tau}^{(j)}\right)\right)^2$, estimates of the bias and MSE of $\hat{C}(\hat{\tau})$ as an estimator of $C(\hat{\tau})$, respectively. These quantities are scaled by the factor $1/C$ and displayed in Table 5.3.

	$C(\tau^*)$	$Av[\hat{C}(\hat{\tau})]$	sd$[\hat{C}(\hat{\tau})]$	$Av[C(\hat{\tau})]$	$b[\hat{C}(\hat{\tau})]$	MSE$[\hat{C}(\hat{\tau})]$
1	1.618	1.481	0.272	1.679	-0.199	0.107
2	1.095	0.904	0.202	1.150	-0.246	0.094
3	0.473	0.272	0.112	0.518	-0.247	0.074
4	1.554	1.420	0.351	1.623	-0.204	0.142
5	1.052	0.860	0.249	1.110	-0.251	0.112
6	0.454	0.256	0.145	0.513	-0.257	0.084
7	1.618	1.516	0.183	1.654	-0.139	0.052
8	1.095	0.960	0.144	1.133	-0.173	0.049
9	0.473	0.318	0.090	0.503	-0.185	0.042
10	1.554	1.465	0.242	1.597	-0.132	0.071
11	1.052	0.921	0.178	1.094	-0.172	0.058
12	0.454	0.300	0.115	0.490	-0.191	0.047

Table 5.3: Small-Sample Performance of $\hat{C}(\hat{\tau})$

As in Table 5.2, by comparing rows 1-6 with rows 7-12 of Table 5.3, we note that increasing the sample sizes results in an increase in the (estimated) accuracy and precision of $\hat{C}(\tau)$ as an estimator of $C(\tau)$. To investigate the effect of a non-uniform $\boldsymbol{p}$ on $\hat{C}(\tau)$, compare rows 1-3 with rows 4-6 and rows 7-9 with rows 10-12. In general, the accuracy and precision of $\hat{C}(\tau)$ decreases slightly, but this effect is reduced as the sample sizes grow.

2. Results of Nesting Simulation

We also use simulation to investigate in more detail the nesting tendency of the policies produced by our procedure. In the general simulation, we used the sequence of cost ratios {1, 0.5, 0.1}; in this simulation we use a more refined sequence {1, 0.9, ...,0.1}. We retain the same slopes and Weibull parameters as in the general simulation. The nesting simulation consists of 4 runs of 20 replications each; for each replication we use a new random number seed. To investigate the effect of sub-sample size and mixing probability on nesting, we vary $\boldsymbol{n}$ and $\boldsymbol{p}$ between runs. The settings for $\boldsymbol{n}$ and $\boldsymbol{p}$ for the four runs coincide with the settings in groups 1-4 in Table 5.1 (i.e., run 1 has the same settings as in Group 1, and so on). In each replication of a given run, we generate n_i Weibull($2,\varphi_i$) lifetimes, $i = 1,2,3$; for this data set we find $\hat{\tau}^{(j)}(r)$ for each r in {1, 0.9, ...,0.1} and we check whether $\hat{\tau}^{(j)}(0.1) \prec \hat{\tau}^{(j)}(0.2) \prec \cdots \prec \hat{\tau}^{(j)}(1)$. For each run, we find that nesting occurs in each of the 20 replications.

C. COMPARISON WITH MIN CV METHOD

We further use simulation to gain insight into the performance of composite policies estimated using our procedure with in comparison with composite policies estimated using the min CV procedure. Here, we compare the true costs of the policies produced by the two procedures using the sample sizes, mixing probabilities, and cost ratios contained in Table 5.1. As in the general simulation, we use devices with usage paths of slope $\theta_1 = 1$, $\theta_2 = 2$, or $\theta_3 = 5$ and that $X \mid \theta_i \sim$ Weibull($2, \varphi_i$), $i = 1,2,3$. But in this simulation, the scale parameters φ_i correspond with distributions for which the min CV method is expected to return reasonable estimates of $(\tau_1^*, \tau_2^*, \tau_3^*)$.

Unlike our procedure, the min CV method is not designed specifically for the purpose of estimating $(\tau_1^*, \tau_2^*, \tau_3^*)$. Nonetheless, for certain families of conditional distributions, the policy based on the min CV method does in fact estimate $(\tau_1^*, \tau_2^*, \tau_3^*)$. Consider, for example, a population of devices on linear usage paths Z whose lifetimes correspond to the model

$$P(X > x \mid \theta) = \exp\left(-\left(\frac{x + \gamma_o \theta x}{\varphi} \right)^{\beta} \right). \tag{5.2}$$

That is, devices have lifetimes corresponding to the linear ITS model with time scale parameter γ_o. The times in the ITS have a Weibull distribution with shape parameter β and scale parameter φ (ex: Duchesne and Lawless, 2000). It can be shown that along paths we have $X \mid \theta \sim$ Weibull$(\beta, \varphi/(1 + \gamma_o \theta))$. Suppose $\beta = 2$, $\varphi = 4$, and $\gamma_o = 3/5$. It follows that $X \mid \theta_i \sim$ Weibull$(2, \varphi_i)$ where $\varphi_1 = 2.5$, $\varphi_2 = 20/11$, and $\varphi_3 = 1$; these scale parameters are used throughout the study. These scale parameters ensure $(\tau_1^*, \tau_2^*, \tau_3^*)$ lies in A for any $r > 0$.

For a given $r > 0$, our procedure always returns a policy with lower estimated cost than any other policy in (4.3). But since the true τ^* in this simulation corresponds to a triangular policy, and min CV restricts attention to such policies, we would expect the policy based on the min CV scale to have lower actual cost than our estimated policy. We find, though, that our procedure compares favorably in terms of true costs also. The

12 runs of this simulation use the $\boldsymbol{n}, \boldsymbol{p}$, and r as described in Table 5.1; each run of the simulation consists of 200 replications. In a given replication, we generate n_i lifetimes from Weibull($2,\varphi_i$), $i = 1,2,3$. From this data set we compute $\hat{a}$, resulting in $\hat{\tau}_{\mathrm{CV}}$, the policy produced by the min CV method. We also compute $\hat{\tau}$ using our method. Hence, the result of each run are pairs $(\hat{\tau}^{(j)}, \hat{\tau}_{\mathrm{CV}}^{(j)}), j = 1, \ldots, 200$. For each run, we compute $C(\tau)$ at each of these values and (due to occasional non-normality) perform a Wilcoxon signed-rank test on the differences $C(\hat{\tau}_{\mathrm{CV}}^{(j)}) - C(\hat{\tau}^{(j)}), j = 1, \ldots, 200$. For every run we reject the null hypothesis that the true mean difference is non-positive; approximate p-values are 0 in each case. In fact, our estimator results in a lower-cost policy in 67% to 85% of the 200 replications for each run.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. POLICIES GIVEN DATA FROM UNKNOWN USAGE PATHS

Assume that (X,Y) has support $\mathcal{X} = (0,\infty)^2$ and that usage paths are unknown. Unlike the setting with known linear usage paths, there is no natural way to write the cost function in terms of one-dimensional cost functions and still be able to compute the cost for any policy M in $\mathcal{M}_{\mathcal{X}}$. Approaches that use combined scales reduce the cost function to a one-dimensional cost function in the combined scale, but they do so by restricting policies to classes of nested policies. Combined scale approaches do not lend themselves to comparison of policies that are not nested. In this chapter, we develop a cost function that is a natural generalization of the one-dimensional cost function (1.1) and can be applied to all policies in $\mathcal{M}_{\mathcal{X}}$.

In the single-scale problem, the cost function (1.1) has the interpretation "long-run average cost per unit of time in use," and arises in a relatively natural way from univariate renewal theory. Under a joint model for (X,Y), it seems reasonable to consider a cost function of the same nature as (1.1), with interpretation "long-run average cost per unit of time in use," where "time in use" can be measured in chronological time or usage (e.g., flight hours or landings). In practice, budgets are often made with respect to chronological time, rather than usage. With this in mind, the cost function we develop has dimension cost per unit of use in chronological time. It does, however, incorporate both scales and could easily be taken to be cost per unit of usage.

As in previous chapters, we consider policies M in $\mathcal{M}_{\mathcal{X}}$ under which a device is replaced upon failure or when its usage path crosses the boundary of M, whichever occurs first. We develop the two-dimensional renewal reward process as the foundation on which we base the cost function for policies in $\mathcal{M}_{\mathcal{X}}$. For a given set of failure times $(x_1,y_1),\ldots,(x_n,y_n)$, we then demonstrate how to estimate an optimal rectangular policy in $\mathcal{M}_{\mathcal{X}}$, and conclude with an example.

A. THE TWO-DIMENSIONAL RENEWAL REWARD PROCESS

The cost function that we develop arises from considering renewal reward processes (see Appendix A) in two dimensions. Let R(u,v) denote the rectangle $[0,u] \times [0,v]$ and $u > 0$, $v > 0$. A stochastic process $\{N(u,v);\, u > 0,\, v > 0\}$ is said to be a *two-dimensional counting process* if $N(u,v)$ represents the total number of events that have occurred in R(u,v). Let $\{(U_i,V_i)\}$ be a sequence of independent and identically distributed (iid) non-negative random vectors, and let $S_n^{(1)} = \sum_{i=1}^{n} U_i$ and $S_n^{(2)} = \sum_{i=1}^{n} V_i$. Define $N(u,v) = max\{n\!: S_n^{(1)} \le u,\, S_n^{(2)} \le v\}$. Then $\{N(u,v);\, u > 0,\, v > 0\}$ is also a *two-dimensional renewal process* (e.g., Hunter 1974a). Both $\{U_i\}$ and $\{V_i\}$ define univariate renewal processes. With $N_u^{(1)} = max\{n\!: S_n^{(1)} \le u\}$ and $N_v^{(2)} = max\{n\!: S_n^{(2)} \le v\}$, it is readily seen that $N(u,v) = min\{N_u^{(1)}, N_v^{(2)}\}$. Let R_n denote the reward earned at the n^{th} renewal. Assume the R_n, $n \ge 1$ are iid; note R_n may depend on (U_n,V_n). Let $Z(u,v) = \sum_{n=1}^{N(u,v)} R_n$ represent the total reward earned in R(u,v). Then $\{Z(u,v);\, u > 0,\, v > 0\}$ is a *two-dimensional renewal reward process.*

Now, we generalize the univariate Renewal Reward Theorem. Let $\mu_1 = E[U_1] < \infty$ and $\mu_2 = E[V_1] < \infty$; suppose also $E[R_1] < \infty$. Given a one-dimensional renewal process $\{N(t); t \geq 0\}$ with mean inter-renewal time μ, it is well known that the total number of renewals $N(\infty)$ is infinite (e.g., Ross, 1997). For a two-dimensional renewal process, let $N(\infty,\infty)$ be the number of renewals in a square of infinite size; that is, $N(\infty,\infty) = \lim_{t\to\infty} N(t,t)$. We show that $N(\infty,\infty)$ cannot be finite.

Lemma 6.1: $N(\infty,\infty) = \infty$ with probability 1.

Proof: This proof is a generalization of Ross's (1997, p. 353) proof for the one-dimensional case.

$$\begin{aligned} P\{N(\infty,\infty)<\infty\} &= P\{X_n=\infty \text{ or } Y_n=\infty \text{ for some } n\} \\ &= P\left(\bigcup_{n=1}^{\infty}\{X_n=\infty \text{ or } Y_n=\infty\}\right) \\ &\leq \sum_{n=1}^{\infty} P\{X_n=\infty \text{ or } Y_n=\infty\}=0. \end{aligned}$$

The result follows by complementation.

Given a renewal process $\{N(t); t \geq 0\}$ with mean inter-renewal time μ, it is also well known that $\lim_{t\to\infty} N(t)/t = 1/\mu$, with probability 1 (e.g., Ross, 1997). That is, the rate at which $N(t)$ goes to infinity is the reciprocal of the mean inter-renewal time, with probability 1. The following result considers the rate at which a two-dimensional renewal process goes to infinity.

Lemma 6.2: $\lim_{t\to\infty} N(t,t)/t = 1/\max\{\mu_1,\mu_2\}$, with probability 1.

Proof: For any fixed t, $N(t,t) = \min\{N_t^{(1)}, N_t^{(2)}\}$. Also, for fixed t, $\min\{N_t^{(1)}, N_t^{(2)}\}/t = \min\{N_t^{(1)}/t, N_t^{(2)}/t\}$. Since $\lim_{t\to\infty} N_t^{(1)}/t = 1/\mu_1$ and $\lim_{t\to\infty} N_t^{(2)}/t = 1/\mu_2$ with probability 1, it follows that $\lim_{t\to\infty} \min\{N_t^{(1)}/t, N_t^{(2)}/t\} = \min\{1/\mu_1, 1/\mu_2\}$ with probability 1.

Next, we generalize the Renewal Reward Theorem.

Theorem 6.1: $\lim_{t\to\infty} Z(t,t)/t = E[R_1]/\max\{\mu_1,\mu_2\}$ with probability 1.

Proof: Decompose $Z(t,t)/t$ as the product of $\sum_{n=1}^{N(t,t)} R_n / N(t,t)$ and $N(t,t)/t$. By Lemma 6.1 and the Strong Law of Large Numbers the first term goes to $E[R_1]$ with probability 1. By Lemma 6.2 the second term goes to $1/\max\{\mu_1,\mu_2\}$ with probability 1.

B. DEVELOPMENT OF COST FUNCTION FOR TWO-SCALE POLICIES

We must modify the above results slightly before they can be applied to the setting in which the components of the two-dimensional inter-renewal times $\{(U_i,V_i)\}$ are measured in different scales. In the case of two parallel time scales, the time units of the mean inter-renewal times in the denominator are not directly comparable. However, if we "convert" time in the usage scale (e.g., landings) to chronological time, we obtain a meaningful denominator. To this end, we prove a corollary to the theorem.

Corollary 6.1: For $a > 0$, $b > 0$, $\lim_{t\to\infty} Z(at,bt)/t = E[R_1]/\max\{\mu_1/a, \mu_2/b\}$ with probability 1.

Proof: From $\{(U_i,V_i\})$ form the new renewal process $\{(W_i,Z_i)\}$ where $W_i = U_i/a$ and $Z_i = V_i/b$. Let $T_n^{(1)} = \sum_{i=1}^n W_i = S_n^{(1)}/a$ and $T_n^{(2)} = \sum_{i=1}^n Z_i = S_n^{(2)}/b$. Let $N(t,t)' = \max\{n : T_n^{(1)} \le t, T_n^{(2)} \le t\}$. As $E[W_i] = \mu_1/a$ and $E[V_i] = \mu_2/b$, we have $\lim_{t\to\infty} N(t,t)'/t = 1/\max\{\mu_1/a, \mu_2/b\}$ with probability 1 from Lemma 6.2. But

$$\begin{aligned} N(t,t)' &= \max\{n : S_n^{(1)} \le at, S_n^{(2)} \le bt\} \\ &= N(at,bt). \end{aligned}$$

This line of reasoning is essentially identical to Hunter's derivation of the limiting growth rate of $E[N(at,bt)]$ (1974b, pp. 555-6). The result follows immediately, using this fact and the decomposition technique from the proof of Theorem 6.1.

Now we are positioned to use the results and discussion above to develop the function with which we can compute the cost for a given member M of set $\mathcal{M}_{\mathcal{X}}$. Consider the one-dimensional case in which a device has lifetime X and operates under the age replacement policy $(0,\tau)$. Recall the interpretation of the objective function (1.1): the long-run average cost per unit of "time in use" of implementing policy $(0,\tau)$. Here, the "time in use" corresponding to lifetime X is simply the replacement time $\min\{X,\tau\}$. Now, consider the two-dimensional case in which a device has lifetime (X,Y) and operates under policy $M \in \mathcal{M}_{\mathcal{X}}$. We seek an objective function with a similar interpretation, but now "time in use" is more problematic. Let (U,V) denote the replacement time under policy M. We consider two cases. First, suppose $(X,Y) \in M$. This means that the device failed before crossing the boundary of M, so clearly $(U,V) = (X,Y)$. Thus, its "time in use" is $U = X$ and $V = Y$, and its two-dimensional replacement time is simply (X,Y). Second, suppose $(X,Y) \notin M$. We know the device begins its life at (0,0). As it ages, it traces out a usage path terminating at (X,Y), which, by assumption, lies outside of M. At some point, its usage curve crossed the boundary of M. Had policy M been in place, its "time in use" in both scales would be the point at which the usage path crossed the boundary of M. But by assumption we only know (X,Y) and M, not its usage path. Since usage paths are often approximated by a straight line, we adopt the following convention: let (U,V) be the point of intersection of the boundary of M and the chord connecting (0,0) to (X,Y). We describe (U,V) in either case as follows:

$$\begin{aligned} U &= \sup\{x \le X : (x, (Y/X)x) \in M\}, \text{ and} \\ V &= (Y/X)U. \end{aligned} \tag{6.2}$$

We now construct the two-dimensional renewal reward process for a device operating under policy $M \in \mathcal{M}_{\mathcal{X}}$. We are given two-dimensional failure times (X_1,Y_1), (X_2,Y_2), … iid from some bivariate lifetime distribution F; thus $\{(U_i,V_i)\}$ are iid. Let $\mathrm{R}(u,v)$ denote $[0,u] \times [0, v]$. Let $N(u,v)$ represent the total number of replacements made in $\mathrm{R}(u,v)$. Since the $\{(U_i,V_i)\}$ are iid, $\{N(u,v);\ u > 0,\ v > 0\}$ is a two-dimensional renewal process. As in the one-dimensional case, let the "reward" (i.e., cost for replacement) R be K if replaced due to age and $(K + C)$ if replaced due to failure. Let $Z(u,v)$ represent the total cost incurred in $\mathrm{R}(u,v)$. Then, $\{Z(u,v);\ u > 0,\ v > 0\}$ is a two-dimensional renewal reward process, with inter-renewal times $\{(U_i,V_i)\}$, rewards $R_i = K + CI[\,(X_i, Y_i) \in M\,]$, and $Z(u,v)=\sum_{i=1}^{N(u,v)} R_i$. Recall the cost of policy $(0,\tau)$ in the one-dimensional case is $C(\tau) = \lim_{t\to\infty} Z(t)/t = E[R_1]/E[U_1]$, as discussed in Appendix A. To obtain a similar limiting result for the situation we have just described, we apply Corollary 6.1. Thus, let $a = 1$ and $b = E[Y]/E[X]$; let $\mu_1(M) = E[U]$ and $\mu_2(M) = E[V]$. From Corollary 6.1,

$$C(M) = \lim_{x\to\infty} Z(x,bx)/x = E[R_1]/\max\{\mu_1(M),\mu_2(M)/b\}, \tag{6.3}$$

with dimension cost per unit of chronological time. The coefficient b in (6.3) is motivated from the "conversion factor" used by Kordonsky and Gertsbakh (1994), and can be interpreted as follows. From a reliability standpoint, one unit of usage is "worth" $E[X]/E[Y]$ units of chronological time, on average.

To "solve" the multiple-scale age replacement problem in this setting, we must find the M^* in $\mathcal{M}_{\mathcal{T}}$ which minimizes this expression. We now demonstrate how to solve the appropriate optimization problem for a specific subset of $\mathcal{M}_{\mathcal{T}}$.

C. FINDING THE BEST RECTANGULAR POLICY

The aim of this section is to search over the set $\mathcal{M}_{\mathrm{R}} = \{\mathrm{R}(s,t)\colon s > 0, t > 0\}$, the set of all "lower rectangular" policies $(0,s) \times (0,t)$. Observe $\mathcal{M}_{\mathrm{R}} \subseteq \mathcal{M}_{\mathcal{T}}$. The set of lower rectangles is attractive since rectangular policies are easily implemented: a device is replaced upon failure or when its elapsed chronological time or cumulative usage reaches some "limit." Hence, rectangular policies are closely akin to automobile warranties. In this section, we derive the form of the cost function for a given rectangle and describe the minimizer of the cost function formed when F is estimated by the empirical distribution on the bivariate data. For the same reasons as in the univariate problem, it is convenient to define $F(x,y) = \mathrm{P}(X < x, Y < y)$ for (x,y) in $\mathcal{X}$. We now calculate the numerator and denominator in (6.3) for $C(s,t)$, the cost when $M = \mathrm{R}(s,t)$.

We find the numerator of $C(s,t)$ in a manner similar to the single-scale case. Define reward R by

$$R = \begin{cases} K + C & \text{if } (X,Y) \in (0,s)\times(0,t) \\ K & \text{if } (X,Y) \notin (0,s)\times(0,t) \end{cases} \tag{6.4}$$

Thus, the numerator is $E[R] = (K + C)\,F(s,t) + K\,(1 - F(s,t)) = K + C\,F(s,t)$.

To compute the denominator, let $\mu_1(s,t) = E[U]$ and $\mu_2(s,t) = E[V]$, where U and V are defined as in (6.2). For a fixed (s,t) in $\mathcal{X}$, let $A_1(s,t) = (0,s) \times (0,t)$, $A_2(s,t) = \{(x,y) \in \mathcal{X}\colon\ y \geq t \text{ and } y \geq (t/s)x\}$, and $A_3(s,t) = \{(x,y) \in \mathcal{X}\colon\ x \geq s \text{ and } y < (t/s)x\}$. In what follows the parameters (s,t) are omitted from these sets to simplify notation. Observe that these regions form a partition of $\mathcal{X}$. From (6.2), we find

$$U = \begin{cases} X, & \text{if } (X,Y) \in A_1 \\ tX/Y, & \text{if } (X,Y) \in A_2\,. \\ s, & \text{if } (X,Y) \in A_3 \end{cases} \tag{6.5}$$

Thus,

$$\mu_1(s,t) = \iint_{A_1} x\,dF(x,y) + \iint_{A_2} (tx/y)\,dF(x,y) + \iint_{A_3} s\,dF(x,y)\,. \tag{6.6}$$

Similarly,

$$V = \begin{cases} Y, & \text{if } (X,Y) \in A_1 \\ t, & \text{if } (X,Y) \in A_2\,, \\ sY/X, & \text{if } (X,Y) \in A_3 \end{cases} \tag{6.7}$$

and it follows that

$$\mu_2(s,t) = \iint_{A_1} y dF(x,y) + \iint_{A_2} t dF(x,y) + \iint_{A_3} (sy/x) dF(x,y). \tag{6.8}$$

Assembling the parts, we find that the cost when $M = \mathrm{R}(s,t)$ is

$$C(s,t) = \frac{K + C\,F(s,t)}{\max\{\mu_1(s,t), \mu_2(s,t)/b\}}. \tag{6.9}$$

When F is estimated by a discrete bivariate distribution with mass p_i on $\{(x_i,y_i),$ $i = 1, \dots, n\}$, such as the empirical distribution, $C(s,t)$ is estimated as follows. Let $I_j(i)$ denote the indicator function on set A_j for $i = 1, \dots, n$ and j in 1,2,3. That is,

$$I_j(i) = \begin{cases} 1, & \text{if } (x_i, y_i) \in A_j \\ 0, & \text{otherwise} \end{cases} \tag{6.10}$$

Then, it can be shown that the quantities $E[R]$, $\mu_1(s,t)$, $\mu_2(s,t)$ and b are estimated by

$$\hat{E}[R] = K + C\sum_{i=1}^{n} I_1(i) p_i \,, \tag{6.11}$$

$$\hat{\mu}_1(s,t) = \sum_{i=1}^{n} [x_i I_1(i) + (tx_i/y_i) I_2(i) + s I_3(i)] p_i \,, \tag{6.12}$$

$$\hat{\mu}_2(s,t) = \sum_{i=1}^{n} [y_i I_1(i) + t I_2(i) + (sy_i/x_i) I_3(i)] p_i \,, \text{ and} \tag{6.13}$$

$$\hat{b} = \sum_{i=1}^{n} y_i p_i \Big/ \sum_{i=1}^{n} x_i p_i \,. \tag{6.14}$$

We substitute these into (6.9), obtaining

$$\hat{C}(s,t) = \frac{\hat{E}[R]}{\max\{\hat{\mu}_1(s,t), \hat{\mu}_2(s,t)/\hat{b}\}}. \tag{6.15}$$

Let us now explain how to find the minimizing value of (6.15). Recall that to solve the one-dimensional problem it suffices to evaluate $\hat{C}(\tau)$ in (1.3) at each of the observations. We apply a similar strategy to find the minimizer of $\hat{C}(s,t)$. Because (1.3) and (6.15) are developed in a similar manner, it is tempting to think that to find the minimizer of $\hat{C}(s,t)$, it suffices simply to evaluate (6.15) at (x_i,y_i), $i = 1, \ldots, n$, and select the two-dimensional failure time with the smallest cost. Upon closer examination, we find that it is necessary to evaluate (6.15) at other points in addition to the two-dimensional failure times. Let $\hat{z}$ be such a minimizer, i.e., $\hat{C}(\hat{z}) \leq \hat{C}(s,t)$ for all (s,t) in $\mathcal{X}$. We now describe how to find $\hat{z}$.

For convenience, suppose that no chronological failure times share the same value, so that the chronological failure times can be strictly ordered $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$, and similarly suppose the usage failure times can be ordered $y_{(1)} < y_{(2)} < \ldots < y_{(n)}$. Let $x_{(0)} = 0 = y_{(0)}$ and $x_{(n+1)} = \infty = y_{(n+1)}$. Form a grid

$$\Gamma = \{ x_{(1)}, x_{(2)}, \ldots, x_{(n)} \} \times \{ y_{(1)}, y_{(2)}, \ldots, y_{(n)} \}. \tag{6.16}$$

Note Γ defines a partition of $= (0,\infty)^2$ into rectangles of the form $(x_{(i)},x_{(i+1)}] \times (y_{(j)},y_{(j+1)}]$, i, $j \in \{0, \ldots, n\}$. Let $n(s,t) = \hat{E}[R]$ and $d(s,t) = \max\{ \hat{\mu}_1(s,t), \hat{\mu}_2(s,t)/\hat{b} \}$ from (6.11), (6.12), (6.13) and (6.14). Consider the numerator. Note that $n(s,t)$ is constant on every $(x_{(i)},x_{(i+1)}] \times (y_{(j)},y_{(j+1)}]$, continuous from the left in s for all t, continuous from the left in t for all s, and non-decreasing in both s and t with jumps that can only occur on the north and east boundaries of the $(x_{(i)},x_{(i+1)}] \times (y_{(j)},y_{(j+1)}]$. Consider the denominator. We have $\hat{\mu}_1(s,t) = \sum_{i=1}^{n} q_i(s,t)p_i$, where $q_i(s,t) = x_i I_1(i) + (tx_i/y_i) I_2(i) + sI_3(i)$. It can be shown that $q_i(s,t)$ and hence $\hat{\mu}_1(s,t)$ is continuous and non-decreasing in both s and t. Likewise, $\hat{\mu}_2(s,t) = \sum_{i=1}^{n} r_i(s,t)p_i$, where $r_i(s,t) = y_i I_1(i) + tI_2(i) + (sy_i/x_i) I_3(i)$. It can also be shown that $r_i(s,t)$, and hence $\hat{\mu}_2(s,t)/\hat{b}$, is continuous and non-decreasing in both s and t. Thus $d(s,t)$ is continuous and non-decreasing in s and t.

On each $(x_{(i)},x_{(i+1)}] \times (y_{(j)},y_{(j+1)}]$, the ratio $n(s,t)/d(s,t)$ is thus continuous and non-increasing in s and t and therefore has minimum value at $(x_{(i+1)}, y_{(j+1)})$. By a careful examination of the cost function it can be shown that $\hat{C}(x_i,y_{(n)}) \le \hat{C}(x_i,y_{(n)} + y)$, $i = 1 \ldots n$ for $y > 0$ and $\hat{C}(x_{(n)},y_j) \le \hat{C}(x_{(n)} + x,y_j)$, $j = 1 \ldots n$ for $x > 0$. As such, it is not necessary to search beyond the outermost point of the grid, namely $(x_{(n)}, y_{(n)})$. These points are gathered into the following result.

Theorem 6.2: Consider the probability distribution which places mass p_i on (x_i,y_i), $i = 1, \ldots, n$. Let $\hat{C}(s,t)$ be defined as in (6.15) and Γ as in (6.16), and $\hat{z} = \text{argmin}\, \hat{C}(s,t)$. Then, $\hat{z} \in \Gamma$.

D. EXAMPLE

Returning to the jet engine and automobile data sets, Table 6.1 contains $\hat{z}$ for the cost ratios $r = 1.0, 0.5$, and 0.1 when F is estimated by the empirical distribution $\hat{F}$. Beneath $\hat{z}$ in each cell is $\hat{F}(\hat{z})$.

	$r = 1$	$r = 0.5$	$r = 0.1$
Jet Engine	(4932,2426) 0.238	(3227,1550) 0.000	(3227,1550) 0.000
Automobile	(330,10300) 0.578	(368,8000) 0.421	(68,8400) 0.053

Table 6.1: Rectangular Policies for Various Cost Ratios. Parenthetical entries in the cells represent the optimal policy corresponding to a particular value of the cost ratio *r*. Beneath each such entry is the value of the empirical distribution at this point.

We make the following observations from Table 6.1. First, as indicated by the values $\hat{F}(\hat{z})$, more conservative policies are selected as r decreases (under more conservative policies, devices have a smaller chance of failure before replacement). However, the policies are not always nested; in particular, for the automobile data, the policy for $r = 0.1$ is not contained in the policy for $r = 0.5$. Also, none of the $\hat{z}$ correspond with observations, thus amplifying the need to evaluate the estimated cost function at all points in the grid Γ. Figure 6.1 depicts the policies for the jet engine data. Note from Table 6.1 that the policy for $r = 0.5$ is identical to the policy for $r = 0.1$, and that this policy is nested within the policy for $r = 1$.

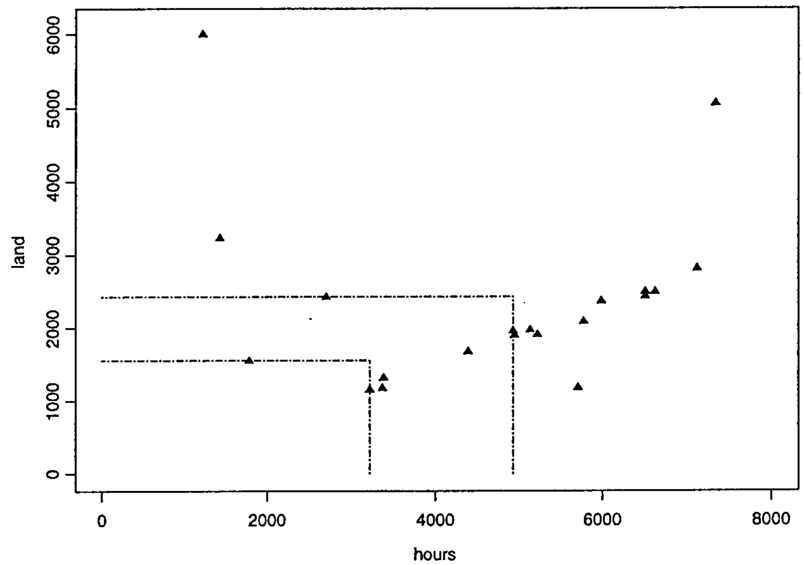

Figure 6.1: Rectangular Policies for Jet Engine Data.
The dashed lines represent the boundaries of the policies for $r = 0.5$ and 1.

E. DISCUSSION AND SUMMARY

In this chapter we developed the two-dimensional renewal reward process, and it served as the foundation on which to build the cost function for policies in $\mathcal{M}_T$ under a joint model for (X,Y). The cost function arises from the analog of the univariate Renewal Reward Theorem, and has dimension cost per unit of chronological time in use, much like (1.1). In the latter half of this chapter, we derived the form of the cost function for rectangular policies and showed how to find the rectangular policy with lowest cost given a set of bivariate failure data. The notions developed in this chapter are easily extended to policies based on more than two scales.

We do not claim the policy $\hat{z}$ produced by this procedure is an estimate of a true optimal z^* for the underlying F. Unlike the case of several linear paths, we have yet to find examples of non-trivial bivariate distributions for which an optimal z^* or an equivalence class of such policies exists. The closest work in the literature is that of Murthy et al (1995) in which the parameters of the optimal rectangular warranty policy are found for certain named bivariate distributions, but the cost functions used to define "optimal" are very different in nature from ours. Perhaps certain bivariate notions of aging (e.g., bivariate IFR, etc.) can be used to identify distributions for which a z^* exists. Also, under additional conditions, it may be possible to show that $\hat{z}$ converges to z^*. If such distributions can be identified, simulation studies can be conducted to verify the small-sample properties of $\hat{z}$.

VII. CONCLUSIONS

In this dissertation, we generalize the classical age replacement policy to the case in which the age of a device is recorded in more than one time scale. We use several case studies to motivate the form of a general replacement policy in multiple scales. The case studies demonstrate the need for careful consideration in developing such policies. In the first two, we notice that in some situations, simply ignoring the usage scale may not be problematic, but in others, failure times in one single scale (e.g., chronological time) may not capture the entire damage accumulation process. The third case study reveals that a naïve (though seemingly sensible) approach for data lying along linear paths can result in a policy that, although "optimal" from the standpoint of (estimated) costs, is not sensible from the standpoint of implementation. Based on these observations, we describe a class of policies that are sensible from the standpoint of implementation. This class generalizes multiple-scale policies found in the literature. Furthermore, we find it is desirable for multiple-scale policies to be nested when considering (in sensitivity analyses, for example) a decreasing sequence of cost ratios; otherwise, the replacement times prescribed by the policies can be inconsistent with the interpretation of the cost ratios.

When failure times are recorded in multiple scales, it becomes readily apparent that identical devices do not operate under identical field conditions. Researchers are grappling with ways to use such lifetime data to produce comprehensive models, and some are seeking to use these models in the arena of optimal preventive maintenance.

Methods for developing preventive maintenance policies for such devices fall on a continuum ranging between two extremes. One extreme, as noted by Kordonsky and Gertsbakh (1997) is to provide an individualized policy for every single device in the population. They note such an approach is totally impractical and, as a result, unacceptable. The other extreme is the "one-size-fits-all" approach, in which the "optimal" policy is based on fitting a single distribution to observations which, in reality, may come from a mixture; this policy is then applied to the entire population. Basing a policy on a combined scale falls in between these extremes in that data in two scales are modeled by a univariate distribution in some "optimal" scale. As expressed by Kordonsky and Gertsbakh (1997), the goal of such approaches is to find a scale in which maintenance actions can be described "in a unified way which would fit all exemplars and would cover all operational conditions." We carefully examine policies based on combined scales arising from three approaches in the literature in light of "desirable" properties. We find that each of the three approaches lacks features important when developing multiple-scale policies. In one approach, the observations are translated into many different scales and the scale corresponding to the minimum value of a "converted" cost function is defined to be "best." This approach, although motivated from the standpoint of minimizing costs, does not guarantee nested policies in the original scales. In the second approach, a combined scale is found in a manner unrelated to maintenance costs. Policies based on this scale have the same "shape" and are nested. The third approach also restricts the form of the policy in a manner unrelated to costs. This approach, although appropriate in some preventive maintenance contexts, does not seem best suited for age replacement.

We consider multiple-scale age replacement in two settings. In the first, since it is common in the literature to approximate unknown usage paths with straight lines, we develop a procedure based on the assumption that devices age along linear paths. Like the scale-combining approaches, our approach lies between the extremes in that it can result in different policies for devices on different usage paths. However, our procedure does not rely on finding an "optimal" scale. Instead, it considers the lifetime distributions corresponding to devices on different paths in a manner resulting in an estimate of the optimal policy among a class of "sensible" policies. We show that under mild conditions, the estimated optimal replacement times are strongly consistent estimators of the true optimal replacement times, and then show by simulation that these estimates are well-behaved in small-sample situations. It is also shown that our procedure tends to produce policies having lower true cost than those based on the min CV method.

In the second setting, device usage paths are unknown. We define the two-dimensional renewal reward process, and prove a two-dimensional version of the Renewal Reward Theorem. Using this result, we develop the cost function by which we can evaluate various policies under the assumption of a joint model for the bivariate failure times. We also derive the form of the cost function for a smaller class of alternatives and present numerical results obtained from solving the corresponding optimization problem for various two-dimensional failure data sets.

We note that our contributions may seem to fall in the area known as "multivariate age replacement." The literature in this realm, however, differs significantly from ours. In this literature, "multivariate age" refers to the ages of several components, where age is measured in a single scale. For example, Ebrahimi (1997) defines MAR($T_1, \ldots, T_k$), the policy for multivariate age replacement for a system of k components which replaces component i, $i = 1, 2, \ldots, k$ either at age T_i or upon its failure. For the case $k = 2$, Ebrahimi explains how to find the optimal MAR(T,T) for both series and parallel systems. Heinrich and Jensen (1996) also discuss optimal replacement in a two-component parallel system, as does Scheaffer (1975).

Numerous extensions to the dissertation research present themselves. Throughout this dissertation our main focus has been on data consisting of ordered pairs representing the chronological age at failure and the cumulative usage at failure. In some cases (e.g., the aircraft wing joint we mention in the Introduction) more than one measure of usage may be available; in other cases, values of other external covariates thought to impact the failure process may be available. The concept of a lower set generalizes to higher dimensions, and the problem of incorporating additional external covariates into policy estimation is worthy of consideration. In fact, as noted in the Introduction, the definition of time scale is general enough to include such cases. In the single-scale realm, Love and Guo (1991) and Kumar and Westberg (1997) present methods for obtaining age replacement policies for a pressure gauge given covariate information (the data set can be found in Appendix B). Both of these use a parametric model to incorporate the effect of the covariate on gauge lifetime. The work of Makis and Jardine (1992, 1999) in the single-scale realm is more comprehensive. They recommend a combination of age replacement and "condition-based" replacement in hopes of obtaining replacement decisions that are more accurate than by employing one approach or the other. The foundation of their work is the Cox proportional hazard model (PHM) with time-dependent covariates. Given a data set of the form considered in this dissertation, we can obtain (in concept) a multiple-scale replacement policy by treating the measurements from the second time scale as the time-dependent covariate. Duchesne (1999), however, remarks that "because models with covariates treat the time variable and the covariates quite asymmetrically, it is not recommended to choose an arbitrary scale as the main scale and the other scale as covariates." Farewell and Cox (1979) issue a similar warning. Of course, one can conceive of a situation where a wealth of information is available at device failure, including time in various scales and numerous condition measurements (some of which may be interval covariates such as measures of wear). In such cases, we echo Duchesne's (1999) call for methods for the systematic identification of information categories for inclusion in models for device failure.

The procedure developed in Chapter IV relies on the assumption that for a given $r > 0$, the collection of conditional distributions $\{F_i\}$ has unique and finite $(\tau_1^*, \tau_2^*, \ldots, \tau_m^*) \in A$. Further investigation is needed to characterize families with this property. This would provide a means for checking model assumptions before applying the procedure. We note that stochastic ordering (or even the stronger failure rate ordering) of the conditional lifetimes is not sufficient to guarantee $(\tau_1^*, \tau_2^*, \ldots, \tau_m^*) \in A$.

In addition, numerous extensions were made to the basic problem with cost function (1.1) in the years following its initial development, as noted in the surveys by McCall (1965), Pierskalla and Voelker (1976), and Valdez-Flores and Feldman (1989). Such extensions as cost discounting, imperfect repair, and others are also viable research topics for the multiple-scale problem.

The cost function (4.2) by which we define the "optimal" composite policy is of the "average of cost functions" form considered by Gertsbakh and Kordonsky (1997). Letting R denote the "reward" (cost) of a replacement and U the replacement time, the estimation of the optimal policy based on a "true" reward functional of the form $E[R]/E[U]$ for the linear path case would also be a worthwhile pursuit. Here $E[R]$ and $E[U]$ could be found by a conditioning approach (e.g., Ross, 1997). This function has a slightly different interpretation than the one in (4.2), and is closely related to (6.3).

Finally, we note much can be built on the foundation created in Chapter VI, where we focus on non-parametric policy estimation for the case in which observations do not fall on linear paths. For example, we concentrate specifically on rectangles. While such policies are easily implemented, it is conceivable that other members of $\mathcal{M}_{\mathcal{X}}$ may result in lower cost than the "best" rectangular policy (if it exists) for a given F. For instance, for some F, the class of policies bounded by the quantile curves of F may be worthy of consideration. Under such a policy (much like the policies based on an ideal time scale) the probability of failure before replacement would be identical for devices on any usage path. However, implementation may be difficult due to the shape of such a policy. It may also be fruitful to consider clustering methods for the case of unknown usage paths.

In such an approach, observations (X,Y) could be clustered by their (estimated) usage rate Y/X and then projected onto the line with slope corresponding to their respective cluster center. With the data in this form, the techniques of Chapter IV could then be applied to the "projected" data. A similar approach was suggested by Duchesne (1999) for non-parametric estimation of the ITS.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A: RENEWAL THEORETIC DEFINITIONS AND DERIVATION OF COST FUNCTION

A. DEFINITIONS

The following renewal theoretic definitions are from Ross (1997). A stochastic process $\{N(t); t \geq 0\}$ is a *counting process* if $N(t)$ represents the total number of events that have occurred up to time t. Let $\{N(t); t \geq 0\}$ be a counting process and let X_n denote the time between the $(n-1)^{st}$ and n^{th} event of this process, $n \geq 1$ (henceforth these times will be called "inter-renewal times"). If the inter-renewal times $\{X_n\}$ are independent and identically distributed (iid), the counting process $\{N(t); t \geq 0\}$ is a *renewal process*; a "renewal" has taken place when an event has occurred. Given a renewal process $\{N(t); t \geq 0\}$ with inter-renewal times $\{X_n\}$, let R_n denote the reward earned at the time of the n^{th} renewal. Assume the R_n, $n \geq 1$ are iid; R_n may depend on X_n. Let $Z(t) = \sum_{n=1}^{N(t)} R_n$ represent the total reward earned up to time t; $\{Z(t); t \geq 0\}$ is a *renewal reward process.*

B. DERIVATION OF SINGLE-SCALE COST FUNCTION

Consider a device which is maintained under an age replacement policy; that is, the device is replaced upon failure or when it reaches age τ, whichever comes first (assume the replacement time is negligible). For example, consider a large supply of identical light bulbs. Upon failure, a light bulb is replaced instantly; operating conditions remain identical from one light bulb to the next. Assume replacement devices are as good as new. Let X_n denote the lifetime of the n^{th} device; assume $X_1, X_2, \ldots$ are iid with distribution function F and survivor function S. For simplicity, assume F is absolutely continuous with density f; Nakagawa and Osaki (1977) discuss the discrete version of this problem. Let $U_n = \min\{X_n, \tau\}$ denote the time between the $(n-1)^{\text{st}}$ and n^{th} replacement; assume a replacement has occurred at time 0. Let $N(t)$ denote the number of replacements to occur in $(0, t]$; by the assumptions made thus far $\{N(t); t \geq 0\}$ is a counting process with times between events iid and is therefore a renewal process. Suppose the cost for replacement is $K > 0$ if replaced due to age (i.e., preventively) and $(K + C)$ if replaced due to failure (assume $C > 0$; this indicates the costly nature of a replacement during operation). Let $Z(t)$ denote the total cost incurred in $(0, t]$; $\{Z(t); t \geq 0\}$ is a renewal reward process with inter-renewal times $\{U_n\}$, where $U_n = \min\{X_n, \tau\}$, $R_n = K + CI[X_n < \tau]$, and $Z(t) = \sum_{n=1}^{N(t)} R_n$. Ross (1997) proves that if $E[R_1] < \infty$ and $E[U_1] < \infty$, the long-run average cost per unit of time in use is $\lim_{t\to\infty} Z(t)/t = E[R_1]/E[U_1]$ with probability 1. If we say a "cycle" is completed every time a replacement occurs, this limit is the "expected reward per cycle" over the "expected cycle length." We now compute $E[R_1]$ and $E[U_1]$. Since $R_1 = K + CI[X_1 < \tau]$, we find $E[R_1] = K + C\,F(\tau)$. Since $U_1 | X_1 = X_1 I(X_1 < \tau) + \tau I(X_1 \geq \tau)$, we find $E[U_1] = \int_0^\tau t f(t)\,dt + \int_\tau^\infty \tau f(t)\,dt$, which reduces to $\int_0^\tau S(u)\,du$. Thus, the long-run average cost per unit of time in use as a function of τ is (1.1).

C. SINGLE-SCALE COST FUNCTION AND SCALE FAMILIES

The following lemma shows that (1.1) behaves "as expected" in scale families.

Lemma A.1 (Optimal Replacement Time Ordering in Scale Families): Let Z and Y denote lifetimes from distributions F_Z and F_Y, respectively, where $Z = aY$, with $a > 0$. Let K and $C > 0$. Let τ_Z^* and τ_Y^* minimize (1.1) when $F = F_Z$ and F_Y, respectively. Then, $\tau_Z^* = a\tau_Y^*$.

Proof: Let $\tau > 0$. Then, by definition

$$C_Z(\tau) = \frac{K + C\,F_Z(\tau)}{\int_0^\tau S_Z(u)du}.$$

It follows that

$$\begin{aligned} C_Z(\tau) &= \frac{K + C\,F_Y(\tau/a)}{a\left\{\int_0^{\tau/a} S_Y(u/a)(1/a)du\right\}} \\ &= \frac{1}{a}\left\{\frac{K + C\,F_Y(\tau/a)}{\int_0^{\tau/a} S_Y(u)du}\right\} \\ &= \frac{1}{a}C_Y(\tau/a). \end{aligned}$$

But then

$$\begin{aligned} \tau_Z^* &= \arg\min C_Z(\tau) \\ &= \arg\min \frac{1}{a}C_Y(\tau/a) \\ &= \arg\min \quad C_Y(\tau/a) \\ &= a\,\tau_Y^*, \end{aligned}$$

where the last two equalities follow by observing that (1) minima are preserved under vertical shrinking, and (2) minima are scaled upon horizontal stretching.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B: DATA SETS

1. Automobile data. This data set consists of 19 failure times in days since purchase and number of miles driven (to the nearest 100 miles) for a particular automobile component. The data set is taken from Wilson (1993, p. 32). The data are presented in the table below.

Failure	Days	Miles
1	146	3200
2	251	11100
3	251	11100
4	470	14100
5	26	8400
6	330	8500
7	187	6800
8	210	9100
9	368	6500
10	68	1200
11	340	11000
12	384	12400
13	286	8000
14	306	10300
15	105	1900
16	24	1100
17	95	2200
18	101	4200
19	187	2400

Table B.1: Auto Data

2. Metal fatigue data. This data set was discussed in Kordonsky and Gertsbakh (1993, p. 240); a summary of their description of the data set follows. A sample of 30 identical steel specimens was divided into six groups of size five; each group was subjected to a cyclic two-level loading regime until failure. The loading regime for group j was a periodic sequence of 5000 loading cycles consisting of $5000\alpha_j$ cycles of small amplitude (i.e., low load) followed by $5000(1-\alpha_j)$ cycles of large amplitude (i.e., high load), $j = 1,\dots,6$. The table below records the cumulative number of low cycles and high cycles at failure for each specimen, scaled by a factor of 10.

Specimen	α_j	Low/10	High/10	Specimen	α_j	Low/10	High/10
1	0.95	25680	1350	16	0.40	3200	4570
2	0.95	23580	1160	17	0.40	4800	7040
3	0.95	37015	1925	18	0.40	4200	6150
4	0.95	33510	1750	19	0.40	4200	6060
5	0.95	38030	2000	20	0.40	5400	8040
6	0.80	15300	3800	21	0.20	1000	3750
7	0.80	17620	4400	22	0.20	1600	6270
8	0.80	16030	4000	23	0.20	1200	4530
9	0.80	15600	3900	24	0.20	1900	7260
10	0.80	10300	2500	25	0.20	1100	4200
11	0.60	8400	5440	26	0.05	300	5390
12	0.60	8100	5230	27	0.05	375	6855
13	0.60	9000	5990	28	0.05	425	7795
14	0.60	5700	3730	29	0.05	332	5795
15	0.60	6600	4270	30	0.05	275	5125

Table B.2: Metal Data

3. Traction motor data. This data set comes from the railroad industry, and is found in Wilson (1993, p. 31). Table B.3 contains the time since inception of service and mileage at failure of forty locomotive traction motors when they were returned to the depot for maintenance.

i	miles	days		*i*	miles	days
1	9766	166		21	5922	128
2	2041	35		22	1974	31
3	12392	249		23	2030	65
4	9889	190		24	12532	221
5	974	27		25	14796	316
6	1594	41		26	979	22
7	2128	59		27	15062	261
8	2158	75		28	2062	32
9	11187	223		29	16888	397
10	47660	952		30	3099	48
11	13827	335		31	28	1
12	5992	164		32	95	27
13	6925	145		33	12600	295
14	7078	170		34	8067	140
15	7553	140		35	41425	827
16	25014	498		36	105	2
17	25380	571		37	12302	209
18	26433	499		38	447	29
19	16494	340		39	9766	166
20	7162	160		40	57304	1200

Table B.3: Traction Motor Data

4. Jet engine failure data. This data set is discussed in Gertsbakh and Kordonsky (1998, p. 1186) and was obtained from the first author. Table B.4 contains the flight hours and number of landings at failure of 21 jet engines.

i	hours	landings	i	hours	landings
1	1216	6000	12	5136	1974
2	1424	3232	13	5224	1913
3	1784	1550	14	5709	1180
4	2712	2426	15	5777	2088
5	2712	2426	16	5988	2367
6	3227	1152	17	6507	2433
7	3374	1176	18	6509	2496
8	3387	1316	19	6623	2495
9	4391	1676	20	7126	2810
10	4932	1960	21	7343	5057
11	4948	1900			

Table B.4: Jet Engine Data

5. Pressure gauge data. The table below contains the failure (or censoring, if marked by an asterisk) time in hours of 15 pressure gauges and the corresponding covariate value "pressure." The data set is from Love and Guo (1991, p. 14). The implication is that the value of the covariate was fixed during each particular life cycle. Thus, for example, the first entry indicates that "medium" (in some sense) pressures were measured from time 0 until failure at 70 hours.

i	Time (hrs)	Pressure
1	70	4
2	53	4
3	77	4
4	42	4
5	61*	4
6	51	5
7	70	5
8	32	5
9	47	5
10	44*	5
11	101	3
12	66	3
13	198	3
14	95	3
15	60*	3

Table B.5: Pressure Gauge Data

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Aras, G. and Whitaker, L., "A Sequential Scheme to Estimate the Optimal Age Replacement Policy," *Communications in Statistics: Theory and Methods Series A*, vol. 21, no. 8, pp. 2325-2338, 1992.

Arunkumar, S., "Nonparametric Age Replacement Policy," *Sankhya Series A*, vol. 34, pp. 251-256, 1972.

Barlow, R. E. and Campo, R., "The Total Time on Test Processes and Applications to Failure Data Analysis," in *Reliability and Fault Tree Analysis*, SIAM, Philadelphia, 1975, pp. 451-81.

Barlow, R. E. and Proschan, F., *Mathematical Theory of Reliability*, John Wiley, NY, 1965.

Bather, J. A., "On the Sequential Construction of an Optimal Age Replacement Policy," *Bulletin of International Statistical Institute*, vol. 47, pp. 253-266, 1977.

Bergman, B., "Some Graphical Methods for Maintenance Planning," *Proceedings 1977 Annual Reliability and Maintainability Symposium*, pp. 468-471.

Bergman, B. and Klefsjö, B., "A Graphical Method Applicable to Age-Replacement Problems," *IEEE Transactions on Reliability*, vol. R-31, no. 5, pp. 478-81, 1982.

Duchesne, T., "Multiple Time Scales in Survival Analysis," Dissertation, University of Waterloo, Ontario, Canada, 1999.

Duchesne, T. and Lawless, J., "Alternative Time Scales and Failure Time Models," *Lifetime Data Analysis*, vol. 6, no. 2, pp. 157-179, 2000.

Ebrahimi, N., "Multivariate Age Replacement," *Journal of Applied Probability*, vol. 34, pp. 1032-1040, 1997.

Farewell, V. T. and Cox, D. R., "A Note on Multiple Time Scales in Life Testing," *Applied Statistics*, vol. 28, no. 1, pp. 73-5, 1979.

Frees, E. W. and Ruppert, D., "Sequential Nonparametric Age Replacement Policies," *The Annals of Statistics*, vol. 13, no. 2, pp. 650-662, 1985.

Gertsbakh, I. B. and Kordonsky, Kh. B., "Parallel Time Scales and Two-Dimensional Manufacturer and Individual Customer Warranties," *IIE Transactions*, vol. 30, pp. 1181-1189, 1998.

Gertsbakh, I. B. and Kordonsky, K. B., "Choice of the Best Time Scale for Preventive Maintenance in Heterogeneous Environments," *Euro. J. Operational Research*, vol. 98, pp. 64-74, 1997.

Geurts, J. H. J., "Optimal Age Replacement versus Condition Based Replacement: Some Theoretical and Practical Considerations," *Journal of Quality Technology*, vol. 15, no. 4, pp. 171-9, 1983.

Glasser, Gerald J., "The Age Replacement Problem," *Technometrics*, vol. 9, no. 1, pp. 83-91, 1967.

Heinrich, G. and Jensen, U., "Bivariate Lifetime Distributions and Optimal Replacement," *Mathematical Methods of Operations Research*, vol. 44, no. 1, pp. 31-47, 1996.

Hunter, J. J., "Renewal Theory in Two Dimensions: Basic Results," *Advances in Applied Probability*, vol. 6, pp. 376-391, 1974.

Hunter, J. J., "Renewal Theory in Two Dimensions: Asymptotic Results," *Advances in Applied Probability*, vol. 6, pp. 546-562, 1974.

Ingram, C. R. and Scheaffer, R. L., "On Consistent Estimation of Age Replacement Intervals," *Technometrics*, vol. 18, no. 2, pp. 213-219, 1976.

Kalfbleisch, J. D. and Prentice, R. L., *The Statistical Analysis of Failure Time Data*, John Wiley, NY, 1980.

Klefsjö, B., "TTT-Transforms—A Useful Tool when Analysing Different Reliability Problems," *Reliability Engineering*, vol. 15, pp. 231-241, 1986.

Kordonsky, Kh. B. and Gertsbakh, I. B., "Choice of the Best Time Scale for System Reliability Analysis," *Euro. J. of Operational Research*, vol. 65, pp. 235-246, 1993.

Kordonsky, Kh. B. and Gertsbakh, I. B., "Best Time Scale for Age Replacement," *International Journal of Reliability, Quality, and Safety Engineering*, vol. 1, no. 2, pp. 219-229, 1994.

Kordonsky, Kh. B. and Gertsbakh, I. B., "System State Monitoring and Lifetime Scales-I," *Reliability Engineering and System Safety*, vol. 47, pp. 1-14, 1995.

Kordonsky, Kh. B. and Gertsbakh, I. B., "Multiple Time Scales and the Lifetime Coefficient of Variation: Engineering Applications," *Lifetime Data Analysis*, vol. 3, pp. 139-156, 1997.

Kumar, D. and Westberg, U., "Maintenance Scheduling Under Age Replacement Policy Using Proportional Hazards Model and TTT-plotting," *Euro. J. Operational Research*, vol. 99, pp. 507-515, 1997.

Lawless, J., Hu, J., and Cao, J., "Methods for Estimation of Failure Distributions and Rates from Automobile Warranty Data," *Lifetime Data Analysis*, vol. 1, pp. 227-240, 1995.

Love, C. E. and Guo, R., "Using Proportional Hazard Modelling in Plant Maintenance," *Quality and Reliability Engineering International*, vol. 7, pp. 7-17 1991.

Makis, V. and Jardine, A. K. S., "Optimal Replacement in the Proportional Hazards Model," *INFOR*, vol. 30, no. 2, pp. 172-183, 1992.

Makis, V. and Jardine, A. K. S., "Condition-Based Maintenance at the University of Toronto," Internet site http://www.mie.utoronto.ca/labs/cbm

McCall, J. J., "Maintenance Policies for Stochastically Failing Equipment: A Survey," *Management Science*, vol. 11, no. 5, pp. 493-524, 1965.

Murthy, D. N. P., Iskandar, B. P., and Wilson, R. J., "Two-Dimensional Failure-Free Warranty Policies: Two-Dimensional Point Process Models," *Operations Research*, vol. 43, no. 2, pp. 356-366, 1995.

Nakagawa, T., "Continuous and Discrete Age-Replacement Policies," *Journal of the Operational Research Society*, vol. 36, no. 2, pp. 147-154, 1985.

Nakagawa, T. and Osaki, S., "Discrete Time Age Replacement Policies," *Operational Research Quarterly*, vol. 28, no. 4, pp. 881-885, 1977.

Oakes, D., "Multiple Time Scales in Survival Analysis," *Lifetime Data Analysis*, vol. 1, pp. 7-18, 1995.

Pierskalla, W. P. and Voelker, J. A., "A Survey of Maintenance Models: the Control and Surveillance of Deteriorating Systems," *Naval Research Logistics*, vol. 23, no. 3, pp. 353-88, 1976.

Robertson, T., Wright, F. T., and Dykstra, R. L., *Order Restricted Statistical Inference*, John Wiley, Chichester, 1988.

Ross, S. M., *Introduction to Probability Models* (*6E*), Academic Press, NY, 1997.

Scheaffer, R. L. "Optimum Age Replacement in the Bivariate Exponential Case," IEEE Transactions on Reliability, vol. R-24, no. 3, pp. 214-5, 1975.

Singpurwalla, N. D. and Wilson, S. P., "Failure Models Indexed by Two Scales," *Advances in Applied Probability*, vol. 30, pp. 1058-72, 1998.

Valdez-Flores, C. and Feldman, R. M., "A Survey of Preventive Maintenance Models for Stochastically Deteriorating Single-Unit Systems," *Naval Research Logistics*, vol. 36, no. 4, pp. 419-46, 1989.

Wilson, S. P., "Failure Modeling with Multiple scales," Dissertation, George Washington University, D.C., August 1993.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   8725 John J. Klingman Rd., Ste 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library (411 Dyer Rd.) 2
   Naval Postgraduate School
   Monterey, CA 93943-5101

3. Professor Robert R. Read Code OR/Re 1
   Naval Postgraduate School
   Monterey, CA 93943-5002

4. Associate Professor Lyn R. Whitaker Code OR/Wh 5
   Naval Postgraduate School
   Monterey, CA 93943-5002

5. Professor Gerald G. Brown Code OR/Bw 1
   Naval Postgraduate School
   Monterey, CA 93943-5002

6. Assistant Professor Samuel H. Buttrey Code OR/Sb 1
   Naval Postgraduate School
   Monterey, CA 93943-5002

7. Assistant Professor Robert A. Koyak Code OR/Kr 1
   Naval Postgraduate School
   Monterey, CA 93943-5002

8. Associate Professor Craig W. Rasmussen Code MA/Ra 1
   Naval Postgraduate School
   Monterey, CA 93943-5002

9. Scott Frickenstein, Captain, USAF 5
   Assistant Professor, Mathematical Sciences
   2354 Fairchild Drive, Suite 6D2A
   USAF Academy, Colorado 80840-6252

10. Professor Ilya B. Gertsbakh 1
Department of Mathematics and Computer Science
Ben Gurion University
P.O. Box 653
Beer-Sheva, Israel

11. Assistant Professor Thierry Duchesne 1
Department of Statistics
University of Toronto at Mississauga
Mississauga, ON L5L 1C6
Canada

12. Dr Robert Sheldon 1
Military Operations Research Society
101 S. Whiting Street, Suite 202
Alexandria, VA 22304-3416