Order Number 9205728

Reliability and maintenance in European nuclear power plants: A structural analysis of a controlled stochastic process

Sturm, Roland, Ph.D.

Stanford University, 1991

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

RELIABILITY AND MAINTENANCE IN EUROPEAN NUCLEAR POWER PLANTS: A STRUCTURAL ANALYSIS OF A CONTROLLED STOCHASTIC PROCESS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ECONOMICS

AND THE COMMITTEE OF GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

By

Roland Sturm

May 1991

I certify that I have read this dissertation and that in my opinion it is fully adquate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Paul A. David (Principal Advisor)

I certify that I have read this dissertation and that in my opinion it is fully adquate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Takeshi Amemiya

I certify that I have read this dissertation and that in my opinion it is fully adquate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Timothy F. Bresnahan

Approved for the University Committee on Graduate Studies:

Dean of Graduate Studies

Für meine Eltern,

Fritz und Lissi Sturm

RELIABILITY AND MAINTENANCE IN EUROPEAN NUCLEAR POWER PLANTS: A STRUCTURAL ANALYSIS OF A CONTROLLED STOCHASTIC PROCESS

Roland Sturm

Stanford University, 1991

This dissertation analyzes the operating performance of nuclear power plants in five European countries using panel data on individual reactors. Two aspects of performance are of main concern: plant availability and plant reliability (defined as the conditional probability of an unplanned shutdown). The goal of the research is a unified framework that combines behavioral models of optimizing agents with models of complex technical systems which take into account the dynamic and stochastic features of the system. In order to achieve this synthesis, two lines of work are necessary. One line requires a deeper understanding of complex production systems and the type of data they give rise to (chapters 2 and 3), the other line involves the specification and estimation of a rigorously specified behavioral model (chapter 4).

Nuclear power plant operations are modeled as a controlled stochastic process and the sequence of up- and downtime spells is analyzed during failure time and point process models. Similar in spirit to work on rational expectations and structural econometric models, the behavioral model of how the plant process is controlled is formulated at the level of basic processes, i.e. the objective

function of the plant manager, technical constraints, and stochastic disturbances. In contrast to much of the rational expectations literature, the specification of the objective function and technical constraints is based on statistical information in the data and theoretical considerations of the properties of a repairable dynamic production system. This "application-oriented" approach therefore differs from the "method-oriented" approach prevalent in that literature, in which the statistical specification has been chosen largely for analytic convenience.

Acknowledgements

All parts of this dissertation benefitted substantially from suggestions and comments by Paul David, who also corrected and re-worded many of my first attempts at putting this material into written form. I am indebted to many people for their comments on chapters 4 and 5. Kathy Bawn, Tim Bresnahan, Paul David, and Geoff Rothwell provided detailed suggestions on both content and presentation; the economic interpretation of chapter 5 was suggested by Paul David and Tim Bresnahan. Comments by Takeshi Amemiya, Bill Lovejoy, and participants at workshops at Stanford, RAND, Bell Labs, Queen's University, Virginia Tech, and UCLA are gratefully acknowledged. Partial financial support was provided through NSF Grant No. IRI-8814179-02 to the project "Informational and Organizational Impacts on Productivity" at the Center for Economic Policy Research, Stanford University.

Table of Contents

List of Tables xi

List of Figures xiii

1. Introduction 1

1.1 Nuclear power - background and previous work 2

1.2 Structural econometric models 8

1.3 An overview 14

2. A descriptive analysis of nuclear power plant operations in Europe 16

2.1 A Semi-Markov Model of Plant Operation 17

2.1.1 The statistical model 19

2.1.2 Results 24

2.2 Testing the renewal assumption 49

2.3 Continuous production processes and linear regression techniques 58

2.4 Estimating learning in a reliability growth model 64

3. Choosing between competing duration models: an analysis of up- and downtime 77

3.1 Downtime 79

3.1.1 unscheduled outages - equipment failures 79

3.1.2 scheduled outages - refueling 88

3.2 Uptime 105

3.2.1 Repairable and nonrepairable systems 105

3.2.2 A censored point process model of outage occurrences 112

3.2.3 Results 119

3.3 Can the statistical model explain the data? A simulation study 135

4. A structural economic model of operating cycle management 145

4.1 Descriptive statistics for the sample 147

4.2 When is the plant refueled? The operator's decision problem 155

4.3 Econometric specification and results 162

4.3.1 The exogeneous plant process 162

4.3.2 The operator's control decision 166

4.4 Discussion 171

5. Economic interpretation 181

5.1 Comparative statics - how do cost changes affect performance measures? 184

5.2 Productivity in an international perspective 188

6. Summary and conclusion 194

Appendix 1 - The data 201

Appendix 2 - Miscellaneous results 208

References 215

List of Tables

Table 1: Parameter Estimates of the Stationary Markov Model, Country Comparison 31

Table 2: Rank Tests of Homogeneity, Country Comparison 32

Table 3: Parameters Estimates of the Stationary Markov Model, New vs. Old Plants 33

Table 4: Rank Tests of Homogeneity, New vs. Old Plants 34

Table 5: Parameter Estimates of the Stationary Markov Model, German BWR vs. PWR 35

Table 6: Rank Tests of Homogeneity, German PWR vs. BWR 35

Table 7: Parameter Estimates of the Stationary Markov Model, Different Technologies 36

Table 8: Rank Test of Renewal Assumption (H_{0a}) 54

Table 9: Rank Test of Renewal Assumption (H_{0b}) 55

Table 10: Power Comparison 56

Table 11: Definition of Variables 73

Table 12: Reliability Growth Model - Results 75

Table 13a-f: Outages due to Equipment Failure 90

Table 14a-f: Refuel Outages 96

Table 15a-f: Failure Rate 126

Table 16: Lenght of Operating Cycle 132

Table 17: Parameter Values for Simulated Process 141

Table 18: Availability and Capacity Factors 152

Table 19: Unplanned Outages 1981-1986 153

Table 20: Planned and Unplanned Outage Duration as Proportions 154

Table 21: Specifications for Plant Process 173
Table 22: The Exponentiated Quadratic Intensity Model 174
Table 23: Mean Log-Likelihood for Different Specifications 175
Table 24: Cost Function Estimates with Model Specification Aa 176
Table 25: Comparative Statics 186
Table 26: Productivity Calculations: Economic and Technical Effects 192
Table 27: Productivity Calculations: Technical Effects only 193
Table 28: Data Excerpts for Plant Doel 1 203
Table 29: Code Explanation 204
Table 30: List of Plants 205

List of Figures

Figure 1: Transitions 17

Figure 2: Survivor Function: Uptime until Equipment Failure 37

Figure 3: -Ln(Survivor Function): Uptime until Equipment Failure 38

Figure 4: Survivor Function: Downtime for Refueling 39

Figure 5: Survivor Function: Downtime for Repair (Equipment Failure) 40

Figure 6: Density Plot: Downtime for Repair (Equipment Failure) 41

Figure 7: Density Plot: Downtime for Refueling 42

Figure 8: Density Plot: State 2 43

Figure 9: Density Plot: State 4 44

Figure 10: Survivor Function: Uptime until Equipment Failure, France New vs. Old 45

Figure 11: Survivor Function: Uptime until Equipment Failure, Sweden New vs. Old 46

Figure 12: Survivor Function: Uptime until Equipment Failure (Reliability), German BWR vs. PWR 47

Figure 13: Survivor Function Downtime for Refueling, German BWR vs. PWR 48

Figure 14: Do Survivor Functions Cross? 57

Figure 15: A Two-State Model 58

Figure 16: Probability Plot Weibull vs. Normal 102

Figure 17: Refuel Outage Durations: France and Sweden 103

Figure 18: Model Predictions for Refuel Outage Durations: Sweden 104
Figure 19: Repairable Systems 133
Figure 20: Failure Time vs. Process Formulation 115
Figure 21: Density Plot: Operating Cycle Duration 134
Figure 22: Time Path: Empirical Point Availability 142
Figure 23: Simulations of Time Path 143
Figure 24: Simulations of Time Path 144
Figure 25: Empirical Failure Rate 177
Figure 26: Predicted Failure Rate 180
Figure 27: Hazard Function for Refuel Outages 187

1. Introduction

Many industrialized countries adopted the highly promising nuclear power technology in the 1950's and 60's to satisfy their growing energy demands. But despite international cooperation and unprecedented financial support from national governments, the operating experience of nuclear power plants has fallen short of expectations in most countries.

This dissertation analyzes the operating experience of plants in five European countries, all of which are heavily dependent upon nuclear power and the Light Water Reactor technology: Belgium, France, Germany, Sweden, and Switzerland. Although much attention has been paid to the safety and economic viability of nuclear power in the literature, these questions have been considered in isolation and have been analyzed using static and deterministic models (see section 1.1). This dissertation focuses upon the dynamic and stochastic features of plant operations and upon the interdependence between plant economics (measured by availability or capacity factors) and plant reliability and safety (measured by the probability of safety-related unplanned shutdowns).

The ultimate goal is a unified framework that combines behavioral models of optimizing agents with models of complex technical systems. In order to achieve this synthesis, two lines of work are necessary. One line requires a deeper understanding of complex production systems and the type of data they give rise to, the other line involves the specification and estimation of a precise and rigorous behavioral model. The ideas developed in the following chapters apply not only to nuclear power plant operations, but may also be useful for understanding and forecasting the failure or success of other new technologies.

Section 1.1 provides some background, including the definition of some technical terms, and reviews the relevant literature on nuclear power plant operations. Conceptual and technical difficulties encountered in structural econometric models are considered in section 1.2. The final section of this chapter, section 1.3, gives an overview of the main parts of the dissertation.

1.1 Nuclear power - background and previous work

Nuclear power plays a central role in many European economies. It generates over 80% of electricity in France, over 50% in Belgium, over 40% in Sweden, and over 30% in Germany and Switzerland, compared to about 20% in the U.S. or Canada. As in the U.S., the Light Water Reactor (LWR) is the prevalent technology in Western Europe outside Great Britain. There are two main types of Light Water Reactors: the Pressurized Water Reactor (PWR) and the Boiling Water Reactor (BWR). LWR's need to be shut down periodically for refueling and preventive maintenance and major repairs are performed during these periodic refuel outages as far as possible[1]. When there are unplanned shutdowns, the plant is generally returned to operations as quickly as possible. The interval between refuel outages is commonly known as a fuel cycle (when the refuel outage is included) or an operating cycle (when it is not included).

Plant and reactor safety is arguably the most extensively investigated problem surrounding nuclear power. The traditional approach to safety engineering[2] has studied the physical behavior of the plant following an assumed initial event or malfunction using deterministic models. The purpose of this deterministic safety analysis is to check whether the design of the plant can withstand the consequences of an assumed event or needs to be modified. Deterministic safety analysis cannot address the inherently stochastic questions of the probability of fault conditions or the possibility that systems may not work as intended. Probabilistic safety (risk) analysis was developed to address these questions and has led to the well known reactor safety studies (Nuclear Regulatory Commission, 1975, Deutsche Risikostudie Kernkraftwerke, 1980). Safety issues have also been studied by social scientists in various disciplines, such as

[1] Gas-Cooled or Heavy Water Reactors (used in Great Britain and Canada, respectively) can be refueled under load and maintenance patterns thus look very different.

[2] Safety engineering could be defined in this context as the development of a reactor design to prevent the release of radionuclides. Pershagen's (1989) monograph covers the engineering principles and practices of LWR safety, including deterministic and probabilistic safety analysis.

sociology (Perrow, 1984), political science (Morone and Woodhouse, 1989), or large interdisciplinary projects such as the German research on the compatibility of nuclear energy and democratic societies (Meyer-Abich and Schefold, 1986).

While the literature has focused on catastrophic events such as core melt or core disruptive accidents[3], many type of failures that are a safety hazard do not necessarily cause a catastrophic accident: McCormack and Gallaher (1982) report that about 3200 events which were classified by the Nuclear Regulatory Commission as safety-related and required a written licensee event report occurred in the U.S. in a single year (1980). This corresponds to an average of 62 safety-related events per BWR and of 42 safety-related events per PWR. Of course, only the most serious of these events led to an unscheduled plant shutdown and the total number of such shutdowns was less than 100. Although reporting practices vary between countries, events leading to unplanned outage, almost always due to <u>reactor scrams</u>[4], are completely reported in all Western European countries. Thus, unplanned outage statistics are an important measure of <u>plant reliability</u>. One of the goals of this study is to analyze how the instantaneous probability (or more correctly, the <u>hazard</u> or <u>intensity</u>) of unplanned outages changes over time and depends on plant characteristics.

---

[3] Reactor safety studies, for example, attempt to estimate the probability of such an event in a year.

[4] a reactor scram is an emergency shutdown in which the control rods are inserted as rapidly as possible into the core to stop fission.

Economic studies typically consider plant productivity, a performance aspect quite different from safety. Two measures are commonly cited: The availability factor measures the percentage of time a plant has been available over a given period, whether or not it has been fully utilized in generating electricity; the capacity factor (or load factor) measures the ratio of electricity generated to potentially feasible generation.

The most common type of study assumes certain prices, productivity levels, and other parameters, and then calculates whether nuclear power is cheaper or more expensive than alternative sources of energy according to some deterministic model[5]. But there also exists a body of empirical economic research devoted to more complex issues. One of the main issues in plant operations that has been addressed in that line of research has been learning during operation (Joskow and Rozanski, 1979, Lester and McCabe, 1988, Krautman and Solow, 1990)[6]. The three studies cited use a reduced form approach which involves regressing a measure of plant productivity on cumulative output, age, or other observed covariates. However, because they ignore the actual dynamics of production, studies following this approach lead to spurious findings. Indeed, it will be shown below (section 2.3) that reduced form models would "discover" nonzero learning effects even in a

[5] See Michaelis (1986) for numerous such calculations.

[6] Other issues include scale economies, construction times, and construction costs. The recent book by Thomas (1988) provides many references, appendix E in David et al. (1988) by Bresnahan and Rothwell reviews empirical economic studies of commercial nuclear power.

stationary continuous time process.

The question how the stochastic and dynamic features of nuclear power plant operations can be modeled raises an immediate difficulty since the standard formulation of economic production theory is not well suited for such a task. As David et al. (1988) have argued, economic production functions are static, deterministic "black box" representations of actual production processes. They suggest, instead, a renewal process or duration model approach for analyzing individual plant operating histories, such as the sequence of "up" and "down" times. Some relevant techniques of duration analysis, originally developed in reliability engineering and popularized in the biomedical literature, have been widely used in studying labor market phenomena and demographic processes. Studies using duration models in the context of nuclear power operations include Rothwell (1989), Rothwell and Jensen (1990), and David, Rothwell, and Maude-Griffin (1991).

Classic duration or failure time models are renewal models concerned with single events ending a spell. The theory of renewal processes is well developed (Cox, 1962) and it is not surprising that the more complex multiple spell models in economics, discussed in Heckman and Singer (1985), are built on the renewal model: time is reset to zero at the beginning of each spell (occurrence of an event). I define time measured in such a way as spell time. In actual applications (e.g. David and Mroz, 1989), dependencies between successive spells may be introduced through mixing distributions (unobserved heterogeneity) or covariates.

Instead of using single spell/renewal models as the building block for multiple spell models, one could generalize hazard function modeling beyond the first event to second and subsequent events. This is a central idea for the formal model of a repairable system in section 3.2, which provides a different conceptual approach to analyze multivariate failure time data. As discussed in section 3.2, this system process approach can model the aging and wear-out of a complex system; the traditional renewal approach can only model the life of components. Time in the system process model is measured from a constant point for several events. When it is measured from the day of first criticality or the day of first commercial production, I call it plant time, if it is measured from the start up from a refuel outage, it is refered to as cycle time. Statistical results show that periodic fuel cycles are an important unit of analysis (section 2.2).

1.2 Structural econometric models

In contrast to purely engineering questions, the successful adoption of a technology involves human decision-making. Thus a behavioral model is necessary in order to assess empirically the impacts of regulation, economic incentives, or new technologies. As I have argued elsewhere (Sturm, 1989a), a model that estimates the strength of regulation by comparing reported "regulated" downtimes, but which ignores the possibility that decision makers can react to changes in their environment, invariably fails to give correct predictions when used for a policy analysis. Similar arguments, familiar to most economists as the "Lucas critique" (Lucas, 1976), can be brought forward against the uncritical use of regression analysis results or reduced form models for predictive purposes. These arguments generally point towards the violation of two principles: the first is that individuals respond to changes in their opportunities; the second is that decisions are made over time in a stochastic environment.

Structural econometric models are consistent with these two principles. The original development of these ideas occurred in not in microeconomics, but in macroeconomics, e.g. Lucas (1976), Sargent and Wallace (1976), Hansen and Singleton (1982). The importance of formulating dynamic economic models at the level of "deep" or "primitive" parameters, i.e. parameters describing the underlying distributions of preferences (utility functions) and transient disturbances, is emphasized in Sargent (1979). Microeconomic applications have only appeared in recent years. Three pioneering papers in microeconomics are Gotz and McCall (1984), Miller (1984), and Wolpin (1984). Gotz and McCall describe and estimate a model of Air Force officers making decisions of whether to stay in or leave the Air Force, Miller estimates a multi-armed bandit model of employment decisions by numerically calculating dynamic allocation indices, and Wolpin estimates a model of decision making by Malaysian women about the number and timing of births. Other structural models have been developed by Pakes (1986) for estimating an optimal stopping model of patent renewal, Wolpin (1987) for the analysis of job acceptances of high school graduates, Ryu (1990) for a replacement model, Fafchamps (1990) for a peasant's labor allocation to different activities in Burkina Faso, and Berkovec and Stern (1991) for a dynamic programming model of job exit behavior of older men. An important contribution is Rust's (1988) development of an algorithm that can be used to estimate structural econometric models for a wide range of applications and which has been employed by Rust (1987), Montgomery (1987), Das (1989), and Kennet (1990).

My definition of a structural econometric model is narrow. Two characteristics are essential:

1. the economic model is specified in terms of "primitive" processes such as preferences and technology, and
2. there is a one-to-one correspondence between the dynamic

economic model and its econometric specification.

Thus a formal economic model that only loosely motivates the econometric implementation is not a structural econometric model. A typical case encountered in the literature is that the economic model is rigorously specified, but the statistical model is chosen for tractability and assumptions are made about the form of the likelihood function rather than about basic economic processes. In other words, the standard econometric model is "method-oriented" rather than "problem-oriented". The distinction is far from trivial and I will now discuss in more detail the two main conceptual and technical difficulties arising in estimating structural models.

The first point is conceptual. In the economic paradigm, an agent's decisions are the result of solving an optimization problem. Theoretical models specify how the decision y (e.g. the demand for certain commodities) depends on the objectives and constraints of an agent x, for example $y=f(x)$. Often the interest is in some constant parameters θ which are unknown and the equation is reformulated as

$$y=f(x,\theta) \tag{1}$$

Note the statistical degeneracy which is typical for almost all decision theoretic models: y is a deterministic function of x and θ. In general, it is impossible to find a θ which satisfies equation (1) -- the economic model is rejected by the data. Thus equation (1) is often amended to a statement about observed decisions

$$y=f(x,\theta)+u \tag{2}$$

with u being some disturbance. Unfortunately, this ad hoc "solution" of the statistical degeneracy problem <u>contradicts the optimizing model</u>![7] The logical inconsistency between economic and statistical model is caused by the economic model assumption that the agent's decisions are described by a unique solution to an optimization problem and the statistical model implication that the agent randomly chooses actions, which are necessarily suboptimal.

Various suggestions to solve this logical difficulty can be found in the papers mentioned above. The common approach is to acknowledge that the information available to the econometrician about an individual's decision problem is incomplete and to formally model this missing information. Differences in behavior reflect differences in some unobserved components whose effects are traced throughout the optimization problem. A specified distribution of "unobservables" is filtered through the economic decision problem and induces a distribution on observed decisions, say g(y).

The difficulty is that the agents response function f(), let

---

[7] The only exception to this rule is the rare case when the only source of u is measurement error, i.e. the agent makes decisions according to (1) but we measure them according to (2),

alone the distribution g(), only has a closed form solution for very special cases of preferences (utility functions), constraints, and the distribution of unobservables. This causes many researchers to switch from a formal theoretical model to a "reduced" form approach by assuming that f() or g() "approximately" equals some tractable function--the "method-oriented" rather than "problem-oriented" statistical implementation. The impossibility of obtaining closed form expressions occurs especially in dynamic models: the solution to a general dynamic programming problem is only defined recursively through Bellman's principle of optimality[8]. It is well known that the time required for a numerical calculation of the solution of a dynamic program can quickly exceed any bounds (Bellman's "curse of dimensionality"). And although obtaining one solution for a set of parameters may solve an agent's problem, it does not help the econometrician who, in order to obtain g(), needs in principle a solution for every possible value of the unobserved components and the unknown parameters.

Among researchers who pursue structural models, two approaches have been adopted. The traditional ("closed form") approach is to look for functional forms of "primitive" processes (preferences, laws of motion, constraints) for which a closed form statistical model can be found. Although such work is often very ingenious

[8] There are a few important exceptions (Sargent, 1987), notably models with linear laws of motion and quadratic preferences.

(e.g. Ryu, 1990), the self-imposed restrictions on functional forms leads to implausible models which cannot be applied to actual problems. It is also impossible to perform a sensitivity analysis or compare different specifications of the basic processes.

Rather than relying on assumptions that guarantee a closed form expression for the likelihood function, the "algorithmic" approach defines the likelihood function only implicitly as the solution to an economic decision problem, but under some constraints which make the problem computationally feasible. Within limits, this approach liberates the researcher from having to make a priori assumptions regarding functional forms. However, the conceptual advantage of precise modeling of the economic problem is always bought at the cost of substantially increased computational demands. The particular approach taken here is to keep the flexibility of functions describing preferences (the cost or utility function) and laws of motion (the technical or plant process). No constraints are imposed on these basic economic processes, but the possible distributions of unobservable states is restricted. The assumption that unobserved states follow an extreme value distribution (as in a static random utility model that gives rise to a logit model) reduces the computational problem by several orders of magnitude. Note that I have performed all calculations on a personal computer, whereas other researchers had to resort to much more powerful machines[9].

---

[9] Fafchamps (1990), for example, performed his calculations on a Cray supercomputer.

It is likely that alternative techniques will be developed in the future. Rather than solving for the distribution of y numerically, one could attempt to simulate it. A related idea has been discussed by Diggle and Gratton (1984) and such a "method of simulated likelihood" may become important to estimate more complex behavioral models. Preliminary work has been successful in that I was capable of obtaining estimates which were close to the maximum likelihood estimates. Of course, simulations were done for a very simple model which could be solved analytically in order to assess the quality of estimates. A discussion of these techniques, their problems and promises, is left for the future and is not attempted in this dissertation. Another approach is to use more ad hoc criteria for estimation, and the "method of simulated moments" (McFadden, 1989, Pakes and Pollard, 1989) has already been proven to be useful in structural econometric estimation (Pakes, 1986).

1.3 An overview

This study is based on a data set containing the operating history of individual reactors in five countries (see appendix 1 for details). This data set permits one to analyze and estimate models of the process that generates the sequence and duration of "up" and "down" times determining performance measures of plant reliability and availability. Chapter 2 estimates a semi-Markov model (2.1) and tests the renewal assumption (2.2). The inconsistency of standard regression approaches with continuous time production processes is demonstrated in section 2.3 and an alternative approach to measuring learning in a reliability growth model is discussed in section 2.4. Parametric models for the durations of the two most important types of outages are estimated in section 3.1, up time durations are analyzed in section 3.2. The simulation study in section 3.3 shows that the process formulation models actual time paths well, whereas Markov renewal models fail to display the observed cyclical behavior. Chapter 4 develops an econometric model of operating cycle management, i.e. the decision of the plant manager to schedule refuel outages. Based on the process model of plant operations developed in chapter 3, this behavioral model brings together various measure of plant performance, such as availability and reliability (unplanned outage rate) in a unified framework. It also demonstrates the existence of trade-offs between different aspects of plant performance and quantifies these trade-offs. Chapter 5 shows how this model can distinguish the effects of economic incentives and technical ("X-efficiency") factors on various aspects of nuclear power plant performance and how this explains observed differences between countries. Chapter 6 discusses extensions of this work and their potential application to other technologies. Several statistical techniques needed throughout the paper are reviewed in appendix 2.

2. A descriptive analysis of nuclear power plant operations in Europe

The ratio of variable (operating) costs to fixed costs is low for nuclear power plants compared to fossil fuel plants. This is a major reason why nuclear power plants have mainly been used to cover base load requirements. There are also technical reasons why utilities are reluctant to change output levels quickly[10]. In the absence of substantial changes in the rate of production, the electricity production model is relatively simple: a plant produces output at a constant rate while running, generally near designed capacity, and it produces nothing during outages. But when is a plant running and when not?

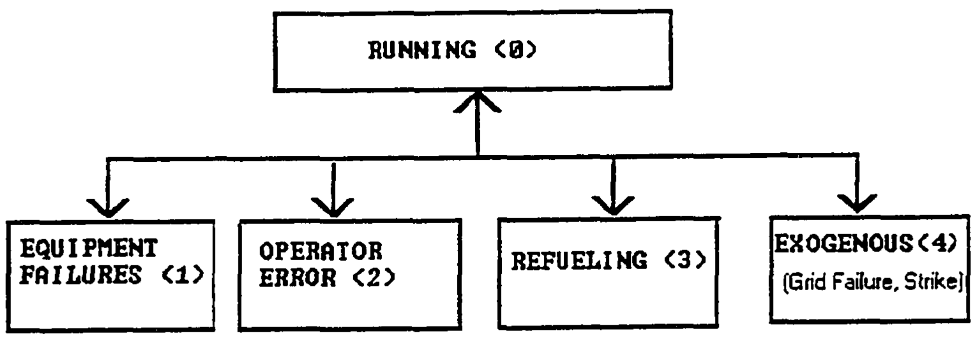
The first section in this chapter gives descriptive statistics and estimates a semi-Markov model (defined in section 2.1.1) of the sequence of up- and downtimes for the sample. The data set covers the years 1981-1986 and all commercial Light Water Reactors in Belgium, France, Germany, Sweden, and Switzerland; details about the data can be found in Appendix 1. Despite being nonparametric, the semi-Markov model has some restrictions, such as the assumption of renewal at every transition. Section 2.2 tests the renewal assumption and alternative hypotheses formally using nonparametric rank tests. The renewal assumption is rejected and this has implications for descriptive and structural models: economic models

[10] Two examples are potential problems with fuel rods and thermal stresses on large components.

built on simple maintenance models from engineering or operations research are inappropriate for a structural model of nuclear power plant operations.

The importance of the stochastic process formulation for continuous production processes is demonstrated in section 2.3 where it is shown that several linear regression models estimated by previous investigators are inconsistent with even the simplest continuous time model. As the goal of much of that literature was directed at estimating learning curves, section 2.4 presents an alternative model for estimating improvements in reliability.

Figure 1: Transitions



2.1 A Semi-Markov Model of Plant Operation

The plant can be in one of several states at any point in time. A simplified representation of the state space S and possible transitions between states are given in figure 1. The International Atomic Energy Agency (IAEA) provides additional information at the reactor system and even component level. However, a finer distinction is not useful for the purpose of this study because there are not enough observations to fruitfully analyze a more extensive state space formulation. One run ("up") state and four "down" states are distinguished: outages due to equipment failure (1), outages due to operator error, repair, or testing (2), planned refuel outages (3), and "exogenous" outages such as lightning, earthquakes, grid failure, or labor disputes (4).

By far the most common cause for an exit from the run state is an equipment failure (state 0→1), almost all of which cause an automatic scram. This transition determines plant reliability and is of major interest for plant safety because, as mentioned before, only the most serious safety-related events will necessitate a plant shutdown. Although outages due to equipment failures occur relatively often, most of the downtime is not spent in state 1, but in state 3 (refueling and maintenance), because the average sojourn time in state 1 (or 2 or 4) is relatively short compared to the average sojourn time in state 3. Thus the transition from state 3 to state 0 is an important determinant of plant availability and productivity. States 2 and 4 are of relatively minor importance during any given interval both as far as interruptions of uptimes and total downtime is concerned, but it is important to separate these effects in order to discuss differences in plant reliability and availability between countries and plant vintages.

Note that the classification scheme used by the IAEA, which focuses on reactor systems and components, differs from engineering practices, which typically classify unplanned outages according to the type of the initiating event. While the IAEA distinguishes equipment failures, operator errors, and external causes, the three corresponding (but different) classes of initiating events in safety analysis (e.g. Pershagen, 1989) are:

- Loss-of-Coolant-Accidents (LOCA), which are events caused by a pipe break or leakage in the primary system;
- transients, a general term for abnormal events other than LOCA's occurring during operations;
- external events, such as lightning, earthquakes, fire (which is considered an external even if it originates within the plant), etc.

2.1.1 The statistical model

Let $\{X(\tau);\ 0 \le \tau < \infty\}$ be a family of random variables where $X(\tau)$ denotes the state of the plant (see figure 1) at time τ. I consider both the classic case where $\{X(\tau)\}$ is a Markov process with stationary transition probabilities, i.e.

$$P_{ij}(t) = P\{X(t+u) = j \mid X(u) = i\} = P\{X(t) = j \mid x(0) = i\} \quad i, j \in \mathcal{S} \tag{3}$$

and the case when the transition probabilities are nonstationary. The standard Markov formulation is not innocuous because equation (3) only holds if transition times between states i and j are exponentially distributed (with constant hazard parameter λ_{ij}). As this may be questionable, I also consider time-dependent transition intensities $\lambda_{ij}(t)$.

For the stationary Markov model with unknown transition probability functions and intensities λ_{ij}, the likelihood function based on a continuously observed sample path of $X(\tau)$ on $[0,T]$, conditional on the initial state $x(0)$, is[11]

$$L=\left\{\prod_{i\neq j}\lambda_{ij}^{n_{ij}}\right\}\left\{\prod_{i}e^{-\lambda_i t^i}\right\} \tag{4}$$

where n_{ij} is the observed number of transitions from state i to state j, t^i is the total amount of time spent in state i, and

$$\lambda_i=\sum_{j|j\neq i}\lambda_{ij} \tag{5}$$

The maximum likelihood estimate for λ_{ij} and its variance are:

$$\hat{\lambda}_{ij}=\frac{n_{ij}}{t^i},\quad V(\hat{\lambda}_{i,j})=\frac{\lambda_{ij}^2}{n_{ij}} \tag{6}$$

Nonparametric estimation of time dependent transition intensities provides additional insights into details of the statistical process and simultaneously provides a means to assess the appropriateness of the stationary Markov assumption. Despite time dependence, the memory of such a semi-Markov process remains limited: spell durations are independent and the process "remembers" only time since the last transition (time is spell time). The model of sojourn times (spell durations) in any one state can be rewritten as what is known in the biometrical

[11] For the following discussion consider everything conditional on the observation period.

literature as an independent competing risk model (Kalbfleisch and Prentice, 1980, Ch. 7). This is a consequence of the limited memory of the semi-Markov model and it permits a nonparametric estimation technique which is very easy to implement[12].

The independent competing risk model assumes that there are several statistically independent causes which can terminate the sojourn in one state, conceptualized commonly with "latent" transition times. Of course, only the smallest of these latent transition times is realized and observed. Generally speaking, the sojourn time in state i has the survivor function

$$\begin{aligned} S_i(t) &= Pr(\text{all latent transition times} > t) \\ &= \prod_j Pr(T_j > t) = \prod_j S_{ij}(t) \end{aligned} \tag{7}$$

where the second equality is due to the independence assumption and the existence of cause specific survivor functions $S_{ij}(t)$ (the third equality) follows from Tsiatis (1975). Statistically, it is impossible to test the independence assumption (Cox, 1959, Tsiatis, 1975) and for descriptive purposes the independent censoring model is equivalent to any model with dependency between the latent transition times. It does not necessarily provide valid answers to policy questions of the type: "what is the hazard rate for equipment failures given the 'removal' of some outage cause, for

[12] See Lagakos, Sommer, and Zelen (1978) for a somewhat different derivation of nonparametric semi-Markov models.

example, operator errors and test outages?"[13] The survivor functions for transitions between states i and j can be expressed in terms of intensities as

$$S_{ij}(t)=\exp\left[-\int_0^t \lambda_{ij}(u)\,du\right] \tag{8}$$

The semi-Markov and independence assumptions together allow to estimate each cause-specific survivor function individually using the product limit or Kaplan-Meier estimator

$$\hat{S}_{ij}(t)=\prod_{\{l|t_l<t\}} 1-\frac{d_{ijl}}{r_{il}} \tag{9}$$

where r_{il} is the size of the riskset (number of observations that are in state i) immediately before t_l and d_{ijl} is the number of transitions from i to j at t_l. The estimated survivor function for the sojourn time in state i then becomes

$$\hat{S}_i(t)=\prod_j \hat{S}_{ij}(t) \tag{10}$$

Nonparametric tests complement plots of cause specific survivor functions. Because only the smallest latent transition time is observed, many cause specific transitions are incomplete (<u>right censored</u>) because a transition due to a different cause ended the sojourn in a certain state. Standard rank methods such as

[13] Kalbfleisch and Prentice, (1980, Ch.7) discuss this problem in detail.

the Wilcoxon test need some modification to deal with the problem of censoring. These modifications have been discussed by Gehan (1965), Breslow (1970), Prentice (1978), and others. The null hypothesis in the two sample case (Gehan's test) is:

$$H_0: F_1 = F_1 \quad G_1 = G_2 \qquad (11)$$

where F_i is the distribution of failures in sample i and G_i is the distribution of censoring values in sample i. The main disadvantage of Gehan's test is its sensitivity to censoring patterns. Because the assumption of identical censoring distributions is unlikely to be satisfied here I prefer Prentice's (1978) versions of the generalized Savage (log-rank) and Wilcoxon tests, which are less sensitive to censoring patterns. These two tests can also be used for multisample comparisons. A potential problem of all rank tests is that they may not be very powerful if $F_1 \neq F_2$, but neither $F_1 > F_2$ nor $F_1 < F_2$. Plotting the survivor functions is an additional useful check if the null hypothesis is not rejected. Two and multiple sample tests for partially censored data are discussed in Kalbfleisch and Prentice (1980) and Cox and Oakes (1984).

As a practical matter, I condition on the first observed transition to avoid the problem of left censoring. Dropping the first incomplete observation, which for almost all plants is a sojourn time in the run state, will not cause a bias of any actual relevance given the relatively large number of observations on complete transitions for each plant. The left censoring problem could be addressed statistically, but it would require rather strong or even implausible assumptions (Heckman and Singer, 1985). A model which deals with left censored observations but requires cycles and a semi-Markov process in equilibrium has been discussed by Dewanji and Kalbfleisch (1987).

The finite observation interval (observations end in December 1986) causes additional statistically independent right censoring for all types of transitions. However, censoring is very minor for outages and can be ignored for practical purposes[14]. Dropping censored observations, I provide nonparametric density plots using kernel estimates (see appendix 2) for outage durations.

2.1.2 Results

Consider first differences across countries using data on all commercial Light Water Reactors operating between 1981 and 1986 (table 1). Time is measured in days and the table reports the reciprocal of the transition intensity for easier readability, i.e. $\hat{\lambda}_{ij}^{-1}$. Under the assumptions of the stationary Markov process, this number can be interpreted as the mean time in days between a transition from i to j, conditional that no transition to a state other than j occurs. Compared to the usual calculation of a mean, a nonparametric method, the results depend on the exponentiality

---

[14] An outage is censored only if the plant was down at the end of the observation period (31.12.1986), thus there could be at most one censored observation per plant. Given that outages other than refueling occur for a very small fraction of total time, censoring is negligible. Even for refuel durations, censoring is not a problem. Of 263 refuel outages, only 2 were censored (Tihange 1, Paluel 3); of 1136 outages due to equipment failure, only 3 were censored (Paluel 4, Tricastin 4, Grafenrheinfeld), all outages for other reasons were uncensored.

assumption of the stationary Markov model. The observed number of transitions between the states n_{ij} are printed in parentheses. The standard errors are easily calculated from this information, according to equation (6), in order to perform tests.

Ignoring transitions intensities that cannot be estimated precisely because of the small number of events in the sample, for example, the intensity of transitions from state 0 to state 4, the most dramatic difference between countries exists in reliability or, equivalently, mean time until an equipment failure: 0→1. The estimated mean time until an equipment failure is 6 times longer in Switzerland than in France. Figure 2 plots the nonparametric estimate of this cause-specific survivor function. The lines correspond to the following countries (from top to bottom): dots and dashes - Switzerland; closely spaced dots - Germany; dashes - Belgium; dots - Sweden; solid - France. Not surprisingly, tests strongly reject the hypothesis that plants in different countries are of the same reliability (table 2). I also perform two-way comparisons which reveal that even for the countries whose reliability is "closest" according to figure 2, the test of homogeneity is rejected at 10% or better for all pairs except France and Sweden.

A very simple graphical check of the stationary Markov assumption is to plot -ln(S^(t) against t. It follows from equation (8) that this should be a straight line if the hazard is constant and the distribution exponential. Figure 3 shows that sample hazards are relatively constant, implying that the exponential model is a fairly good description. The slight concavity suggests some negative duration dependence, but no deeper structural interpretation can be attached to this finding. In particular, one would suspect the existence of heterogeneity and the near constancy of the empirical hazards could very well be due to the effects of positive duration dependence for the hazards of each particular spell and heterogeneity in the population canceling each other. This well-known phenomenon is discussed in Barlow and Proschan (1975), Heckman and Singer (1985).

Another major difference between countries exists in the distribution of refuel durations (3→0). Again, French plants appear to be substantially worse than others. In contrast, the distribution of the duration of repair outages, i.e. recoveries from equipment failures (1→0), appears to be very similar. The corresponding plots appear in figures 4 and 5. Although the hypothesis of homogeneity across countries is rejected for the case of the distribution of refuel durations, this is not true for the distribution of repair durations (table 2).

The latter finding is consistent with the density plots of figures 6-9. Figure 6 gives nonparametric density estimates for the duration of spells in state 1 (equipment failures) using an Epanechnikov kernel with bandwidths h=1.0 in a) (which is likely to undersmooth the data), and h=10.0 in b) (which may oversmooth the data). As before, the lines correspond to the following countries:

dashes - Belgium, solid - France, closely spaced dots - Germany, dots - Sweden, dots and dashes - Switzerland. There appears to be a mode between 1 and 2 days and a smaller second mode around one week. There may be a third mode further out in the tail (not plotted), but observations there are sparse. The bandwidths in figure 7 (refuel outages) are h=10 and h=30. The density plots clearly reveal the differences between countries discovered by the Markov model and by rank tests. There are few observations for states 2 and 4 (figures 8 and 9), which makes a detailed interpretation questionable. Note that the dotted and dashed line for Switzerland in figure 9 corresponds to 1 observation. The bandwidths are h=0.3 in a) and h=1.0 in b).

A question arising from this is to what extent differences in reliability or availability reflect differences in the composition of plants regarding vintages or technologies (Pressurized Water Reactors, PWR, vs. Boiling Water Reactors, BWR). France, for example, is the only country that has far more "new" plants than "old" plants[15]; Sweden has more BWR than PWR, Germany and Switzerland more PWR than BWR. I therefore stratify the sample for a particular country first according to vintage and then according to technology. The results for a comparison for vintages is

[15] "Old" plants are defined as plants that started commercial operations between 1970 and 1980, "new" plants started commercial operation after 1980.

reported for France and Sweden[16], the comparison of technology PWR-BWR is limited to Germany[17].

Table 3 reports the estimates under the stationary Markov model for "old" and "new" plants in France and Sweden. Of particular interest is plant reliability for which one might expect two competing effects to be at work: although newer plants may be inherently more productive or reliable because of a more modern technical design, the data may also reflect startup problems since these plants have yet to undergo an initial "shakedown" period in which installation errors and inferior components are detected and rectified[18]. A third effect, namely overall improvements in managing nuclear power plants better over time, does not cause a bias since we observe plants only over a relatively small window, i.e. the effect of calendar time does not matter. If one had complete plant histories, then it would be important to take into account that observations for older plants tend to occur earlier in time than observations for newer plants.

---

[16] There were not enough data on recent reactors for this analysis to be useful for Germany, Belgium, and Switzerland.

[17] Because country-specific effects are of major importance, pooling observation across countries is inappropriate. Sweden has only 2 PWR, Switzerland only 2 BWR, and neither France nor Belgium have BWR.

[18] A study by the Institute of Nuclear Power Operations (1984) shows that reactors which have been in operations for 3 years or more have a lower scram rate in any given year than new plants. McCormack and Gallaher (1982) calculate a mean number of 82 safety-related events per plant which are in a state of power ascension in 1980 compared to 42 (62) per PWR (BWR) in commercial operation.

As in the earlier discussion, I plot the nonparametric estimates of the cause specific survivor function for reliability and perform tests of homogeneity (figures 10 and 11, table 4). The null hypothesis of no differences in reliability between new and old plants cannot be rejected using nonparametric tests.

German BWR were much less reliable than German PWR (figure 12), but this may be particular to Germany and the design problems with the 69 BWR series. While all pressurized plant systems are subject various degradation mechanisms, intergranular stress corrosion cracking (Danko and Stahlkopf, 1982) has been a major problem in the piping system of BWR and contributed to the lower reliability of BWR compared to PWR in Germany. Swedish BWR were less susceptible to stress corrosion cracking due to the choice of a stainless steel material with low carbon content (Pershagen, 1989, p.380) than German or U.S. BWR. Table 5 also shows a substantial difference regarding refuel outage durations. However, this finding is caused by a few very atypical durations. Neither the plot of the nonparametric survivor functions (figure 13) nor a rank test (table 6) reveal a major difference between the distribution of refuel outage durations of PWR and BWR.

There are several plants representing two very different technologies. The first nuclear power plants constructed in France, many of which were permanently shut down in the 1980s, were Gas-Cooled Reactors (GCR). Canada has developed its unique design of

Heavy Water Reactors (HWR). In contrast to LWR which need to be shut down periodically for refueling, GCR and HWR can be refueled under load. While preventive maintenance and major repairs at LWR are performed during refuel outages as far as possible, planned repair outages assume this role for GCR and HWR. Therefore state 3 (refuel outage) is replaced by state 3' (planned maintenance/repair outage) in table 7.

Table 1: Stationary Markov Model: Estimated Mean Waiting Time in Days (number of observations in parentheses)

Transition	Belgium	France	Germany	Sweden	Switzerland
0→1	110.27 (70)	75.39 (531)	252.46 (64)	87.94 (197)	476.05 (16)
0→2	321.67 (24)	232.75 (172)	577.05 (28)	666.30 (26)	1523.4 (5)
0→3	350.86 (22)	333.61 (120)	343.78 (47)	333.15 (52)	317.37 (24)
0→4	1929.7 (4)	727.88 (55)	4039.4 (4)	1082.7 (16)	7616.8 (1)
1→0	5.45 (73)	4.84 (552)	3.42 (69)	4.08 (204)	3.46 (17)
2→0	6.16 (26)	4.49 (189)	7.79 (32)	9.85 (30)	2.42 (6)
3→0	43.02 (23)	63.86 (121)	62.39 (56)	45.99 (53)	39.36 (27)
4→0	1.30 (4)	4.09 (56)	3.71 (4)	0.85 (16)	0.29 (1)

Table 2: Rank Tests of International Homogeneity

(χ^2-statistic, p-value)

comparison (degrees of freedom)	log rank test	generalized Wilcoxon test
Reliability (0→1) all countries (4)	89.40 0.000*	80.54 0.000*
Reliability (0→1) Belgium-Sweden (1)	2.73 0.099	2.20 0.138
Reliability (0→1) Sweden-France (1)	1.76 0.18	0.97 0.32
Reliability (0→1) Belgium - France (1)	7.06 0.008*	5.35 0.021*
Reliability (0→1) Germany - Switzerland (1)	3.16 0.075	3.37 0.066
Reliability (0→1) Germany - Belgium (1)	10.63 0.001*	6.49 0.011*
Repair Duration (1→0) all countries (4)	2.56 0.634	3.92 0.417
Refuel duration (3→0) all countries (4)	89.40 0.000*	80.54 0.000*

* significant at 5% level.

Table 3: Stationary Markov Model: Estimated Mean Waiting Time in Days (number of observations in parentheses)

Transition	France, old plants	France, new plants	Sweden, old plants	Sweden, new plants
0→1	70.87 (256)	79.60 (275)	83.18 (150)	103.14 (47)
0→2	201.59 (90)	266.95 (82)	623.81 (20)	807.92 (6)
0→3	323.99 (56)	354.49 (64)	319.90 (39)	372.88 (13)
0→4	1008.0 (18)	591.62 (37)	959.71 (13)	1615.8 (3)
1→0	5.99 (261)	3.81 (291)	3.30 (153)	6.43 (51)
2→0	4.91 (96)	4.06 (93)	2.36 (23)	34.40 (7)
3→0	63.97 (56)	64.00 (65)	46.42 (40)	44.66 (13)
4→0	3.36 (18)	4.43 (38)	0.85 (13)	0.86 (3)

"Old" plants are defined as plants that started commercial operations between 1970 and 1980, "new" plants started commercial operation after 1980.

Table 4: Rank Tests of Homogeneity for Different Plant Vintages (χ^2-statistic, p-value)

comparison (degrees of freedom)	log rank test	generalized Wilcoxon test
Reliability (0→1) France new vs. old (1)	0.94 0.332	1.34 0.247
Reliability (0→1) Sweden new vs. old (1)	0.73 0.392	1.02 0.313

"Old" plants are defined as plants that started commercial operations between 1970 and 1980, "new" plants started commercial operation after 1980.

Table 5: Stationary Markov Model: Estimated Mean Waiting Time in Days for West German plants (number of observations in parentheses)

$1/\lambda_{ij}$	0→1	0→2	0→3	0→4	1→0	2→0	3→0	4→0
PWR	393.3 (27)	816.9 (13)	312.3 (34)	10619.0 (1)	4.19 (30)	14.48 (14)	50.27 (40)	1.29 (1)
BWR	149.7 (37)	369.2 (15)	426.0 (13)	1846.1 (3)	2.82 (39)	2.59 (18)	92.68 (16)	4.51 (3)

Table 6: Rank Tests of Technological Homogeneity, West German Plants (χ^2-statistic, p-value)

comparison (degrees of freedom)	log rank test	generalized Wilcoxon test
refuel duration (3→0) BWR - PWR Germany (1)	1.51 0.218	0.44 0.505
Reliability (0→1) BWR - PWR Germany (1)	10.53 0.001*	8.72 0.003*

* significant at 5% level.

Table 7: Different Technologies: Canadian HWR and French GCR

Estimated Mean Waiting Time in Days for Stationary Markov Model

(number of observations in parentheses)

$1/\lambda_{ij}$	France GCR	Canada old	Canada new	Canada all
$0 \to 1$	76.44 (79)	116.73 (183)	94.11 (135)	107.99 (318)
$0 \to 2$	1006.5 (6)	504.28 (25)	399.98 (16)	463.58 (41)
$0 \to 3'$	232.27 (26)	573.05 (22)	355.53 (18)	475.17 (40)
$0 \to 4$	506.20 (11)	525.29 (24)	266.65 (24)	395.97 (48)
$1 \to 0$	7.56 (81)	7.42 (111)	4.45 (73)	6.24 (184)
$2 \to 0$	35.53 (6)	7.02 (26)	1.78 (18)	4.87 (44)
$3' \to 0$	60.86 (26)	33.46 (24)	28.01 (18)	31.16 (42)
$4 \to 0$	4.84 (12)	7.25 (24)	6.57 (26)	6.89 (50)

Figure 2: Survivor Function: Uptime until Equipment Failure

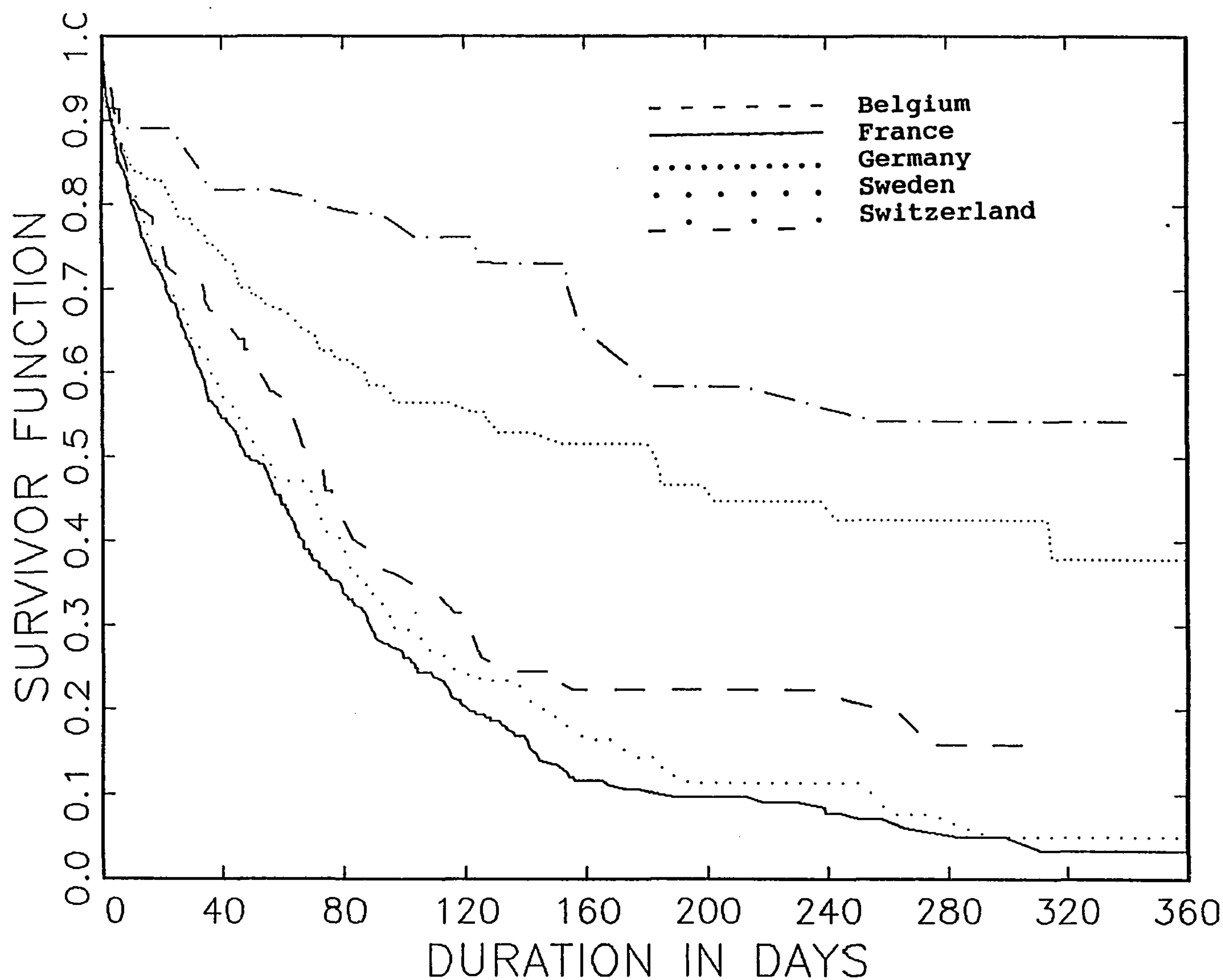


Figure 3: -Ln(Survivor Function): Uptime until Equipment Failure

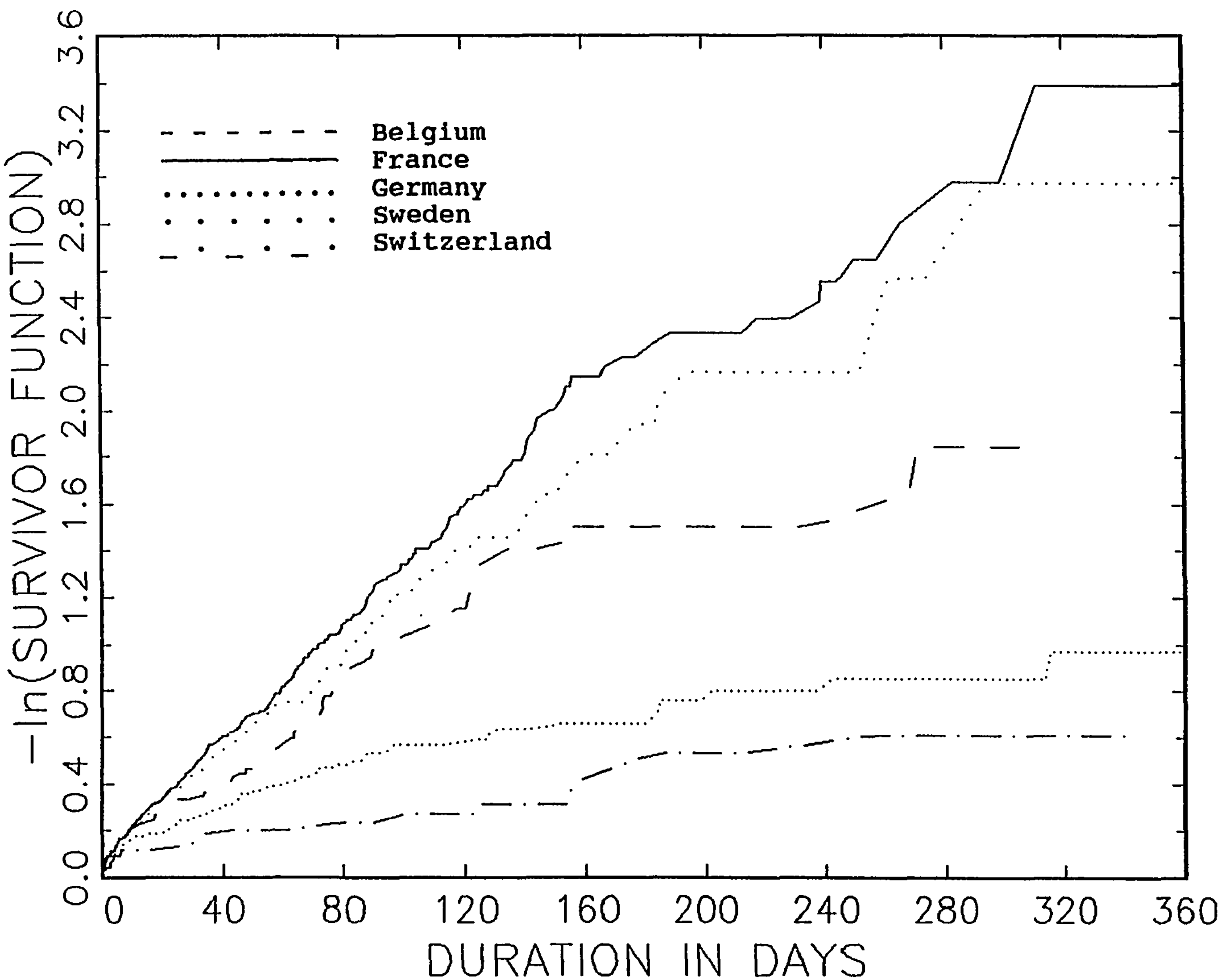


Figure 4: Survivor Function: Downtime for Refueling

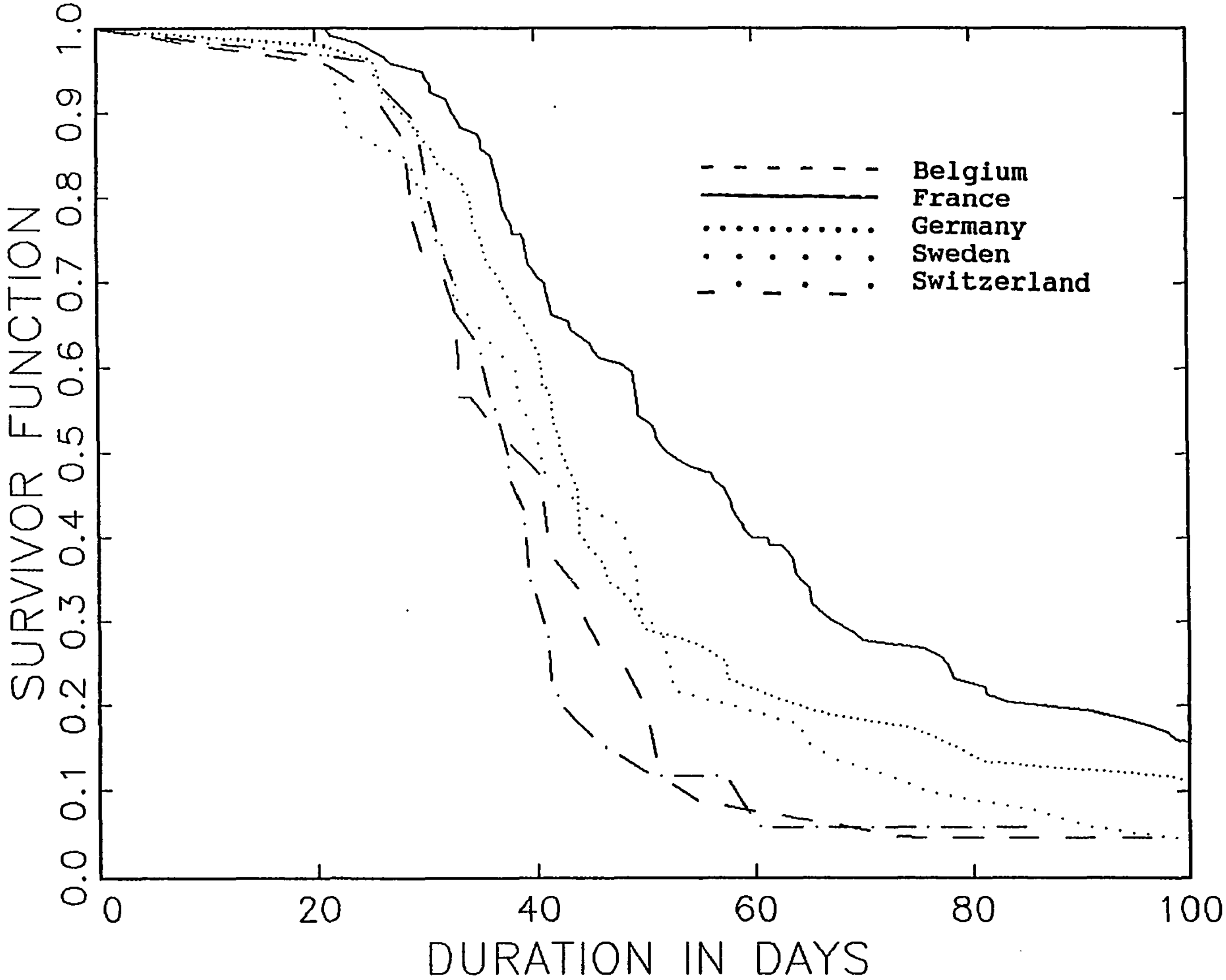


Figure 5: Survivor Function: Downtime for Repair

(Equipment Failure)

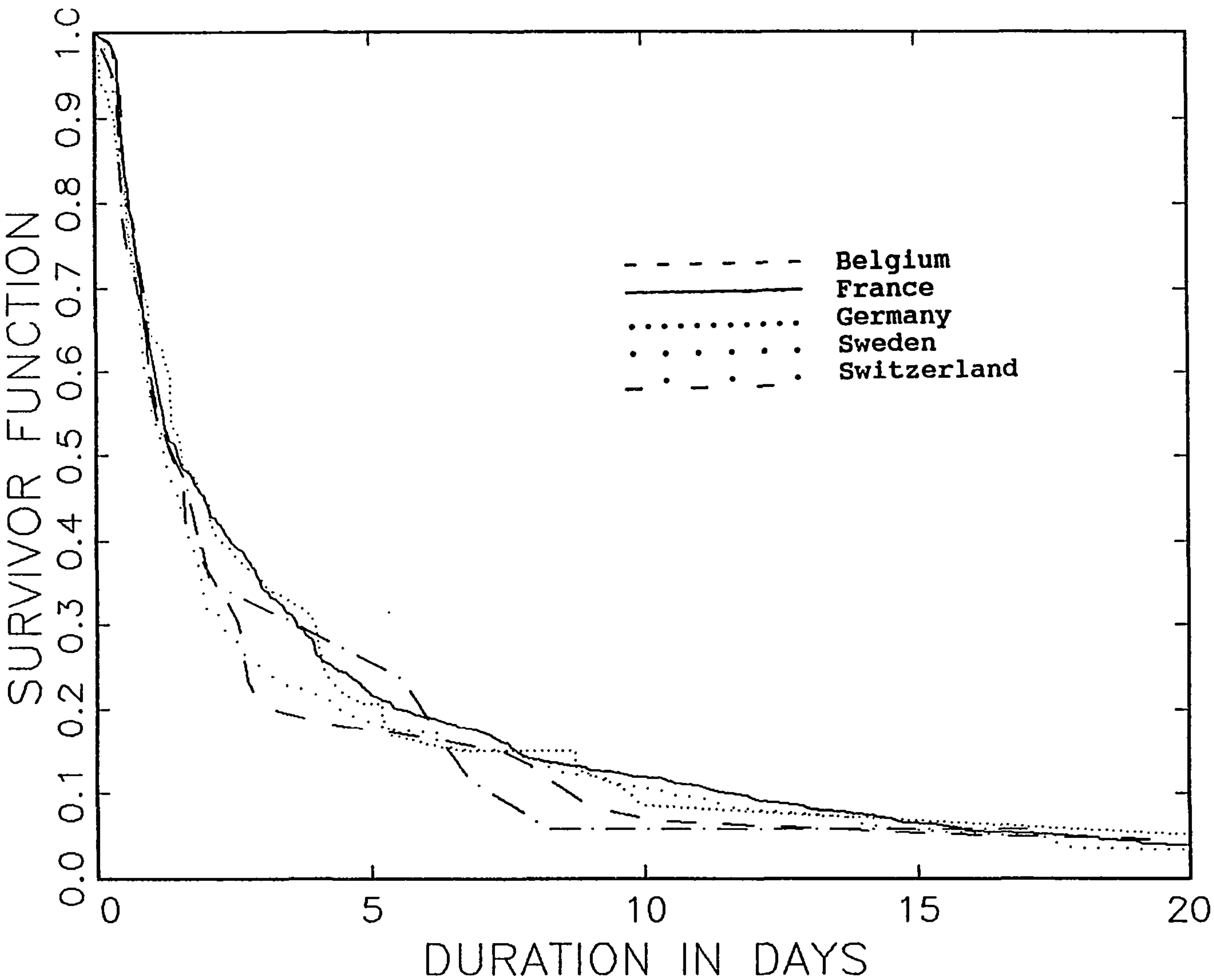
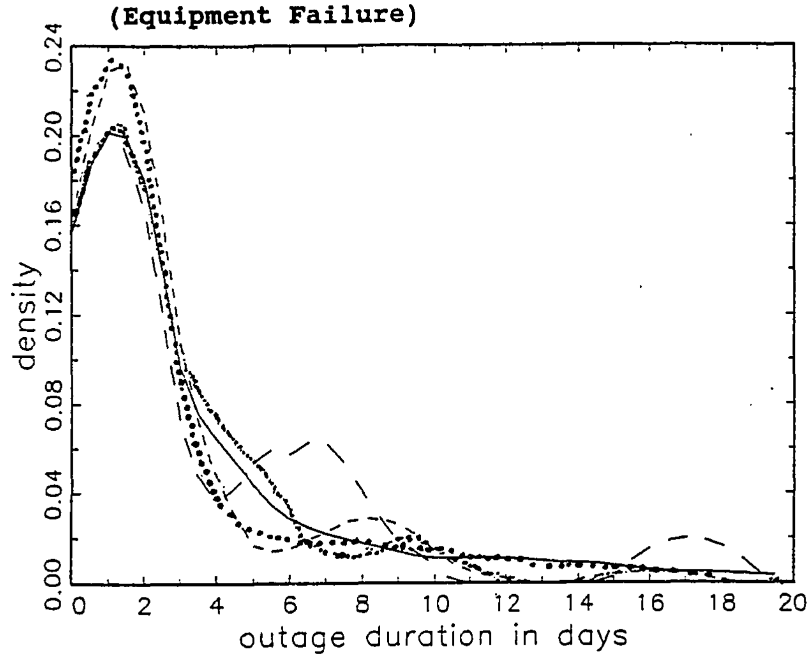


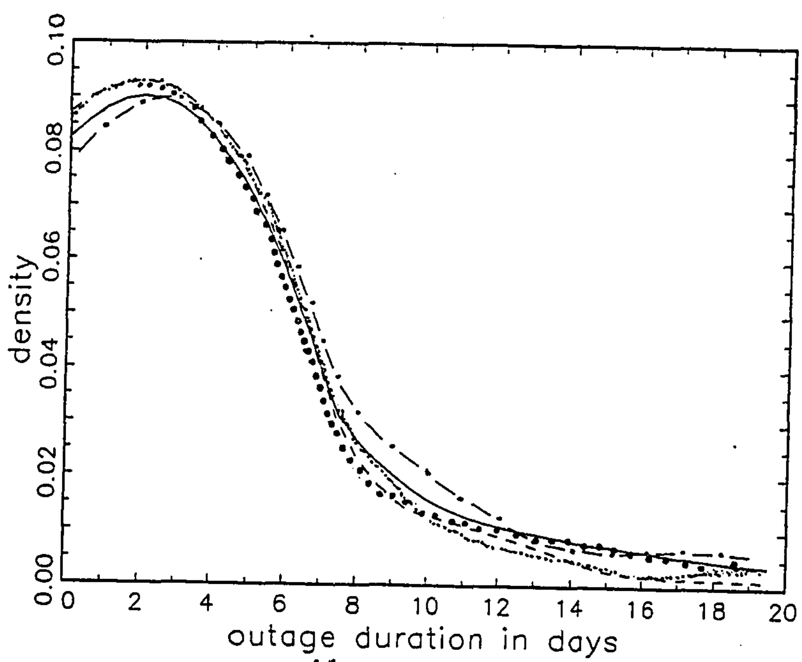
Figure 6: Density Plot: Downtime for Repair

(Equipment Failure)



h=1.0

- - - - - - Belgium
———— France
.... Germany
. . . . . . Sweden
_ . _ . _ . Switzerland



h=10.0

Figure 7: Density Plot: Downtime for Refueling

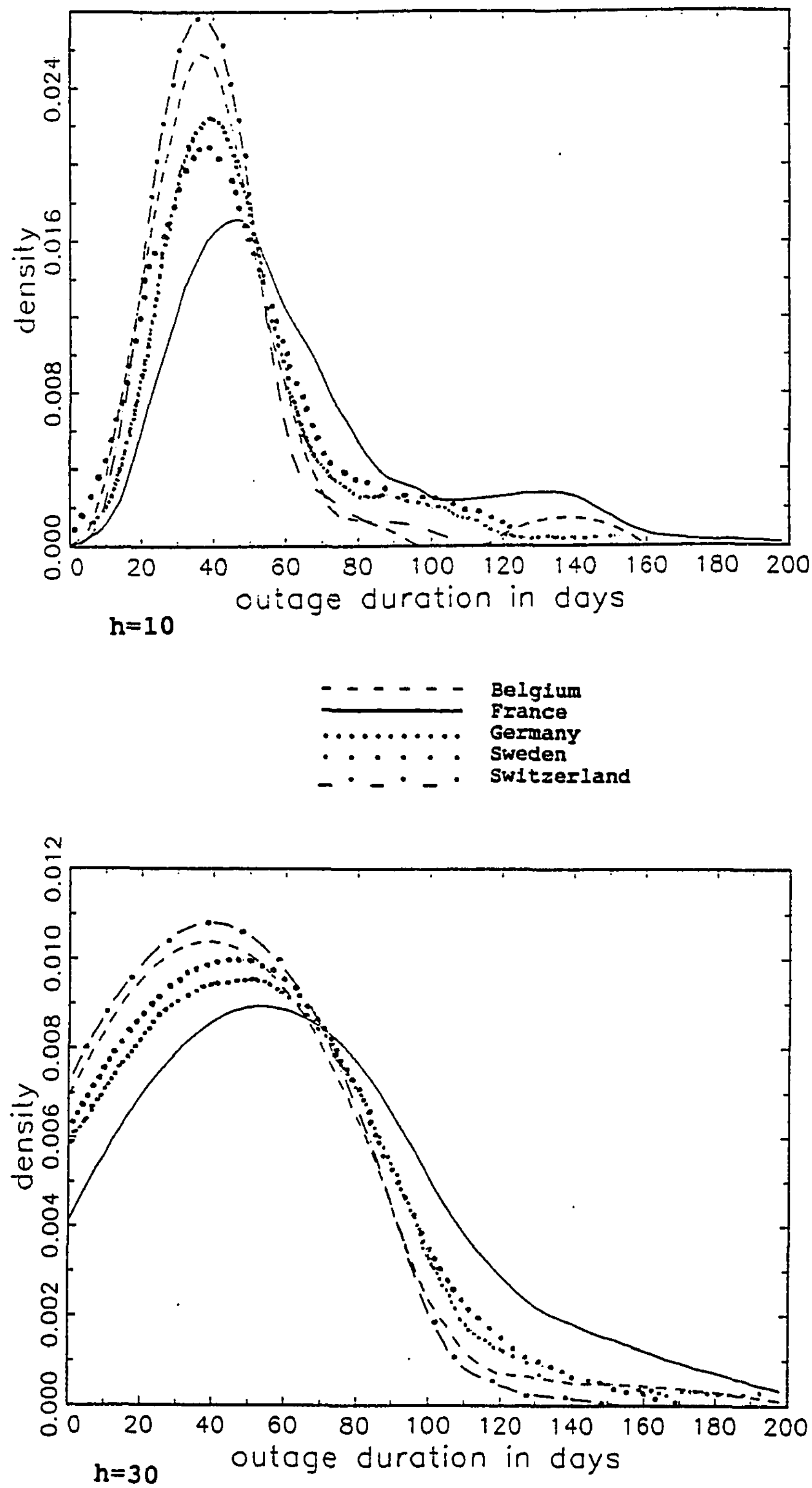


Figure 8: Density Plot: State 2

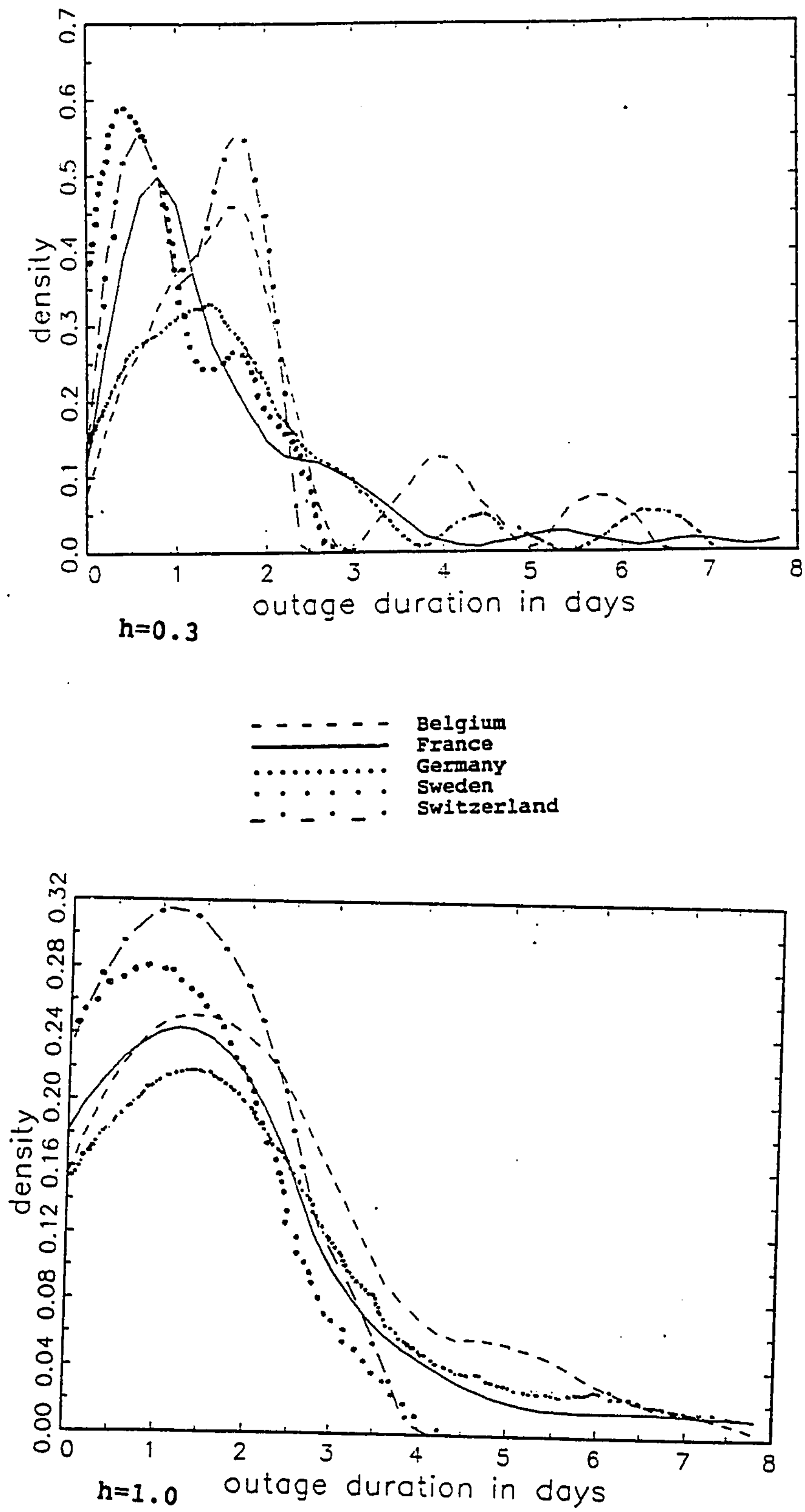


Figure 9: Density Plot: State 4

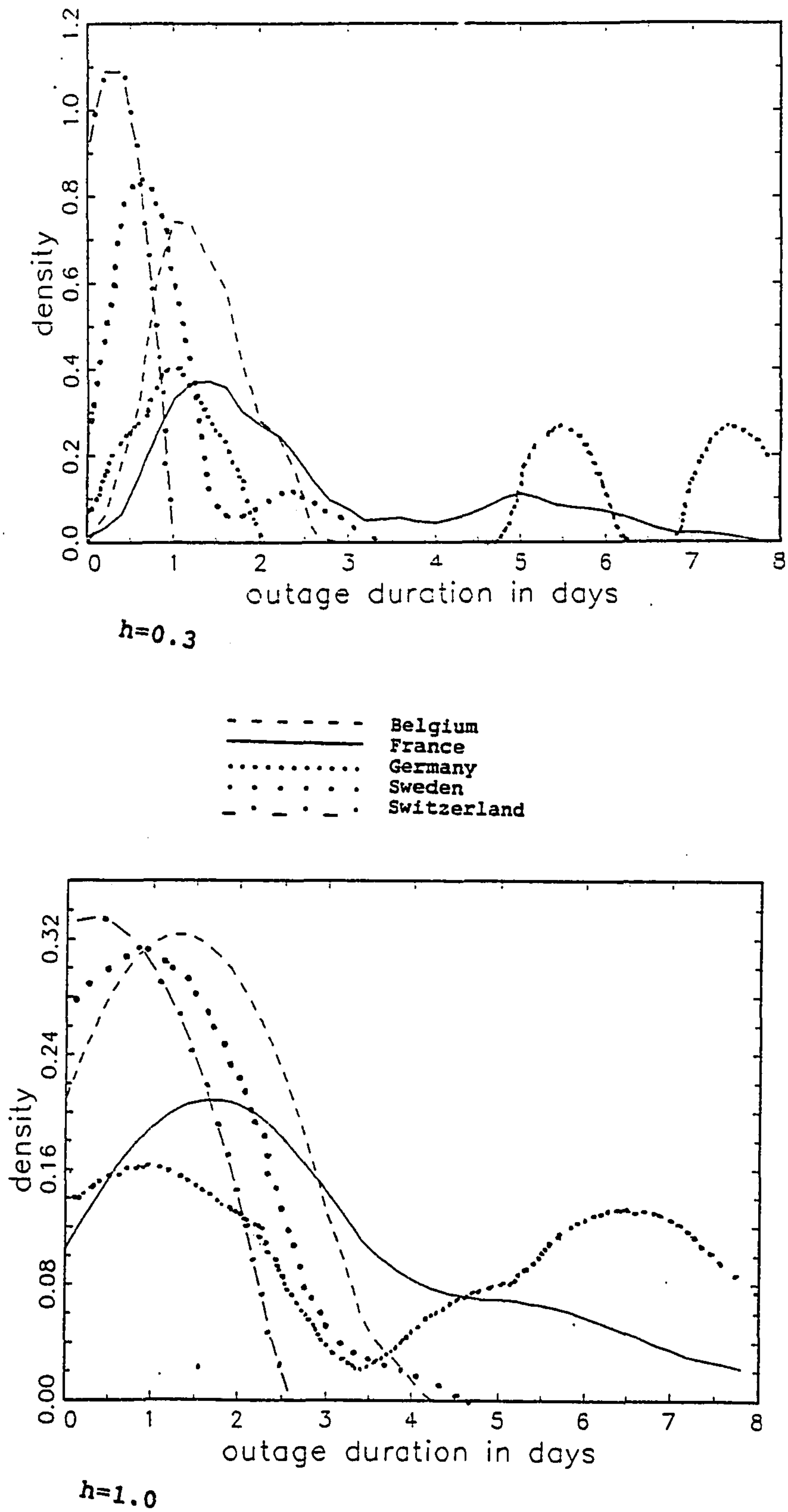


Figure 10: Survivor Function: Uptime until Equipment Failure, France New vs. Old

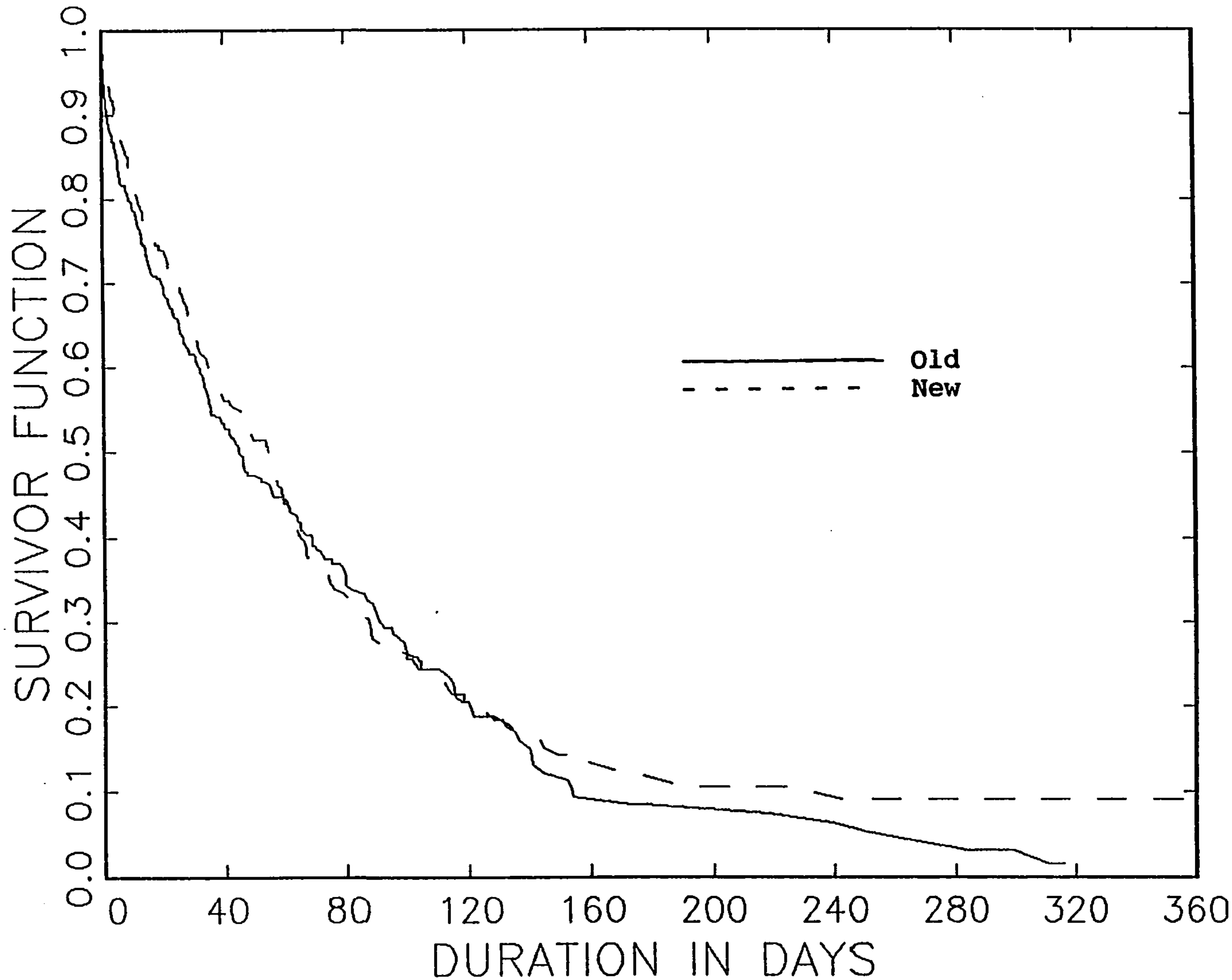


Figure 11: Survivor Function: Uptime until Equipment Failure, Sweden New vs. Old

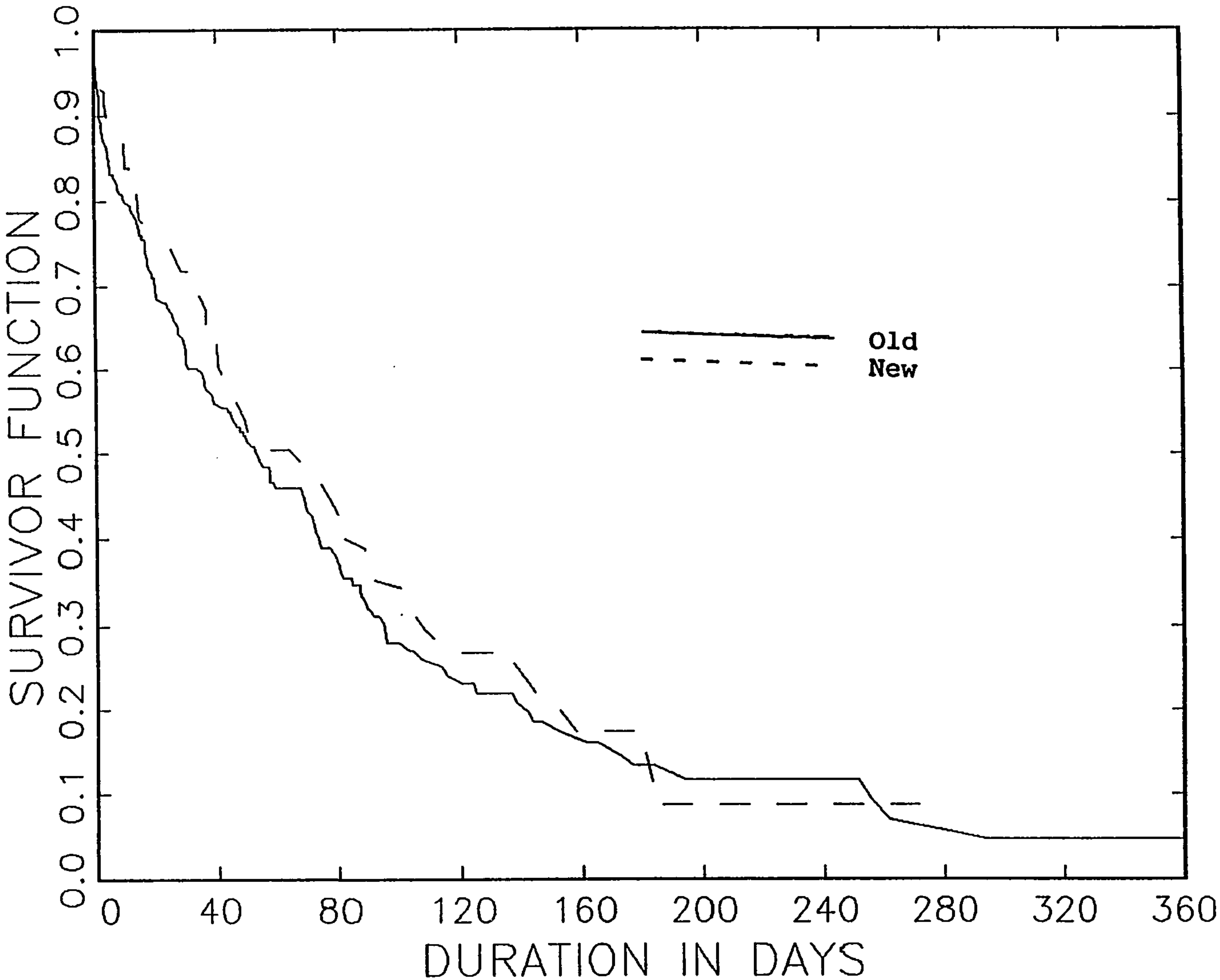


Figure 12: Survivor Function: Uptime until Equipment Failure (Reliability), German BWR vs. PWR

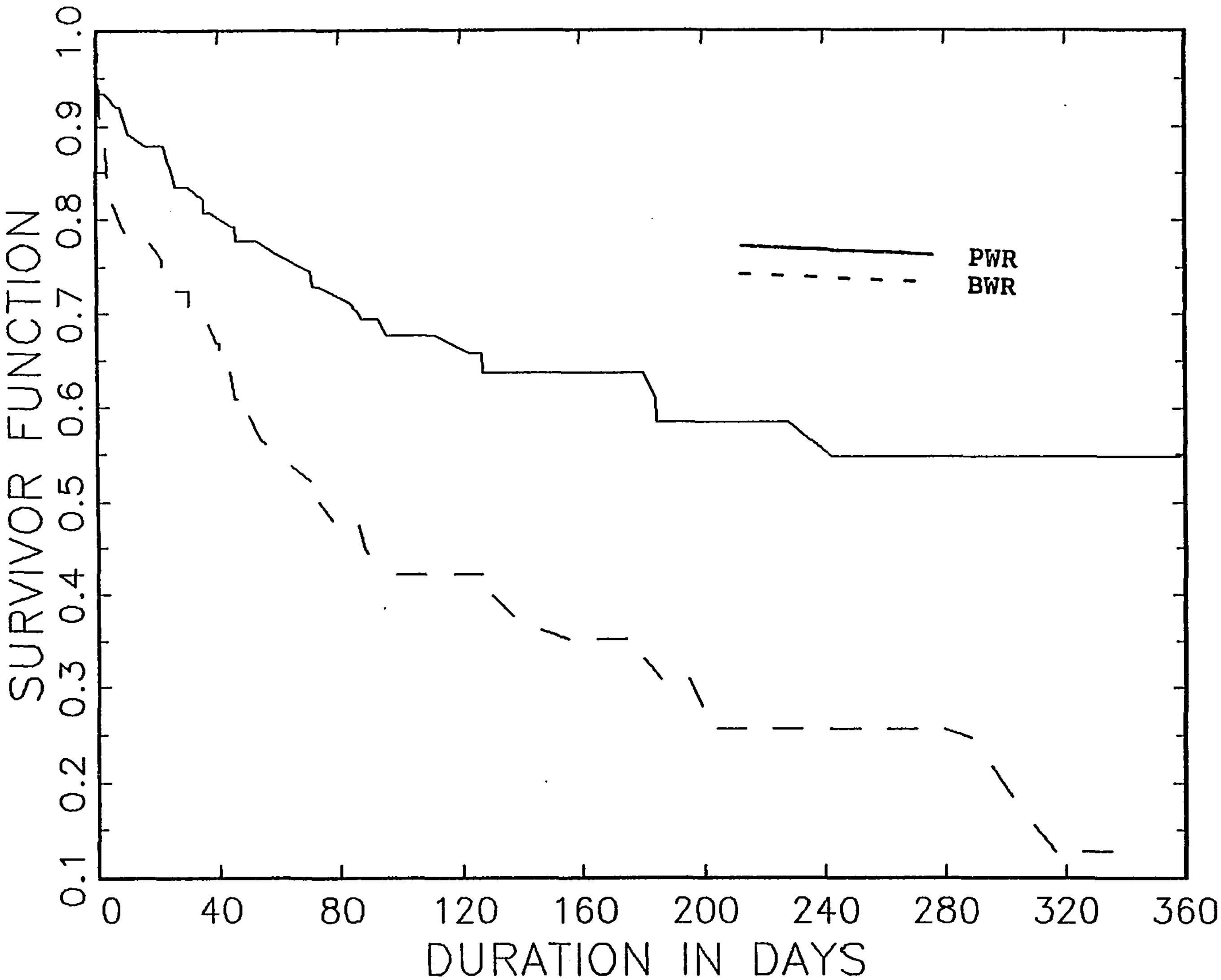
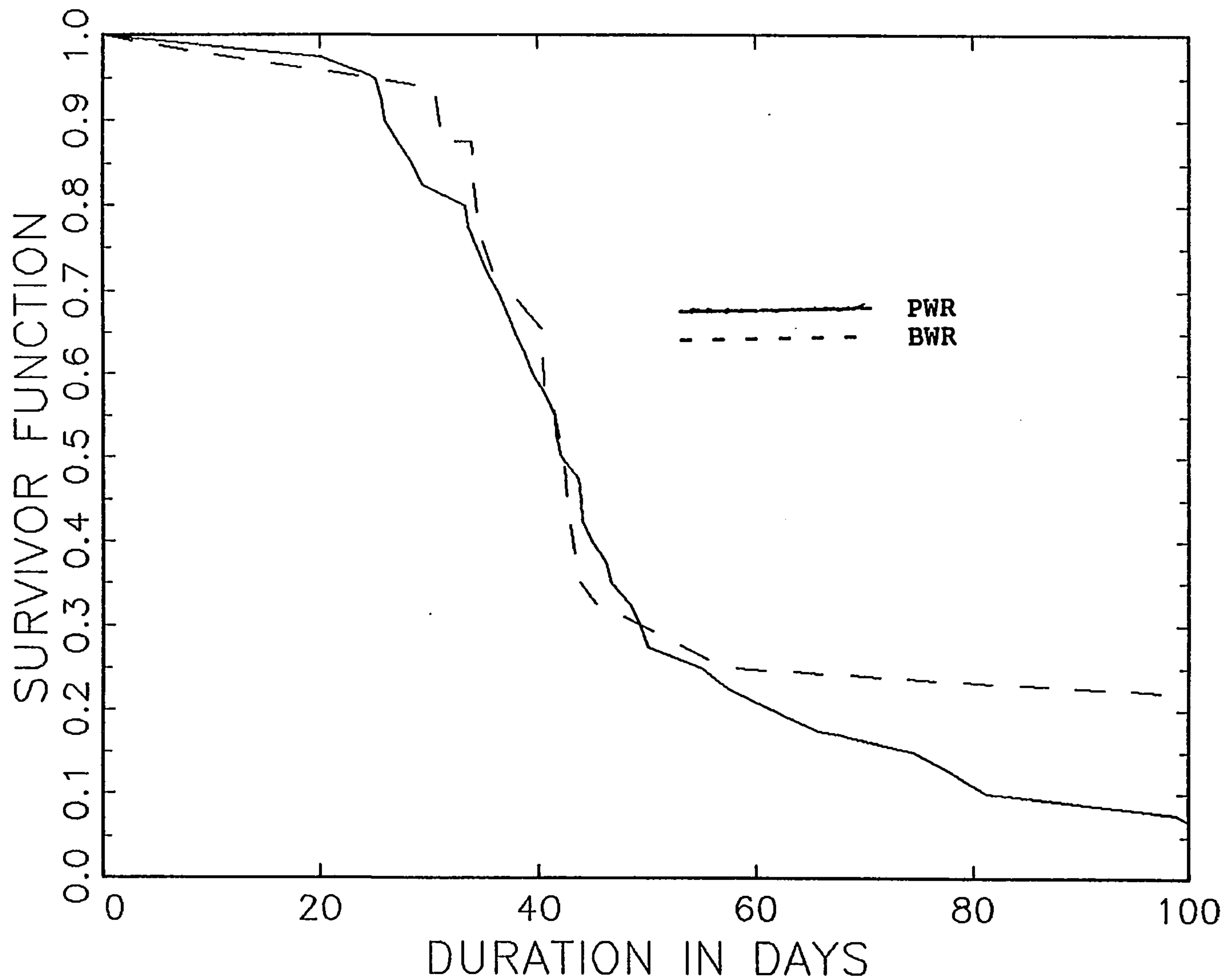


Figure 13: Survivor Function Downtime for Refueling

German BWR vs. PWR



2.2 Testing the renewal assumption

An assumption implicit in the preceding section as well as in other descriptive models of plant operations using duration analysis (Rothwell, 1989, Rothwell and Jensen, 1990, David, Rothwell, and Maude-Griffin, 1991) is that successive spells (the interval between two transitions) are independent. This is a questionable assumption which should be tested. Uptimes can be ended by a number of different causes (competing risks) and the plant manager's decision will generally not be independent of the condition of the plant[19]. Thus a behavioral model describing the relationship between the occurrence of planned and unplanned outages is desirable and will be developed in later chapters. However, the semi-Markov model and many behavioral models impose restrictions that can be tested without having to assume particular parametric forms for the distribution of spell durations.

The classic maintenance model in the field of operations research has the central assumption that every action regenerates the system completely, i.e. resets the hazard function[20]. In a stationary environment where the technology and the operator's preferences do not change, this assumption implies that both the

---

[19] The question may be less urgent for outage durations because outages always end with a return to the run state.

[20] Pierskalla and Voelker (1976) and Valdez-Flores and Feldman (1989) review the large literature on maintenance models for stochastically deteriorating systems. An early model that distinguishes between two types of action is the minimal repair model of Barlow and Hunter (1960).

distribution of failures and the distribution of censoring times (preventive maintenance/refueling) are stationary, i.e. statistically we are dealing with a semi-Markov model. Two such models in economics are Rust (1987) and Ryu (1990)[21].

Rank methods, briefly discussed in the previous section, can be an appropriate technique to test this assumption. These tests even work if the model is not completely stationary, for example, when energy prices or climatic conditions (plants are not refueled in the winter because of high electricity demand) influence the operators decision to take a plant down, as long as this source of nonstationarity does not systematically bias one sample. A systematic bias is unlikely since approximately the same number of observations from each year are contained in each sample in a comparison between two spells of different cycles. Among the countries in the data set, only France and Sweden have enough observations to make a non-parametric test of this kind worth considering. For France, it is possible to consider the two subsamples "old" and "young" plants. Observations on "young" plants--plants beginning commercial operation after January 1, 1981--are complete, the observations for "old" plants--plants beginning commercial operation before 1981--are incomplete for the years before 1981. The first operating cycle in a nuclear power plant is known to be different from the following cycles, and I

[21] Rust only considers one type of event (engine replacement), whereas Ryu considers planned (preventive maintenance) and unplanned (repair) events.

therefore only used observations following the first refueling outage after 1/1/1981[22].

The first null hypothesis tested is the one implied by the complete regeneration model, e.g. semi-Markov models or the behavioral models of Rust or Ryu:

H_{0a}: all operating spells have the same distribution

Table 8 reports the results of the test, which rejects the hypothesis H_{0a}.

A hypothesis weaker than complete regeneration at every event is regeneration during refuel outages. This statistical assumption is much more plausible than the assumption of regeneration at every event: utilities attempt to return the plant to the running state as quickly as possible in the case of an unscheduled event and such outages often last only a few hours. Repair and maintenance performed during such an outage is unlikely to affect more than a very small part of the complex system. During refuel outages, however, we observe extensive maintenance and inspection activities in which sometimes more than one thousand people, often from outside the plant, are involved. One testable hypothesis implied by the assumption of regeneration during refueling is the following:

[22] For old plants, this also avoids the problem of left censoring (considering the fuel cycle as one unit of analysis).

H_{0b}: the operating spell in each cycle has the same distribution

The samples compared are the first spell in one of four cycles. The null hypothesis in this case is not rejected, see table 9. Of the 36 possible two-way comparisons of cycles 1,2,3,4, for Swedish and new and old French plants and log-rank and Wilcoxon tests, only one rejects the hypothesis at the 10% level (log rank test, France old, cycles 2 vs 3). Testing all 4 samples simultaneously leads to one rejection at the 10% level (log rank test, Sweden).

Of course, this test has less power than the test of the first hypothesis because there are fewer observations. In order to assess the loss of power, I tested the first hypothesis again, but only used observations from cycle 1 for spells 1 and 2 and from cycle 2 for spells 3 and 4. This gives a comparable number of observations. Compare table 10 with table 8 which used all observations. It appears that the failure to reject hypothesis H_{0b} was not simply due to the loss in power.

As mentioned before, the rank tests may not be powerful if distributions are different, but one is not consistently larger than the other. However, this does not appear to be the cause for not rejecting the hypothesis either, as may be seen from figure

14[23]. Thus we may conclude that the assumption of regeneration following refuel outages is consistent with the empirical evidence and that fuel cycles are a relevant unit of analysis. However, the regeneration assumption at every event has been rejected and this raises the question: what is the relationship between successive failures? This question is taken up again in chapter 3 in the development of a parametric model.

[23] Only two representative plots are shown. I found only one crossing of survivor functions (new French plants, cycle 2 vs. cycle 4).

Table 8: Rank Tests of H_{0a} (χ^2-statistic, p-value)

comparison (degrees of freedom)	log rank test	generalized Wilcoxon test
France old spells 1-7 (6)	15.44 0.017*	16.88 0.010*
France old spells 1 vs 3 (1)	6.84 0.009*	6.73 0.010*
France old spells 2 vs 4 (1)	0.01 0.92	0.69 0.407
France new spells 1-7 (6)	7.31 0.293	9.48 0.148
France new spells 1 vs 3 (1)	1.21 0.271	2.96 0.086
France new spells 2 vs 4 (1)	4.08 0.043*	4.58 0.032*
Sweden spells 1-7 (6)	14.83 0.022*	24.47 0.000*
Sweden spells 1 vs 3 (1)	3.58 0.058	8.98 0.003*
Sweden spells 2 vs 4 (1)	1.51 0.220	3.82 0.051

* significant at 5% level.

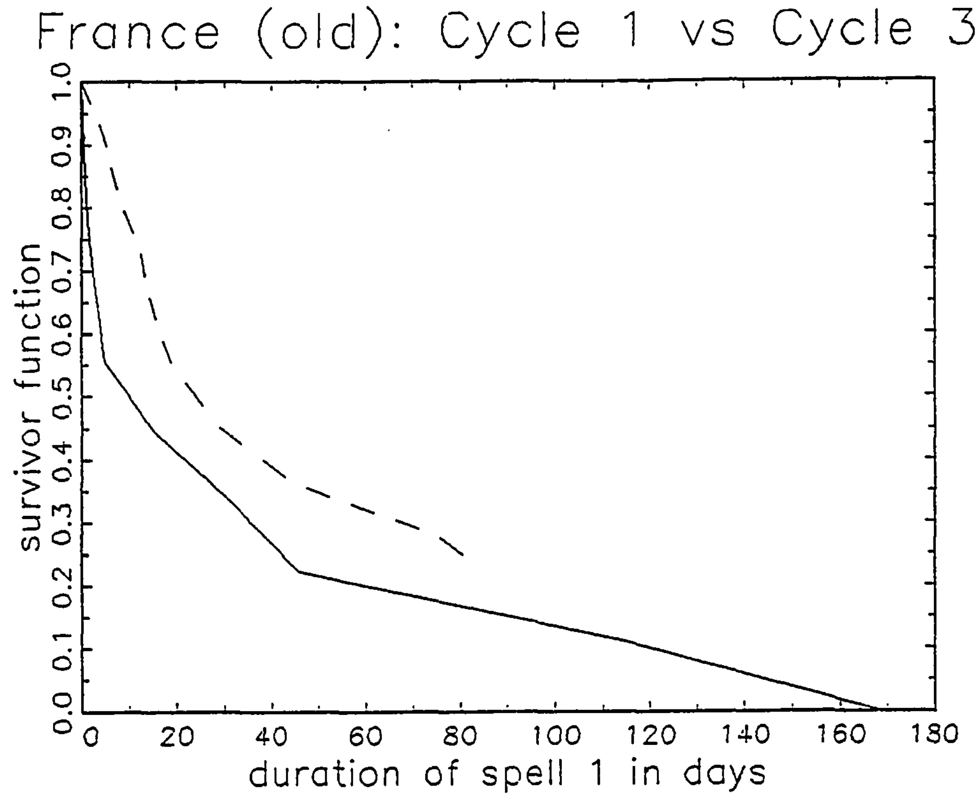
Table 9: Rank Tests of H_{0b} (χ^2-statistic, p-value)

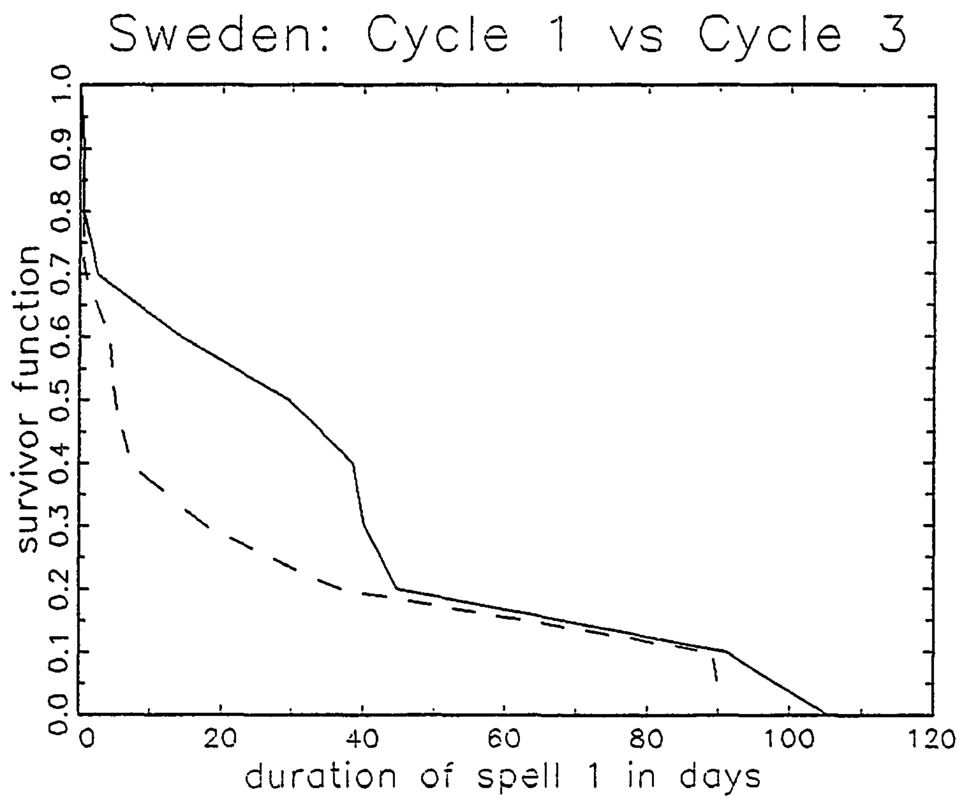
comparison (degrees of freedom)	log rank test	generalized Wilcoxon test
Sweden cycles 1-4 (3)	6.47 0.091	3.96 0.266
Sweden cycles 1 vs 3 (1)	1.32 0.250	0.95 0.331
Sweden cycles 2 vs 4 (1)	2.39 0.122	1.42 0.233
France new cycles 1-4 (3)	2.93 0.402	2.92 0.403
France new cycles 1 vs 3 (1)	2.17 0.141	1.41 0.235
France new cycles 2 vs 4 (1)	0.11 0.738	0.55 0.458
France old cycles 1-4 (3)	2.37 0.500	2.54 0.467
France old cycles 1 vs 3 (1)	0.03 0.87	0.68 0.410
France old cycles 2 vs 4 (1)	0.46 0.498	0.01 0.92

Table 10: Power Test

comparison (degrees of freedom)	log rank test	generalized Wilcoxon test
France old spells 1 vs 3 (1)	1.96 0.162	3.81 0.051
France old spells 2 vs 4 (1)	0.30 0.58	0.19 0.660
Sweden spells 1 vs 3 (1)	0.02 0.893	0.07 0.786
Sweden spells 2 vs 4 (1)	3.17 0.075	2.62 0.105

Figure 14: Do Survivor Functions Cross?



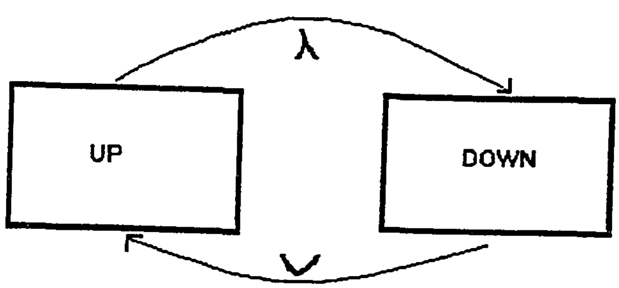


2.3 Continuous production processes and linear regression techniques

Many previous authors (e.g. Joskow and Rozanski, 1979, Lester and McCabe, 1988, Krautman and Solow, 1990) have analyzed power plant operations in a linear regression framework, usually by regressing availability or capacity factors on some variables of interest. A critical review of this literature can be found in David et al. (1988, Appendix E), which suggests that ignoring the dynamics of production may lead to very misleading results. This will be demonstrated formally in the present section.

Consider a simple two state Markov model[24]. This is sufficient to show analytically the inconsistency of the regression framework with a continuous production process.

Figure 15: A Two-State Model



[24] David in appendix B of David et al. (1988) considers such a model.

The point availability A(t) is the probability that the plant is up at time t. The interval availability therefore is defined as

$$A^*(t,T) := \frac{1}{T}\int_t^{t+T} A(u)\,du \qquad (12)$$

and the asymptotic, long-term, or steady-state availability

$$A^*_\infty := \lim_{t\to\infty}\frac{1}{t}\int_0^t A(u)\,du \qquad (13)$$

The state space S has two states, namely up and down, and the transition between up and down is assumed to follow a pure Markov jump process. The constant failure rate λ is the hazard of a transition from up to down, i.e. the probability density function for operating spells is that of an exponential distribution:

$$f(u) = \lambda e^{-\lambda u} \qquad (14)$$

Similarly, there is a constant repair rate ν describing the hazard from down to up. Solving for the probability of being in any one state is a simple exercise in Markov models (e.g. Hoel, Port, and Stone, 1972, Ch.3). In terms of availability, we obtain the differential equation

$$\frac{dA(t)}{dt} = \nu(1-A(t)) - \lambda A(t) \qquad (15)$$

Using the integrating factor $e^{\lambda+\nu}$ and the initial condition that the plant begins operation at time 0 in the up state (A(0) = 1), the formula for availability becomes

$$A(t)=\frac{\nu}{\lambda+\nu}+\frac{\lambda}{\lambda+\nu}e^{-(\lambda+\nu)t} \tag{16}$$

This yields the interval availability

$$A^{*}(t,T)=\frac{\nu}{\lambda+\nu}+\frac{\lambda e^{-(\lambda+\nu)t}}{(\lambda+\nu)^{2}T}[1-e^{-(\lambda+\nu)T}] \tag{17}$$

and the asymptotic availability

$$A^{*}_{\infty}=\frac{\nu}{\lambda+\nu} \tag{18}$$

Now consider how the existing regression literature deals with data generated by continuous production processes. The most recent study by Krautmann and Solow (1990) estimates a linear regression model of the following form, using yearly data on individual plants in the U.S.[25]

$$\frac{1}{A_{it}}=\alpha' Z_{i}+\beta' Z_{i}e^{-t}+\epsilon_{it} \tag{19}$$

[25] The usual metric of learning in all the regression papers are changes in the capacity factor. The capacity factor is not an ideal measure since it also measures other economic effects. Joskow and Rozanski (1979) are aware of this problem, but they claim that "differences in capacity factors across plants due to 'load following' is not a problem." The availability factor would nevertheless be a more appropriate measure in these papers.

where A_{it} is the actual interval availability of plant i in [t-1,t], α and β are vectors of parameters to be estimated, Z_i is a vector of characteristics corresponding to plant i, and ε_{it} are iid random variables with mean zero. For the present purpose it is enough to consider a panel of identical plants (Z_i =1) observed over a finite time period [0,T] when the data is generated by the stationary two-state model with hazards ν and λ. If the "reduced" form (19) were appropriate, the coefficient measuring learning should *indicate no learning effects* (β=0). However, this is *not the case*: asymptotically (as the number of plants goes to infinity), the estimate of β converges to

$$plim\beta = \frac{T\sum_{t=1}^{T} \frac{e^{-t}}{A^*(t-1,t)} - \sum_{t=1}^{T} e^{-t} \sum_{t=1}^{T} \frac{1}{A^*(t-1,t)}}{T\sum_{t=1}^{T} e^{-2t} - \left(\sum_{t=1}^{T} e^{-t}\right)^2} \tag{20}$$

which is never zero unless the interval availability $A^*(t-1,t)$ does not depend on t. For a numerical example, let λ=.5 and ν=1, which implies an asymptotic availability of 67%. For a 2-year panel, the probability limit of β is -1.10, for a 10-year panel, it is -0.81, etc.. Thus the regression framework leads to the "discovery" of a strong negative relationship between experience and availability.

Joskow and Rozanski (1979) estimate "learning" both by operators and by suppliers of nuclear power plants. They assume the existence of an "asymptotic capacity factor" which is approached by an increasing annual plant capacity factor ("learning during operations"). Suppliers build better plants over time, i.e. plants with higher asymptotic capacity factors ("learning in production"). Ignoring the effect of other covariates such as vintage, their model is of the form:

$$\ln(A_i) = \alpha + \frac{\beta}{X_i} + \epsilon_i \qquad (21)$$

where X_i is the cumulated output of plant i up to but not including the year for which an average availability factor A_i was observed. In the absence of learning effects, this model implies that X_i and A_i are statistically independent. But this is not the case in a continuous process. A simple example with discrete duration distributions should suffice to point out the problem, an analytic derivation of the relationship between X and A is rather involved for continuous distributions. Consider a large number of plants that have been in operation for two periods, i.e. X is experience measured as availability in period 1, A is the availability in period 2. Plants start operation in the up state, up times have a duration of.5 with probability .5 and of 1 with probability .5, downtimes always last .5. Although this is a stationary technology and no learning occurs, ß estimated by OLS in equation (21) converges in probability to .35.

Why do these regression models fail to recognize the stationarity of the underlying technical process? The reason is that initial conditions influence the results of regression analysis. The foregoing calculations assumed that observations begin with a startup (as is the case for the data set used in this study). On the other hand, assuming that observations start at the beginning of a downtime changes the expected value of the estimated parameter measuring "learning". Clearly, this is a highly unsatisfactory situation: whether one finds that availability increases with experience or decreases with experience depends on how the data is collected.

Spurious findings can be avoided by using models that allow for the dynamics of the process. Standard duration and failure time models, used by Rothwell (1989), Rothwell and Jensen (1990), or David, Rothwell, and Maude-Griffin (1991), for example, are not sensitive to starting conditions and recognize the stationarity of the technology in this environment. Note that the sensitivity of regression analysis results to starting condition is a more fundamental problem than the criticism in David et al., which considers spurious regression results due to the possibility of unobserved heterogeneity (essentially a missing variable problem) and the corresponding correlation of experience and availability.

2.4 Estimating learning in a reliability growth model

The previous section showed that models which fail to address the dynamic aspects of production may give rise to spurious results. This section suggests a different model to estimate learning effects which complements the line of research in several regards. First of all, the focus is on reliability improvements over time, measured by the hazard of experiencing emergency shutdowns, rather than on plant availability studied by previous researchers[26]. Secondly, this section attempts to disentangle the effects of different possible causes of reliability improvements such as weeding out installation or design errors, "spill over" effects from the operation of other plants, and improvements due to experience as measured by plant age or total output. Thirdly, data generated by stochastic processes comes in two forms: duration between events and event counts. The duration between events has been analyzed in the preceding sections, here I use count data on unplanned shutdowns to investigate changes in plant reliability.

This section qualifies the results of section 2.2 in which the renewal assumption following refuel outages could not be rejected. I find that new plants beginning commercial operations display "learning" effects in that the reliability increases. This introduces an element of nonstationarity apparently inconsistent with renewal during refueling. In other words, plant time, not just

[26] The use of the words "reliability" and "availability" here is consistent with the engineering literature (e.g. Lewis, 1987). In the economics literature, the word "reliability" has occasionally been used incorrectly to denote "availability" (e.g. Joskow and Rozanski, 1979).

cycle time, matters. However, note that the model in this section implicitly assumes a strict Markov model for uptime durations within each fuel cycle, which is only an approximation to a more complex model (chapter 3). On the other hand, the test of the renewal assumption in section 2.2 used the duration of the first spell and thus may not have picked up effects occurring later in the fuel cycle. The question of reliability growth is mainly of interest for the first years of plant operations, more experienced plants do not display similarly statistically significant effects and the renewal assumption at the beginning of a fuel cycle appears to hold without qualifications for them.

The reliability of a plant describes how likely the plant is to experience an unplanned shutdown (the intensity or hazard of a shutdown), generally an equipment failure leading to a scram. Reliability is therefore closely associated with plant safety, an important performance goal in nuclear plant operations. Because emergency shutdowns are exits from the running state, a statistical model of reliability could alternatively be formulated in terms of uptime durations interrupted by both planned and unplanned outages. Unplanned outages have very short average durations. Their overall contribution to availability is therefore minor (table 20, Chapter 4). Thus the model sheds light on a neglected aspect of plant performance. Chapter 4 will combine both aspects of plant performance (reliability and availability) in a unified behavioral model.

The dependent variable of interest is the number of unplanned shutdowns during an operating cycle (the time between two successive planned refuel outages). The length of operating cycles and the uptime within a cycle (shutdowns can only happen when the plant is up) differ from cycle to cycle and this sets the model apart from previous economic applications of event count data which had identical durations for all observations (Hausman, Hall, and Griliches, 1984, Arora and Gambardella, 1990). The data set analyzed here contains information on 164 fuel cycles of nuclear reactors which began commercial operations in December 1980 or later in five European countries[27].

Most studies of learning curves or progress function consider plant specific learning as a function of a single measure of experience. However, plant specific experience could be measured in several ways, all of which may be important. The number of previous unplanned shutdowns, for example, is relevant when unplanned shutdowns are caused by rectifiable installation and design errors. A general assumption in software reliability models (e.g. Jelinski and Moranda, 1972, Musa, 1975), is that program crashes lead to the detection and permanent removal of a bug, thus improving the program's reliability. By analogy, the number of previous fuel cycles matters if major retrofits (or "bug" removals) can only be performed during refuel outages; total up time in nuclear power

---

[27] See table 11 for definitions of variables and summary statistics.

plants, which are typically baseload plants, provides a measure of output -- often equated with experience. Similar arguments could be given for plant age, total unplanned downtime, or total downtime (plant age minus total uptime). In addition, reliability improvements may not only be a function of plant specific experience but also of total experience within a country or within a cluster of reactors at one site. For each of the plants in the sample, I constructed all the measures of plant specific experience just mentioned. Because there is little information on plants that came on line before 1981, only two variables could be constructed to test for spill-over effects: reactor years for a multiple unit site and reactor years for a country.

Unplanned or emergency shutdowns occur randomly and independently in time while the plant is running. Furthermore, this intensity differs between plants and fuel cycles, but is constant within one fuel cycle[28]. There cannot be any unplanned shutdowns during outages and the intensity is zero while the plant is down. Assuming that major changes in the intensity of shutdowns occur only between fuel cycles is not unreasonable as a first approximation: many substantial changes affecting reliability can only occur during the long downtimes for planned refueling and maintenance and they cause discrete changes in reliability, for example

[28] This assumption ignores the possibility of regular changes within a fuel cycle (see chapter 3) and is only made to investigate the broader question of reliability improvements over time.

- the installation of new (and probably better and more reliable) hardware
- the installation of new monitoring devices (with the aim of improving the control of the production process).

These assumptions give rise to a Poisson model. Let $\lambda(i,x_{in})$ denote the intensity of unplanned outages in the i'th fuel cycle for plant n where x_{in} are covariates. The contribution of an observed fuel cycle with k unplanned outages and a total uptime of T to the likelihood function is

$$L(k)=\frac{e^{-\lambda_{in}T_{in}}(\lambda_{in}T_{in})^k}{k!} \tag{22}$$

The intensity of unplanned outages is parametrized as

$$\lambda_{in}=\exp(x'_{in}\beta) \tag{23}$$

The gradient and Hessian of the log likelihood function are

$$\begin{aligned}\frac{\partial L}{\partial\beta_l}&=\sum_i\sum_n\left[kx_{lin}-x_{lin}T_{in}\exp(x'_{in}\beta)\right]\\ \frac{\partial^2 L}{\partial\beta_l\partial\beta_m}&=\sum_i\sum_j\left[-x_{lin}x_{min}T_{in}\exp(x'_{in}\beta)\right]\end{aligned} \tag{24}$$

Equation (24) shows that the log likelihood function is globally concave and thus all standard maximum likelihood algorithms work very well.

Table 12 lists results for the Poisson model for various subsets of regressors. The first goal was to determine which of the plant specific measures of experience are most important. If only one variable measuring experience is included, as in many previous studies, then the best one in terms of increasing the likelihood function is total up time (SUMUP, model A). Both the number of the fuel cycle (CYCLE) and plant age (AGE, model B) perform well too, but not the number of previous outages (SUMFAIL) or total downtime (SUMDOWN). The coefficient is negative, thus indicating that more experience is correlated with higher reliability. An argument that might be raised against total up time as a regressor is its correlation with plant quality if there is uncontrolled plant heterogeneity. In particular, a less reliable plant has less uptime that a more reliable plant of the same age. But because the contribution of unplanned outages to unavailability is small, this effect is likely to be minor if it exists at all. Nevertheless, all regressions were done twice to avoid spurious results: once with the unquestionable exogenous variable AGE and once with SUMUP. As in models A and B or models E and F, the difference in the likelihood between the using either SUMUP or AGE as regressors is relatively small (although the models with SUMUP have a larger likelihood), the magnitude of all other coefficients and standard errors remains virtually unchanged, and the predicted effects of SUMUP and AGE are always proportional to each other, negative, and statistically significant (at 5% or 1%, depending on the number of additional regressors). The correlation between SUMUP and AGE is very high and if both are included simultaneously, one of the coefficients becomes positive and both coefficients become statistically insignificant (at 10% or 5%, depending on the number of regressors, see model C for a typical example). If CYCLE is included the coefficient of SUMUP changes signs, otherwise the coefficient of AGE changes signs. Although CYCLE has a negative coefficient if neither AGE nor SUMUP are included, the sign can change when AGE or SUMUP are added. The reason for this is quite clear: SUMUP, AGE, and CYCLE are good measures of experience, but they are highly correlated and the effects on reliability are unlikely to be linear.

Using interaction terms of experience and country specific dummies allows to test whether learning paths differ across countries. There are very substantial differences in the levels (all country dummy variables have negative coefficients, indicating that France has the least reliable plants), but the null hypothesis that reliability improvements are the same in all countries cannot be rejected[29]. Furthermore, the results are very similar if only data on French reactors is selected (the country with the largest number of observations). For example, compare models A' and B', which only use observations on French reactors, with models A and B.

Although both plant age and the number of the fuel cycle are

---

[29] Under model B (experience is measured by plant age), the mean log likelihood with interaction terms is -2.1814. The likelihood ratio statistic, asymptotically Chi-squared distributed with 4 degrees of freedom, has a value of 3.13.

important variables by themselves, the second most important variable once total up time is included is country specific experience (EXPER, for example, model D). The coefficient is significantly negative (at the 1% level for almost all specifications) and indicates the importance of learning from the experience in other plants. Site specific experience (EXPERSITE) may or may not be important as an additional variable. The negative coefficient indicates that there are improvements associated with the experience of other plants at one site in addition to their contribution to country experience. Although the t-statistic is larger than one in absolute value, it was not significant at the 10% level for any specification. Not important are total downtime or the number of previous outages; the null hypothesis that they are uncorrelated with reliability cannot be rejected for any model estimated. This does not exclude the possibility that an error detection model is useful to model reliability changes within a fuel cycle which is investigated in the next chapter.

Larger plants are less reliable than smaller plants regardless of learning effects. The coefficient on capacity (CAP) is significantly negative (1%) even after controlling for plant age and experience[30]. Models E and F contain additional quadratic terms to allow for nonlinear effects of experience. None of the results discussed are affected.

---

[30] This is conventional wisdom, but evidence so far had been based on descriptive statistics and was therefore not convincing: more recent plants tend to be larger and the results show that reliability improves.

It is unlikely that the observed covariates capture all the differences between plants and fuel cycles. Fixed and random effects models can deal with this problem of unobserved heterogeneity, although the fixed effects model is not useful for this question because many plants contribute only one or two observations. However, before introducing unobserved heterogeneity in may be important to reconsider the implicit assumption of stationarity within the fuel cycle. This is done in chapter 3 which investigates time dependence of the intensity function of unplanned outages within the fuel cycle, as well as the dependence of downtime durations on observed plant characteristics.

TABLE 11: Definitions of Variables

variable name	definition	mean (std.dev)
FAIL	number of unplanned shutdowns in cycle (except exogenous causes)	3.63 (3.12)
CYCLE	number of cycle	2.65 (1.49)
UPTIME	total up time in cycle (months)	9.57 (4.32)
DOWNTIME	total down time in cycle (months)	0.62 (1.28)
SUMUP	total up time during life time (months)	17.20 (15.93)
SUMDOWN	total unplanned down time during life time (months)	1.56 (2.46)
SUMFAIL	total number of unplanned outages during life time (except exogenous causes)	8.09 (8.83)
AGE	age of plant at beginning of cycle (years)	1.88 (1.70)
DATE	calendar date of beginning of refueling	84.66 (1.65)
EXPER	experience with commercial PWR reactors in country at beginning of cycle (reactor years)	80.03 (38.82)
EXPERSITE	only for multiunit sites: experience at plant site (reactor years)	8.08 (10.52)
CAP	generating capacity in 1000 MW	0.96 (0.13)
BWR	= 1 if Boiling Water Reactor = 0 otherwise	0.17 (0.38)
SITE	= 1 if several units on site = 0 otherwise	0.93 (0.25)
B	= 1 if Belgian reactor = 0 otherwise	0.05 (0.23)
F	= 1 if French reactor = 0 otherwise	0.70 (0.46)

D	= 1 if German reactor = 0 otherwise	0.08 (0.27)
S	= 1 if Swedish reactor = 0 otherwise	0.15 (0.36)
CH	= 1 if Swiss reactor = 0 otherwise	0.02 (0.13)

164 observations for all countries (114 observations for France)

Table 12: Reliability Growth Model - Results

variable	A	B	C	D
mean logl	-2.1778	-2.1909	-2.1672	-2.0579
Constant	-0.5109 (0.0602)	-0.5116 (0.0616)	-0.5418 (0.0629)	-2.9297 (0.4794)
CYCLE				0.2199 (0.1528)
SUMUP	-0.0192 (0.0030)		-0.0566 (0.0202)	-0.0460 (0.0216)
AGE		-0.1689 (0.0270)	0.3463 (0.1843)	0.2251 (0.2068)
EXPER				-0.0072 (0.0016)
EXPERSITE				-0.0102 (0.0087)
CAP				2.6805 (0.4830)
BWR				-0.0306 (0.2881)
B	-0.2829 (0.1753)	-0.3137 (0.1757)	-0.2108 (0.1797)	-0.3686 (0.1982)
D	-2.0041 (0.3200)	-2.0307 (0.3202)	-1.9358 (0.3221)	-2.6532 (0.3674)
S	-0.4633 (0.1331)	-0.4812 (0.1331)	-0.4332 (0.1338)	-0.4727 (0.2437)
CH	-0.4748 (0.3814)	-0.4831 (0.3815)	-0.4383 (0.3819)	-0.5796 (0.4682)

164 observations for all countries (114 observations for France)

Table 12: Reliability Growth Model - Results, continued

variable	E	F	A' France only	B' France only
mean logl	-2.0566	-2.0679	-2.3736	-2.3945
Constant	-2.8163 (0.5164)	-2.6564 (0.5289)	-0.5055 (0.0627)	-0.5133 (0.0641)
CYCLE	0.3043 (0.1418)	0.1441 (0.1489)		
SUMUP	-0.0372 (0.0173)		-0.0197 (0.0032)	
SUMUP^2	0.1922 (0.2074)			
AGE		-0.1833 (0.1538)		-0.1677 (0.0294)
AGE^2		15.72 (18.58)		
EXPER	-0.0102 (0.0049)	-0.0100 (0.0050)		
EXPER^2	0.0186 (0.0320)	0.0166 (0.0324)		
EXPERSITE	-0.0078 (0.0084)	-0.0122 (0.0087)		
CAP	2.5996 (0.4939)	2.6040 (0.4961)		
BWR	-0.1000 (0.2814)	-0.1284 (0.2826)		
B	-0.4296 (0.1905)	-0.4476 (0.1946)		
D	-2.5847 (0.3740)	-2.6219 (0.3725)		
S	-0.4584 (0.2477)	-0.4358 (0.2448)		
CH	-0.5190 (0.4694)	-0.4844 (0.4680)		

164 observations for all countries (114 observations for France)

3. Choosing between competing duration models: an analysis of up- and downtimes

The previous chapter described the production in nuclear power plants in Europe as a stochastic process with alternating "up" and "down" times. A difficulty discovered in chapter 2, and which will be demonstrated in the simulation study in section 3.3, is that actual data exhibits more structure than consistent with Markov models. This chapter probes deeper into the structure by considering theoretical models of complex production systems, such as nuclear power plants, by developing a detailed parametric representation individual spell durations, and by comparing the properties of parametric models with actual data through simulations. It reaches the following main conclusions:

a) Downtime durations

In the representation of individual plant data on the duration of outages, it is quite difficult to find models that pass goodness-of-fit tests even when one distinguishes among the main outage causes. It is not sufficient to choose a parametric hazard model that allows for duration dependence in the hazard rate of terminating an outage, such as the Weibull model. Unobserved heterogeneity is substantial, in particular where outages caused by equipment failures are concerned. Unless more specific information becomes available, such as a technical description of the type of equipment failure, this heterogeneity needs to be controlled by random coefficient models. The exponentiated quadratic polynomial is the best baseline hazard function for durations of equipment failure outages. By contrast, the Weibull distribution performs well in describing the observed pattern of variability in refuel durations.

b) Uptime durations

Standard failure time models are unsatisfactory as a statistical description of the durations of sequential operating spells in nuclear power plants. The "process" model introduced in this chapter as an alternative to the "failure time" model of duration analysis is theoretically and empirically more appealing. Its theoretical advantage is that it can describe the sequence of failures of repairable systems that can be restored to a satisfactory operating condition without having the properties of a completely new system. The likelihood function derived here corresponds to a nonstationary point process with regressors. The estimation in the chapter goes one step further by taking into account that events (unplanned outages) are not of zero durations.

c) Simulations

Simulating plant operations using the process model developed in this chapter gives rise to a sequence of up- and downtimes which closely resembles observed sequences. The time paths implied by the Markov model of chapter 2 are, by contrast, not consistent with the data. In particular, the point availability converges too quickly to its stationary value.

3.1 Modeling Downtime Durations

Since nuclear power plants are mainly used to cover base load requirements, the availability of a station is a main determinant of its productivity. Availability itself is a direct result of both planned and unplanned outages. The probability of outages is governed by the reliability of the plant, analyzed in the next section (3.2). In general, econometric estimation is much simpler for outage durations than for run durations because there are essentially no censoring or competing risk problems (see section 2.1). Since little is known about the statistical properties of nuclear power plant operations, it is important to compare the results of different distributional specifications, which is the purpose of this section.

A detailed analysis of individual spells, giving rise to much more complex distributions for aggregate data than the descriptive models of chapter 2, uses regression type duration models with and without unobserved heterogeneity. A sensitivity analysis is important because estimates may not be robust and conclusions based on restrictive parametric models may be misleading. Downtime durations in U.S. nuclear power plants have been studied by Rothwell and Jensen (1990), who analyze the impact of the information structure in the plant on outage durations under the assumption of a Weibull proportional hazard model. In contrast to the U.S., where a large number of utilities with vast differences in size and technological capabilities operate nuclear power plants, plant management in Europe does not display much variation within a country[31]. Therefore, the choice of regressors here focuses on the plant history and is complementary to the analysis of Rothwell and Jensen.

The states and possible transitions in the data have been graphed in chapter 2, figure 1. There is one run state (state 0) and a number of outage states. The most important "down" states, both in terms of occurrence and total time spent in these states, are states 1 and 3 (unplanned outage due to equipment failure and planned refuel outage), and the sojourn times in these two states are investigated in detail here. The major portion of unavailability is due to planned outages, generally refuel outages. For example, from 1984 to 1986, approximately 78% of nonavailability time of plants constructed by the German company KWU was due to refueling and inspection outages, 14% was due to backfitting measures and repairs that prolonged refueling outages and only about 8% was due to unplanned shutdowns (Brettschuh, 1988). Since the average availability in this time period has been 85.2% for PWR and 86.5% for BWR manufactured by KWU, unplanned downtime durations (but not necessarily the occurrence of unplanned outages) are almost negligible from an economic point of view. Some descriptive statistics and density plots were given in chapter 2.

---

[31] But because differences between countries may be substantial I chose not to pool data from different countries.

One important issue is the dependence of durations on the plant history and exogenous regressors. Path dependence invalidates the use of many non- or semiparametric methods that rely on some form of (semi-) Markov assumption[32]. A typical form of path dependence is occurrence dependence where the hazard for the k'th duration depends on the number of previous spells. This is a form of dependence that often appears in fertility studies which consider repeated events of the same kind (here we have different types of events). For example, David and Mroz (1989) include the number of boys and girls (among many other variables) in analyzing birth intervals.

The effect of covariates are introduced in this section in the standard multiplicative form (proportional hazard) where the intensity of leaving the down state at time t is

$$\lambda(t) = \lambda_0(t)\exp(x'\beta) \tag{25}$$

The main effects considered are:

a) the number of the spell in the fuel cycle (e.g. how many outages have occurred since the last refueling)

b) the duration of the last up spell

c) the total up time since the last refuel outage

d) the last refuel duration

e) the age of plant (measured in time since its first commercial

---

[32] The term "path dependence" as used here refers to higher order statistical dependence in the transitions. A broader meaning has been attached to this term in economics (Arthur, 1989, David, 1985, 1988).

operation)

f) the capacity in MW (not history dependent)

g) a dummy variable for BWR (not history dependent).

A number of different baseline hazard $\lambda_0(t)$ specifications are estimated:

a) exponential hazard

$$\lambda_0(t)=\lambda \tag{26}$$

b) Weibull hazard

$$\lambda_0(t)=\lambda\rho t^{\rho-1} \tag{27}$$

c) Gompertz hazard

$$\lambda_0(t)=\lambda\exp(\gamma_1 t) \tag{28}$$

d) exponentiated quadratic polynomial

$$\lambda_0(t)=\lambda\exp(\gamma_1 t+\gamma_2 t^2) \tag{29}$$

In addition to the effect of covariates, one has to consider the possibility of unobserved heterogeneity (e.g. omitted variables) which would bias inference, especially about the shape of the hazard function. Unobserved heterogeneity is modeled as a scalar random variable with either a gamma or a discrete distribution, e.g. for the gamma mixing model, the parameter λ is a random variable with density

$$f(\lambda)=\frac{\alpha^{r}}{\Gamma(r)}\lambda^{r-1}e^{-\alpha\lambda} \tag{30}$$

If a constant term is included as a regressor, the gamma distribution is reparametrized to have mean 1 and only the variance is estimated.

The assumptions of the proportional hazard model are not innocuous (*e.g. hazards for different subpopulations cannot cross*), but I did not encounter obvious violations of this assumption when performing a preliminary nonparametric analysis with stratified subsamples. As discussed in section 3.2, however, the proportional hazard model in its failure time formulation is not adequate to analyze uptime durations and to investigate plant reliability.

In a fertility study, Heckman and Walker (1987) found that no model passed their goodness-of-fit criteria and the commonly found nonrobustness of estimates may be due to an inadequate fit of any model to the data. Surprisingly, goodness-of-fit tests are rarely reported in duration analysis (David, Mroz, and Wachter, 1985, being a notable exception). Given the results of Heckman and Walker, I generally use some chi-square test using k discrete time intervals to check the adequacy (or rather inadequacy) of models in addition to plots of predicted and actual frequencies. In the presence of censoring, the classical goodness-of-fit tests, Pearson's Chi-square test and the likelihood ratio test, need to be modified. This can be done easily under the assumption that censoring occurs only at the end of an interval. Censoring of downtime durations is very light and this assumption is therefore unproblematic. I base a goodness-of-fit test on the hazard rather than on the probability of a failure occurring in a specific interval. The null hypothesis and its alternative are:

$$H_0:\ h_{0i}=h_i,\ i=1,2,\ldots,k \qquad (31)$$
$$H_1:\ h_{01}\neq h_i,\ i=1,2,\ldots,k$$

where h_i is the empirical hazard of interval i and h_{0i} is the hazard in interval i implied by the estimated model. Thus the null hypothesis tests whether the hazard implied by the estimated model is consistent with the empirical hazard in every period. The log-likelihood function is

$$l=\sum_i [D_i logh_i+(R_i-D_i)\log(1-h_i)] \qquad (32)$$

where D_i and R_i are the number of failures and the size of the risk set in interval i. The likelihood ratio test statistic becomes:

$$\Lambda=-2[l(H_0)-\max l]=2\sum_i\left[D_i\log\left(\frac{h_i}{h_{0i}}\right)+(R_i-D_i)\log\left(\frac{1-h_i}{1-h_{0i}}\right)\right] \qquad (33)$$

which under the assumption about independence of censoring and durations--satisfied because downtime durations are only time censored--has the standard limiting $X^2_{(k-q)}$ distribution, where q is the number of parameters estimated under H_0, and k-q are the degrees of freedom[33].

In regression type duration models, regressors are normalized to have mean 0 and variance 1. The normalization substantially improves numerical convergence, avoids over- and underflow problems, and makes the size of parameter estimates for different regressors comparable. The means and standard deviation of the regressors before normalization are reported with the regression results.

3.1.1 unscheduled outages - equipment failures

The results for five model specifications are reported in tables 13a-f. Variables that were not found to be statistically significant in a preliminary analysis are excluded. The excluded variables are the number of the spell in a fuel cycle (measuring occurrence dependence under the assumption of regeneration during refueling), the previous refuel duration, and generating capacity (Germany only). The last row contains the p-value for a chi-squared goodness-of-fit test which compares the predicted probabilities of durations falling into one of 20 intervals of equal length with the observed probabilities. Equal spacing makes the test very tough because a few unusually long durations will put a lot of weight on

---

[33] On a more technical note, the limiting distribution is of degree k-q only if the maximum likelihood estimates use discretized (grouped) data, not when they are based on the raw continuous (ungrouped) data (eg. Kendall and Stuart, 1967, Ch.30). Thus the test is slightly conservative with ungrouped data.

the tails (compared to using intervals with the same number of observed durations).

The goodness-of-fit test reveals that simple hazard specification such as the exponential or the Weibull model are inconsistent with the data for Belgium, France, and Sweden at the 1% significance level. It appears that the exponentiated quadratic polynomial hazard is the most appropriate functional form for equipment failure outages. It is the only model that passes the goodness-of-fit test at the 1% level for 4 out of 5 countries, and the test does not reveal any inconsistency of the model with the data even at the 30% level for Germany and Switzerland. The exponentiated quadratic polynomial is also the best specification reported for Belgium, the only country where all of the reported specifications are rejected at the 1% level[34].

Although the regressor variables play an important role, unobserved heterogeneity is likely to be substantial: outage durations caused by equipment failures depend on the particular type of failure[35]. Cattaneo et al. (1988) have analyzed unplanned outages with respect to problem categories and their findings substantiate this impression. They calculate a mean outage time for reactor vessel equipment problems of 3467 hours and a mean outage time for turbine problems of only 65 hours. In the lower tail of

[34] Note that passing the test at 1% means that the model is rejected at the 5% or 10% level.

[35] Unfortunately, after comparing the information in the IAEA publications with detailed descriptions of events in trade journal, it appears that the data is not reliable enough on this aspect to be used in any quantitative analysis.

the duration distribution, a physical reason, xenon poisoning (e.g. Pershagen, 1989), may contribute to the failure of modeling outage durations by standard distribution function. Xenon-135 is a strong neutron absorber and it is no longer lost through neutron absorption once the reactor is shut down. But because iodine-135 continues to decay to xenon, the xenon concentration <u>increases</u> once the neutron flux disappears. Unless the reactor can be restarted very quickly (within the first 2-5 hours), it may become necessary to wait until the xenon concentration, which reaches a peak about 10-12 hours after shutdown, has decreased sufficiently.

Both discrete mixing (the technique underlying estimates in David and Mroz, 1989) and gamma mixing models (only gamma mixing for the Weibull model is reported in tables 13a-f) were successful in dealing with the problem of unobserved heterogeneity, success being defined as passing goodness-of-fit tests.

Parameter estimates are sensitive to the specification, although sign changes are rare. One bad case of nonrobustness occurs for France: the duration dependence becomes positive and three estimated parameters change signs when unobserved heterogeneity is permitted in the Weibull model. "Significant" results may easily become insignificant in more flexible models, and they may be even less robust than that! This illustrates the importance of a sensitivity analysis before reaching conclusions. No uniform picture about the effect of regressors across countries emerges from the analysis, even though many parameter estimates are significantly different from zero. Thus variables may not capture fixed technological relationships, but characteristics of plant operations subject to the operator's control.

3.1.2 scheduled outages - refueling

Refuel outage durations are without any doubt the main determinant of availability. Tables 14a-f contain the result for several model specifications. The exponential hazard was again found to be very inadequate compared to more general hazards, it was never consistent with the data at the 1% significance level. Although the overall fit (all countries and models) was not better than for unplanned outages due to equipment failure, some models (with Weibull hazards) described the data very well for several countries. For example, the Weibull model (even without unobserved heterogeneity) was not rejected at the 12.5% significance level for Sweden under the goodness-of-fit test described in the previous subsection. Note the strong positive duration dependence. With a shape parameter that high, the Weibull distribution becomes almost indistinguishable from a normal distribution (compare figure 16). That is, refuel durations (not log(durations)) could be considered to be normally distributed, conditional on the regressors. Figure 17 plots the actual cell probabilities for Sweden and France. Note the far right cells, which according to one's taste may be labeled "due to unobserved heterogeneity" or "outliers". Except for these cells, the Weibull model without unobserved heterogeneity captures the general distribution fairly well. Figure 18 plots the predicted cell probabilities under the exponential model without heterogeneity and the Weibull model with gamma mixing for Sweden. The goodness-of-fit test compares the probabilities of figures 17 and 18.

Table 13: Outages due to equipment failure

dependent variable is duration in days
continuous regressors are normalized to mean 0 and standard deviation 1
regressors: constant
# of spell in cycle
last up duration
up time since last refuel outage
last refuel duration
age of plant (time since first commercial operation)
capacity (MW)
BWR dummy (no BWR's in France and Belgium)

Table 13a: Means and Standard Deviations of Regressors

country	last up (hours)	up since last refuel (hours)	age (hours)	capacity (MW)	BWR (dummy)
Belgium	1540 (1628)	3211 (2436)	65924 (29646)	634 (254)	no BWR
France	1179 (1335)	3011 (2433)	32093 (16435)	907 (42)	no BWR
Germany	1537 (1857)	3597 (3249)	64338 (27810)	not used	0.47 (0.50)
Switzer-land	2076 (2013)	2966 (2414)	75372 (54479)	584 (297)	0.27 (0.46)
Sweden	1293 (1404)	2941 (2425)	69030 (29787)	671 (177)	0.80 (0.40)

Table 13b: France 400 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-2.5535	-2.4985	-2.3966	-2.1874	-2.3056
last up		0.1835 (0.0655)	0.1109 (0.0638)	-0.0691 (0.2268)	0.0799 (0.0650)
up since ref		0.1877 (0.0623)	0.1061 (0.0617)	-0.1966 (0.2421)	0.0464 (0.0623)
age		0.0528 (0.049)	0.0292 (0.04928)	0.3774 (0.2030)	0.0203 (0.0500)
BWR		0.0804 (0.0486)	0.0502 (0.0486)	-0.0817 (0.1709)	0.0435 (0.0459)
constant*	-1.5535 (0.0500)	-1.4985 (0.0500)	-0.9498 (0.0718)	2.3817 (1.0772)	-0.9248 (0.0643)
shape1			0.7482 (0.0263)	5.5267 (1.211)	-0.1041 (0.0115)
var/ shape2*				8.0993 (2.0704)	0.0712 (0.0102)
p-value	<0.0001	<0.0001	<0.0001	0.01884	0.0296

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 13c: Sweden 169 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-2.1914	-1.9705	-1.9665	-1.8990	-1.9270
last up		0.2776 (0.1006)	0.2539 (0.1020)	0.2899 (0.2267)	0.2131 (0.1009)
up since ref		0.2392 (0.1003)	0.2175 (0.1010)	0.1657 (0.2265)	0.1754 (0.0990)
age		-0.4297 (0.1204)	-0.3960 (0.1236)	-0.2799 (0.3033)	-0.3129 (0.1247)
capacity		-0.3342 (0.1454)	-0.3106 (0.1459)	-0.2885 (0.3399)	-0.2462 (0.1452)
BWR		0.3970 (0.0988)	0.3667 (0.1018)	0.4356 (0.2430)	0.2900 (0.1018)
constant*	-1.1914 (0.0769)	-0.9704 (0.0769)	-0.8772 (0.1092)	0.1413 (0.5010)	-0.6451 (0.1099)
shape1			0.9368 (0.0536)	2.4480 (0.5848)	-0.1229 (0.0374)
var/ shape2*				2.4835 (0.9547)	0.2892 (0.1222)
p-value	<0.0001	0.0001	0.0002	0.0026	0.0103

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 13d: Germany 51 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-2.1279	-1.9515	-1.936	-1.9121	-1.8576
last up		-0.2865 (0.1802)	-0.2050 (0.1834)	0.0308 (0.2574)	-0.0943 (0.1834)
up since ref		0.5070 (0.1814)	0.3948 (0.1883)	0.2711 (0.2748)	0.3026 (0.1867)
age		0.3263 (0.1615)	0.2302 (0.1625)	0.27765 (0.2423)	0.2302 (0.1621)
BWR		0.3182 (0.1758)	0.2389 (0.1710)	0.4694 (0.3027)	0.2771 (0.1700)
constant*	-1.1705 (0.1414)	-0.9959 (0.1416)	-0.6966 (0.1867)	-0.4507 (0.2996)	-0.4463 (0.2090)
shape1			0.8144 (0.0875)	1.1269 (0.2628)	-0.2598 (0.0951)
var/ shape2*				0.6041 (0.4867)	0.9917 (0.4371)
p-value	<0.0001	0.0001	0.0165	0.0083	0.5079

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 13e Switzerland 15 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-2.2432	-1.8031	-1.6484	not	-1.5894
last up		0.1723 (0.5799)	0.2772 (0.7079)	converged	0.8818 (0.7509)
up since ref		0.3064 (0.6663)	0.4276 (0.8499)		0.0895 (0.7549)
age		-1.9362 (1.3580)	-3.0783 (1.7733)		-6.3572 (3.2119)
capacity		-1.3131 (1.2432)	-1.9073 (1.5341)		-4.6464 (2.5789)
BWR		0.1058 (0.5297)	0.1767 (0.6319)		-0.0692 (0.5680)
constant*	-1.2432 (0.2582)	-0.8031 (0.2582)	-1.5570 (0.5077)		-1.7951 (0.7356)
shape1			1.6978 (0.3861)		0.2325 (0.2781)
var/ shape2*					1.2703 (1.7585)
p-value	0.0024	0.0087	0.0213		0.3571

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 13f: Belgium 54 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-2.1266	-2.0328	-2.0322	-1.9373	-2.0143
last up		-1.0328 (0.1361)	-0.1910 (0.1581)	-0.0916 (0.4707)	-0.1641 (0.1689)
up since ref		-0.0086 (0.1716)	-0.0121 (0.1723)	0.2847 (0.4992)	0.0205 (0.1713)
age		-0.0461 (0.2288)	-0.0505 (0.2302)	-0.1707 (0.6099)	-0.0153 (0.2281)
capacity		-0.3652 (00.2273)	-0.3828 (0.2377)	0.1477 (0.6092)	-0.2522 (0.2362)
constant*	-1.1266 (0.1361)	-1.0328 (0.1361)	-1.0747 (0.2108)	-0.2477 (0.6024)	-0.7910 (0.2104)
shape1			1.0281 (0.1076)	2.8852 (0.8916)	-0.1094 (0.0832)
var/ shape2*				2.7350 (1.3397)	0.3688 (0.3701)
p-value	<0.0001	<0.0001	<0.0001		0.0047

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 14: Refuel outages

dependent variable is duration in months
continuous regressors are normalized to mean 0 and standard deviation 1
regressors: constant
# of spell in cycle
last up duration
up time since last refuel outage
last refuel duration
age of plant (time since first commercial operation)
capacity (MW)
BWR dummy (no BWR's in France and Belgium)

Table 14a: Means and Standard Deviations of Regressors

country	up since last refuel (hours)	last refuel duration (hours)	age (hours)	capacity (MW)	BWR (dummy)
Belgium	7869 (1742)	874 (223)	70008 (26070)	622 (252)	no BWR
France	7563 (988)	1658 (946)	35111 (14803)	903 (14)	no BWR
Germany	7869 (2100)	1412 (1505)	70168 (30081)	931 (300)	0.21 (0.41)
Switzer-land	7390 (1675)	882 (198)	88369 (42282)	522 (280)	0.27 (0.46)
Sweden	7775 (1805)	1167 (562)	65774 (29534)	713 (169)	0.78 (0.41)

Table 14b: France 85 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-1.4211	-1.4515	-0.7938	-0.6580	-0.7395
up since last ref		0.0062 (0.1057)	-0.1498 (0.0979)	0.4079 (0.2746)	-0.0129 (0.1009)
last ref duration		-0.0902 (0.1305)	-0.3755 (0.1628)	-0.3925 (0.2566)	-0.3671 (0.1617)
age		-0.5361 (0.1302)	-0.1415 (0.1523)	-0.4572 (0.2879)	-0.1412 (0.1481)
capacity		0.0128 (0.1131)	0.1000 (0.1242)	-0.0074 (0.2190)	0.0468 (0.1210)
constant*	-0.4547 (0.1085)	-0.4115 (0.1085)	-1.7005 (0.2037)	-2.2938 (0.3137)	-4.6314 (0.6472)
shape1			3.0134 (0.2359)	6.5622 (1.4859)	4.9598 (0.7649)
var/ shape2*				1.4658 (0.6976)	-104.26 (21.82)
p-value	<0.0001	<0.0001	<0.0001	0.0046	<0.0001

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 14c: Sweden 41 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-1.3660	-1.3360	-0.7719	-0.3243	-0.7512
up since last ref		0.0447 (0.1712)	0.2024 (0.1562)	-2.1369 (1.5170)	0.1550 (0.1533)
last ref duration		-0.0635 (0.1899)	-0.1981 (0.2443)	-1.9304 (1.2726)	-0.2661 (0.2393)
age		0.0248 (0.2812)	0.0737 (0.3105)	-2.4205 (1.4230)	0.2280 (0.3055)
capacity		0.1154 (0.3212)	0.1326 (0.3689)	2.6526 (1.5598)	0.3014 (0.3615)
BWR		0.2285 (0.2008)	0.3908 (0.2121)	7.5134 (3.1646)	0.5210 (0.2196)
constant*	-0.3660 (0.1562)	-0.3360 (0.1562)	-1.2098 (0.2464)	-1.5608 (0.7345)	-3.7518 (0.7769)
shape1			2.6592 (0.2917)	21.52 (8.10)	4.5879 (1.0114)
var/ shape2*				6.5190 (3.091)	104.53 (30.02)
p-value	<0.0001	<0.0001	0.1252	0.1350	0.0002

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 14d: Germany 38 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-1.5933	-1.5649	-1.2205	-0.9790	-1.2939
up since last ref		-0.2534 (0.2282)	-0.6181 (0.2442)	-1.1054 (0.5510)	-0.5808 (0.2459)
last ref duration		0.1431 (0.1807)	0.4623 (0.1920)	-0.1152 (0.3852)	0.3454 (0.1887)
age		-0.2278 (0.2914)	-0.6467 (0.3393)	-1.2256 (0.8210)	-0.6719 (0.3445)
capacity		-0.0742 (0.2567)	-0.2078 (0.2709)	-1.0101 (0.9024)	-0.2661 (0.2715)
BWR		-0.0278 (0.2538)	-0.2588 (0.2938)	0.4761 (0.5284)	-0.0667 (0.2815)
constant*	-0.5933 (0.1622)	-0.5649 (0.1622)	-1.4776 (0.2916)	-1.9993 (0.4929)	-2.1993 (0.4922)
shape1			2.1235 (0.2636)	7.5601 (2.1643)	1.6960 (0.4675)
var/ shape2*				2.9451 (1.2297)	-24.39 (9.055)
p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 14e: Switzerland 22 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-1.2816	-1.2744	-0.3362	not	-0.2533
up since last ref		0.0226 (0.2709)	0.6203 (0.7871)	converged	0.3320 (0.6738)
last ref duration		-0.02274 (0.2467)	-0.0195 (0.2881)		-0.0620 (0.2749)
age		-0.3270 (0.6683)	-1.7004 (0.7501)		-1.5248 (0.7602)
capacity		-0.3205 (0.6726)	-1.6181 (0.7107)		-1.5299 (0.7329)
BWR		-0.1102 (0.2636)	-0.6611 (0.3035)		-0.3954 (0.2916)
constant*	-0.2816 ((0.2132)	-0.2744 (0.2132)	-1.6015 (0.4220)		-7.2197 (1.9681)
shape1			4.2906 (0.7604)		9.6851 (2.6877)
var/ shape2*					-252.99 (89.42)
p-value	<0.0001	<0.0001	<0.0001		<0.0001

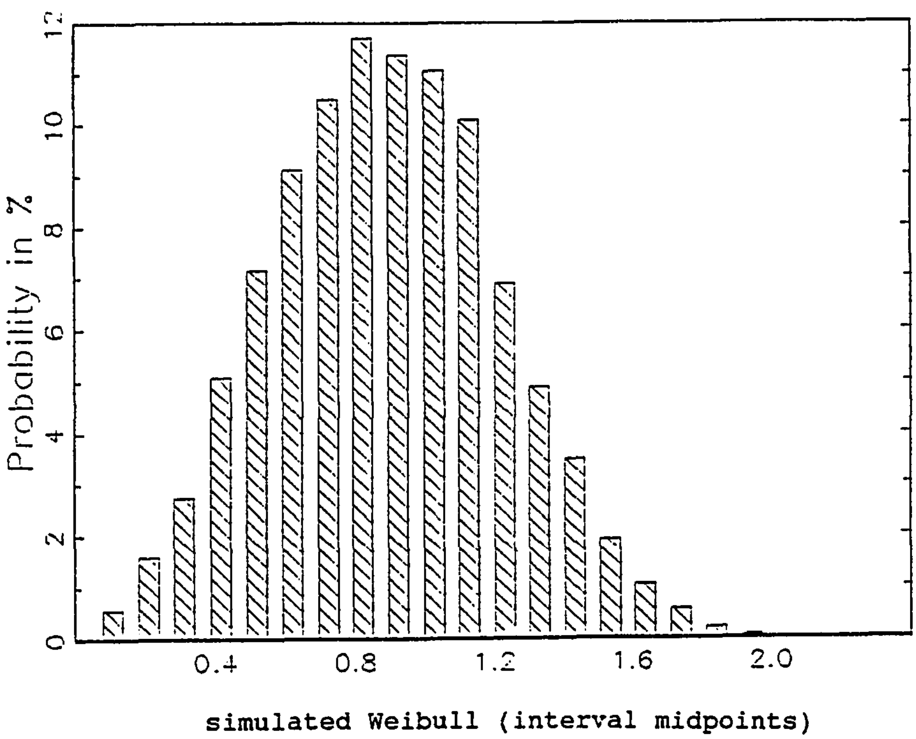
*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Table 14f: Belgium 17 observations

	expo	expo	Weibull	Weibull with Gamma mixing	expoquad
mean logl	-1.4111	-1.3620	-0.5257	-0.5241	-0.6472
up since last ref		-0.1877 (0.2922)	-0.5766 (0.3089)	-0.6750 (0.5170)	-0.5357 (0.3078)
last ref duration		0.1387 (0.2660)	0.7402 (0.3689)	0.7773 (0.4046)	0.6889 (0.3755)
age		-0.2299 (0.3181)	-1.2883 (0.4648)	-1.3312 (0.5239)	-1.1971 (0.4924)
capacity		-0.1144 (0.3553)	-0.6783 (0.4316)	-0.7528 (0.5415)	-0.5847 (0.4612)
constant*	-0.4211 (02425)	-0.3620 (0.2425)	-1.8419 (0.4999)	-1.8981 (0.5730)	-3.7653 (1.0296)
shape1			3.9801 (0.7737)	4.3631 (1.7398)	3.6772 (1.0301)
var/ shape2*				0.1542 (0.5992)	-48.19 (23.35)
p-value	0.0095	0.0007	0.0739	0.0436	<.0001

*: The constant is $\ln(\lambda)$, var is variance of gamma distribution, shape 2 is $100\gamma_2$; see equations (26)-(30)

Figure 16: Probability Plot Weibull vs. Normal



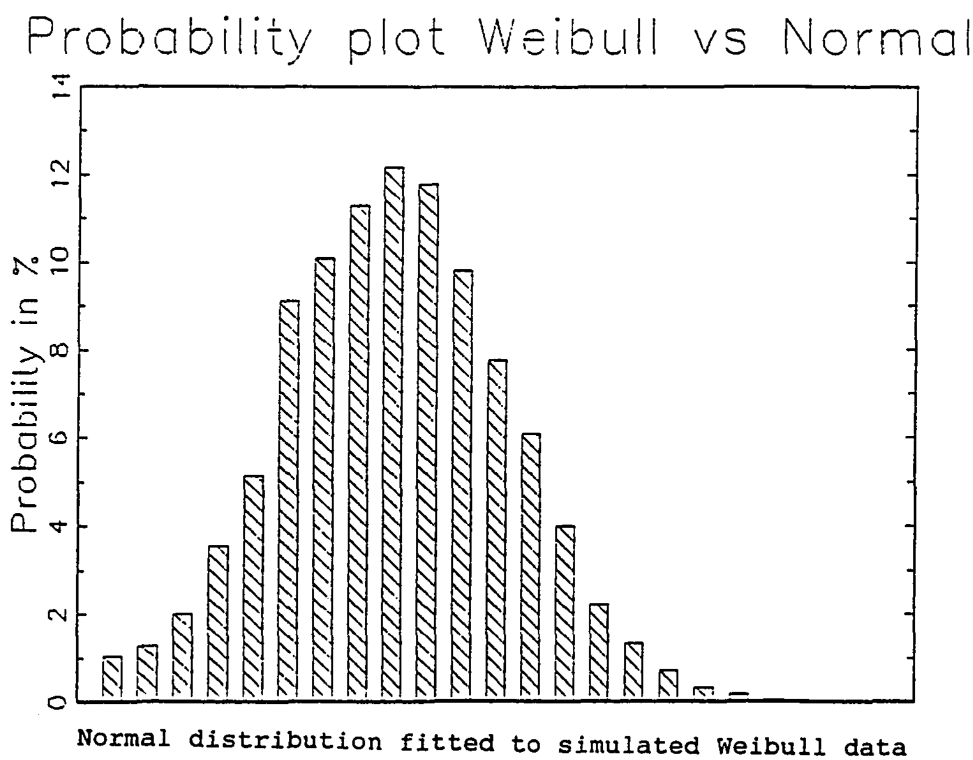
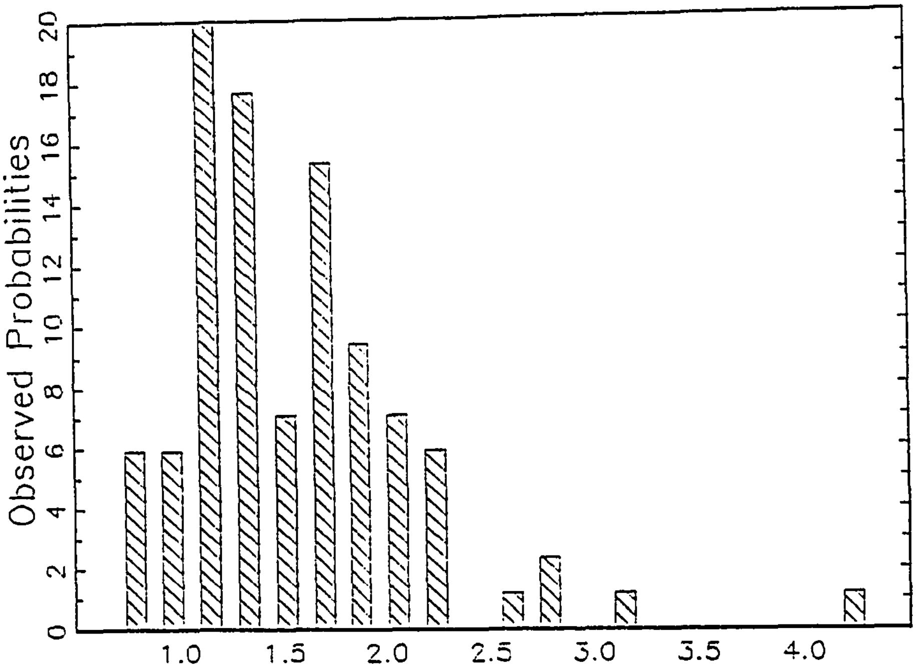
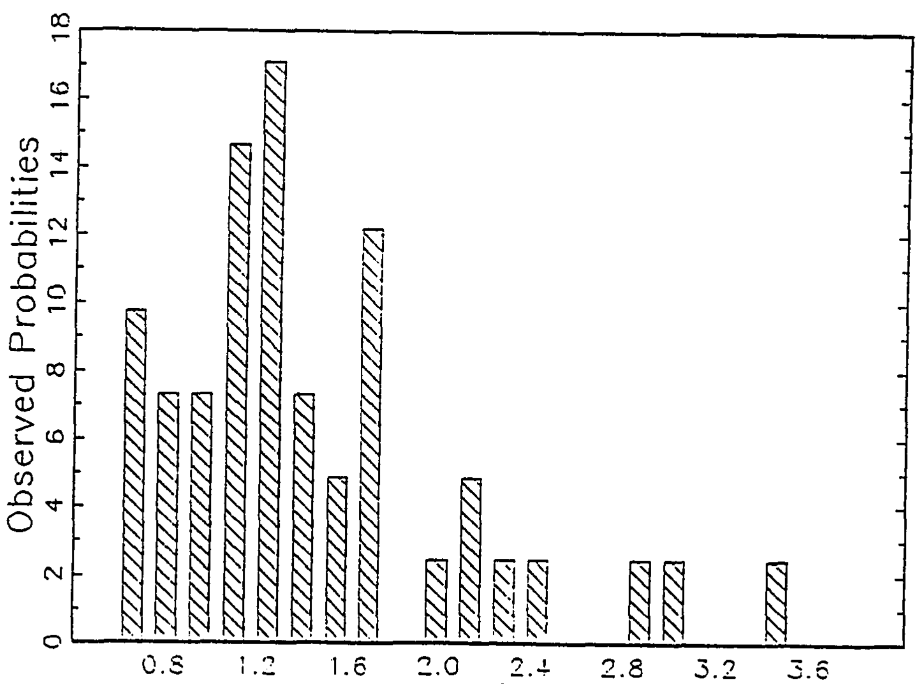


Figure 17: Refuel Outage Durations: France and Sweden
time in months

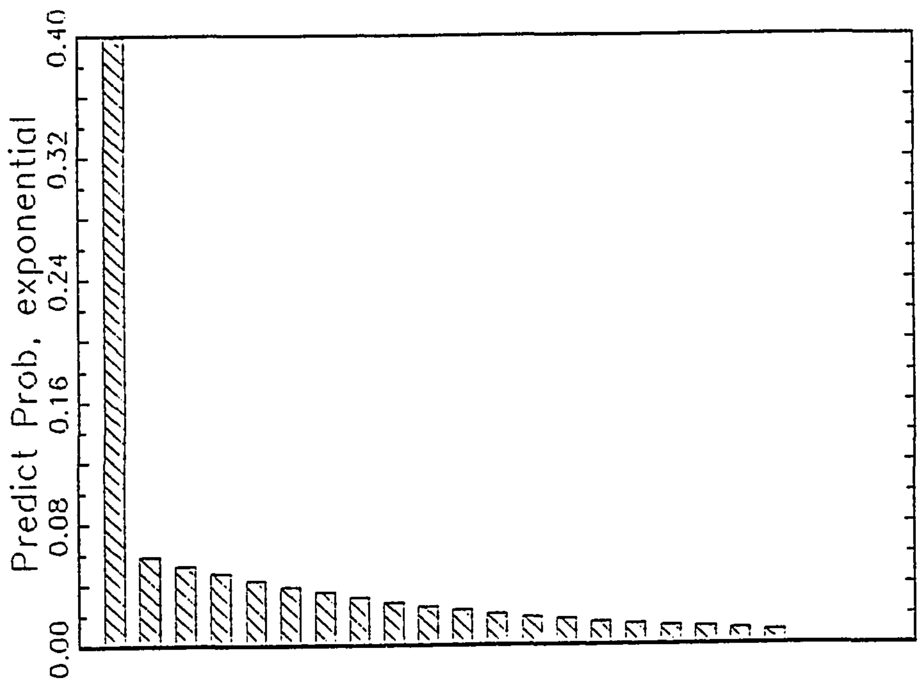


France

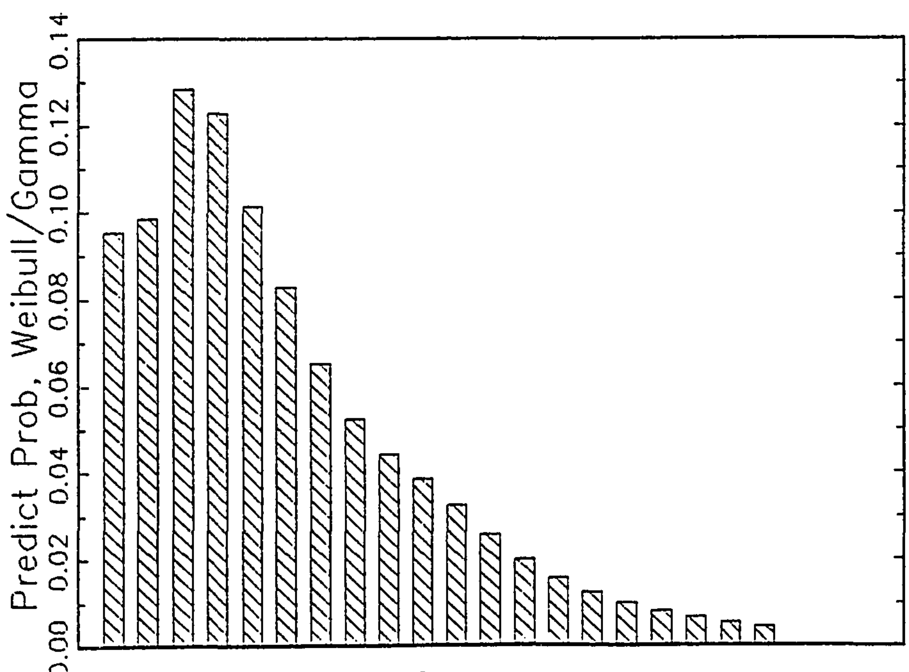


Sweden

Figure 18: Model Predictions for Refuel Outage Durations: Sweden



same scale as figure 17b)



same scale as figure 17b)

3.2 Uptimes

Real systems can be repaired and restored to satisfactory performance without necessarily having the properties of a completely new system. Despite this obvious fact, almost all statistical models only consider the occurrence of the first failure, or put differently, they only consider nonrepairable systems. Nonrepairable systems are replaced upon failure which leads to the renewal assumption at every event, an assumption that was rejected in chapter 2[36]. Both the models by Rust (1987) and Ryu (1990) belong to this class, as do standard failure time models (see section 3.2.2). The situation is not better in the engineering and statistics literature where renewal models have been used even if the application clearly was inconsistent with the assumption of a nonrepairable system[37]. But how can the relationship between successive failures of a repairable system be modeled? This section derives such a model and estimates a parametric version of it.

3.2.1 Repairable and nonrepairable systems

A complex system consists of different units/components, some of which may be redundant. I consider a system in which components

---

[36] An excellent introduction to reliability analysis of complex nonrepairable system is Barlow and Proschan (1975).

[37] Asher and Feingolds (1984) provide a book long criticism of the literature. See Frees (1988) for an example of an incorrect application of the renewal assumption.

are connected in series and in which the failure of any component causes a system failure. Assume that each component is immediately replaced upon failure and that failures of components are independent. A mode is the location of a component. The flow of system failures is the superposition of the renewal processes of the failure modes m=1,...,M.

Consider a completely new system and let T^* denote the time of the first system failure and T_m the "latent" failure time in mode m. The probability of no system failure in the interval [0,t] is

$$P(T>t)=P(T_1>t)\,P(T_2>t)\,...\,P(T_m>t) \tag{34}$$

Each failure mode m has an associated hazard function $\lambda_m(t)$ and survivor function

$$S_m(t)=P(T_m>t)=\exp[-\int_0^t \lambda_m(u)\,du] \tag{35}$$

Thus the system survivor function (for the first failure) in terms of mode hazard functions becomes

$$\begin{aligned} S(t) &= \exp[-\int_0^t \lambda(u)\,du] \\ &= \exp[-\int_0^t \lambda_1(u)\,du]\exp[-\int_0^t \lambda_2(u)\,du]\ldots\exp[-\int_0^t \lambda_M(u)\,du] \\ &= \exp[-\int_0^t \sum_{m=1}^{M} \lambda_m(u)\,du] \end{aligned} \tag{36}$$

and the system hazard

$$\lambda(t)=\frac{f(t)}{S(t)}=-\frac{d\log S(t)}{dt}=\sum_{m=1}^{M}\lambda_m(t) \tag{37}$$

is simply the sum of individual failure mode terms. This standard failure time formulation for independent competing risks underlies the semi-Markov model of chapter 2 and the analysis in David, Rothwell, and Maude-Griffin (1991).

Now consider the failure intensity after the first failure, say in mode j, has occurred at T^* and the component in mode j has been replaced by a new component:

$$\lambda(t|t>T^*)=\lambda_j(t-T^*)+\sum_{m=1,m\neq j}^{M}\lambda_m(t) \tag{38}$$

At the next failure T^{**}, another component is replaced and the hazard function in this mode is reset to time 0, but not the hazard function in the other modes. Clearly, the failure rate of the system (failure intensity or system hazard) is a complicated stochastic function with discontinuous jumps at each failure time. The sequence of failures is a nonstationary point process but it is not a nonstationary Poisson process because the system failure rate depends on which component was renewed. Probabilistically, we have the following relationship. Let $N_m(t)$ be the random number of failures in mode m until time t. The number of system failures N(t) and its expected value (the renewal function) H(t) is obviously

$$N(t)=\sum_{m}^{M}N_m(t)$$
$$EN(t)=H(t)=\sum_{m}^{M}EN_m(t)=\sum_{m}^{M}H_m(t) \tag{39}$$

The expected intensity of failures is the derivative of the renewal function

$$h(t)=\frac{d}{dt}H(t) \tag{40}$$

Note that this is a deterministic function of time and thus different from the actual failure intensity $\lambda(t)$ which depends on the pattern of failures. However, if the contribution of each component hazard to the overall failure rate is small, the failure and renewal of an individual component has only a small effect on the system failure intensity. Thus for a complex systems with a "large" number of modes, the system failure intensity approaches that of a deterministic function and failures follow a nonstationary Poisson process. Figure 19 demonstrates this for two cases with identical modes and components. The number of modes are m=2,20, and 200; component lifetimes have a Weibull distribution with parameters 1.0/m (scale) and 1.5 (shape) in a) and 2.0/m (scale) and 0.75 (shape) in b). Note how fast the convergence occurs and how close a deterministic function of time could approximate the stochastic intensity of a relatively small system with 200 modes. Real system consists of a much larger number of components, probably in the thousands, that could cause a failure and the approximation does not appear to be too unreasonable. The asymptotic result can be made precise, see Cinlar (1972) and Cox and Isham (1980).

Consider an important special case of failures, namely failures due to installation errors or substandard components. The renewal argument of the preceding section does not apply because an installation error is generally removed, not renewed. Similarly, a substandard component is general replaced by a standard component[38]. A number of models dealing with increasing reliability over time due to the detection of errors and learning have been developed in the engineering literature as software reliability models. These models are among the few exceptions in the literature that considered multiple failures of the same system. Errors in computer programs are not due to an ageing and wear-out process, and thus only one failure type is analyzed. An early software reliability model was developed by Jelinski and Moranda (1972) who suggested the following model.

---

[38] Although an "imperfect" repair might make it possible that a substandard component is renewed with probability p and replaced by a new regular component with probability (1-p). The imperfect repair model (Brown and Proschan, 1983) was introduced somewhat differently in the literature. Brown and Proschan considered a component which is renewed with probability p (perfect repair) and returned to operations without changing its degradation (the hazard is not reset) with probability 1-p (imperfect repair).

Assumptions:

1. The successive interfailure times of the program t_i are independently exponentially distributed with a hazard proportional to the number of errors remaining

$$\lambda(t)=\phi(N-i+1) \qquad (41)$$

where ϕ is an unknown constant and N is the number of errors initially present in the program.

2. Each failure cause is perfectly and immediately removed after causing a failure.

These assumptions imply a likelihood function for n observed interfailure times $t_1,t_2,..,t_n$

$$L(N,\phi)=\prod_{i=1}^{n}\phi(N-i+1)\exp[-\phi(N-i+1)t] \qquad (42)$$

A number of variants and extensions of this model exist (Shooman, 1972, Musa, 1975, Littlewood, 1981). Littlewood and Verrall (1973) describe a Bayesian reliability growth model for computer software in which repair actions diminish the failure rate probabilistically.

Although these models can be very useful when reliable information on outages at the component level is available, the problem with the IAEA data is that the precise cause is often unknown and it is never known if the failure cause has been found and removed. Does there exist a similar approximation for such model as for a complex system with many renewable components?

Consider the following generalization of a software reliability model: "bugs" (installation errors, software errors, substandard components) need not have the same hazard of causing a failure, but this (constant) hazard λ_m may depend on the location (mode) of the "bug". Each "bug" has a constant hazard d of being detected. Thus the system hazard is a doubly stochastic process with intensity:

$$\lambda(t) = \sum_{m=1}^{M} I_{\{T_m > t\}} \lambda_m \tag{43}$$

where T_m is the time at which a "bug" in mode m is detected and removed. The expected value of the system hazard $\lambda(t)$ is:

$$E\lambda(t) = \sum_{m=1}^{M} \lambda_m E I_{\{T_m > t\}} = \sum_{m=1}^{M} \lambda_m \exp[-dt] \tag{44}$$

What happens as the number of "bugs" grows large while for each "bug" the hazard of causing a failure becomes small?

Proposition:

Consider the sequence

$$\sum_{m=1}^{M} \sum_{l=1}^{K} \frac{\lambda_{ml}}{K} I_{\{T_{ml} > t\}}, \quad K=1,2,\ldots, \tag{45}$$

where λ_{ml} are iid random variables with finite mean λ and variance. As the number of bugs, all of which have the same detection hazard d<∞ grows large (K→∞), the system hazard converges pointwise (for any fixed t) to Mλexp(-dt).

Proof:

The system hazard is

$$\lambda(t)=\sum_{m=1}^{M}\sum_{l=1}^{K}\frac{\lambda_{ml}}{K}I_{\{T_{ml}>t\}} \tag{46}$$

with expected value

$$\begin{aligned}E\lambda(t)&=\sum_{m=1}^{M}\sum_{l=1}^{K}E\frac{\lambda_{ml}}{K}EI_{\{T_{ml}>t\}}\\&=\sum_{m=1}^{M}\sum_{l=1}^{K}\frac{\lambda}{K}\exp[-dt]=M\lambda\exp[-dt]\end{aligned} \tag{47}$$

Because of the law of large numbers

$$\lim_{K\to\infty}\sum_{m=1}^{M}\sum_{l=1}^{K}\frac{\lambda_{ml}}{K}I_{\{T_{ml}>t\}}=M\lambda\exp[-dt]\quad a.s. \tag{48}$$

This is an interesting result because it implies that for certain types of "learning" a Gompertz system hazard is the correct parametric form rather than the much more common Weibull hazard.

3.2.2 A point process formulation

The preceding discussion indicated that point processes (Cox and Isham, 1980) provide a powerful framework to model the occurrence of successive events in time. With a large number of failure modes where each individual failure hazard is small, the system failure intensity may be approximated by a nonstationary

Poisson process, even if the system is not a superposition of individual renewal processes. Consider the following statistical mechanism:

A system i is installed at time 0 and failures (which are repaired immediately) occur randomly according to an intensity $\lambda(t,x_i(t),\Theta)$ where $x_i(t)$ is a vector of covariates and Θ is a vector of unknown parameters. At time $\bar{T}_i$ the system is shut down ($\bar{T}_i$ is a censoring time). To simplify notation, I drop the dependence on x and Θ. The Poisson assumption implies that the increments of the associated counting process N_{it} are independent. However, the intervals between points are not independent since the integrated hazard and the survivor function of t conditional on starting at t_j both depend on the starting point.

$$\Lambda(t_j,t)=\int_{t_j}^{t}\lambda(u)\,du \qquad (49)$$
$$S(t_j,t)=\exp\{-\Lambda(t_j,t)\}$$

For a sample of n such systems with fixed censoring times $\bar{T}_i$ and m_i observed events at times $0=t_{i0}<t_{i1}<t_{i2}<\ldots<t_{imi}<\bar{T}_i$ for process i, the log-likelihood function is:

$$l(.)=\sum_{i=1}^{n}\sum_{j=1}^{m_i}\{\log\lambda(t_{ij})-\Lambda(t_{ij-1},t_{ij})\}-\sum_{i=1}^{n}\Lambda(t_{im_i},\overline{T_i}) \qquad (50)$$
$$=\sum_{i=1}^{j}\sum_{j=1}^{m_i}\log\lambda(t_{ij})-\sum_{i=1}^{n}\Lambda(0,\overline{T_i})$$

If $\bar{T}_i$ is a noninformative independent variable rather than a constant, the complete log-likelihood only involves an additional term that does not contain any unknown parameters of interest. This model is different from many existing models which assume regeneration at each event (the failure time formulation resets time to 0 after each event). It is useful to compare the statistical properties of a sequence of interfailure times under the assumption of a repairable system and under the renewal/regeneration assumption. Letting t_m denote the time of the m'th failure, the renewal assumption implicit in failure time models implies that the survivor function for the m'th failure is

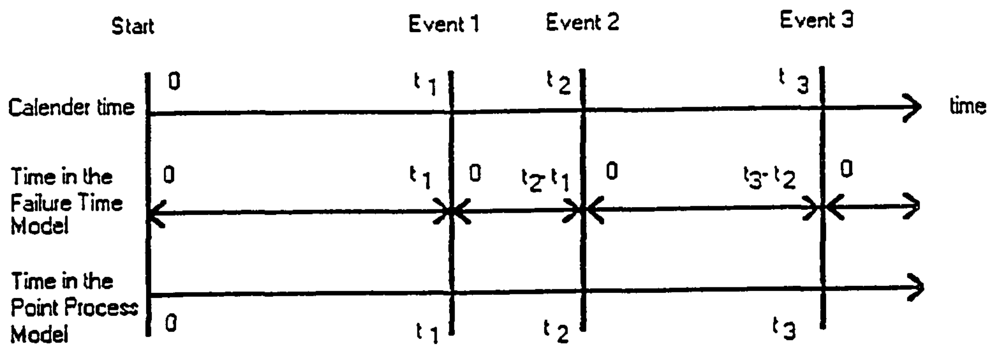
$$S_m^{renewal}(t|t_{m-1}) = S(0, t-t_{m-1}) = \exp\{-\Lambda(0, t-t_{m-1})\} \tag{51}$$

The incorrect renewal assumption would lead to the likelihood function

$$l^{renewal}(.) = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\{\log\lambda(t_{ij}-t_{ij-1}) - \Lambda(0, t_{ij}-t_{ij-1}\} - \sum_{i=1}^{n}\Lambda(0, \overline{T_i}-t_{im_i}) \tag{52}$$

which yields inconsistent estimates unless the intensity function is constant (the exponential model). The multiple spell models in economics, discussed in Heckman and Singer (1985), are built on the renewal model, although dependence between successive events is introduced through covariates in actual applications. Whereas the renewal/failure time model resets time to 0 at the beginning of each spell (spell time), the point process formulation measures time from one constant point for the life of the system (cycle or plant time). A graphical comparison of the two ideas with respect to their treatment of time is given in figure 20.

Figure 20: Failure Time vs. Process Formulation



The following example illustrates the argument. Proschan (1963) presented his famous data set on interfailure times of the air-conditioning systems of 13 Boeing 720 jet airplanes with the goal to obtain information about the distribution of failure intervals. This data set has been analyzed by many researchers, including Barlow et al. (1972), who perform a statistical test using the interval between failures to form a "cumulative total time on test" statistic. Their null hypothesis is that failures are exponentially distributed, the alternative hypothesis is that the hazard of failures increases. Time is measured from the last repair (spell time). But the equipment aging trend they discuss cannot measure aging of the air-conditioning system because their test ignores the <u>order</u> of occurrence of interfailure intervals. This implies that the authors implicitly assume that each repair completely renews the air-conditioning system and that a repaired system is indistinguishable from a new air-conditioning system, regardless of its age!

Cox and Lewis (1978) have also analyzed the air-conditioner data set. Their null hypothesis is that failures occur according to a homogeneous Poisson process, the alternative is that failures occur according to a nonstationary Poisson process with intensity $\lambda=\exp(a+bt)$, where t measures the age of the system. In other words, Barlow et al. assume that their is no aging of the system, only the effect of repair wears out, whereas Cox and Lewis correctly consider the system wear-out. On the other hand, the failure intensity in the model of Cox and Lewis is independent from the time since the last repair, i.e. there is no wear-out effect of repairs. The confusion between renewal models and models that do not reset the process at each event appears to continue in the literature (see Asher and Feingold, 1984). The statistical point of view taken here is similar to Cox and Lewis and differs from Barlow et al. or David et al. (1988), who suggest a renewal theory framework to analyze nuclear power plant operations. Summarizing, we can say that renewal theory cannot model system behavior because the order of occurrence of events is ignored.

The most common way to introduce regressors in failure time models is the so-called proportional intensity model where a function of the covariates acts multiplicatively on the baseline hazard. A particularly convenient form is

$$\lambda(t, x_i, \theta) = \lambda_0(t, \theta_1) \exp(x_i' \theta_2) \tag{53}$$

and this model extends immediately to point processes (Lawless, 1987). Since observed data is necessarily incomplete, much research in economic duration analysis has been devoted to address the problem of "unobserved" heterogeneity (Lancaster, 1979, Flinn and Heckman, 1982, Heckman and Singer, 1984), mainly restricted to the case of scalar heterogeneity in a proportional hazard model. The analogue for a point process is an intensity

$$\lambda(t, x_i, \theta, \nu) = \lambda_o(t, \theta_1) \exp(x_i' \theta_2) \exp(\nu) \tag{54}$$

where n is unobserved with a probability density function k(ν). The likelihood function requires integration with respect to ν

$$L(.) = \prod_{i=1}^{n} \left\{ \int_0^{\infty} \left[\prod_{j=1}^{m_i} \lambda(t_{ij}, x_i, \theta, \nu) \right] \exp[-\Lambda(0, \overline{T_i}, x_i, \theta, \nu | \nu] \, dk(\nu) \right\} \tag{55}$$

The gamma distribution is the most commonly used distribution for mixing distributions because it is analytically manageable. As long as there is a constant term in the covariate function, we can assume n to have mean 1 and variance α. This model was originally estimated by Lancaster (1979) for single durations. Under the proportional hazard assumption with a gamma mixing distribution we can perform the integration analytically to obtain the likelihood function:

$$L(.)=\prod_{i=1}^{n}\left\{\prod_{j=1}^{m_i}\left\{\frac{\lambda_0(t_{ij},\theta_1)}{\Lambda_0(0,\overline{T_i},\theta_1)}\right\}\times\right. \tag{56}$$

$$\left.\frac{\Gamma(m_i+\alpha^{-1}}{\Gamma(\alpha^{-1})}\frac{\left[\alpha\exp(x_i'\theta_2)\Lambda_0(0,\overline{T_i},\theta_1)\right]^{mi}}{\left[1+\alpha\exp(x_i'\theta_2)\Lambda_0(0,\overline{T_i},\theta_1)\right]^{m_i}}\right\}$$

Finding a good model for the baseline intensity may be of substantial practical importance and a factor limiting the applicability of standard parametric intensities is the empirically well documented "bathtub" shape of the failure rate of both repairable and nonrepairable systems. Some authors even claim that a complex system "... will invariably have the general characteristics of a bathtub curve. The bathtub curve is a ubiquitous characteristic both of inanimate, complex engineering devices and of living creatures." (Lewis, 1987, p.84)

The Weibull hazard, the most common nonstationary model, is either monotonically increasing or monotonically decreasing, the log-logistic hazard is either monotonic or has an "inverse bathtub" shape, as are other tractable parametric forms[39]. I therefore estimate flexible intensity models, one class being exponentiated polynomial and spline function, for example

$$\lambda(t)=\exp\left[\sum_{k=0}^{K}\theta_k t^k\right] \tag{57}$$

[39] Glaser (1980) discusses the conditions under which a parametric form can model the bathtub shape.

The problem with polynomials without exponentiation, which would lead to simpler estimation, is how to incorporate the non-negativity constraint. Another class of models, several of which are estimated in the next chapter, are additive intensities. As discussed earlier, there are different failure types. If we can approximate each type by a particular form, for example a Gompertz model for installation errors and substandard components and a Weibull model for wear-out, the system failure rate is the sum of both intensities:

$$\lambda(t)=\exp(\theta_1+\theta_2 t)+\theta_3\theta_4 t^{\theta_4-1} \tag{58}$$

3.2.3 Results

The performance of the failure time and the point process model is compared in an analysis of up time durations in European nuclear power plants (table 15a-f). In the point process model, the dependent variable is time in months since startup from the last refuel outage (i.e. the fuel cycle is the life of the system), in the failure time model, it is time in months since startup from the last outage. There are two reasons for considering fuel cycles as the unit of analysis in the process model. The first reason is based on a casual analysis of plant operations: we can observe that utilities attempt to return the plant to the running state as quickly as possible if an unscheduled event occurs and such outages often last only a few hours. Repair performed during these outages cannot affect more than a minor part of the complex system[40].

During refuel outages, however, we observe extensive maintenance and inspection activities in which sometimes more than one thousand people, often from outside the plants, are involved. Compared to the minimal repair of equipment failures, the extensive maintenance activities affect a substantial part of the plant. I therefore consider two stylized actions: one is to repair the plant upon failure which returns the plant to its operating state but does not change the current state of degradation, the other is to shut the plant down for refueling and maintenance and this resets the degradation process (renews the plant).

This assumption is not necessarily contradicted by the existence of plant aging (e.g. the conference papers published in International Atomic Energy Agency, 1983, 1988). The focus of the research on plant aging is on plant systems that cannot be replaced or were not constructed to be replaced (such as the pressure vessel and the containment) and on degradation problems over the whole life of a plant, i.e. 50 years or more. These degradation mechanisms therefore change the statistical properties of the plant very slowly and can be considered constant over the relative short duration of a fuel cycle and even over the observation interval of the data. This is consistent with a statistical analysis of the data, the second reason for choosing a "fuel cycle" model: the

---

[40] In fact, more detailed descriptions of outages in trade journals show that typically only the failure cause is removed and that no major preventive maintenance of other parts of the plant is performed.

nonparametric test of the renewal assumption following refuel outages was not rejected in chapter 2, whereas the renewal assumption at every event was. Nevertheless, I account for differences between older and newer plants in this section by including plant age as a regressor.

The regressor variables are the last refuel duration, age of the plant (at the occurrence of the event), capacity, and a dummy variable equal to one for BWR. The failure time model uses the time since the last refuel outage in addition[41]. The operator's decision to shut the plant down for refueling and maintenance is treated as independent censoring in both models. The distribution of censoring times (operating cycle lengths) are plotted in figure 21, table 16 gives descriptive statistics[42]. Chapter 4 develops a behavioral model for the distribution of operating cycle lengths.

Consider first the failure time model. The coefficient of time since last refueling is significantly (1%) negative for all countries and both the Weibull and the exponentiated quadratic polynomial baseline hazard. This implies that failures are higher at the beginning of the fuel cycle while the operators detect and

---

[41] This regressor is necessary in the failure time model to address the nonstationary detected in section 2.2 (rejection of renewal at every event). It could also be included in the process formulation unless the baseline hazard is exponential (in which both models coincide). Although the variable is ideally treated as continuously varying, its value was taken at the occurrence of the event (see the criticism of this treatment of time varying covariates in Heckman and Singer, 1985).

[42] The statistics in brackets include two very unusual fuel cycles. These two cycles were excluded in the analysis in chapter 4.

resolve problems caused by incorrectly installed equipment, imperfect maintenance, or substandard components installed during the refuel outages. The same phenomenon in the process model is measured by the shape parameter(s) of the baseline hazard. The other (statistically) significant regressors in the failure time formulation are the BWR dummy in Germany, and capacity in Sweden. The positive estimate of the BWR coefficient in Germany (implying more outages) reflects the problems with the design of BWR in Germany (and the U.S.). In particular, many of these outages were caused by problems in the piping system due to stress corrosion cracking. The negative coefficient on capacity in Sweden implies fewer problems with larger units. Since the larger units incorporate design changes following operating experience with smaller units, the coefficient reflects successful design changes[43]. The estimated baseline hazards in the failure time model do not exhibit substantial duration dependencies, which does not change when controlling for the possibility of unobserved heterogeneity in the Weibull baseline hazard model. Almost all of the duration dependence is captured by the (time-varying) covariate "time since the last refueling outage".

Although the failure time formulation is the standard way of modeling multiple spells, it compares poorly to the process

[43] A priori, I expected this effect to show up in the plant age coefficient. However, reconsidering the actual data, design changes are much clearer reflected in capacity for Sweden or France (but not, for example, for Canada) than in plant age.

formulation when comparing the values of the likelihood function[44]. The exponentiated quadratic polynomial baseline intensity in the process formulation is a very dramatic improvement over all other specifications[45]. Not surprisingly, this formulation is the only one in accordance with an error ("bugs") detection model when "bugs" in a plant can be detected and removed without necessarily causing a failure (see above). Contrary to the usual experience with duration analysis, estimates are relatively robust: the signs on almost all estimated coefficients are the same as in the other three model specifications. However, many more estimates become significantly different from zero and the following discussion refers to the process formulation with the exponentiated polynomial intensity.

Previous refuel duration has a significant positive coefficient in France, Sweden, and Switzerland (and it is positive, although not significant, in Belgium, too). This implies that unplanned outages are more likely to occur after a long refuel outage. Two structural explanations are consistent with this finding: a long refuel outage may indicate the existence of

---

[44] Although the log-likelihood for the Weibull baseline hazard is higher in the failure time formulation, this changes when including "up since refueling" as a regressor in the process formulation. I do not report this specification since I want to emphasize the conceptual distinction between the failure time and the process formulation.

[45] It may be somewhat disconcerting to see a positive value of the log-likelihood function, but this is quite possible for certain parameter values. In contrast to the exponentiated quadratic model or to a Gompertz hazard, an exponential model always gives rise to a negative log-likelihood function value.

particular problems that are likely to cause future outages. A typical example are problems with the piping system or the steam generator. A long refuel outage may also indicate that major repairs have been performed or new equipment has been installed and the higher outage rate is due to initial "teething" problems or a "shakedown" period.

The coefficient on capacity is negative and significant in Germany, Belgium, and Switzerland (and in Sweden for all other specifications except this one), but positive and significant in France. As discussed before, the negative coefficient may indicate successful design changes. The effect of plant age is not clear, but there is some indication (France, Germany, Switzerland) that newer units are less reliable, after controlling for plant design with the capacity regressor. This increase in unplanned outages is most likely due to an initial "shakedown" period new plants have to go through.

As discussed before, BWR's have been less reliable than PWR's in Germany, and the higher outage rate is reflected in a significantly positive coefficient on the BWR dummy. Sweden, on the other hand, has successfully developed her own design of BWR's. Having accumulated substantial experience with BWR's before the begin of the sample period (7 BWR had been on line before 1981, but only 1 PWR), it is not surprising to see that BWR's were more reliable as reflected in the significantly negative coefficient on BWR. The negative coefficient on the BWR dummy in Switzerland may not be very meaningful, other than as an indication of the extremely good performance of the BWR Muehleberg (the second Swiss BWR, Leibstadt only contributed observations for one year).

Table 15: Comparing Failure Time and Process Models

regressors are normalized to mean 0 and standard deviation 1
regressors: constant
up time since last refuel outage (only failure time)
last refuel duration
age of plant (time since first commercial operation)
capacity (MW)
BWR dummy (no BWR's in France and Belgium)

Table 15a: Means and Standard Deviations of Regressors

country	up since last refuel (hours)	last refuel duration (hours)	age (hours)	capacity (MW)	BWR (dummy)
Belgium	4313 (3084)	921 (367)	66044 (29021)	631 (254)	no BWR
France	4055 (2919)	1669 (921)	32036 (16498)	908 (45)	no BWR
Germany	5119 (3428)	1428 (1401)	67633 (31201)	915 (296)	0.36 (0.48)
Switzer-land	5250 (2920)	977 (290)	83259 (48435)	560 (289)	0.30 (0.46)
Sweden	3973 (2985)	1113 (490)	68817 (29912)	680 (177)	0.80 (0.40)

Table 15b: France 585 spells

	Process Model		Failure Time Model	
Baseline Hazard	Expoquad	Weibull	Expoquad	Weibull
mean logl	+1.065	-1.3798	-1.1725	-1.175
up since last refuel			-1.0915* (0.0728)	-0.9413* (0.0653)
last refuel duration	0.0727* (0.0252)	0.1179* (0.0566)	0.1410* (0.0479)	0.1276* (0.0485)
age	-1.5163* (0.0298)	-0.1414* (0.0566)	-0.0106 (0.0545)	-0.0001 (0.0546)
capacity	0.0945* (0.0369	-0.0258 (0.0501)	-0.0649 (0.0503)	-0.0651 (0.0508)
constant	-1.9695* (0.1206)	-0.1019 (0.0872)	-0.8475* (0.0926)	-0.6657* (0.0573)
shape1	0.463* (0.028)	0.6914* (0.0315)	0.0253 (0.0678)	0.9158* (0.0351)
shape2	-0.8458* (0.1582)		0.9383 (0.8248)	

Note: shape 2 is $100\gamma_2$ (see equation (5) in part 1) * is significantly different from 0 (from 1 for shape parameter in the Weibull hazard) at the 5% level.

Table 15c: Germany 116 spells

	Process Model		Failure Time Model	
Baseline Hazard	Expoquad	Weibull	Expoquad	Weibull
mean logl	-0.3361	-1.5629	-1.2626	-1.252
up since last refuel			-1.5057* (0.2236)	-1.222* (0.1806)
last refuel duration	-0.3570* (0.0451)	-0.1341 (0.1111)	-0.0446 (0.1077)	-0.0274 (0.1078)
age	-0.0050 (0.1132)	-0.0988 (0.2189)	-0.0686 (0.2130)	-0.0721 (0.2129)
capacity	-0.3396* (0.1019)	-0.2383 (0.2156)	-0.2042 (0.2047)	-0.1946 (0.2048)
BWR	0.2044* (0.0832)	0.4117* (0.1502)	0.3418* (0.1488)	0.3243* (0.1489)
constant	-4.0526* (0.4197)	-1.1197* (0.2171)	-1.9187* (0.3037)	-1.602* (0.1906)
shape1	0.4742* (0.0835)	0.6391* (0.0757)	-0.0933 (0.1557)	0.8014* (0.0877)
shape2	-0.3989* (0.4067)		1.7292 (1.5178)	

Note: shape 2 is $100\gamma_2$ (see equation (5) in part 1) * is significantly different from 0 (from 1 for shape parameter in the Weibull hazard) at the 5% level.

Table 15d: Sweden 232 spells

	Process Model		Failure Time Model	
Baseline Hazard	Expoquad	Weibull	Expoquad	Weibull
mean logl	+0.0150	-1.3911	-1.1299	-1.159
up since last refuel			-1.4792* (0.1391)	-1.192* (0.1172)
last refuel duration	0.1510* (0.0418)	0.0629 (0.0815)	0.0405 (0.0836)	0.03560 (0.0845)
age	0.3700* (0.0588)	-0.0765 (0.1177)	-0.0194 (0.1172)	-0.0020 (0.1168)
capacity	0.0269 (0.0715)	-0.4708* (0.1510)	-0.3742* (0.1408)	-0.3121* (0.1386)
BWR	-0.1633* (0.0497)	-0.1353 (0.1014)	-0.0655 (0.1019)	-0.04380 (0.1012)
constant	-2.6497* (0.1938)	-0.02487 (0.1278)	-1.2455* (0.1683)	-0.8622* (0.0995)
shape1	0.5815* (0.0404)	0.5864* (0.0435)	0.1402 (0.1050)	0.9590 (0.5962)
shape2*	-1.8101* (0.2048)		0.8277 (1.1519)	

Note: shape 2 is $100\gamma_2$ (see equation (5) in part 1) * is significantly different from 0 (from 1 for shape parameter in the Weibull hazard) at the 5% level.

Table 15e: Belgium 84 spells

	Process Model		Failure Time Model	
Baseline Hazard	Expoquad	Weibull	Expoquad	Weibull
mean logl	-0.5228	-1.5839	-1.2175	-1.254
up since last refuel			-1.5488* (0.2400)	-1.290* (0.1967)
last refuel duration	0.0592 (0.1054)	-0.1516 (0.1939)	-0.2024 (0.1828)	-0.1803 (0.1836)
age	0.1019 (0.0769)	-0.0335 (0.1862)	0.0326 (0.1904)	-0.1803 (0.1892)
capacity	-0.2556* (0.0773)	-0.2065 (0.1758)	-0.0029 (0.1839)	0.0102 (0.1848)
constant	-2.9037* (0.3415)	-0.6619* (0.2249)	-1.5181* (0.3171)	-1.217* (0.1914)
shape1	0.5822* (0.0697)	0.7489* (0.0936)	0.0964 (0.1906)	1.072 (0.1130)
shape2	-1.6773* (0.3382)		1.2040 (1.9788)	

Note: shape 2 is $100\gamma_2$ (see equation (5) in part 1) * is significantly different from 0 (from 1 for shape parameter in the Weibull hazard) at the 5% level.

Table 15f: Switzerland 43 spells

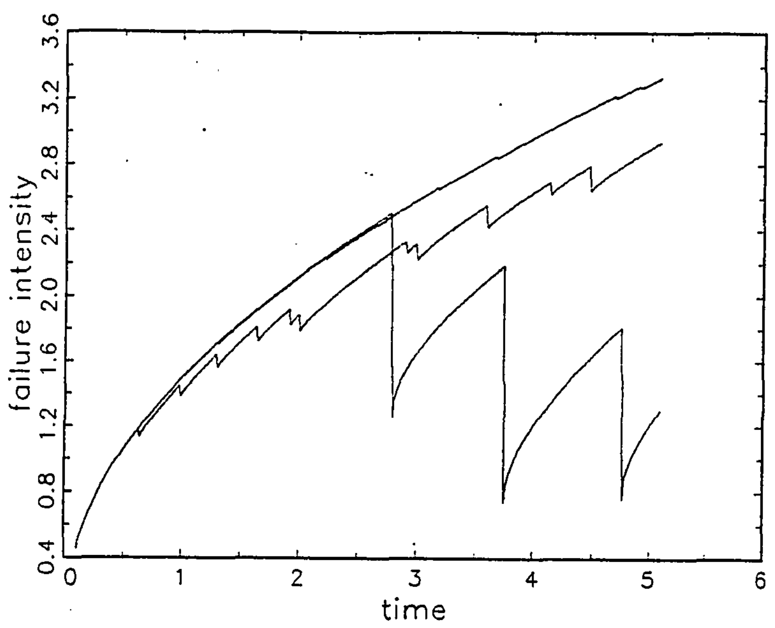
	Process Model		Failure Time Model	
Baseline Hazard	Expoquad	Weibull	Expoquad	Weibull
mean logl	-1.4281	-1.2063	-0.6296	-0.7151
up since last refuel			-3.8197* (0.7960)	-2.516* (0.5138)
last refuel duration	1.0069* (0.1583)	0.4945 (0.2730)	0.0542 (0.3038)	0.04607 (0.3048)
age	-1.4723* (0.4927)	-2.1182* (0.9476)	-1.5054 (0.8800)	-1.528 (0.8754)
capacity	-1.1297* (0.4305)	-1.5176 (0.8286)	-1.4233 (0.8407)	-1.472 (0.8463)
BWR	-0.4861* (0.1909)	-0.5004 (0.3435)	-0.5214 (0.4006)	-0.5815 (0.4088)
constant	-2.8801* (0.6925)	-1.8171* (0.4155)	-5.4990* (1.3310)	-3.123* (0.5944)
shape1	-0.4951* (0.2298)	0.5624* (0.1353)	0.6120 (0.3691)	1.048 (0.2070)
shape2	7.95116* (1.7482)		-0.9607 (3.1224)	

Note: shape 2 is $100\gamma_2$ (see equation (5) in part 1) * is significantly different from 0 (from 1 for shape parameter in the Weibull hazard) at the 5% level.

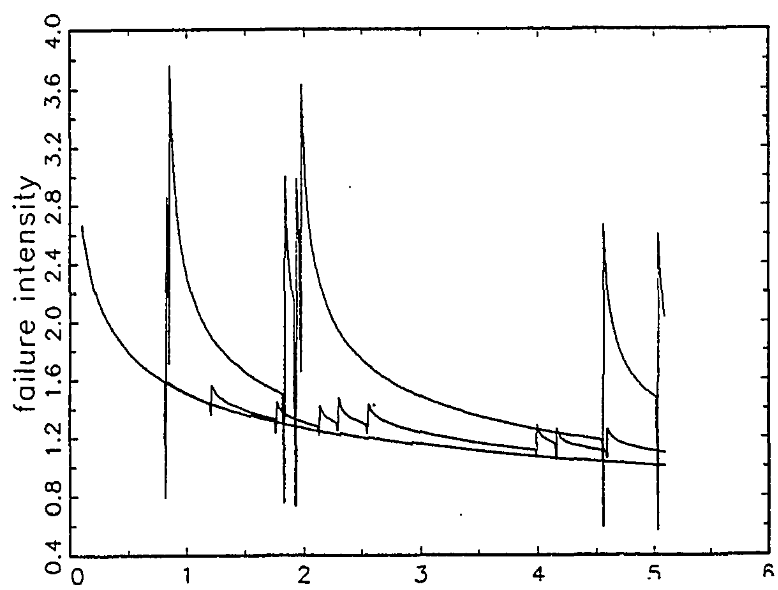
Table 16: Length of Operating Cycle (in days)

Country	Mean	Standard Deviation	Minimum	Maximum	N
Belgium	326	47	283	475	14
France	340 (351)	43 (68)	235 (235)	459 (660)	43 (45)
Germany	307	78	38	448	26
Sweden	347	81	223	586	28
Switzerland	305	81	6	342	16

Figure 19: Repairable Systems

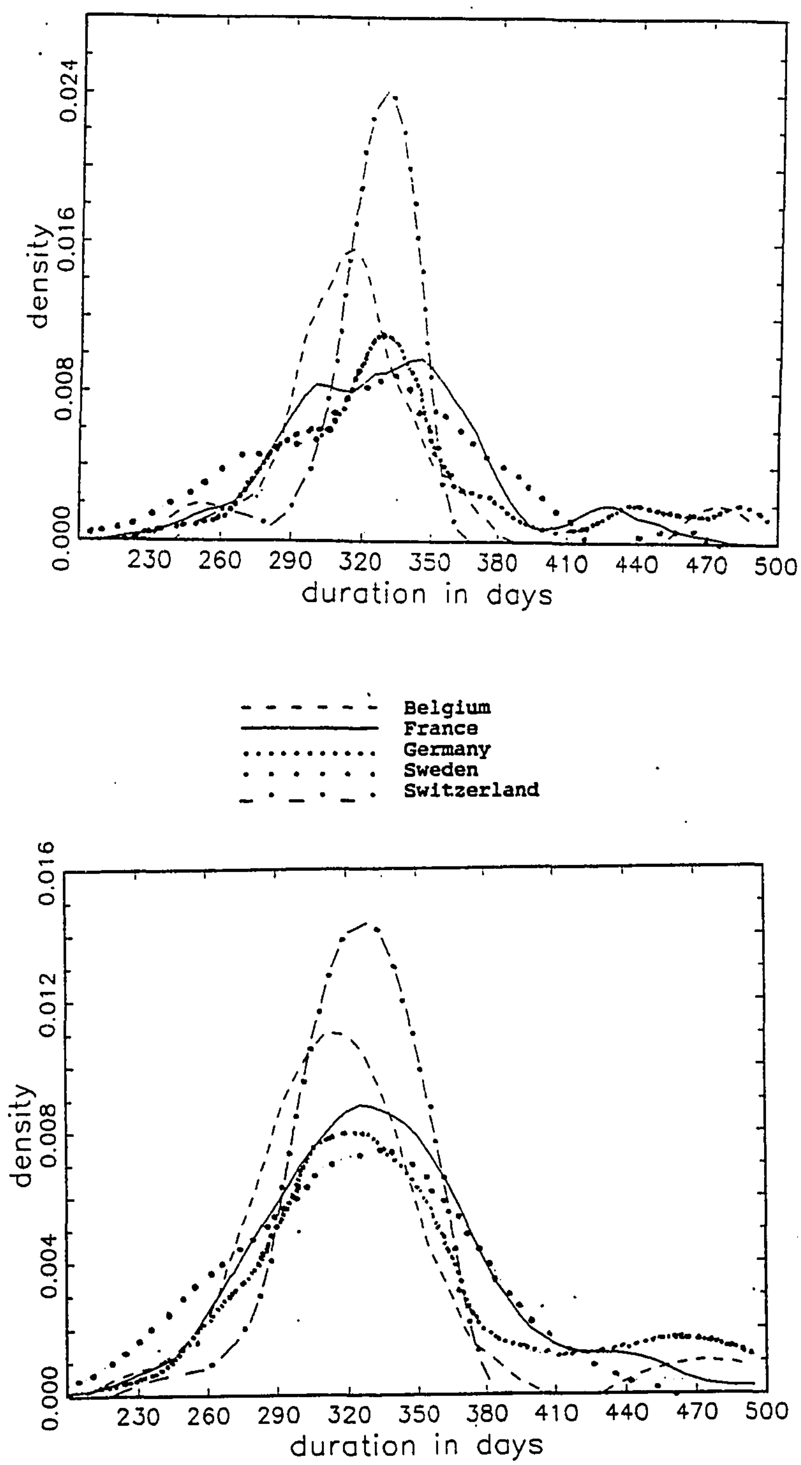


systems with 2, 20, 200 components, increasing hazard rate



systems with 2, 20, 200 components, decreasing hazard rate

Figure 21: Density Plot: Operating Cycle Duration



3.3 Can the statistical model explain time paths? A simulation study

Goodness-of-fit statistics, used in section 3.1, allow to evaluate how well different statistical models describe population data. A different way to evaluate the performance of statistical models is to compare their predictions of time paths with actual time paths. The focus of this alternative approach is to simulate medium to long run behavior and the dynamics of successive spells, rather than analyzing short run behavior, central to an analysis of individual spells.

The simulations are only done for France; the reason was that I could not obtain a useful estimate of the empirical time paths for a reasonably long horizon in any other country. To obtain the empirical time paths I used the following method: set time to zero at the first observed start up from refueling for each plant. Evaluate the state of the plant (states 0,1,2,3,4 as in figure 1, chapter 2) at t=1,...,60 where time is in months and calculate the probability of being in any particular state by dividing the number of plants found in each state by the total number of plants observed in that month. This gives the empirical point probability of being in any particular state. Consistent with section 2.3, I call the point probability of being in the running state "point availability". The time is plant time, but it is measured from the first observed restart from refueling, not from first criticality or the begin of commercial operation[46].

The ragged empirical estimate of the point availability in figure 22 obscures the main features. Therefore I smooth the raw estimate using a cubic spline smoother (see appendix 2) with smoothing parameter $\alpha=10$. The smoothed empirical time path is compared to two simulated time paths in two figures, one showing the first 24 months (figure 23), the second showing the first 60 months (figure 24). The number of plants in the sample drops substantially and the estimates become rather unreliable after 24 months; there are between 30 and 40 observations in the first 12 months and between 4 and 6 in the last 12 months. A confidence band would be very wide throughout and I therefore omitted it[47]. The two simulated time paths, both based on 500 replications and smoothed, correspond to the Markov model of chapter 2 (dashed line) and a point process model (dots and dashes). The parameters for the Markov model are the ones in table 1, column 2 (chapter 2). Clearly, the Markov model fails to model the long run waves and the

[46] Although plant time could be measured from first criticality or first commercial production, it is not particularly useful. The first cycle length has a very high variance for plant specific reasons, obscuring regularities of plant operations. Another alternative would be calendar time which evaluates the impact of structural breaks in regulation or the effect of changing fuel prices. This could be of interest for some data sets, but my data set is short and spans the (relatively quiet) years between the two major nuclear accidents (1981-1986).

[47] An approximate confidence interval could be calculated by considering each point availability as sampling from a binomial. For an availability of 70% and a sample of 40 plants an approximate 95% confidence interval would be [0.56,0.84], for 4 plants it would not fit on the plot anymore.

system converges to an almost steady state much too quickly.

The point process model is based on the simplest specification possible that captures the main features: the length of the operating cycle (the operators decision to shut down the plant for refueling), the refuel duration, and the duration of unplanned outages (only one type is considered) are random draws from gamma distributions with two parameters, the point process intensity is an exponentiated quadratic polynomial. Thus the point process model has only one more parameter than the Markov model; parameter values are in table 17. The time path is very close to the empirical time path, the most salient discrepancy is a small phase shift, and it could be made even closer by using more elaborate distributions. Alternatively, different simulations can be compared by the sum of squared deviations from the empirical point availability[48].

Figure 24 plots the complete time path for 5 years. Unfortunately, the empirical estimate is very unreliable. The cubic spline smoother allows to take into account the higher variance towards the tail, but I face a dilemma: heteroskedastic smoothing decreases the amplitude and makes the empirical path look like the simulated point process. However, this might be an artifact since

---

[48] This was suggested by Frank Wolak. I summed the squared deviation calculated at the end of each of the 60 months without weighting (one could also use the integral of squared deviations and/or weight by the number of plants observed). The sums of squared deviation from the empirical point availability are: 0.77 for the "naive" model of descriptive statistics, the mean point availability, 0.75 for the Markov model, and 0.71 for the point process model.

the point estimates, although unreliable, indicate that the amplitude does not decrease. I decided to use homoskedastic smoothing which shows up as a discrepancy between the point process path and the empirical path in figure 24.

Discrepancies between simulated statistical models and empirical time paths in the plots are have three causes: amplitudes frequencies, and levels. Each of them may point to a different inconsistency between statistical model and reality.

<u>Amplitude</u>

A decreasing amplitude is typical for an ergodic process. Markov or semi-Markov models are ergodic, but inconsistent with the data as demonstrated by the rejection of the renewal assumption (section 2.2) and the plots here. The implementation of the point process is ergodic, but this could easily be changed. For example, the plant manager might choose the operating cycle length such that the plant is refueled at constant intervals. The next chapter develops a behavioral model of how the operator chooses the operating cycle length which gives rise to a time path whose amplitude is declining more slowly. It is not clear from the data which assumption is correct. The homoskedastic smoothing does not show decreasing amplitudes, the heteroskedastic smoothing shows decreasing amplitudes very similar to the simulation of the process model in figure 24. A longer panel would be necessary in order to decide this problem.

Frequency

There is a slight difference between the phases of the simulated process path and the empirical path in the first two years (the difference increases substantially afterwards, but this is likely to be noise). I attribute this difference to the effect of plant time. Many of the plants in France came on line during the observed period, the main reasons for the sharp drop in observations after two years. Fuel cycles in the early years of the plant are different than later fuel cycles when the core has reached an equilibrium configuration. The process model was estimated treating all fuel cycles the same (no regressors were used in the simple simulation). Another possibility, irrelevant for France, but probably of significance in the U.S., could be the effect of calendar time. In the U.S., fuel cycles have become longer over time (EPRI, 1987,1989), but not in Europe. In fact, some German plants have even shortened their fuel cycles in order to improve fuel efficiency (Biblis A and B).

Levels

The empirical point availability lies above the simulated point availability for most of the latter part of figure 24. The likely reason for this is the change in the sample composition. The plants which contribute observations to the tail finished their first fuel cycle before 1980. If there exists learning effects in plant age (see section 2.4), we expect these plants to be more reliable than the "average" plant for which the process model was estimated. Again there might be an additional effect of calendar time, plant availability has risen over time, but I expect it to be minor. In a longer panel, however, the calendar time effect may be important, especially if the period of observations covers structural breaks such as the TMI or Chernobyl accidents.

Table 17: Parameter values for simulating point process

gamma distribution	time measured in	beta	rho	
unplanned outage durations (spell time)	days	6.93	0.65	
refuel durations (spell time)	days	16.23	3.93	
operating cycle lengths (cycle time)	days	6.41	52.36	
point process $\exp(\gamma_1+\gamma_2 t+\gamma_3 t^2)$		γ_1	γ_2	γ_3
failure intensity (cycle time)	months	-0.265	-0.240	0.015

Figure 22: Time Path: Empirical Point Availability

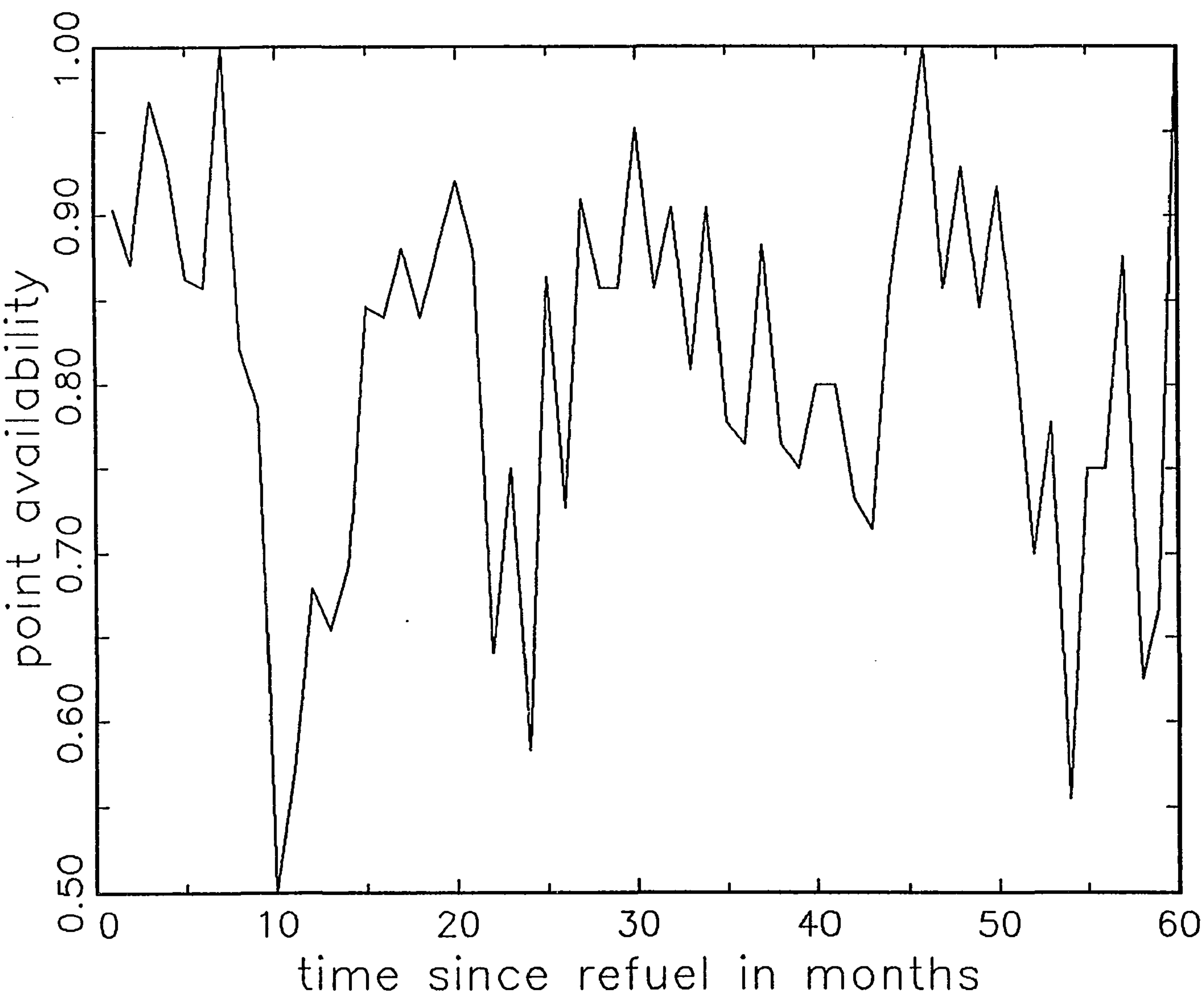
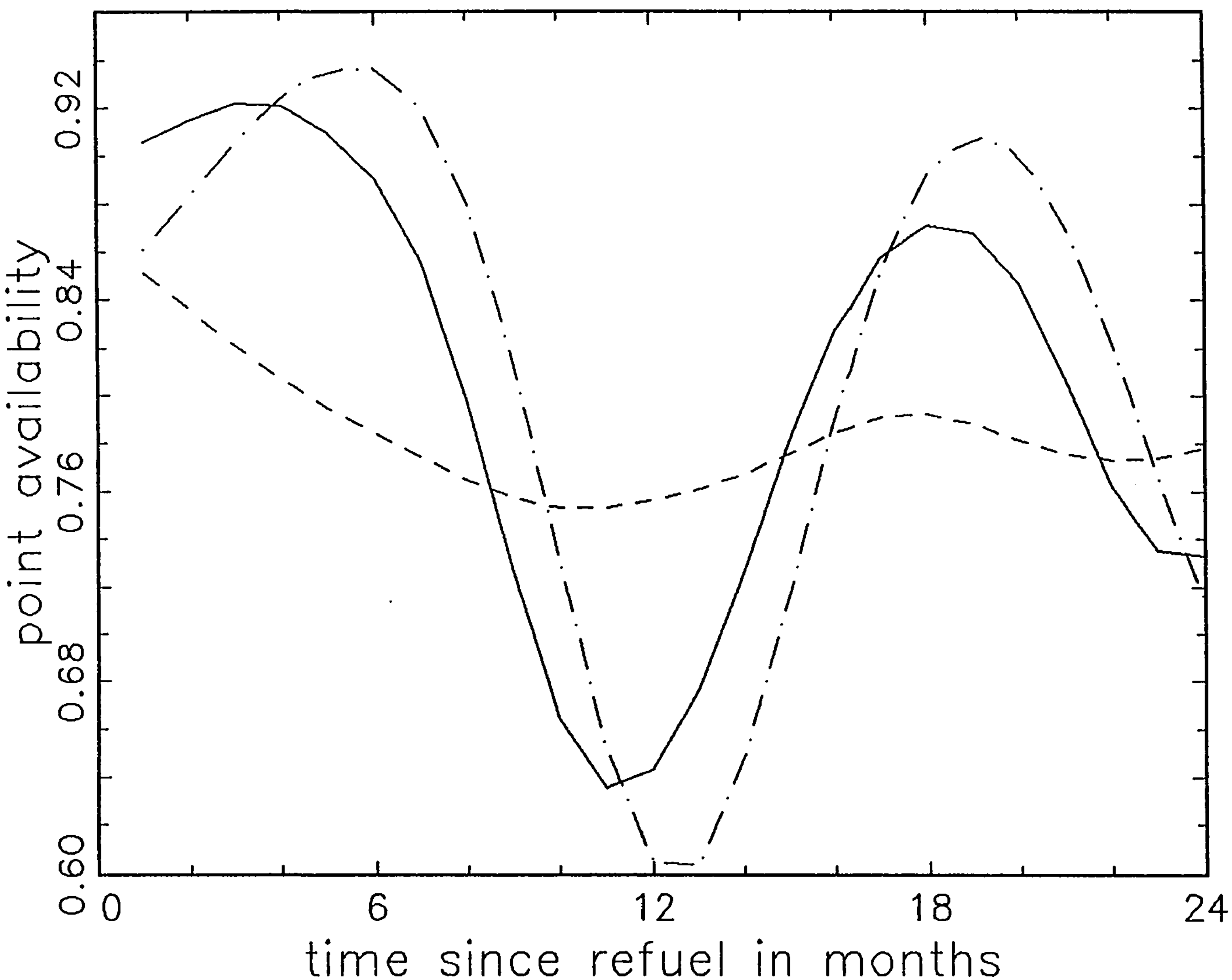


Figure 23: Simulations of Time Path



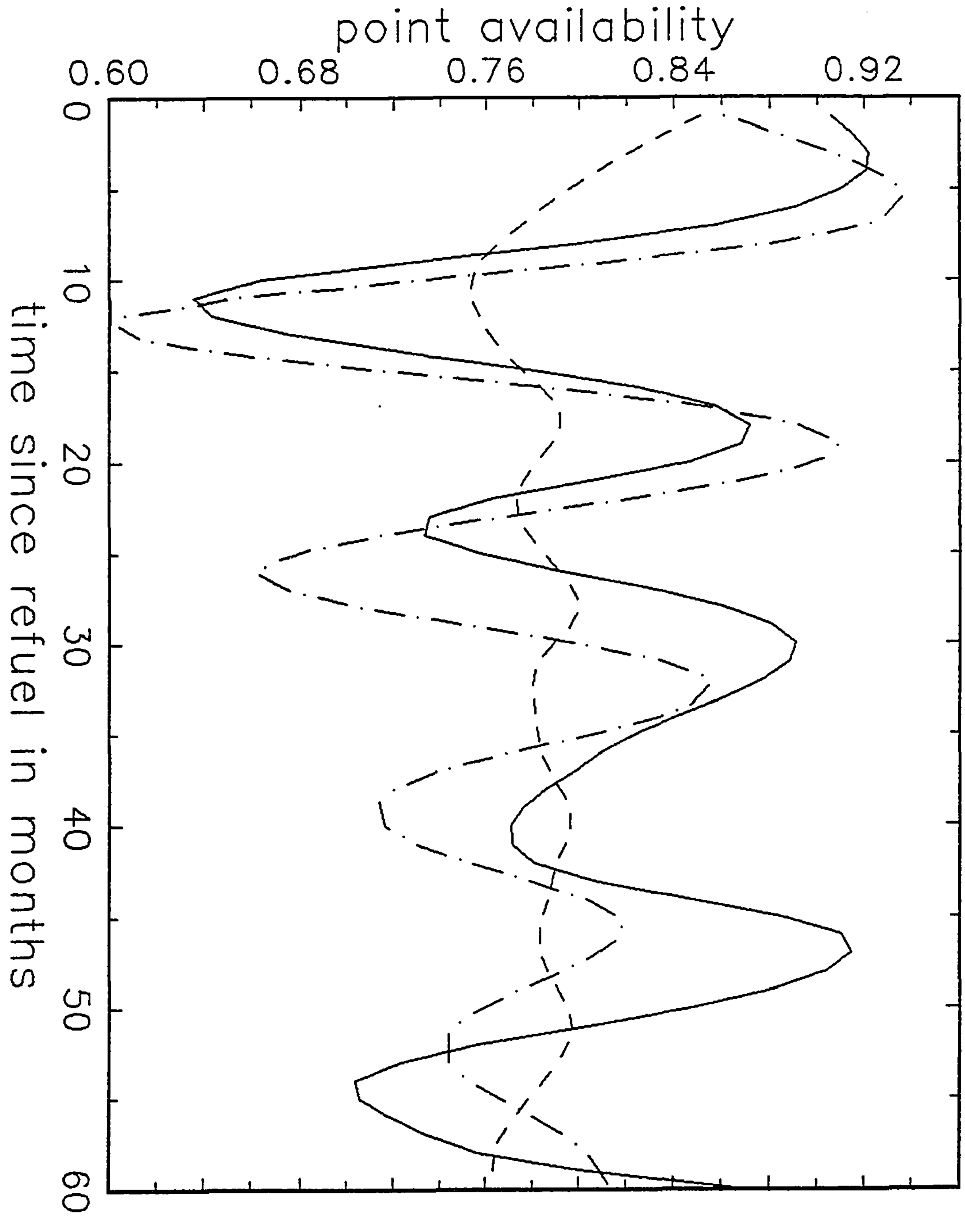


Figure 24: Simulations of Time Path

4. A structural economic model of operating cycle management

The aim of this chapter is to model and estimate the production process in nuclear power plants in five European countries at a structural economic level. Rather than abstracting from the actual production process as in the standard "black box" approach embodied in neoclassical production theory, I explicitly analyze a plant operator's production decisions in the context of controlling a stochastic process. Focusing on operating cycle management, this approach brings together various measures of the performance of the nuclear power industry such as availability, unplanned outage rates, and planned outage durations in a unified dynamic economic model. Under the assumption that plant operating experience reflects the optimal solution of a stochastic control problem, I estimate a plant operator's utility (cost) function from the operating history of the plants instead of relying on exogenous cost measures. The advantages of this approach, which is analogous to a revealed preference analysis in the theory of consumer demand, are two-fold: not only does it circumvent the problem of having little reliable information about the operating costs in European nuclear power stations, but important intangible cost components, such as the public's or the regulatory authorities' reactions to unplanned failures, can only be captured in this way. Using the resulting "cost function" parameter estimates, chapter 5 discusses how the model relates the variation in different performance measures across countries to differences in their economic/regulatory environment. There it will be shown how one can calculate the policy effects of changing relative prices through taxation or regulatory penalties within each country.

Rather than relying on assumptions that guarantee a closed form expression for the likelihood function, I use an "algorithmic" approach, i.e. the likelihood function is defined only implicitly as the solution to an economic decision problem (see chapter 1). The approach liberates the researcher from having to make a priori assumptions regarding functional forms. Unfortunately, as previous authors following this recently introduced methodological approach have found (e.g. Miller, 1984, Pakes, 1986, Rust, 1987, Wolpin, 1987), the conceptual advantage of precise modeling of the economic problem is bought at the cost of increased computational demands which themselves tend to be limiting.

The model is related to duration analysis and behavioral economic duration models[49], but there are several conceptual differences. While duration models typically deal with a single failure cause[50], the production process in power plants may be interrupted by a number of different events (competing risks) with

---

[49] Examples of this type of research are Wolpin (1987), who considers the decision of when to accept the first full time job after leaving school, Rust (1987), who models the decision of a manager when to replace a bus engine, and David and Mroz (1989), who estimate a model of fertility regulation from data on sequences of birth intervals.

[50] In Wolpin (1987), the "failure" that ends a spell is the decision to take a job, in Rust (1987), it is the replacement of the engine, in David and Mroz (1989), it is conception leading to a live birth. Ryu's model (1990) is an exception; he derives an estimator for a bivariate duration model, but his work is theoretical and offers no application.

different effects on the statistical properties of the system. As discussed in chapter 3, most events will not regenerate the system as they do in renewal processes, i.e. a power plant is not "new" after starting up from an equipment failure and intervals between successive events are not independent. Furthermore, plant managers have a choice between different repair and maintenance actions, and downtime durations for refueling are not negligible. The typical assumption in microeconomic studies of replacement/maintenance problems (e.g. Rust, 1987, Ryu, 1990) has been that only a *single* action is available to the agent and that this action immediately regenerates the system.

4.1 Descriptive statistics for the sample

I select a restricted sample of plants for the analysis in this chapter. Some reactors are prototypes representing different technologies (Fast Breeders, High Temperature Reactors), some operators are unreliable in reporting their outages (e.g. KKP, the operator of the two Philippsburg units in Germany), and focusing on operating cycle management eliminates technologies that do not have such cycles such as the old gas-cooled French units and Canada's CANDU reactors. I further restrict the sample by only considering commercial LWR's that began commercial operation between 1970 and 1980 and for which there is information on all years. Therefore we have technically very homogeneous plants in each country. Although the renewal assumption during refuel outages was not rejected (section 2.2), the reliability growth model of section 2.4

indicated, not surprisingly, that new plants starting operation undergo a learning period. The learning process leads to complex behavioral models which I plan to investigate in the future, but which are outside the scope of this dissertation. One such behavioral model has been analyzed in Sturm (1989b). It is also well known that the first fuel cycle has very different characteristics than the following cycles for technical reasons; already the second and third cycle are very similar to later ones. Finally, it seems inappropriate to pool data from different countries; all preliminary tests of homogeneity convincingly rejected the null hypothesis that different countries have the same operating experience (see both the descriptive statistics of chapter 2 and the following tables for the sample analyzed here). Since I could not pool, I had to ignore a few orphans: 1 BWR each in Switzerland and Germany, 1 PWR in Sweden. The remaining sample contains 28 plants (3 in Belgium, 11 in France, 5 in Germany, 6 in Sweden, and 3 in Switzerland). To avoid the problem of left censoring, the operating histories were considered from the start up after the first observed refueling.

To indicate the magnitudes involved, I have calculated the commonly reported descriptive statistics. Availability and capacity factors, defined in chapter 1, are reported in table 18. The only major difference between these two measures is for France in 1986. France brought many new plants on line in the 80's and the overcapacity resulting by the end of the decade has reduced capacity factors by more than five percentage points below availability. The first effects of generating capacity outstripping demand became visible around 1985/86.

Table 19 reports statistics on unplanned outages that were safety- or reliability-related, and for which the plant manager was held responsible. The unplanned outage rate addresses a different aspect of plant performance. Almost all of these outages involved scrams initiated by equipment failures during plant operation or testing, or by operator error. Some outages due to more controlled shutdowns for unplanned repairs are also included. Outages unrelated to reactor operations per se, such as labor disputes or grid failures, have been excluded[51]. In Germany and Switzerland, there were 1 or 2 unplanned outages per reactor year; the unplanned outage rate in Belgium averaged about 3, it was about 4 in Sweden, and even higher in France. This is even true for all plants, not just the sample selected for this analysis (see the survivor function for reliability in chapter 2).

Table 20 shows that the durations of unplanned outages account for only a small proportion of total time (the sum of all up- and downtimes) and are therefore almost negligible from an economic point of view. However, the occurrence of unplanned outages is likely to remain important (the reliability issue) since failures

---

[51] A small discrepancy (3 to 5 percentage points) exists between the availability factor in table 18 and its breakdown into planned and unplanned unavailability in table 20. There are two reasons: the availability and capacity factors in table 18 include a) the effects of such exogenous outages and b) partial unavailabilities, i.e. output losses which are not due to a full shutdown and which are not reported. These losses are therefore not included in the breakdown in table 20.

are highly visible and, due to their relationship to actual and perceived safety, could affect the whole industry. These intangible costs accrue in addition to the direct costs of repairing defective equipment and the substantial wear on the plant caused by a scram. Compared to unplanned outage durations, the durations of planned outages account for a substantial part of total time.

Utilities in Europe have operated on planned fuel cycles of approximately 12 months. This is in contrast to U.S. utilities, many of which have lengthened their planned cycles to 18 or even 24 months (see EPRI, 1987, 1989). In recent years (after the end of the sample period), some utility companies in Europe, in particular in Belgium, have introduced longer cycles, mainly for newer plants (the sample only contains plants that were on line by 1980). Sweden, where there may have been longer planned cycles even within the sample period (some realized cycles were unusually long), might be an exception. However, Sweden's reactor technology (BWR) is somewhat different from that in other countries (all other reactors in the sample are PWR). In contrast to planned cycles, realized cycles (excluding refueling) have a substantial variation, the standard deviation is approximately 1 1/2 months in Belgium and France and over 2 1/2 months in the other countries (see table 16, chapter 3); this does not change when the refuel duration is included. Some cycles were truncated very early whereas some cycles lasted over one year. Why does this happen? Do plant manager randomly deviate from the carefully planned the original cycle? The approach taken here is that some conditions have changed and that the variation reflects new information about the environment that the plant manager takes into account but which we cannot observe. For example, an unexpected period of very low replacement energy costs (e.g. low demand due to mild weather) or discovering that the costs of an equipment failure are higher than usual (maybe due to technical problems in other plants and a suspicious regulator) may make it opportune to refuel earlier. Similarly, manpower constraints (additional shift or overtime expenses), high energy replacement costs, or less than expected use of fuel can make it desirable to operate the plant longer than planned. Of course, the cycle length will be limited by the amount of burnable fuel available and the need to comply with (or at least not violate blatantly) regulations such as authorized limits for fuel burn up. Unfortunately, the main causes which allow stretching the cycle are not observable. They include the effects of partial outages and operator action (so-called core-preservation operations) and we will have to incorporate them into an "error" term describing our ignorance about important decision variables. Unplanned full outages, which also save fuel, are reported, but their overall effect is smaller and will therefore be modeled with the unobserved effects. As always, a rule must have its exceptions and, indeed, there are two fuel cycles in France where some very unusual unplanned outages have a clear causal effect on fuel cycle length. In table 16, chapter 3, the statistics in brackets for France include these two cycles, but the two cycles are excluded in the analysis of this chapter.

Table 18: Availability and Capacity Factors

Country	Availability in 1986	Capacity Factor in 1986	Cumulative Availability	Cumulative Capacity Factor
Belgium 3 units	67.8	65.8	78.1	78.0
France 11 units	79.7	73.4	73.9	69.6
Germany 5 units	72.3	69.6	77.1	75.4
Sweden 6 units	82.8	80.9	76.5	73.4
Switzer-land 3 units	84.9	84.5	82.6	82.1

Table 19: Unplanned Outages 1981-1986

Country	Unplanned Outages	Unplanned Outage Rate per Reactor Year	Number of Reactors
Belgium	53	2.94	3
France	335	5.08	11
Germany	30	1.00	5
Sweden	138	3.90	6
Switzerland	12	0.67	3

Table 20: Planned and Unplanned Outage Duration as Proportions of Calendar Time

Country	Unavailability due to Planned Outages in 1986	Lifetime Unavailability due to Planned Outages	Unavailability due to Unplanned Outages in 1986	Lifetime Unavailability due to Unplanned Outages
Belgium	0.28	0.15	0.02	0.03
France	0.12	0.17	0.06	0.07
Germany	0.23	0.16	0.00	0.03
Sweden	0.09	0.13	0.03	0.05
Switzerland	0.12	0.12	0.02	0.03

4.2 The conceptual model of plant operations and the plant manager's decision problem

There are two conceptually different parts to the economic model of fuel cycle management. One part deals with the exogenous technical system, which I call the plant model, the other part is the plant manager's control of this mechanism. The model of the exogenous mechanism, which can be estimated using standard techniques (see section 4.3), is built on the results of chapter 3 and I only describe it briefly before considering the control decision. Being able to estimate the exogenous process separately permits a two-stage procedure which substantially simplifies estimation of the plant manager's cost (or utility) function. The technical plant process is assumed to be exogenously given and the manager cannot change its characteristics. This assumption, made throughout the chapter, is appropriate for existing production facilities in a stationary environment. Nevertheless, differences in the plant process between countries exist and these differences reflect institutional and environmental factors that are not captured by the behavioral model.

We can observe that utilities attempt to return the plant to the running state as quickly as possible if an unscheduled event occurs and such outages often last only a few hours. Minimal necessary repairs are performed during these outages which affect a small part of a complex system. During refuel outages, however, we observe extensive maintenance and inspection activities in which sometimes more than one thousand people, often from outside the plant, are involved. To distinguish the two activities, I consider two stylized actions: one is to repair the plant upon failure which returns the unit to its operating state but does not change the current state of degradation, the other is to shut the plant down for refueling and maintenance which resets the degradation process. Unless the plant is shut down for a major overhaul (refueling), it is always better to repair broken equipment immediately than leaving the plant down. The minimal repair then becomes part of the plant model and is not considered as a control variable any longer. The observed operating history can be modeled as a sequence of censored point processes. A point process, causing the (immediately rectified) unplanned outages, begins when a plant restarts after refueling and is censored when the plant manager decides to take the plant down for the next refueling. The statistical assumptions in this model have been spelled out and justified in chapters 2 and 3.

Fuel cycles consist of two durations which are controlled by the plant manager within limits. The first is the duration of the operating cycle, i.e. the interval between the startup after one refuel outage and the beginning of the next refuel outage. The second is the duration of the refuel and maintenance outage. Although the plant manager can always decide to shut down the plant immediately for refueling, the decision to bring it up is subject to regulatory and technical constraints. The combined effects of the plant manufacturer's recommendations, "sound engineering practice", and regulatory constraints may be such that the refuel duration is a random variable beyond the immediate control of the plant manager[52].

When Is The Plant Refueled?

Although it is convenient to estimate the exogenous plant process in continuous time, as I have done in the previous chapters, the decision problem, which is the central element of the complete model, is best set up as a dynamic programming problem in discrete time. Conceptually, this entails no loss of generality because a continuous (Markov) decision model can be approximated arbitrarily closely by a discrete model (van Dijk, 1984). As a practical matter, Bellman's "curse of dimensionality" will require restricting the size of the state space to keep the problem computationally manageable.

In each time period, the plant manager observes that the plant is in one state of a state space S. After observing the state of the process, the plant manager must choose an action from the control set A. Independent of the past, the plant manager incurs a reward (costs) R(s,a), a function of state $s \in S$ and action $a \in A$. Then the next state of the system is chosen according to a transition function $P: S \times S \times A \rightarrow [0,1]$ such that for each $s' \in S$, $P(s',.,.)$ is a

[52] This will be the assumption throughout this chapter. The data must be discretized to solve the dynamic programming model and this discretization makes it impossible to distinguish a structure in refuel durations. Given the data and computer limitations, I have to focus on the phenomenon with the highest variation and that is without any doubt the duration of the operating cycle; see the density plots for operating cycle and refuel durations.

probability measure on (S,S,A).

The choice of actions is between repairing malfunctions that occur randomly or shutting down the plant for refueling and major overhauls, i.e. A has two elements for all states. The action has no effect when the plant is in a refuel state[53]. The agent chooses a policy (or plan) Π which is a sequence of functions $\Pi=\{\Pi_n\}$, $n=0,1,\ldots$ from S to A. The economic problem is to find a policy that maximizes the plant manager's optimality criterion. A stationary policy Π is defined as a time invariant function Π^* mapping S into A.

I assume that the plant manager maximizes a time separable utility function with discount factor $\beta<1$:

$$V_\Pi(s_0) = E_\Pi\left[\sum_{n=0}^{\infty} \beta^n R(s_n, a_n)\right] \tag{59}$$

One of Blackwell's (1965) results is that if A is finite and R() is bounded, there exists an optimal stationary policy. The optimal value function V() = $\sup_\Pi V_\Pi()$ under such a policy satisfies the functional equation and the solution to the functional equation is unique.

$$V(s) = \max_a \left[R(s,a) + \beta \int_{s'} V(s')\, dP(s'|s,a)\right] \tag{60}$$

This result is important since it allows calculating the optimal value function by finding the fixed point of the functional

---

[53] Remember that refuel duration is assumed to be a random variable beyond the plant manager's control (footnote 52).

equation. When considering unobservable states to accommodate the fact that we ignore some important information, it may be desirable to allow for unbounded returns and costs and we need conditions under which the value function under the optimal solution remains the unique fixed point. Rust (1988, Theorem 3.1) gives conditions under which there exists an optimal stationary Markovian policy satisfying the functional equation. Lippman (1975) provides an alternative set of conditions.

No matter how detailed our observations are, they will never be complete to the extent of including all important time varying decision variables such as energy replacement costs or scheduling constraints. These unobserved variables are the cause for observed deviations from the initially planned annual operating cycles and they trace out the plant manager's cost function conditional on the initial plan. We would be able estimate the "envelope" cost function only if the initial plan had no economic impact, i.e. in the implausible case that there is no cost for deviating from it. Since the planned cycle length was the same in all countries, the estimates are comparable.

Partition the state space into two subspaces $\mathbf{S} = \mathbf{X} \times \mathbf{Y}$, where the data provides information only about $\mathbf{X}$. With exception of the refuel state, there are two dimensions in the observable state space $\mathbf{X}$: one dimension represents the time since the last startup from refueling and the other dimension counts the number of unplanned outages that occurred in the discrete time interval. The rewards/costs incurred during the operation of the plant are functions of both the observed and the unobserved states and of the action taken by the plant manager R(x,y,a). I will assume that the reward function has the additive form

$$R(x,y,a)=R(x)+g(y,a) \tag{61}$$

The unobserved state that enters the reward is a random function of the action chosen. Furthermore, the unobserved states only affect the reward function, not the transition function. This last assumption is not innocuous since it rules out unobserved states in the technical process such as plant specific problems. I minimize this potential misspecification by considering groups of homogenous plants. There remains the possibility of fuel cycle specific technical heterogeneity, but modeling this possibility by using a nonparametric random coefficient model is not successful: the variance of observed durations requires a very large number of points of support relative to the number of cycles and the maximum likelihood estimates imply a zero probability for cycle durations that are not observed. A model allowing for a large range of different values (such as a Gamma mixing model) creates a dimensionality problem which renders estimation impossible on any computer that I have access to.

The dynamics of the model are the following: The plant manager observes the states x and y and then chooses an action a. Based on the states and the action chosen, the plant manager receives the reward or pays the costs. Then time is advanced by one unit, here one month, and the exogenous process moves to a new state according to the transition function corresponding to the chosen action. If the plant manager decides to refuel, the plant is in refueling with probability 1 in the next period; if the plant manager decides to keep the plant up, the state is chosen according to the exogenous process. The monthly (conditional) probability of leaving refueling is a constant (<1) since only one refuel state is considered.

There is a vector of constant parameters θ known to the plant manager, which represent the discount factor, parameters of the cost function and parameters of the point process. The likelihood function of the observed data is the product of a transition probability and the probability that the plant manager chooses action a. Let $P(x'|x,a,\theta)$ be the marginal transition probability function and $Q(a|H_n,\theta)$ be the conditional probability of choosing action a given the <u>observed</u> history up to and including state n[54]. The contribution to the log-likelihood function for a plant which is observed for N stages is

$$l(\theta)=\sum_{n=1}^{N}\{\log Q(a_n|H_n,\theta)+\log P_x(x_n|x_{n-1},a_{n-1},\theta)\} \qquad (62)$$

with x_0 and a_0 being given constants; we begin observations with a startup from refueling. If we did not observe the startup, the left-censoring problem would make estimation impossible without extremely strong and unappealing assumptions. This is similar to

[54] When the plant is in the refuel state and the plant manager does not take the plant down for refueling (i.e. does not censor the point process of unplanned outages), the marginal transition probability P() is the discrete version of the point process discussed before.

left-censoring in failure time models (e.g. Heckman and Singer, 1985). The difficult problem is to obtain P() and especially Q(), which is implicitly defined and needs to be calculated from a numerical solution of the dynamic program.

4.3 Econometric Specification

None of the previous discussion depended on any particular parametric assumption. Several different parametric models for the exogenous plant process and the plant manager's cost function have been estimated for one parametric model for the unobserved states. Notationally, it is convenient to split the vector of parameters to be estimated (θ) into the parameters of the exogenous process (γ), the parameters of the cost or reward function (r,c), and the discount factor β.

4.3.1 The exogenous plant process

A factor limiting the applicability of standard parametric intensities, such as the Weibull or the log-logistic hazard, to estimate the point process generating the unplanned outages is the empirically documented "bathtub" shape of the intensity function of both repairable systems and nonrepairable systems (see section 3.2). Because ignoring the empirical regularity of a non-monotone hazard may cause very misleading statistical estimates, I have also estimated flexible intensity models to ascertain the robustness of the results[55]. The additive models have a theoretical justification in complex systems with different failure modes and can estimate the relative importance of different modes. In particular, the sum of a Weibull and a Gompertz intensity models the effects of wear-out failures and of initial "shakedown" or "teething" problems (see section 3.2)[56]. Since the discretization of the data is only necessary to solve the control problem, I first estimate the parameters of the discrete transition function P() corresponding to the occurrence of unplanned outages using the original data and the continuous point process models. Although separating the exogenous process and the economic decision in this two-step procedure may cause some efficiency losses, these losses are likely to be more than offset by the gains of using the additional information in the non-discretized data which would not be possible in a full maximum likelihood model. The failure rate is modeled as follows:

Each operating cycle i has an intensity $\lambda_i(t,\gamma)$ of having unplanned outages, where γ is a vector of unknown parameters, and is observed over the finite interval $(\bar{T}_i)$. To simplify notation, I drop the dependence on γ. For a sample of n such operating cycles

---

[55] If the intensity function is indeed bathtub-shaped, an intensity function imposing monotonicity can give negative, positive, or no duration dependence, depending on when the plant manager decides to interrupt the process. The estimates of the exogenous process have a major impact on the estimates of the cost function and an unjustified restriction in the plant model could lead to meaningless cost function estimates.

[56] New or repaired units tend to suffer from lower performance at the beginning as installation errors are rectified and substandard components replaced.

with fixed censoring time $\bar{T}_i$ and m_i observed unplanned outages at times $0=t_{i0}<t_{i1}<t_{i2}<\ldots<t_{im}<\bar{T}_i$ for process i, we have the log-likelihood function:

$$
\begin{aligned}
l(.) &= \sum_{i=1}^{n}\sum_{j=1}^{m_i}\{\log\lambda(t_{ij}) - \Lambda(t_{ij-1},t_{ij})\} - \sum_{i=1}^{n}\Lambda(t_{im_i},\overline{T_i}) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m_i}\log\lambda(t_{ij}) - \sum_{i=1}^{n}\Lambda(0,\overline{T_i})
\end{aligned}
\qquad (63)
$$

This assumes that the censoring variable $\bar{T}_i$ was constant or at least noninformative (Kalbfleisch and Prentice, 1980) in which case l(.) can be interpreted as a conditional likelihood for which all the usual properties hold (Cox and Hinkley, 1974). However, $\bar{T}_i$ is a control and therefore a function of the environment the plant manager faces, including economic incentives and the process itself. Then the full likelihood will involve a term describing the distribution of $\bar{T}_i$ as a function of γ and other parameters which reflect the plant managers cost function. Nevertheless, it is possible under appropriate assumptions to interpret equation (63) as a partial likelihood (Cox, 1975) and estimate the parameters γ separately (Rust, 1988). The "appropriate assumptions" in this context require that unobserved information and missing variables only affect the cost function and not the plant process (see Rust, 1988).

Table 21 summarizes how the following parametric models for the exogenous plant process compare:

A) an exponentiated quadratic intensity

$$\lambda(t)=\exp(\gamma_1+\gamma_2 t+\gamma_3 t^2) \tag{64}$$

B) a Weibull intensity

$$\lambda(t)=\exp(\gamma_1)\gamma_2 t^{\gamma_2-1} \tag{65}$$

C) the sum of two Weibull intensities

$$\lambda(t)=\exp(\gamma_1)\gamma_2 t^{\gamma_2-1}+\exp(\gamma_3)\gamma_4 t^{\gamma_4-1} \tag{66}$$

D) the sum of Gompertz and Weibull intensities

$$\lambda(t)=\exp(\gamma_1+\gamma_2 t)+\exp(\gamma_3)\gamma_4 t^{\gamma_4-1} \tag{67}$$

In model A), the estimated coefficient on time squared is positive for all countries, implying an eventually increasing unplanned outage rate. I find the same phenomenon for the additive intensity of two Weibull intensities (C) and, for Belgium and Switzerland, of Gompertz and Weibull intensities (D): the outage rate decreases following the start up from refueling, but increases eventually. As expected from the plot of empirical intensities (figure 25), the simple Weibull model (B) shows a negative duration dependence. Deciding between different parametric forms is difficult since all specifications describe the occurrence of unplanned outages well despite their apparent differences. To illustrate the results, figure 26 plots the estimated intensities for the four different specifications for two countries (A: solid line, B: long dashes, C: dots, D: short dashes); table 22 presents the parameter estimates for model A (standard errors in parentheses).

The plant manager's control decision

The only difficulty created by unobserved state variables is computational. But this is a rather serious problem because estimating the decision problem for a general statistical model for the distribution of these states is outside the range of today's computational possibilities. In two important papers, Rust (1987, 1988) showed that the assumption of conditional independence which restricts the dependence between unobserved states in different periods circumvents the computational difficulty. I will follow Rust's suggestion to model unobserved states. This assumption requires that the transition function factors as:

$$P(x',y'|x,y,a)=q(y'|x')\,p(x'|x,a) \tag{68}$$

The value function is a function of all states V(x,y). The expected value as a function of the action a is now of the form:

$$EV(x,y,a)=\sum_{x'}\int_{y} V(dy|x')\,p(x'|x,a) \tag{69}$$

and the optimal value function is the unique solution to the functional equation

$$V(x,y)=\max_{a}\,[R(x,y,a)+\beta EV(x,y,a)] \tag{70}$$

Let the vector of unobserved states y in any one period have a bivariate extreme value distribution with mean 0 ($\mu\approx0.577$) and density

$$q(y_1,y_2)=\prod_{i=1}^{2}\exp(-y_i+\mu)\exp(-\exp(-y_i+\mu)) \qquad (71)$$

If the part of the reward function R(x,y,a) = R(x) + g(y,a) corresponding to the unobserved states is

$$g(y,a)=\begin{cases} y_1 & \text{if the plant is shut down} \\ y_2 & \textit{otherwise} \end{cases} \qquad (72)$$

the conditional probability of choosing action a_n, $Q(a_n|H_n)$, has the relatively simple form

$$Q(a|H)=Q(a|x)=\frac{\exp\{R(x,a)+\beta EV(x,a)\}}{\sum_{b\in A}\exp\{R(x,b)+\beta EV(x,b)\}} \qquad (73)$$

where EV is given as the unique solution to the functional equation[57]

$$EV(x,a)=\sum_{x'}\log(\sum_{b\in A}\exp[R(x',b)+\beta EV(x',b)])p(x'|x,a) \qquad (74)$$

For estimation we have to keep in mind that both R() and P() depend on parameters. The log-likelihood function is

$$l(\theta)=\sum_{n=1}^{N}\{\log Q(a_n|x_n,\theta)+\log p(x_n|x_{n-1},a_{n-1},\theta)\} \qquad (75)$$

where θ includes the parameters of the plant process, the plant manager's cost function, and the discount factor.

As mentioned before, I first estimate P() and then use the estimated parameter values as constants. In the second step, I

---

[57] In order to simplify notation, I left out the dependence on constant parameters.

maximize the function

$$l_1(\theta) = \sum_{n=1}^{N} \{\log Q(a_n|x_n,\theta)\} \tag{76}$$

a procedure which yields consistent estimates (Rust, 1987, 1988).

For all cost specifications, one period of refueling incurs a constant cost r and there is no cost if the plant is running without incidents. In the simplest model, the plant manager incurs a cost c for each unplanned outage.

a)

$$c(t) = c_0 \tag{77}$$

It may be desirable to relax the restrictions of constant costs since the relative costs of unplanned outages to refuel outages could change over the duration of the fuel cycle; for example, unplanned outage costs might be highest in the middle of the cycle. I use the following dynamic functional forms for the cost of an unplanned outage:

b)

$$c(t) = c_0 + c_1 t \tag{78}$$

c)

$$c(t) = c_0 + c_1 t + c_2 t^2 \tag{79}$$

d)

$$c(t) = c_0 + \exp(c_1 t) \tag{80}$$

Different failure modes (different types of unplanned outage) are likely to incur different costs. In particular, outages occurring during testing or due to replacement of substandard components might be less costly than outages due to failure of worn-out components. Unfortunately, there is no reliable data on failure modes since reporting conventions regarding the classification of outages are inconsistent. One of the major advantages of the additive intensity specification is that it can distinguish probabilistically between wear-out failures and "teething" problem following a start-up. The expected cost of an unplanned outage is the sum of the costs of all types of unplanned outages weighted by their probabilities.

e)

$$c(t)=\sum_{m} p_m(t)\, c_m \tag{81}$$

where m is the failure mode.

Table 23 summarizes the results for various cost functions and the exponentiated quadratic intensity plant model.

Allowing for a dynamic cost function improves the fit of the model significantly. Unfortunately, the data set is too small to permit reasonable precise estimates for these functions: estimates become unstable and standard errors increase dramatically for models with more than two parameters in the cost function[58]. The more sophisticated cost functions do not predict uniformly better

---

[58] For example, the standard error for the parameter estimate of the cost of refueling per period in Switzerland increases from 0.99 to 1832.97 when switching from cost specification a) to specification b). Similar, although somewhat smaller, increases occur for other countries and specifications.

than the simple specification a) with two parameters corresponding to a constant cost for each period of refueling and a different constant cost for each unplanned outage, although they fit the data better. In fact, because of their sensitivity, the cost functions b) through e) do strictly worse than cost function a) for at least one country when used to predict the unplanned outage rate and the unplanned unavailability factor. For that reason, and because some implications of the behavioral model are most easily discussed by considering the cost function with two parameters (a), I use the estimates from model specification Aa) in the following discussion. The results for this specification are reported in table 24 (estimates from the exponentiated quadratic polynomial in table 22 were used to calculate the transition probabilities). The standard errors assume a fixed known transition function. Although conceptually one could calculate the correct standard errors by maximizing the full likelihood instead of using the two-step procedure, the additional computational requirements are prohibitive. The nature of the model makes it also infeasible to obtain simple corrections. The estimates for Sweden did not converge. This may simply be due to the randomness of the data, but it may also reflect the restrictiveness of the functional form and the fact that Sweden uses a somewhat different technology; it is the only country in the sample using BWR rather than PWR.

I cannot estimate the discount factor precisely, a problem that appears to be quite common in this type of models (Rust, 1987, Ryu, 1990) and which is not surprising given the size of the data set. Even for the simplest cost function, the likelihood function is almost completely flat and standard errors increase in order of magnitude. Nevertheless, it is worth emphasizing that I was capable of estimating a discount factor for one country (France), whereas Rust (1987) and other researchers were not capable of estimating any discount factor for their models. Fixing the discount factor does not change the estimates or the value of the likelihood function to any substantial degree[59]. More important than leaving the discount factor unrestricted is how the exogenous process and the cost function are specified. The results reported in tables 23 and 24 correspond to a monthly discount factor of .99 (approximately 13% annually), but even an unreasonably low discount factor of .9 (approximately 250% annually) gives similar results.

4.4 Discussion

The behavioral model describing the distribution of planned refuel outages (of the transition from the up state to the refuel state) complements the analysis of chapter 3 on the distribution of unplannned outages over time. In order to simulate time paths in section 3.3, a descriptive statistical model (a gamma distribution) of the distribution of planned shutdowns was previously used.

---

[59] For example, for France and the model specification Aa, the changes in the estimated cost parameters between the model with the annual discount rate fixed at approximately 13% (monthly discount factor 0.99) and the model with estimated discount rate (approximately 3% annual) were less than 0.2. A 95% confidence interval for the discount rate ranges from 0% to 30%, and this is without taking into account the additional variation due to the fact that the exogenous process is estimated.

Although the behavioral model cannot improve much upon the descriptive gamma distribution in terms of fitting actual data (the simplest cost specification of the structural model has the same number of parameters as the descriptive gamma distribution), it provides additional insights about the economic environment that cannot be obtained by a descriptive statistical model. The interpretation of estimated parameters in an analysis of differences between countries is discussed in the next chapter. A further strong argument in favor of a structural model is that the actions of the plant manager are not random events, in contrast to the occurrences of equipment failures, which are unforeseen and stochastic. Therefore, the analysis of purposeful behavior proceeds differently than the analysis of mechanical systems.

A major advantage of the algorithmic approach taken here is that neither the form of the plant process nor of the utility (cost) function needed to be specified in advance. This was desirable since there was no a priori knowledge regarding functional forms. However, I had to specify a priori a functional form for the distribution of unobserved states. Although this constraint may be relaxed when much more powerful computers become available, it is arguably a major shortcoming of the model in its current form.

Table 21: Specifications for Plant Process (mean log-likelihood)

country	Belgium 63	France 311	Germany 58	Sweden 164	Switzer-land 30
Model A)	-1.5917	-1.4155	-1.4765	-1.5102	-1.3493
Model B)	-1.5940	-1.3836	-1.3889	-1.4354	-1.2442
Model C)	-1.5940	-1.3746	-1.3856	-1.4323	-1.2210
Model D)	-1.5883	-1.3654	-1.3256	-1.4112	-1.1516

Table 22: The Exponentiated Quadratic Intensity Model:

country	Belgium 63	France 311	Germany 58	Sweden 164	Switzer-land 30
mean log-likelihood	-1.5917	-1.4155	-1.4765	-1.5102	-1.3493
$\gamma 1$	-0.6370 (0.3254)	-0.0431 (0.1377)	-1.2477 (0.3709)	-0.2362 (0.1775)	-1.8709 (0.6607)
$\gamma 2$	-0.1983 (0.1305)	-0.2806 (0.0592)	-0.41478 (0.1905)	-0.2153 (0.0679)	-0.3473 (0.3492)
$\gamma 3$	0.0112 (0.0107)	0.01830 (0.0049)	0.02335 (0.0176)	0.0080 (0.0050)	0.0217 (0.0339)

Table 23: Mean Log-Likelihood for Different Specifications[60]

country N	Belgium 173	France 624	Germany 322	Sweden 386	Switzer- land 195
Model Aa)	-0.2152	-0.1297	-0.2207	*)	-0.1307
Model Ab)	-0.1339	-0.1228	-0.1677	-0.1676	-0.0582
Model Ac)	*)	-0.1223	-0.1673	-0.1672	*)
Model Ad)	-0.1478	-0.1278	-0.1689	-0.1674	-0.0385

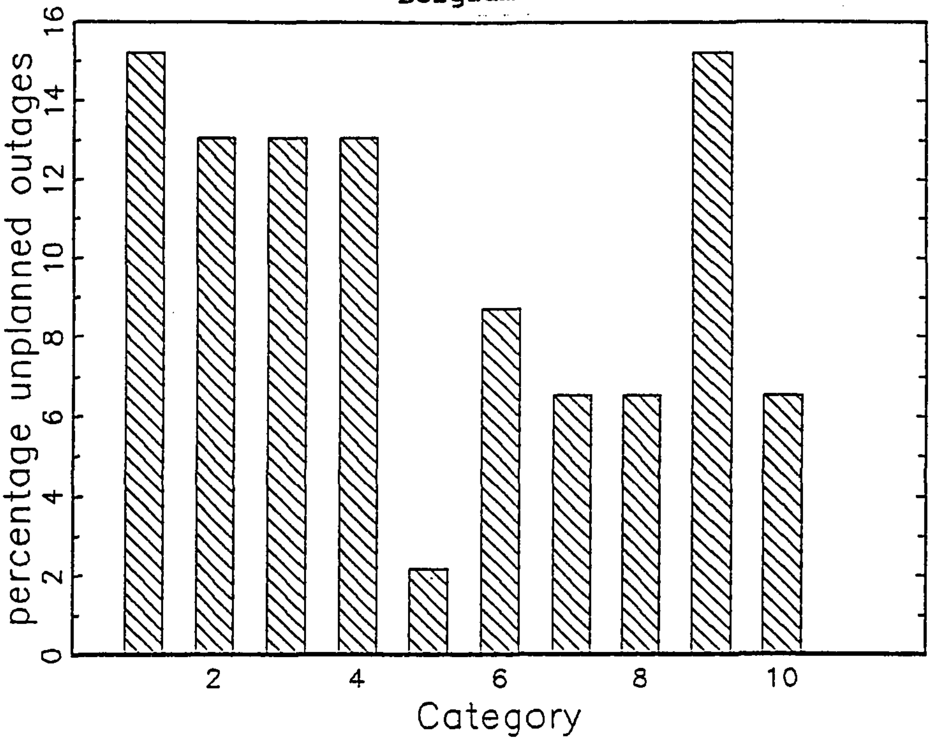
[60] N is the number of observed periods, the log-likelihood is N times the mean log-likelihood. Specifications that did not converge are denoted by *); the discount factor is fixed at 0.99.

Table 24: Cost Function Estimates with Model Specification Aa

country	Belgium 173	France 624	Germany 322	Switzer-land 195
mean log-likelihood	-0.2162	-0.1297	-0.2207	-0.1307
cost per refuel period (r)	2.5684 (0.4198)	5.4914 (0.6952)	1.5878 (0.3899)	1.6173 (0.9947)
cost per unplanned outage (c_0)	6.1012 (2.7313)	8.9677 (1.8802)	9.7133 (3.3425)	101.76 (36.92)

Figure 25: Empirical Failure Rate

Belgium



France

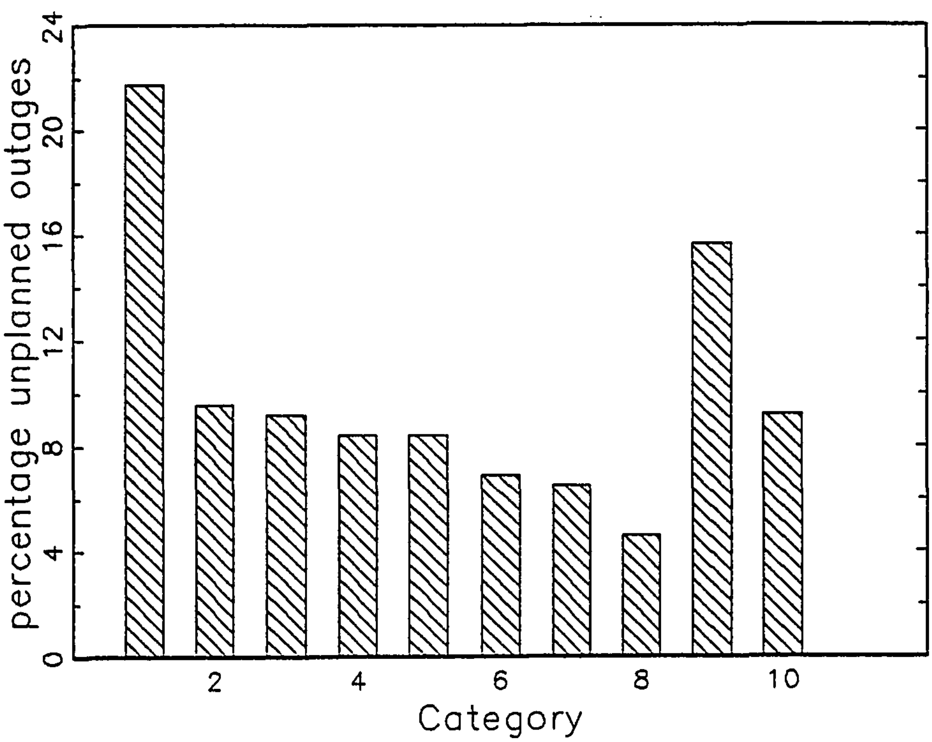
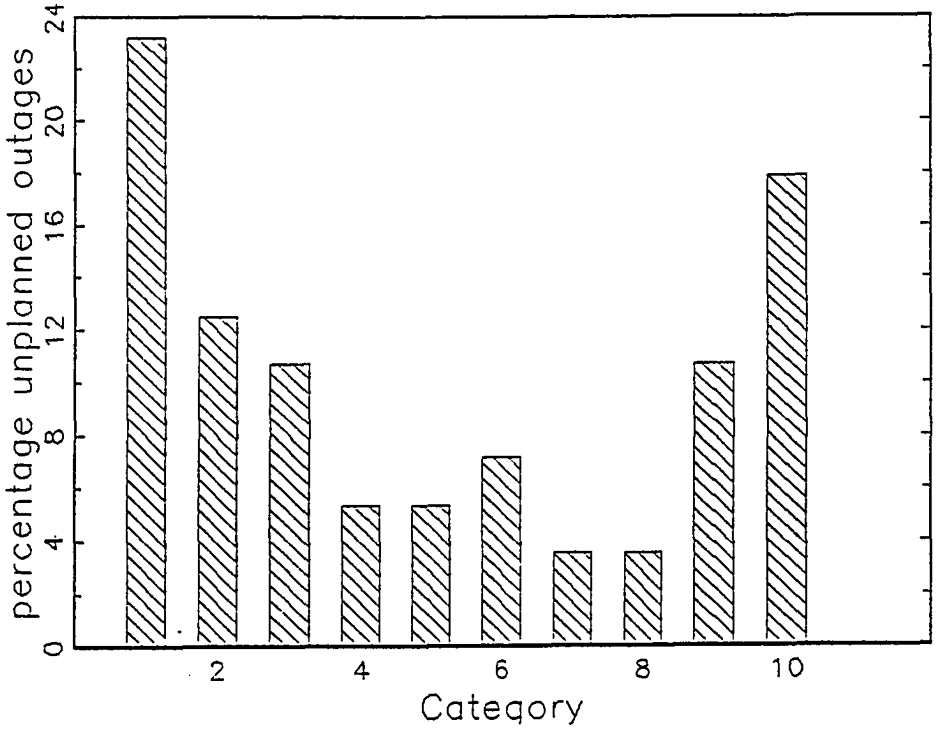


Figure 25 continued: Empirical Failure Rate

Germany



Sweden

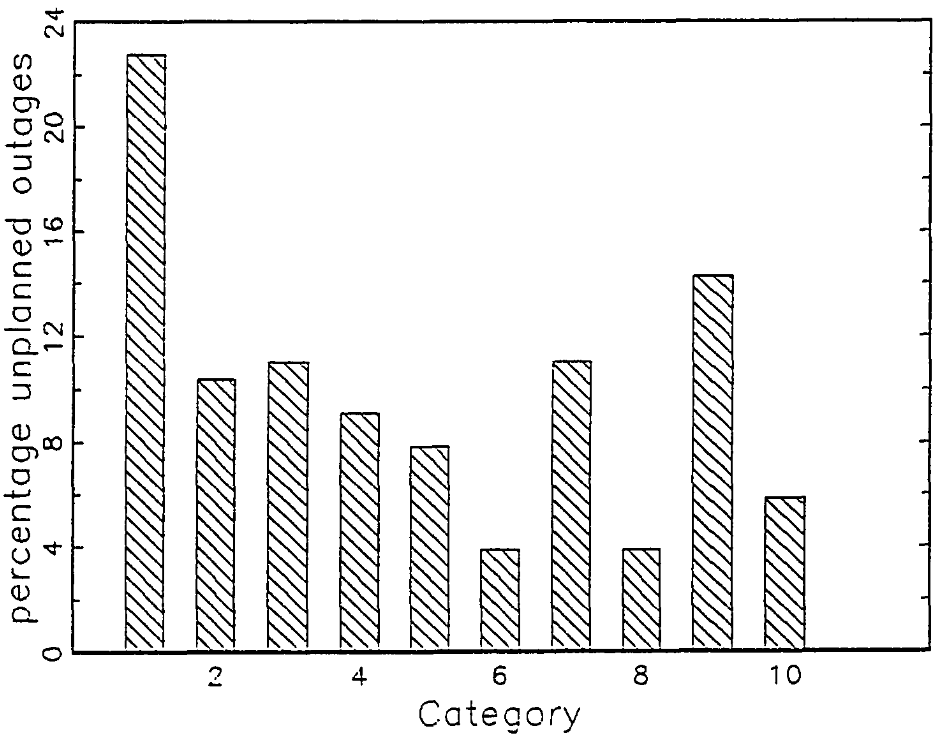


Figure 25 continued: Empirical Failure Rate

Switzerland

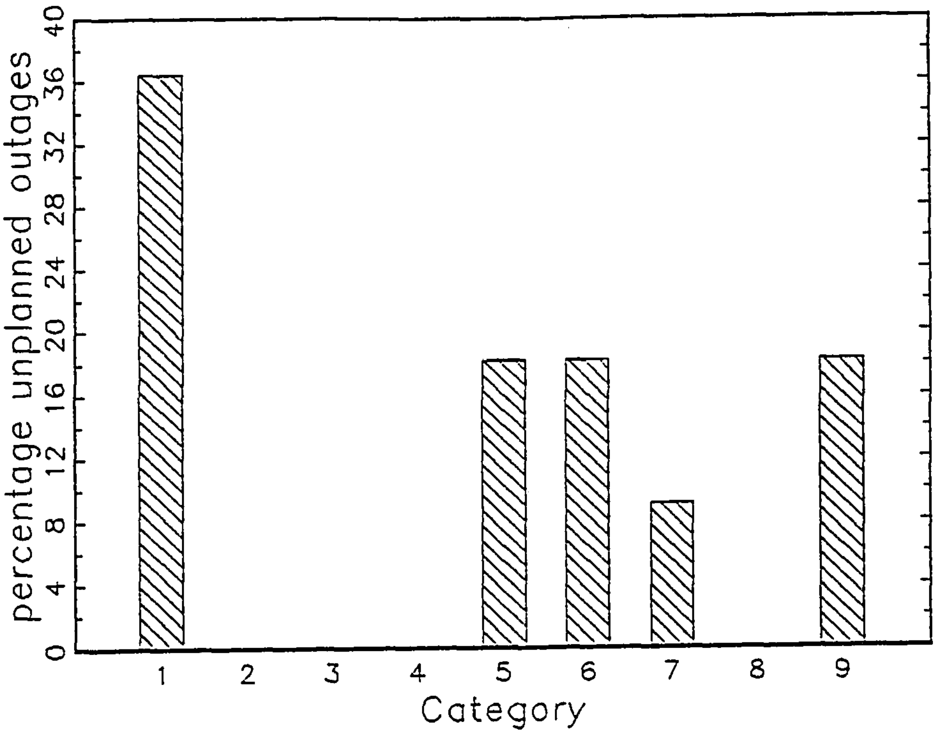
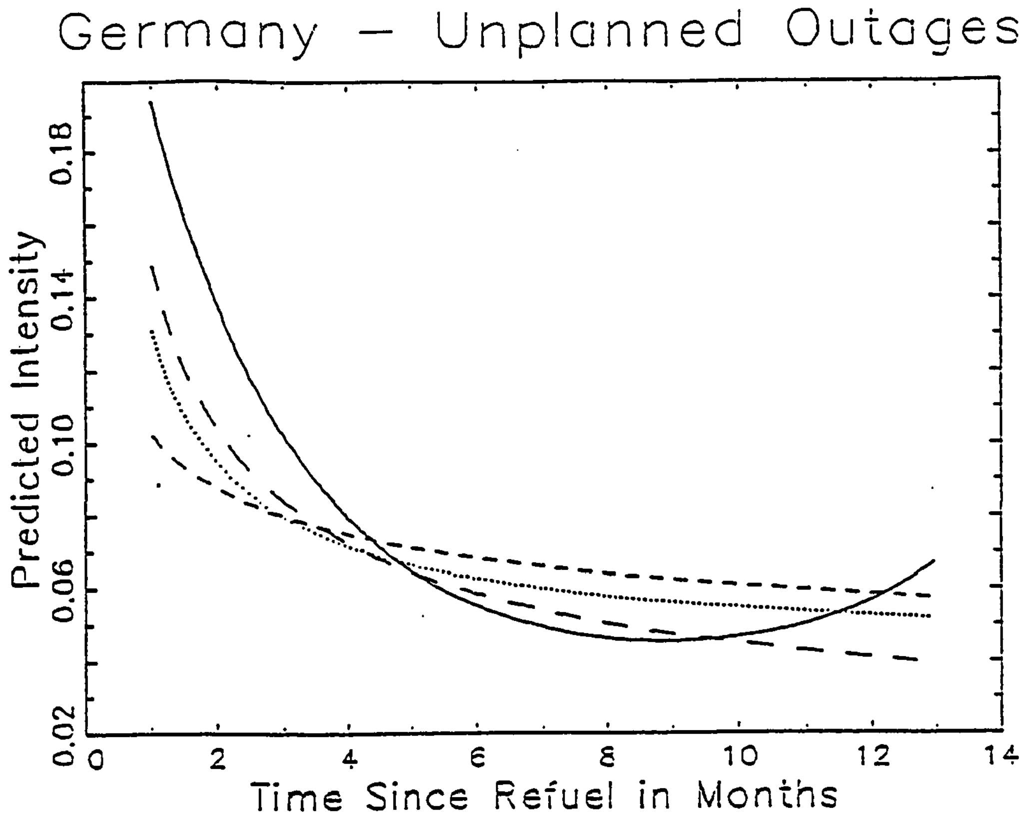
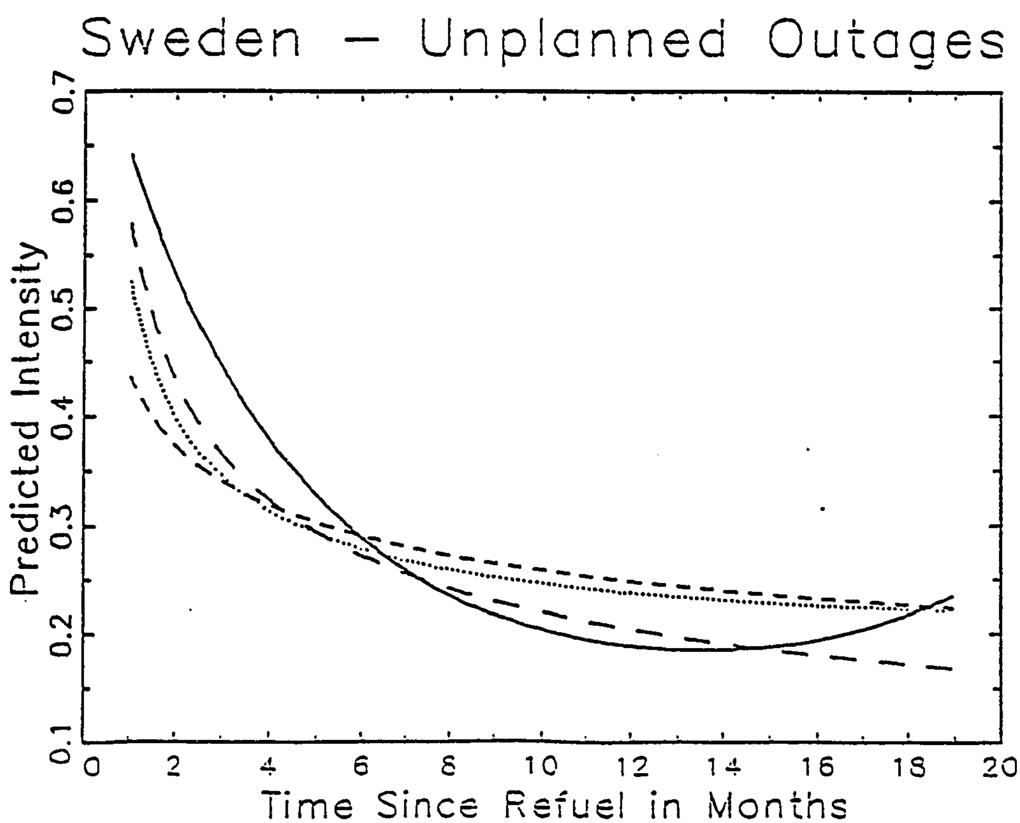


Figure 26: Predicted Failure Rate





5. Economic Interpretation

The differences in the performance between countries reflected in the descriptive statistics of section 4.1 have two sources. One source is different incentives for operations, reflected in the "cost" function parameters, which lead to different performance characteristics for a given technical plant process. This source, which has been ignored in the literature, has been explicitly modeled in chapter 4. However, the model requires that the other source, namely differences in the exogenous plant process between countries, is taken into account. Different countries' plants may differ in technological operating characteristics because of different abilities to manage technology, institutional constraints, licensing procedures, or other factors. For convenience, this broad collection of factors impinging on the "exogenous" plant process can be labeled "X-efficiency" factors.

In theory, one could attempt to formulate an economic model to describe the "choice" of a particular plant process as a function of the foregoing institutional constraints, technological capabilities, or the organizational and informational structure of the utility company operating the plant. For example, the unplanned outage intensity in the plant process is low in German nuclear generating stations (compare table 19) due to a required design concept which automatically invokes corrective measures at the early stage of a transient to prevent reaching a scram[61]. The adoption of the more reliable plant process in German reflects a legal requirement: the 1976 amendment to the German Atomic Law of 1960 requires all plants to embody best practice technology. The plant design in a country is determined to a substantial extent by licensing procedures which require the demonstration that the design is capable of handling certain assumed events such as a Loss-of-Coolant-Accident using prescribed models for calculations. These arguments lead me to believe that relative price differences (compared to differences in institutional constraints) could only play a minor role in explaining differences among the plant processes between countries.

The distinction between process control and choice of process is related to the distinction between operation and construction. Because I am concerned with plant operations in existing production facilities, I did not consider the more complex model, which can at best be done in the form of a qualitative case study. The current model demonstrates that overemphasis in the literature on observed availability or capacity factors as measures of productivity is misguided. The "productivity" or "inefficiency" measured by observed availability factors alone confuses the effects of economic differences (different cost functions) on the control of the plant and differences in the plant process being controlled.

---

[61] The "phased safety concept", see Morimoto (1986) for this and other reasons for international differences in scram rates.

The present analysis permits these to be disentangled under the maintained hypothesis as will be shown below.

The estimated cost (utility function) parameters capture both pecuniary and nonpecuniary "psychic" costs ("tastes", "attitudes", risk aversion) because they are estimated from observed behavior. The estimated parameters can be related to actual monetary values. There are two possible approaches. One could fix the scale factors[62] by a direct comparison between actual costs, which would show how actual tax payments or regulatory penalties affect the cost parameters. Another approach is to enlarge the state space formulation by including actual costs for different types of events as states. If the plant manager is cost minimizing, the residual should become smaller[63]. Both approaches would provide a partition into actual (pecuniary) costs and nonpecuniary costs which permits to test the strength of the economic theory: do plant managers minimize pecuniary costs, or does "taste" matter?

---

[62] The normalizations that affect the scale factors are the variance of the unobserved states and setting the cost of generating electricity without incidents for one unit to zero. Rust (1987) discusses the connection between actual costs and the normalized parameter estimates in more detail.

[63] In this formulation, a state is the cost or reward associated with a certain event (whereas in the current formulation it is the physical state). If the formulation is correct and the plant manager is cost minimizing, the estimated weights should be close to 1. A deviation shows that the plant manager puts more or less emphasis on the occurrence on certain events than a cost minimizing strategy.

5.1 Comparative statics - how do cost changes affect performance measures?

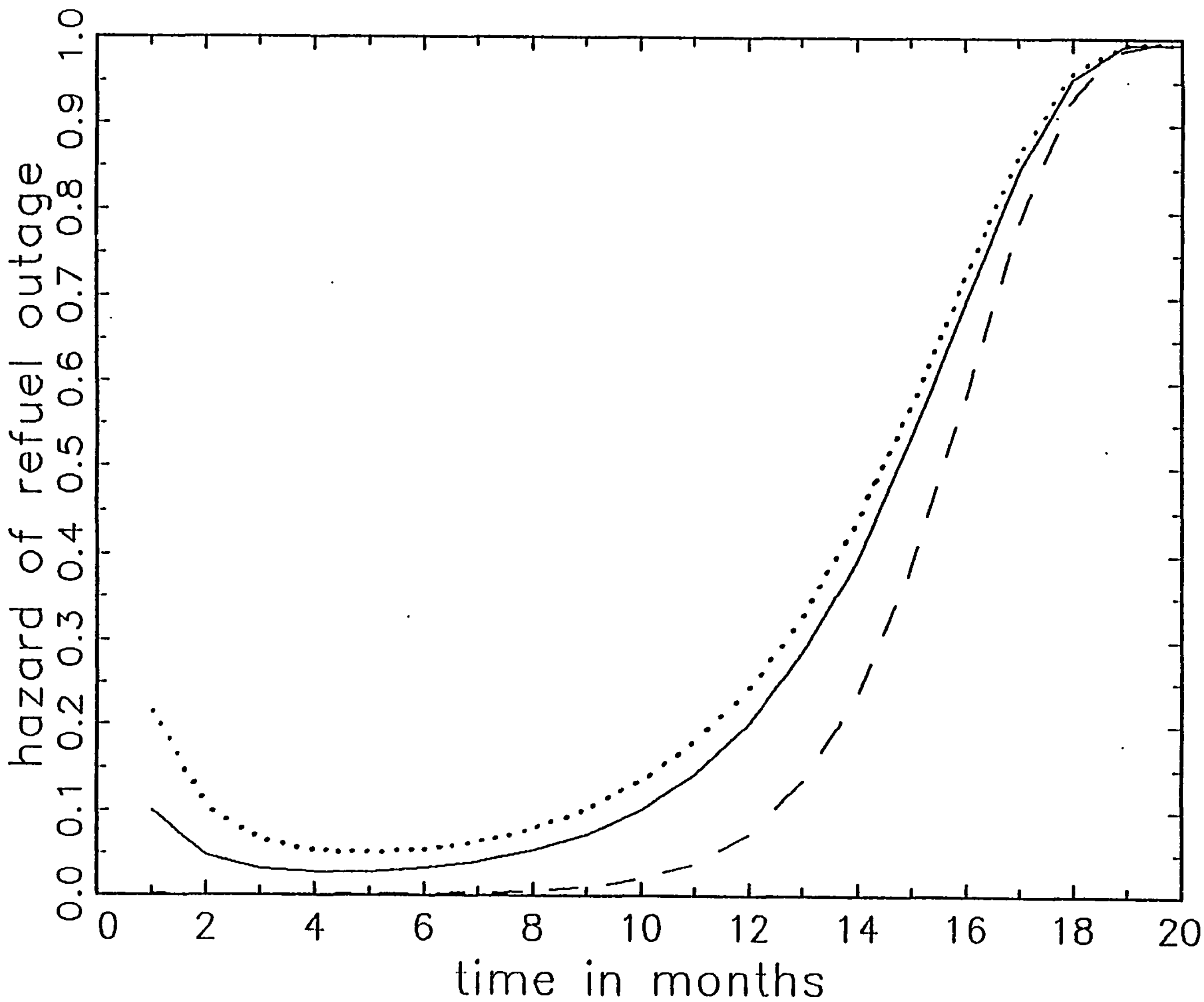
The comparative static results show the magnitude and type of changes in measures of different aspects of nuclear power performance that can be explained by the behavioral model. Different economic incentives for plant operations and their associated change in the parameters of the cost (utility) function change the plant manager's value function. This in turn alters the probability of observing a plant shutdown and has a direct effect on the operating cycle length, the unplanned outage rate, and the plant availability. Note that a descriptive statistical model of the distribution of refuel shutdowns cannot address such a question.

The effects of three different sets of cost parameters are illustrated for Germany and the model specification Aa (the two parameter cost function and an exponentiated quadratic polynomial for the unplanned outage intensity). The expected equilibrium values of the operating cycle length, the outage rate per year, and the unavailability due to refueling are shown in table 25 as cost function parameters vary from the estimated ones. The first set of values corresponds to the estimates, the next two columns correspond to hypothetical "policy experiments". In figure 27, which plots the hazard function of an observed refuel outage derived from solving the plant manager's maximization problem, the solid line corresponds to the estimated parameter values and the dashed and the dotted lines to the first and second "policy experiments". Figure 27 shows how the distribution of refuel shutdowns varies with changes in parameters. Relatively higher unplanned outage costs (or relatively lower refuel costs) lead to a lower unplanned outage rate and a higher proportion of unavailability due to planned outages. The economic model appears to capture the actual performance of the nuclear power industry; compare with the descriptive statistics in section 4.1 how well the structural model using estimates predicts the actual values in Germany. Even with this simple specification (model Aa), I found a fairly close agreement between model predictions and actual performance in the other countries. None of the more sophisticated model specifications allowing for dynamic changes in the costs that the plant manager faces provided uniformly better predictions, even though some of them predicted better for one or two countries.

Table 25: Comparative Statics

parameters	r=1.59 c=9.7	r=5 c=10	r=1 c=10
mean operating cycle length in months	10.5	15.0	8.1
Unavailability due to Planned Outages	0.15	0.11	0.19
Unplanned Outage Rate per reactor year	0.83	1.18	0.66

Figure 27: Hazard Function for Refuel Outages



5.2 Productivity in an International Perspective

Many descriptive studies have demonstrated the existence of international differences in the performance of nuclear power industries. But one can only determine which of these differences can be eliminated by changing an plant manager's incentives after explicitly modeling and estimating a structural behavioral model. France is often considered a successful user of the nuclear technology because of a relatively high availability factor, especially when compared to the U.S. However, the behavioral model estimated in the previous chapter reveals that the incentive structure in France in skewed in favor of a high availability (low planned unavailability) factor: the relative costs of unplanned outages (with short average durations) compared to planned refuel and maintenance outages (with long average durations) are the lowest of all countries.

I consider two aspects of performance: the planned unavailability factor and the unplanned outage (scram) rate. The assumption underlying the calculations is that the estimated parameters in different countries are directly comparable and that there is no important causal link between cost function parameters and differences in the plant process, i.e. that the differences unexplained by the behavioral model are "noise". This assumption essentially reflects the distinction between operating existing production facilities and constructing new ones. A causal link biasing the calculations could exist, for example, if the plant manager can choose between plant processes, i.e. if the plant process is not a fixed datum but a control variable, as in models that consider the adoption of a technology. I have taken Germany as the reference point for these calculations.

Table 26 reports the ratio between the performance measure for a given country and the corresponding one in Germany. A higher value means "lower productivity" and I therefore call these ratios "inefficiency" indices. Comparing the model predictions of the "inefficiency" indices under the estimated specification Aa with the actual values of these indices (calculated from tables 19 and 20) provides an informal assessment of the goodness-of-fit of the model. The predictions of the unplanned outage rate and planned unavailability match the actual values well, with the exception of the German-French comparison of unavailability. In terms of planned unavailability, France and Belgium appear to be comparable to Germany, but in terms of unplanned outages, they appear to be doing substantially worse. Switzerland is performing best with respect to both performance measures.

The indices of table 26 include both the differences in productivity as measured by the exogenous plant process and the impact of different economic environments. There are two possible ways to split economic and technical effects: one can ask what happens if plant managers in a country face different incentives while controlling the same plants as before (the approach taken in the previous section) or one can ask what happens if plant managers have to control a different process while the incentives structure remains constant. The first approach considers the effect of economic changes, the second approach considers the effect of technical changes. Here, the second approach is taken and the economic environment in a country is held constant in calculating the "inefficiency" indices of table 27. For a comparison with table 27, the predicted values in table 26 should be used. The second and fourth columns contain the values of the performance measure predicted from the estimated model specification Aa. The first number corresponds to the country's estimated plant process and its "cost" function, the number in parentheses corresponds to the same "cost" function, but under the adoption of the German plant process. The indices in the third and fifth columns divides the first number by the number in parentheses and answers the question: after taking into account that we face different incentives than German plant managers, how much worse (or better) are the unavailability factor and the unplanned outage rate than they could be if the technical process were as efficient as in Germany?

The index for both measures is larger than 1 for France and Belgium, which means that both countries could improve their performance uniformly (i.e. with respect to both aspects of performance) without changing the incentive structure but by adopting a technical process as efficient as in Germany (or Switzerland). In terms of planned unavailability, the model predicts that slightly over one week per reactor year is lost due to technical inefficiency in France and Belgium compared to Germany. Similarly, about 3.5 and 1.5 unplanned outages, respectively, can be attributed to this source. Switzerland appears to be the most productive country and the argument is reversed in this case: adopting the German production process, but controlling it according to the Swiss incentive structure, would cause about 1 additional week per reactor year lost in planned outages and about 0.4 additional unplanned outages in Switzerland.

Table 26: Productivity Calculations: Economic and Technical Effects

	"Inefficiency" Indices			
	Planned Un-availability (predicted)	Planned Un-availability (observed)	Unplanned Outage Rate (predicted)	Unplanned Outage Rate (observed)
Belgium	1.00	0.94	2.95	2.94
France	0.87	1.06	5.67	5.08
Germany	1.00	1.00	1.00	1.00
Switzer-land	0.63	0.75	0.74	0.67

Table 27: Productivity Calculations, Technical (X-efficiency) Effects Only

	Planned Un-availability	"Ineffi-ciency" Index	Unplanned Outage Rate	"Ineffi-ciency" Index
Belgium	0.15 (0.13)	1.15	2.46 (0.99)	2.48
France	0.13 (0.11)	1.19	4.73 (1.23)	3.85
Germany	0.15	1.00	0.83	1.00
Switzer-land	0.10 (0.12)	0.83	0.62 (1.01)	0.61

6. Summary and Conclusion

This dissertation analyzes the operating performance of nuclear power plants in five European countries using panel data on individual reactors. While the duration of a single "up" or "down" spell might be analyzed with standard duration models, the sequence of up and down spells cannot be analyzed in this fashion. One reason for the failure of duration models is their implicit renewal assumption and their focus on spell time: the argument in the hazard function is reset to zero at the beginning of the spell. The multiple spell models in economics analyzing sequences of unemployment/employment spells or fertility histories are built on single spell duration models.

The process formulation suggested in this dissertation generalizes the hazard function approach of duration models beyond the first event and is a natural formulation for complex systems, which are generally built to be repaired rather than replaced (renewed) upon failure. The time argument in the hazard or intensity function is measured from a constant point for several events, such as the first day a plant came on line (plant time)[64]. A particular feature of nuclear power plant operations are cyclical patterns due to major maintenance and refuel outages and I found fuel cycles to be particularly useful units of analysis. When each fuel cycle is treated as a system, time is measured as cycle time from the startup from the last refuel outage.

---

[64] "Time" need not be measured in hours or days, it may very well be a measured in units of output, or units of usage.

According to a nonparametric rank test, cycle time (but not spell time) by itself accounts for much variation found in the data (the hypothesis that the first up spell in each fuel cycle has the same distribution was not rejected). Of course, no plant is completely "renewed" during refueling; many components are the same and cannot be changed during the lifetime of the plant (the concrete containment structure being the most salient example). Why does the cycle model work for the data? The main reason, I believe, is that the properties of "middle age" plants change relatively slowly with plant age, compared to the length of a fuel cycle or the length of the planning horizon of a utility company or plant manager.

As the current stock of plants ages, it may become important to analyze long run (30-50 year) changes in plant reliability. This question has only recently been addressed and there are very few results and even less data available (IAEA, 1988). Should we be interested in new plants, plant age might be important because of learning effects which contributed to the findings in the reliability growth model of section 2.4. There I abstracted from the nonstationarity in the fuel cycle to analyze changes in plant reliability over several years. In other words, the process of interest may depend on several time scales simultaneously. One might even want to add an additional one, calendar time, which is an important measure of changes exogenous to the process analyzed, such as changes in regulation or overall technical improvements. For my data set, which spans a relatively short time period (6 years) of relatively "quiet" years (between the two major nuclear accidents) for nuclear power (1981-1986), however, calendar time is not particularly interesting to analyze.

Due to the almost exclusive focus on spell time in the literature, statistical models are not well developed to deal with multiple time scales. The only way in which multiple time scales have been used is in a proportional hazard model with spell time as the only argument of the hazard function and calendar time and plant time (age) as time varying covariates. For many purposes, this can be adequate, but for questions such as the failure of complex systems where some components are replaced regularly and others remain the same over a long period of time, this may be problematic. My empirical attempts at modeling multiple time scales simultaneously by using additive intensity models with arguments corresponding to different time scales were not successful. The statistical problem is very similar to the estimation of mixtures: a relatively large number of parameters has to be estimated and the likelihood function becomes very flat in certain regions. Nevertheless, this may be an interesting problem to be investigated in larger data sets, such as maintenance data on aircraft, automobiles.

The nonstationarity caused by learning or long run aging without any periodic resets poses difficult (and unsolved) questions for behavioral models. For dynamic optimization models, the most important technique for empirical work, the state space explodes, making everything but very short horizon models impossible. Only one very special model has a known solution, the multi-armed bandit problem, and Miller (1984) estimates a model of employment choice in this framework. I plan to investigate the effects of learning on productivity and reliability in a behavioral model in future research.

Although stationary in the sense that there are periodic returns to one state, the behavioral model based on the statistical results of chapters 2 and 3 introduces several extensions to microeconomic models of replacement/maintenance problems[65], which take essential features of this application into account. In particular, it is necessary to consider more than one maintenance action and to allow for maintenance actions with durations that are not negligible. Although there are models in the operations research literature[66] addressing the first point ("minimal" repair, "imperfect" repair), the second point, which is of crucial importance in any practical application, has not been acknowledged before. It also may be necessary to consider that the costs for different actions change over time. Unfortunately, the data set in this application may have been too small to evaluate the benefits of the more general specifications: although the dynamic cost functions substantially improved the fit over the model with

---

[65] Three such studies are Rust (1987), Kennet (1990), and Ryu (1990).

[66] The literature on maintenance and repair models is comprehensively reviewed in Pierskalla and Voelker (1976) and Valdez-Flores and Feldman (1989).

constant costs, they did not provide uniformly better predictions.

The information of the econometrician is limited compared to the information of the operator and the dynamic program that needs to be solved in the structural model is therefore related to the work on partially observable Markov decision models (POMDP), surveyed in Monahan (1982). The essential difference between that line of research and the present paper is that the econometrician ignores essential information in this application, not the operator. Given the well-known computational problems for partially observable models, I opted for a simple statistical specification for the unobservable component. Although the specification of the utility function and the exogenous process was guided by statistical information, unobservables were assumed be distributed according to an iid extreme value distribution. Although this particular specification is very standard and is generally used without any justification--how many applied papers justify their use of a logit model?--I would have preferred to be able to specify a more general model, especially allowing serial correlation in the unobserved components. Unfortunately, the practical usefulness of the model depends on being able to solve the maximization problem analytically and to express the functional equation for the expected value function (eq. 74) without requiring numerical integration. It is a trivial exercise to write down the general model (eq. 69 and 70), but a numerical solution is impossible without access to a supercomputer even when a very coarse discretization for the space of unobservables is used (say approximate the real line by 100 points). Future work may attempt to integrate some results from the research on POMDP. As of now, the algorithms developed for POMDP are very special and cannot be used for models as complex as the one considered here.

The structure of behavioral models, which are specified at the level of "primitive" parameters is invariant to broad classes of policy changes, whereas the structure of "reduced" form models depends on the current policy implementation (the Lucas critique). While models in operations research and the behavioral model consider an action such as repair or maintenance as a fixed basic unit, in reality there is a substantial range of how maintenance is performed. Thus the model cannot consider policy changes that would affect the type of maintenance performed and it is a "reduced" form model in that sense. For example, imagine the case that new regulations drastically limit radiation exposure. One would expect the utility company to substitute capital (robots) for human labor and since the previous situation was optimal, the relative cost for maintenance compared to repair has risen (assuming that more people are exposed for a longer time during maintenance/refueling than during repairs). However, the model predictions given the observed change in relative costs are incorrect because it fails to take into account the "deeper" aspects of changes in the type of maintenance.

Within its range of applicability, he model is capable of distinguishing the effects of economic incentives (the relative costs of a fixed type of repair and maintenance) and technical ("X-efficiency") factors on various aspects of nuclear power plant performance empirically. In particular, the results indicate that France and Belgium's relatively good showing in an international comparison with respect to one performance measure, namely availability, is to some extent due to the incentive structure (the estimated "cost" function) that the operators in these two countries face. The comparative static calculations for Germany illustrate that changing the incentives for operations can have a substantial impact on all aspects of plant performance. Alternatively, by keeping the incentive structure constant and by varying the technical plant process, one obtains a measure of technical efficiency. This provides a better assessment of a country's success with using a technology than the direct comparison of observed performance measures.

Appendix 1

The International Atomic Energy Agency publishes yearly reports on the individual performance of nuclear power plants in member states. I have extracted duration data for commercial nuclear power plants in several European countries (Belgium, France, Germany, Sweden, Switzerland) and Canada. Before 1981, only the outage duration and the cause is provided, not the date of the outage. From 1981 onwards, the date of the outage is reported as well.

Table 28 gives an example of the data set, taken from unit 1 of the Belgian Doel complex. The entries are explained in table 29. This data structure allows to distinguish 28 potentially different states (RUN, U-A, P-A, X-A, U-B, ..., U-K, P-K, X-K), but many of these states do not occur or have too few occurrences to be useful (eg. X-A, P-A). There are some inconsistencies between reporting practices and I have recoded numerous outages after referring to the description of the outage in the IAEA reports and trade journals. In particular, several French plants reported failures of safety equipment as an exogenous or regulated outage, instead of an outage due to equipment failure as all other plants did. Even less reliable or consistent is the information about the subsystem or components involved and I have decided not to use this information.

Information was available on 107 different plants, although information on several plants was incomplete or incorrect. I did not use any data that appeared questionable. The two German Philippsburg units, for example, never reported an outage even though it was clear from production data that refuel outages occurred. The list of plants is given in table 30. The second column gives the type of plant. Five different technologies are represented in the data set: Pressurized Light Water Reactors (PWR), Boiling Water Reactors (BWR), Heavy Water Reactors (PHWR) (all of which are Canadian reactors), Gas Cooled Reactors (GCR) (which represent the first generation of French reactors), and Fast Breeders (FBR) (only data on Phenix is available, Super Phenix and Kalkar had not started commercial operation by the end of 1986). Column 3 contains the electrical generating capacity in MW (which may change during operation within a small range). The dates in column 4 and 5 and 6 are of the form (yymmdd) and correspond to the date of first commercial production and start of plant construction. Column 7 and 8 correspond to the operator and the constructor of the plant. This information was taken from the IAEA 1986 book.

Table 28: Excerpt from data for Belgian plant Doel 1

DOEL1	BE	53592.0	1068.0	0	P	F	C	810328
DOEL1	BE	54660.0	7476.0	1	X	X	X	0
DOEL1	BE	62136.0	617.0	0	P	F	C	820319
DOEL1	BE	62753.0	1591.0	1	X	X	X	0
DOEL1	BE	64344.0	16.0	0	U	F	A	820619
DOEL1	BE	64360.0	1832.0	1	X	X	X	0
DOEL1	BE	66192.0	36.0	0	U	F	A	820904
DOEL1	BE	66228.0	4332.0	1	X	X	X	0
DOEL1	BE	70560.0	1199.0	0	P	F	C	830305
DOEL1	BE	71759.0	337.0	1	X	X	X	0
DOEL1	BE	72096.0	19.0	0	U	F	K	830508
DOEL1	BE	72115.0	461.0	1	X	X	X	0
DOEL1	BE	72576.0	42.0	0	U	F	D	830528
DOEL1	BE	72618.0	102.0	1	X	X	X	0
DOEL1	BE	72720.0	12.0	0	U	F	D	830603
DOEL1	BE	72732.0	492.0	1	X	X	X	0
DOEL1	BE	73224.0	101.0	0	U	F	D	830624
DOEL1	BE	73325.0	1123.0	1	X	X	X	0
DOEL1	BE	74448.0	49.0	0	U	F	D	830814
DOEL1	BE	74497.0	1367.0	1	X	X	X	0
DOEL1	BE	75864.0	17.0	0	U	F	D	831012
DOEL1	BE	75881.0	2887.0	1	X	X	X	0
DOEL1	BE	78768.0	707.0	0	P	F	C	840210
DOEL1	BE	79475.0	157.0	1	X	X	X	0
DOEL1	BE	79632.0	21.0	0	U	F	A	840317
DOEL1	BE	79653.0	1491.0	1	X	X	X	0
DOEL1	BE	81144.0	32.0	0	U	F	A	840519
DOEL1	BE	81176.0	2320.0	1	X	X	X	0
DOEL1	BE	83496.0	10.0	0	U	F	A	840825
DOEL1	BE	83506.0	3350.0	1	X	X	X	0
DOEL1	BE	86856.0	527.0	0	P	F	C	850112
DOEL1	BE	87383.0	313.0	1	X	X	X	0
DOEL1	BE	87696.0	49.0	0	U	F	A	850216

Table 29: Code Explanation

Column 1: Name of the plant

Column 2: Country
- BE Belgium
- CAN Canada
- D Germany
- F France
- S Sweden
- CH Switzerland

Column 3: Age of the plant in hours since the day of commercial operation

Column 4: Length of Duration

Column 5:
- 0 down
- 1 up

Column 6:
- P planned outage
- U unplanned outage
- X outage reason exogenous to the plant or run

Column 7:
- F full outage
- P partial outage (only full outages were retained)
- X run

Column 8: outage cause
- A: Equipment failure
- B: Operator error
- C: Refueling and maintenance
- D: Scheduled inspection, maintenance, repair (no refuelling)
- F: Training
- G: Stretch-out, Coast-down
- H: Regulatory
- J: Grid Failure
- K: Others
- X: Run

Column 9: day of outage (yymmdd), 0: run

Table 30 Nuclear Power Plants

PLANT	TYPE	CAP	COMMOP	CRITICAL	CONSTR	OPER	CONTRACTOR
Belgium							
DOEL 1	PWR	392	750215	740700	690600	EBES	ACECOWEN
DOEL 2	PWR	392	751201	750804	710900	EBES	ACECOWEN
DOEL 3	PWR	897	821001	820614	750100	EBES	FRAMACEC
DOEL 4	PWR	1006	850901	850331	781200	EBES	ACECOWEN
TIHANGE 1	PWR	870	751001	750221	690900	INTC	FRAM
TIHANGE 2	PWR	902	830300	821005	750600	INTC	FRAMACEC
TIHANGE 3	PWR	1006	850901	850605	771200	INTC	ACECOWEN
Canada							
BRUCE 1	PHWR	740	770901	761217	720100	OH	AECL
BRUCE 2	PHWR	740	770901	760727	701200	OH	AECL
BRUCE 3	PHWR	740	780201	771128	730200	OH	AECL
BRUCE 4	PHWR	740	790118	781210	740200	OH	AECL
BRUCE 5	PHWR	750	850301	841115	780600	OH	AECL
BRUCE 6	PHWR	750	840914	840529	780200	OH	AECL
BRUCE 7	PHWR	750	860410	860107	790400	OH	AECL
DOUGLAS POINT	PHWR	220	680900		600200	OH	AECL
GENTILLY 2	PHWR	645	831001	820911	740400	HQ	AECL
PICKERING 1	PHWR	508	710721	710225	650600	OH	AECL
PICKERING 2	PHWR	508	711230	710915	650600	OH	AECL
PICKERING 3	PHWR	508	720601	720424	660600	OH	AECL
PICKERING 4	PHWR	508	730617	730516	660600	OH	AECL
PICKERING 5	PHWR	516	830510	821023	740700	OH	AECL
PICKERING 6	PHWR	516	840201	831015	740700	OH	AECL
PICKERING 7	PHWR	516	850101	841022	740700	OH	AECL
PICKERING 8	PHWR	516	860228	851217	750700	OH	OH/AECL
POINT LEPREAU	PHWR	630	830100	820725	750500	NBEPC	AECL
France							
BLAYAIS 1	PWR	920	811201	810520	770100	EDF	FRAM
BLAYAIS 2	PWR	910	830201	820728	770100	EDF	FRAM
BLAYAIS 3	PWR	910	831114	830729	780400	EDF	FRAM
BLAYAIS 4	PWR	910	831001	830501	780400	EDF	FRAM
BUGEY 1	GCR	540	720800	720321	680400	EDF	VARIOUS
BUGEY 2	PWR	925	790301	780420	721199	EDF	FRAM
BUGEY 3	PWR	925	790301	780831	730300	EDF	FRAM
BUGEY 4	PWR	900	790701	790217	740500	EDF	FRAM
BUGEY 5	PWR	900	800103	790715	740600	EDF	FRAM
CATTENOM 1	PWR	1360		861024	791000	EDF	FRAM
CHINON A2	GCR	170	650308		580100	EDF	VARIOUS
CHINON A3	GCR	480	670815	660301	610000	EDF	VARIOUS
CHINON B1	PWR	870	840100	821028	770300	EDF	FRAM
CHINON B2	PWR	870	840801	830923	770300	EDF	FRAM
CHINON B3	PWR	921		860918	801000	EDF	FRAM
CHOOZ	PWR	266	670415	661019	620100	SENA	ACEC
CRUAS 1	PWR	880	840402	830402	780700	EDF	FRAM
CRUAS 2	PWR	880	850401	840801	781100	EDF	FRAM

CRUAS 3	PWR	917	840910	840409	790500	EDF	FRAM
CRUAS 4	PWR	917	850211	841001	791000	EDF	FRAM
DAMPIERRE 1	PWR	890	800910	800315	750200	EDF	FRAM
DAMPIERRE 2	PWR	890	810216	801205	750400	EDF	FRAM
DAMPIERRE 3	PWR	890	810521	810125	750900	EDF	FRAM
DAMPIERRE 4	PWR	890	811120	810805	751200	EDF	FRAM
FESSENHEIM 1	PWR	890	771230	770307	710800	EDF	FRAM
FESSENHEIM 2	PWR	890	780318	770627	711100	EDF	FRAM
FLAMANVILLE 1	PWR	1363	861201	850929	791100	EDF	FRAM
FLAMANVILLE 2	PWR	1363		860612	800600	EDF	FRAM
GRAVELINES B1	PWR	910	801201	800221	750200	EDF	FRAM
GRAVELINES B2	PWR	910	801201	800802	750300	EDF	FRAM
GRAVELINES B3	PWR	910	810601	801130	751200	EDF	FRAM
GRAVELINES B4	PWR	910	811001	810531	760400	EDF	FRAM
GRAVELINES C5	PWR	964	850115	840805	791000	EDF	FRAM
GRAVELINES C6	PWR	963	851025	850721	791000	EDF	FRAM
PALUEL 1	PWR	1363	851201	850513	770800	EDF	FRAM
PALUEL 2	PWR	1363	851201	840811	780300	EDF	FRAM
PALUEL 3	PWR	1363	860201	850807	781200	EDF	FRAM
PALUEL 4	PWR	1363	860601	860329	791200	EDF	FRAM
PHENIX	FBR	233	740714	730831	681100	CEA	CEA/TECH
ST. ALBAN 1	PWR	1363	860501	850804	790300	EDF	FRAM
ST. ALBAN 2	PWR	1363		860607	790800	EDF	FRAM
ST. LAURENT A1	GCR	390	690815	690106	630800	EDF	VARIOUS
ST. LAURENT A2	GCR	450	710815	710704	660100	EDF	VARIOUS
ST. LAURENT B1	PWR	880	830803	810104	760400	EDF	FRAM
ST. LAURENT B2	PWR	880	830803	810512	760700	EDF	FRAM
SUPER PHENIX	FBR	1200		850907	761200	EDF	NOVATOME
TRICASTIN 1	PWR	915	801201	800221	741100	EDF	FRAM
TRICASTIN 2	PWR	915	801201	800722	741100	EDF	FRAM
TRICASTIN 3	PWR	915	810511	801129	750400	EDF	FRAM
TRICASTIN 4	PWR	915	811001	810531	750500	EDF	FRAM
Germany							
BIBLIS A	PWR	1146	750226	740716	700100	RWE	KWU
BIBLIS B	PWR	1178	770131	760325	720200	RWE	KWU
BROKDORF	PWR	1307	861222	861008	810401	KBR	KWU
BRUNSBUETTEL	BWR	770	770209	760622	700415	KKB	KWU
GRAFENRHEINFELD	PWR	1225	820616	811209	750100	BW	KWU
GROHNDE	PWR	1289	850201	840800	760600	KWG	KWU
GRUNDREMMINGEN B	BWR	1244	840719	840309	760720	KGB	KWU
GRUNDREMMINGEN C	BWR	1249	850118	841026	760720	KGB	KWU
ISAR 1	BWR	870	790321	771120	720200	KKI	KWU
KNK II	FBR	18	790303	771010	740900		
KRUEMMEL	BWR	1260	840328	830914	740100	KKK	KWU
MUEHLHEIM-KAERL	PWR	1227		860301	750116	RWE	BBR
NECKARWESTHEIM	PWR	805	761201	760526	710100	GKN	KWU
OBRIGHEIM	PWR	283	690330	680922	650300	KWO	SIEMENS
PHILIPPSBURG 1	BWR	864	800326	790309	701000	KKP	KWU
PHILIPPSBURG 2	BWR	1268	850418	841213	770707	KKP	KWU
STADE	PWR	630	720519	720108	671200	KKS	SIEMENS
THTR-300	HTGR	296	870616	830913	710500	HKG	HRB

UNTERWESER	PWR	1230	790906	780916	720700	KKU	KWU
WUERGASSEN	BWR	640	751111	711022	680100	PE	AEG
Sweden							
BARSEBAECK 1	BWR	570	750700	750118	710200	SYDK	ASEA
BARSEBAECK 2	BWR	570	770900	770220	730100	SYDK	ASEA
FORSMARK 1	BWR	890	801210	800423	730600	SSPB	ASEA
FORSMARK 2	BWR	890	810701	801116	750100	SSPB	ASEA
FORSMARK 3	BWR	1050	850818	841028	790100	SSPB	ASEA
OSKARSHAMN 1	BWR	440	720200	701212	660800	OKG	ASEA
OSKARSHAMN 2	BWR	580	750100	740306	690900	OKG	ASEA
OSKARSHAMN 3	BWR	1050	850815	841229	800500	OKG	ASEA
RINGHALS 1	BWR	760	760101	741014	690200	SSPB	ASEA
RINGHALS 2	PWR	820	750500	740619	701000	SSPB	WEST
RINGHALS 3	PWR	915	810909	800729	710900	SSPB	WEST
RINGHALS 4	PWR	915	831100	820519	731100	SSPB	WEST
Switzerland							
BEZNAU 1	PWR	350	690900	690630	650900	NOK	WEST/BBC
BEZNAU 2	PWR	350	711200	711016	680100	NOK	WEST
GOESGEN	PWR	920	791100	790120	731200	KKG	KWU
LEIBSTADT	BWR	942	841215	840309	740100	KKL	GE
MUEHLEBERG	BWR	306	721106	710308	670300	BWK	GETSCO

Appendix 2

This appendix collects a number of results on less well known techniques used in chapters 2 and 3.

Cubic spline smoothing

The polynomial smoothing spline is well suited for a statistical model of the following type

$$y_j = f(t_j) + e_j \tag{82}$$

where we observe y_j and t_j. The function of interest f() is unknown. The error terms are independent with mean zero and variance

$$V(e_j) = w_j^2 \sigma^2 \tag{83}$$

This is a typical problem in curve fitting approximation theory. The simplest class of functions for this purpose is the class of polynomial functions. Unfortunately, high order polynomials are only of limited usefulness because the necessary flexibility for complicated functions causes severe oscillations and the notoriously ill-conditioned Vandermondian matrix presents additional numerical problems. Dividing the region for which an approximation is desired into several subregions and fitting different polynomials for each subregion avoids both of these problems. If the approximating function g(x) satisfies certain continuity restrictions on its derivatives ($g(x) \varepsilon C^{m-1}$), the piecewise polynomials belong to the spline space $S^m(K_n)$, where K_n is a system of knots (or mesh) with

$$a=t_0<t_1<t2,\ldots,<t_n=b \tag{84}$$

The cubic spline smoother, in its current form due to Schoenberg (1964) and Reinsch (1967), is the function that minimizes

$$SP(g)=\sum_j [y_j-g(t_j)]^2+\alpha\int_t [g''(x)]^2dx \tag{85}$$

It is calculated in the following way. Let $\Delta_j=t_{j+1}-t_j$. The continuity and boundary conditions imply:

$$\begin{gathered} c_1=c_n=0 \\ d_n=0 \\ d_j=\frac{1}{3}\left(\frac{c_{j+1}}{\Delta_j}-\frac{c_j}{\Delta_j}\right) \\ b_j=\frac{1}{3}\Delta_{j-1}c_{j-1}+\frac{2}{3}(\Delta_j+\Delta_{j-1})c_j+\frac{1}{3}\Delta_j C_{j+1} \end{gathered} \tag{86}$$

Q is an N+2 x N matrix, R an N x N matrix, with elements

$$q_{j,j}=\frac{1}{\Delta_j}$$

$$q_{j+1,j}=-\frac{1}{\Delta_j}-\frac{1}{\Delta_{j+1}}$$

$$q_{j+2,j}=\frac{1}{\Delta_{j+1}} \tag{87}$$

$$r_{j,j}=\frac{2}{3}(\Delta_j+\Delta_{j+1})$$

$$r_{j+1,j}=r_{j,j+1}=\frac{1}{3}\Delta_{j+1}$$

Two equations for a and c determine the smoothing spline. Including a diagonal weighting matrix W containing the inverse of the standard deviation of the error terms, these equations are

$$a=y-\alpha W^2 Qc$$

$$(\alpha Q'W^2Q+R)\,c=B_\alpha c=Q'y \tag{88}$$

For more details on the cubic spline and recent references see Silverman (1984, 1985) or Wahba (1975,1983).

Kernel density estimation

Estimating a continuous density function is a well studied topic in statistics and a general reference for results on density estimation is Silverman (1986). The best understood method for estimating a density function nonparametrically is kernel estimation. A kernel estimator of a density function is defined as

$$\hat{f}(y) = \frac{1}{Sh}\sum_{i=1}^{S} K\left(\frac{y-X_i}{h}\right) \tag{89}$$

where K is the kernel function, h the window width, and X_i the i'th simulated observation. Provided that the kernel function K(.) is non-negative and that

$$\int_{-\infty}^{\infty} K(t)\,dt = 1 \tag{90}$$

f^(.) is a probability density function. In addition, f^(.) inherits all the smoothness properties of K().

A large number of kernel estimates have been suggested in the literature. Among the most efficient kernels (Silverman, 1986 ,p.43) are the Epanechnikov

$$K(t)\begin{cases} =0.75(1-0.2t^2)/\sqrt{5} & -\sqrt{5}<t<\sqrt{5} \\ =0 & otherwise \end{cases} \tag{91}$$

and the biweight kernel.

$$K(t)\begin{cases} =\frac{15}{16}(1-t^2)^2 & if -1<t<1 \\ = 0 & otherwise \end{cases} \tag{92}$$

Another easy computable function is the parabolic kernel

$$K(t)\begin{cases} =0.75(1-t^2) & \textit{if } -1\le t\le 1 \\ =0 & \textit{otherwise} \end{cases} \qquad (93)$$

The problem of choosing the smoothing parameter is of major importance. With reference to a normal distribution and a Gaussian kernel, it has been established that the optimal value of the window width h which minimizes the mean integrated square error is proportional to

$$h=1.06\sigma S^{-1/5} \qquad (94)$$

If the underlying distribution is bimodal, this window width would oversmooth the data. Several techniques for choosing the window width automatically have been suggested, the best known being cross-validation. Density estimates are used in dissertation mainly for an exploratory data analysis and I therefore prefer choosing the smoothing parameter by plotting out several curves. Examining several different plots also gives more insight into the data than a single curve with an automatically chosen bandwidth.

Simulating nonstationary Poisson processes and random variate generation

The stationary Poisson process with intensity λ can easily be simulated by adding iid exponential variates which are generated by inversion, U is a uniform [0,1] variate

$$T=\frac{-\ln(U)}{\lambda} \tag{95}$$

The easiest way to generate a time dependent process is by redefining the time scale

$$\tau=\int_0^T \lambda(u)\,du \tag{96}$$

Events occur are generated according to a homogeneous Poisson process ($\lambda=1$) and times are then inverted. Unfortunately, this does not work for many complicated intensities for which the inversion is impossible, such as the exponentiated quadratic polynomial. In this case, I use the following thinning algorithm. Prospective events are generated by a stationary Poisson process with $\lambda^*>\lambda(t)$ and an event at time t is accepted with probability $\lambda(t)/\lambda^*$. Although the efficiency of generating data is not very high, it is still substantially better than the cumbersome approximation by a piecewise uniform function. Other suggestions can be found in Dagpunar (1988).

The gamma distribution has been used for many purposes in this dissertation. Depending on the problem, I have used one of the following two parametrizations. They are completely equivalent, $\alpha=1/\beta$, but numerically it may be desirable to choose the version with the parameter to be larger than 1.

$$f(x)=\frac{\alpha^{r}}{\Gamma(r)}x^{r-1}e^{-\alpha x} \quad or$$
$$f(x)=\frac{\beta^{-\rho}}{\Gamma(\rho)}x^{\rho-1}e^{-\frac{x}{\beta}} \qquad (97)$$

The generalized gamma distribution is parametrized as follows:

$$f(x)=\frac{\delta}{\beta^{\rho\delta}\Gamma(\rho)}x^{\rho\delta-1}\exp\left[-\left(\frac{x}{\beta}\right)^{\delta}\right] \qquad (98)$$

There are many algorithms to generate gamma deviates, see Dagpunar (1988), Devroye (1986). Depending on α, I use three different algorithms. For $\alpha=1$, I use the inversion method for exponential variates, for $\alpha>1.0$ envelope rejection is used, generating a prospective variate by rescaling a T-variate which can be obtained by inversion, for $\alpha<1.0$ envelope rejection is used again and the target distribution is a probabilistic mixture of different distributions covering mutually exclusive regions. My implementation corresponds to the algorithms G4 and G6 in Dagpunar (1988).

References

Amemiya, T. (1985) Advanced Econometrics, Cambridge: Harvard University Press

Anderson, S.C. (1987) "An international comparison of the impact of safety regulation on LWR performance," MIT-EL 87-006, Center for Energy Policy Research, MIT

Arora, Ashish and Alfonso Gambardella (1990) "Acquiring Technology: External Linkages and Internal Knowledge in Biotechnology," MS, Stanford University

Arrow, K.J., 1962, 'The economic implication of learning by doing', Review of Economic Studies, 29, pp. 155-173

Arthur, W.B. (1989) "Competing Technologies, Increasing Returns, and Lock-in by Historical Events," Economic Journal, 99, 116-131

Ascher, H. and H. Feingold (1984) Repairable Systems Reliability: Modeling, Inference, Misconceptions and Their Causes, New York: Marcel Dekker

Aven, Terje (1983) "Optimal Replacement under a Minimal Repair Strategy - A General Failure Model," Advances in Applied Probability, 15, 198-211

Bain, I.J. and M. Engelhardt (1980) "Inferences on the Parameters and Current System Reliability for a Time Truncated Weibull Process," Technometrics, 22, 421-426

Barlow, R.E. and L. Hunter (1960) "Optimum Preventive Maintenance Policies," Operations Research, 8, 90-100

Barlow, R.E., D.J. Bartholomew, J.M Bremner, and H.D. Brunk (1972) Statistical Inference under Order Restrictions, New York: Wiley

Barlow, R.E. and F. Proschan (1975) Statistical Theory of Reliability and Life Testing, New York: Holt, Rinehart and Winston

Beckjord, E., M. Golay, E. Gyftopoulos, K. Hansen, R. Lester, and D.K. Winje (1987), "International Comparison of LWR Performance," MIT-EL 87-004, Center for Energy Policy Research, MIT

Bellman, R. (1961) Adaptive Control Processes: A Guided Tour, Princeton: Princeton University Press

Berkovec, James and Steven Stern (1991) "Job Exit Behavior of Older Men," Econometrica, 59, 189-210

Berry, D. A. and B. Fristedt (1985), Bandit Problems: Sequential

Allocation of Experiments, London: Chapman & Hall

Bertsekas, D. and S. Shreve (1978) Stochastic Optimal Control: The Discrete Time Case, New York: Academic Press

Blackwell, D. (1965) "Discounted Dynamic Programming," Annals of Mathematical Statistics, 36, 226-235

Breslow, N.E. (1970) "A Generalized Kruskal-Wallis Test for Comparing K Samples Subject to Unequal Patterns of Censorship," Biometrika, 57, 579-594

Brettschuh, W. (1988) "Measures Designed to Reduce Refueling and Maintenance Outages and BWR and PWR Nuclear Power Plants," in: Nuclear Power Performance and Safety (IAEA-CN-48/137), Vol.2, 95-109

Brown, Mark and Frank Proschan (1983) "Imperfect Repair," Journal of Applied Probability, 20, 851-859

Cattaneo, F. S. de Sactis, H. Kalfsbeek, G. Mancini, L Marciano, G. Volta (1988) "Analysis of Operational Experience in Nuclear Power Plants by the European Reliability Data System _ERDS," in: Nuclear Power Performance and Safety (IAEA-CN-48/137), Vol.2, 213-230

Cinlar, E. (1972) "Superposition of Point Processes," in: Stochastic Point Processes, P.A.W. Lewis (ed.), 549-606, New York: Wiley

Cowan, Robin (1987) Backing the Wrong Horse: Sequential Technology Choice Under Increasing Returns, Ph.D. dissertation, Stanford University

Cox, D.R. (1959) "The Analysis of Exponentially Distributed Life-Times With Two Types of Failure," Journal of the Royal Statistical Society, B, 21, 411-421

Cox, D.R. (1962) Renewal Theory, London: Methuen

Cox, D.R.(1975) "Partial Likelihood," Biometrika, 62, 269-276

Cox, D.R. and D.V. Hinkley (1974) Theoretical Statistics, London: Chapman & Hall

Cox, D.R. and P.A.W. Lewis (1978) The Statistical Analysis of Series of Events, London: Chapman and Hall (reprinted)

Cox, D.R. and V. Isham (1980) Point Processes, London: Chapman & Hall

Cox, D.R. and D. Oakes (1984) Analysis of Survival Data, London: Chapman and Hall

Crow, L.H. (1982) "Confidence Interval Procedures for the Weibull Process with Applications to Reliability Growth," Technometrics, 24, 67-72

Dagpunar, John (1988) Principles of Random Variate Generation, Oxford: Clarendon Press

Danko, J.C. and K.E. Stahlkopf (1982) Status of Research on Pipe Cracking in BWR, Nuclear Safety, 23, 653-664

Das, Sanghamitra (1989) "An Estimable Discrete Choice Dynamic Model of Capital Utilization and Retirement," MS, Indiana University

David, P.A. (1985) "Clio and the Economics of QWERTY," American Economic Review, 75, 332-337

David, P.A., T.A. Mroz, and K.W. Wachter (1985) "Rational Strategies of Birth-Spacing and Fertility Regulation in Rural France during the Ancien Regime," Working Paper #14, Stanford Project on the History of Fertility Control, Stanford University

David, P.A. (1988) "Path Dependence: Putting the Past into the Future of Economics," Technical Report #533, IMSSS, Stanford University

David, P.A., T. Amemiya, K. Arrow, T. Bresnahan, E. Pate-Cornell, T.J. Connolly, G. Rothwell, W.E. Steinmueller (1988) "Informational and Organizational Impacts on Productivity: The Economics of Control and Reliability in Complex Production Processes", Center for Economic Policy Research, Stanford University

David, P.A. (1988b) "The Simple Economics of Reliability, Control, and Capacity Utilization," in: David et. al. (1988), Appendix B

David, P.A. and T. Mroz (1989) "Evidence of Fertility Regulation Among Rural French Villagers, 1749-1789", European Journal of Population, 5, 1-26 (part 1), 173-206 (part 2)

David, P.A., G. Rothwell, and R. Maude-Griffin (1990) "Survival Distribution Functions for Spells of Operation in U.S. Nuclear Power Plants," IOIP Research Memorandum, Stanford University

Denardo, E.V. (1967) "Contraction Mappings in the Theory Underlying Dynamic Programming," SIAM Review, 9, 165-177

Derman, C. (1970) Finite State Markovian Decision Processes, New YorK: Academic Press

Deutsche Risikostudie Kernkraftwerke (1980), Fachband 8, Koln: Verlag TUV Rheinland

Devroye, L. (1986) Non-Uniform Random Variate Generation, New York: Springer

Dewanji, Anup, and John D. Kalbfleisch (1987) "Estimation of Sojourn Time Distributions for Cyclic Semi-Markov Processes in Equilibrium," Biometrika, 74, 281-288

Diggle, Peter J. and Richard J. Gratton (1984) "Monte Carlo Methods of Inference for Implicit Statistical Models," Journal of the Royal Statistical Society, B, ,193-227

van Dijk, N.M. (1984) Controlled Markov Processes: Time-Discretization, CWI Tract 11, Centrum voor Wiskunde en Informatica, Amsterdam

Duane, J.J. (1964) "Learning Curve Approach to Reliability Modeling," IEEE Trans. Aerospace, 2, 563-566

Dynkin, E.B. and A.A. Yushkevich (1979), Controlled Markov Processes, Berlin: Springer

Easley, David and Kiefer, Nicholas M. (1988) "Controlling A Stochastic Process With Unknown Parameters", Econometrica, 56, 1045-1064

Easley, David and Kiefer, Nicholas M. (1989) "Optimal Learning with Endogenous Data," International Economic Review, 30, 963-978

Electric Power Research Institute (1986) Nuclear Unit Operating Experience - 1983 and 1984 Update, EPRI Report NP-4368

Electric Power Research Institute (1987) The Influence of Fuel-Cycle Duration on Nuclear Unit Performance, EPRI Report NP-5042

Electric Power Research Institute (1989) The Influence of Fuel-Cycle Duration on Nuclear Unit Performance - An Update, EPRI Report NP-6333

Fafchamps, Marcel (1990) "Estimation of an Intertemporal Expected Utility Model without Closed-Form Solution," MS, Stanford University

Frees, E.W. (1988) "Correction: Nonparametric Renewal Function Estimation," Annals of Statistics, 16, 1741

Fullwood, R.R. and R.E. Hall (1988) Probabilistic Risk Assessment in the Nuclear Power Industry, New York: Pergamon

Galambos, J. (1978) The Asymptotic Theory of Extreme Order Statistics, New York: Wiley

Gehan, E.A. (1965) "A Generalized Wilcoxon Test for Comparing Arbitrarily Sinly-Censored Samples, Biometrika, 52, 203-223

Gihman, I.I. and A.V. Skorohod (1979) Controlled Stochastic Processes, New York: Springer

Gittens, John C. (1989) Multi-armed Bandit Allocation Indices, Chichester: Wiley

Glaser, R.E. (1980) "Bathtub and Related Failure Rate Characterizations," Journal of the American Statistical Association, 75, 667-672

Gnedenko, B.V., Yu Belyayev, and A.D. Solovyev (1969) Mathematical Methods in Reliability Theory, New York: Academic Press

Gotz, Glenn A. and John J. McCall (1984) "A dynamic retention model for Air Force officers," Rand Publication R-3028-AF

Hansen, L. and K. Singleton (1982) "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," Econometrica, 50, 1269-1286

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches (1984) "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," Econometrica, 52, 909-938

Heckman, James J. and Burton Singer (1984a) "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," Econometrica, 52, 271-320

Heckman, James J. and Burton Singer (1984b) "The Identifiability of the Proportional Hazard Model," Review of Economic Studies, 60, 231-243

Heckman, James J. and Burton S. Singer (1985) "Social Science Duration Analysis," in: Longitudinal Analysis of Labor Market Data (James J. Heckman and Burton S. Singer ed.), New York: Cambridge University Press, 39-110

Heckman, James J. and J.R. Walker (1987) "Using Goodness of Fit and Other Criteria to Choose Among Competing Duration Models: A Case Study of Hutterite Data," in: Sociological Methodology, C.Clogg (ed.), 18, 247-307, American Sociology Association, New York

van Hee, K.M. (1978) Bayesian Control of Markov Chains, Mathematisch Centrum, Amsterdam, Tract 95

Heyman, D.P. and M.J. Sobel (1984) Stochastic Models in Operations Research, 2 vols., New York: McGraw-Hill

Hirsh, Richard F. (1989) Technology and Transformation in the American Electric Utility Industry, Cambridge: Cambridge University Press

Hjorth, J. (1980) "A Reliability Distribution with Increasing, Decreasing, Constant, and Bathtub-Shaped Failure Rates," Technometrics, 27, 99-108

Hoel, Paul G., Sydney C. Port, and Charles J. Stone (1972) Introduction to Stochastic Processes, Houghton Mifflin

Hulkower, S.D. (1986) "The effects of regulation on the performance of nuclear power in the United States and the Federal Republic of Germany," MIT-EL 86-008, Center for Energy Policy Research, MIT

International Atomic Energy Agency (1988) Safety Aspects of the Ageing and Maintenance of Nuclear Power Plants, SM-295

International Atomic Energy Agency (1983) Reliability of Reactor Pressure Components, Vienna

International Atomic Energy Agency (various years) Operating Experience with Nuclear Power Stations in Member States, Vienna

Jelinski, Z. and P.B. Moranda (1972) "Software Reliability Research," in: Statistical Computer Performance Evaluation (W. Freiberger, ed.), Academic Press, 465-484

Jewell, W.S. (1984) "A General Framework for Learning Curve Reliability Growth Models," Operations Research, 32, 547-558

Joskow, P. and Rozanski, G.A. (1979) "The Effects of Learning By Doing On Nuclear Power Plant Operating Reliability", Review of Economics and Statistics, 61, 161-168

Kalbfleisch, J.D. and R.L. Prentice (1980) The Statistical Analysis of Failure Time Data, New York: Wiley

Kendall, M.G. and A. Stuart (1979) The Advanced Theory of Statistics, 4th edition

Kennet, D. Mark (1990) "Airline Deregulation and Aircraft Engine Maintenance: An Empirical Policy Analysis," MS, Tulane University

Kirk, J. and J.R. Harrison (1987) "The Approach to Safety for Sizewell B," Nuclear Energy, 26-3 (June)

Krautman, A.C. and J.L. Solow (1990) "Regulation and Capacity Utilization: Evidence from Nuclear Power Generation," MS

Lagakos, S.W., C.J. Sommer, and M. Zelen (1978) "Semi-Markov Models for Partially Censored Data," Biometrika, 65, 311-317

Lancaster, T. (1979) "Econometric Methods for the Duration of Unemployment," Econometrica, 47, 939-956

L'Ecuyer, Pierre and Alain Haurie (1987) "The Repair vs. Replacement Problem: A Stochastic Control Approach," Optimal Control Applications & Methods, 8, 219-230

Lawless, J.F. (1987) "Regression Methods for Poisson Process Data," Journal of the American Statistical Association, 82, 808-815

Leclercq, Jacques (1986) "Getting Improved Performance from Electricite de France's PWRS," Nuclear Engineering International, July, 34-37

Lester, R.K. and M.J. McCabe (1988) "The Effect of Industrial Structure on Learning by Using in Nuclear Power Plant Operation, MIT-EL-88-024WP, Center for Energy Policy Research, MIT

Lewis, E.E. (1987) Introduction to Reliability Engineering, New York: John Wiley

Lippman, S. (1975) "On Dynamic Programming with Unbounded Rewards," Management Science, 1225-1233

Littlewood, B. and J. Verrall (1973) "A Bayesian Reliability Growth Model for Computer Software," Applied Statistics, 22, 332-346

Littlewood, B. (1981) "Stochastic Reliability-Growth: A Model for Fault-Removal in Computer-Programs and Hardware-Designs," IEEE Transactions on Reliability, R-30 313-320

Lucas, R.E. (1976) "Econometric Policy Evaluation: A Critique," in K. Brunner and A.H. Meltzer (eds.), The Phillips Curve and Labor Markets, North Holland, Amsterdam

Manski, C.F. and D. McFadden (eds.) (1981) Structural Analysis of Discrete Data with Econometric Applications, Cambridge: MIT Press

McCormack, K.E. and R.B. Gallaher (1982) "Review of Safety-Related Events at Nuclear Power Plants in 1980, Nuclear Safety, 23, 320-328

McFadden, Daniel (1981) "Econometric Models of Probabilistic Choice," in C.F. Manski and D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications, 198-272, Cambridge: MIT Press

McFadden, Daniel (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," Econometrica, 57, 995-1026

Meyer-Abich, Klaus Michael and Bertram Schefold (1986) Die Grenzen der Atomwirtschaft, Munchen: C.H.Beck

Michaelis, Hans (1986) Handbuch der Kernenergie, Düsseldorf: ECON

Miller, R. (1984) "Job Matching and Occupational Choice," Journal of Political Economy, 92, 1086-1120

Monahan, G.E. (1982) "A Survey of Partially Observable Markov Decision Processes," Management Science, 28, 1-16

Montgomery, M. (1988) "Lifetime Fertility as a Controlled Stochastic Process: An Application of Rust's Estimation Method," Princeton University

Montgomery, M. and J. Trussell (1986) "Models of Marital Status and Childbearing," in: Handbook of Labor Economics, Vol. 1, Ch.3, edited by O. Ashenfelter and R. Layard, New York: North Holland

Morimoto, K. (1986) "Comparing International Scram Frequency Statistics," Nuclear Engineering International, September, 42-43

Morone, J.G and E.J. Woodhouse (1989) The Demise of Nuclear Power? New Haven: Yale University Press

Musa, John D. (1975) "A Theory of Software Reliability and Its Application," IEEE Transactions on Software Engineering, SE-1, 312-326

Muth, J.F. (1986) "Search Theory and the Manufacturing Progress Function," Management Science, 32, 948-962

Nuclear Regulatory Commission (1975) Reactor Safety Study: An Assessment of Accident Risks in Commercial Nuclear Power Plants, NUREG-75/014 (WASH-1400)

Nuclear Regulatory Commission (1983) PRA Procedures Guide: A Guide to the Performance of Probabilistic Risk Assessments for Nuclear Power Plants Review, NUREG-2300, (ANS/IEEE)

Nuclear Regulatory Commission (1985) Nuclear Plant Aging Research (NPAR) Program Plan, NUREG-1144

Olsen, R., D.A. Smith, and G. Farkas (1986) "Structural and Reduced-Form Models of Choice among Alternatives in Continuous Time: Youth Employment Under a Guaranteeed Jobs Program," Econometrica, 54, 375-394

Pakes, A. (1986) "Patents as Options: Some Estimates on the Values of Holding European Patent Stocks," Econometrica, 54, 755-785

Pakes, Ariel and David Pollard (1989) "Simulation and the Asymptotics of Optimization Estimators," Econometrica, 57, 1027-1058

Perrow, Charles (1984) Normal Accidents: Living with High-Risk Technologies, New York: Basic Books

Pershagen, Bengt (1989) Light Water Reactor Safety, Oxford: Pergamon Press

Phung, D.L. and W.B. Cottrell (1983) "Pressure Vessel Thermal Shock: Experience at U.S. Pressurized Reactors 1963-1981," Nuclear Safety, 24, 520-533

Pierskalla, W.P. and J.A. Voelker (1976) "A Survey of Maintenance Models: The Control and Surveillance of Deteriorating Systems," Naval Research Logistics Quarterly, 23, 353-388

Pollock, S.M. (1968) "A Bayesian Reliability Growth Model," IEEE Transactions in Reliability, R-17, 187-198

Prentice, R.L., B.J. Williams, and A.V. Peterson (1981) "On the Regression Analysis of Multivariate Failure Time Data," Biometrika, 68, 373-379

Prentice, R.L. (1978) "Linear Rank Tests with Right Censored Data," Biometrika, 65, 1 67-179

Press, W.H., B.P. Flannery, S.A. Teukolsky, W.T. Vetterling (1988) Numerical Recipes in C, The Art of Scientific Computing, Cambridge: Cambridge University Press

Proschan, F. (1963) "Theoretical Explanation of Observed Decreasing Failure Rate," Technometrics, 5, 375-383

Ramlau-Hansen, H. (1983) "Smoothing Counting Process Intensities by Means of Kernel Functions," Annals of Statistics, 453-466

Reinsch, C. (1967) "Smoothing by spline functions," Numerische Mathematik, 10, 177-183

Rothwell, G. (1989) "Stop and Start: A Duration Analysis of Nuclear Reactor Operations," Proceedings of the International Association of Energy Economists' Annual North American Conference, 309-317

Rothwell, G. (1991) "Utilization and Service: Decomposing Nuclear Reactor Capacity Factors," forthcoming Resources and Energy

Rothwell, G. and J. B. Jensen (1990) "Information Structures for Operating Nuclear Reactors", to be presented at the American Economics Association Annual Meeting, 1990

Rust, John (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," Econometrica, 999-1033

Rust, John (1988) "Maximum Likelihood Estimation of Discrete Control Processes," SIAM Journal of Control and Optimization, 26, 1006-1024

Ryu, Keungkwan (1990) A Structural Approach to Bivariate Duration Analysis, Ph.D. Dissertation, Dept. of Economics, Stanford University

Sargent, Thomas J., (1987), Dynamic Macroeconomic Theory, Cambridge: Harvard University Press

Sargent, Thomas J. (1979) "Interpreting Economic Time Series," Journal of Political Economy, 89, 213-248

Sargent, Thomas J. and N. Wallace (1979) "Rational Expectations and the Theory of Economic Policy," Journal of Monetary Economics, 2, 169-183

Schenk, H., E. Pickel, and H. Bartsch (1983) "Experience and Status of Components of the Primary System after 100 000 Hours of Operation," in: Reliability of Reactor Pressure Components, IEAE, 3-28

Schoenberg, I.J. (1964) "Spline Functions and the Problem of Graduation," Proc. Nat. Acad. Sci., 52, 947-950

Sheshinski, E. (1967) "Optimal Accumulation with Learning By Doing", in: K. Shell (ed.) Essays on the Theory of Optimal Economic Growth, Cambridge: MIT Press, 31-52

Shooman, M. (1972) "Probabilistic Models for Software Reliability and Prediction," in: Statistical Computer Performance Evaluation (W. Freiberger, ed.), Academic Press, 485-502

Silverman, B.S. (1982) "On the extimation of a probability density function by the maximum penalized likelihood method," Annals of Statistics, 12, 898-916

Silverman, B.W. (1985) "Some Aspects of the Spline Smoothing Approach to Non-parametric Regression Curve Fitting," Journal of the Royal Statistical Society, B, 47, 1-52

Silverman, B.W. (1986) Density Estimation for Statistics and Data Analysis, London: Chapman & Hall

Spence, A. Michael (1981) 'The Learning Curve and Competition', Bell Journal of Economics, 12, 49-70

Sturm, R. (1989b) "The Learning Curve and Adaptive Control of a Production Process with Unknown Parameters," IOIP Research Memorandum, Stanford University, presented at the Australasian meeting of the Econometric Society

Sturm, R. (1989a) "Can Measured Regulatory Losses Explain Differences in the Performance of the Nuclear Power Industry Between Countries?," IOIP Research Memorandum

Suzuki, T. and K.F. Hansen (1988) "A Comparison of Nuclear Safety Regulation Systems and their Impact on LWR Performance in Japan and the United States," MIT-EL-99-019WP, Center for Energy Policy Research, MIT

Thomas, S.D. (1988) The Realities of Nuclear Power, Cambridge: Cambridge University Press

Tsiatis, A. (1975) "A Nonidentifiability Aspect in the Problem of Competing Risks," Proc. Natl. Acad. Sci., 72, 20-22

Valdez-Florez, Ciriaco and Richard M. Feldman (1989) "A Survey of Preventive Maintenance Models for Stochastically Deteriorating Single-Unit Systems," Naval Research Logistics, 36, 419-446

Wahba, Grace (1975) "Smoothing Noisy Data with Spline Functions," Numerische Mathematik, 23, 383-393

Wahba, Grace (1983) "Bayesian Confidence Intervals for the Cross-Validated Smoothing Spline," Journal of the Royal Statistical Society, B, 45, 133-150

Wolpin, K. (1984) "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," Journal of Political Economy, 92, 1086-1120

Wolpin, K. (1987) "Estimating a Structural Search Model: the Transition from School to Work," Econometrica, 55, 801-817

Zimmerman, Martin B. (1982) "Learning Effects and the Commercialization of New Energy Technologies: The Case of Nuclear Power," Bell Journal of Economics, 13,