SD07

Time to Failure in UPS Data Collector and UPS Thermal Printer

Rebecca Dennison, University of Louisville, Louisville, KY

ABSTRACT

The purpose of this study is to model the effect of equipment age and model on the time to failure in "UPS Data Collector" and "UPS Thermal Printer" units repaired in 2003. The data were analyzed using a generalized linear model. The generalized linear model assumed a Poisson distribution with a log link function and an offset based on the natural log of the population. Life tables were calculated for the time to failure using equipment model and age as strata. Proportional hazards were also calculated using equipment age and model predicted time to failure.

The generalized linear model appeared statistically significant overall (p-value <0.001). The effects of model and equipment age also appeared statistically significant (p-value <0.001 for both). However, 17.6% of the units had not been returned for repair by 4/30/2006, the end of the study, undermining the validity of these results.

The life tables indicated that the strata were significant (Log-Rank and -2Log(LR) with p-value<0.0001 and Wilcoxon with p-value=0.0151). The proportional hazards model was statistically significant overall (p-value <0.001). The effect of the model was shown to be statistically significant (p-value=0.0009), though the effect of age was not (p-value=0.3251). The survival curves for "UPS Data Collector" were consistent with the time to failure predicted by the generalized linear model. The survival curves for the "UPS Thermal Printer" predict a much longer time between failures compared to the generalized linear model.

Further research would be required to determine the significance of the effect of age in equipment that experiences less severe censoring. Further study would also be required to determine if other thermal printers and data collectors experience similar survival rates. The effect of age might prove to be significant in these other models.

INTRODUCTION

This paper uses the generalized linear model, life tables, and proportional hazards to examine the effect of age and model number on the time to failure for the "UPS Data Collector" and "UPS Thermal Printer". The dataset to be used contains one year of repair information. The data were obtained from the UPS equipment repair database with the permission of the department that maintains the data. The specific model number of the equipment has been replaced with commodity class names to protect proprietary information. The "UPS Data Collector" is a data collector used to capture information and transfer that information via radio to the UPS network. The second model, "UPS Thermal Printer", is a thermal label printer used to print "smart" labels for packages to be shipped through the UPS system. The model was applied to units returned for repair from February 1, 2003 to January 31, 2004. A random sample of 250 units was drawn from the repairs of each model.

The generalized linear model assumes a Poisson distribution with a log link function offset by the natural log of the population. The overall model appeared to be statistically significant (p-value <0.001). The effect of model number and age also appeared to be statistically significant (p-value <0.001 for each). However, 17.6% of the units had not been returned for repair by 4/30/2006, the end of the study, undermining the validity of these results.

The life tables indicated that the age and model strata were significant (Log-Rank and -2Log(LR) with p-value<0.0001 and Wilcoxon with p-value=0.0151). The proportional hazards model was statistically significant overall (p-value <0.001). The effect of model was shown to be statistically significant (p-value=0.0009), though the effect of age was not (p-value=0.3251). The survival curves for "UPS Data Collector" were consistent with the time to failure predicted by the generalized linear model. The survival curves for the "UPS Thermal Printer", the model which experienced 92% of the censoring, predict a much longer time between failures compared to the generalized linear model.

The survival curves for the "UPS Thermal Printer" terminate at 35 months, the age of the oldest unit, without dropping below 30%. An additional study would need to be performed at a later date to approximate the complete survival curves. Further research would be required to determine the significance of the effect of age in equipment that experiences less severe censoring. Further study would also be required to determine if other thermal printers and data collectors experience similar survival rates. The effect of age might prove to be significant in these other models.

METHOD

Part number, serial number, the date of installation, and the date of equipment repair were obtained from the UPS equipment repair database. Installation began in September of 1998 for the "UPS Data Collector" and in September of 2001 for the "UPS Thermal Printer". A total of 77,499 "UPS Data Collectors" and 109,211 "UPS Thermal Printers" were in use on February 1, 2003. An additional 4,716 "UPS Data Collectors" and 93,217 "UPS Thermal Printers" were deployed by January 31, 2004. A total of 185,776 "UPS Data Collectors" and 69,127 "UPS Thermal Printers" were returned for repair between February 1, 2003 and January 31, 2004. Units returned more than once between February 1, 2003 and January 31, 2004 were considered duplicate returns. A random sample of 250 units was taken from the returns for each model. The resulting sample did not contain any duplicate returns.

Days in the field were calculated as the number of days between the initial return in the period between February 1, 2003 and January 31, 2004, and the first subsequent return prior to May 1, 2006. Months in the field were calculated by dividing days in the field by 30. The resulting number was rounded down to the nearest integer. Age was calculated by dividing the number of days between the installment and the initial return by 365. The resulting number was rounded down to the nearest integer.

Data visualization was performed using a box plot showing the months in the field by age for each model and a kernel density estimation comparing the months in the field of replacements by model. A generalized linear model was applied using the age and model of the equipment to predict the number of months between returns. The model assumed a Poisson distribution and used a log link function. An offset based on the natural log of the population of the given model at the time of the initial return was employed.

RESULTS

A generalized linear model was fitted to the return information using age and equipment model as predictors for the number of days until the next failure. The model assumed a Poisson distribution with a log link function offset by the natural log of the population.

Figures 1 and 2 display the distribution of months between returns by equipment age using box plots for the “UPS Data Collector” and “UPS Thermal Printer” respectively.

FIGURE 1. UPS DATA COLLECTOR MONTHS IN FIELD BY YEARS OLD

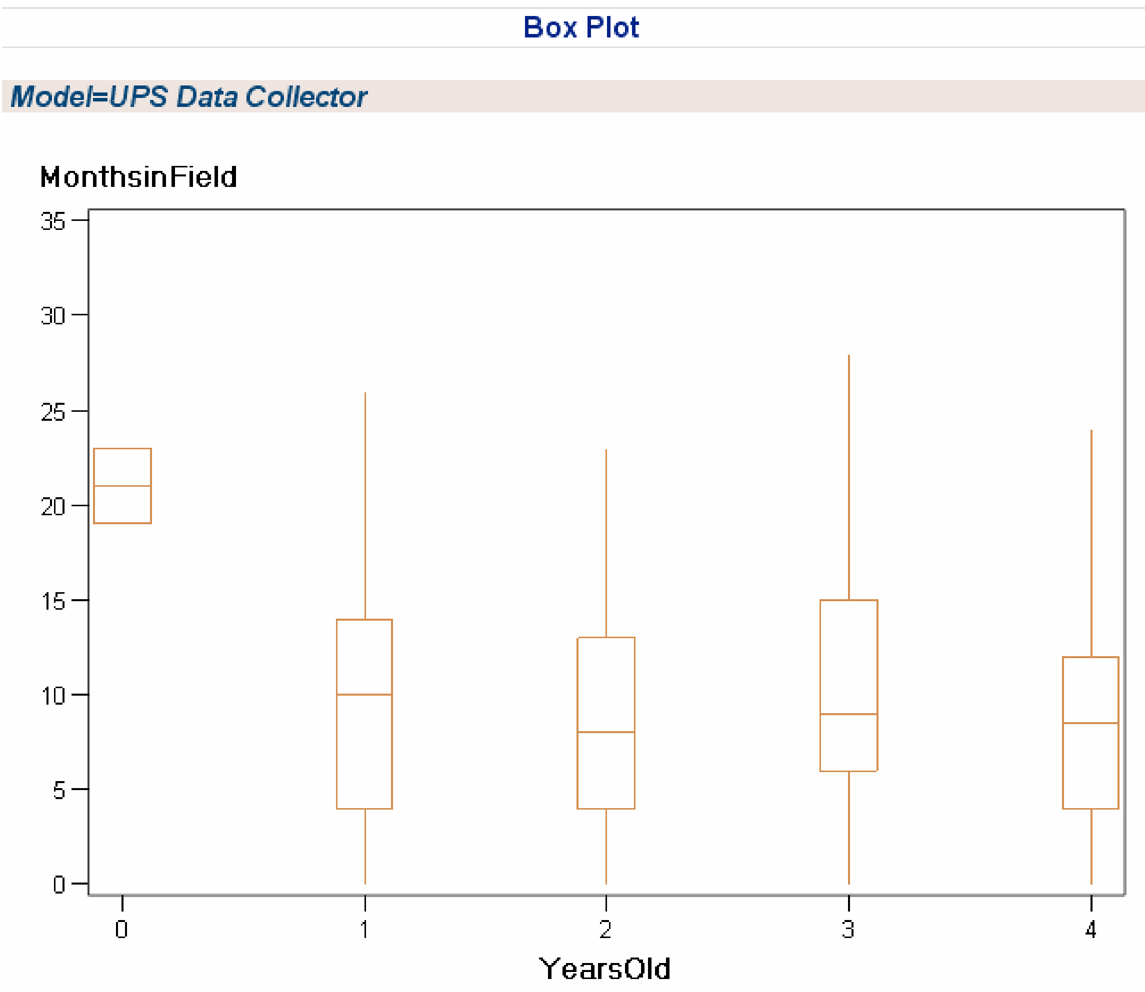
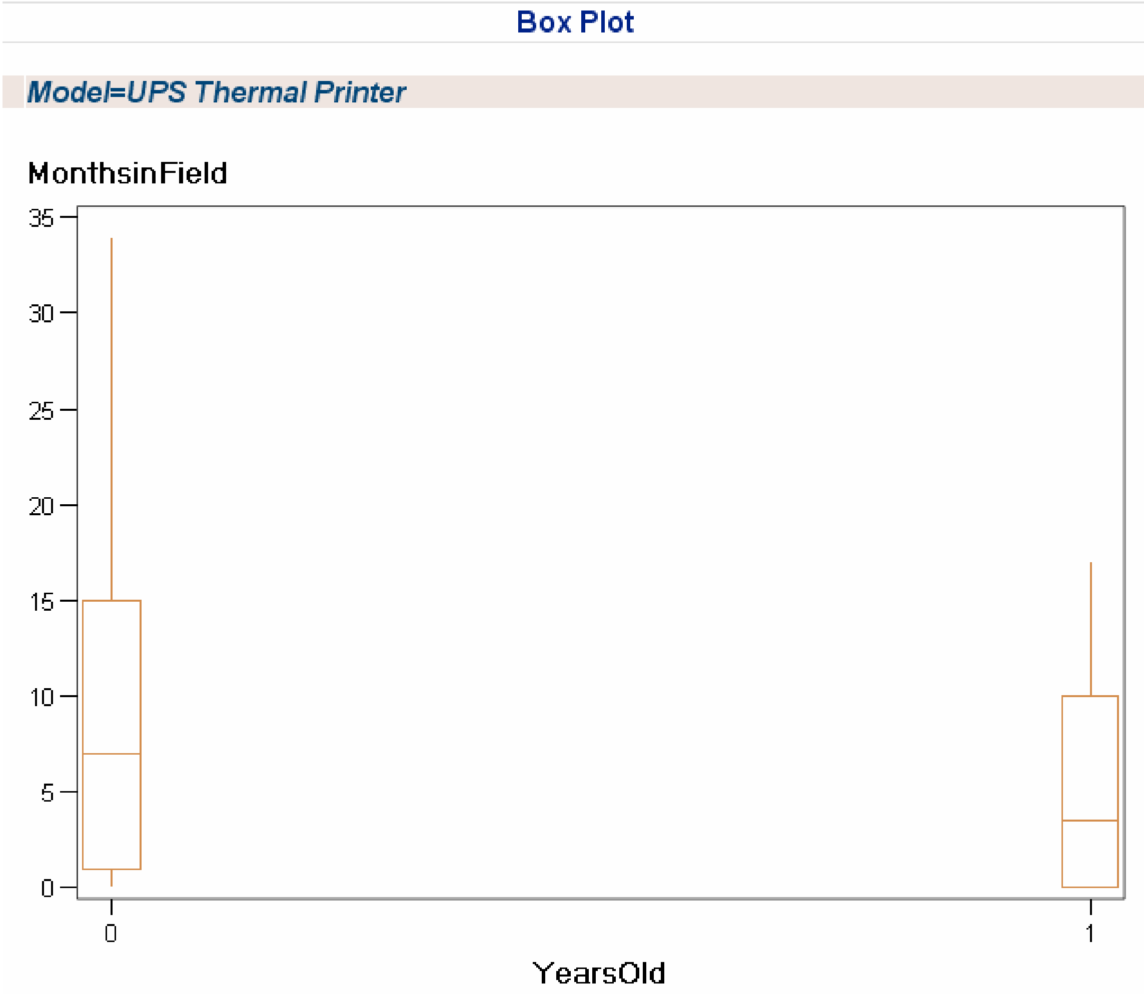


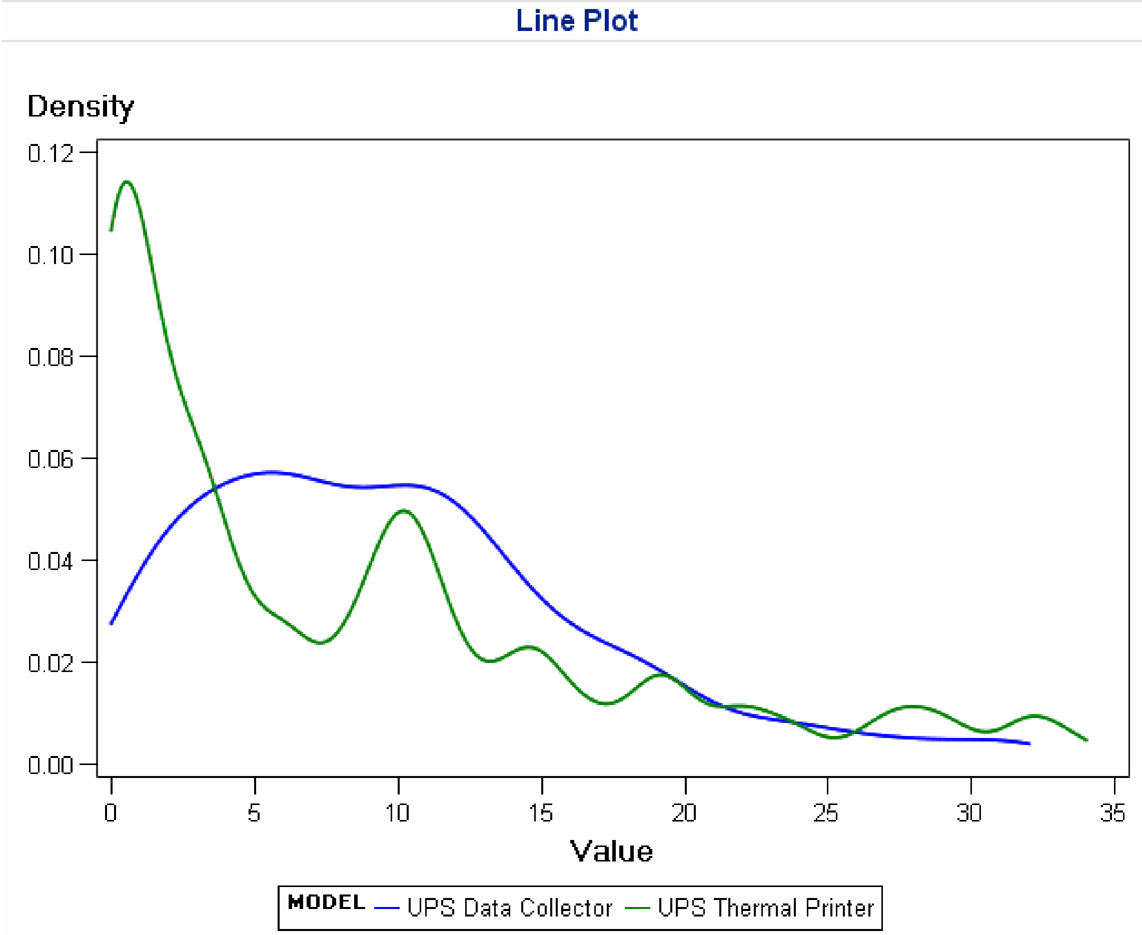
FIGURE 2. UPS THERMAL PRINTER MONTHS IN FIELD BY YEARS OLD



It can be observed that there is a significant variance in the average number of months in the field between the models and between “UPS Data Collectors” in the first year of use and in subsequent years.

Figure 3 is a kernel density estimate plot of the number of months between returns by model. Over 50% of the units returned by April 30, 2006 were returned within 12 months of the repair in the period of February 1, 2003 to January 31, 2004. The “UPS Thermal Printer” shows a concentration at 1 and 10 months between returns.

FIGURE 3. MONTHS IN FIELD BY MODEL NUMBER



The model experienced significant censoring with 7 “UPS Data Collectors” and 81 “UPS Thermal Printers” still in use on April 30, 2006. The “UPS Thermal Printer” experienced the highest degree of censorship with 32.4% of units still in use as of April 30, 2006.

Figure 4 shows the chi-square values for the various goodness-of-fit criteria. The generalized linear model applied to these variables appeared statistically significant with a p-value <0.0001.

FIGURE 4. OVERALL SIGNIFICANCE OF GENERALIZED LINEAR MODEL

Criteria For Assessing Goodness Of Fit			
Criterion	DF	Value	Value/DF
Deviance	406	2778.6056	6.8439
Scaled Deviance	406	2778.6056	6.8439
Pearson Chi-Square	406	2803.9788	6.9064
Scaled Pearson X2	406	2803.9788	6.9064
Log Likelihood		4778.1686	

Algorithm converged.

Figure 5 shows the statistical significance of the effect of age and model. Both effects appeared statistically significant (p-value <0.001 for each).

FIGURE 5. SIGNIFICANCE OF AGE AND MODEL NUMBER IN THE GENERALIZED LINEAR MODEL

LR Statistics For Type 3 Analysis			
Source	DF	Chi-Square	Pr > ChiSq
YEARSOLD	4	125.57	<.0001
MODEL	1	397.91	<.0001

Figure 6 shows the statistical significance of the predictions for each individual year of age and model. These also appeared statistically significant except for two years of age. Four years of age and “UPS Thermal Printer” were found to be linearly dependent on the preceding values.

FIGURE 6. SIGNIFICANCE OF PARAMETER VALUES IN THE GENERALIZED LINEAR MODEL

Analysis Of Parameter Estimates								
Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept		1	-10.5322	0.0898	-10.7082	-10.3562	13754.8	<.0001
YEARSOLD	0	1	0.8688	0.0926	0.6873	1.0504	87.99	<.0001
YEARSOLD	1	1	0.2539	0.0700	0.1167	0.3911	13.16	0.0003
YEARSOLD	2	1	0.0736	0.0742	-0.0718	0.2190	0.98	0.3213
YEARSOLD	3	1	0.2510	0.0657	0.1221	0.3798	14.57	0.0001
YEARSOLD	4	0	0.0000	0.0000	0.0000	0.0000	.	.
MODEL	UPS Data Collector	1	1.3372	0.0688	1.2024	1.4720	377.92	<.0001
MODEL	UPS Thermal Printe	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.0000	0.0000	1.0000	1.0000		

Note: The scale parameter was held fixed.

Table 1 shows the distribution of the predicted number of months between returns by equipment age for the "UPS Data Collector" and "UPS Thermal Printer" respectively. It can be seen that the predicted distribution is very similar to that shown in Figures 1 and 2. The "UPS Data Collector" population increased by 4,716 units and the "UPS Thermal Printer" population increased by 93,217 between February 1, 2003 and January 31, 2004. The variance in the distribution of the predicted number of months between returns is a result of this change in population. The "UPS Thermal Printer" distribution displays much greater variance due the more significant change in population.

TABLE 1. DISTRIBUTION OF PREDICTED MONTHS IN FIELD BY AGE AND MODEL NUMBER

Model	Year s Old	Average Predicted Months in Field	Minimum Predicted Months in Field	Maximum Predicted Months in Field
UPS Data Collector	0	19.7409	19.5776	19.9042
UPS Data Collector	1	10.6231	10.2251	10.7615
UPS Data Collector	2	8.8733	8.5266	8.9861
UPS Data Collector	3	10.5359	10.1714	10.7303
UPS Data Collector	4	8.3333	8.2116	8.3486
UPS Thermal Printer	0	9.6775	7.0044	12.8742
UPS Thermal Printer	1	5.5576	4.0012	6.9606

Units that had not been returned by 4/30/2006 were considered censored. The percentage of units censored is shown in Figure 7.

FIGURE 7. SUMMARY OF CENSORSHIP

Summary of the Number of Censored and Uncensored Values						
Stratum	YEARSOLD	MODEL	Total	Failed	Censored	Percent Censored
1	0	UPS Data Collector	2	2	0	0.00
2	0	UPS Thermal Printe	184	125	59	32.07
3	1	UPS Data Collector	57	57	0	0.00
4	1	UPS Thermal Printe	66	44	22	33.33
5	2	UPS Data Collector	56	52	4	7.14
6	3	UPS Data Collector	99	96	3	3.03
7	4	UPS Data Collector	36	36	0	0.00
Total			500	412	88	17.60

Due to the degree of censoring, life tables were calculated and a proportional hazards model was applied to the data. Life Tables were calculated for the entire sample using the Kaplan-Meier estimation method. Figure 8 shows statistical significance of the age and model strata. The strata were determined to be significantly significant.

FIGURE 8. SIGNIFICANCE OF STRATA UNDER KAPLAN-MEIER ESTIMATION

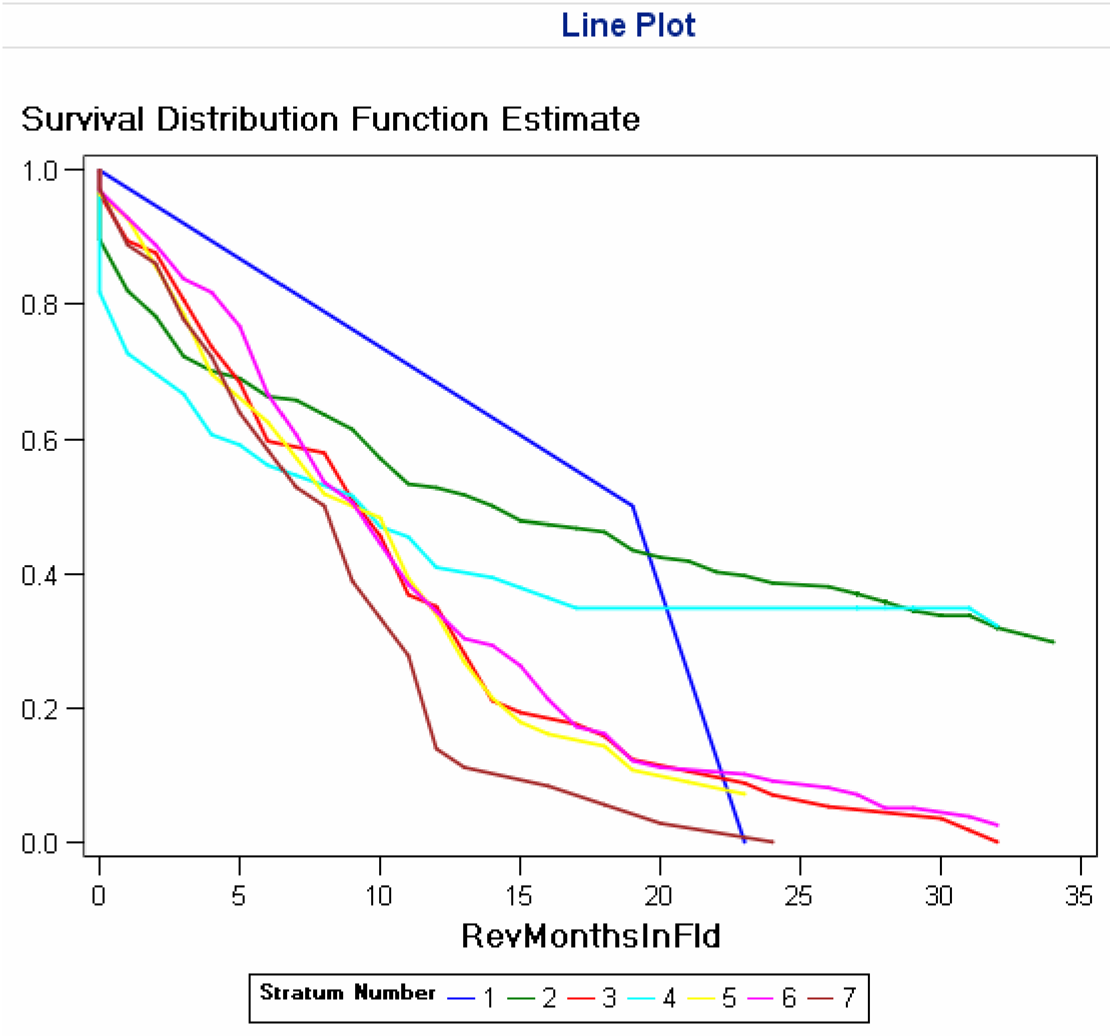
Test of Equality over Strata			
Test	Chi-Square	DF	Pr > Chi-Square
Log-Rank	52.4127	6	<.0001
Wilcoxon	15.7614	6	0.0151
-2Log(LR)	69.4875	6	<.0001

Figure 9 displays the survival curves of the strata described in Table 2 It can be observed that the survival curves for "UPS Data Collector" are consistent with the values predicted by the Generalized Linear Model. The survival curves for the "UPS Thermal Printer", the model which experienced more significant censoring, imply much longer life than that predicted by the Generalized Linear Model. The survival curves for the "UPS Thermal Printer" terminate at 35 months, the age of the oldest unit, without dropping below 30%.

TABLE 2. DESCRIPTION OF STRATA

Model	Years Old	Stratum
UPS Data Collector	0	1
	1	3
	2	5
	3	6
	4	7
UPS Thermal Printer	0	2
	1	4

FIGURE 9. KAPLAN-MEIER SURVIVAL CURVES BY STRATA



The proportional hazards model was fitted to the return information using age and equipment model as explanatory variables for the number of days until the next failure. “UPS Data Collector” was labeled “0” and “UPS Thermal Printer” was labeled “1” to create the new numeric variable ModelNum. The model was shown to be statistically significant overall (p-value <0.0001) as shown in Figure 10.

FIGURE 10. OVERALL SIGNIFICANCE OF THE KAPLAN-MEIER ESTIMATE

Testing Global Null Hypothesis: BETA=0			
Test	**Chi-Square**	**DF**	**Pr > ChiSq**
Likelihood Ratio	45.2203	2	<.0001
Score	45.6580	2	<.0001
Wald	44.1815	2	<.0001

The maximum likelihood estimates were found to be statistically significant for equipment model but not equipment age. The analysis of the maximum likelihood estimates is shown in Figure 11.

FIGURE 11. SIGNIFICANCE OF AGE AND MODEL NUMBER IN THE KAPLAN-MEIER ESTIMATE

Analysis of Maximum Likelihood Estimates									
Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits		Variable Label
ModelNum	1	-0.55617	0.16799	10.9608	0.0009	0.573	0.413	0.797	ModelNum
YEARSOLD	1	0.05929	0.06024	0.9685	0.3251	1.061	0.943	1.194	YEARSOLD

DISCUSSION

Figures 1 through 3 of the box plots and kernel density estimation appear to support the initial theory that the number of months between returns is affected by the age and model of the equipment.

The generalized linear model based on this assumption also appeared statistically significant (p-value<0.001). However, 17.6% of the units had not been returned for repair by 4/30/2006, the end of the study, undermining the validity of these results.

The life tables indicated that the strata were significant (Log-Rank and -2Log(LR) with p-value<0.0001 and Wilcoxon with p-value 0.0151). The proportional hazards model was statistically significant overall (p-value <0.001). The effect of model was shown to be statistically significant (p-value 0.0009) though the effect of age was not (p-value 0.3251). This contradicts the finding of the generalized linear model which found each of the effects included (age and model) to be statistically significant (p-value < 0.001).

The time between failures predicted by the generalized linear model was very similar to the actual time between failures. The survival curves for the “UPS Data Collector” were consistent with the predictions made by the generalized linear models. The survival curves for the “UPS Thermal Printer”, the model which experienced 92% of the censoring, predict a much longer time between failures than the generalized linear model.

The survival curves for the “UPS Thermal Printer” terminate at 35 months, the age of the oldest unit, without dropping below 30%. An additional study would need to be performed at a later date to approximate the complete survival curves.

Further research would be required to determine the significance of the effect of age in the presence of less severe censoring. Further study would also be required to determine if other thermal printers and data collectors experience similar survival rates. The effect of age might prove to be significant in these other models.

APPENDIX

The analysis performed in this study was completed using SAS Enterprise Guide version 3. The code created to perform the analysis follows.

GENERALIZED LINEAR MODEL

```
proc genmod data=sasuser.qury9059;
class yearsold model ;
model monthsinfi= yearsold model / dist=poisson
                                   link=log
                                   offset=ln
                                   type3;
output out      = Residuals
        pred      = Pred
        resraw    = Resraw
        reschi    = Reschi
        resdev    = Resdev
        stdreschi = Stdreschi
        stdresdev = Stdresdev
        reslik    = Reslik;
run;
```

LIFE TABLES

```
/* Run PROC LIFETEST to perform the analysis. */
PROC LIFETEST DATA=WORK.EGT6866
        OUTSURV=SASUSER.LTOS420(LABEL="Survival Estimates and Confidence Limits for
WORK.QURY9059_17284")
        OUTTEST=SASUSER.LTOT6167(LABEL="Rank Test Statistics for WORK.QURY9059_17284")
        ALPHA=0.05;
```

```
	STRATA YEARSOLD MODEL;
	TIME RevMonthsInFld * Censored (1);
RUN;
```

PROPORTIONAL HAZARDS

```
/* Run PROC PHREG to perform the analysis. */
PROC PHREG DATA=WORK.EGT6329
	OUTEST=SASUSER.PRG36757(LABEL="Regression Coefficient Estimates for WORK.QURY9059_17284")
	COVOUT;
	MODEL RevMonthsInFld * Censored (1) = ModelNum YEARSOLD /
		TIES=BRESLOW
		CORRB
		COVB
		RISKLIMITS ALPHA=0.05
		SELECTION=NONE;
	OUTPUT OUT=SASUSER.PRG13617(LABEL="Survival Statistics for WORK.QURY9059_17284")
		XBETA=_PRED
			RESDEV=_RESDEV
			RESMART=_RESMAR
			RESSCH=_RESSCH1 - _RESSCH1
			RESSCO=_RESSCO1 - _RESSCO1
			STDXBETA=_STDERR
			SURVIVAL=_SURV
			LOGSURV=_LSURV
			LOGLOGS=_LLSURV
			NUM_LEFT=_LEFT
			LMAX=_LMAX
			LD=_LD
			DFBETA=_DFBETA1 - _DFBETA1  ;
	BASELINE OUT=SASUSER.PRG27106(LABEL="Baseline Survivor Function Estimates for 
WORK.QURY9059_17284")
		SURVIVAL=_SURVIV_
			UPPER=_SDFUCL_
			LOWER=_SDFLCL_
			LOGLOGS=_LOGLOGS_
			LOGSURV=_LOGSURV_
			STDERR=_STDERR_
			STDXBETA=_STDXBETA_
			XBETA=_XBETA_ ;

RUN;
```

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Rebecca Dennison
University of Louisville
4703 S 6th St
Louisville Kentucky 40214
Work Phone: 502.762-8528
Email: REDenn01@louisville.edu
Web: www.louisville.edu/~redenn01