Faculty Paper Series

Faculty Paper 02-3 June, 2002

Survival Analysis of U.S. Meat and Poultry Recalls, 1994-2001

By

Victoria Salin, Neal H. Hooker, and Ratapol Teratanavat

Department of Agricultural Economics
2124 TAMU
Texas A&M University
College Station, Texas 77843-2124

Survival Analysis of U.S. Meat and Poultry Recalls, 1994-2001

Victoria Salin,[1] Neal H. Hooker, and Ratapol Teratanavat

Introduction

There were hundreds of voluntary recalls of meat and poultry in the United States during 1994-2001. What can be learned from this experience to improve the industry's ability to safeguard the food supply? This study applies statistical methods that explore the dynamics of food recalls, rather than examining a particular point in time or aggregating over periods of time. It is an approach that adds to the research base in economics and business decision-support and will contribute to the effort to use risk analysis as the foundation for U.S. food safety policy.

Risk assessment is the first step in the risk analysis process that underlies the U.S. policy approach to food safety. Risk assessments are followed by risk management and risk communication programs to complete the three-part risk analysis process. Both government and industry participate in the risk analysis process. Given the limitations on government's direct role in food handling and processing, it is important to fully understand business incentives for food safety enhancement. Business managers, surveillance personnel, and policymakers can only implement effective risk management programs if they have complete information about the probability of harm and the severity of food contamination incidents at the business and industry level.

[1] Contact: Dr. Victoria Salin, Asst. Prof., Dept. of Agric. Econ., Texas A&M University, College Station, TX 77843-2124, (979) 845-8103, v-salin@tamu.edu

Risks related to bacterial contamination of food are dynamic in nature. Conditions at any stage in the farm-to-table food chain could generate contamination. For example, the factors controlling microbial populations in seafood are temperature, pH, organic acid levels, water activity, and preservatives (Ross, Dalgaard, and Tienungoon), many of which change with time, particularly given the differences in handling technology along the food chain. A model in the field of predictive microbiology would incorporate some or all of the factors and provide information about exposure assessment, in terms of time required for a 1,000-fold increase in pathogen numbers, or total microbial populations, for example. In order for the information on exposure to be most useful in economic decision-making, it must be translated from microbiological terms into disease incidence and, finally, into costs of illness and death. There is substantial uncertainty regarding the relationship between pathogen levels and health, and the economic benefits of improved health (Antle). Given the current state of knowledge, economic researchers should explore a variety of measures that may be useful in assessing costs or benefits of food safety programs. At the firm level, one set of readily available information reflecting business risks related to food safety is the recall. Recalls are likely to be a function of all the microbiological and food production factors that arise during handling, storage, and preparation. Business management and regulatory decisions also likely affect the probability of recalls. For example, investments in testing technology or the effectiveness of surveillance programs would impact recalls. Therefore, we would expect that the dynamics of food recalls are complex and multifaceted.

Other researchers have analyzed the descriptive statistics on food recalls, using Food Safety and Inspection Service (FSIS) data (Teratanavat and Hooker). Descriptive statistics are a necessary starting point in risk assessment or evaluations of risk management activities, but they cannot accurately portray the complete time series characteristics. Means and standard deviations are derived from data pooled over time, thus obscuring information about timing. A researcher can further subset the data and compute descriptive statistics to isolate important time periods, but the disaggregation procedure may lead to the loss of degrees of freedom and can impair statistical inference. Thus, the aggregate statistics must be supplemented with other approaches.

The analysis of survival data provides the basis for such an alternative perspective on food safety risk assessment and risk management. This paper fills a gap in our knowledge about the dynamic properties of food recalls. The objective is to contribute to the understanding of the time series processes that underlie the risks businesses face as a result of food contamination.

The method is to use models for duration data and estimate survival functions. Survival functions describe the time until failure or, in its broadest sense, the time before occurrence of an event. In engineering statistics, and in many statistics texts, survival functions are applied to such things as light bulbs and machinery (Pitman). In medical science and biostatistics, survival functions are commonly used to describe effectiveness of treatments (Lee). In social science, survival analysis has been used to analyze worker strikes, unemployment spells, time until business failure, and intervals between purchases. (See Greene for a general description; Agarwal and Mahmood present applications to small business failures.)

In this application to FSIS meat and poultry recall data, the time before failure is defined as the time elapsed before a recall occurs. We examine time before a recall, for the firms that experienced a recall within the 1994-2001 period. This application of survival data analysis to the FSIS recall database gives statistical results that must be interpreted in a precise way. An analogy to medical science will clarify. Consider a group of patients with terminal illness. If some patients have received an experimental medical treatment, survival data analysis can determine if the treatment had an effect on survival times. If all patients remain in the study until death, exact survival times for all the subjects are known. It is useful to know whether the treatment had an effect, but the results say nothing about the health of people not included in the study.

The implications we can glean from the recall database are likewise limited to the subjects being observed—only those firms that had a recall. In spite of this limitation, some information can be obtained with analysis of times before recall. First, dynamics can be examined directly. The dynamics could answer questions such as whether recalls are coming faster in certain seasons, or in particular years. Second, if a policy change has occurred, the distributions of time until failure with and without the policy can be compared, analogous to the evaluations of medical treatment. It should not be overlooked that hundreds of firms are included in this dataset, which is a useful source for a preliminary examination of the effectiveness of HACCP. While the data set clearly does not constitute a random sample, it is not necessarily non-representative and may be helpful in drawing conclusions about food safety.

Keeping in mind the limitations on our data, the objectives of this investigation are to:

- Improve understanding of the dynamics of food recall risk, and
- Investigate whether policy regimes are associated with any differences in the time series process of food recall risk.

This paper demonstrates the use of non-parametric methods to estimate survival functions for meat and poultry recalls during 1994-2001. The results from non-parametric and parametric methods are compared to determine the relative utility of the two approaches in the context of the economics of food safety. The results of this research shed some light on food safety policy and can serve as useful information to drive other economic analysis of food safety. Dynamics are important to many business decisions, as the financial event studies in the literature on the economics of food safety have demonstrated (Salin and Hooker, Thomsen and McKenzie, Wang, et al., and Henson and Mazzochi). Other predictive economic and business models that rely on risks related to food contamination may also be developed and applied using the results from this survival analysis.

Models of Survival Data

In this section, the conceptual basis for the duration models of food recalls is explained. The statistical foundation for the estimations is also presented, drawing mainly upon the work of Lee. Procedures to estimate functions of interest are presented both parametric models that are estimated with maximum likelihood procedures and non-parametric models that are simpler to estimate and well suited for analysis of relatively small samples.

Conceptual framework

Analysis of survival times can be applied to any processes in which the duration of a condition is of interest. The survival time, or time before failure in this research, is defined as the time that passes before a Class 1 food recall event occurs. The data that measure the time to the event of interest is survival time (T), measured in days.

We define two regimes for consideration of the effects of food safety policy. The first begins January 1, 1994, with modern/industrial food processing under FSIS continuous inspection. Thus survival time T=24 signifies a food recall case that opened on January 25, 1994. In early1994, a major incident of food contamination was attracting the attention of U.S. businesses, policy makers, and consumer activists. The policy regime in place at that time was the FSIS system of continuous visual inspection at every meat processing plant. Notices of the regulatory changes to come under Hazard Analysis Critical Control Points (HACCP) were published in the Federal Register. Large meat processing plants (those with more than 500 employees) were required to have HACCP systems in place by January 26, 1998. On that date the second policy regime was defined to begin with the HACCP requirement in large meat packing facilities. The regulation for smaller plants had implementation dates of January 25, 1999 or January 25, 2000 (Teratanavat and Hooker).

Functions of survival time

Three equivalent functions are used to describe the distribution of survival time: the survivorship function, the probability density function, and the hazard function. These functions can be derived from each other mathematically. The survivorship function (also called survival function or cumulative survival rate) is defined as:

$$S(t) = Pr\text{ (time before failure is at least } t)$$
$$= Pr\ (T > t\ )$$

The survivorship function depicts the probability that the failure has not occurred at time t. By definition of the cumulative distribution function (CDF) of any random variable, survivorship functions are related to CDFs such that

$$S(t) = 1 - Pr\ (T < t)$$
$$= 1 - F(t), \text{where } F(t) \text{ denotes the CDF}$$

Graphically, survivorship functions are usually decreasing in time, and in biostatistics applications, they are often found to be decreasing at an increasing rate (convex). Steep survival curves represent short survival times.

The probability density function for survival times is defined in the usual way for a density of a continuous random variable:

$$f(t) = \lim_{\Delta t \to 0} \frac{P\{\text{failure in interval}(t\ , t + \Delta t)\}}{\Delta t}$$

Density curves illustrate the proportion of failures that occur in any time interval, as well as any peaks in the number of failures during the time period under study.

Hazard functions describe the conditional failure rate, or the probability of failure during a very small interval, given survival up to the beginning of the interval. The hazard function is defined as:

$$h(t) = \lim_{\Delta t \to 0} \frac{P\ \{\text{subject with survival to age } t \text{ fails in the time interval } (t, t + \Delta t)\}}{\Delta t}$$

Another definition of the hazard function uses the relationship between the probability density and CDF of survival time, as follows:

$$h(t) = \frac{f(t)}{1 - F(t)}$$

Data

The data are from the U.S. Department of Agriculture FSIS database (Teratanavat and Hooker).[2] The incidence of Class 1 meat and poultry recalls from 1994 to 2001 is shown in figure 1. The scatter plot shows that most recalls are relatively small in terms of pounds of product affected. It also appears that the density of recalls is greater in the years following the HACCP requirement for large plants. Aggregate statistics show evidence of increasing number of recalls over time. There were 86 Class 1 recalls in the 1994-January 25, 1998 period, the period before HACCP was required for large meat and poultry processors. From January 26, 1998, to the end of 2001, there were 219 Class 1 recalls, two and one-half times more recall events than in the pre-HACCP period. These data on the increasing number of recalls over time could lead to concerns that the food supply is becoming less safe.

This study includes only the firms that have experienced a recall, which eliminates the potential problem of censoring. Censoring occurs when subjects have not experienced the failure event during the time of the study. For example, in medical studies, if treatment was successful for a number of patients, their true time before

---

[2] The data used here differ from that used in Teratanavat and Hooker in the following ways: Only Class 1 recalls are used in this research; the full set of observations for 2001 became available; the date on one observation was corrected, and observations for a given firm that were dated within a 5-day period were considered a single recall event.

failure is not known exactly. Special statistical procedures are required to account for censoring (Lee, and Greene, chapter 22.3).

Estimation Procedures

Parametric estimation requires that the researchers assume a functional form, then estimate the parameters of the function using maximum likelihood procedures. For instances in which no particular distribution is known to fit the data, non-parametric techniques are useful in estimating functions of survival time. Both non-parametric and parametric estimation are undertaken in this research. Because there are no censored observations, the techniques are straightforward.

Non-parametric methods for estimating survival distributions

According to Lee, non-parametric methods are "...less efficient than parametric methods when survival times follow a theoretical distribution and more efficient when no suitable theoretical distributions are known" (pg. 66). It is suggested to use the non-parametric methods before attempting to fit a theoretical distribution.

The observations are not grouped, following Lee's recommendation for a simple case with no censored data. Defining $t_1, t_2, \ldots, t_n$ $(i = 1, 2, \ldots, n)$ to be the exact times before failure of the n recall events, the survivorship function $\hat{S}(t_i)$ is the estimated probability that the time before failure is at least t. Because there are no censored observations, the estimated survival function is defined as the proportion of subjects for which time before failure is greater than t.

The survival distributions were estimated in SAS using the Lifetable procedure. This employs the Kaplan-Meier product limit method of estimating survival functions based on actual survival times, without grouping times into intervals. The Kaplan-Meier product limit procedure is based on the fact that the probability of surviving k or more periods from the beginning of the time observed is the product of k survival rates, defined as the proportion of subjects surviving at least to year k. The equation for the survivorship function is:

$$\hat{S}(t) = \prod_{t_{(r)} \le t} \frac{n-r}{n-r+1},$$

where n is the total number of observations and r is the rank of the n survival times, in increasing order in this application because there are no censored observations. Once the survivorship function is estimated, hazard rates can be obtained from the mathematical relationship between the survivorship and hazard functions.

Parametric methods for estimating survival distributions

Analytical methods are used to estimate the parameters of some common statistical functions and assess their fit to the time before food recalls. The exponential distribution is often used in biomedical applications. The Weibull distribution has the advantage of being flexible to accommodate a variety of shapes and it is considered here along with the normal and exponential distributions. The equations of the normal distribution are well known and not presented in this section.

Exponential. The exponential distribution is characterized by one parameter, λ. Its density function is:

$$f(t)=\begin{cases} \lambda e^{\lambda t}, & t \ge 0,\ \lambda>0 \\ 0, & t<0 \end{cases}$$

The cumulative distribution function is:

$$F(t) = 1 - e^{\lambda t} t \ge 0,$$

and the survivorship function is then:

$$S(t) = e^{-\lambda t}, t \geq 0.$$

Using the mathematical relationships developed previously, the hazard function is constant:

$$h(t) = \lambda,\ \ t \geq 0$$

Weibull. The Weibull distribution is flexible to accommodate multiple possible shapes. The probability density function and cumulative distribution functions are, respectively,

$$f(t) = \lambda\gamma(\lambda t)^{\gamma - 1} e^{-(\lambda t)^{\gamma}} \quad t \geq 0,\ \gamma,\ \lambda, > 0$$

$$F(t) = 1 - e^{-(\lambda t)^{\gamma}}$$

Given that *S(t) = 1- F(t),* the survivorship function is easy to calculate once the estimated parameters are provided.

Results

Overall, the survival data analysis does not consistently support the concern that the food supply is becoming less safe, based on pre- and post-HACCP survival functions estimated with non-parametric methods. There is evidence of clustering of recalls and some seasonal patterns across the years examined. Fitting of theoretical distributions led to rejection of many statistical distributions, due in part to positive skewness in the distribution of times before recall. The best-fitting distributions provide evidence that risk of recall is increasing with time.

Pre- and Post-HACCP Dynamics of Recalls Using Non-Parametric Estimation

Initial evidence that the food supply is increasingly contaminated can be obtained from the graph of the survival function estimated with non-parametric methods for the entire 1994-2001 period (figure 2). The horizontal axis shows time before failure in days. The vertical axis is the probability of survival. The concave shape, with steeper portions of the survival function occurring later, suggests that failures occurred faster in the years following HACCP.

After subdividing the sample and estimating survival functions for the pre- and post-HACCP periods, a more detailed picture of the dynamics emerges. The pre-HACCP survival function (figure 3) is steepest in the early part of the period and flattens beginning in 1996, suggesting a high rate of failure around 1994-1995. Recalls occur less often beginning in late 1995, as is shown by the flatter portion of the survival function. This improvement in the duration of time before failure could have resulted from firms preparing to implement HACCP.

The post-HACCP period begins with a similarly flat portion of the survival function (figure 4), again possibly suggesting some success of the program in slowing failures. During 2000 (days 700-1000), however, the picture changes, with an increasing pace of recalls for the year. The cluster of recalls during 2000 could have been an anomaly that contributed to higher aggregate numbers of recalls post-HACCP, since it is clear that the survival function does not maintain that steep slope through the end of 2001.

A non-parametric statistical test was used to evaluate the hypothesis that survival distributions pre-HAACP were below those following HACCP implementation. The Wilcoxon test is based on a comparison of the ranks of each observation in the distributions being compared. The test statistic W is large when survival times, post-HACCP, are larger than the survival times pre-HACCP. A normal distribution is assumed for W (asymptotically), thus Z scores $(Z = \frac{W}{\sigma_W})$ are used as the statistical test. We use the Mantel procedure for calculating Gehan's Generalized Wilcoxon test statistic, recommended in Lee (page 106-108). Results are reported in table 2, and provide a strong basis for rejection of the null hypothesis that the distributions are equal. Thus, on the basis of non-parametric estimation, there is support for the hypothesis that times before recall are longer following implementation of the HACCP program.

In addition to the statistical test comparing the full distributions, two other statistics from the survival estimation support the improved food supply hypothesis. According to Lee, median survival time is the single most commonly used statistic to describe a survival distribution. Median survival time corresponds to the 50^{th} percentile of the survival distribution. Median time before recall is 846 days following the HACCP program, compared with median time pre-HACCP of 514 days, which is well below the 95% confidence interval around the corresponding post-HACCP median time. Another summary statistic, the percent that survive one year, gives a similar result, with the probability of surviving one year at 83% post-HACCP and 66% pre-HACCP.

The date used for subdividing the sample into the pre- and post-HACCP periods reflects regulatory requirements for large firms. The use of a cut-off date implies that HACCP is turned on immediately, which of course is a simplification. It is possible that many of the large firms implemented HACCP before the date in question, while smaller firms were given longer time to adopt HACCP. Because smaller firms had more time to adopt HACCP, it may be more appropriate to contrast the distributions only for large firms. Most firms that experienced recalls during 1994-2001 were not large, thus the sample size falls considerably when attention is limited to large firms (table 3). The relatively small sample size for large firms results in survival distributions with wide confidence intervals, and thus we can expect statistical evidence to be less compelling than with the preceding comparison based on the full dataset of Class 1 recalls.

The pre-HACCP survival function for large firms, estimated with the non-parametric approach (figure 5), is flatter before 1996 than in the years approaching full implementation. Immediately following HACCP (figure 6), recalls were relatively slow in coming, but a period of rapid failure for large firms occurred in 1999. Median survival times post-HACCP were 717 days for large firms, below the median of 929 days in the period prior to HACCP. However, the 95% confidence interval around the pre-HACCP estimate is from 502 to 1,248 days, so there is no statistically significant difference in the two estimates. The means of the pre-and post-HACCP distributions for large firms are likewise not statistically different, based on a z-test for comparison of the means of two independent populations. Overall, the subsample of large firms does not offer strong evidence regarding a difference in dynamics of recalls for these two periods.

Hazard rates derived from the survivorship functions estimated with the non-parametric methods are reported in table 4. Recall that hazard rates are defined as the probability of a recall, conditional on the time elapsed since a food safety regime was in effect. They provide information about the likelihood of a recall as a function of time. For example, machinery that wears out with age would have an increasing hazard rate. Hazard rates associated with accidental death or other purely random factors are typically estimated as constant. Figures 7-8 indicate that there is some variation in hazard rates for meat and poultry recalls, but for most of the period, a constant hazard rate model would approximate the data fairly well. The reason for the rapid increase in hazard rates at the end of the two time periods is not clear. Annual hazard functions (figure 9) suggest more clearly a pattern of increasing hazard rates associated with food recalls.

Pre- and Post-HACCP Dynamics of Recalls Using Parametric Estimation

Probability density functions were fit to the data on survival times before and after HACCP to identify which, if any, distributions represent the data well. The parameters from the well-known mathematical forms can be utilized in a variety of models for risk analysis (table 5).

Normality was rejected for all of the distributions, based on the tests calculated in SAS (Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling test statistics). The goodness of fit tests for the distributions of all Class 1 recalls did not provide strong support for any theoretical distribution. The peak in the right tail of the histogram (figure 10) can explain the poor fit . This tendency toward a bi-modal empirical distribution is not well represented by the theoretical forms that were fitted. Based on the test procedures described by Lee, the Weibull, lognormal, and exponential distributions are rejected for the distributions for all firms. Only the Weibull was not rejected as a model of the distribution for large firm's subsample. Using the Kolmogorov test in SAS, the Weibull was not rejected for the pre-HAACP distribution of time before failure, for all firms.

A closer look at the estimated parameters of the fitted Weibull distributions is warranted. The graphs of the density functions for the fitted distributions are unimodal and slightly positively skewed. The shape of the Weibull distribution curve is determined by the value of γ. When $\gamma =1$, the Weibull distribution becomes the exponential case, with a constant hazard rate. The parameters estimated all find $\gamma > 1$ (table 6), which would come from a population with increasing risk, or "positive aging." This result is consistent with the hazard rates estimated with the non-parametric procedures, in which hazard rates increased at the end of the time periods. Very small values of λ, such as those found here, are a result of the scaling being over fairly large numbers which occurs when time until failure is denoted in days.

The plot of the survival function from the fitted Weibull distribution is contrasted with the corresponding function estimated from the non-parametric Kaplan Meier procedure (figures 11-12). The results of decreasing survival are fairly consistent, but the rate of decline in survival is increasing in the parametric estimation, but not according to the non-parametric result. This result occurs because the fitted Weibull cannot accommodate the skewness evident in the distributions, and the non-parametric method is flexible to accommodate the data while a theoretical distribution is not.

Seasonal Dynamics of Recalls

Survival functions for each calendar year from 1994-2001 were examined in order to investigate seasonal patterns in the timing of food recalls. Flat survival functions during the mid-winter (days 0-60) indicate longer time before failure in winter, for six of the 8 years examined (1994, 1996, 1997, 1998, 2000, and 2001). There is a steep drop in the graph, meaning more frequent recalls, during the second half of the year (days 250-350) for 1994, 1995, 1997, 1998, 2000 (weakly) and 2001.

Statistical tests of annual survivorship distributions were conducted to determine if there is evidence of increasing or decreasing risk with time. The only statistically significant result is that the distribution of recalls in 1994 is below that of 2001, indicating that time before recall was longer in 2001 (table 2). This is another piece of evidence in support of an improvement in food safety. Median survival time was greatest in 1998, the year that HACCP was required for large plants, but median survival times did not increase consistently in each year following HACCP (table 7).

Conclusions

This research is an initial inquiry into the application of survival data analysis to the industry statistics on recalls of meat and poultry. To our knowledge, it is the first such analysis conducted in the literature on food safety and quality. The methods are well-known in biostatistics and can easily be transferred from the medical and engineering applications to allow investigation of the dynamics of food recalls, which is the food safety event that is most relevant to business decisions.

The analysis of the recall data provided some insight into the effectiveness of food safety programs by examining dynamics of recalls before and after implementation of HACCP. Even if the recall is initiated because of a “zero tolerance” or an unrealistically low pathogen level that would not have caused significant illness, the recall means that the businesses in the distribution system experience direct costs and perhaps less tangible costs as a result of loss of reputation with customers. The most significant finding from a policy perspective is that survival data analysis does not consistently support the concern that the food supply is becoming less safe. The pre- and post-HACCP survival functions estimated with non-parametric methods indicate that times before failure are longer since the program went into place. The differences in survivorship functions are statistically robust and provide some support for the effectiveness of HAACP.

The efforts to fit theoretical distributions to the time before recalls were less successful. Weibull distributions emerged as the most plausible among the functional forms considered, and their parameters led to the conclusion that probability of recall is increasing with time. The lack of statistical support for many functional forms occurred because the distribution of times before failure had bimodal shapes, skewness, and fat tails, features that most mathematical forms used in statistics do not share. This result illustrates the importance of developing predictive models that are flexible and not tied to a particular mathematical distribution.

One can envision many ways in which survivorship functions of recalls or other industry-level indicators can be useful for surveillance programs, decision support models, and evaluation of regulations. For example, the evidence of clustering of recalls and seasonal patterns would be useful in formulating a baseline against which monitoring information can be compared in real time. Data such as these could also be used in some form of dynamic performance standards to evaluate benefits of food safety programs.

The main limitation of survival data analysis is that it is not possible to define an event with a differential measure of severity. That is, the recall event is defined as a failure whether the recall encompassed 25,000 pounds or 25 million pounds. By limiting the analysis to Class 1 recalls, this study targeted only those recalls considered most significant by FSIS. It would be possible to re-define failure as recalls over a certain threshold severity in terms of pounds, or in terms of illness and death. But the researcher would potentially face problems with statistical comparisons of the two survival functions estimated with small sample size, similar to the problems experienced with the subsample of large firms examined in this paper.

Next steps in this research area include refinements of the data and the application of econometric procedures such as the Cox regression to investigate the relationship of hazard rates to explanatory factors other than time. The data set of FSIS recalls could be augmented with FDA recalls. A broader industry-level perspective would be obtained by including all plants, in addition to those experiencing recalls.

Table 1. Pounds Recalled in Class 1 Recalls, 1994-2001

	N	Mean	Std Dev	Minimum	Maximum
1994 - Jan. 25, 1998	86	390,848	2,711,501	36	25 million
Jan. 26, 1998 - 2001	219	565,421	3,649,998	4	35 million

Table 2. Results From the Comparison of Survival Distributions Before and After HACCP, Using Wilcoxon Test

	W	Var(W)	Z	Test Result
All class 1	4,753	1,965,024	3.39	Reject Ho
Large firms	-69	11,530	-0.643	Cannot reject Ho
2001 versus 1994	1,972	64,414	7.77	Reject Ho

Ho : S1(t)=S2(t), Ha : S1(t)<S2

Table 3. Summary Statistics for Class 1 Recalls by All Firms and Large Firms.

	Recalls by All Firms		Recalls by Large Firms		
	Pre-HACCP	Post-HACCP	Pre-HACCP	Post-HACCP	Total Large
From raw data:					
N	86	219	16	36	52
Mean pounds	390,847	565,424	1.7 million	1.3 million	1.4 million
Maximum pounds	25 million	35 million	25 million	35 million	35 million
From non-parametric estimation:					
Median days before failure	514	846[1]	929	717[2]	
Mean days before failure	638	812[1]	843	730[2]	

[1] Statistically different from the corresponding pre-HACCP estimate at the .05 level.
[2] Not statistically different from the corresponding pre-HACCP estimate.
Source: Author's calculations using FSIS data

Table 4. Estimated Hazard Functions, Annual 1994 - 2001

	Interval of Time Before Failure, in Days					
	0~0	1~90	91~180	181~270	271~360	361~450
		Hazard Rate				
1994	0	0.004	0.002	0.004	0.022	
1995	0	0.004878	0.006222	0.006349	0.022222	
1996	0	0.002299	0.006667	0.016667	0.022222	
1997	0	0.001481	0.003704	0.009524	0.022222	
1998	0	0.001481	0.003704	0.007407	0.022222	
1999	0	0.001481	0.003704	0.007407	0.022222	
2000	0	0.002222	0.007046	0.009687	0.022222	
2001	0	0.002593	0.002573	0.006496	0.018667	0.011111

Table 5. Results From Parametric Density Estimation

	All Class 1 Recalls							
	Pre-HACCP (n = 86)				Post-HACCP (n = 219)			
	μ/θ	σ	Mode	Z	μ/θ	σ	Mode	Z
Normal	637.60	428.87	637.60		812.47	389.36	812.47	-
Lognormal	0	1.02	158.92	6.11	0	0.70	413.05	6.52
Exponential	0	637.60	0	-	0	812.47	0	-
Weibull	0	696.58	292.85	1.42	0	912.46	685.67	2.17
	Class 1 Recalls by Large Firms (n = 16)				Post-HACCP (n = 36)			
	μ/θ	σ	Mode	Z	μ/θ	σ	Mode	Z
Normal	843.31	423.81	843.31	-	729.72	434.35	729.72	-
Lognormal	0	1.28	114.30	6.39	0	0.89	247.22	6.31
Exponential	0	843.31	0	-	0	729.72	0	-
Weibull	0	915.03	544.92	1.70	0	810.17	459.49	1.65

Note: Results are in days before failure.

Table 6. Parameters of Weibull distributions, time before recall, in days

	Shape (γ)	Scale (λ)
All firms, pre-HACCP	1.42	.0014356
Large firms, pre-HACCP	1.70	.0010929
All firms, post-HACCP	2.17	.0010959
Large firms, post-HACCP	1.65	.0012343

Table 7. Recall Data by Year

Year	Number of recalls (Class 1)	Average pounds recalled	Average duration of case in days	Average survival time in days	Median survival time in days
1994	28	50,068	117	199	186
1995	25	192,944	163	147	134
1996	16	21,962	188	169	165
1997	16	1,687,875	246	206	212
1998	32	1,348,484	218	218	235
1999	55	712,273	NA	181	167
2000	66	308,361	NA	172	165
2001	67	316,104	NA	204	210

NA: not available because some cases are still open as of February, 2002

Figure 1. Meat and poultry recalls, Class 1, 1994-2001.

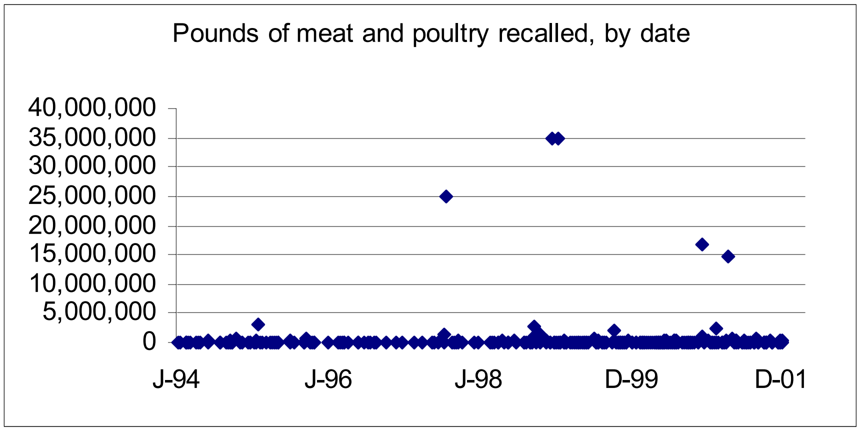


Figure 2. Survival function for all Class 1 recalls, in days, 1994-2001.

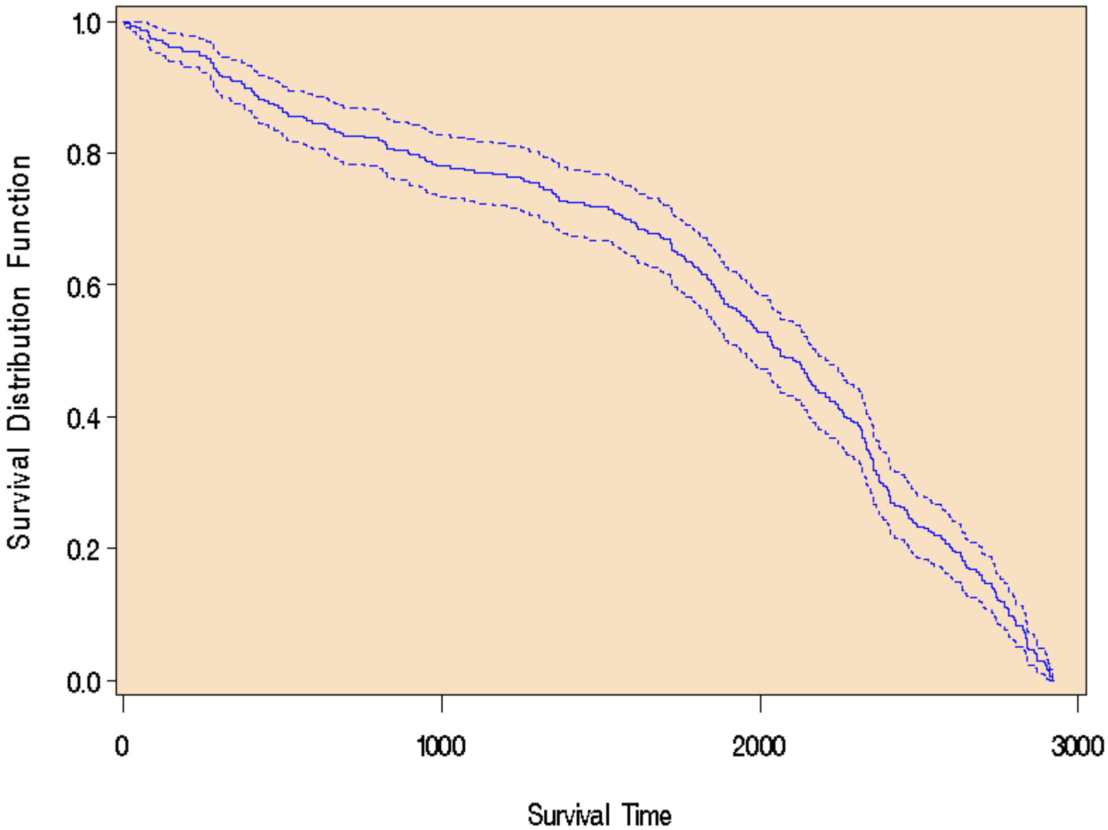


Figure 3. Survival function for pre-HACCP period, all Class 1 recalls, 1994-Jan. 25, 1998.

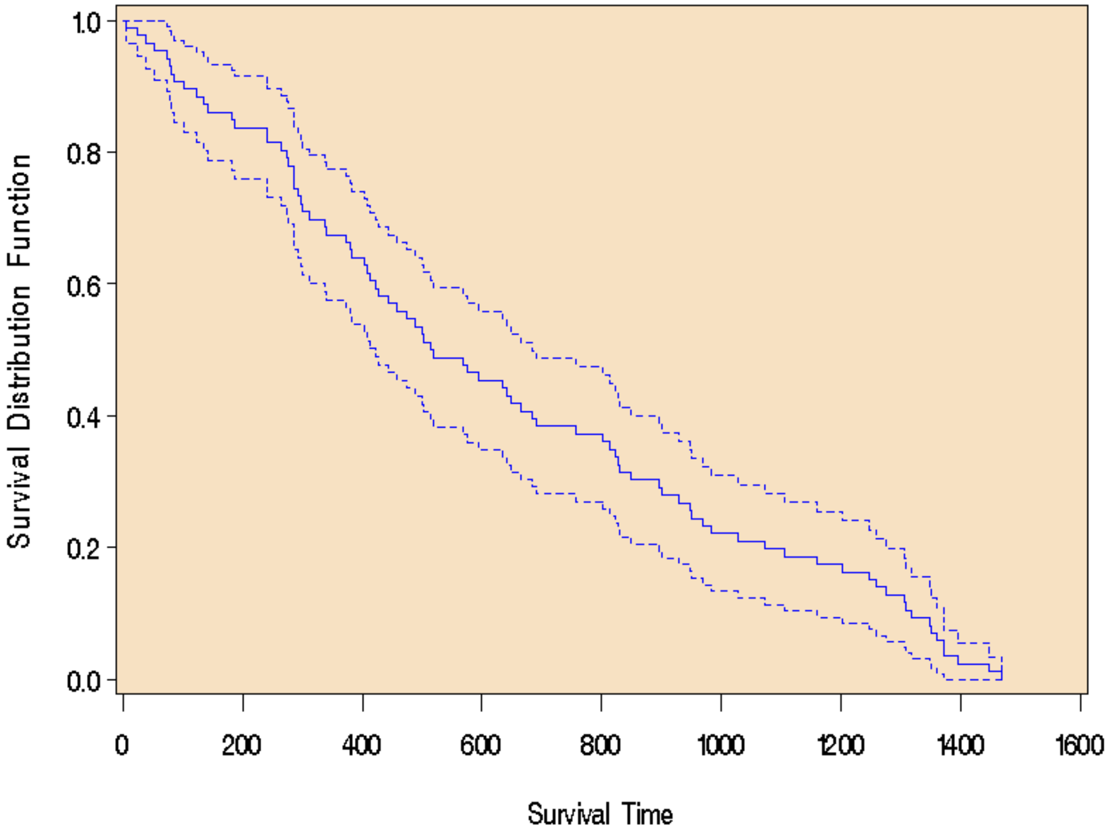


Figure 4. Survival function for post-HACCP period, all Class 1 recalls, Jan. 26,1998-2001.

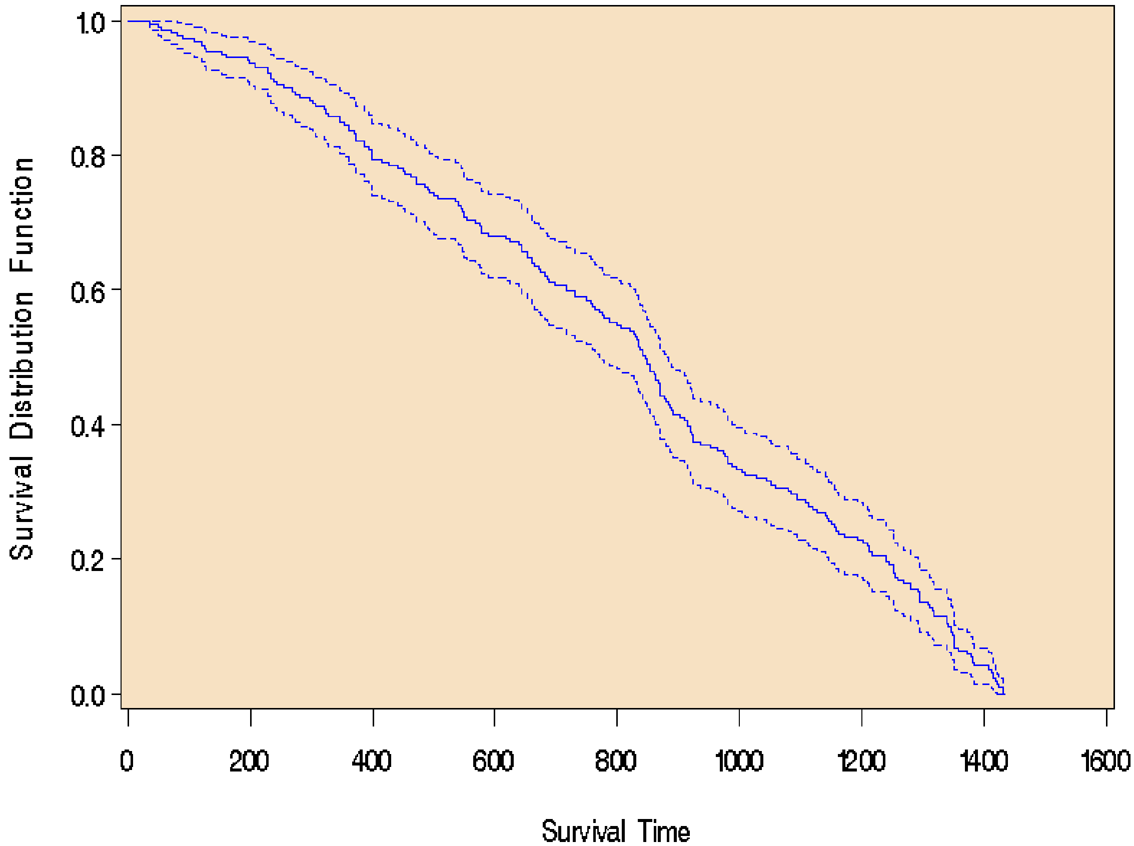


Figure 5. Survival function for large firms, pre-HACCP period, 1994-Jan. 25, 1998.

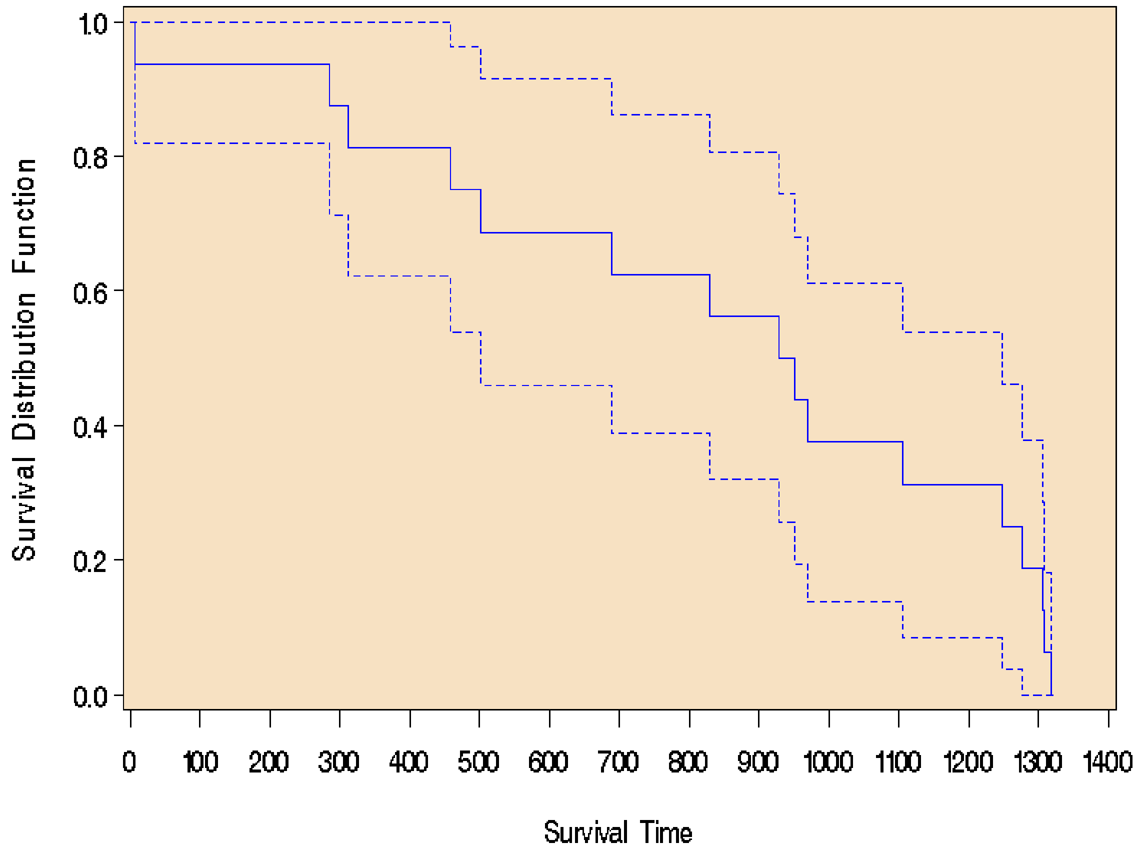


Figure 6. Survival function for large firms, post-HACCP period, Jan. 26, 1998 – 2001.

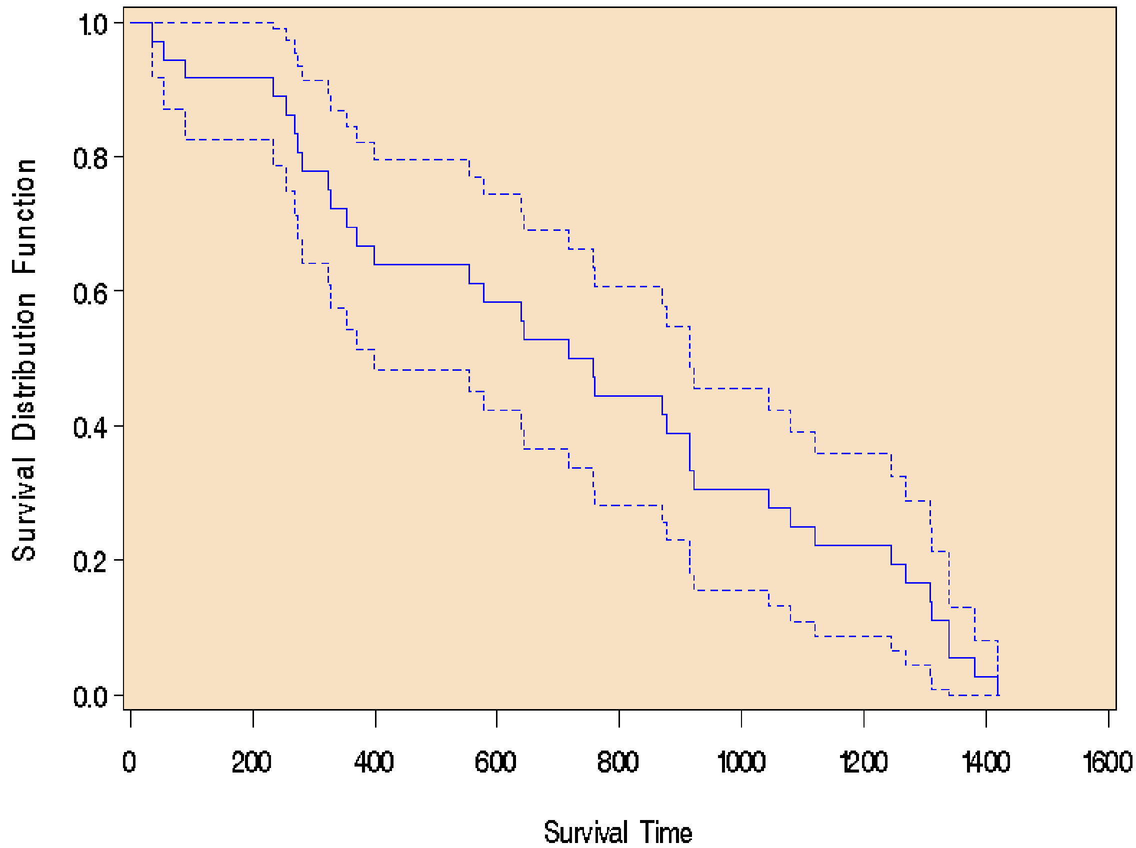


Figure 7. Hazard function for pre-HACCP period, all Class 1 recalls, 1994-Jan. 25, 1998.

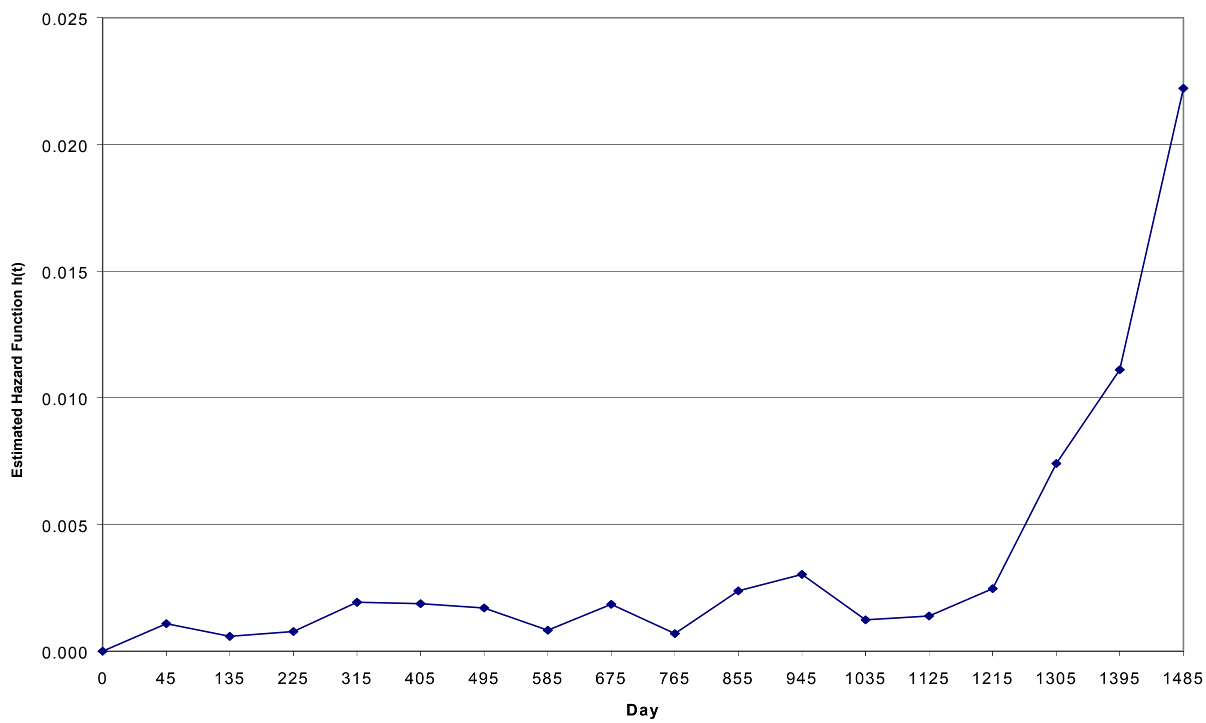


Figure 8.Hazard function for post-HACCP period, all Class 1 recalls, Jan. 26, 1998-2001.

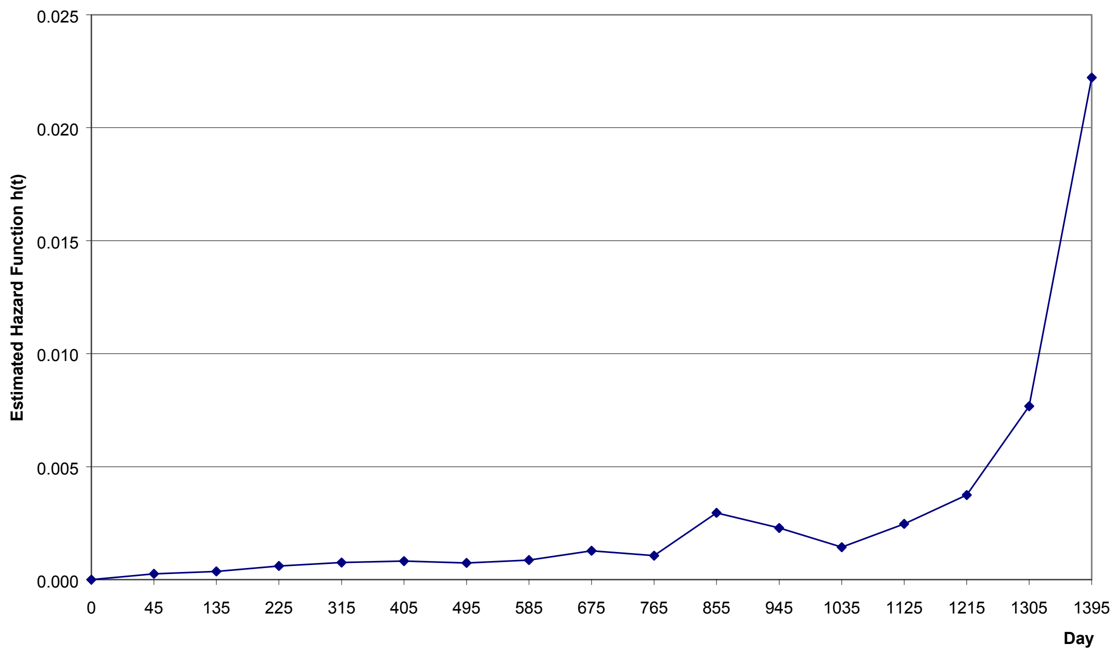


Figure 9. Hazard functions for all Class 1 recalls, 1994-2001, by year.

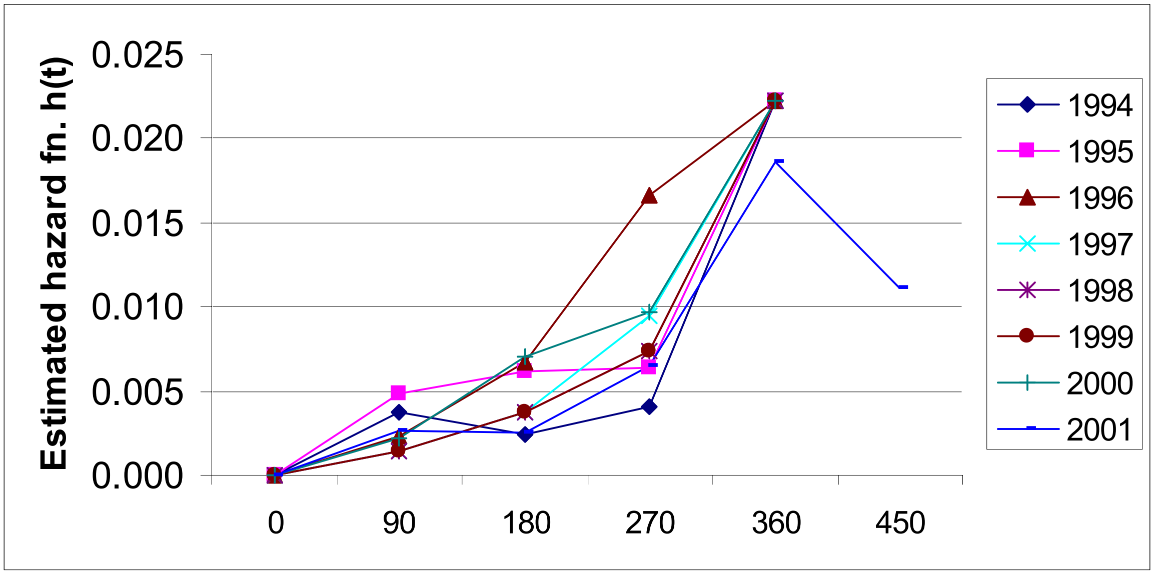


Figure 10. Parametric distributions fit to survival times for pre-HACCP period, all Class 1 recalls, 1994-Jan. 25, 1998.

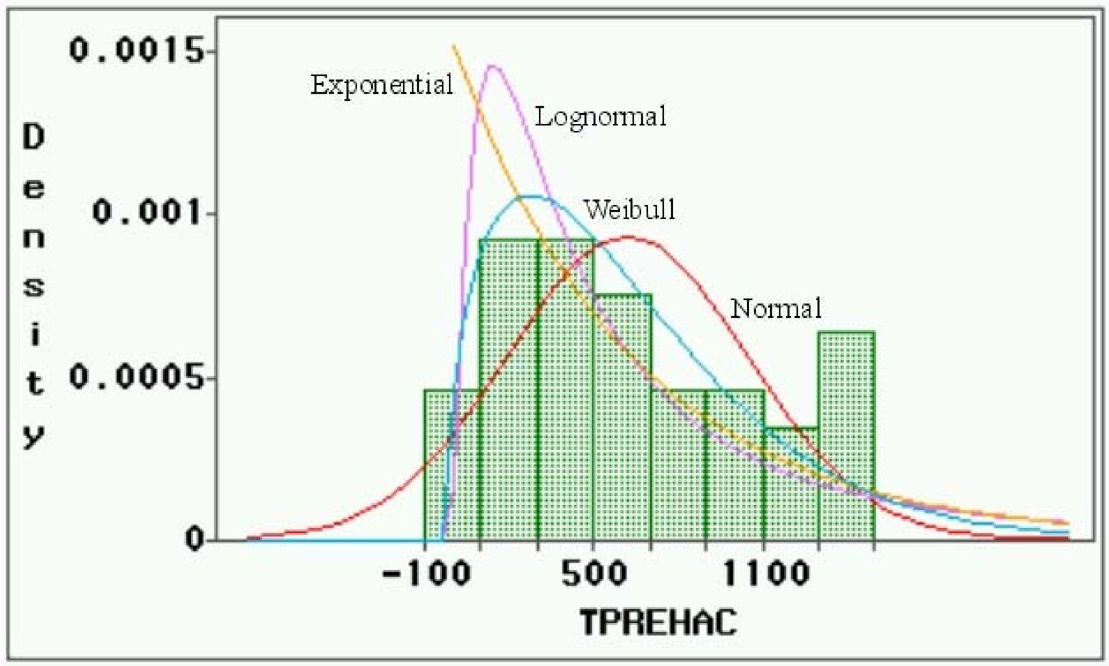


Figure 11. Survival function from fitted Weibull distribution compared with non-parametric estimation, for Class 1 recalls by large firms, 1994-Jan. 25, 1998.

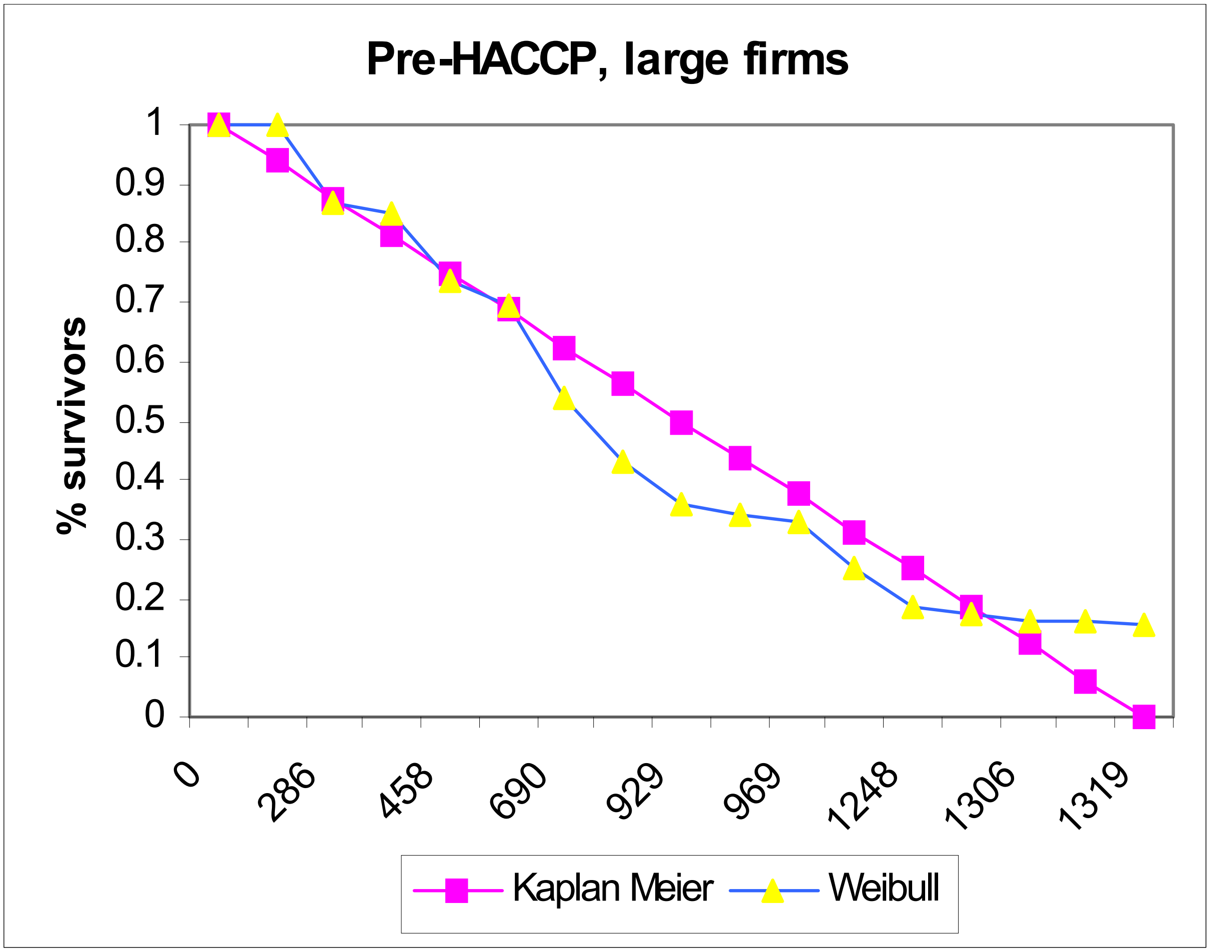
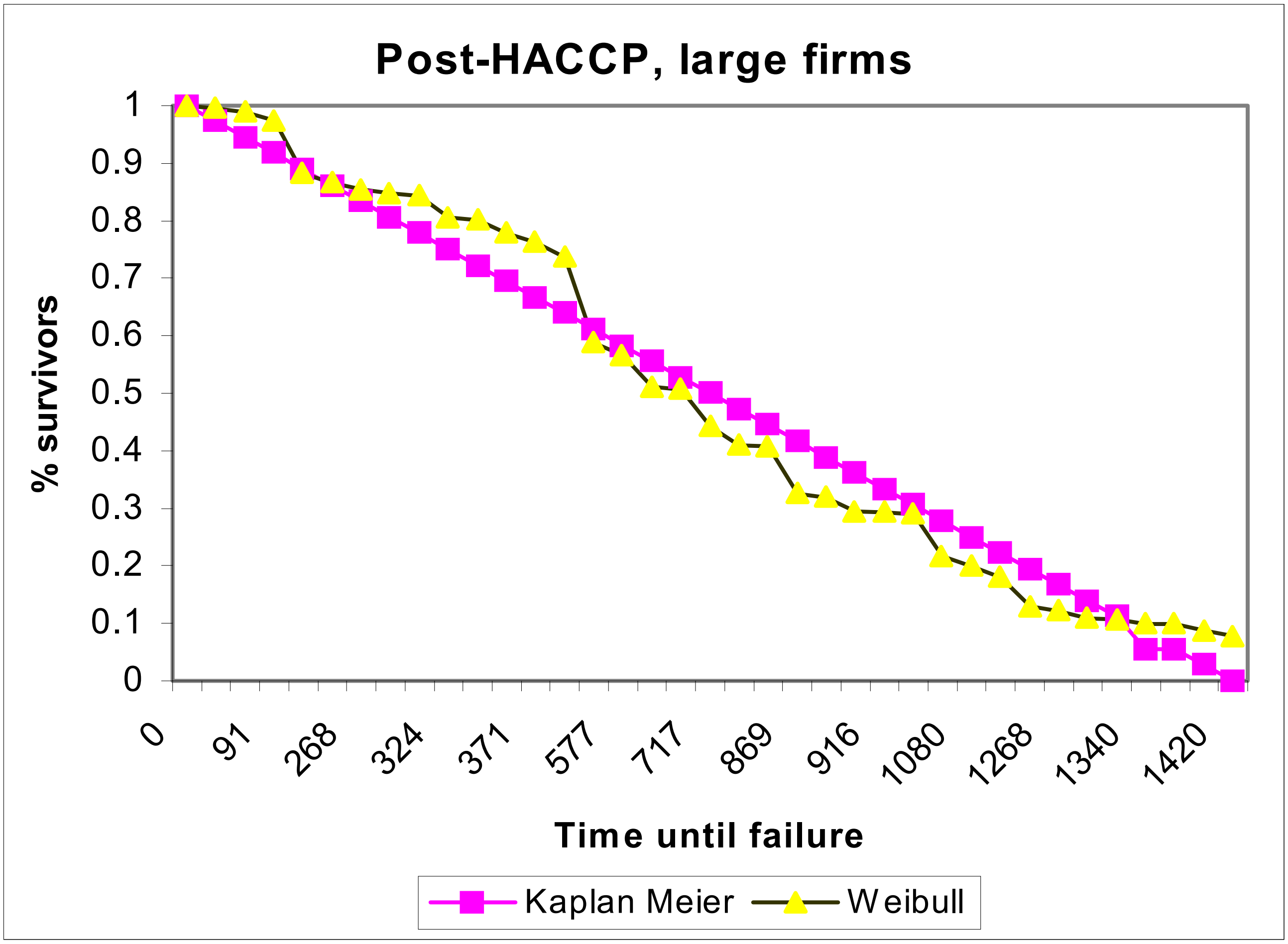


Figure 12. Survival function from fitted Weibull distribution compared with non-parametric estimation, for Class 1 recalls by large firms, Jan. 26, 1998-2001.



References

Agarwal, Rajshree. “Survival of Firms over the Product Life Cycle,” *Southern Econ. Jnl.,* 63 (1997), Jan. p. 571-584.

Antle, John M. “Benefits and Costs of Food Safety Regulation,” *Food Policy*, 24 (1999) 605-623.

Greene, William H. *Econometric Analysis*, 2nd edition. MacMillan: New York, 1993.

Henson, Spencer, & Mario Mazzuchi. "Impact of Bovine Spongiform encephalopathy on Agribusiness in the United Kingdon: Results of an Event Study of Equity Prices," Am. Jnl. Of Agric. Econ., 84 (2) (May 2002): 370-386.

Lee, Elisa T. *Statistical Methods for Survival Data Analysis,* 2nd edition. Wiley: New York, 1992.

Mahmood, Talat. “Does the Hazard Rate for New Plants Vary Between Low and High-Tech Industries,” *Small Business Economics,* 4 (1992):201-209.

Pitman, Jim. *Probability*. Springer Verlag: New York, 1993.

Ross, Tom, Paw Dalgaard, and Suwunna Tienungoon. “Predictive Modelling of the Growth and Survival of *Listeria* in Fishery Products,” *Intl Jnl of Food Microbiology,* 62(2000), pp. 231-245.

Salin, Victoria, and Neal H. Hooker. “Stock Market Reaction to Food Recalls,” *Rev. of Agricultural Econ.*, 23 (2001):1, Spring/Summer, pp. 33-46.

Teratanavat, Ratapol and Neal H. Hooker. “Meat and Poultry Recalls: Trend Analysis and Implications,” paper presented at the 3rd International Food Safety Conference, 18th National Conference of Food Microbiology, Hygiene and Toxicology, Guadalajara, Mexico, Nov. 9-10, 2001.

Thomsen, Michael R., and Andrew M. McKenzie. “Market Incentives for Safe Foods: An Examination of Shareholder Losses from Meat and Poultry Recalls,” *Amer J of Agricultural Econ,* 82:3 (Aug), pp. 526-538, 2001.

Wang, Zijun, Victoria Salin, Neal H. Hooker, and David Leatham. “Stock Market Reaction to Food Recalls: A GARCH Application,” forthcoming in *Applied Economics Letters*.