Behind the Diffusion Curve: An Analysis of ATM Adoption*

SUNIL SHARMA†

Department of Economics, UCLA
and
International Monetary Fund

First draft September 1991
Revised September 1992

WORKING PAPER #686
DEPARTMENT OF ECONOMICS
UNIVERSITY OF CALIFORNIA, LOS ANGELES
JANUARY 1993

---

* I would like to thank Neil Ericsson, Marnik Dekimpe, Harold Demsetz, Keunkwan Ryu, and Rajiv Sinha for discussions, and Sven Thommessen for superb research assistance. The paper has also benefitted from seminars at the Division of Monetary Affairs of the Federal Reserve Board and the UCLA Business Economics workshop.

† Address: International Monetary Fund, European I Department, 700 19th Street N.W., Washington, D.C. 20431.

Abstract

The paper examines the impact of firm characteristics, market structure and state regulations on the adoption of automated teller machines (ATMs) by banking organizations. A grouped duration data framework is used to investigate the effect of these factors on the hazard rate of adoption. The analysis shows that larger firms, especially those owned by bank holding companies, were earlier adopters. At any given time, the proportion of prior adopters in a market increased the likelihood that a non-adopter would do so. Market concentration increased the hazard rate if only a small proportion of firms in that market were using ATMs. Firms operating in urban, high wage markets and experiencing relatively rapid growth of deposits, especially demand deposits, were more likely to introduce ATMs. The mandatory sharing of ATM systems required by some states had a negative, though not significant impact. However, the conditional probability of adoption was higher in states where branching was either prohibited or restricted and off-premise ATMs were allowed. The study also shows that the effect of the explanatory variables changed across different phases of ATM diffusion. Further, it is demonstrated that not accounting for the grouping of the data and/or the use of restrictive parametric hazards, as in some past analyses, affects both the estimated coefficients and their standard errors, and may even lead to incorrect qualitative conclusions regarding the impact of certain covariates.

KEY WORDS: Duration Models, Grouped data, Time-varying parameters, Diffusion of technology, Automated Teller Machines.

1 Introduction

Technological developments in Electronic Fund Transfer (EFT) systems since the mid-1960s have revolutionized the nature of banking operations. Automated clearinghouses, regional and nationwide automatic teller machine (ATM) systems, and developments in credit authorization and homebanking products have, besides making retail banking more attractive to commercial banks, transformed the the way transactions are conducted. Since their appearance in 1969, ATMs have been one of the most successful EFT services. This paper examines the micro-aspects of ATM diffusion in the US with a view to evaluating the impact of firm characteristics, market structure and state regulations on the adoption of this new banking technology.

The analysis shows that larger firms, especially those owned by bank holding companies, were earlier adopters. At any given time, the proportion of prior adopters in a market increased the likelihood that a non-adopter would do so. Market concentration increased the hazard rate if only a small proportion of firms in that market were using ATMs. Firms operating in urban, high wage markets and experiencing relatively rapid growth of deposits, especially demand deposits, were more likely to introduce ATMs. The mandatory sharing of ATM systems required by some states had a negative, though not significant impact. However, the conditional probability of adoption was higher in states where branching was either prohibited or restricted and off-premise ATMs were allowed. It is also shown that the effect of many of the covariates changed across different phases of ATM diffusion.

The hazard specification for analyzing the time to first adoption of ATMs is based on Prentice and Gloeckler (1978). The empirical model can be interpreted as a natural generalization of the probit/logit specifications used in the vast multi-disciplinary literature on adoption and diffusion of technical change (see, Rogers (1983), Thirtle and Ruttan (1987), Feder, Just and Zilberman (1985) for surveys). It allows for the grouped nature of the available data, time variation in the explanatory variables, and also the possibility that the impact of the explanatory variables may change over the course of technology diffusion. The baseline hazard is constant over each (grouping) interval but may differ across intervals and represents a step function approximation to the underlying true hazard. The results demonstrate that failure to account for grouping of the data and/or the use of restrictive baseline hazards, as in some past analyses, dramatically affects coefficient estimates and their standard errors, and may even lead to incorrect qualitative conclusions regarding the impact of certain covariates.

Hannan and Mcdowell (1984,1987), Karshenas and Stoneman (1990), Rose and Joskow (1990) and Saloner and Shepard (1991) utilize a duration modeling framework to examine technology adoption. While these empirical studies represent advances in modelling diffusion, they use continuous time approximations and/or restrctive parametric hazards to estimate models that are based upon grouped duration data. When the grouping intervals are very small compared to the rate of event occurrence it may be acceptable to ignore the discrete nature of the data and treat them as if they were truly continuous. However, when the intervals are large such an approximation may not be justified. Further, the partial likelihood method of Cox (1975) is difficult to implement in the presence of large time intervals due to the occurrence of ties in the data and may even be inappropriate (see Farewell and Prentice 1980; Prentice and Gloeckler 1978). Although methods to handle ties in the data have been developed in the literature (see, for example, Breslow 1974) they are generally ad hoc in nature and/or yield approximate likelihoods that are computationally not feasible. An additional problem with the diffusion studies mentioned above is that they make no attempt to control for the effects of unobserved heterogeneity. This may be a serious problem since they use restrictive parametric baseline hazard functions to model the timing of adoption.

ATMs have been introduced by banks for many reasons: (1) to increase their share of the retail banking market and to attract new customers by offering more flexible and convenient services. It was hoped that ATMs would lead to higher levels of individual account balances and make small credit loans more easily available; (2) it was envisaged that these machines could perform many deposit, withdrawal and transfer operations at lower cost than human tellers. Further, they could act as surrogate branches and decrease the number of hours the regular branches needed to be open; (3) they could be used for marketing purposes to test the demand for services in a particular area before a regular branch was established. On the cost side, besides the expense of setting up and maintaining an ATM system (or obtaining access to one), the banks have to deal with problems of malfunction, fraud, robbery and vandalism (see Baker and Brandel (1988)).

ATM systems typically require high-capital investments and have high fixed costs (Baker and Brandel (1988)). To a large extent such investment expenditure is a sunk cost and hence irreversible; further, these expenditures can be delayed allowing the firm to accumulate more information about costs, benefits and market conditions before committing resources. Pindyck (1991) is an excellent survey of models of irreversible investment. He argues that the nature of such investments invalidates the usual net present value rule and that irreversibility "makes investments especially sensitive to various forms of risk, such as uncertainty over future product prices and operating costs that determine cash flows, uncertainty over future interest rates, and uncertainty over the cost and timing of the investment itself." For example, see Favero, Pesaran and Sharma (1991) where an irreversible investment model is developed to examine the decision of oil companies to tap known oil reserves, and a duration framework is used to empirically analyze the time lag between discovery and exploitation of oil fields on the UK Continental Shelf.

The theoretical literature on technology diffusion argues that potential adopters of a new innovation in a market (or industry) will generally have different adoption dates. At each point in time, firms evaluate the innovation's profitability (i.e. compare the expected streams of benefits and costs, factoring in the associated uncertainties) and make a decision regarding adoption. The heterogeneity in the adoption dates reflects different valuations of the technology relative to costs (see, David (1969), Davies (1979), Mansfield (1961,1968) and the surveys by Reinganum (1989), Stoneman(1983) and Thirtle and Ruttan (1987)). This may be due to differing firm characteristics (for example, size, managerial willingness to take risks, product mix, organizational factors, information about the technology), market structure (for example, concentration and number of firms in the market) and state regulations pertaining to the use of the new technology.

In the next section, a general duration model is specified to analyze the impact of the various factors mentioned above on the adoption of ATMs. The third section contains the data description and the fourth section discusses estimation results. The last section presents some caveats and concluding remarks.

2 Model Specification

There are two general approaches to handling the grouped nature of observed duration data. The simplest is to treat time as if it were truly discrete. An alternative approach is to use a continuous time model taking account of the fact that the data on the duration of an event is grouped into intervals. This approach put forward by Prentice and Gloeckler (1978) and Thompson (1977) in the biometrics literature, has been recently used in economics by Kiefer (1988), McCall (1990), Meyer (1990), and Sueyoshi (1991a,b). While both these methods may at times yield similar results, the former method leads to inferences that are sensitive to choice of (grouping) interval length, whereas the latter does not suffer from this defect. The specification given below uses the latter approach.

Consider the case when time to adoption of a new technology is grouped into intervals $[t_{k-1}, t_k), k = 1, 2, \ldots, m+1, t_0 = 0$ and $t_{m+1} = \infty$, with adoption in the interval $[t_{k-1}, t_k)$ recorded as t_k. These intervals are the same for each firm and t_{k_i} means that firm $i, i = 1, 2, \ldots, N$, adopted the technology in interval $k_i \in 1, 2, \ldots, m+1$. In addition, information is available on a $1 \times \ell$ vector of covariates $X_i(k), i = 1, 2, \ldots, N, k = 1, 2, \ldots, m+1$, for each firm in each interval of time. These covariates, comprising both firm and market characteristics, can vary over the intervals but are assumed fixed within each interval.

For notational simplicity, in the rest of the paper we will quite often drop the dependence of the covariate vector on time, it being understood that X_i changes across time intervals.

Let the underlying density of durations conditional on regressors X_i and unknown parameters θ be given by $f_i(t, X_i, \theta)$, with associated hazard function $\lambda(t, X_i, \theta)$. The probability of firm i not adopting in the k^{th} interval given that it did not adopt in the first $(k-1)$ intervals is given by

$$\begin{aligned} \Pr(t \geq t_k | t \geq t_{k-1}; X_i, \theta) &= \frac{S(t_k, X_i, \theta)}{S(t_{k-1}, X_i, \theta)} \\ &= exp\left[\int_{t_{k-1}}^{t_k} \lambda(u, X_i, \theta)\, du\right] = \alpha_k(X_i, \theta) \end{aligned} \tag{1}$$

where $S(t, X_i, \theta)$ is the survivor function. For a completed observation, the probability of firm i adopting in the k^{th} interval $[t_{k-1}, t_k)$ is

$$\begin{aligned} &\left[1 - \frac{S(t_k, X_i, \theta)}{S(t_{k-1}, X_i, \theta)}\right] \prod_{j=1}^{k-1} \left[\frac{S(t_j, X_i, \theta)}{S(t_{j-1}, X_i, \theta)}\right] \\ &= [1 - \alpha_k(X_i, \theta)] \prod_{j=1}^{k-1} [\alpha_j(X_i, \theta)] \end{aligned} \tag{2}$$

The likelihood for a sample of N firms is the product over i of terms in (2). Censoring in the data is easily handled by introducing an indicator variable δ_i that equals one if the observation is complete and zero if it is censored. Besides assuming that the censoring mechanism is independent of the parameters θ of interest, we make the assumption that censoring, if it occurs, does so at the beginning of the last interval of observation. Clearly some such assumption is required given that adoption of ATMs are grouped into annual intervals. The log-likelihood for the data $(t_i, \delta_i, X_i), i = 1, \ldots, N$ is given by

$$L(\theta) = \sum_{i=1}^{N} \left\{ \delta_i\, ln\, [1 - \alpha_{k_i}(X_i, \theta)] + \sum_{j=1}^{k_i - 1} ln\, [\alpha_j(X_i, \theta)] \right\} \tag{3}$$

Kiefer (1988) examines interesting links between the grouped duration data models and the ordered discrete choice framework. These relationships provide insights into estimation, testing and interpretation of such models (see Sueyoshi (1991a,b) for discussion and extensions).

The time dependence of the hazard and the impact of covariates are embodied in specifications for α_k. I use the Cox (1972) specification for the hazard function

$$\lambda(t, X_i, \theta) = \lambda_0(t)\, exp(X_i(t)\, \beta) \tag{4}$$

where $\lambda_0(t)$ is the baseline hazard and the β vector a subset of θ. Note, this is strictly speaking a proportional hazard only if the covariates are the same function of time for all firms. The conditional probability of surviving interval k can now be written as

$$\alpha_k(X_i, \theta) = exp\{-exp(X_i\beta + \gamma_k)\} \tag{5}$$

where $\gamma_k = \int_{t_{k-1}}^{t_k} \lambda_0(u)\, du$ and $\theta = (\beta', \gamma')$. The above specification leads to the following grouped data log-likelihood

$$L(\theta) = \sum_{i=1}^{N} \left\{ \delta_i\, ln[1 - exp\{-exp(X_i\beta + \gamma_k)\}] - \sum_{j=1}^{k_i - 1} exp(X_i\beta + \gamma_j) \right\} \tag{6}$$

Model estimation involves numerically solving the likelihood equations formed by setting the score vector equal to zero. With fixed time intervals, only mild restrictions are required on the censoring mechanism and covariate vectors for the score to have an asymptotic normal distribution (see Kalbfleisch and Prentice 1980). The estimated γs allow us to calculate a step function approximation to the underlying baseline hazard.

Note that Han and Hausman (1990) specify the hazard function as

$$ln\left[\int_0^{t_i} \lambda_0(u)\, du\right] = X_i\beta + \epsilon_i \tag{7}$$

where assuming ϵ_i takes the extreme-value distribution leads to an ordered-logit specification. They put $ln\,[\int_0^{t_k} \lambda_0(\tau)\, d\tau] = \delta_k$ and estimate these constants along with the β coefficients. It is easy to show that these δ_k are related to the γs used above by the relation $\delta_k = ln\,[\sum_{j=1}^{k} exp(\gamma_j)]$ and that the Han and Hausman approach in this case merely leads to a reparametrization of the likelihood in (6).

The above approach is "semi-parametric" in the sense that the effect of covariates takes a particular functional form while the baseline hazard is specified in a non-parametric fashion. The estimation of the model is based only upon variations in the covariates across observations. The temporal variation in the mean of the covariates is absorbed by the baseline hazard. Consequently, in the likely situation that the explanatory variables differ more across observations than over time, the semi-parametric model insures consistency at the expense of a small loss in efficiency (see Meyer (1990)). This is a powerful feature of this approach, given the lack of a priori knowledge about the true form of the baseline hazard. A parametric form for the baseline hazard would provide inconsistent estimates if the assumed specification is incorrect. Also, the baseline hazard implied by the γ parameters is easily examined graphically. For example, if the Weibull hazard is an appropriate specification, the natural logarithm of the integrated hazard $ln[\sum_{i\leq k} exp(\gamma_i)]$ plotted against $ln(t_k)$ should yield a straight line.

The hazard rate conditional on unobserved heterogeneity ζ is factored as

$$\lambda(t, X_i, \theta \mid \zeta) = \zeta \, \lambda_0(t) \, exp(X_i \beta) \tag{8}$$

A convenient and frequently used distribution for ζ is the gamma since it yields a closed form expression for the likelihood (Meyer (1990)). If ζ is distributed gamma with mean one (an innocuous normalization) and variance σ^2, then the log-likelihood function becomes

$$L^*(\theta, \sigma^2) = \sum_{i=1}^{N} \left\{ \left[1 + \sigma^2 \sum_{j=1}^{k_i - 1} exp(X_i(j)\, \beta + \gamma_j) \right]^{\frac{-1}{\sigma^2}} - \delta_i \left[1 + \sigma^2 \sum_{j=1}^{k_i} exp(X_i(j)\, \beta + \gamma_j) \right]^{\frac{-1}{\sigma^2}} \right\} \tag{9}$$

It is likely that the effects of different firm and market characteristics on the probability of adopting a new technology change over the course of the diffusion process. The above hazard framework is flexible enough to allow for some or all β coefficients to vary across intervals (see section 4.2 below). The time structure of the β parameters can be evaluated using likelihood ratio or Lagrange multiplier tests (Sueyoshi (1991)).

Heckman and Singer (1984) have criticized the use of parametric heterogeneity in the context of parametric hazard specifications on the grounds that the coefficient estimates may be very sensitive to the choice of a mixing distribution. They advocate instead, the use of non-parametric specifications for heterogeneity. However, recent research indicates that specification of the mixing distribution is not as important whereas an appropriate (flexible) choice for the base-line hazard is crucial and the estimates of the parameters are typically more sensitive to the latter. For instance, Manton, Stallard and Vaupel (1986) find that the specification of the baseline hazard function is far more critical for estimation than is the specification of the heterogeneity distribution (see, also Han and Hausman (1990), Ridder(1986) and the discussion in Lancaster (1990)).

3 Data

The empirical analysis uses data pertaining to the adoption of ATM's by a sample of 3,689 individual banking firms operating in either an SMSA or a county which has been judged to approximate a local banking market. From the extensive population surveys conducted by the Federal Deposit Insurance Corporation (FDIC), we have information on the timing of adoption of the first ATM (by year) for the period 1971-1979, which represented the first nine years of ATM usage by banking firms. All 3,689 banks in the sample were in existence for the entire 1971-1979 period and 739 of these banks had adopted ATMs by the end of 1979. This represents an adoption proportion of 20%, as compared to an overall rate of 12% reported for the entire national population of 14,314 insured commercial banks in 1979. [My sample is slightly smaller than that used by Hannan and Mcdowell (1987). They had data on 3834 firms of which 750 had adopted ATMs by the end of 1979. Despite best efforts the information on the firms representing the difference in the samples was not retrievable.]

In addition to year of ATM adoption, the dataset contains detailed information on many firm and market/environment characteristics that may be relevant for a firm-level analysis of ATM adoption. The following (time-varying) covariates are used in the study :

SIZE firm size as measured by total assets (in millions of dollars)

GROWTH annual growth in market deposits in the previous year

MIX product mix, defined as ratio of each bank's demand deposits to its total deposits

BHC dummy variable indicating ownership by a bank holding company

URBAN dummy variable indicating operation in an SMSA

WAGE prevailing hourly market wage rate for each year

ATMSHAR a dummy variable to indicate operation in a state which requires some form of sharing of ATM systems

OFFPRM a dummy variable to indicate operation in a state where branching is either prohibited or restricted but where off-premise ATM's are allowed

CR the three-firm concentration ratio, measured as the proportion of total market deposits accounted by the three largest banks

PROPN the proportion of banks headquartered in the market that had introduced an ATM system as of the end of the previous year. Note that the variable PROPN refers to proportion of firms in a specific banking market not to the industry as a whole.

Interstate banking in the US is restricted and commercial bank operations are geographically constrained. Hence, the above data set documents a convenient experiment which enables us to examine the adoption decision by different firms operating under different market conditions *in the same industry.* One drawback is that data on supply-side factors (e.g prices, changes in technology) is not available.

The pattern of adoptions is shown in table 1 and the empirical hazard function which is the fraction of non-adopters at the beginning of a year that adopt ATM technology in that year, is displayed in figure 1. The empirical distribution function and it's 95% confidence band are depicted in figure 2.

Table 2 contains the time-varying mean values for the firm covariates. Table 2(A) presents them for all the 3689 banks in the dataset, where as Table 2(B) does the same for those banks who had *not* adopted ATMs by the beginning of a particular year. This summary itself is revealing. For example, a comparison of the values for SIZE over the period suggests that the larger banks were earlier adopters since the mean values in Table 2(A) are greater than the corresponding entries in Table 2(B).

4 Results

4.1 Time-Invariant Coefficients

The estimates of the β coefficients obtained from maximizing the likelihoods in (6) and (9) are presented in tables 3 and 4 respectively. It is quite striking that the coefficient estimates and the likelihood values in these two tables are virtually identical. The only difference is that taking account of unobserved heterogeneity leads to the same or lower standard errors for *all* covariates in table 4. The results show that besides ATMSHAR, the dummy for operation in states which mandated some form of sharing of ATM systems, all the other included covariates had statistically significant impacts. Model 5 reveals that the coefficients of size, the wage rate, dummy for whether the bank was owned by a bank holding company, the concentration index, proportion of prior adopters and the interaction of the last two factors are significant at the 1% level, and the coefficients of the others are significant at the 5% level.

The effects of firm size and concentration are decidedly Schumpeterian. Much of the theoretical and empirical work on technology diffusion at a disaggregated level suggests that firm size is a critical factor as it serves as a proxy for risk aversion, economies of scale and research and development activities. However, empirical results on the question of whether larger firms are more "innovative" than their smaller counterparts (i.e., early adopters) have been mixed (see Thirtle and Ruttan (1987)) . While these differences in results may be attributed to alternative indicators of firm size (e.g., number of employees versus total assets) in addition to methodological differences, an increasing amount of evidence suggests that these results will differ depending upon the innovation under consideration. The results here indicate that both firm size and the concentration index have a positive impact on the hazard. Further, differences in the structure of banking institutions are important - subsidiaries of holding companies as opposed to "independent" banking firms had a significantly higher conditional probability of adoption. Firms with a higher proportion of demand deposits to total deposits, operating in urban markets experiencing faster growth were early adopters of ATMs. Banks that have a high value of the MIX variable are liable to benefit more from the introduction of a flexible retail banking technology. The urban dummy is possibly a proxy for greater demand for ATM services as well as less resistance to new ways of doing things.

A key reason for using ATMs is that they can perform a wide variety of retail banking services at lower cost than human tellers. Clearly, this substitution effect will produce greater cost savings in high wage banking markets. The positive and highly significant coefficient on WAGE is in accord with the above reasoning. This is in contrast to the Hannan and McDowell (1987) result where the wage rate had a negative influence on the hazard rate.

As noted earlier the variable PROPN refers to prior adopters in a specific banking market in which a bank is headquartered. It captures more the strategic interaction among competing firms, and in our context should not be interpreted as a proxy for the usual communicaton and informational effects that result as the use of a technology spreads in the industry. In our empirical specification the latter effects are captured by the base-line hazard. The estimates indicate that PROPN is an important explanatory variable and for a particular firm the conditional probability of introducing ATMs increases as it's competitors adopt the technology. The interaction between PROPN and CR reveals that the concentration index has a positive effect on the hazard rate if the proportion of prior adopters in the market is less than 0.24(=1.814/7.62). For larger values of PROPN, the concentration index has a negative effect. [Note that this result is similar to that obtained by Hannan and Mcdowell (1987) although it is based on different values for the coefficients]. Similarly, the interaction between PROPN and ATMSHAR suggests that for PROPN $\leq$ 0.24(=0.207/0.848) requiring banks to share their ATM systems has a negative impact on the hazard since part of the first mover advantage is eroded [see Fudenberg and Tirole (1985)].

Government regulations affect the adoption and use of new technologies. The positive coefficient on OFFPRM implies that firms operating in states where branching is prohibited attempt to circumvent this restriction by the use of ATMs. Also, the mandated sharing of ATM systems in some markets slowed down the rate of diffusion and especially (as argued above) had a negative effect on the earlier adopters.

Table 5 reports the estimates from commonly used parametric hazard models, both when the annual grouping of the data is taken into account and when this feature of the data is ignored. The exponential models give similar estimates of the coefficients for almost all the covariates. However, this is not true for the Weibull models. A comparison of models 9 and 10 shows that the coefficients of five covariates - GROWTH, URBAN, WAGE, OFFPRM, PROPN - and the estimate of the duration dependence parameter change quite substantially. In fact the effect of PROPN more than doubles, and that of GROWTH and WAGE reverse their sign. [Notice that model 7 is nested in model 9, and model 8 is nested in model 10. The results indicate that in terms of log-likelihood values the Weibull models do better. However, GROWTH has negative coefficients in the exponential models but positive ones in the Weibull models; WAGE has a positive, highly significant effect in models 7 and 8, a much smaller but still significant positive effect in model 10, but a negative effect under model 9.]

A comparison of model 2 (non-parametric base-line hazard) with model 8 (exponential accounting for grouping) shows that the coefficients of GROWTH, URBAN, WAGE, ATMSHAR, OFFPRM, and PROPN are substantially different; model 10 (Weibull accounting for grouping fares slightly better in such a comparison with the effects of GROWTH, ATMSHAR, and OFFPRM showing dramatic differences. The results clearly indicate that not accounting for the grouping of the adoption data and/or using parametric specifications for the base-line hazard may lead to severely biased estimates of the coefficients.

The baseline hazard may be interpreted as capturing the effect of time after correcting for covariates. To the extent that we do not include supply side factors (like price changes, or introduction of new generation of machines), the baseline hazard also partly reflects the impact of such factors. Table 6 contains the estimates of the γ parameters from some of the models. The baseline hazard calculated from model 5 and it's 95% confidence band is also shown. Again note that the effect of allowing for unobserved heterogeneity in model 5 as compared to model 2 is to decrease the standard errors while leaving the coefficients virtually unchanged.

The estimated baseline hazard function from model 5 is plotted in figure 3. The diagram begs the question: Why was there an increase in the hazard rate in 1976? One possible explanation is that around 1974-1976 the retail banking capabilities of ATM systems dramatically improved and "since 1976, in addition to dispensing cash, ATMs have performed loan transactions, deposits to any account, transfers between accounts.

and payments for mortgages and other debts" [see Baker and Brandel, Ch 6, pg 10]. Also, this increase in adoptions took place despite an increase in the real price of an ATM in 1976. An index for real ATM prices calculated by taking the average price quoted by the four largest ATM companies (source: various issues of The Magazine of Bank Administration) deflated by the CPI, shows that if 1971 is taken as the base year, the index was 62.2 in 1975, 67.7 in 1976, 63.7 in 1977 and 43.7 in 1980. Further, 1976 was the only year in which this index increased over the preceding year.

For comparison the estimates of the baseline hazard from the parametric specifications and (semi-parametric) model 5 are also displayed in figure 3. The parametric forms uniformly overestimate the impact of time on the conditional probabilities of adoption. A graphical analysis of the integrated baseline hazard from model 5 can also be quite revealing. Under the exponential model ($\gamma_1 = \gamma_2 = \ldots = \gamma_9$) a plot of the integrated baseline hazard, $\sum_{i \leq k} exp(\gamma_i)$, versus t_k should give a straight line through the origin. Similarly, the Weibull model imposes the constraints $\gamma_k = constant + ln\{t_{k-1}{}^{\rho} - t_k{}^{\rho}\}$ on the γ parameters. A graph of the natural logarithm of the integrated hazard $ln[\sum_{i \leq k} exp(\gamma_i)]$ against $ln(t_k)$ should give points on a straight line [see Cox and Oakes (1984), Lancaster (1990)]. Figures 4 and 5 depict the above plots using estimates of the integrated baseline hazard. They clearly suggest that the exponential and Weibull assumptions are inappropriate for modelling the conditional probabilities of ATM adoption in this data set.

4.2 Time-varying Coefficients

The diffusion of a technology over time can be generally divided into three broad phases: Phase I when the "innovators" adopt, followed by Phase II in which the majority begins use of the technology, and lastly Phase III when the laggards adopt. Many diffusion studies attempt a categorization of the adopters (see, Thirtle and Ruttan) to isolate organizational characteristics that may help us to predict whether certain firms are likely to be early or late adopters. However, as far as I know there has been no systematic study of how firm characteristics, market structure and government regulations affect the adoption of a technology in *different* phases of the diffusion process.

An examination of Table 1 suggests that for ATMs, the adoption rate began to accelerate in 1979. Hence, a good prior guess is that Phase I was 1971-1978 and Phase II from 1979 onwards. In our empirical analysis this is captured by allowing the β coefficients to vary across phases. The model of section 3 was estimated with each year from 1973 to 1979 taken to be the dividing line between the first and second phases. The log-likelihood values obtained are presented below:

Loglikelihood values as Phase II beginning is varied							
Year	1973	1974	1975	1976	1977	1978	1979
Loglikelihood	-3180.3	-3176.3	-3173.7	-3175.5	-3169.1	-3175.2	-3168.2

Table 7 presents the results obtained with β coefficients varying over Phase I (1971-78) and Phase II (1979 onwards). A likelihood ratio test (yielding a value of 40.6 for a chi-square variate with 11 d.f.) rejects the hypothesis that the β coefficients are the same across the two phases. The estimates reveal that the impact of many of the covariates changed over the course of ATM diffusion. The size of the banking firm and the growth of deposits was much more important in the second phase of diffusion - the coefficient of these covariates are much larger (two and a half to three times larger) and statistically significant in the second phase.

Firms with a high proportion of demand deposits to total deposits operating in high wage urban banking markets, especially in states that had branching restriction (OFF-PREM=1), were the innovators. It is clear that these firm and market characteristics are the ones for which the value of ATM adoption is likely to be the highest. Further, these characteristics were not important in Phase II, implying that the reasons for adoption and associated calculations may be different as the technology is more widely used. The results also suggest that as expected, regulations regarding sharing of ATM systems are liable to have a greater impact in the initial stage of diffusion.

The concentration index had a relatively greater effect in the first phase which one can think of as the experimental period. Further, CR increases the hazard of adoption in Phase I if PROPN $\leq$ 0.25, and in Phase II if PROPN $\leq$ 0.13. The proportion of prior adopters in a specific market had a relatively greater impact in the second phase. This is what one would expect given that once the experimental period for the technology is over, and its benefits become more established and known, strategic considerations are liable to become more important. The estimates also indicate that PROPN increased the probability of adoption in Phase II, whereas this was true in Phase I only if CR $\leq$ 0.75.

Banking firms that are subsidiaries of holding companies were earlier adopters in both phases of diffusion, highlighting the fact that organizational factors may be of paramount importance and that further research using more detailed information on organizational structures is clearly warranted.

5 Concluding Remarks

This paper attempts to go behind the well-known S-shaped curve for the diffusion of technology to unravel the effect of firm characteristics and market structure. The observed adoption rate in the aggregate depends on the impact of time and the time-varying distribution of relevant firm and market variables in the population of potential adopters. The paper shows that even with low penetration of the technology (the adoption rate over the period covered was only 20% in our dataset) and only a year's information on the second phase of diffusion, behavioural hypotheses about adoption can be fruitfully examined. Many datasets record the adoption data using reasonably large (e.g. annual) intervals. It was demonstrated that it is important to use flexible baseline hazards, and that the coefficients of interest may be quite sensitive to whether one accounts for the grouping of data or not.

Clearly a more disaggregated analysis is required to answer questions like: Why larger firms were quicker adopters? Was there an "informational advantage" in being owned by a bank holding company? Or was some other organizational pecularity of bank holding companies the key factor? For example, it is possible that the large fixed costs involved in setting up an ATM system could only be profitably borne by big firms who benefitted from economies of scale. Sharing of ATM systems mandated by some states did not come into effect till 1975 and 19 states had such laws in 1979. This allowed easier and cheaper access to the technology for relatively smaller banks and such institutions may have adopted ATMs rapidly in the second phase of diffusion.

The US experienced an explosive growth in ATM installation starting in 1981. An estimated 50,000 of the 64,000 machines in use by the end of 1986 were installed in the period 1981-1986 [see Baker and Brandel (1988)]. The rate of adoption decelerated in the late 1980s and reached near saturation levels especially in certain urban areas. If the current data set could be augmented by information on the spurt and the subsequent levelling off in ATM adoptions one could perform an analysis of this new technology in all three phases of diffusion.

References

Baker, D. I. and R. E. Brandel R.E. (1988), *The Law of Electronic Fund Transfer Systems*, Warren, Gorham & Lamont, Inc.

Breslow, N. E. (1974), "Covariance Analysis of Censored Survival Data," *Biometrics*, 30, 89-100.

Cox, D.R. (1972), "Regression Models and Life-Tables," *Journal of the Royal Statistical* Society, B, 34, 187-202.

—— (1975), "Partial Likelihood," *Biometrika*, 62, 269-276.

—— and Oakes D. (1984), *Analysis of Survival Data*, Chapman and Hall, London and New York.

David, P. (1969), "A Contribution to the Theory of Diffusion," Stanford Centre for Research in Economic Growth, WP # 70, Stanford University.

Davies, S. (1979), *The Adoption of Process Innovations*, Cambridge, MA: Cambridge University Press.

Farewell, V.T., and Prentice R. L. (1980), "The Approximation of Partial Likelihood with Emphasis on Case-Control Studies," *Biometrika*, 67, 273-278.

Favero, C., Pesaran, M. H. and Sharma S. (1991), "Uncertainty and Irreversible Investment: An Empirical Analysis of Development of Oil Fields on the UK Continental Shelf," mimeo, University of Cambridge and UCLA. Feder, G., Just R. and Zilberman D. (1985), "Adoption of Agricultural Innovations in Developing Countries: A Survey," *Economic Development and Cultural Change*, 33, 255-298.

Fudenberg, D. and Tirole J. (1985), "Pre-emption and Rent Equalization in the Adoption of New Technology," *Review of Economic Studies*, 52, 383-401.

Han, A. and Hausman J. (1990), "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 5(June), 1-28.

Hannan, T. and McDowell J. D. (1984), "The Determinants of Technology Adoption: The Case of the Banking Firm," *RAND Journal of Economics*, 15, 328-335.

—— (1987), "Rival Precedence and the Dynamics of Technology Adoption: an Empirical Analysis," *Economica*, 54, 155-171.

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

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

Karshenas, M. and Stoneman P. (1990), "Rank, Stock, Order and Epidemic Effects in the Diffusion of New Process Technologies : An Empirical Model." Working Paper # 358, Department of Economics, University of Warwick.

Kiefer, N. M. (1988), "Analysis of Grouped Duration Data," in N.U. Prabhu (ed.), *Statistical Inference From Stochastic Processes*, Providence: American Mathematical Society.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge.

Mansfield, E. (1961), "Technical Change and the Rate of Imitation," *Econometrica*, 29(October), 741-66.

--- (1968), *Industrial Research and Technological Innovation: an Econometric Analysis*. New York: W.W. Norton and Co.

Manton, K.G., Stallard E. and Vaupel J. W.(1986), "Alternative Models for the Heterogeneity of Mortality Risks Among the Aged," *Journal of the American Statistical Association*, 81,635-44.

McCall, B. (1990), "Occupational Matching: A Test of Sorts," *Journal of Political Economy*, 98, 45-69.

Meyer, B. D. (1990), "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58(4), 757-782.

Pindyck, R. (1991), "Irreversibility, Uncertainty and Investment," *Journal of Economic Literature*, 29, 1110-1148.

Prentice, R. L. and Gloeckler L. A. (1978), "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data," *Biometrics*, 34(March), 57-67. Reinganum, J. F. (1989), "The Timing of Innovation: Research, Development, and Diffusion," *Handbook of Industrial Organization, Volume I*, edited by R. Schmanlensee and R.D. Willig, 849-908.

Ridder, G. (1986), "The Sensitivity of Duration Models to Misspecified Unobserved Heterogeneity and Duration Dependence," unpublished manuscript, University of Amsterdam.

Rogers, E. M. (1983), *Diffusion of Innovations, 3rd ed.*, Macmillan, New York.

Rose, N. L. and Joskow P. L. (1990), "The Diffusion of New Technologies: Evidence from the Electric Utility Industry," *RAND Journal of Economics*, 21(3), 354-73.

Saloner, G. and Shepard A. (1991), "Adoption of Technologies with Network Effects: An Empirical Examination of Automated Teller Machines," WP # 577, M. I. T., Cambridge.

Stoneman, P. (1983), *The Economic Analysis of Technical Change*, Oxford University Press, London.

Sueyoshi, G. (1991a), "Evaluating Simple Alternatives to the Proportional Hazard Model: Unemployment Insurance Receipt and the Duration of Unemployment," WP, UCSD.

Sueyoshi, G. (1991b), "A Class of Binary Response Models for Grouped Duration Data," WP, UCSD.

Thirtle, C. G. and V. W. Ruttan (1987), *The Role of Supply and Demand in the Generation and Diffusion of Technical Change*, Harwood Academic Publishers.

Thompson, W.A. (1977), "On Treatment of Grouped Observations in Life Studies," *Biometrics*, 34, 57-67.

Table 1
Empirical Hazard

Year	Potential Adopters	Adopters	Hazard Rate	Standard Error
1971	3689	14	0.00380	0.0010
1972	3675	24	0.00655	0.0013
1973	3651	78	0.02159	0.0024
1974	3573	55	0.01551	0.0021
1975	3518	78	0.02242	0.0025
1976	3440	102	0.03009	0.0030
1977	3338	90	0.02733	0.0029
1978	3248	86	0.02683	0.0029
1979	3162	212	0.06937	0.0048

Note: All 3689 firms in the sample were observed over the 1971-79 period. By the end of this period 739 firms had adopted ATMs.

Table 2(A)
Data Summary

	1971	1972	1973	1974	1975	1976	1977	1978	1979
SIZE	10.033	11.184	13.228	15.755	16.349	17.315	19.246	21.854	24.812
	(86.970)	(95.459)	(114.786)	(149.676)	(157.476)	(164.420)	(183.612)	(211.259)	(241.759)
GROWTH	1.15	1.15	1.12	1.10	1.08	1.07	1.10	1.12	1.07
	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
MIX	0.437	0.420	0.418	0.396	0.378	0.357	0.344	0.349	0.337
	(0.125)	(0.123)	(0.118)	(0.116)	(0.114)	(0.110)	(0.107)	(0.106)	(0.105)
BHC	0.226	0.246	0.271	0.299	0.314	0.327	0.335	0.350	0.371
N	3689	3689	3689	3689	3689	3689	3689	3689	3689

Table 2(B)
Data Summary

	1971	1972	1973	1974	1975	1976	1977	1978	1979
SIZE	10.033	10.696	11.959	13.660	14.048	14.591	15.123	17.235	19.349
	(86.970)	(94.242)	(111.626)	(145.153)	(155.086)	(162.719)	(176.659)	(206.729)	(238.028)
GROWTH	1.15	1.15	1.12	1.10	1.08	1.07	1.10	1.12	1.07
	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
MIX	0.437	0.420	0.418	0.396	0.378	0.356	0.343	0.348	0.335
	(0.125)	(0.123)	(0.118)	(0.116)	(0.114)	(0.111)	(0.108)	(0.108)	(0.107)
BHC	0.226	0.244	0.266	0.287	0.297	0.306	0.308	0.320	0.337
N	3689	3675	3651	3573	3518	3440	3338	3248	3162

Note: Standard errors are shown in parenthesis

Table 3
Non-parametric Baseline Hazard Models

	Model 1	Model 2	Model 3
SIZE	0.247	0.254	0.309
	(0.078)	(0.077)	(0.156)
GROWTH	1.669	1.589	1.565
	(0.745)	(0.753)	(0.754)
MIX	0.774	0.751	0.736
	(0.333)	(0.333)	(0.334)
BHC	0.919	0.920	0.917
	(0.075)	(0.075)	(0.075)
URBAN	0.284	0.241	0.233
	(0.121)	(0.122)	(0.122)
WAGE	0.093	0.093	0.091
	(0.041)	(0.041)	(0.041)
ATMSHAR	-0.109	-0.085	-0.207
	(0.165)	(0.165)	(0.207)
OFFPRM	0.381	0.316	0.324
	(0.163)	(0.165)	(0.166)
CR	1.238	1.836	1.814
	(0.257)	(0.314)	(0.316)
PROPN	1.712	7.826	7.454
	(0.364)	(1.801)	(1.859)
PROPN × CR	.	-7.884	-7.620
	.	(2.324)	(2.358)
PROPN × SIZE	.	.	-0.345
	.	.	(0.901)
PROPN × ATMSHAR	.	.	0.848
	.	.	(0.860)
LOG-LIKELIHOOD	-3194.465	-3188.511	-3187.942

Note: Standard errors are shown in parenthesis

Table 4
Non-parametric Baseline Hazard Models with Unobserved Heterogeneity

	Model 4	Model 5	Model 6
SIZE	0.247	0.254	0.309
	(0.078)	(0.078)	(0.156)
GROWTH	1.669	1.589	1.565
	(0.731)	(0.718)	(0.723)
MIX	0.774	0.751	0.736
	(0.335)	(0.319)	(0.319)
BHC	0.919	0.920	0.917
	(0.075)	(0.075)	(0.075)
URBAN	0.284	0.241	0.233
	(0.118)	(0.113)	(0.117)
WAGE	0.093	0.093	0.091
	(0.040)	(0.037)	(0.040)
ATMSHAR	-0.109	-0.085	-0.207
	(0.153)	(0.136)	(0.202)
OFFPRM	0.381	0.316	0.324
	(0.155)	(0.148)	(0.163)
CR	1.238	1.836	1.814
	(0.254)	(0.309)	(0.310)
PROPN	1.712	7.825	7.454
	(0.364)	(1.757)	(1.817)
PROPN × CR	.	-7.883	-7.620
	.	(2.268)	(2.304)
PROPN × SIZE	.	.	-0.346
	.	.	(0.483)
PROPN × ATMSHAR	.	.	0.848
	.	.	(0.865)
HETEROGENEITY	-9.056	-9.653	-9.044
PARAMETER	(16.154)	(11.806)	(11.148)
LOG-LIKELIHOOD	-3194.465	-3188.511	-3187.942

Note: Standard errors are shown in parenthesis

Table 5
Parametric Models: Exponential and Weibull Baseline Hazards

	Model 7	Model 8	Model 9	Model 10
ALPHA	.	.	2.416	1.569
	.	.	(0.118)	(0.094)
CONSTANT	-6.215	-6.241	-9.650	-7.875
	(0.834)	(0.831)	(0.870)	(0.942)
SIZE	0.209	0.239	0.223	0.243
	(0.072)	(0.076)	(0.075)	(0.076)
GROWTH	-0.942	-0.972	0.793	-0.027
	(0.680)	(0.677)	(0.700)	(0.769)
MIX	0.767	0.804	0.918	0.884
	(0.334)	(0.333)	(0.327)	(0.330)
BHC	0.921	0.937	0.876	0.912
	(0.075)	(0.075)	(0.075)	(0.076)
URBAN	0.136	0.134	0.342	0.230
	(0.121)	(0.121)	(0.122)	(0.122)
WAGE	0.231	0.234	-0.051	0.105
	(0.032)	(0.032)	(0.038)	(0.038)
ATMSHAR	-0.207	-0.213	-0.151	-0.187
	(0.177)	(0.177)	(0.159)	(0.168)
OFFPRM	0.537	0.554	0.311	0.457
	(0.172)	(0.171)	(0.155)	(0.164)
CR	1.845	1.885	1.782	1.893
	(0.308)	(0.309)	(0.324)	(0.316)
PROPN	8.984	9.379	4.451	7.395
	(1.667)	(1.684)	(1.854)	(1.782)
PROPN × CR	-8.801	-9.191	-4.934	-7.395
	(2.195)	(2.216)	(2.397)	(2.315)
LOG-LIKELIHOOD	-3263.721	-3245.747	-3163.900	-3222.826

Note: Standard errors are shown in parenthesis.
Model 7 assumes an Exponential hazard and does *not* account for grouping of data.
Model 8 assumes an Exponential hazard and accounts for grouping of data.
Model 9 assumes a Weibull hazard and does *not* account for grouping of data.
Model 10 assumes a Weibull hazard and accounts for grouping of data.

Table 6

Estimates of γ Parameters and the Non-Parametric Baseline Hazard

Year	Estimates of γ			Model 5	
	Model 1	Model 2	Model 5	Hazard × 100	Approx 95% Band
1971	-9.518	-9.779	-9.779	0.0057	(0.0010, 0.0338)
	(0.951)	(0.963)	(0.912)		
1972	-9.006	-9.264	-9.264	0.0095	(0.0016, 0.0549)
	(0.937)	(0.950)	(0.897)		
1973	-7.834	-8.099	-8.100	0.0304	(0.0056, 0.1655)
	(0.907)	(0.920)	(0.866)		
1974	-8.215	-8.497	-8.497	0.0204	(0.0038, 0.1092)
	(0.897)	(0.910)	(0.856)		
1975	-7.891	-8.181	-8.181	0.0280	(0.0054, 0.1456)
	(0.884)	(0.896)	(0.842)		
1976	-7.654	-7.960	-7.960	0.0502	(0.0098, 0.2584)
	(0.878)	(0.891)	(0.836)		
1977	-7.937	-8.287	-8.288	0.0252	(0.0046, 0.1399)
	(0.906)	(0.921)	(0.865)		
1978	-8.069	-8.415	-8.415	0.0222	(0.0039, 0.1255)
	(0.930)	(0.944)	(0.855)		
1979	-7.157	-7.522	-7.522	0.0541	(0.0101, 0.2893)
	(0.900)	(0.915)	(0.856)		

Note: Standard errors are shown in parenthesis

Table 7
Non-parametric Baseline Hazard Model with Time-varying Coefficients

	Phase I: 1971-78	Phase II: 1979
SIZE	0.176	0.445
	(0.124)	(0.175)
GROWTH	0.982	3.158
	(0.878)	(1.599)
MIX	0.637	1.097
	(0.387)	(0.671)
BHC	0.888	1.019
	(0.088)	(0.143)
URBAN	0.433	-0.099
	(0.148)	(0.224)
WAGE	0.106	0.093
	(0.052)	(0.068)
ATMSHAR	-0.200	-0.055
	(0.280)	(0.222)
OFFPRM	0.587	0.100
	(0.275)	(0.219)
CR	2.097	1.415
	(0.357)	(0.704)
PROPN	6.249	10.955
	(2.688)	(2.748)
PROPN × CR	-8.387	-9.140
	(3.528)	(3.581)
LOG-LIKELIHOOD	-3168.2	

Note: Standard errors are shown in parenthesis

FIGURE 1
Empirical Hazard

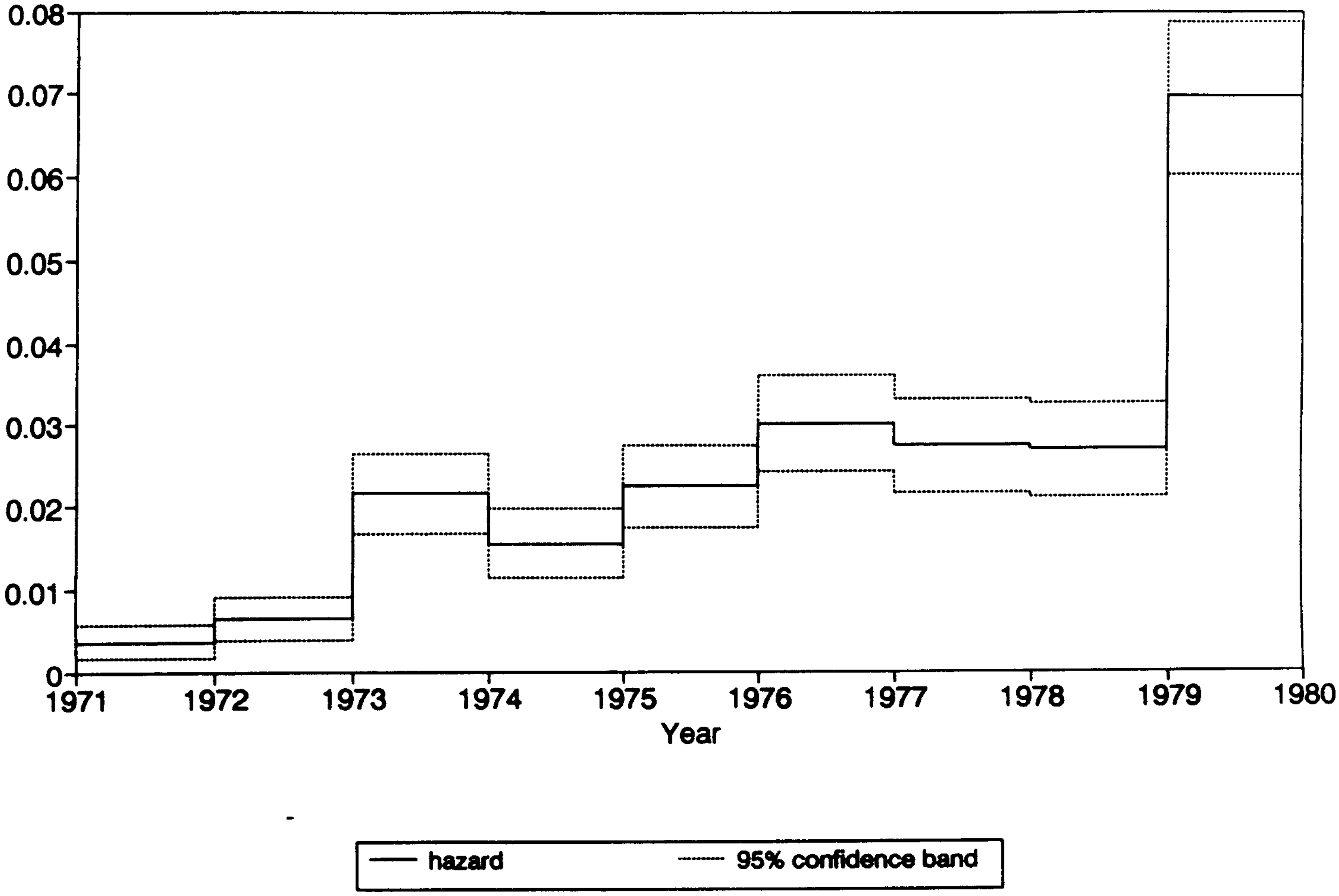

FIGURE 2
Empirical Distribution Function

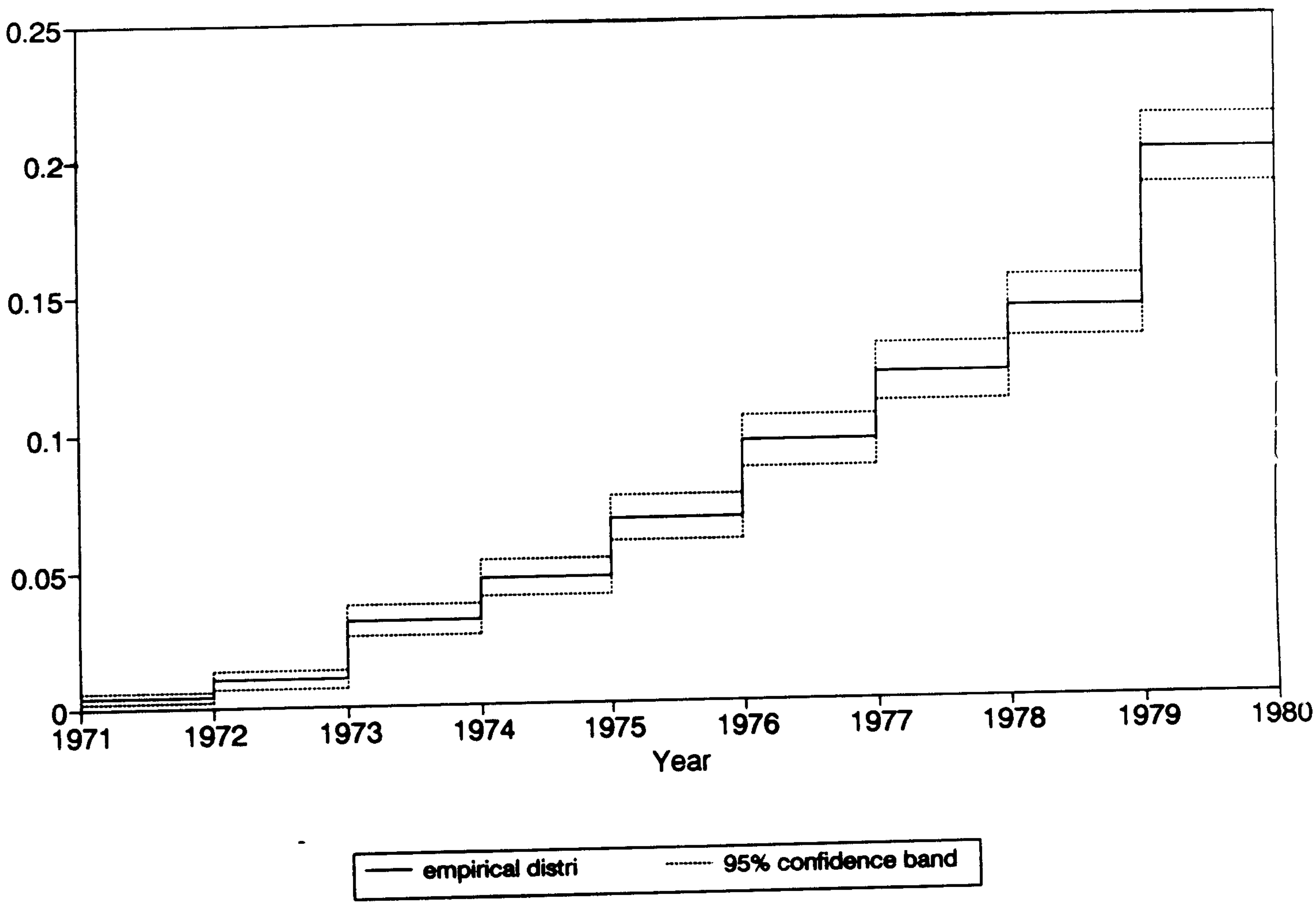

FIGURE 3
Baseline Hazard

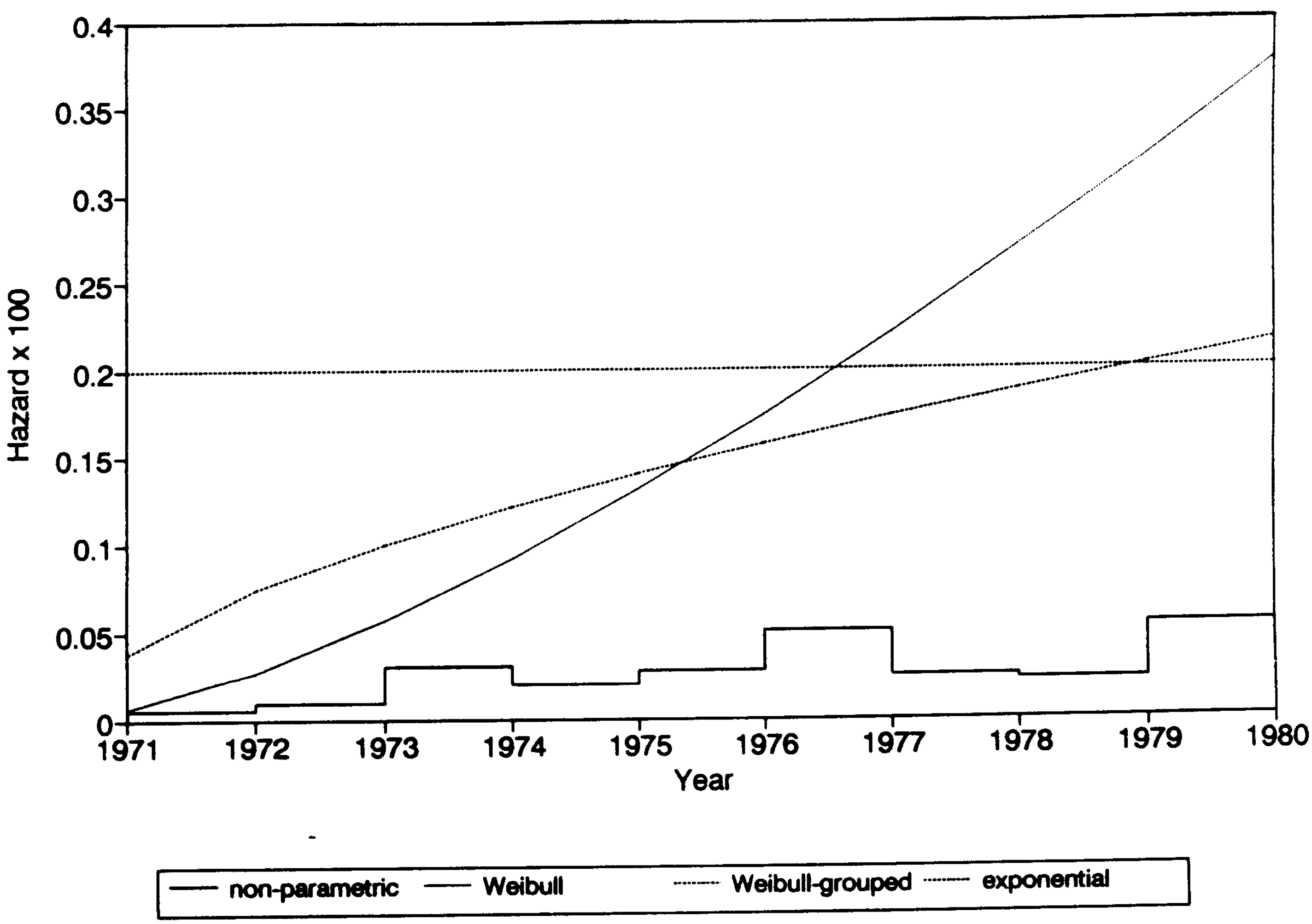

FIGURE 4
Plot of Integrated Hazard vs t

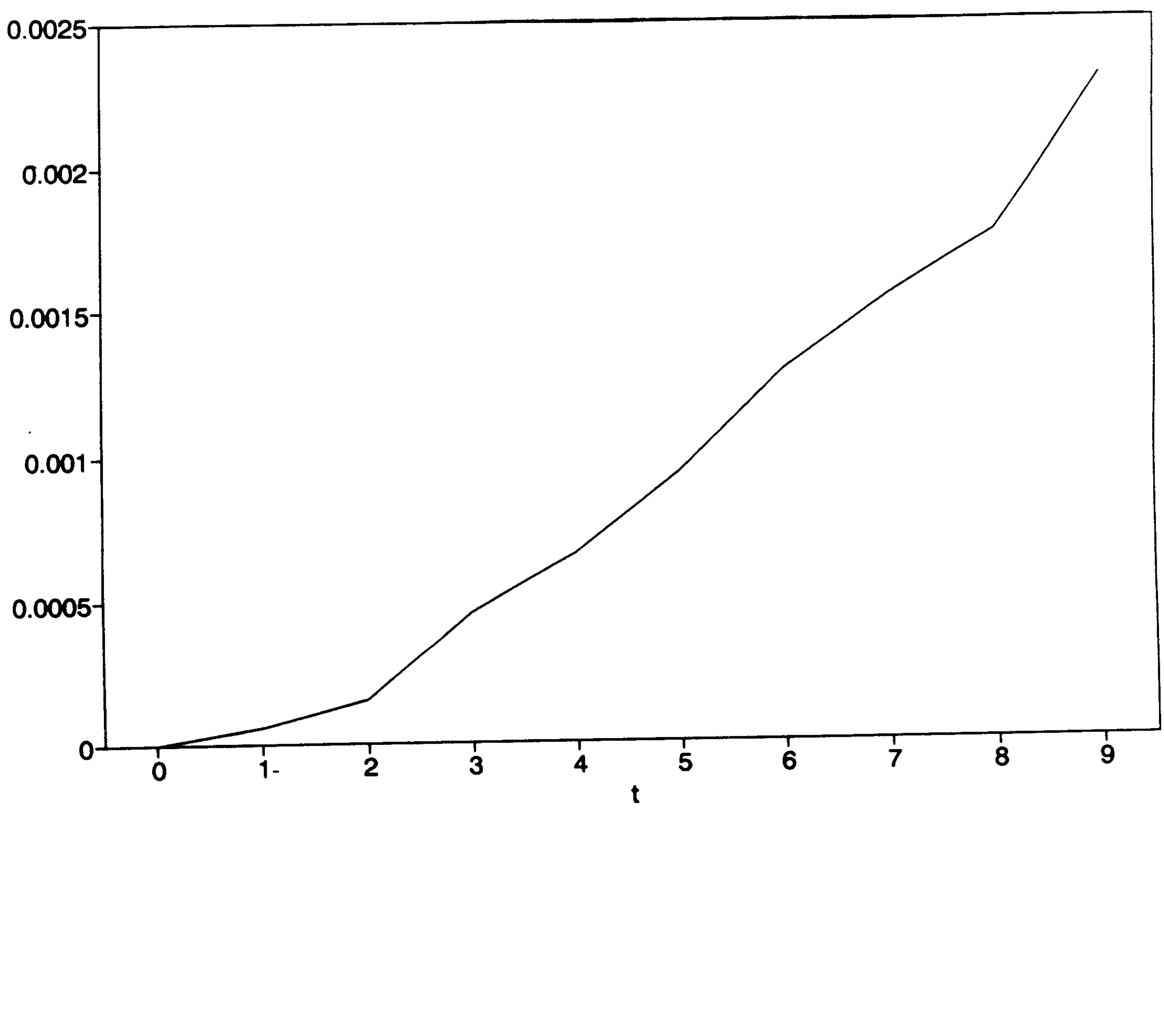

FIGURE 5
Plot of ln(integrated hazard) vs ln(t)