Adverse Selection, Moral Hazard and the Term Structure of Default

Koresh Galil*

Center for Financial Studies, Goethe University of Frankfurt

Eitan Berglas School of Economics, Tel-Aviv University

March 2004

Abstract

This paper estimates the term structure of the hazard rate to default by using two hazard models - one ignoring and another allowing unobserved heterogeneity and annual shocks to the hazard rate. Diamond (1989) predicts a declining hazard rate to default due to adverse selection and moral hazard. The adverse selection effect is confirmed. After controlling for adverse selection, the hazard rate shows to be increasing over time and hence the moral hazard effect cannot be confirmed. The paper contributes also to the credit risk literature by providing a tool to gap between the immediate and long-term default probabilities.

Keywords: Credit Risk, Adverse Selection, Moral Hazard, Survival Analysis, Unobserved Heterogeneity

JEL classification: **G10, G12, G14, G20**

---

* Center for Financial Studies (galil@ifk-cfs.de) and Eitan Berglas School of Economics, Tel-Aviv University (koresh@post.tau.ac.il). This paper is part of my Ph.D. dissertation under the supervision of Oved Yosha and Simon Benninga.

Introduction

Credit risk is one of the major financial risks faced by commercial banks and portfolio managers. However, credit risk was not so thoroughly explored as market risk. While we know much about the properties of market risk, we know very little about properties of credit risk. One of the major reasons for that is the curse of rare events. The frequency of defaults or bankruptcies is too low to enable empirical investigation in the same scale of investigations regarding market risk. One consequence of this phenomenon is a lack of stylized facts. For example, when evaluating theoretical or empirical models concerning market risk, the phenomena of fat tails and negative skewness can be used as validating facts. However such stylized facts hardly exist when discussing credit risk models.

This study attempts to investigate one aspect of credit risk, which is the term structure of the hazard rate to default. The estimation of the term structure has an importance in providing additional insight toward a better understanding of the properties of default risk. However, the results can shed light also on the properties of the bond markets, and phenomena such as adverse selection and moral hazard. Diamond (1989), shows that in a debt market with heterogeneous credit quality of borrowers, the hazard rate to default should decline over time. This result has two deriving forces. The first, firms with lower credit quality tend to go bankrupt faster, and hence the remained population tends to become safer over time. The second, the lenders that know that the remained population is safer demand lower interest and hence the borrowers who have gained reputation (an intangible asset that cannot be liquidated in the case that the firm goes bankrupt) tend to invest in less risky assets.

Therefore, according to Diamond (1989), both the adverse selection and the moral hazard phenomena will bring to a declining term structure of hazard rate.

Gorton (1996) brings an evidence for one of the consequences of Diamond (1989) theory, i.e. firms with shorter (or no) credit history pay higher interest rate. However, this result cannot be traced to the two deriving forces described in Diamond (1989), which are the diminishing adverse selection and the moral hazard.

Fons (1994) and Carty (2000) focus on the term structure of the hazard rate itself. Fons, using average cumulated default rates shows that for higher graded firms, the hazard rates first increases over time and only then it decreases. The paper also shows that low-graded firms (firms having a Moody's rating of Ba) have a decreasing hazard rate to default. Carty (2000) however, once controlling for particular annual macro-economic shocks and industrial classification and some other possible heterogeneities, shows that the hazard rate first increases and then decreases even for low graded firms (such as firms graded Baa by Moody's).

The appearance of an upward slope in the term structure of the hazard rate, suggests that Diamond's model ignores some relevant forces. Such a possible force might be the accessibility to cash (cash-flow effect). After raising new funds, a firm can experience several years of low default risk due to the new reserves of cash. However, as these reserves diminish over time, the more volnurable the firm becomes to possibility of default.

The empirical investigation employs two types of hazard models.[1] The first model follows a Biometric model introduced by Prentice and Gloeckler (1978). In this model, the hazard rate is proportional, i.e. the term structure is separable from other factors that affect the hazard risk, and the time measure is discrete. The

[1] Hazard models are also called in duration models or survival models.

model allows also covariates (i.e. the explanatory variables to change over time). This model is semi-parametric and can be easily estimated using a maximum likelihood approach. However the model does not allow unobserved heterogeneity, since there is no source of noise in the hazard function. Therefore, ratings, when used as explanatory variables are assumed to be a fully efficient measurement of the determinant of the firm's credit risk.

The second model follows a model implemented by Meyer (1990). This model simply ads to the Prentice and Gloeckler (1978) hazard function a source of noise. On assuming this noise to be Gamma distributed, it is possible to draw the likelihood function and estimate the hazard function using a maximum likelihood approach. This model allows both covariates and unobserved heterogeneity. Therefore, it is suitable for the case that ratings are assumed to be just a proxy for the creditworthiness of the firm due to unobserved heterogeneity or annual shocks.

A simple and naive calculation of the hazard rate to default, using cumulated historical default probabilities leads the observer to suspect that the hazard rate of default increases in the first years after issuing a new bond and then decreases again to a level equal to the one immediately after issuance. However, such a calculation is misleading. Lancaster (1990) points out that if the hazard rate is stochastic over time, or in a case of missing variables, ignoring the unobserved heterogeneity would lead to downward estimates of all coefficients, including the term structure. The intuition behind this can be easily illustrated using a simple example. Consider a population that consists of two equal-sized groups of firms that differ in their constant hazard rate of default, with 8 percent and 6 percent respectively. The observed hazard rate would be 7 percent (the average) at the beginning. Over time the remained population would consist of firms with lower default risk, and therefore, the observed hazard rate would tend to decrease toward 6 percent. However, conditioned on the information available at the time of rating, the hazard rate to default is constant over time for each firm.

This effect of unobserved heterogeneity can be to some extent controlled for by using a transition matrix for calculating the implied hazard rate. A transition matrix is a common way to describe the stochastic process of the credit quality of the firm and is used in reduced form models (such as Jarrow, Lando & Turnbull (1997)) for pricing of credit spreads. The matrix describes not only the annual probability of default for each rating classification but also the probability of transition to another rating classification. Hence, the matrix takes into account existence of unobserved heterogeneity (or adverse selection). Yet, according to S&P methodology, the ratings aim to look through the cycle, and hence the transition matrix does not take into account (or slightly takes into account) possible transitions due to aggregate shocks.

The term structure of the hazard rate to default derived from the transition matrix is shown to be increasing over time for most rating classification. This exercise shows that on controlling for adverse selection, there is almost no diminishing moral hazard effect visible. Though such an effect might still exist, it seems to be overshadowed by another effect that induces increasing hazard rate over time (possibly cash-effect).

The database used for this paper is based three sources. A list of 10,000 new corporate bonds issued in the US during the years 1983-1993 is linked with lists of default occurrences during the years 1983-2000, obtained mainly from *Moody's Investor Services* publications. After eliminating financial corporations, multiple issues by single issuers within a calendar year, a database with 2596 bonds of 1013 issuers is left. The long-term horizon that features the survival analysis enables 235 cases of default by 155 firms to be identified. However, this methodology also uses each year of exposure to default risk as an individual observation. Therefore the total number of observation in the sample is 27906 of which 235 are observation of default.

The estimation of the term structure, while ignoring unobserved heterogeneity reveals similar results to those when using Standard and Poor's statistics for cumulated frequencies of default. The term structure appears to be first increasing, and then decreasing. This term structure appears to be significant. Yet, on allowing unobserved heterogeneity and annual shocks to the hazard rate, the decreasing pattern of the term structure is replaced with a slightly increasing pattern. The hazard rate in the year following issuance is still significantly low and from the second year and on, it is slightly increasing. The noise-variance is significant, and therefore the case of absence of unobserved heterogeneity (adverse selection within each rating category) is rejected.

An interesting observation (yet not significant) is a drop in the hazard rate in the years 4, 6, 8, and 11 from issuance. These drops might further support the cash-flow-effect thesis. Since firms tend to issue new bonds in intervals of 3, 5, 7 and 10 years, then we can expect their default risk to drop in the respective following years (i.e. 4, 6, 8, and 11). Nevertheless, this correlation of the term structure of the hazard rate to default with the time to maturity of the bonds is not the driving force of the results but only an indicator for the relevance of the cash-flow effect.

The remainder of the paper is organized as follows. Section I provides an analysis based on aggregate data. Section II describes the methodology used. Section III describes the data and Section IV the results. Section V contains the conclusions.

I Aggregate Observations

A. Using Cumulative Default Probabilities[2]

Table I shows the historical average cumulative default probabilities of the main rating categories up to fifteen years after issue as documented by S&P.[3] Let $F_r(t)$ denote the average cumulative probability of default of rating r, t years after assigning the rating. Table I-b describes $f_r(t) = F_r(t) - F_r(t-1)$ - the average probability of default of rating r between time $t-1$ and time t and table I-c shows $\theta_r(t) = f_r(t) / [1 - F_r(t-1)]$ - the average hazard rate of default between time $t-1$ and time t.

(Insert Table I about here)

Figure 1 shows $\theta_r(T)$ of each rating category up to 15 years after assigning the rating. These tables and figures suggest that the average hazard rate first increases over time and then decreases. The lower the rating, the faster the hazard rate reaches its maximum. However, it must be noted that such calculation is biased due to heterogeneity within each rating category. Such heterogeneity will induce the average hazard rate to be decreasing.

(Insert Figure 1 about here)

B. Using Transition Matrix

Suppose $S \in \{1,...k\}$ represents a firm's rating at some period, where 1 is the highest rating (the lowest default risk) and $k-1$ the lowest rating (the highest default risk

[2] This calculation appears also in Appendix A of Galil (2003).
[3] See "Ratings Performance 2000", Standard and Poor's. These statistics are based on all bonds rated by S&P during the years 1981 to 2000.

but not default and k a state of default. Suppose the rating of the firm follows a Markovian process that can be specified by the following $k \times k$ transition matrix,

$$Q = \begin{Bmatrix} q_{11} & q_{12} & \cdots & q_{ik} \\ q_{21} & q_{22} & \cdots & q_{2k} \\ \vdots & & & \\ q_{k-1,1} & q_{k-1,2} & \cdots & q_{k-1,k} \\ 0 & 0 & \cdots & 1 \end{Bmatrix}$$

where $q_{rs} \geq 0$ for all $r,s,\ r \neq s$ and $q_{rr} = 1 - \sum_{s=1,\ s \neq r}^{k} q_{rs}$ for all r. q_{rs} represents the probability of a change from rating r to rating s in one year. It is assumed that the default rating is an absorbing state, so that the probability of a change from a state of default to a non-default rating is $q_{kr} = 0$ for all r and $q_{kk} = 1$. Note that the k-th column of the matrix Q represents the annual probability of default for all rating classes. However, for estimation of the term structure of the hazard rate to default, it is necessary to take into account that the rating evolves over time.

Let $q_{rs}(0,t)$ denote the probability of transition from rating r at time 0 to rating s at time t. The t-years $k \times k$ transition matrix, $Q_{0,t}$ whose (r,s) entry is $q_{rs}(0,t)$ satisfies $Q_{0,t} = Q^t$. The k-th column represents the set of the cumulated probabilities of default till time t. Let $F_r(t) \equiv q_{rk}(0,t)$, then the probability of default at time t of a firm rated r at time 0 can be calculated by $f_r(t) \equiv F_r(t) - F_r(t-1)$ and the hazard rate to default at time t of a firm rated r at time 0 can be also calculated by $\theta_r(t) \equiv f_r(t) / \left[1 - F_r(t)\right]$.

Table II-a shows an S&P annual transition matrix.[4] The category of RW resembles cases that the rating is withdrawn. Carty (1997) shows that 92 percent of the rating withdrawals by Moody's were because the issues had matured or had been called. Assuming that S&P's withdrawals of ratings follow the same reason, it is reasonable to add the cases of withdrawals to cases that the rating was unchanged. Table II-b shows the transition matrix when the cases of withdrawals are added to the cases that the rating was unchanged and the cases of default are assumed to be terminal (bonds that default do not recover).

(Insert Table II about here)

Figure 2 shows the term structure of the hazard rate to default for each of the rating categories computed using the transition matrix. The hazard rate to default for ratings AA, A, BBB, BB, and B show to be monotonically increasing. Rating AAA appears to be increasing to a pick at years 6 and 7 after issuance and then slightly decreasing. The decreasing patterns in the term structure of default that appear in figure disappear in the ratings AA, A, BBB, BB and B, and weaken in the case of AAA. In the case of CCC, the term structure remains declining. However, it should be noted that the transition matrix approach would always exclude a case of monotonically increasing term structure of default for the lowest non-default rating category.

The results' differences might be attributed to the effect of adverse selection. The first approach totally ignores existence of unobserved heterogeneity. The

[4] See Standard & Poor's (1997), "Rating Performance: Stability and Transition. Special Report". These statistics are based on changes in ratings of all S&P-rated bonds during the years 1981-1996.

transition matrix, however, allows for unobserved heterogeneity. While the one year default probability for each rating category is homogeneous, the transition probabilities induce heterogeneous probabilities of default over time for each rating category. It should be noted that S&P ratings follow the rule of 'looking through the cycle'[5] and therefore the transition matrix still does not take into the account the whole possibilities of changes in default probability.

II Methodology

A. A Discrete Time Proportional Model for Default Occurrence with no Unobserved Heterogeneity (Prentice-Gloekler Model)

A.1 General Discrete-time Hazard Model with Covariates

Assume that all firms that are exposed to default risk experience default at some time in future. Let i denote an observation on a firm issuing a new bond.[6] Let also T_i^D denote the time till the first time the firm defaults, and let $\theta(t)$ denote the hazard rate to default at time t (the probability to default at time t conditioned on survival till time t). x_{it} is a vector of variables, whether describing the firm or its industry or the economy at time t. Suppose that the time is observed in discrete intervals. It is possible to integrate the hazard rate over the time interval $(t-1,t]$. The integrated hazard rate is

[5] A temporary change in the credit quality of the firms does not necessarily lead to a change in its rating.

[6] Firms might issue multiple bonds annually and even concurrently. The inclusion of all bonds, and their respective ratings and time to default, would create a severe problem of autocorrelation. On the other hand, we can expect the expected term structure of a firm to vary over time. Therefore, the database for this study includes only one observation per firm per year (the first bond each year) and each observation is treated separately. However, this assumption is not needed for drawing the estimation model, but merely for addressing the problem of autocorrelation.

$$h(t;x_{it}) = \int_{t-1}^{t} \theta(u;x_{it})du. \tag{1}$$

Note that this is not the discrete hazard rate. Now $\bar{F}\left(t;x_{i1},...,x_{it}\right)$, the associated survivor probability till time t is

$$\begin{aligned}\bar{F}\left(t;x_{i1},...,x_{it}\right) &\equiv \exp\left(-\int_{0}^{t}\theta\left(u;x_{iu}\right)du\right) = \exp\left(-\sum_{s=1}^{t}h\left(s;x_{is}\right)\right) \\ &= \prod_{s=1}^{t}\exp\left[-h\left(s;x_{is}\right)\right]\end{aligned}. \tag{2}$$

The probability to survive till time t conditional on surviving till time $t-1$ is

$$prob(T_i^D \geq t; T_i^D \geq t-1, x_{i1},...,x_{it}) = \frac{\bar{F}\left(t;x_{i1},...,x_{it}\right)}{\bar{F}\left(t-1;x_{i1},...,x_{it-1}\right)}. \tag{3}$$

Incorporating equations (2), it can be rephrased,

$$prob(T_i^D \geq t; T_i^D \geq t-1, x_{i1},...,x_{it}) = \exp\left[-h\left(s;x_{it}\right)\right]. \tag{4}$$

And the hazard rate for the discrete time t is,

$$prob(T_i^D \leq t; T_i^D \geq t-1, x_{i1},...,x_{it}) = 1-\exp\left[-h\left(s;x_{it}\right)\right]. \tag{5}$$

Define T_i as the period during which the firm i is known to have been exposed to default risk. Each period can end whether due to default, or censorship (when T_i^D is not observed due to a maximum length of observation or lack of information on availability of outstanding bonds). Let $\delta_i = 1$ in the case the period T_i ended with default ($T_i = T_i^D$), and $\delta_i = 0$ otherwise ($T_i < T_i^D$). Then the likelihood function for a sample of N observation is,

$$l(.) = \prod_{i}^{N} \left\{ \prod_{s=1}^{T_i - 1} \exp\left[-h\left(s; x_{is}\right) \right] \right\} \cdot \left\{ 1 - \exp\left[-h\left(s; x_{it}\right) \right] \right\}^{\delta_i}. \tag{6}$$

A.2 Proportional Hazard Function

For estimating the model, it is needed to fully or partially parameterize the hazard rate. Following Cox (1972), assume that the hazard function is of the proportional hazard form and has the following construction

$$\theta(t; x_{it}) = k_1(x_{it}) \cdot k_2(t). \tag{7}$$

This construction assumes that term structure of the hazard rate $k_2(t)$ is separable from $k_1(x_{it})$. The Cox (1972) is a semi-parametric model, which enables estimating the first component $k_1(x_{it})$ without parameterizing the term structure of the hazard rate, $k_2(t)$ which is also called the baseline hazard. On assuming that x_{it} is constant during the interval $(t-1, t]$, the integrated hazard at the discrete time t, $h(t; x_{it})$, has the following form

$$h(t; x_{it}) = \int_{t-1}^{t} \theta(u; x_{it}) du = k_1(x_{it}) \cdot \exp\left(\gamma(t)\right). \tag{8}$$

where $\gamma\left(t\right) \equiv \ln\left[\int_{t-1}^{t} k_2\left(u\right) \cdot du \right]$. On incorporating this equation in the survival function presented in equation (2), the survival function will be

$$\bar{F}\left(t; x_{i1}, ..., x_{it}\right) = \prod_{s=1}^{t} \exp\left[-k_1(x_{is}) \cdot \exp\left(\gamma(s)\right) \right]. \tag{9}$$

And the probability to survive till time t conditional on surviving till time $t-1$ is

$$prob(T_i^D \geq t; T_i^D \geq t-1, x_{i1}, ..., x_{it}) = \exp\left[-k_1(x_{it}) \cdot \exp\left(\gamma(t)\right) \right]. \tag{10}$$

Hence, the hazard rare to default at the discrete time t (the probability to default at the discrete time t conditional on surviving till time $t-1$) is

$$prob(T_i^D \le t; T_i^D \ge t-1, x_{i1},...,x_{it}) = 1-\exp\left[-k_1(x_{it})\cdot\exp\left(\gamma(t)\right)\right]. \tag{11}$$

Therefore this likelihood function in the case of Cox (1972) proportional hazard model, can be written as

$$l(.) = \prod_i^N \left\{ \prod_{s=1}^{T_i-1} \exp\left[-k_1(x_{is})\cdot\exp\left(\gamma(s)\right)\right] \right\} \cdot \left\{ 1-\exp\left[-k_1(x_{iT_i})\cdot\exp\left(\gamma(T_i)\right)\right] \right\}^{\delta_i}. \tag{12}$$

This approach enables estimating the parameters of $k_1(x_{it})$, the first component of the hazard function, and a vector of parameters, $\gamma = \left[\gamma(0), \gamma(1), ...\right]$, describing the term structure of the hazard rate without parameterizing $k_2(t)$, the baseline of the hazard function. A simple and common case is where $k_1(x_{it})$ is linear. In such a case, $k_1(x_{it}) = \exp(x_{it}'\beta)$ where β is a vector of parameters corresponding to x_{it}. The hazard function for the continuous term in such a case is

$$\theta(t; x_{it}) = \exp(x_{it}'\beta)\cdot k_2(t). \tag{13}$$

And the integrated hazard at the discrete time t, is

$$h(t; x_{it}) = \exp\left(x_{it}'\beta + \gamma(t)\right). \tag{14}$$

The probability to survive till time t is $\bar{F}\left(t; x_{i1},...,x_{it}\right) = \prod_{s=1}^{t} \exp\left[-\exp\left(x_{is}'\beta + \gamma(s)\right)\right]$.

And the corresponding likelihood function for estimating equation (14) is

$$l(\beta,\gamma) = \prod_i^N \left\{ \prod_{s=1}^{T_i-1} \exp\left[-\exp\left(x_{is}'\beta + \gamma(s)\right)\right] \right\} \cdot \left\{ 1-\exp\left[-\exp\left(x_{iT_i}'\beta + \gamma(T_i)\right)\right] \right\}^{\delta_i}. \tag{15}$$

Note that the formation of the hazard function allows for the determinants of the hazard rate to vary over time. However, it is also possible to assume constant variables over time. This case draws a special attention since it enables to estimate the hazard function conditioned on information available only at the time of rating. In such a case the vector x_{it} in equations (13), (14), and (15) has to be simply replaced by the vector x_i.

B. A Mixture Model for Default Occurrence (Meyer Model)

B.1 General Model

Meyer (1990) enhances the hazard model by Prentice & Gloeckler (1978) by introducing unobserved heterogeneity into the proportional hazard rate function. The introduction of the random variable in the hazard rate function enables considering unobserved heterogeneity among firms and the possibility of stochastic hazard rate due to unexpected changes in the economy or unexpected changes in the creditworthiness of the issuer itself. Let $\theta(t; x_{it}, v_i)$ denote the hazard rate to default at time t where v_i is a random variable with density $\mu(v)$. v_i resembles the affect of all omitted variables or unexpected events on the exact realization of T_i^D. The integrated survival function till time t can be constructed by conditioning on the unobserved v_i:

$$\bar{F}(t; x_{i1}, \ldots, x_{it}) \equiv \int_v \exp\left(-\int_0^t \theta(u; x_{iu}, v)\, du\right) d\mu(v). \qquad (16)$$

The likelihood function for the sample is

$$l(.)=\prod_{i=1}^{N}\left\{\begin{array}{l}\delta_i\cdot\left[\bar{F}\left(T_i-1;x_{i1},\ldots x_{iT_i-1}\right)-\bar{F}\left(T_i;x_{i1},\ldots x_{iT_i}\right)\right]\\+\left(1-\delta_i\right)\cdot\left[\bar{F}\left(T_i;x_{i1},\ldots x_{iT_i}\right)\right]\end{array}\right\}. \quad (17)$$

It is possible to write (17) differently:

$$l(.)=\prod_{i=1}^{N}\left\{\bar{F}\left(T_i;x_{i1},\ldots x_{iT_i}\right)-\delta_i\cdot\bar{F}\left(T_i-1;x_{i1},\ldots x_{iT_i-1}\right)\right\}. \quad (18)$$

Now, incorporating equation (16), it will have the following form:

$$l(.)=\prod_{i=1}^{N}\int_v \exp\left(-\int_0^{T_i}\theta\left(u;x_{iu},v\right)du\right)d\mu\left(v\right)-\delta_i\cdot\int_v \exp\left(-\int_0^{T_i-1}\theta\left(u;x_{iu},v\right)du\right)d\mu\left(v\right). \quad (19)$$

B.2 The Case of Proportional Hazard Rate

Assume again that the hazard rate has a proportional form $\theta(t;x_{it})=k_1(x_{it})\cdot k_2(t)\cdot v_i$. Then the survival function (16) can be written as,

$$\bar{F}\left(t;x_{i1},\ldots,x_{it}\right)\equiv\int_v \exp\left(-\int_0^t k_1(x_{iu})\cdot k_2(u)\cdot v\cdot du\right)d\mu\left(v\right) \quad (20)$$

On observing discrete times (where a discrete time t is the time-interval $(t-1,t]$), and assuming again that x_i is constant during the time-interval $(t-1,t]$, the likelihood function is:

$$\bar{F}\left(t;x_{i1},\ldots,x_{it}\right)\equiv\int_v \exp\left(-v\cdot\sum_{u=0}^{t}k_1(x_{iu})\cdot\exp\left[\gamma(u)\right]\right)d\mu\left(v\right), \quad (21)$$

where again $\gamma\left(t\right)=\ln\left[\int_{t-1}^{t}k\left(u\right)\cdot du\right]$. In the linear case where $k(x_j)=\exp\left(x_{it}'\beta\right)$, the survival function has the following form,

$$\bar{F}\left(t;x_{i1},...,x_{it}\right)\equiv\int_{v}\exp\left(-v\cdot\sum_{u=0}^{t}\exp\left[x_{it}'\beta+\gamma(u)\right]\right)d\mu\left(v\right) \tag{22}$$

For estimation of a duration model using equation (22), it is typical to assume a shape for the distribution $\mu\left(v\right)$. A convenient and commonly used distribution for v_j is the gamma distribution with unit mean (normalized) and variance $\frac{1}{\eta}$.[7] This distribution gives a closed form expression for the likelihood function, avoiding numerical integration. Meyer (1990) shows that under such conditions $\bar{F}\left(t;x_{i1},...,x_{it}\right)$ - the survival probability till time T can be denoted as:

$$\bar{F}\left(t;x_{i1},...,x_{it}\right)=\left[1+\frac{1}{\eta}\cdot\sum_{u=0}^{t}\exp\left\{x_{iu}'\beta+\gamma\left(u\right)\right\}\right]^{-\eta}. \tag{23}$$

And the likelihood function of the sample is

$$l\left(\beta,\gamma,\eta\right)=\prod_{i=1}^{N}\left\{\begin{array}{l}\left[1+\frac{1}{\eta}\cdot\sum_{u=0}^{T_i}\exp\left\{x_{iu}'\beta+\gamma\left(u\right)\right\}\right]^{-\eta}\\ -\delta_i\cdot\left[1+\frac{1}{\eta}\cdot\sum_{u=0}^{T_i-1}\exp\left\{x_{iu}'\beta+\gamma\left(u\right)\right\}\right]^{-\eta}\end{array}\right\}, \tag{24}$$

[7] For further details concerning Gamma distribution see appendix A.

III Data

A. Database

The database is based on three sources. A list of more than 10,000 corporate bonds issued during the years 1983-1993 was obtained from the *Capital Division of Federal Reserve*.[8] Each issue in this database is detailed with name of issuer, date of issue, S&P and Moody's rating at date of issue and other characteristics of the bond. A list of default events was mainly obtained from *Moody's Investor's Service* publications.

After combining these sources and eliminating financial corporations, multiple issues within each year, companies with no S&P rating, 2596 bonds of 1013 non-financial corporations remained. Of which 235 bonds belong to 155 firms that default at some point after appearance of their issues in the sample. Many corporations issued more than one bond during the sample period. For the estimation of the term structure of the hazard function each year of exposure to default risk is defined as an observation. Hence the total number of observations (exposures to default risk) is 27906 of which 235 are observations of default.

B. Data Definition

First, T_i the time that firm i has been exposed to default risk since the issuance.[9] This period depends not only on the time to maturity of a bond issued at this time but also on bonds issued before and after. For example, if the time of maturity of a bond issued at year 1991 is year 1999 and the time of maturity of the

[8] This dataset is used by Guedes & Opler (1996) and is in the public domain.
[9] For a thorough explanation of this variable, including various examples, see Appendix A in Galil (2003).

bond issued at time 1992 is 1998, then it is clear that the firm has been exposed to default risk since 1991 through time 1992 till 1999. Therefore, and unless default occurred before 1999, for the issuance of 1991, $T_i = 8$ and for the issuance of 1992, $T_j = 7$. If the firm defaulted during this period then the final period T_{it} was calculated from its date of issue till date of default. In such a case (and only in such case) the observation is considered to be uncensored ($\delta_i = 1$). For all observations, where the period of exposure to default risk has not ended with default, the observation is considered to be censored ($\delta_i = 0$). An observation is also considered censored if the time of exposure to risk is beyond year 2000. The reason for that is that it is not known at what exact time (after year 2000) the firm defaults.

The observations are taken over 11 years (1983-1993). Some firms appear in the sample several times since they issued bonds in several different years, while other firms only appear in the sample once. Since this paper focuses on estimation of the term structure of default conditioned on the information at the time of issuance, each year of issuance by the same firm is considered as a separate observation. It should be noted that this is a common practice in reports concerning default experience.

C. Data Description

Table I shows the shows the distribution of annual observation of exposure to default risk by the number of years passed from issuance. It obviously shows that the number of annual observations decreases the larger is the distance from the time of issuance. It is due to censorship and defaults that occur over time and also the end of observation of defaults after year. An especial attention should be paid to the number of defaults (which does not have to be automatically decreasing over time). The number of defaults is especially low after 8 years. There are no cases of defaults at years 16 and 17. Therefore it is not possible to estimate the term structure at years 16 and 17. The low number of defaults at the later years has to get special attention.

(Insert Table III about here)

Table II shows the distribution of the annual observations of exposure to default risk by rating of the bond at the time of issuance. The number of defaults in the sample for the higher rated bonds is quite low. However, it should be noted that rating is used here just for controlling and the coefficients of the rating dummy variables are not the target of this paper. Table II shows that the sample includes quite diverse and includes both investment graded and speculative graded bonds.

(Insert Table IV about here)

IV Results

Table V shows the results of estimations of the two models using the micro data; one ignoring unobserved heterogeneity (UH) and annual shocks (AS) to the hazard rate and the other not. The significance of the variance of the Gamma distributed noise in the second model, leads to the rejection of the hypothesis of absence of unobserved heterogeneity.

Insert Table V about here

The analysis should focus not on the levels of the coefficients but on the differences between them. Figure 3 describes the estimated term structure for each model. When ignoring unobserved heterogeneity, the estimated term structure is quite similar to the one generally observed when looking at the aggregate cumulated frequencies of default. When ignoring UH and AS, the term structure is low at the year following issuance, increasing to a maximum and then decreasing slowly over time to a level equal to the one in the first year following issuance. However, the control for UH and AS makes clear changes. All coefficients are larger than in the first model. This result is very expected in presence of observed heterogeneity (see Lancaster (1990)). The hazard rate to default is still smaller in the first year following the issuance compared to the following years. Then the term structure shows an increasing pattern. The gap created between the two graphs widens. The significance of the noise variance confirms existence of adverse selection (after controlling for ratings), and the effect on the term structure of the hazard rate to default is consistent with Diamond (1989). However, after controlling for the adverse selection (by introducing UH), the term structure does not show a declining pattern. This result is not consisted with Diamond (1989) that predicts that the accumulation of reputation by the issuer would weaken the problem of moral hazard and encourage the firms to take less risky investments.

Insert Figure 3 about here

An interesting result is the correlation between the pattern of the term structure to default and the time maturity of the issues that were used for creating the estimation sample. Figure 4 shows the frequencies of the time to maturity of 2596 bonds. The highest frequencies (in the horizon of 15 years) are time to maturities of 10 years (27.9 percent), 7 years (7.6 percent), 5 years (6.8 percent), 15 years (6.4 percent), 12 years (5.5 percent), and 3 years (2.4 percent). Excluding the year 15, in all these years the hazard rate to default is slightly higher than the year before and the year after. Significant drops are in the years 4 and 11. These drops might reflect existence of a cash-effect. Yet, the general pattern of the term structure as reflected in the model allowing UH and AS cannot be explained by the time to maturity of the bonds. As an example while the frequencies of bonds of 1 year to maturity is 1.0 percent and 11 years to maturity is 0.5 percent, the hazard rate to default at year 11 is significantly higher than at year 1.

Insert Figure 4 about here

To address the question of the significance of the term structure, it is necessary to group the dummy variables. As was shown in table III, the number of observations for the 9th year from issuance and alter is relatively small. Figure 3 shows the term structure has a small slope during from year 2 and later. Hence, for easing the significance tests, it is desirable to group the dummy variables for the years from issuance into 4 dummy variables – 1 year after issuance, 2 or 3 years after issuance, 4 to 8 years from issuance, and 9 to 17 years from issuance. The results of estimation of the two models (the one ignoring unobserved heterogeneity and the one allowing unobserved heterogeneity) when including these dummy variables are reported in table VI.

Insert Table VI about here

Table VII shows the t statistics for the differences between the dummy variables for the time from issuance. The term structure in model 1 (ignoring UH and AS) is as described in the unconstrained case. The hazard rate is significantly lower in year 1 compared to years 2 and 3 and 4 to 8 but not so comparing with years 9 to 17. The hazard rate during the years 2 and 3 is significantly higher comparing to years 4 to 8 and 9 to 17. And the hazard rate during the years 4 to 8 is also significantly higher than the years 9 to 17.

Insert Table VII about here

The t statistics for the second model (allowing UH and AS) show that only the hazard rate of 1 year after issuance is significantly lower compared to years 2 and 3, 4 to 8 and 9 to 17. And the hazard rates during the years 2 to 17 are not significantly different from each other. It should be emphasized that the models are estimated using the same sample and with identical variables. Furthermore, it should be noted that adding the noise does not only make the difference between years 2 to 17 insignificant but also the default risk at year 1 becomes significantly different from the one in years 9 to 17. Therefore, the change in the estimated term structure cannot be only regarded to the increased standard deviation of the coefficients. Figure 5 illustrates again the change in the term structure of default between the two models.

Insert Figure 5 here

IV Conclusions

This study has several contributions to understanding of default risk. The paper confirms existence of unobserved heterogeneity (adverse selection) in the bond markets. It also shows that adverse selection and another effect (possibly cash-flow effect) determine the term structure of the default risk. The results cannot provide a solid evidence on the moral hazard effect projected by Diamond (1989).

The decomposition if the term structure to its deriving force can be insightful in both risk management and credit spread pricing. Ratings are shown to be merely a noisy signal of the default risk state of a firm and not its state of the nature. It is also shown that ignoring the unobserved heterogeneity (i.e. the fact that ratings are merely noisy signals) can bring to some biased results.

The findings concerning the term structure of default amplify the need for a more precise definition of ratings' objective. If ratings try the assess the instantaneous rate of default, then ratings might be subject to a downgrade drift in the second year following issuance of bonds. It might be that the downgrade drift documented since the late 1970's is partially due to that. On the other hand, if ratings are assessments of a firm's long term default risk, then current risk assessment should involve a measure for the time passed since the last bond-issuance in addition to the rating.

The paper presents a comprehensive framework for studying default risk, but several enhancements are still possible and even desired. For example, it is interesting to try other distributions than the Gamma distribution. In the current state of the research on default risk, a Gamma distribution is as legitimate as others. However, it was employed in this research merely due to its convenient properties and not due to its fit to previous observations or theoretical insights. Another desired enhancement is introduction of heteroscedasticity. It is reasonable that the larger the distance from the rating assignment, the lower it's relevance and the greater the unobserved heterogeneity in the default risk. Allowing heteroscedasticity might show that the term structure of default is significantly non-decreasing or even significantly increasing.

Implementation of the methodology presented in this paper with a larger sample, can shade a light on many open questions. For example, it would be possible to estimate the term structure of the hazard rate for each rating category and to test the separability of the hazard rate function (as assumed in this paper and others). It would be also possible to identify the covariates that determine a firm's hazard rate to default including macroeconomic variables. The adjustment of ratings to business cycles and the declining quality in the credit markets can be also tested by using this framework.

Appendix A

The Gamma Distribution

The Gamma function $\Gamma(\alpha)$ is defined

$$\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} dz, \qquad \alpha > 0$$

where $\Gamma(\alpha) > 0$ and $\Gamma(1) = 1$. It can easily be proven that $\Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$ and therefore if α is an integer, $\Gamma(\alpha+1) = \alpha!$. The first derivative of the logarithm of the gamma function is called digamma function and denoted by-

$$\psi(\alpha) = \frac{d \log \Gamma(\alpha)}{d\alpha}.$$

Consider a variable y where $z = y \cdot \eta$ $(\eta > 0)$ and accordingly $dz = \eta \cdot dy$. Then it can easily be shown that,

$$\Gamma(\alpha) = \eta^{\alpha} \int_0^\infty y^{\alpha} e^{-\eta \cdot y} dy.$$

Since the integrated is non-negative, the function

$$f(y) = \frac{\eta^{\alpha} y^{\alpha-1} e^{-\eta \cdot y}}{\Gamma(\alpha)}, \qquad y \geq 0,$$

is non-negative as well. Furthermore $0 \leq f(y) \leq 1$ and $\int_0^\infty f(y) dy = 1$. $f(y)$ is the probability density function of the family of Gamma distributions. It can be shown that $E(y)=\alpha\cdot\eta$, $\operatorname{var}(y)=\alpha\cdot\eta^2$, $E(\log y)=-\log\frac{1}{\eta}+\psi(a)$, and $\operatorname{var}(\log y)=\psi'(a)$.

The Gamma distribution of unit mean $\varsigma(1,\eta)$ has $E(y)=\alpha\cdot\eta=1$ and $\operatorname{var}(y)=\eta$.

References

Altman, Edward I., 1968, Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance* 23, 189-209

Altman, Edward I., and D. L. Kao, 1992, Rating Drift of High Yield Bonds, *Journal of Fixed Income* March, 15-20

Blume, M., F. Lim and C. Mackinlay, 1998, 'The Declining Credit Quality of U.S. Corporate Debt: Myth or Reality?, *Journal of Finance* 53(4), 1389-1413

Carty, L., 1997, Moody's Rating Migration and Credit Quality Correlation, *Moody's Sepcial Report* July

Carty, L., and J. Fons, 1994, Measuring Changes in Credit Quality, *Jouornal of Fixed Income* June, 27-41

Cox, David R., 1972, Regression Models and Life Tables, *Journal of the Royal Statistical Society B* 34, 187-220

Duffie, D., and K. J. Singleton, 1999, Modeling Term Structure of Defaultable Bonds, *Review of Financial Studies* 12, 687-720

Galil, Koresh A., 2003, The Quality of Corporate Credit Rating: an Empirical Investigation, Miemo, Tel-Aviv University

Jarrow, R. A., David Lando, and S. Turnbull, 1997, A Markov Model for the Term structure of Credit Risk Spreads, *Review of Financial Studies* 10(2), 481-523

Jarrow, R. A, and S. Turnbull, 1995, Pricing Derivatives on Financial Securities Subject to Credit Risk, *Journal of Finance* 50, 53-86

Lando, David and Torben M. Skoedeberg, 2002, Analyzing Rating Transitions and Rating Drift with Continuous Observations, *Journal of Banking & Finance* 26, 423-444

Lancaster, Tony, 1990, The Econometric Analysis of Transition Data, Econometric Society Monograph No.17. Cambridge: Cambridge University Press

Merton, Robert C., 1974, On the Pricing of Corporate debt: The Risk Structure of Interest Rates, *Journal of Finance* 29, 449-470

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

Prentice, Ross L. and Lynn A. Gloeckler, 1978, Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data, *Biometrics* 34, 57-67

Shumway, T., 2001, Forecasting Bankruptcy More Accurately: a Simple Hazard Model, *Journal of Business*

Tables

Table I

The Time structure of Average Hazard Rate using S&P Historical Average Cumulative Default Probabilities

Table I-a describes the historical average cumulative default rates (in percents) of the main rating categories as documented by S&P. Denote by $F_r(t)$ the average cumulative probability of default of rating r from time of rating till time t. Table I-b describes $f_r(t) = F_r(t) - F_r(t-1)$ - the average probability of default of rating r between time $t-1$ and time t. Table I-c describes $\theta_r(t) = f_r(t)/[1 - F_r(t-1)]$ the average hazard rate of default between time $t-1$ and time t.

Table I-a – Average cumulative default rate - $F_r(t)$

	Years from rating														
Rating	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
AAA	0.00	0.00	0.03	0.06	0.10	0.18	0.26	0.40	0.45	0.51	0.51	0.51	0.51	0.51	0.51
AA	0.01	0.04	0.09	0.16	0.25	0.37	0.53	0.63	0.70	0.79	0.85	0.92	0.96	1.01	1.07
A	0.04	0.11	0.19	0.32	0.49	0.65	0.83	1.01	1.21	1.41	1.56	1.65	1.70	1.73	1.83
BBB	0.22	0.50	0.79	1.30	1.80	2.29	2.73	3.10	3.39	3.68	3.91	4.05	4.22	4.37	4.48
BB	0.98	2.97	5.35	7.44	9.22	11.11	12.27	13.35	14.29	15.00	15.65	16.00	16.29	16.36	16.36
B	5.30	11.28	15.88	19.10	21.44	23.20	24.77	26.01	26.99	27.88	28.48	28.96	29.34	29.68	29.96
CCC	21.94	29.25	34.37	38.24	42.13	43.62	44.40	44.82	45.74	46.53	46.84	47.21	47.66	48.29	48.29
Investemnt Grade	0.08	0.19	0.31	0.51	0.72	0.95	1.17	1.37	1.54	1.71	1.84	1.93	2.00	2.06	2.14
Speculative Grade	4.14	8.34	11.93	14.67	16.84	18.64	19.98	21.09	22.05	22.85	23.46	23.88	24.22	24.45	24.58

Table I-b – Average probability of default - $f_r(t)$

	Years from rating														
Rating	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
AAA	0.00	0.00	0.03	0.03	0.04	0.08	0.08	0.14	0.05	0.06	0.00	0.00	0.00	0.00	0.00
AA	0.01	0.03	0.05	0.07	0.09	0.12	0.16	0.10	0.07	0.09	0.06	0.07	0.04	0.05	0.06
A	0.04	0.07	0.08	0.13	0.17	0.16	0.18	0.18	0.20	0.20	0.15	0.09	0.05	0.03	0.10
BBB	0.22	0.28	0.29	0.51	0.50	0.49	0.44	0.37	0.29	0.29	0.23	0.14	0.17	0.15	0.11
BB	0.98	1.99	2.38	2.09	1.78	1.89	1.16	1.08	0.94	0.71	0.65	0.35	0.29	0.07	0.00
B	5.30	5.98	4.60	3.22	2.34	1.76	1.57	1.24	0.98	0.89	0.60	0.48	0.38	0.34	0.28
CCC	21.94	7.31	5.12	3.87	3.89	1.49	0.78	0.42	0.92	0.79	0.31	0.37	0.45	0.63	0.00
Investemnt Grade	0.08	0.11	0.12	0.20	0.21	0.23	0.22	0.20	0.17	0.17	0.13	0.09	0.07	0.06	0.08
Speculative Grade	4.14	4.20	3.59	2.74	2.17	1.80	1.34	1.11	0.96	0.80	0.61	0.42	0.34	0.23	0.13

Table I-c – Average hazard rate of default - $\theta_r(t)$

	Years from rating														
Rating	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
AAA	0.00	0.00	0.03	0.03	0.04	0.08	0.08	0.14	0.05	0.06	0.00	0.00	0.00	0.00	0.00
AA	0.01	0.03	0.05	0.07	0.09	0.12	0.16	0.10	0.07	0.09	0.06	0.07	0.04	0.05	0.06
A	0.04	0.07	0.08	0.13	0.17	0.16	0.18	0.18	0.20	0.20	0.15	0.09	0.05	0.03	0.10
BBB	0.22	0.28	0.29	0.51	0.51	0.50	0.45	0.38	0.30	0.30	0.24	0.15	0.18	0.16	0.12
BB	0.98	2.01	2.45	2.21	1.92	2.08	1.30	1.23	1.08	0.83	0.76	0.41	0.35	0.08	0.00
B	5.30	6.31	5.18	3.83	2.89	2.24	2.04	1.65	1.32	1.22	0.83	0.67	0.53	0.48	0.40
CCC	21.94	9.36	7.24	5.90	6.30	2.57	1.38	0.76	1.67	1.46	0.58	0.70	0.85	1.20	0.00
Investemnt Grade	0.08	0.11	0.12	0.20	0.21	0.23	0.22	0.20	0.17	0.17	0.13	0.09	0.07	0.06	0.08
Speculative Grade	4.14	4.38	3.92	3.11	2.54	2.16	1.65	1.39	1.22	1.03	0.79	0.55	0.45	0.30	0.17

Table II

The Average S&P One-Year Rating Transition Matrix, 1981-1996

Table IV-a describes the average S&P one-year transition matrix during the years 1981-1996.[•] Table IV-b describes the corresponding S&P one-year transition matrix when withdrawal of rating is assumed to be a no-change in the rating and default is considered to be a terminal state.

Table II-a – Average S&P One-Year Transition Matrix with State of Rating Withdrawal (RW)

Rating	AAA	AA	A	BBB	BB	B	CCC	D	RW
AAA	88.5	8.1	0.7	0.1	0.1	0.0	0.0	0.0	2.6
AA	0.6	88.5	7.6	0.6	0.1	0.1	0.0	0.0	2.4
A	0.1	2.3	87.6	5.0	0.7	0.2	0.0	0.4	3.6
BBB	0.0	0.3	5.5	82.5	4.7	1.0	0.1	0.2	5.7
BB	0.0	0.1	0.6	7.0	73.8	7.6	0.9	1.0	8.9
B	0.0	0.1	0.2	0.4	6.0	72.8	3.4	4.9	12.2
CCC	0.2	0.0	0.3	1.0	2.2	9.6	53.1	19.3	14.2

Table II-b – Average S&P One-Year Transition Matrix, Where Rating Withdrawal (RW) Is Assumed as a Non-change in Rating and Default is Assumed as a Terminal State

Rating	AAA	AA	A	BBB	BB	B	CCC	D
AAA	91.1	8.1	0.7	0.1	0.1	0.0	0.0	0.0
AA	0.6	90.9	7.6	0.6	0.1	0.1	0.0	0.0
A	0.1	2.3	91.2	5.0	0.7	0.2	0.0	0.4
BBB	0.0	0.3	5.5	88.2	4.7	1.0	0.1	0.2
BB	0.0	0.1	0.6	7.0	82.7	7.6	0.9	1.0
B	0.0	0.1	0.2	0.4	6.0	85.0	3.4	4.9
CCC	0.2	0.0	0.3	1.0	2.2	9.6	67.3	19.3
D	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

[•] Source: Standard & Poor's (1997), "Rating Performance: Stability and Transition. Special Report".

Table III

Distribution of Annual Observations of Exposure to Default Risk by the Number of Years from Issuance and Default

This table presents the distribution of annual observations of exposure to default risk according to the number of years past from the issuance and the occurrence of default in that year. Note that the number of issuances is 2596, from which 235 ended with default during the subsequent 17 years. Since in most cases the time to default was larger then one, the number of years of exposure to default risk is 27906 and exceeds the number of issuances.

Year from Issue	*Not Default*	*Default*	*Total*
1	2584	12	2596
2	2545	38	2583
3	2508	37	2545
4	2483	22	2505
5	2455	27	2482
6	2414	19	2433
7	2386	23	2409
8	2108	16	2124
9	1787	10	1797
10	1556	9	1565
11	1315	5	1320
12	1139	7	1146
13	957	5	962
14	718	3	721
15	408	2	410
16	205	0	205
17	103	0	103
Total	**27671**	**235**	**27906**

Table IV

Distribution of Annual Observations of Exposure to Default Risk by Rating at Issuance and Default

This table presents the distribution of annual observations of exposure to default risk by the rating at the issuance and the occurrence of default during that year. Note that the number of issuances is 2596, from which 235 ended with default during the subsequent 17 years. Since in most cases the time to default was larger then one, the number of years of exposure to default risk is 27906 and exceeds the number of issuances.

Rating at Issuance	*Not Default*	*Default*	*Total*
AAA	1073	1	1074
AA	4710	9	4719
A	7999	10	8009
BBB	5761	24	5785
BB	1979	21	2000
B	5404	137	5541
CCC	740	30	770
CC	5	3	8
Total	*27671*	*235*	*27906*

Table V

Estimation of the Survival Function

The table contains results of two regressions for estimation of the survival function. The first model (ignoring unobserved heterogeneity) assumes that the hazard function is proportional with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t)$. The second model (allowing unobserved heterogeneity) also assumes that the hazard rate is proportional but with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t) \cdot v_i$ where v_i is Gamma distributed with unit mean (standardization). The parameters β, γ (a set of parameters which each integrates $k_2(t)$ the term-structure-component of the hazard rate), and the variance of v_i (in the second model) are estimated using a maximum likelihood approach. In these regressions x_{it} includes only dummy variables indicating the issue's rating at the time of issuance (x_{it} can be notated x_i).

	Ignoring Unobserved Heterogeneity and Aggregate Shocks		Allowing for Unobserved Heterogeneity and Aggregate Shocks	
Explanatory Variable	**Coefficient**	**t Statistic**	**Coefficient**	**t Statistic**
Term Structure (Years from Issuance):				
1 year	-3.822***	(-11.610)	-3.384***	(-3.200)
2 years	-2.642***	(-11.630)	-1.944*	(-1.720)
3 years	-2.627***	(-11.440)	-1.758	(-1.530)
4 years	-3.116***	(-11.600)	-2.090*	(-1.770)
5 years	-2.887***	(-11.450)	-1.607	(-1.350)
6 years	-3.210***	(-11.410)	-1.603	(-1.310)
7 years	-2.996***	(-11.300)	-1.145	(-0.920)
8 years	-3.217***	(-10.810)	-1.117	(-0.870)
9 years	-3.511***	(-9.910)	-0.902	(-0.680)
10 years	-3.526***	(-9.560)	-0.761	(-0.560)
11 years	-3.931***	(-8.280)	-1.161	(-0.820)
12 years	-3.465***	(-8.460)	-0.786	(-0.550)
13 years	-3.585***	(-7.530)	-0.987	(-0.670)
14 years	-3.727***	(-6.180)	-1.191	(-0.780)
15 or 16 or 17 years	-4.091***	(-5.600)	-1.543	(-0.950)

*, **, *** are significant at the 10 percent, 5 percent and 1 percent respectively.

Table V - *continued*

	Ignoring Unobserved Heterogeneity and Aggregate Shocks		Allowing for Unobserved Heterogeneity and Aggregate Shocks	
Explanatory Variable	**Coefficient**	**t Statistic**	**Coefficient**	**t Statistic**
Calendar Year:				
1985	-	-	-0.135	(-0.120)
1986	-	-	0.536	(0.490)
1987	-	-	0.384	(0.350)
1988	-	-	-0.943	(-0.840)
1989	-	-	-0.464	(-0.420)
1990	-	-	0.095	(0.090)
1991	-	-	0.523	(0.480)
1992	-	-	-0.933	(-0.830)
1993	-	-	-1.216	(-1.080)
1994	-	-	-2.491**	(-2.070)
1995	-	-	-1.405	(-1.240)
1996	-	-	-2.468**	(-2.070)
1997	-	-	-2.147*	(-1.820)
1998	-	-	-1.780	(-1.520)
1999	-	-	-1.448	(-1.240)
2000	-	-	-1.081	(-0.920)
Rating:				
AAA	-3.829***	(-3.770)	-4.492***	(-3.970)
AA	-3.094***	(-8.220)	-3.781***	(-6.360)
A	-3.529***	(-9.770)	-4.184***	(-7.170)
BBB	-2.340***	(-8.720)	-2.903***	(-5.740)
BB	-1.401***	(-5.020)	-1.893***	(-3.950)
B	-0.546***	(-2.820)	-0.741**	(-2.210)
Unobserved Heterogeneity:				
Variance	-	-	2.924**	(1.827)
No. of observations	27906		27906	
Log Likelihood	-1181.762978		-1116.062604	

Table VI

Estimation of the Hazard Function with Grouping of Years from Issuance

The table contains results of two regressions for estimation of the survival function. The first model (ignoring unobserved heterogeneity) assumes that the hazard function is proportional with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t)$. The second model (allowing unobserved heterogeneity) also assumes that the hazard rate is proportional but with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t) \cdot v_i$ where v_i is Gamma distributed with unit mean (standardization). The parameters β, γ (a set of parameters which each integrates $k_2(t)$ the term-structure-component of the hazard rate), and the variance of v_i (in the second model) are estimated using a maximum likelihood approach. In these regressions x_{it} includes dummy variables indicating the issue's rating at the time of issuance and dummy variables indicating the year of observation of exposure to default risk. In these regressions, the years from issuance are grouped into dummy variables for, 1 year after issuance, 2 to 3 years from issuance, 4 to 8 years from issuance and 9 to 17 years from issuance.

	Ignoring Unobserved Heterogeneity and Aggregate Shocks		Allowing for Unobserved Heterogeneity and Aggregate Shocks	
Explanatory Variable	**Coefficient**	**t Statistic**	**Coefficient**	**t Statistic**
Term Structure (Years from Issuance):				
1 year	-3.818***	(-11.600)	-3.488***	(-3.360)
2 or 3 years	-2.630***	(-13.350)	-2.088*	(-1.890)
4 to 8 years	-3.068***	(-16.220)	-2.054*	(-1.820)
9 to 17 years	-3.624***	(-16.120)	-1.821	(-1.500)
Calendar Year:				
1985	-	-	-0.146	(-0.130)
1986	-	-	0.507	(0.470)
1987	-	-	0.278	(0.260)
1988	-	-	-0.963	(-0.860)
1989	-	-	-0.448	(-0.410)
1990	-	-	0.099	(0.090)
1991	-	-	0.571	(0.530)
1992	-	-	-0.812	(-0.730)
1993	-	-	-1.008	(-0.910)
1994	-	-	-2.223*	(-1.880)
1995	-	-	-1.082	(-0.980)
1996	-	-	-2.124*	(-1.820)
1997	-	-	-1.774	(-1.550)
1998	-	-	-1.379	(-1.220)
1999	-	-	-0.984	(-0.880)
2000	-	-	-0.646	(-0.580)

Table VI – *continued*

Explanatory Variable	Ignoring Unobserved Heterogeneity and Aggregate Shocks		Allowing for Unobserved Heterogeneity and Aggregate Shocks	
	Coefficient	t Statistic	Coefficient	t Statistic
Rating:				
AAA	-3.834***	(-3.780)	-4.171***	(-3.840)
AA	-3.101***	(-8.250)	-3.454***	(-6.670)
A	-3.535***	(-9.790)	-3.857***	(-7.600)
BBB	-2.346***	(-8.740)	-2.604***	(-6.030)
BB	-1.408***	(-5.040)	-1.654***	(-3.980)
B	-0.551***	(-2.840)	-0.653**	(-2.270)
Unobserved Heterogeneity:				
Variance	-	-	1.653	(1.350)
No. of observations	27906		27906	
Log Likelihood	-1183.3783		-1121.9761	

*, **, *** are significant at the 10 percent, 5 percent and 1 percent respectively.

Table VII
Statistics for Differences between Years from Issuance

The table shows the t statistics for the differences between the parameters describing the term structure of the hazard rate to default. The statistics are based on estimation of two models (reported in table V), where the first model (ignoring unobserved heterogeneity) assumes that the hazard function is proportional with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t)$. The second model (allowing unobserved heterogeneity) also assumes that the hazard rate is proportional but with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t) \cdot v_i$ where v_i is Gamma distributed with unit mean (standardization). The parameters γ (a set of parameters which each integrates $k_2(t)$ the term-structure-component of the hazard rate) are estimated using a maximum likelihood approach. In these regressions, the years from issuance are grouped into dummy variables for, 1 year after issuance, 2 to 3 years from issuance, 4 to 8 years from issuance and 9 to 17 years from issuance.

a. Ignoring Unobserved Heterogeneity and Annual Shocks

Years after Issuance	1	2 and 3	4 to 8
2 and 3	14.590***		
4 to 8	6.070**	8.440***	
9 to 17	0.350	26.160***	9.160***

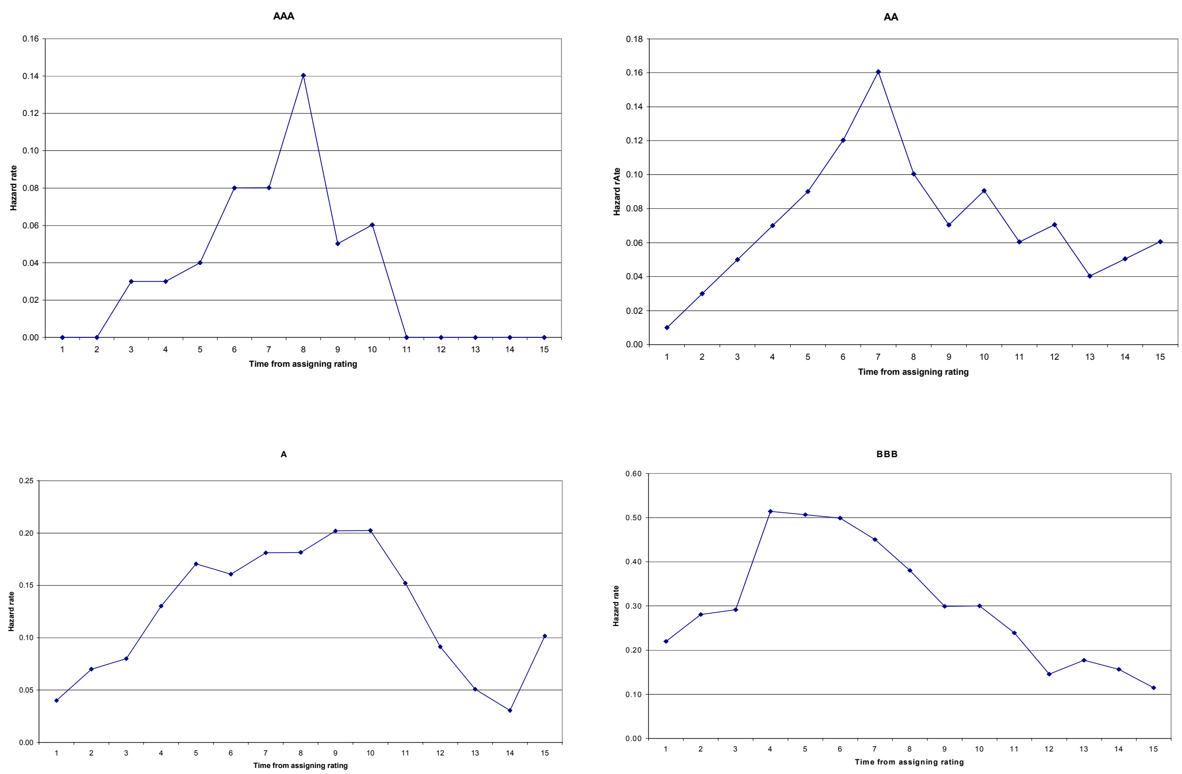
b. Allowing Unobserved Heterogeneity and Annual Shocks

Years after Issuance	1	2 and 3	4 to 8
2 and 3	16.960***		
4 to 8	13.130***	0.030	
9 to 17	8.700***	0.400	0.560

*, **, *** these statistics reflect significance of 10 percent, 5 percent and 1 percent respectively.

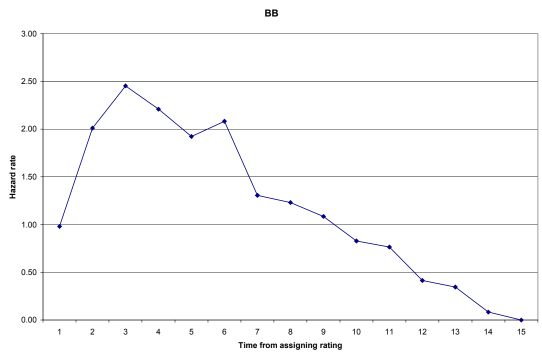
Figure 1 - The Term Structure of Average Hazard Rate Using S&P Historical Average Cumulative Default Probabilities

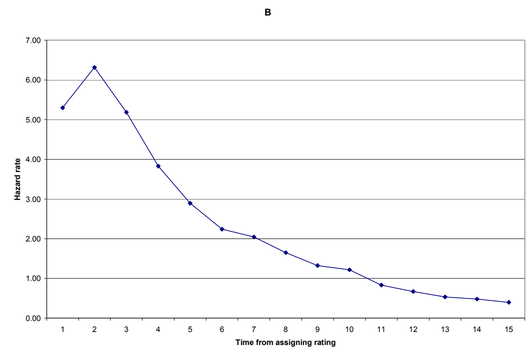
Figures A-1a to A-1i describe the average hazard rate of default for each rating category as a function of time from rating. These figures are based on average cumulative default probability as documented by S&P.[•] Let $F_r(t)$ denote the average cumulative probability of default of rating r, t years after assigning the rating. Then $f_r(t) = F_r(t) - F_r(t-1)$ denote the average probability of default of rating r between time $t-1$ and time t and $\theta_r(t) = f_r(t)/[1 - F_r(t-1)]$ the average hazard rate of default between time $t-1$ and time t.



[•] "Ratings Performance 2000", Standard & Poor's.

Figure 1 – *continued*





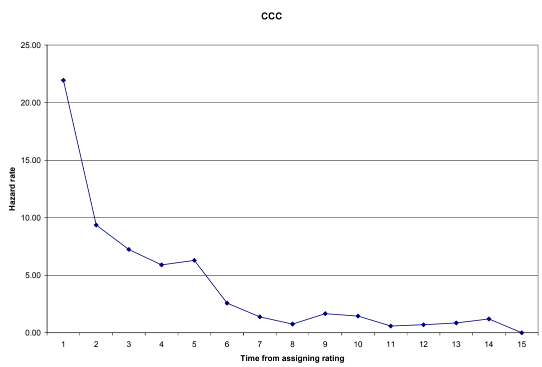
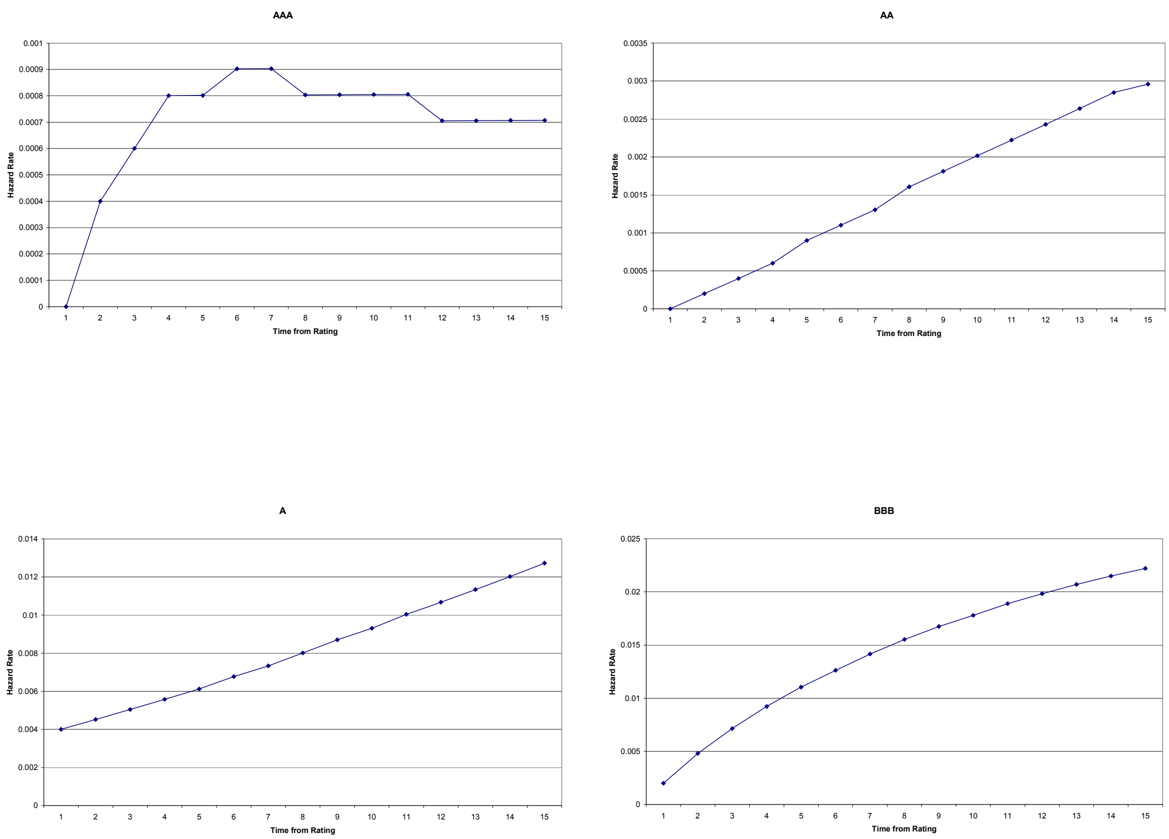
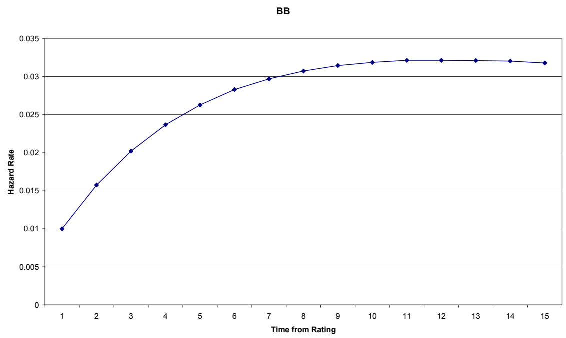
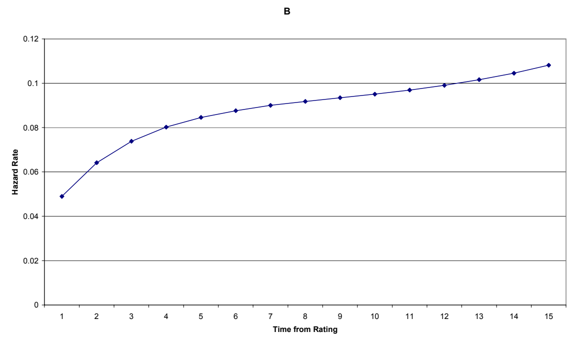


Figure 2 - The Term Structure of the Average Hazard Rate Using One-Year S&P Ratings Transition Matrix

This figure shows term structures of the hazard rate to default computed by using the S&P one-year rating transition matrix during the years 1981-1996. Suppose the rating of the firm follows a Markovian process that can be specified by a $k \times k$ transition matrix Q where each $0 \le q_{rs} \le 1$ is the probability of transition from rating r to rating s within one year. The t-years transition matrix, $Q_{0,t}$ whose (r,s) entry is $q_{rs}(0,t)$ satisfies $Q_{0,t} = Q^t$. The k-th column represents the set of the cumulated probabilities of default till time t. Let $F_r(t) \equiv q_{rk}(0,t)$, then the probability of default at time t of a firm rated r at time 0 can be calculated by $f_r(t) \equiv F_r(t) - F_r(t-1)$ and the hazard rate to default at time t of a firm rated r at time 0 can be also calculated by $\theta_r(t) \equiv f_r(t)/\left[1 - F_r(t)\right]$.




BB
Hazard Rate
Time from Rating


B
Hazard Rate
Time from Rating

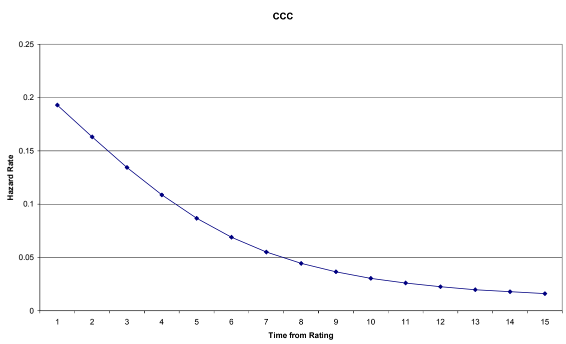

CCC
Hazard Rate
Time from Rating

Figure 3 - The Estimated Term Structure of the Hazard Rated to Default

This figure shows the estimated term structures of two hazard models (see Table IV). The first model (ignoring unobserved heterogeneity) assumes that the hazard function is proportional with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t)$. The second model (allowing unobserved heterogeneity) also assumes that the hazard rate is proportional but with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t) \cdot v_i$ where v_i is Gamma distributed with unit mean (standardization). This figure shows the estimates for γ - a set of parameters which each integrates $k_2(t)$ the term-structure-component of the hazard rate within each year.

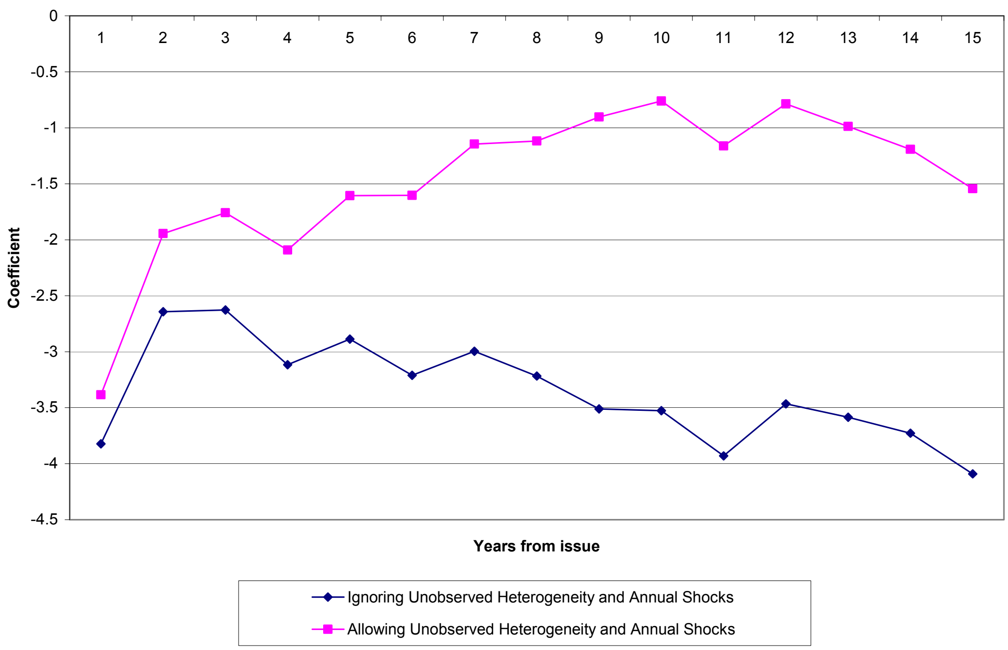


Figure 4 – Frequencies of the Time to Maturity of the Issues

The figure describes the frequency (in percent) of the time to maturity a sample of 2596 bonds that were used to create the estimation sample.

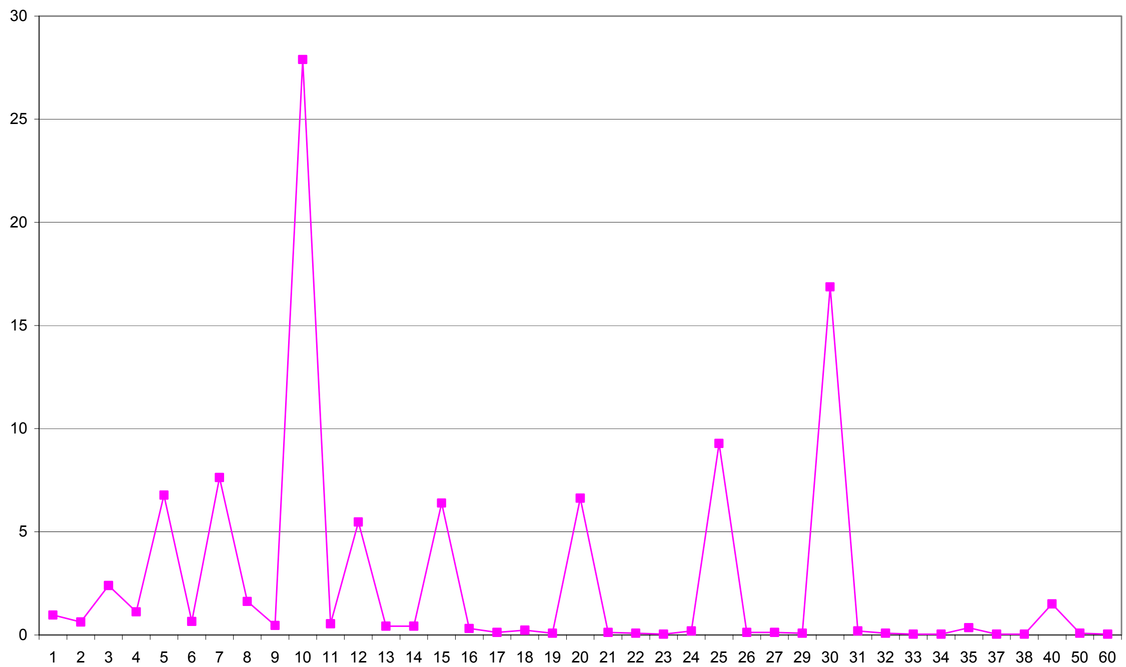


Figure 5 - The Estimated Term Structure of the Hazard Rated to Default when Using Grouped Dummy Variables for the Year from Issuance

This figure shows the estimated term structures of two hazard models (see Table VI) when the dummy variables for the year from issuance are grouped into 4 dummy variables - 1 year, 2 or 3 years, 4 to 8 years, and 9 to 17 years from issuance. The first model (ignoring unobserved heterogeneity) assumes that the hazard function is proportional with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t)$. The second model (allowing unobserved heterogeneity) also assumes that the hazard rate is proportional but with the form $\theta(t, x_{it}) = \exp(x_{it}'\beta) \cdot k_2(t) \cdot v_i$ where v_i is Gamma distributed with unit mean (standardization). This figure shows the estimates for γ - a set of parameters which each integrates $k_2(t)$ the term-structure-component of the hazard rate within each year.