TITLE: Sirket Temerrütleri ve Diger Türden Sirket Çikislarinin Tahmini için Bir Sirali Nitel Tepki Modelleme Yaklasimi

AUTHORS: Sinan ALÇIN,T Sabri ÖNCÜ

PAGES: 159-183

ORIGINAL PDF URL: https://dergipark.org.tr/tr/download/article-file/358244

An Ordered Qualitative Response Modeling Approach for the Estimation of Corporate Defaults and Other Forms of Exit

Sinan ALÇIN[*], T. Sabri ÖNCÜ[**]

Abstract

We propose a new approach for the estimation of defaults and other forms of exit of borrowers. Our approach is based on the ordered qualitative response model. We first show that any ordered qualitative response model is equivalent to the competing risks model – commonly employed in the estimation of corporate defaults and other forms of exit – in continuous-time. We then construct the continuous-time likelihood function of the models and further present its discrete-time simplification. Lastly, we compare and contrast the competing risks and ordered qualitative response models through numerical experiments in a two-state setting, and demonstrate that none of the alternatives necessarily dominates the others. Our results indicate that it may be worthwhile to estimate the models in continuous-time.

Keywords: Banking, Risk Management, Finance, Econometrics, Financial Econometrics

Jel Classification: E50, G32, C58

Şirket Temerrütleri ve Diğer Türden Şirket Çıkışlarının Tahmini için Bir Sıralı Nitel Tepki Modelleme Yaklaşımı

Öz

Temerrütler ve diğer türden şirket çıkışlarının tahmini için yeni bir yaklaşım öneriyoruz. Yaklaşımımız sıralı nitel tepki modeli üzerine kuruluyor. Önce, sıralı nitel tepki modelinin – şirket temerrütleri ve diğer türden şirket çıkışları tahmininde sıkça kullanılan – rakip riskler modeline sürekli zamanda denk olduğunu gösteriyoruz. Sonra, modellerin sürekli zaman olabilirlik fonksiyonunu kuruyor ve daha sonra, bu fonksiyonun basitleştirilmiş kesikli zaman şeklini sunuyoruz. Son olarak, iki durumlu bir uzayda yaptığımız sayısal deneylerle rakip riskler ve sıralı nitel tepki modellerini

[*] Assoc. Prof. Dr., Istanbul Kultur University, Department of Economics, Istanbul, Turkey, E-mail: s.alcin@iku.edu.tr ORCID ID http://orcid.org/0000-0002-2330-0693
[**] Visiting Research Professor, New York University, Stern School of Business, New York, United States, E-mail: sabri.oncu@gmail.com
ORCID ID http://orcid.org/0000-0002-1092-7497

karşılaştırıyoruz ve seçeneklerden hiçbirinin diğerine yeğlenir olması gerekmediğini gösteriyoruz. Elde ettiğimiz sonuçlar modellerin sürekli zamanda tahminlerinin yararlı olabileceğini gösteriyor.

Anahtar Kelimeler: Bankacılık, Risk Yönetimi, Finans, Ekonometri, Finansal Ekonometri

Jel Kodları: E50, G32, C58

1. Introduction

We are motivated by the ongoing push on the banks by their regulators around the globe to adopt Basel II – and, now, its revision, Basel III – to meet their capital adequacy requirements for the loans they make. According to the Basel Capital Accord, banks can employ one of two approaches: standard and advanced. If a bank chooses the standard approach, then the domestic banking regulator requires the bank to employ the standard rules of the Basel Capital Accord, applied uniformly across all participating banks to determine the risk weights for various loan categories to set aside capital to meet their regulatory obligations.

If a bank chooses the advanced approach, on the other hand, although it is still subject to the rules of the Accord, it can determine its own risk weights to various loan categories based on its internal credit risk models. This requires, among other things, the development of an internal-rating system to assess the default probabilities and the potential loss given default – that is, the credit risk – of the loans. However, passing from the standard to the advanced approach is usually subject to an assessment of and the approval by the domestic banking regulator of the bank's internal credit risk models.

It is with this in mind that we propose a qualitative response modeling approach for defaults and other forms of exit of borrowers that can be used to assess the quality of internal credit risk models employed by financial institutions. The model we propose is an alternative to the commonly employed duration models and can help the banks to choose from a larger menu of models to build their internal-rating models.

In a seminal article in 1975, Amemiya defined the qualitative response models "generically as models that involve one or more discrete random variables whose conditional probability distribution given the values of the independent variables is specified up to a finite number of parameters".[1] The most well-known qualitative response models are Logit and Probit, although there are others. In the rating transitions and defaults literature, multi-period Logit and Probit have been the most popular among the qualitative response

---

[1] Takeshi Amemiya, "Qualitative Response Models", *Annals of Economic and Social Management*, 1974, 4(3), 363-372

models. These models are generally referred to as "static" models in finance literature mainly because they have been employed traditionally in pooled estimations with the data from each firm in each period as if it were a separate observation.[2]

The simplest duration model is the survival model, which has just two states: alive and dead. Another well-known duration model is the competing risks model, which has many states: one alive, the rest are dead for different reasons, such as a heart attack or a traffic accident and the like. While the survival model is used in the modeling of defaults extensively, the competing risks model is used in the modeling of defaults as well as other forms of exit such as delisting from the exchanges or leveraged buyouts or mergers and acquisitions, to name but a few forms of exits other than defaults.[3]

The main proposition of this paper is that the ordered qualitative response models are equivalent to the competing risks model in continuous-time. Our proposition takes as special case the proposition of Shumway (2001) that "a multi-period 'Logit' model is equivalent to a discrete-time hazard model" and extends it to continuous-time. An important implication of our proposition is that ordered qualitative response models are not "static". We then construct the likelihood function of our models in continuous – as well as in discrete – time, and compare and contrast the duration and ordered qualitative response models through numerical experiments in a two-state setting.

Although our discussion will have to be mathematical for inevitable reasons, we will try to keep the level of our mathematical exposition as easily accessible to bankers and other practitioners as we can. Therefore, we assume once and for all that all regularity conditions necessary for the derivation of our results hold. We organize the rest of the paper as follows. In Section 2, we discuss the competing risks formulation of defaults and other forms of exit, and show that any K-state ordered qualitative response model is equivalent to the K-state competing risks model in continuous-time. In Section 3, we construct our 3-state qualitative response model likelihood function in continuous-time – as well as in discrete-time – for the estimation of defaults and other forms of exit, and give a brief discussion of how the borrower specific variables, and the default and other exit processes can be estimated

---

[2] Tyler Shumway, "Forecasting bankruptcy more accurately: A simple hazard model", *Journal of Business*, 2001, Vol. 74, 101–124

[3] See, for example, Sanjiv Das, Darrel Duffie, Nikunj Kapadia and Leandro Saita, "Common failings: How corporate defaults are correlated", *Journal of Finance*, 2007, 62, 93–117; Darrel Duffie, Andreas Eckner, Guillaume Horel and Leandro Saita, "Frailty correlated default", *Journal of Finance*, 2009, 64, 2089–2123; Darrel Duffie, Leandro Saita and Ke Wang, "Multi-period corporate default prediction with stochastic covariates", *Journal of Financial Economics*, 2007, 83(3), 635–665

jointly in our framework. We devote Section 4 to our numerical experiments and, finally, in Section 5, we conclude.

2. The Competing Risks Formulation of Defaults and Other Forms of Exit

2.1 Preliminaries

The competing risks model is a particular duration model. Duration models are characterized by a stochastic process, say, $Z(t)$, which takes values from a set of states, say, $S = \{1,2,\dots,K\}$, over a period $\mathcal{T} = [a,T]$ or $\mathcal{T} = [a,T)$ with $T \leq \infty$, where K is the number of states. Let Z_t denote the history generated by the stochastic process $Z(t)$ up to time t. Suppose that there is a collection of some external variables, say, $X(t)$, that influences the time evolution of the stochastic variable $Z(t)$. Lastly, let X_t be the history up to time t of the external covariates $X(t)$ and denote by $Z_t \vee X_t$ the combined history of $Z(t)$ and $X(t)$.

With the above definitions, the state *transition probabilities* from time s to time t, $\boldsymbol{P}(s,t|Z_s \vee X_s) = \big[P_{ij}(s,t|Z_s \vee X_s)\big]$, relative to the history $Z_s \vee X_s$, are defined as

$$P_{ij}(s,t|Z_s \vee X_s) = Prob(Z(t) = j|Z(s) = i, Z_s \vee X_s) \tag{1}$$

for all $i,j \in S$ and $s < t \in \mathcal{T}$. If the transition probabilities depend on the combined history $Z_s \vee X_s$ only through the value of external covariates at time s, that is, if

$$P_{ij}(s,t|Z_s \vee X_s) = Prob(Z(t) = j|Z(s) = i, X(s)) = P_{ij}(s,t|X(s)), \tag{2}$$

then the process $Z(t)$ is called *Markov*. We assume now that the process $Z(t)$ is Markov and suppress the $X(t)$ dependence in the transition probabilities for convenience. Since we assumed that the process is Markov, it follows from the above that in this period

$$\boldsymbol{P}(s,u) = \boldsymbol{P}(s,t)\boldsymbol{P}(t,u), \qquad s < t < u \in \mathcal{T}. \tag{3}$$

Associated with the above transition probabilities are the *transition intensities* $\boldsymbol{\lambda}(t) = \big[\lambda_{ij}(t)\big]$ given by the derivatives

$$\lambda_{ij}(t) = \lim_{\Delta t \to 0} \frac{P_{ij}(t,t+\Delta t)}{\Delta t}, i \neq j \tag{4}$$

and we set

$$\lambda_i(t) = -\lambda_{ii}(t) = \sum_{j \neq i} \lambda_{ij}(t). \tag{5}$$

It is clear from the above that the transition intensities measure the speeds with which the transition probabilities change. If for a state $i \in S$, $\lambda_{ij}(t) = 0$ for all $t \in \mathcal{T}, j \in S$, then this state is called *absorbing*, because this indicates that once the stochastic process $Z(t)$ enters that state, it remains there forever. The state $i \in S$ is called *transient*, otherwise.

2.2 The Competing Risks Model

In the case of competing risks model, only one of the states, say, state 1, is transient whereas all other states are absorbing. For convenience in what follows, we set $K = 3$ and note that our results readily generalize to any K. Therefore, our transition probability matrix $\boldsymbol{P}(s,t)$ takes the form

$$\boldsymbol{P}(s,t) = \begin{bmatrix} P_{11}(s,t) & P_{12}(s,t) & P_{13}(s,t) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

so that from the equations (3) and (6) we have

$$P_{11}(a,t)P_{1j}(t,t+\Delta t) = P_{1j}(a,t+\Delta t) - P_{1j}(a,t). \tag{7}$$

It then follows from the equations (4), (5) and (7) that

$$\lambda_1(t) = -\frac{d}{dt} \ln P_{11}(a,t) \tag{8}$$

and

$$\lambda_{1j}(t) = \frac{1}{P_{11}(a,t)} \frac{d}{dt} P_{1j}(a,t) \tag{9}$$

for any $t \in \mathcal{T}$ whereas for all $i, j \in S$ and $i \neq 1$ we have $\lambda_{ij}(t) = 0$. This luxury arises for two reasons. Firstly, we have assumed that the process is Markov over $\mathcal{T}$. Secondly, since state 1 is the only transient state, transitioning from state 1 to state 1 from time a to time t cannot include any transition from state 1 to any other state, for otherwise, the process $Z(t)$ is absorbed.

Solving the equations (8) and (9) subject to initial conditions $P_{11}(a,a) = 1$ and $P_{1j}(a,a) = 0$ for all $j = 2, ..., K$, we ge

$$P_{11}(a,t) = \exp\{- \int_a^t \lambda_1(u) du\} \tag{10}$$

and

$$P_{1j}(a,t) = \int_a^t \lambda_{1j}(z) \exp\{- \int_a^z \lambda_1(u) du\} dz,$$
$$j = 2, ..., K. \tag{11}$$

If $K = 2$, then the above are the well-known equations of the survival model with the hazard function $\lambda_1(t) = \lambda_{12}(t)$ and the survival function $P_{11}(a, t)$. We will refer to all of the transition intensities $\lambda_{1j}(t)$ where $j \neq 1$ as the hazard functions of the competing risks model.

2.3 The Equivalence of Ordered Qualitative and Competing Risks Models

Let us set $K = 2$ for a while so that

$$P_{12}(a, t) = 1 - P_{11}(a, t) \tag{12}$$

and let $\mu_{a,1}(t - a)$ be the scalar valued function which solves the equation

$$\Phi\left(\mu_{a,1}(t - a)\right) = P_{11}(a, t), \tag{13}$$

where $\Phi(\cdot)$ is the cumulative probability distribution function of any qualitative response model. By the properties of $\Phi(\cdot)$, such a solution exists, is unique, and $\lim_{t \to a} \mu_{a,1}(t - a) = \infty$. Hence, we proved the proposition of Shumway (2001) in continuous-time.

Proposition 1. Any two-state qualitative response model is equivalent to the survival model in continuous-time.

Let us now go back, set $K = 3$, and suppose that $\mu_{a,1}(t - a)$ still solves the equation (13) with $K = 3$. In this case, we have

$$P_{13}(a, t) = 1 - \left(P_{11}(a, t) + P_{12}(a, t)\right). \tag{14}$$

Let us now suppose that $\mu_{a,2}(t - a)$ is the scalar valued function that solves the equation

$$\Phi\left(\mu_{a,2}(t - a)\right) = P_{11}(a, t) + P_{12}(a, t). \tag{15}$$

Again, by the properties of $\Phi(\cdot)$, such a solution exists, is unique and $\lim_{t \to a} \mu_{a,2}(t - a) = \infty$. Furthermore, it is evident that we have $\mu_{a,2}(t - a) > \mu_{a,1}(t - a)$ for all $t \in \mathcal{T}$. It then follows from the above that

$$P_{11}(a, t) = \Phi\left(\mu_{a,1}(t - a)\right),$$

$$P_{12}(a, t) = \Phi\left(\mu_{a,2}(t - a)\right) - \Phi\left(\mu_{a,1}(t - a)\right), \tag{16}$$

$$P_{13}(a, t) = 1 - \Phi\left(\mu_{a,2}(t - a)\right).$$

Hence, we proved that any three-state ordered qualitative response model is equivalent to the three-state competing risks model in continuous-time. Proceeding in the same manner, we have the following proposition.

Proposition 2. Any K-state ordered qualitative response model is equivalent to the K-state competing risks model in continuous-time.

We will refer to the functions $\mu_{a,j}(t-a), j = 1,2, \dots, K-1$, as the cutoff functions of the K-state ordered qualitative response model.

2.3 The Relationship Between Hazard and Cutoff Functions

In this section, we are interested in connecting the hazard functions of the competing risks model to our cutoff functions. To this end, let us denote by $\phi(\cdot)$ the probability distribution function associated with the cumulative probability distribution function $\Phi(\cdot)$. From (9) and (16) we get for any $t > a$

$$\lambda_{1j}(t)$$
$$= \frac{\phi\left(\mu_{a,j}(t-a)\right)\mu'_{a,j}(t-a) - \phi\left(\mu_{a,j-1}(t-a)\right)\mu'_{a,j-1}(t-a)}{\Phi\left(\mu_{a,1}(t-a)\right)}, \qquad " \qquad (17)$$

$j = 2, \dots, K-1$, and we have

$$\lambda_{1K}(t) = -\frac{\phi\left(\mu_{a,K-1}(t-a)\right)\mu'_{a,K-1}(t-a)}{\Phi\left(\mu_{a,1}(t-a)\right)} . \qquad (18)$$

It is clear from the equation (18) that when $K = 2$, we are back to the usual equation of the survival model. Since the hazard functions are positive, this equation shows also that $\mu'_{a,K-1}(t-a)$ is negative, that is, $\mu_{a,K-1}(t-a)$ is a decreasing function of its argument. Furthermore, since $\lim_{t\to\infty}\phi(t) = 0$, we also have $\lim_{t\to a}\mu'_{a,K-1}(t-a) = -\infty$. Indeed, using the equations (17) and (18), and rearranging, we get

$$\phi\left(\mu_{a,j}(t-a)\right)\mu'_{a,j}(t-a)$$
$$= -\sum_{k=j+1}^{K}\lambda_{1k}(t)\Phi\left(\mu_{a,1}(t-a)\right). \qquad (19)$$

This shows that all of the $\mu_{a,j}(t-a)$ must be decreasing functions of time with $\lim_{t\to a}\mu'_{a,j}(t-a) = -\infty, j = 1, \dots, K-1$.

To give a simple example, let us set $K = 2$ and suppose that the hazard function of the survival model is constant, that is, $\lambda_{12}(t) = \lambda > 0$. Since in this case there is just one cutoff function, let us drop its subscript and set

$a = 0$ for convenience. Let us now look at the Logit model whose cumulative probability distribution function is

$$\Phi\big(\mu(t)\big) = \frac{1}{1 + \exp\{-\mu(t)\}}. \tag{20}$$

Since we assumed that the hazard function of the survival model is constant, from (10) and (19) we get

$$\frac{1}{1 + \exp\{-\mu(t)\}} = \exp\{-\lambda t\} \tag{21}$$

so that $\mu(t) = -\ln[\exp(\lambda t) - 1]$, which is clearly a decreasing function of time, $\lim_{t \to 0} \mu(t) = \infty$ and $\lim_{t \to 0} \mu'(t) = -\infty$.

3. The Likelihood Function

3.1 A Single Borrower

In this subsection, we will look at a single borrower and suppose that the external covariates $X(t)$ are sampled with intervals of equal length Δt. This is a realistic assumption, since in the case of default estimation the external covariates would include borrower specific accounting and macroeconomic variables that can be observed at best monthly, if not quarterly (see, for example, Duffie et al, 2007). Suppose that the observations were made in the interval $[0, T]$ where $T = N\Delta t$ and N is the number of intervals. Let us set $t_i = i\Delta t, i = 0,1, \dots, N$. Lastly, we have two absorbing states: the other types of exit state, state 2, and the default state, state 3. Let $\tau = \min(T, T_1, T_2)$ where T_1 and, T_2 are the other exit and default times, respectively.

Let us now suppose that $Z^*(t)$ is a latent variable that generates the default and other forms of exit process $Z(t)$ for any $t \in [t_{i-1}, t_i), i = 1, \dots, N$ such that

$$Z(t) = 1, \ if \ \ Z^*(t) \le \mu_{t_{i-1},1}(t - t_{i-1}),$$

$$Z(t) = 2, \ if \ \mu_{t_{i-1},1}(t - t_{i-1}) < \ Z^*(t)$$
$$\le \mu_{t_{i-1},2}(t - t_{i-1}), \tag{22}$$

$$Z(t) = 3, \ if \ \ Z^*(t) > \mu_{t_{i-1},2}(t - t_{i-1}).$$

We define $Y_j(t) = 1\{Z(t) = j\}, j = 1,2,3$ where $1\{\cdot\}$ is the indicator function which returns the value one if its argument is true and zero,

otherwise. We choose $\mu_{t_{i-1},j}(t_i - t_{i-1}) = \mu_j(t_{i-1}) = \alpha_j(X(t_{i-1})), j = 1,2$ where $\alpha_2(X(t_{i-1})) > \alpha_1(X(t_{i-1}))$ are functions to be specified.

Lastly, we assume for any $t \in [t_{i-1}, t_i), i = 1, ..., N$

$$Z^*(t) = \varepsilon(t) \tag{23}$$

where $\varepsilon(t)$ is either identically and independently distributed standard normal for Probit or identically and independently distributed logistic for Logit. Under these assumptions, the one-period transition probabilities in each interval $[t_{i-1}, t_i), i = 1, ..., N$ are given by

$$P_{11}(t_{i-1}, t_i) = \Phi(\mu_1(t_{i-1})),$$

$$P_{12}(t_{i-1}, t_i) = \Phi(\mu_2(t_{i-1})) - \Phi(\mu_1(t_{i-1})), \tag{24}$$

$$P_{13}(t_{i-1}, t_i) = 1 - \Phi(\mu_2(t_{i-1})).$$

The above fixes our ordered qualitative response function formulation of the one-period transition probabilities. Our second alternative is specifying the hazard function. So we set $\lambda_{1j}(t) = \lambda_{1j}(t_{i-1}) = \gamma_{j-1}(X(t_{i-1})), j = 2,3$ for any $t \in [t_{i-1}, t_i), i = 1, ..., N$ where $\gamma_j(X(t_{i-1}))$ are functions to be specified and recall that $\lambda_1(t) = \lambda_{12}(t) + \lambda_{13}(t)$. In this case, the one-period transition probabilities given by the equations (10) and (11) take the form

$$P_{11}(t_{i-1}, t_i) = \exp(-\lambda_1(t_{i-1})\Delta t)$$

$$P_{12}(t_{i-1}, t_i) = \frac{\lambda_{12}(t_{i-1})}{\lambda_1(t_{i-1})}\{1 - \exp(-\lambda_1(t_{i-1})\Delta t)\}, \tag{25}$$

$$P_{13}(t_{i-1}, t_i) = \frac{\lambda_{13}(t_{i-1})}{\lambda_1(t_{i-1})}\{1 - \exp(-\lambda_1(t_{i-1})\Delta t)\}.$$

Suppose now that at some time $\tau \in [t_{k-1}, t_k)$ for some $k, 0 < k \le N$, that is, in the kth period, either an exit or a default occurred. Then, the process gets absorbed in this interval and, depending on the type of "failure", one of the following probabilities $P_{11}(t_0, \tau)P_{1j}(\tau, \tau + \kappa)$ for $j = 2,3$ needs to be computed for small κ. Since κ is small, from the equation (11) through straightforward calculations we get

$$P_{1j}(\tau, \tau + \kappa) = \lambda_{1j}(t_{k-1}) \kappa + O(\kappa^2), \qquad j = 2,3 \tag{26}$$

where we used the usual big-O notation.

Since

$$P_{11}(t_0, t_{k-1}) = \prod_{n=1}^{k-1} P_{11}(t_{n-1}, t_n), \tag{27}$$

it remains to determine $P_{11}(t_{k-1}, \tau)$ to write down the equation for $P_{11}(t_0, \tau)$. It is evident from the equation (10) that

$$P_{11}(t_{k-1}, \tau) = \exp\{-\lambda_1(t_{k-1})(\tau - t_{k-1})\}, \tag{28}$$

so we are done with the case when the one-period transition probabilities are modelled through the hazard function.

Let us now suppose that the one-period transition probabilities are modelled through the ordered qualitative response models and observe from the equations (24) and (25) that

$$\lambda_1(t_{k-1}) = \frac{1}{\Delta t} \ln\left[\frac{1}{\Phi(\mu_1(t_{k-1}))}\right], \tag{29}$$

$$\lambda_{12}(t_{k-1})$$
$$= \frac{1}{\Delta t} \ln\left[\frac{1}{\Phi(\mu_1(t_{k-1}))}\right] \left\{\frac{\Phi(\mu_2(t_{k-1})) - \Phi(\mu_1(t_{k-1}))}{1 - \Phi(\mu_1(t_{k-1}))}\right\}, \tag{30}$$

$$\lambda_{13}(t_{k-1})$$
$$= \frac{1}{\Delta t} \ln\left[\frac{1}{\Phi(\mu_1(t_{k-1}))}\right] \left\{\frac{1 - \Phi(\mu_2(t_{k-1}))}{1 - \Phi(\mu_1(t_{k-1}))}\right\}. \tag{31}$$

Lastly, we note from the equations (28) and (29) that

$$P_{11}(t_{k-1}, \tau) = \Phi(\mu_1(t_{k-1}))^{\frac{\tau - t_{k-1}}{\Delta t}}. \tag{32}$$

Therefore, the likelihood function for this borrower is

$$\mathcal{L}_C(\theta) = P_{11}(t_0, \tau)\{Y_1(\tau) + Y_2(\tau)\lambda_{12}(t_{k-1}) + Y_3(\tau)\lambda_{13}(t_{k-1})\} \tag{33}$$

where θ is the parameter vector to be estimated after either the functions $\alpha_j(X(t_{i-1})), j = 1,2$ or the functions $\gamma_j(X(t_{i-1})), j = 1,2$ are specified. Of course, it is evident from the above that specifying the functions $\alpha_j(X(t_{i-1})), j = 1,2$ is one possible way of specifying the functions $\gamma_j(X(t_{i-1})), j = 1,2$ in this framework. Note that if no failure occurred, then $\tau = T$ so that $\mathcal{L}_C(\theta) = P_{11}(t_0, T)$.

This is the first alternative. The second alternative is to ignore that the default or exit occurred in the interior of the interval $[t_{k-1}, t_k)$, set $\tau = t_k$ and, instead, use the likelihood function

$$\mathcal{L}_D(\theta) = P_{11}(t_0, t_{k-1})\{Y_1(t_k)P_{11}(t_{k-1}, t_k) + Y_2(t_k)P_{12}(t_{k-1}, t_k) + Y_3(t_k)P_{13}(t_{k-1}, t_k)\} \qquad (34)$$

This amounts to discretizing the continuous exit and default process $Z(t)$. Although this discretization leads to some loss of information, the likelihood function becomes straightforward.

3.2 Many Borrowers

For convenience in what follows, we suppose that there is no *left truncation*, that is, no late entry of some of the borrowers into the data set. Similarly, we suppose that there is no *right censoring*, that is, no early departure of some borrowers from the data set except because of defaults or other forms of exit. In reality, however, borrowers may get left truncated, right censored, then left truncated again, then right censored again and so forth. That is, borrowers may come late, then drop out, then come back again and so forth without defaulting or exiting, because of missing data or some other reason. Although all of these can be handled within the ordered qualitative model formulation of defaults and other forms of exit with no essential difficulty, we ignore these possibilities for convenience in the following discussion.

With the above caveat, let us now suppose that there are M borrowers $m = 1,2,\ldots,M$, and subscript the latent process that generates the default and other forms of exit of the mth borrower by m so that it is $Z_m^*(t)$. Suppose now that $Z_m^*(t)$ are independent across borrowers and denote by $\mathcal{L}_m(\theta)$ either the continuous-time likelihood function (31) or the discrete-time likelihood function (32) for the mth borrower. Also let $X_m(t_i)$ denote the external covariates for the mth borrower, sampled at the discrete time points $t_i = i\Delta t, i = 0,1,\ldots,N$, and assume that $Z_m^*(t_i)$ and $X_m(t_i)$ are independent for each m. The $X_m(t_i)$ consists of two components, the economy-wide covariates $E(t_i)$ and the borrower specific covariates $B_m(t_i)$, that is, $X_m(t_i) = (E(t_i), B_m(t_i))$.

Lastly, let $\tau_m \in [t_{k(m)-1}, t_{k(m)})$ be the last time the mth borrower is observed in the $k(m)^{th}$ interval due to failure either in the form of default or other exit, or just the observation period ended. Suppose that failures occurred in L intervals beginning at times $t_{k(l)-1}$ where $l = 1,2,\ldots,L$, and set $t_{k(0)} = t_1$. Further, set $t_{k(L+1)} = t_{N+1} > t_N$, which is any time after the end of the observation period so that we can count the remaining borrowers, if any, on the last date of the observation period conveniently.

Let us now order the borrowers according to their failure times from the latest to the earliest, denote by $J(l), l = 1,2,\ldots,L$, the number of remaining borrowers at time $t_{k(l)}$ and set $J(0) = M$, the total number of borrowers in the beginning. Let us now define $M(t_i), i = 1,2,\ldots,N$, which counts the number of borrowers at time t_i as follows:

$$M(t_i) = J(l) \quad if \ t_i \in \left[t_{k(l)}, t_{k(l+1)}\right), l = 0,1,2,\ldots,L, i = 1,2,\ldots,N.$$

With these definitions, we now introduce the combined external covariates vector at time t_i defined as $C(t_i) = (E(t_i), B_1(t_i), B_2(t_i), \ldots, B_{M(t_i)}(t_i))$, and let $\Psi_{t_i}(\cdot \,|C(t_{i-1}), \vartheta)$ denote the joint probability density function of $C(t_i)$ conditioned on $C(t_{i-1})$ where ϑ is the parameter vector of this density function.

Under the above set up, the likelihood function of the above observations over the period $[0, T]$ is given by

$$\mathcal{L}(\vartheta, \theta) = \mathcal{L}(\vartheta)\mathcal{L}(\theta) \tag{35}$$

where

$$\mathcal{L}(\vartheta) = \prod_{i=1}^{N} \Psi_{t_i}(C(t_i)|C(t_{i-1}), \vartheta) \tag{36}$$

whereas

$$\mathcal{L}(\theta) = \prod_{m=1}^{M} \mathcal{L}_m(\theta) \tag{37}$$

This completes the construction of our likelihood function. Under our independence assumptions, it is clear that we can estimate the likelihood functions given by (36) and (37) separately. Once the model is estimated, the forward survival, other exit and default probabilities can be calculated through Monte Carlo simulations.

3.3 A Brief Note on the Possibility of Joint Estimation of Defaults, Other Exits and External Covariates

One advantage of our ordered qualitative response modeling approach is that it provides a simple tool for the joint estimation of defaults, other forms of exit and external covariates, if we choose our ordered qualitative response model as the ordered Probit. Recall that we assumed that the latent processes $Z_m^*(t)$ and the external covariates $X_m(t)$ are independent for each borrower m as well as that $Z_m^*(t)$ are independent across borrowers, $m = 1,2,\ldots,M$.

While the latter of these assumptions is plausible – with which we continue – the former may be questioned because the external covariates $X_m(t)$ contain borrower specific variables. Recall that $X_m(t) = (E(t), B_m(t))$

where $E(t)$ are the economy wide covariates and $B_m(t)$ are the borrower specific covariates. Unless there is a latent factor such as a crisis that influences all of the borrowers collectively, and leads to correlated defaults and contagion, an individual borrower's impact on the overall economy may be ignored. Therefore, in the absence of correlated defaults and contagion, it is plausible that economy wide covariates are independent from the individual borrower variables, which we continue to assume, also.[4] In the absence of correlated defaults and contagion, it is plausible also that the borrower specific external variables are independent across borrowers, so suppose that $B_m(t)$ are independent across borrowers, as well.

Let us now focus on a single borrower and drop the subscript m for convenience. Also for convenience, suppose that there is one borrower specific variable, say, $B(t)$. Lastly, suppose that there are no other forms of exit so that a borrower either defaults or survives. Let state 1 be the survival and state 2 be the default state, and focus on two observations in discrete-time, at times $t = t_0$ and $t = t_1 > t_0$. The process $Z(t)$ is initialized at $t = t_0$ as $Z(t_0) = 1$ and the latent process $Z^*(t)$ determines the next state according as

$$Z(t_1) = 1, \quad if \quad Z^*(t_1) \le a \quad and \quad Z(t_1) = 2, \quad if \quad Z^*(t_1) > a \tag{37}$$

For further simplicity, consider the following processes:

$$B(t_1) = b + \eta \tag{38}$$
$$Z^*(t_1) = \epsilon \tag{39}$$

where b is some constant, and η and ϵ are some random variables. Since both $B(t_1)$ and $Z^*(t_1)$ are variables associated with the same borrower, it may be desirable to consider the possibility that they may be dependent. This possibility can be addressed in our framework with no essential difficulty and without the curse of dimensionality, as we demonstrate below.

To this end, let us suppose that η and ϵ are jointly normally distributed with mean zero and the covariance matrix

$$\Sigma = \begin{bmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{bmatrix}$$

---

[4] We address the issue of correlated defaults and contagion elsewhere.

In this case, the joint probability distribution is given by

$$\phi_2(\epsilon, \eta)$$

$$= \frac{1}{2\pi\sqrt{(1-\rho^2)\sigma^2}} exp\left\{-\frac{\eta^2 - 2\rho\sigma\eta\epsilon + \sigma^2\epsilon^2}{2(1-\rho^2)\sigma^2}\right\}. \tag{40}$$

Suppose now that the borrower did not default at time t_1. Then the likelihood of this observation is

$$\mathcal{L}(a,b) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{(B(t_1)-b)^2}{2\sigma^2}\right\} \Phi\left(\frac{a}{\sqrt{(1-\rho^2)}}\right.$$

$$\left. -\frac{\rho(B(t_1)-b)}{\sqrt{(1-\rho^2)\sigma^2}}\right). \tag{41}$$

whereas if the default occurred at time t_1, then the likelihood of this observation is

$$\mathcal{L}(a,b) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{(B(t_1)-b)^2}{2\sigma^2}\right\}\left\{1\right.$$

$$\left. -\Phi\left(\frac{a}{\sqrt{(1-\rho^2)}} -\frac{\rho(B(t_1)-b)}{\sqrt{(1-\rho^2)\sigma^2}}\right)\right\} \tag{42}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Since in most applications one works with a handful of borrower specific variables, the above can be extended to several borrower specific variables without much difficulty.

4. Numerical Experiments

In this section, we compare and contrast the predictions of the competing risks and ordered qualitative response models in a two-state setting for eight simulated data sets of 50,000 firm-periods. State 1 is the survival state whereas state 2 is the default state so that the competing risks model is simply the survival model. We use a version of the Cox model to model the hazard function of the survival model and work with the Probit model as a representative of the qualitative response models. We then set the data sampling period length as Δt=1 for convenience, suppose there is an external variable X that drives the defaults, and draw 50,000 values for X assuming that it is identically and independently distributed standard normal.

Finally, we generate our default data for two sets of parameter values from the following four models that give the survival probabilities as follows.

a) Probit: $P_{11}(0,1) = \Phi(\alpha_{PROBIT}(X))$, $\alpha_{PROBIT}(X) = A_{PROBIT} + X$,

b) Cox: $P_{11}(0,1) = \exp(-\lambda_{COX}(X))$, $\lambda_{COX}(X) = \exp(A_{COX} - X)$,

c) Chi Squared: $P_{11}(0,1) = X^2(\alpha_{CHI}(X))$, $\alpha_{CHI}(X) = \exp(A_{CHI} + X)$,

d) Beta: $P_{11}(0,1) = B(\alpha_{BETA}(X))$, $\alpha_{BETA}(X) = 1/(1 + \exp(A_{BETA} + X))$.

In the above, $\Phi(\cdot)$, $X^2(\cdot)$ and $B(\cdot)$ are the cumulative distribution functions of the standard normal, Chi Squared with unit degree of freedom and Beta with unit shape parameters, respectively. We choose the free model parameters A_{PROBIT}, A_{COX}, A_{CHI} and A_{BETA} to generate two data sets from each of the models in such a way that the resulting in-sample unconditional default probabilities in the first data sets are 5%, whereas they are 15% in the second data sets. Table 1 summarizes the chosen free parameter values.

Table 1. Summary of Simulated Data

This table summarizes the models and selected free parameter values that are used to generate the data of 50,000 firm-periods for each of the eight simulated data sets.

Simulation 1: Unconditional Default Probability: 5%

Data Generating Model	Probit	Cox	Chi Squared	Beta
Free Parameter Value	2.326	-3.41913	1.94021	-3.34253

Simulation 2: Unconditional Default Probability: 15%

Data Generating Model	Probit	Cox	Chi Squared	Beta
Free Parameter Value	1.4646	-2.17793	1.00125	-2.04379

Next, we estimate the models for each of the data sets under the assumption that the observations are made in discrete-time. Table 2 summarizes the estimation results. All parameter estimates and models are statistically significant at better than 1%. Although a comparison of models based on the likelihood ratios is not a proper comparison for non-nested models, we nevertheless see from this comparison in Panels A and B that when Probit is the data-generating model, Probit appears to fit the data better than Cox, whereas when Cox is the data-generating model, Cox appears to fit

the data better than Probit. These results are expected and included only as a check on our results.

Table 2. Comparison of Probit and Cox Models in Discrete-time

This table presents the estimation results of Probit and Cox models for eight simulated data sets of 50,000 firm-periods. All parameter estimates and models are statistically significant at better than 1%.

Panel A. Data Generating Model: Probit

	Simulation I		Simulation II	
Model	Probit	Cox	Probit	Cox
Constant	2.311	-4.168	1.462	-2.530
Slope	0.989	-1.668	0.997	-1.422
Loglikelihood				
Null	-9925.76	-9925.76	-21135.45	-21135.45
Model	-6827.81	-6912.10	-14943.32	-15126.97
Likelihood Ratio	6195.91	6027.33	12384.26	12016.97

Panel B. Data Generating Model: Cox

	Simulation I		Simulation II	
Model	Probit	Cox	Probit	Cox
Constant	1.843	-3.435	1.388	-2.194
Slope	0.507	-1.008	0.581	-0.999
Loglikelihood				
Null	-9925.76	-9925.76	-21135.45	-21135.45
Model	-8718.25	-8706.93	-18215.03	-17833.21
Likelihood Ratio	2415.02	2437.67	5840.86	6604.48

Panel C. Data Generating Model: Chi Squared

	Simulation I		Simulation II	
Model	Probit	Cox	Probit	Cox
Constant	2.250	-3.995	1.328	-2.295
Slope	0.933	-1.517	0.796	-1.115
Loglikelihood				

|				
Null	-9925.76	-9925.76	-21135.45	-21135.45
Model	-7082.03	-7259.74	-16578.74	-16876.59
Likelihood Ratio	5687.47	5332.05	9113.43	8517.72

Panel D. Data Generating Model: Beta

	Simulation I		Simulation II	
Model	Probit	Cox	Probit	Cox
Constant	1.819	-3.382	1.182	-2.116
Slope	-0.470	0.941	-0.547	0.870
Loglikelihood				
Null	-9925.76	-9925.76	-21135.45	-21135.45
Model	-8860.21	-8852.18	-18500.08	-18500.97
Likelihood Ratio	2131.11	2147.15	5270.76	5268.96

However, the results in Panels C and D are interesting. In these panels, the data generated from the Chi Squared and Beta models represent "real world" data whose distribution is "unknown". From these panels – again based on a comparison of the likelihood ratios – we see that when the data-generating model is Chi Squared, Cox appears to fit the data better than Probit, whereas when the data-generating model is Beta, the roles are reversed (although the models appear to tie for the second set of simulations in this case).
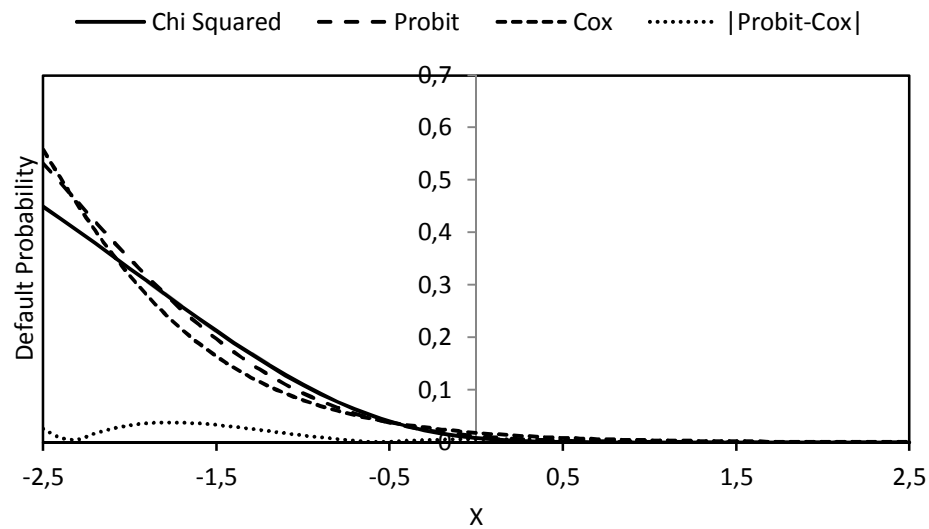
Figure 1 depicts the predicted one-period default probabilities from the Probit and Cox models as functions of the underlying external variable X, and compares their predictions with the "true" default probabilities for the first simulated data sets where the unconditional default probability is 5%.
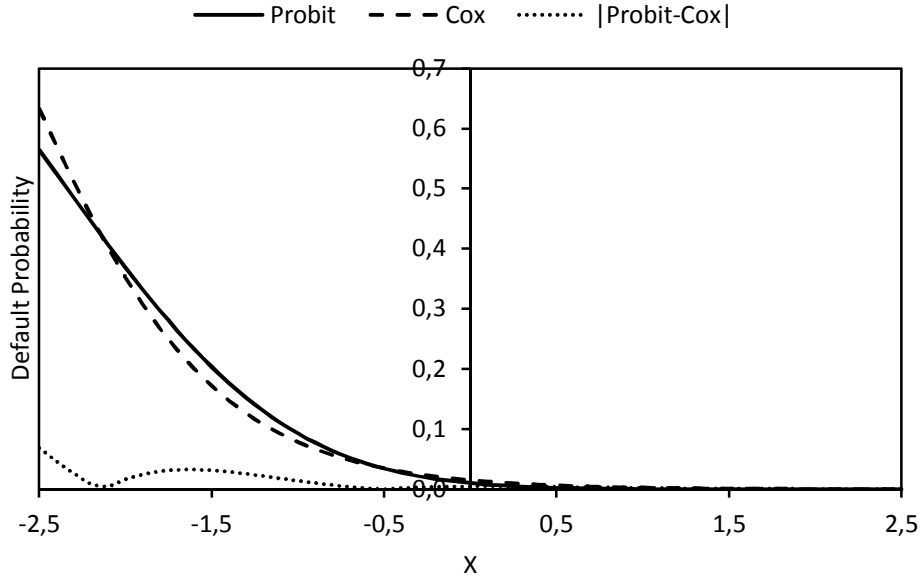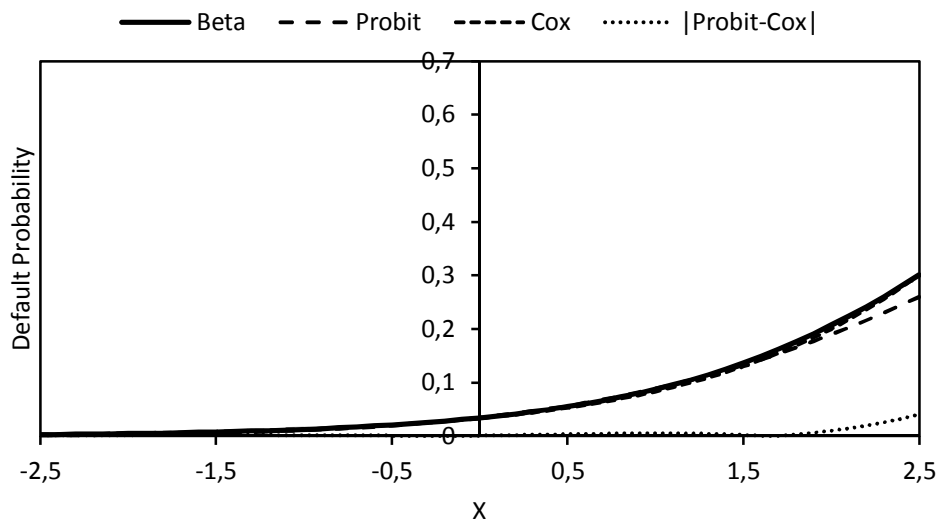
a) Data Generating Model: Probit
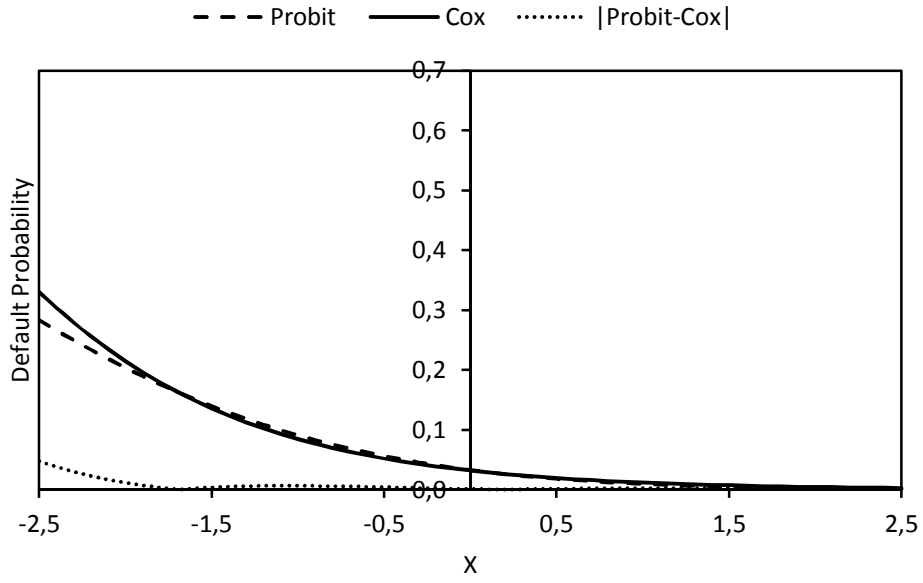
b) Data Generating Model: Cox

c) Data Generating Model: Chi Squared

d) Data Generating Model: Beta

Figure 1. Comparison of Probit and Cox Models in Discrete-time – In Sample Unconditional Default probability: 5%

The figures plot one-period default probabilities as estimated by discrete-time Probit, Cox and data-generating models as functions of the simulated firm-period variable X for the simulated data of 50,000 firm-periods. In Figures 1.a and 1.b, only the estimated models are included because the data-generating model and its estimation produce indistinguishable graphs.

In Figures 1.a and 1.b, we compare the Probit and Cox models when the data-generating model is one of them, and see that although the models agree in their predictions for the most part, the deviations at the left extreme – where the predicted default probabilities are large – can be as large as 7% for our simulated data sets. However, this does not constitute a problem for Basel capital requirements, because at that level of one-period of default probabilities, all boundaries are crossed despite the deviations.

In Figures 1.c and 1.d, we compare the Probit and Cox models when the data-generating processes are "unknown" (that is, when they are Chi Squared and Beta, respectively). We see from these figures that not only the Probit and Cox models generally agree with each other, but also they do not deviate from the "true" models significantly except at the extremes.

Lastly, we turn our attention to continuous-time, and focus on the last two data-generating models, that is, Chi Squared and Beta. We keep our original simulated data sets generated by these models, but randomly assign to the defaults a default time between zero and one. The results of our estimations based on these data are reported in Table 3.

Table 3. Comparison of Probit and Cox Models in Continuous-time

This table presents the estimation results of Probit and Cox models for four simulated data sets of 50,000 firm-periods. All parameter estimates and models are statistically significant at better than 1%.

Panel A. Data Generating Model: Chi Squared with Randomized Default Times

	Simulation I		Simulation II	
Model	Probit	Cox	Probit	Cox
Constant	2.278	-3.967	1.335	-2.304
Slope	1.014	-1.573	0.780	-1.088
Loglikelihood				
Null	-9905.86	-9905.86	-21048.58	-21048.58
Model	-6917.36	-7103.42	-16534.63	-16831.65
Likelihood Ratio	5977.00	5604.88	9026.74	8433.86

Panel B. Data Generating Model: Beta with Randomized Default Times

	Simulation I		Simulation II	
Model	Probit	Cox	Probit	Cox
Constant	1.817	-3.369	1.197	-2.136
Slope	-0.479	0.944	-0.545	0.860

Loglikelihood				
Null	-10011.31	-10011.31	-20936.61	-20936.61
Model	-8921.64	-8925.30	-18331.93	-18350.47
Likelihood Ratio	2179.34	2172.02	5209.36	5172.28

All of the parameter estimates and models reported in Table 3 are statistically significant at better than 1%. As before, we note that although a comparison between two non-nested models is not appropriate based on such criteria, our results show that our Probit formulation appears to fit the data better than the Cox model in all of the samples based on these criteria in our simulated samples.

Since in these sets of simulated data the "true" data generating processes are truly unknown, we compare predictions of the Probit and Cox models without any reference to any "true" default probabilities in Figure 2.
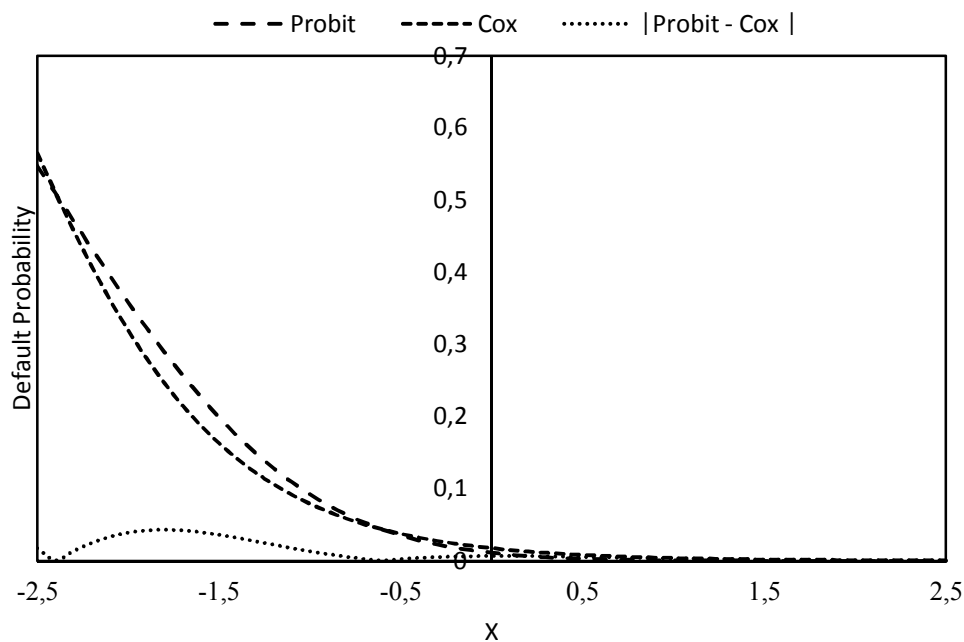
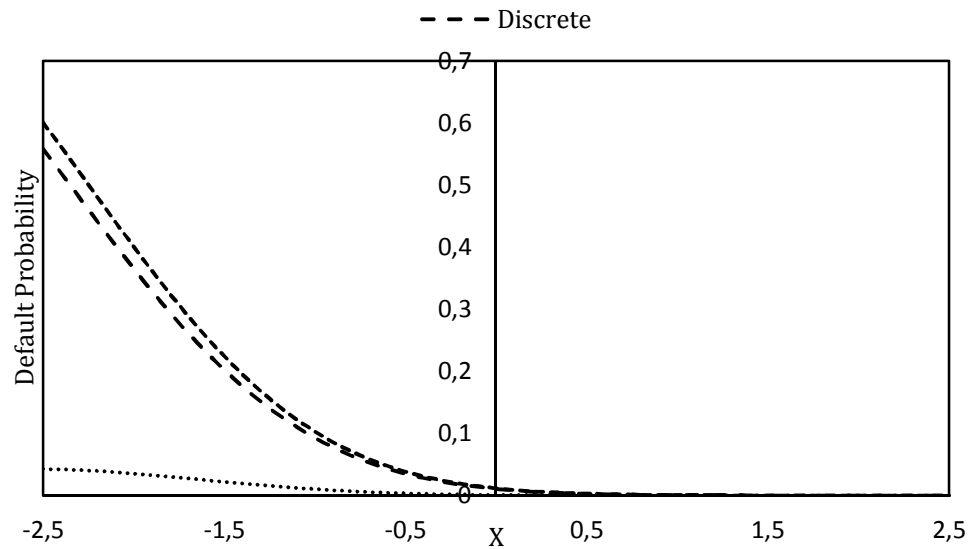a) Data Generating Model: Chi Squared with randomized default times
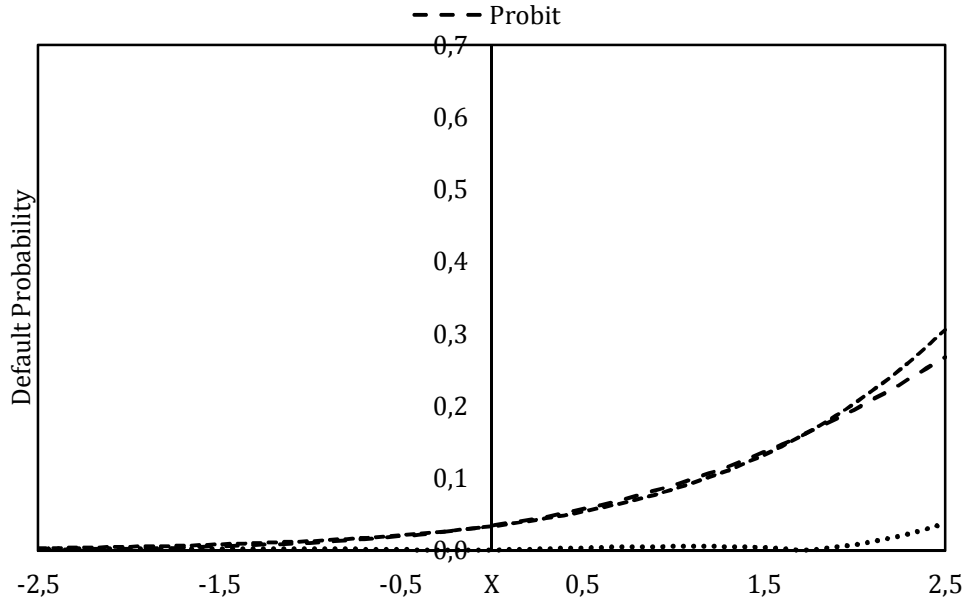
b) Data Generating Model: Beta with randomized default times

Figure 2. Comparison of Probit and Cox Models in Continuous-time – In Sample Unconditional Default Probability: 5%
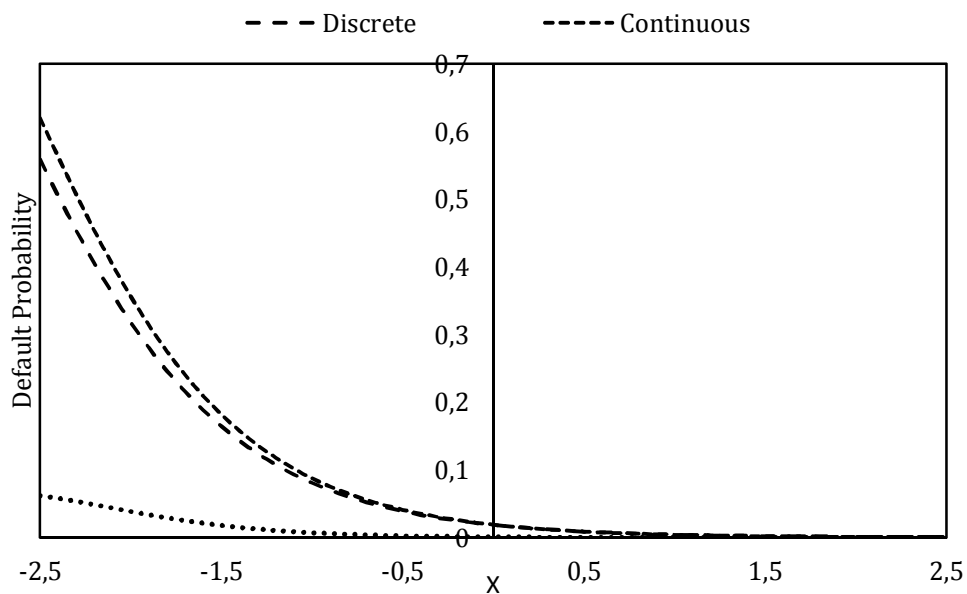
The figures plot one-period default probabilities as estimated by continuous-time Probit and Cox Models as functions of the simulated firm-period variable X for the simulated data of 50,000 firm-periods. The data-generating models are Chi Squared and Beta with randomized default times.

This figure shows for our simulated data that the Probit model appear to predict lower default probabilities than the Cox model, at least, at the extremes of the distribution. However, since the "true" data-generating processes are unknown in this case, it is not possible to reach any conclusion based on this figure regarding which of the models does better than the other. This is because the results depicted in this figure are specific to our simulated data sets. Therefore, there is no assurance that the predictions of the models depicted in this figure can be generalized to other data sets for which the underlying data-generating processes are different.

Notwithstanding this, however, the results depicted in Figure 3 show that not accounting for defaults that occur between discretized sampling times may lead to the underestimation of the default probabilities. Figures 3.a and 3.b compare the continuous-time and discrete-time predictions of the Probit and Cox models, respectively, and provides support for this.

a) Probit

b) Cox

Figure 3. Comparison of Discrete-time Probit and Cox Models with Their Continuous-time Equivalents – In Sample Unconditional Default Probability: 5%

The figures plot one-period default probabilities as estimated by Probit and Cox Models as functions of the simulated firm-period variable X for the simulated data of 50,000 firm-periods in continuous-time and discrete-time. The data-generating model is Chi Squared with randomized default times.

Let us close this section by summarizing our observations from the numerical experiments of this section as follows.

1) As long as the underlying data generating process for the defaults is not known, it is not possible to decide whether they are the qualitative response or duration models that are the better suited to the task, at least, in discrete-time;

2) No matter which class of models is chosen, it may be worthwhile to estimate the models in continuous-time, for otherwise, default probabilities may be underestimated.

5. Conclusions

Motivated by the ongoing push on the banks by their regulators around the globe to adopt Basel II – and, now, its revision, Basel III – to meet their capital adequacy requirements for the loans they make, we developed an approach for the estimation of defaults and other forms of exit of borrowers. Our approach is based on the ordered qualitative response models.

We first showed that ordered qualitative response models are equivalent to the commonly employed competing risks model in continuous-time, indicating that competing risks could be modeled using ordered qualitative response models in addition to duration models. We then constructed the likelihood function of ordered qualitative response models for the estimation of defaults and exits in continuous-time and presented a simpler discrete-time version. Furthermore, since borrower specific variables are expected to be dependent, we gave a brief discussion of how the borrower specific external variables, and the default and other exit processes can be jointly estimated in our framework. Lastly, we compared and contrasted the competing risks and ordered qualitative response models through numerical experiments in a two-state setting, and demonstrated that none of the alternatives necessarily dominates the others. Further, we demonstrated that it may be worthwhile to estimate the models in continuous-time.

Under the ongoing push of the banking regulators, the banks are expected to move from the standard approach to the advanced approach of the Basel Capital Accord. While the standard approach is easy to implement, the advanced approach requires the banks to develop their internal credit risk models. However, passing from the standard to the advanced approach is subject to an assessment of and the approval by the domestic banking regulator of the banks' internal credit risk models. We hope that the ordered qualitative response modeling approach we proposed provides the banks with a broader menu of models to choose from and that it offers a convenience to the banks and their regulators for the assessment of internal credit risk models.

REFERENCES

AMEMIYA, Takeshi, "Qualitative Response Models", *Annals of Economic and Social Management*, 4 (3), 1975, pp. 363-372.

DAS, Sanjiv, Darrel DUFFIE, Nikunj KAPADIA, and Leandro SAITA, "Common Failings: How Corporate Defaults are Correlated", *Journal of Finance*, 62, 2007, pp. 93–117.

DUFFIE, Darrel, Andreas ECKNER, Guillaume HOREL, and Leandro SAITA, "Frailty Correlated Default", *Journal of Finance*, 64, 2009, pp. 2089–2123.

DUFFIE, Darrel, Leandro SAITA, and Ke WANG, "Multi-period Corporate Default Prediction with Stochastic Covariates", *Journal of Financial Economics*, 83 (3), 2007, pp. 635–665.

SHUMWAY, Tyler, "Forecasting Bankruptcy More Accurately: A Simple Hazard Model", *Journal of Business*, 74, 2001, pp. 101–124.

Özet

Tüm dünyada, bankacılık düzenleyicilerinin bankaları yönelttikleri Basel II – ve şimdilerde revizyonu Basel III – sermaye yeterliliği düzenlemelerinden alınan ilham ile şirket temerrütleri ve diğer türden şirket çıkışlarının tahmini için bir yaklaşım önerilmektedir. Yaklaşımımız sıralı nitel tepki modelleme yaklaşımıdır. İlk olarak, bu amaçla kullanılan rekabet eden riskler modeli ile sıralı nitel tepki modellerinin denk oldukları gösterilmiştir. Yapılan, rekabet eden risklerin sıralı nitel tepkiler olarak da modellenebileceğini göstermektir.

Bunu takiben, şirket temerrütleri ve diğer türden şirket çıkışlarının tahmini için sürekli zamanda ihtimal fonksiyonu kurulmuş ve sonrasında bu fonksiyon, bir de basitleştirilmiş kesikli zaman ihtimal fonksiyonuna indirgenmiştir. Yaptığımızın en önemli yanı ise genel değişkenler ile temerrüt ve diğer çıkış değişkenlerinin nasıl birlikte tahmin edilebileceği konusunda yol göstermektir ki yapılan, bu konuda ilktir.