

Assessing bankruptcy prediction models via information content of technical inefficiency

Ruey-Ching Hwang · Jhao-Siang Siao ·
Huimin Chung · C. K. Chu

Published online: 13 March 2011
© Springer Science+Business Media, LLC 2011

Abstract We use a stochastic frontier model with firm-specific technical inefficiency effects in a panel framework (Battese and Coelli in *Empir Econ* 20:325–332, 1995) to assess two popular probability of bankruptcy (PB) measures based on Merton model (Merton in *J Financ* 29:449–470, 1974) and discrete-time hazard model (DHM; Shumway in *J Bus* 74:101–124, 2001). Three important results based on our empirical studies are obtained. First, a firm with a higher PB generally has less technical efficiency. Second, for an ex-post bankrupt firm, its PB tends to increase and its technical efficiency of production tends to decrease, as the time to its bankruptcy draws near. Finally, the information content about firm's technical inefficiency provided by PB based on DHM is significantly more than that based on Merton model. By the last result and the fact that economic-based efficiency measures are reasonable indicators of the long-term health and prospects of firms (Baek and Pagán in *Q J Bus Econ* 41:27–41, 2002), we conclude that PB based on DHM is a better credit risk proxy of firms.

Keywords Discrete-time hazard model · Merton model · Robust Wald test · Stochastic frontier model

R.-C. Hwang (✉)
Department of Finance, National Dong Hwa University,
Hualien 974, Taiwan
e-mail: rchwang@mail.ndhu.edu.tw

J.-S. Siao · C. K. Chu
Department of Applied Mathematics, National Dong Hwa
University, Hualien 974, Taiwan

H. Chung
Graduate Institute of Finance, National Chiao Tung University,
Hsinchu 300, Taiwan

JEL Classification D24 · G20

1 Introduction

Bankruptcy prediction has been routinely studied by academics, practitioners, and regulators. The well-known prediction models include discriminant analysis model (Altman 1968), Merton model (Merton 1974; Vassalou and Xing 2004), logit model (Ohlson 1980), and probit model (Zmijewski 1984), to name only a few. The common principal of these approaches is that the models are developed using single-period data of firms. Shumway (2001) pointed out that such prediction processes are static in nature, since they ignore the changing characteristics of firms through time. In order to avoid possible loss of prediction power due to using static models, Shumway (2001) and Chava and Jarrow (2004) proposed a discrete-time hazard model (DHM) using multiple-period data for bankruptcy prediction. Their novel model applies the idea of survival analysis (Cox and Oakes 1984), and has the advantage of using all available information of firms to build up a prediction system so that each firm's bankruptcy risk at each time point can be determined. Thus DHM is a dynamic forecasting model.¹

The performance of bankruptcy prediction models was mainly assessed in the literature by performing prediction-

¹ For other bankruptcy prediction approaches based on multiple-period data, see for example Duffie (2005) and Duffie et al. (2007) making use of different idea in point processes. Approaches based on neural networks (Atiya 2001), support vector machines (Härdle et al. 2008), Bayesian networks (Sun and Shenoy 2007), etc. were also introduced in the literature for bankruptcy prediction. Basically, these methods are more complicated in computation and interpretation than that based on DHM.

oriented tests; see for example the above references. Recently, Hillegeist et al. (2004) proposed a different approach to do it. They compared the information content about credit risk provided by out-of-sample values of probability of bankruptcy (PB) based on Merton model, Z-Score based on discriminant analysis model, and O-Score based on logit model. Their results show that PB based on Merton model provides significantly more information than Z-Score and O-Score. In contrast, Agarwal and Taffler (2008) pointed out that there is little difference between PB based on Merton model and Z-Score, in terms of predictive ability and information content. The results in Hillegeist et al. (2004) and Agarwal and Taffler (2008) were developed by comparing “static” bankruptcy prediction models.

In this paper, PB derived from the “static” Merton model (PB-Merton) is compared with that obtained from the “dynamic” DHM (PB-DHM) on their information content about firm’s technical inefficiency. To do it, a stochastic frontier model with firm-specific technical inefficiency effects in a panel framework (Battese and Coelli 1995) is considered. A firm is characterized as technically inefficient if it is not able to reach maximum output given its available resources and technology. By two reasons, analyzing the relationship between PB and economic efficiency is particularly important for firms. First, economic-based efficiency measures are reasonable indicators of the long-term health and prospects of firms (Baek and Pagán 2002). Second, by the finding in Becchetti and Sierra (2003) that ex-post failed firms are ex-ante significantly more technically inefficient, there is a linkage between PB and technical inefficiency.

Stochastic frontier analysis is a method of economic modeling. It has been widely used to estimate technical inefficiencies of production of firms; see for example, the monographs by Kumbhakar and Lovell (2003) and Coelli et al. (2005). It also has been used in studies of corporate financial decision and bank efficiency; see for example, Hunt-McCool et al. (1996), Baek (2004), Fries and Taci (2005), Lensink et al. (2008), and Kauko (2009).

To compare PB-Merton and PB-DHM on their ability in explaining technical inefficiency, they are “separately” used in the stochastic frontier model. In the model, the explanatory variables for technical inefficiency effects are taken as the out-of-sample PB² and control variable Age.³

² The purpose of this paper is to compare the two bankruptcy prediction models, DHM and Merton model, on their out-of-sample performance. Similar to Hillegeist et al. (2004), we employ relative information content tests to compare the out-of-sample performance of their PB measures.

³ The variable Age denotes a firm’s age which is defined as the number of calendar years it has been listed on the NYSE/AMEX/NASDAQ (Shumway 2001). Battese and Coelli (1995) use the variable Age to show the improvement of technical efficiency over

The out-of-sample values of PB-Merton and PB-DHM were computed for the 1996–2005 period. After deriving those out-of-sample values of PB-Merton and PB-DHM, the stochastic frontier model was estimated for the 1996–2005 period. The studied data were collected from both COMPUSTAT and CRSP databases. Our final sample for building the stochastic frontier model consists of 6,228 firms with 35,080 firm-year observations.

Our empirical results show that, after controlling for firm’s age, each of PB-Merton and PB-DHM is significantly positively correlated with technical inefficiency. The result indicates that a firm with a higher PB generally has less technical efficiency. Combining this result with the finding in Becchetti and Sierra (2003) that technical inefficiency has significant explanatory power in predicting bankruptcy, there is a causal relationship in either direction between credit risk and technical inefficiency. Our empirical results also suggest that, for an ex-post bankrupt firm, each of its PB-Merton and PB-DHM tends to increase and its technical efficiency of production tends to decrease, as the time to its bankruptcy draws near. This result coincides with our finding that PB is significantly positively correlated with technical inefficiency and the finding in Becchetti and Sierra (2003) that ex-post failed firms are ex-ante significantly more technically inefficient. Finally, our empirical results demonstrate that the information content about firm’s technical inefficiency provided by PB-DHM is significantly more than that generated by PB-Merton. The result indicates a potential increase of information content by using PB-DHM instead of PB-Merton as the credit risk proxy of firms.

The paper is organized as follows. Section 2 introduces the formulations of DHM, Merton model, and stochastic frontier model with firm-specific technical inefficiency effects. Section 3 describes the data under study. Section 4 presents our empirical results. Finally, concluding remarks are given in Sect. 5.

2 Methodology

In this section, we first describe the basic idea of DHM and the methodology for estimating PB-DHM. Next, we introduce an algorithm for computing PB-Merton. Finally, the formulation of the stochastic frontier model with firm-specific technical inefficiency effects is given.

Footnote 3 continued

time. The time dummy variable similarly shows the time effect on the improvement of technical efficiency, and thus it is not considered as the control variable in this paper. On the other hand, we do not suggest using the firm size as the control variable, since its related variables have been used in computing PB-DHM and PB-Merton.

2.1 DHM

The DHM can be formally defined from the log-likelihood function of the panel data. It has the advantage of using all available information to predict each firm’s bankruptcy risk at each point in time. In the following, we describe the structure of the panel data used in the prediction model based on DHM.

The panel data are determined by two factors: the sampling period and the sampling criterion. The sampling criterion in the paper is that all firms listed on the NYSE/AMEX/NASDAQ during the sampling period are recruited in the sample. All information occurred at the discrete time points during the sampling period are collected from both COMPUSTAT and CRSP databases. Suppose that there are n selected companies under the particular sampling scheme. We denote the panel data by

$$\{(Y_{i,j}, x_{i,j}) : j = 1, \dots, t_i, i = 1, \dots, n\}.$$

Here we denote $t_i \in \{1, \dots, \zeta\}$ to be the length of data collected from the i th firm during the sampling period, where ζ is a positive integer standing for the total length of the sampling period. At the last observation time t_i , $Y_{i,t_i} = 1$ indicates that the i th company is bankrupt, and $Y_{i,t_i} = 0$, otherwise. At the observation time $j < t_i$, we always have $Y_{i,j} = 0$. Finally, we let $x_{i,j}$ be values of a $d \times 1$ explanatory variable collected from the i th firm at time j .

Note that our panel data $\{(Y_{i,j}, x_{i,j}) : j = 1, \dots, t_i, i = 1, \dots, n\}$ are composed of left-truncated or right-censored data. By Klein and Moeschberger (2003), if a firm’s entry time into the sampling period and its bankrupt time are independent, then the hazard function of left-truncated data is equivalent to that of nontruncated data. By the result, the log-likelihood function of our panel data can be expressed as:

$$\ell_{\text{DHM}} = \sum_{i=1}^n Y_{i,t_i} \log \left\{ \frac{h(t_i, x_{i,t_i})}{1 - h(t_i, x_{i,t_i})} \right\} + \sum_{i=1}^n \sum_{j=1}^{t_i} \log \{1 - h(j, x_{i,j})\}. \tag{1}$$

It also has been given in (21) of Allison (1982). Here $h(j, x_{i,j})$ is the value of hazard function standing for the PB instantly happened at time j for the i th company which is non bankrupt before time j for each $j = 1, \dots, t_i$ and $i = 1, \dots, n$.

Note that the hazard function $h(t, x_t)$ in ℓ_{DHM} can be of any functional form with values in the interval (0,1). We consider a linear logistic function for $h(t, x_t)$ (Shumway 2001; Chava and Jarrow 2004; Campbell et al. 2008):

$$h(t, x_t) = \frac{\exp(\alpha + \beta x_t)}{1 + \exp(\alpha + \beta x_t)}, \tag{2}$$

where α and β are 1×1 and $1 \times d$ vectors of parameters, respectively. Given the linear logistic hazard function, the resulting log-likelihood of our panel data becomes:

$$\ell = \sum_{i=1}^n Y_{i,t_i} (\alpha + \beta x_{i,t_i}) - \sum_{i=1}^n \sum_{j=1}^{t_i} \log \{1 + \exp(\alpha + \beta x_{i,j})\}. \tag{3}$$

The maximum likelihood estimates of parameters α and β in (3) can be simply obtained by solving the normal equations:

$$0 = \sum_{i=1}^n Y_{i,t_i} \begin{bmatrix} 1 \\ x_{i,t_i} \end{bmatrix} - \sum_{i=1}^n \sum_{j=1}^{t_i} \frac{\exp(\alpha + \beta x_{i,j})}{1 + \exp(\alpha + \beta x_{i,j})} \begin{bmatrix} 1 \\ x_{i,j} \end{bmatrix}. \tag{4}$$

Using (2) and replacing the unknown parameters α and β in (2) with their maximum likelihood estimates $\hat{\alpha}$ and $\hat{\beta}$, if a firm has the predictor value x_t at time t , then its predicted instant PB-DHM can be given by:

$$\hat{h}(t, x_t) = \frac{\exp(\hat{\alpha} + \hat{\beta} x_t)}{1 + \exp(\hat{\alpha} + \hat{\beta} x_t)}. \tag{5}$$

Cox and Oakes (1984) showed that the maximum likelihood estimates $\hat{\alpha}$ and $\hat{\beta}$ are consistent for α and β , respectively. Thus the resulting predicted instant PB-DHM $\hat{h}(t, x_t)$ converges to the true instant PB-DHM $h(t, x_t)$. The result shows that DHM is a reliable and efficient prediction model.

2.2 Merton model

In this section, we describe an algorithm for computing PB-Merton. The PB-Merton defined by

$$N \left\{ \frac{\ln(V/B) + (\mu - \sigma_V^2/2)T}{\sigma_V \sqrt{T}} \right\} \tag{6}$$

can be computed using an iterative procedure suggested by Vassalou and Xing (2004) and Bharath and Shumway (2008). Here $N(\cdot)$ is the standard normal distribution function, V the market value of the firm’s assets, B the face value of the firm’s debt, μ the expected continuously compounded return on V , σ_V the volatility of firm’s value, and T the maturing period of the firm’s debt. The formulation of PB-Merton in (6) has been given in (7) of Bharath and Shumway (2008).

The iterative procedure for computing PB-Merton includes the following steps. First, take the initial value of σ_V as $\sigma_E E / (E + B)$ and apply the value of σ_V to the Black–Scholes-Merton option valuation equation

$$\begin{cases} E = VN(d_1) - e^{-rT}BN(d_2), \\ d_1 = \frac{\ln(V/B) + (r + \sigma_V^2/2)T}{\sigma_V\sqrt{T}}, \quad d_2 = d_1 - \sigma_V\sqrt{T}, \end{cases} \quad (7)$$

to derive the value of V at the end of every month for the previous year. Here E is calculated from the CRSP database as the product of share price at the end of the month and the number of shares outstanding, B the sum of the short-term debt (debt in current liabilities) and one half of long term debt, σ_E the annualized percent standard deviation of returns and estimated from the prior year stock return data for each month, r the risk-free interest rate, and $T = 1$. In our empirical study, we took r as one-year treasury constant maturity rate provided by the Board of Governors of the Federal Reserve system available at the website <http://research.stlouisfed.org/fred/data/irates/gsl>. Thus, the implied log return on assets of each month can be computed, and the selected returns are used to generate new estimates of σ_V and μ . Iterate on σ_V in this manner until it converges (so the absolute difference in adjacent σ_V is less than 10^{-3}).

2.3 Stochastic frontier model

To examine the relationship between PB and technical inefficiency for firms, the stochastic frontier model proposed by Battese and Coelli (1995) for panel data is considered. Using a translog specification, the model is given by:

$$\ln(Q_{i,t}) = (\gamma_0 + \gamma_1 w_{i,t}) + \sum_{j=1}^m (\eta_{0,j} + \eta_{1,j} w_{i,t}) \text{SIC}_j + V_{i,t} - U_{i,t}, \quad (8)$$

for $t = 1, \dots, t_i$ and $i = 1, \dots, n$. Here $\ln(Q_{i,t})$ is the natural logarithm of the dollar value of production $Q_{i,t}$ for the i th firm at time t , and $Q_{i,t}$ is estimated as the sum of sales and inventory increase during the period from the time point $t - 1$ to t ; $w_{i,t} = \{\ln(L_{i,t}), \ln(K_{i,t}), \ln(L_{i,t})^2, \ln(K_{i,t})^2, \ln(L_{i,t}) \ln(K_{i,t})\}'$ is a 5×1 vector of explanatory variables related to the stochastic frontier production function for the i th firm at time t ; $\ln(L_{i,t})$ is the natural logarithm of the total number of employees for the i th firm at time t ; $\ln(K_{i,t})$ is the natural logarithm of the value of fixed assets for the i th firm at time t ; SIC_j is an industry dummy variable, for each $j = 1, \dots, m$, which is designed for capturing

differences in the production technology across $m + 1$ industry categories⁴; $V_{i,t}$ s are independent and identically distributed $N(0, \sigma_v^2)$ random errors and independently distributed of $U_{i,t}$ s;

$U_{i,t}$ s are nonnegative random variables, associated with technical inefficiency effects, which are assumed to be independently distributed, such that $U_{i,t}$ is obtained by truncation (at zero) of the normal distribution with mean $\delta_0 + \delta_1 z_{i,t}$ and variance σ_u^2 ;

$z_{i,t}$ is a $q \times 1$ vector of explanatory variables related to technical inefficiency effect $U_{i,t}$ for the i th firm at time t ;

$\gamma_0, \gamma_1, \eta_{0,j}, \eta_{1,j}, \delta_0$, and δ_1 are unknown $1 \times 1, 1 \times 5, 1 \times 1, 1 \times 5, 1 \times 1$, and $1 \times q$ vectors of parameters to be estimated, respectively.

The unknown parameter vector $\theta = (\gamma_0, \gamma_1, \eta_{0,1}, \dots, \eta_{0,m}, \eta_{1,1}, \dots, \eta_{1,m}, \delta_0, \delta_1)$ and variance parameters $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\zeta = \sigma_u^2 / \sigma^2$ in (8) can be estimated by maximizing the log-likelihood function of the panel data (Battese and Coelli 1993). Their maximum likelihood estimates $\hat{\theta} = (\hat{\gamma}_0, \hat{\gamma}_1, \hat{\eta}_{0,1}, \dots, \hat{\eta}_{0,m}, \hat{\eta}_{1,1}, \dots, \hat{\eta}_{1,m}, \hat{\delta}_0, \hat{\delta}_1)$, $\hat{\sigma}^2$, and $\hat{\zeta}$ can be simply computed using the free software FRONTIER Version 4.1 written by Professor Tim Coelli and available at the website <http://www.uq.edu.au/economics/cepa/frontier.htm>. On the other hand, to allow for non-independence of technical inefficiency over time, the Wald test using “robust” standard error of estimated parameter (Alvarez et al. 2006) was employed to test the significance of parameter in the empirical study.

To study possible sources of technical inefficiency effect $U_{i,t}, z_{i,t}$ used in Sect. 3 is a 2×1 vector, and consists of the out-of-sample PB-Merton (or PB-DHM) and control variable Age. The technical efficiency of production for the i th firm at time t is defined by $TE_{i,t} = \exp(-U_{i,t})$, which can be estimated by

$$\hat{E}\{\exp(-U_{i,t}) | \varepsilon_{i,t} = \hat{\varepsilon}_{i,t}\} = \exp(-\mu_{i,t} + 0.5\phi^2) \times \frac{\Phi\{(\mu_{i,t}/\phi) - \phi\}}{\Phi(\mu_{i,t}/\phi)}, \quad (9)$$

where

$$\begin{aligned} \varepsilon_{i,t} &= V_{i,t} - U_{i,t}, \hat{\varepsilon}_{i,t} \\ &= \ln(Q_{i,t}) - \left\{ \hat{\gamma}_0 + \hat{\gamma}_1 w_{i,t} + \sum_{j=1}^m (\hat{\eta}_{0,j} + \hat{\eta}_{1,j} w_{i,t}) \text{SIC}_j \right\}, \end{aligned}$$

$$\mu_{i,t} = (1 - \hat{\zeta})(\hat{\delta}_0 + \hat{\delta}_1 z_{i,t}) - \hat{\zeta} \hat{\varepsilon}_{i,t}, \phi = \left\{ \hat{\zeta}(1 - \hat{\zeta}) \right\}^{1/2} \hat{\sigma},$$

and $\Phi(\cdot)$ represents the standard normal distribution function.

⁴ Nine industry categories are defined in Table 3 of Sect. 3 for our empirical studies.

3 Data and variable estimation

Bankruptcy filings between 1984 and 2005 were identified from the CRSP database according to the delisting codes 400–490, 572, and 574. The sampled firms consist of all industrial firms listed on the NYSE/AMEX/NASDAQ, and have complete predictor values for calculating the PB-DHM collected from both CRSP and COMPUSTAT databases during the period from 1983 to 2004. The financial services firms with SIC codes 6000–6999 were eliminated from the sample due to the unique capital requirements and regulatory structure in that industry group. Our final sample for computing out-of-sample PB-DHM consists of 9,150 solvent firms, 245 bankrupt firms, and 69,285 firm-year observations.

The predictors considered in DHM for computing out-of-sample PB-DHM were those accounting and market-driven variables suggested by Shumway (2001) and Campbell et al. (2008) including:

NIMTA	Net income divided by market-valued total assets, ⁵
TLMTA	Total liabilities divided by market-valued total assets,
EXRETAVG	Average of 12 monthly log excess returns ⁶ with geometrically declining weights, ⁷
RSIZE	Logarithm of each firm’s market equity value divided by the total NYSE/AMEX/NASDAQ market equity value,
SIGMA	Annualized square root of the average of squared deviations in firm’s daily stock returns from 0 over the past 3 months,
CASHMTA	Cash and short-term investments divided by market-valued total assets,
MB	Market equity value divided by book equity value, and
PRICE	Logarithm of stock price if the price is below \$15, and logarithm of \$15 otherwise.

The values of each of these explanatory variables were winsorized using a 5/95 percentile interval in order to eliminate outliers except SIGMA and PRICE (Campbell et al. 2008). The summary statistics of the predictor values are given in Table 1.

⁵ The market-valued total assets is the sum of the book value of liabilities and the market value of equities.

⁶ Our log excess return (EXRET) is based on the firm’s equity relative to the value-weighted NYSE/AMEX/NASDAQ index.

⁷ For year t , $EXRETAVG_t = \frac{1-\omega}{1-\omega^{12}}(EXRET_{t,12} + \dots + \omega^{11}EXRET_{t,1})$, where $\omega = 2^{-1/3}$.

Table 1 Summary statistics of the panel dataset used to compute out-of-sample PB-DHM

Predictor	Mean	Median	SD	Minimum	Maximum
NIMTA	−0.014	0.021	0.098	−0.305	0.091
TLMTA	0.371	0.330	0.259	0.030	0.882
EXRETAVG	−0.005	−0.004	0.053	−0.112	0.098
SIGMA	0.803	0.623	0.710	0.025	27.437
RSIZE	−11.054	−11.135	1.911	−14.281	−7.397
CASHMTA	0.097	0.050	0.114	0.002	0.416
MB	2.776	1.829	2.763	0.035	10.951
PRICE	1.891	2.209	0.954	−4.159	2.708

The predictors considered in DHM were those accounting and market-driven variables suggested by Shumway (2001) and Campbell et al. (2008). Their definitions have been given in Sect. 3. The sampled firms consist of all industrial firms listed on the NYSE/AMEX/NASDAQ, and have complete predictor values for calculating the PB-DHM collected from both CRSP and COMPUSTAT databases during the period from 1983 to 2004. Each of these explanatory variables was winsorized using a 5/95 percentile interval in order to eliminate outliers except SIGMA and PRICE (Campbell et al. 2008). Our final sample consists of 9,150 solvent firms, 245 bankrupt firms, and 69,285 firm-year observations

Based on the above predictors, we calculated out-of-sample PB-DHM using an expanding rolling window approach (Hillegeist et al. 2004; Chava et al. 2008). For the first window, we estimated updated coefficients of DHM using the firm-year data from 1983 to 1994 and bankruptcy outcomes from 1984 to 1995. These updated coefficients were combined with predictor values in 1995 to evaluate out-of-sample PB-DHM in 1996. For the second window, we used the firm-year data from 1983 to 1995 and bankruptcy outcomes from 1984 to 1996 to estimate the second set of updated coefficients of DHM. This second set of coefficients was combined with predictor values in 1996 to produce out-of-sample PB-DHM in 1997. The process was continued so that the last set of updated coefficients of DHM used to generate out-of-sample PB-DHM in 2005 was based on the firm-year data from 1983 to 2003 and bankruptcy outcomes from 1984 to 2004. Hence, we collected out-of-sample PB-DHM from 1996 to 2005. The maximum likelihood estimates of parameters in DHM for computing out-of-sample PB-DHM in each year between 1996 and 2005 are given in Table 2. Table 2 shows that, in each window, the predictors NIMTA, TLMTA, and PRICE are all significant at 1% level and the values of their estimated coefficients all agree with their expected signs. On the other hand, we derived PB-Merton as discussed in Sect. 2.2 from 1995 to 2004, and the result was taken as out-of-sample PB-Merton from 1996 to 2005. Here the values of out-of-sample PB-DHM and PB-Merton were produced using the Eqs. (5) and (6), respectively, and they were further used in the Eq. (9) to produce the values of technical efficiency of production for firms.

Table 2 Maximum likelihood estimates of parameters in DHM for computing out-of-sample PB-DHM

Predictor	Sample period									
	1983–1994	1983–1995	1983–1996	1983–1997	1983–1998	1983–1999	1983–2000	1983–2001	1983–2002	1983–2003
Intercept	-8.426**	-7.945**	-7.490**	-7.736**	-8.046**	-7.872**	-8.269**	-8.096**	-8.249**	-8.259**
NIMTA	-4.832**	-4.897**	-4.554**	-4.978**	-4.616**	-4.365**	-4.248**	-4.627**	-4.399**	-4.530**
TLMTA	3.349**	3.397**	3.108**	3.505**	3.588**	3.622**	3.793**	3.936**	3.962**	4.001**
EXRETAVG	-0.174	0.269	0.346	-0.274	0.130	-1.038	-2.103	-0.151	-0.615	-1.195
SIGMA	0.007	0.011	0.034	-0.009	0.002	0.003	-0.010	-0.013	0.055	0.042
RSIZE	-0.076	-0.032	-0.016	-0.024	-0.047	-0.041	-0.064	-0.046	-0.047	-0.046
CASHMTA	0.706	0.796	0.744	0.239	-0.063	-0.144	-0.377	-0.073	-0.437	-0.475
MB	0.027	0.026	0.036	0.028	0.034	0.015	0.038	0.039	0.036	0.038
PRICE	-0.645**	-0.701**	-0.692**	-0.636**	-0.576**	-0.518**	-0.448**	-0.453**	-0.412**	-0.406**
Observations	23838	28383	33338	38427	43331	48158	52900	57252	61371	65311
Bankruptcies	73	77	91	106	119	144	180	207	225	237
Log-likelihood	-365.75	-391.25	-475.46	-551.53	-636.11	-775.90	-955.65	-1087.86	-1174.92	-1237.54

Using the panel dataset in Table 1, the out-of-sample PB-DHM in each year between 1996 and 2005 was computed using an expanding rolling window approach. For the first window, the coefficients of DHM were estimated using the firm-year data from 1983 to 1994 and bankruptcy outcomes from 1984 to 1995. The estimated coefficients were combined with predictor values in 1995 to evaluate out-of-sample PB-DHM in 1996. For the second window, the firm-year data from 1983 to 1995 and bankruptcy outcomes from 1984 to 1996 were used to estimate the second set of coefficients of DHM. This second set of estimated coefficients was combined with predictor values in 1996 to produce out-of-sample PB-DHM in 1997. The process was continued so that the last set of estimated coefficients of DHM was based on the firm-year data from 1983 to 2003 and bankruptcy outcomes from 1984 to 2004. This last set of estimated coefficients was combined with predictor values in 2004 to produce out-of-sample PB-DHM in 2005. The notation ** indicates the significance of test at 1% level

To study the relationship between PB and technical inefficiency for firms, a panel dataset for building the stochastic frontier production function was collected from the COMPUSTAT database during the period from 1996 to 2005. We used the firm's age as the control variable related to technical inefficiency. The firm's age is defined as the number of calendar years that the firm has been listed on the NYSE/AMEX/NASDAQ (Shumway 2001), and denoted as Age. The out-of-sample PB-Merton (or PB-DHM)

and Age are explanatory variables related to technical inefficiency effects in the stochastic frontier model. After deleting observations with missing values, our final sample for estimating the stochastic frontier model consists of 6,228 firms with 35,080 firm-year observations.

Table 3 shows the definitions of industry dummy variables and the frequency distributions of industry categories corresponding to those industry dummy variables. The industry dummy variables are included for studying

Table 3 The definitions of industry dummy variables and the frequency distributions of the sampled companies

Industry dummy variable	SIC code	Industry	Frequency
SIC ₁	100–999	Agriculture, forestry, and fishing	25
SIC ₂	1000–1999	Mining, construction	335
SIC ₃	2000–2999	Food, tobacco, textile, and paper manufacturing etc.	1,057
SIC ₄	3000–3999	Rubber, leather, stone, and industrial manufacturing etc.	1,928
SIC ₅	5000–5999	Wholesale trade, retail trade	725
SIC ₆	7000–7999	Hotel, personal, business, and automotive repair services etc.	1,156
SIC ₇	8000–8999	Health, legal, education, and social services etc.	317
SIC ₈	9000–9999	Public administration	32
Reference level	4000–4999	Transportation, communications, electric, gas, and sanitary services	653
Total firms			6,228

The eight industry dummy variables were defined according to the first digit in the four-digit SIC codes provided by the COMPUSTAT database, and used in the stochastic frontier model for capturing differences in the production technology across nine industry categories. The SIC code 4000–4999 was used as the reference level for studying the industry effects in stochastic frontier model. The financial services firms with SIC codes 6000–6999 were excluded from the study, since they are subject to regulations and adopt different sing conventions. The frequency distributions are based on SIC codes of the sampled companies in the panel dataset for estimating the stochastic frontier model

Table 4 Summary statistics of the panel dataset used to estimate the stochastic frontier model

Variable	Mean	Median	SD	Minimum	Maximum
ln(Q)	5.323	5.376	2.067	0.824	8.799
ln(L)	0.120	0.105	1.956	-3.612	3.505
ln(K)	3.804	3.750	2.408	-0.749	8.033
PB-Merton	0.061	0.000	0.190	0.000	1.000
PB-DHM	0.002	0.001	0.006	0.000	0.158
Age	15.123	9.000	15.015	2.000	81.000

The values of variables ln(Q), ln(L), and ln(K) were collected from the COMPUSTAT database during the period from 1996 to 2005. Here ln(Q) is the natural logarithm of the dollar value of production and Q is estimated as the sum of sales and inventory increase for a year, ln(L) is the natural logarithm of the total number of employees, and ln(K) is the natural logarithm of the value of fixed assets. The values of PB-Merton were derived from 1995 to 2004, and used as out-of-sample PB-Merton from 1996 to 2005. The values of PB-DHM were out-of-sample PB based on DHM, and produced for the period from 1996 to 2005 using an expanding rolling window approach. The variable Age denotes a firm’s age which is defined as the number of calendar years it has been listed on the NYSE/AMEX/NASDAQ. Our final sample consists of 6,228 firms and 35,080 firm-year observations

industry effects by capturing differences in the production technology across nine industry categories. Table 4 reports summary statistics for the predictor values employed in estimating the stochastic frontier model. It shows that the sample average of PB-DHM is much smaller than that of PB-Merton. This result is due to the fact that the former was produced by a reduced-form model designed to calibrate real default experience. Similar result for the magnitude of PB based on the reduced-form model can also be seen in Table 3 of Hillegeist et al. (2004).

4 Empirical results

The empirical results based on the stochastic frontier model with different specifications of the inefficiency effect equation are given in Table 5 and Fig. 1. In Table 5 and Fig. 1, Model 1 uses PB-Merton and Age as explanatory variables, and Model 2 employs PB-DHM and Age as explanatory variables in the inefficiency effect equation.

Panel A of Table 5 shows that most of parameters of the production function under each of Models 1 and 2 are statistically significant at 5% level. The result of statistical significance for most of industry dummy variables indicates that the production technology indeed varies across industries. The estimated coefficients of the production function under each of Models 1 and 2 were used to calculate marginal products of the two inputs, labor and fixed assets, for each sample point. The average values of the two marginal products under Model 1 are 0.650 and 0.474,

Table 5 Maximum likelihood estimates of parameters in the stochastic frontier model with different specifications of the inefficiency effect equation

Variable	Model 1		Model 2	
	Coefficient	Robust SE	Coefficient	Robust SE
Panel A: Production function				
Intercept	6.005**	0.126	6.055**	0.134
ln(L)	1.374**	0.049	1.388**	0.052
ln(K)	-0.487**	0.051	-0.513**	0.055
ln(L) ²	0.076**	0.008	0.076**	0.008
ln(K) ²	0.086**	0.005	0.089**	0.006
ln(L) ln(K)	-0.152**	0.008	-0.155**	0.008
SIC ₁	-0.043	0.662	-0.586	0.824
SIC ₂	2.008**	0.190	1.953**	0.202
SIC ₃	0.109	0.176	0.056	0.183
SIC ₄	-1.685**	0.191	-1.724**	0.200
SIC ₅	-0.795**	0.169	-0.830**	0.177
SIC ₆	-1.227**	0.147	-1.290**	0.155
SIC ₇	-1.508**	0.214	-1.541**	0.219
SIC ₈	-3.624**	1.153	-3.662**	1.129
ln(L)SIC ₁	0.037	0.378	-0.179	0.457
ln(L)SIC ₂	0.636**	0.082	0.618**	0.085
ln(L)SIC ₃	-0.079	0.080	-0.111	0.081
ln(L)SIC ₄	-1.132**	0.107	-1.135**	0.111
ln(L)SIC ₅	-0.846**	0.071	-0.859**	0.073
ln(L)SIC ₆	-0.764**	0.062	-0.776**	0.064
ln(L)SIC ₇	-0.839**	0.109	-0.855**	0.109
ln(L)SIC ₈	-1.713*	0.782	-1.705*	0.769
ln(K)SIC ₁	-0.181	0.399	0.097	0.451
ln(K)SIC ₂	-0.641**	0.081	-0.608**	0.086
ln(K)SIC ₃	-0.075	0.075	-0.039	0.077
ln(K)SIC ₄	0.980**	0.098	1.003**	0.102
ln(K)SIC ₅	0.596**	0.079	0.623**	0.082
ln(K)SIC ₆	0.732**	0.068	0.760**	0.071
ln(K)SIC ₇	0.809**	0.124	0.827**	0.125
ln(K)SIC ₈	1.730**	0.522	1.738**	0.506
ln(L) ² SIC ₁	-0.037	0.061	-0.047	0.071
ln(L) ² SIC ₂	0.002	0.013	0.002	0.014
ln(L) ² SIC ₃	-0.076**	0.013	-0.080**	0.013
ln(L) ² SIC ₄	-0.193**	0.019	-0.191**	0.019
ln(L) ² SIC ₅	-0.073**	0.013	-0.071**	0.013
ln(L) ² SIC ₆	-0.123**	0.011	-0.122**	0.011
ln(L) ² SIC ₇	-0.116**	0.020	-0.117**	0.020
ln(L) ² SIC ₈	-0.194	0.136	-0.188	0.134
ln(K) ² SIC ₁	0.032	0.060	-0.002	0.064
ln(K) ² SIC ₂	0.050**	0.008	0.046**	0.009
ln(K) ² SIC ₃	0.022*	0.009	0.017*	0.009
ln(K) ² SIC ₄	-0.120**	0.013	-0.123**	0.014

Table 5 continued

Variable	Model 1		Model 2	
	Coefficient	Robust SE	Coefficient	Robust SE
$\ln(K)^2SIC_5$	-0.048**	0.010	-0.051**	0.010
$\ln(K)^2SIC_6$	-0.089**	0.008	-0.092**	0.008
$\ln(K)^2SIC_7$	-0.108**	0.018	-0.111**	0.018
$\ln(K)^2SIC_8$	-0.211*	0.087	-0.207*	0.087
$\ln(L)\ln(K)SIC_1$	0.003	0.086	0.051	0.101
$\ln(L)\ln(K)SIC_2$	-0.095**	0.014	-0.090**	0.015
$\ln(L)\ln(K)SIC_3$	0.011	0.014	0.020	0.014
$\ln(L)\ln(K)SIC_4$	0.281**	0.027	0.283**	0.028
$\ln(L)\ln(K)SIC_5$	0.106**	0.012	0.109**	0.013
$\ln(L)\ln(K)SIC_6$	0.169**	0.012	0.173**	0.012
$\ln(L)\ln(K)SIC_7$	0.204**	0.028	0.208**	0.028
$\ln(L)\ln(K)SIC_8$	0.415*	0.175	0.404*	0.169
Panel B: Inefficiency effect equation				
Intercept	-0.585**	0.135	-1.774**	0.193
PB-Merton	0.265*	0.122		
PB-DHM			16.946**	5.134
Age	-0.090**	0.013	-0.102**	0.016
Panel C: Variance parameter				
σ^2	1.224**	0.034	1.793**	0.048
ζ	0.745**	0.001	0.829**	0.001
Panel D: Model fit test				
Log-likelihood	-35424.943		-35388.310	
Likelihood-ratio	1313.928**		1387.194**	
Panel E: Vuong test				
Z-statistic	3.590**			

Model 1 uses PB-Merton and Age as explanatory variables in the inefficiency effect equation. Model 2 employs PB-DHM and Age as explanatory variables in the inefficiency effect equation. Panels A and B show respectively the estimation result of production function and that of inefficiency effect equation in Models 1 and 2. Panel C presents the estimation result of variance parameters in Models 1 and 2. To allow for non-independence of technical inefficiency over time, the Wald test using robust standard error of estimated parameter (Alvarez et al. 2006) is employed in each of Panels A, B, and C to test the significance of parameter. Panel D contains the result of model fit test for the inefficiency frontier model under Models 1 and 2. Panel E gives the Z-statistic of Vuong test. The notations ** and * indicate the significance of test at 1% and 5% levels, respectively

and those under Model 2 are 0.653 and 0.484 for the inputs labor and fixed assets, respectively in each case. By the positive signs of the two marginal products, both Models 1 and 2 appear to attain the implications in economics. Hence, the estimated coefficients of the production function under each of Models 1 and 2 have appropriate signs.

Panel B of Table 5 shows that the estimated coefficient of control variable Age under each of Models 1 and 2 is

negative and statistically significant at 1% level, which indicates that the old firms are more technically efficient than the young ones. After controlling for firm's age, each of PB-Merton in Model 1 and PB-DHM in Model 2 is positively correlated with technical inefficiency since the sign of its coefficient is positive at 5% level of significance. The result indicates that if a firm has a higher PB, then it is less technically efficient. Combining this result with the finding in Becchetti and Sierra (2003) that technical inefficiency has significant explanatory power in predicting bankruptcy, there is a causal relationship in either direction between credit risk and technical inefficiency. It is possible that a firm with high credit risk may not be able to get financing, and thus can not do things to increase its technical efficiency. Conversely, technical inefficiency may cause poor profitability which may cause high bankruptcy risk, and thus economic-based efficiency measures are reasonable indicators of the long-term health and prospects of firms.

Panel C of Table 5 presents that the estimated ratios of variance parameter ζ are 0.745 and 0.829 under Models 1 and 2, respectively. Both of them are statistically significant at 1% level. The result indicates that the inefficiency effect is likely to be important because most of the composite error variance σ^2 is accounted for by the variance of the inefficiency effect in each model. Further, by performing a likelihood-ratio test, the null hypothesis that the values of parameters in inefficiency effect equation and ζ are all equal to zero is rejected at 1% level of significance under each of Models 1 and 2. The test result is given in Panel D of Table 5, and indicates that the inefficiency frontier model is appropriate under each model.

Table 5 also shows that PB-DHM provides more information content about firm's technical inefficiency than PB-Merton. This superiority can be directly observed from Panel D that Model 2 has a larger log-likelihood statistic than Model 1. To test the null hypothesis that the information content about firm's technical inefficiency provided by Model 2 is not more than that generated by Model 1, a non-nested hypothesis test procedure in Vuong (1989) based on the log-likelihood statistic was performed. Since PB-Merton and PB-DHM were separately used in the inefficiency effect equation, their corresponding Models 1 and 2 are non-nested models. Panel E shows that the null hypothesis of interest was rejected at 1% level of significance. The test result indicates that the information content about firm's technical inefficiency provided by PB-DHM is more than that generated by PB-Merton at 1% level of significance.

Overall, the results in Table 5 show that PB-DHM has the advantage of having better ability in explaining firm's technical inefficiency over PB-Merton. Combining the

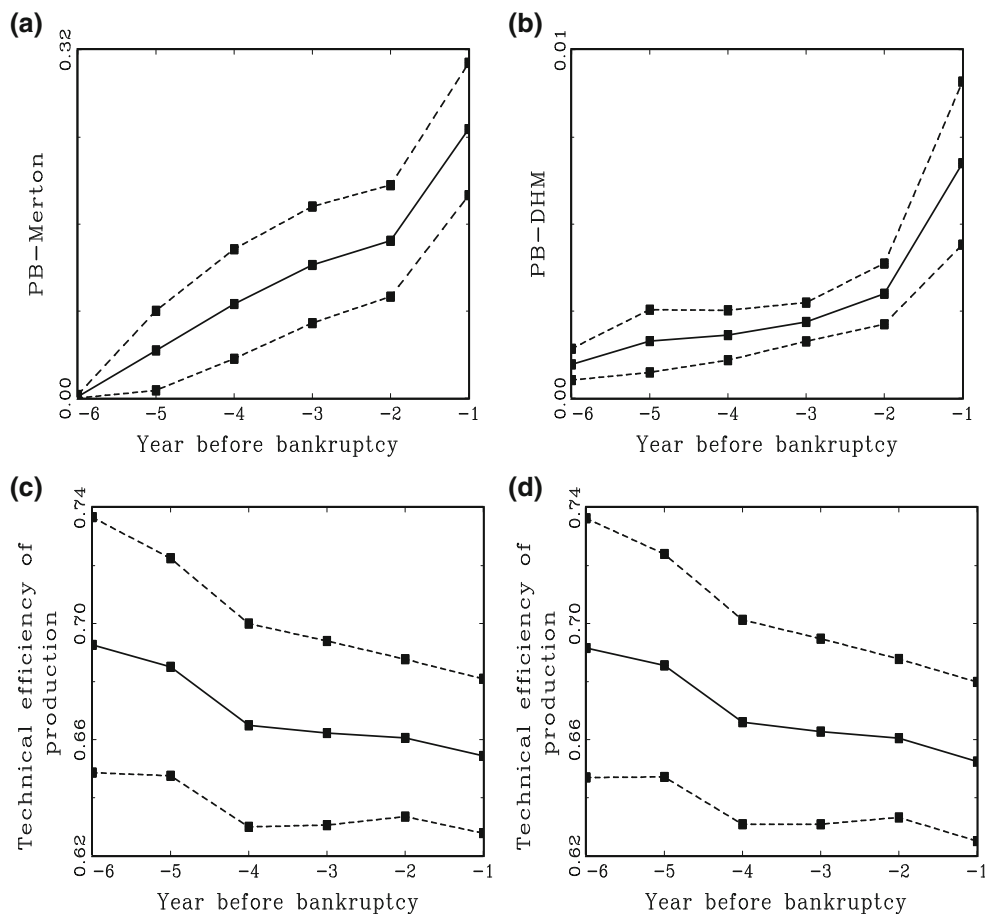


Fig. 1 Plot of PB and technical efficiency of production of ex-post bankrupt firms. Among the 6,228 firms for analyzing the stochastic frontier model during the period from 1996 to 2005, there were 133 firms bankrupt within 1 year after their last observations in the panel dataset were made. Among the 133 ex-post bankrupt firms, **a** shows the sample average (*solid curve*) as well as the corresponding 95%

confidence interval (*dashed curves*) of estimates of PB-Merton, and **c** presents those of estimates of technical efficiency of production under Model 1, in each year during the last 6 years preceding their bankruptcy. The captions of **b** and **d** are the same as those of **a** and **c** with PB-Merton and Model 1 replaced by PB-DHM and Model 2, respectively in each case

result with the fact that economic-based efficiency measures are reasonable indicators of the long-term health and prospects of firms, we conclude that PB-DHM is a better credit risk proxy than PB-Merton.

On the other hand, among the 6,228 firms for analyzing the stochastic frontier model during the period from 1996 to 2005, there were 133 firms bankrupt within 1 year after their last observations in the panel dataset were made. Among the 133 ex-post bankrupt firms, Fig. 1 shows clearly that the sample averages of PB-Merton and PB-DHM increase and those of technical efficiency of production under Models 1 and 2 decrease, as the time to their bankruptcy draws near. The relationship between PB and technical efficiency of production shown in Fig. 1 for ex-post bankrupt firms coincides with our result obtained from Panel B of Table 5 that PB is significantly positively correlated with technical inefficiency. The decrease of the sample average of technical efficiency of production shown

in Fig. 1 for ex-post bankrupt firms coincides with the finding in Becchetti and Sierra (2003) that ex-post failed firms are ex-ante significantly more technically inefficient.

5 Concluding remarks

The performance of bankruptcy prediction models was mainly assessed in the literature by performing prediction-oriented tests. Recently, Hillegeist et al. (2004) proposed a different approach to do it. They compared the information content about credit risk provided by out-of-sample values of PB-Merton, Z-Score, and O-Score. Their results show that PB-Merton provides significantly more information than Z-Score and O-Score. In contrast, Agarwal and Taffler (2008) pointed out that there is little difference between PB-Merton and Z-Score, in terms of predictive ability and information content.

In this paper, PB-Merton is compared with PB-DHM on their information content about firm's technical inefficiency. To do it, their out-of-sample values were computed for the 1996–2005 period. After deriving those out-of-sample values of PB-Merton and PB-DHM, they were separately used in the stochastic frontier model with firm-specific technical inefficiency effects. The model was estimated for the 1996–2005 period. The studied data were collected from both COMPUSTAT and CRSP databases. Our final sample for building the stochastic frontier model consists of 6,228 firms with 35,080 firm-year observations.

Three important results based on our empirical studies are obtained. First, after controlling for firm's age, each of PB-Merton and PB-DHM is significantly positively correlated with technical inefficiency. The result indicates that if a firm has a higher PB, then it is less technically efficient. By this result and the finding in Becchetti and Sierra (2003) that technical inefficiency has significant explanatory power in predicting bankruptcy, there is a causal relationship in either direction between credit risk and technical inefficiency. Second, for an ex-post bankrupt firm, each of its PB-Merton and PB-DHM tends to increase, and its technical efficiency of production generated from the corresponding stochastic frontier model tends to decrease, as the time to its bankruptcy draws near. This result coincides with our finding that PB is significantly positively correlated with technical inefficiency and the finding in Becchetti and Sierra (2003) that ex-post failed firms are ex-ante significantly more technical inefficiency. Finally, PB-DHM provides significantly more information content about firm's technical inefficiency than PB-Merton. By the result and the fact that economic-based efficiency measures are reasonable indicators of the long-term health and prospects of firms, we conclude that PB-DHM is a better credit risk proxy than PB-Merton.

There is one possible extension of the methods considered in this paper. Credit ratings are based on both public information and private information conveyed to the rating agencies by firms. It is of interest to compare PB-DHM with credit ratings on their information content about firm's technical inefficiency.

Acknowledgments The authors thank the reviewers and the editor for their valuable comments and suggestions which greatly improve the presentation of this paper. The research was supported by National Science Council, Taiwan, Republic of China.

References

- Agarwal V, Taffler R (2008) Comparing the performance of market-based and accounting-based bankruptcy prediction models. *J Bank Financ* 32:1541–1551
- Allison PD (1982) Discrete-time methods for the analysis of event histories. *Sociol Methodol* 13:61–98
- Altman EI (1968) Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *J Financ* 23:589–609
- Alvarez A, Amsler C, Orea L, Schmidt P (2006) Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics. *J Prod Anal* 25:201–212
- Atiya AF (2001) Bankruptcy prediction for credit risk using neural networks: a survey and new results. *IEEE Trans Neural Netw* 12:929–935
- Baek HY (2004) Corporate diversification and performance: evidence on production efficiency. *J Multinatl Financ Manag* 14:135–152
- Baek H, Pagán J (2002) Executive compensation and corporate production efficiency: a stochastic frontier approach. *Q J Bus Econ* 41:27–41
- Battese GE, Coelli TJ (1993) A stochastic frontier production function incorporating a model for technical inefficiency effects. Working paper, University of New England
- Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empir Econ* 20:325–332
- Becchetti L, Sierra J (2003) Bankruptcy risk and productive efficiency in manufacturing firms. *J Bank Financ* 27:2099–2120
- Bharath ST, Shumway T (2008) Forecasting default with the Merton distance to default model. *Rev Financ Stud* 21:1339–1369
- Campbell JY, Hilscher J, Szilagyi J (2008) In search of distress risk. *J Financ* 63:2899–2939
- Chava S, Jarrow RA (2004) Bankruptcy prediction with industry effects. *Rev Financ* 8:537–569
- Chava S, Stefanescu C, Turnbull S (2008) Modeling the loss distribution. Available at <http://faculty.london.edu/cstefanescu/research.html>
- Coelli TJ, Rao DSP, O'Donnell CJ, Battese GE (2005) An introduction to efficiency and productivity analysis. Springer, New York
- Cox DR, Oakes D (1984) Analysis of survival data. Chapman and Hall, London
- Duffie D (2005) Credit risk modeling with affine processes. *J Bank Financ* 29:2751–2802
- Duffie D, Saita L, Wang K (2007) Multi-period corporate default prediction with stochastic covariates. *J Financ Econ* 83:635–665
- Fries S, Taci A (2005) Cost efficiency of banks in transition: evidence from 289 banks in 15 post communist countries. *J Bank Financ* 29:55–81
- Härdle W, Moro RA, Schäfer D (2008) Graphical data representation in bankruptcy analysis. In: Chen C, Härdle W, Unwin A (eds) Handbook of data visualization. Springer, Berlin, pp 853–873
- Hillegeist SA, Keating EK, Cram DP, Lundstedt KG (2004) Assessing the probability of bankruptcy. *Rev Acc Stud* 9:5–34
- Hunt-McCool J, Koh SC, Francis BB (1996) Testing for deliberate underpricing in the IPO premarket: a stochastic frontier approach. *Rev Financ Stud* 9:1251–1269
- Kauko K (2009) Managers and efficiency in banking. *J Bank Financ* 33:546–556
- Klein JP, Moeschberger ML (2003) Survival analysis: techniques for censored and truncated data. Springer, New York
- Kumbhakar SC, Lovell CAK (2003) Stochastic frontier analysis. Cambridge University Press, Cambridge
- Lensink R, Meesters A, Naaborg I (2008) Bank efficiency and foreign ownership: do good institutions matter? *J Bank Financ* 32:834–844
- Merton RC (1974) On the pricing of corporate debt: the risk structure of interest rates. *J Financ* 29:449–470
- Ohlson J (1980) Financial ratios and the probabilistic prediction of bankruptcy. *J Acc Res* 18:109–131
- Shumway T (2001) Forecasting bankruptcy more accurately: a simple hazard model. *J Bus* 74:101–124
- Sun L, Shenoy PP (2007) Using Bayesian networks for bankruptcy prediction: some methodological issues. *Eur J Oper Res* 180:738–753

- Vassalou M, Xing Y (2004) Default risk in equity returns. *J Financ* 59:831–868
- Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57:307–333
- Zmijewski ME (1984) Methodological issues related to the estimation of financial distress prediction models. *J Acc Res* 22:59–82