Modeling Strategies for Prediction of Hedge Fund Failure: A Comparison of Parametric and Semi-Parametric Approaches

Jose Faias*
Universidade Católica Portuguesa, Lisbon, Portugal - jfaias@ucp.pt

Leila Amorim
Universidade Federal da Bahia, Salvador, Brazil - leiladen@ufba.br

Abstract

The proportion of hedge funds failing every year is on average of 10%. Past literature has addressed this phenomenon using a ad-hoc model without understanding which fits the data better. In this study, we run a comprehensive study of parametric and semi-parametric models to fit this event. The data is constituted by more than 3,000 funds over the last 15 years using either only static covariates or allowing some to be time-varying. The time-varying covariates allow a possibly more precise description of the impact of hedge fund's characteristics on their prediction to failure.

Keywords: Cox model; parametric survival models; proportional odds model; time-varying covariates.

1. Introduction

Hedge funds are investment pools, typically organized as a private partnership, that face few regulatory restrictions on their portfolio transactions. In quest for returns, many hedge funds rely on alternative risk factors which include leverage, investments in illiquidity markets, potentially high portfolio turnover and non-vanilla OTC contracts. There are a few papers trying to understand hedge fund failure which adopted survival analysis or qualitative response models. For example, Brown et al. (2001), Gregoriou (2002), Grecu et al. (2007) and Liang and Park (2010) use the Cox proportional hazards (CPH) model to examine factors affecting hedge fund failure while Baquero et al. (2005), and Malkiel and Saha (2005) use logit or probit models to investigate hedge fund survival. They also considered different databases and time frames for the analysis, making difficult, if no impossible, to compare the model contributions for evaluation of hedge fund's performance.

There are two major regression model strategies used for right censored data: proportional hazards (PH) model (Cox) (Cox , 1972) and accelerated failure time (AFT) models as parametric models, including Generalized Gamma, among other distributions for modeling time to event. Although Cox's semi-parametric model is well known, easy to interpret and very flexible – allows the inclusion of time-dependent covariates, multiple failure times, recurrent events, left truncation, among other additional data structures – the parametric models may lead to more efficient estimates under certain circumstances (Cox and Oakes, 1984). When PH assumption is violated, however, Cox regression estimates become biased. Thus, we also considered the proportional odds (PO) model as a special case of the transformation survival models to fit this data (Bennett, 1983; Zeng and Lin, 2007).

In order to make a comprehensive discussion of strategies to predict hedge fund's failure, we analyze two settings: (i) all characteristics of interest are fixed or averaged over time (static setting), and (ii) some characteristics are fixed and others are treated as time-dependent (time-varying setting). We describe the methods adopted and discuss their advantages and limitations for analysis of hedge funds survival in order to contribute to a more consistent discussion about strategies to evaluate hedge fund's performance.

In Section 2 we describe our data. In Section 3, we review the statistical models. Section 4 presents the results and finally we conclude.

2. Data

We extract monthly fund level data from The Tremont Advisors Statistical Services (TASS) hedge fund database obtained from Thomson Reuters from January 1995 until April 2013. TASS provides two databases: live and graveyard. The live database includes the currently existing funds, whereas the graveyard database has the so-called dead funds - funds that either no longer exist or have stopped reporting to the TASS service but still exist. By incorporating both databases we eliminate the survivorship bias [Ackermann et al. (1999)]. We require additional criteria for the inclusion of funds in our dataset. We drop funds of hedge funds (FOHFs) and managed futures funds (CTAs) since we want to focus only on individual hedge funds. Funds with less than 24 months of return history are not included because we delete the first 2 years of return data to mitigate the instant history bias. After the screening procedure, our final sample consists of 3,115 hedge funds. Among them, 683 funds are live as of April 2013, and the remaining 2,432 funds are considered defunct.

In our survival analysis, we use variables that may impact the survival of hedge funds. Prior literature suggests that performance, risk, HWM, managerial investment and lockup provision affect the attrition of hedge funds (see Brown et al., 2001 and Liang and Park, 2010). *Av Ret* is the average hedge fund's Rate of Return in the previous year in %. The poorer the performance, the more likely is the fund to terminate. *Std Ret* is the standard deviation of hedge fund's Rate of Return over the previous 60 months in % or if not available at least 24 months. The higher the standard deviation, the higher the hazard rate of the fund. *HWM* (High Water Mark) is a dummy variable equal to 1 if the hedge fund has a high water mark provision, 0 otherwise, and *Man Inv* is a dummy variable equal to 1 if the hedge fund manager has invested his/her personal capital, 0 otherwise. Funds with HWM or managerial investment are more likely to survive since this is indication of greater managerial quality, since fund's managers only receive incentives after recovering from negative performance. *Lockup* is a dummy variable equal to 1 if the hedge fund has a lockup period. Funds with a lockup provision can retain investors even after poor performance so long as the lockup period is not over, and thus these funds are more likely to survive. For the static setting, we take the average of the entire history of the fund for the continuous variables. All categorical variables are fixed over time.

3. Modelling Strategies

We define T as the continuous event time measured from the hedge fund inception to the last month of reported performance to TASS database. When follow up ends and the event (failure or ceasing reporting) did not occur it results in right censoring. Let U be the follow up time. When T exceeds U, then the time to event is censored. Let X be the $n \times (p+1)$ covariate matrix whose influence on the distribution of T is of interest. Due to the longitudinal feature of the data gathering some covariates are time-independent (e.g, managerial investment) while others are time-varying (e.g., average return). Two functions of great interest in duration analysis are the survival function [S (t)] and the hazard function [$\lambda(t)$], which can be defined, for time-fixed covariates, as: $S(t|X) = P(T > t|X) = exp(-\int_0^t \lambda(u|X).du)$ and $\lambda(t|X) = lim_{\Delta t \to 0} \frac{P(t \leq T < t+\Delta t|T \geq t, X)}{\Delta t}$, where $\lambda(t|X)\Delta t$ may be interpreted as a conditional probability since $\lambda(t|X)\Delta t \approx P(t \leq T < t + \Delta t|T \geq t, X)$. Thus, assuming that Δt is very small, $\lambda(t|X)$ represents an instantaneous risk of failure at time t, conditional on surviving up to t. The probability density function of T can be defined as $f(t) = \lambda(t)S(t)$. The behavior of the hazard function might be useful to describe the distribution of time to event data, showing how the instantaneous hazard rate changes in time. Both functions are used in modeling strategies for time to event data. Next, we briefly define the parametric and semi-parametric survival models adopted.

3.1 Parametric models

We considered five parametric distributions for the accelerated failure-time (AFT) model: Exponential, Weibull, Log-logistic, Log-normal and Generalized Gamma. Using this approach, the natural logarithm of the survival time is expressed as a linear function of the covariates, yielding the linear model $ln(T) = X\tilde{\beta} + \epsilon$, where $\tilde{\beta}$ is a $(p+1)$ vector of regression coefficients, and ϵ is the error with density $f()$. The distributional form of the error term determines the regression model. The effect of the AFT model is to change the time scale by a factor of $\exp(-X\beta)$, called "acceleration factor". The parameters of these models are estimated via maximum likelihood, where the log likelihood is defined by $logL = \Sigma_{i=1}^{n}\{\delta_i log f_i(t_i) + (1-\delta_i) log S_i(t_i) - log S_i(t_0)\}$, considering $i = 1, ..., n$ observations, the trivariate response, (t_{0i}, t_i, δ_i), representing a period of observation $(t_{0i}, t_i]$, ending in either failure ($\delta_i = 1$) or right-censoring ($\delta_i = 0$) and assuming independent censoring. This structure allows the analysis of a wide variety of models, including time-varying covariates. The distributions for the parametric models are implicit because of the chosen $S(t)$ and its corresponding density function $f(t)$.

3.2 Semi-Parametric models

A class of models that has been widely used is the Cox proportional hazards (PH) model (Cox, 1972), which assumes $\lambda(t|X) = \lambda_o(t|X)\exp(X\beta)$, where $\lambda_o(t|X)$ is an unspecified baseline hazard function (non-parametric), describing the risk for individuals with $\mathbf{X} = 0$, and $\exp(X\beta)$ is the relative risk associated to the covariates. Differently from the parametric models, PH model does not assume a shape for $\lambda_o(t|X)$ and instead captures its underlying shape based on the data. The partial likelihood approach is used to estimate the parameters in the PH model. Alternatively we also use a proportional odds (PO) model. We implement PO under the specification of transformation survival models. Nonparametric maximum likelihood, where an unknown failure distribution is treated as an infinite-dimensional parameter, is used for parameter estimation. Expectation-maximization (EM) algorithm is adopted for the computation of NPMLE for the model parameters (Zeng and Lin, 2007). With inclusion of time-dependent covariates, we may use an extension of the Cox regression model without assumption of proportionality of risks. Here, interpretation of estimated effects is associated to time. This type of covariate can be easily incorporated into the described modelling framework through counting process notation, with appropriate definition of start and stop times.

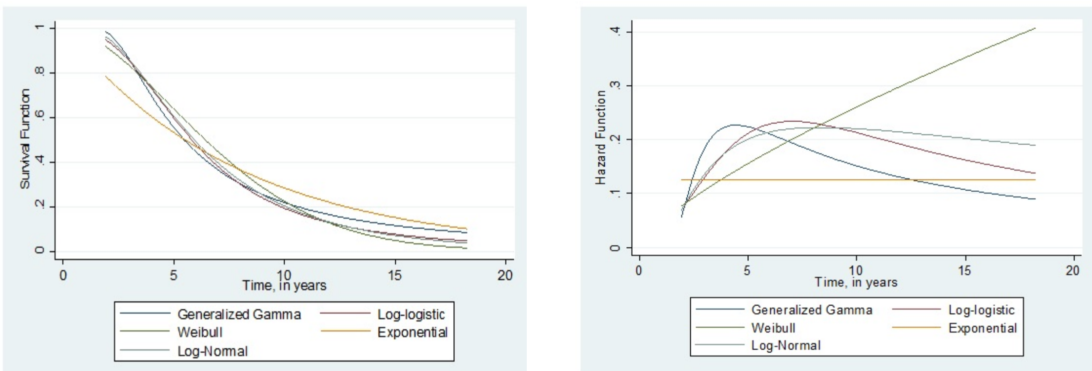
3.3 Models estimation and evaluation

We implement the parametric and semi-parametric survival models using STATA v.10. When considering all covariates as fixed, we compared parametric models with different distributions, and semi-parametric models: proportional Cox hazards (PH) model and proportional odds (PO) model. The PO model, however, is still not available in software for handling time-varying covariates. Therefore, when using time-varying covariates we considered only the parametric and PH models. Parametric models involve stronger assumptions than semi-parametric models, mainly due to the additional requirement of checking the appropriateness of the chosen distribution. Residuals are often used for model diagnostics, with a important role on discriminating models and judging their goodness of fit. Several methods based on residuals have been used for evaluation of survival models (Kalbfleisch and Prentice, 2002). We also consider graphical methods based on the Cox-Snell residuals for discriminating alternative parametric models. For Cox model, plots of martingale and Schoenfeld residuals were also used. As the scales of parametric and semi-parametric models differ, parameter estimates and estimated variances can not be directly compared. Nardi and Schemper (2003) had proposed the use of Wald-type χ^2 values and a standardized measure of variability ($sv = \hat{\sigma}_\beta/|\hat{\beta}|$), which is analogous to the coefficient of variation, to be used for evaluation of relative efficiency of parameter estimates between the two approaches. They highlight that neither measure is superior, with the standardized measure of variability tending to overemphasizing cases where parameter estimates are close to zero and Wald-type χ^2 values exaggerating the gain in precision for highly significant effects. When models are nested (e.g. Weibull vs. exponential), the likelihood-ratio or Wald test can be used to discriminate between them. When models are not nested, these tests are inappropriate. We then use the Akaike information criterion (AIC) or the Bayesian information criterion (BIC). AIC can be defined by: $AIC = -2$ (log likelihood)$+2 \times (p + q + 1)$, where p is the number of covariates in the model and q is the number of ancillary parameters of the model. BIC, on its turn, may be defined as $BIC = -2$ (log likelihood)$+ln(n) \times (p + q + 1)$, where n refers to the number of observations. The preferred model is the one with the smallest value of either criteria. For both AIC and BIC, however, the likelihood functions must be conformable; that is, they must be measuring the same event. Furthermore, BIC can only be compared for models using the same number of observations. Their only difference in practice is the size of the penalty, in which BIC penalizes model complexity more heavily. The only way they should disagree is when AIC chooses a larger model than BIC.

4. Results

4.1 Static Setting

The first step is to fit the exponential, Weibull, loglogistic, lognormal, and gamma models separately, without covariates. We estimated the survival and hazard functions for the five parametric survival models (Figure 1). Note the different shapes of the hazard functions for different models, which demonstrates that it does make a difference which model is selected. The generalized gamma model is the most suitable for these data. There are a few results pointing out in this direction. (1) It achives the highest log likelihood and smallest AIC and BIC. (2) Wald tests on gamma versus either lognormal or Weibull clearly reject the former distributions (*p-value* = 0.000), and (3) Cox-Snell residuals only apart from expected when time is large.



(a) Predicted Survival (b) Hazard Function

Figure 1: Cox Snell Residuals and Cumulative Hazard Function for Generalized Gamma Models

The second step is to include fixed covariates in the parametric models. Gamma is again the best model to fit the data following the same reasoning as before. The first columns of Table 1 present the estimated results for the generalized gamma. In AFT models, the sign of the coefficient indicates how a covariate affects the logged survival times. A one unit increase in the return standard deviation ($StdRet$) leads to a decrease of 0.611 in the logged survival time or equivalently one unit increase in the return standard deviation decreases the survival time by 46%. Lockup has also a negative effect on HF survival, such that HFs that has a lockup period leads to a decrease of 14% on the survival time.

The next step is to fit the Cox models. Table 1 presents the estimated PH and PO models in the last four columns. Within semi-parametric models, PH presents the lowest AIC and BIC values. The signs of all coefficients are consistent across parametric and semi-parametric models. Note that the parametric approach explains the time to failure, whereas Cox models explain the time to survival. Average return was the only statistically significant factor, at 5% level, for all models, suggesting that hedge funds with higher average returns have lower likelihood to fail. Managerial investment was only significant for the Cox model, indicating that there is a reduced likelihoood of hedge fund's failure if the hedge fund manager has invested his/her personal capital. We use the χ^2 and the sv values of the covariates to compare across different models. There is higher significance in covariates in the semi-parametric models compared to the gamma model, exception for lockup. Higher efficiency is obtained through Cox proportional hazards model, producing lower standardized variability in general than the other two models. We also compare the Cox and gamma models using Cox-Snell residuals. Again Cox fits the best. Overall, the PH Cox model was the one with best fit to the hedge funds data in the static setting.

Table 1: Estimates of Risk Effects from Generalized Gamma, Proportional Hazards and Proportional Odds Models for Time to Hedge Fund Failure.

Variable	**Models**					
	Gen Gamma		**Prop Hazards**		**Prop Odds**	
	$\hat{\beta}$	**SE**($\hat{\beta}$)	$\hat{\beta}$	**SE**($\hat{\beta}$)	$\hat{\beta}$	**SE**($\hat{\beta}$)
Av Ret	7.019*	(1.167)	−36.479*	(3.425)	−40.052*	(4.493)
Std Ret	−0.611	(0.337)	1.445	(0.758)	2.000	(1.100)
HWM	0.031	(0.026)	−0.078	(0.047)	−0.125	(0.075)
Man Inv	0.025	(0.023)	−0.106*	(0.044)	−0.125	(0.069)
Lockup	−0.015	(0.023)	0.008	(0.044)	0.031	(0.069)
AIC	5,593.81		39,874.52		39,915.66	
BIC	5,642.16		39,904.74		39,945.88	

*Significant at 5% significant level.

4.2 Time-Varying Covariates Setting

After inclusion of time-varying covariates – return average and standard deviation – we re-evaluate the best parametric model, which is again the generalized gamma model (results not shown for space reasons). Table 2 presents the results only for the generalized gamma and the extended Cox models, both allowing the inclusion of time-varying covariates.

Table 2: Estimates of Risk Effects from Generalized Gamma and Extended Cox Models for Time to Hedge Fund Failure, including Time-Varying Covariates.

Variable	**Models**							
	Generalized Gamma				**Extended Cox**			
	$\hat{\beta}$	**SE**($\hat{\beta}$)	**Wald** χ^2	**SV**	$\hat{\beta}$	**SE**($\hat{\beta}$)	**Wald** χ^2	**SV**
Av Ret	4.327*	(0.924)	21.930	0.214	−12.047*	(2.473)	23.731	0.205
Std Ret	−0.393	(0.435)	0.816	1.107	−0.969	(0.913)	1.126	0.942
HWM	0.075*	(0.029)	6.688	0.387	−0.117*	(0.052)	5.063	0.444
Man Inv	0.029	(0.025)	1.346	0.862	−0.107*	(0.047)	5.183	0.439
Lockup	−0.018	(0.025)	0.518	1.389	0.014	(0.047)	0.089	3.357
AIC	5,003.63				29,570.43			
BIC	5,076.65				29,616.06			

*Significant at 5% significant level.

The interpretation of the estimated effects is associated to time t, i.e, in the Cox model, for instance, they represent the effect on the logarithm of hazard function when changing one unit for time-varying covariates at time t, keeping the other variables constant in the model at time t. Average return remains being statistically significant, at 5% level, for all models, indicating that given the same values of the other included factors for two hedge funds, the hazard for failure at time t decreases for those funds with higher average returns before time t. Differently from the static setting, here the HWM is statistically significant in both models, suggesting that there is a reduction on the likelihoood of hedge fund's failure if it has a high water mark provision. For time-varying covariates, the standardized variability is lower in the Cox model, while for the fixed covariates they do not show a clear pattern. The same is observed by comparing the Wald χ^2 statistics in the two models. We then compare the time-varying against static models. We can only do this for the same type of models, i.e., time-varying gamma versus static gamma and time-varying cox versus static cox. When using either AIC or BIC the results are the same. As expected, there is always an improvement of using time-varying variables.

5. Conclusions

Interpretability of the estimates is a very important property when choosing a model. Many researchers are familiar to the RR interpretation of parameters in Cox's model and have some difficulty interpretating in AFC metric. If the PH assumption does not hold, however, PH Cox regression estimates become biased and there is a loss of power for estimation or hypothesis testing. Nevertheless, it is rare to find a study in hedge funds literature that discuss their attempt to verify the underlying assumptions of the models. One way to handle the violation of PH assumption is by including time-varying effects using an extended Cox model. Inclusion of this type of covariates can be important to evaluate dynamic effects. In our study we used the continuous variables, which are measured montly, as fixed (using as an average) and as time-varying to compare and describe their effects on the evaluation hedge fund's performance. The time-varying covariates allow a possibly more precise description of the impact of hedge fund's characteristics on their prediction to failure.

References

Ackermann, C., McEnally, R., & Ravenscraft, D. (1999) The performance of hedge funds: Risk, return, and incentives, Journal of Finance 54: 833-874.

Baquero, G., Horst, J., Verbeek, M. (2005) Survival, look-ahead bias, and persistence in hedge fund performance, Journal of Financial and Quantitative Analysis 40: 493-517.

Bennett, S. (1983). Analysis of survival data by the proportional odds model, Statistics in Medicine 2: 273-277.

Brown, S.J., Goetzmann, W.N., Park, J. (2001) Careers and survival: Competition and risk in the hedge fund and CTA industry, Journal of Finance 56: 1869-1886.

Cox, D. (1972). Regression models and life-tables. (with discussion), Journal of the Royal Statistical Society B 34: 187-220.

Cox, D., & Oakes, D. (1984). Analysis of Survival Data. London, England, Chapman and Hall.

Grecu, A., Malkiel, B.G., & Saha, A. (2007) Why do hedge funds stop reporting performance?, The Journal of Portfolio Management 34: 119-126.

Gregoriou, G.N. (2002) Hedge fund survival lifetimes, Journal of Asset Management 3: 237-252.

Haghani, S. (2014) Modeling hedge fund failures: A dependent competing risks framework with latent exit types, Journal of Empirical Finance 28: 291-320

Kalbfleisch, JD., & Prentice, RL. (2002) The Statistical Analysis of Failure Time Data, Second Edition, New York, USA, Wiley.

Lee, H.S., & Kim, T.Y (2014) Dynamic prediction of hedge fund survival in crisis-prone financial markets, Journal of Banking & Finance 39: 57-67.

Liang, B., & Park, H. (2010) Predicting hedge fund failure: A comparison of risk measures, Journal of Financial and Quantitative Analysis 45: 199-222.

Malkiel, B.G., Saha, A. (2005) Hedge funds: risk and return, Financial Analyst Journal 61: 8088.

Nardi, A., & Schemper M. (2003) Comparing Cox and parametric models in clinical studies, Statistics in Medicine 22: 3597-3610.

Rossini, A.J., & Tsiatis, A.A. (1996) A Semiparametric Proportional Odds Regression Model for the Analysis of Current Status, Journal of the American Statistical Association, 91: 713-721.

Zeng, D. and Lin, D. Y. (2007) Maximum likelihood estimation in semiparametric regression models with censored data (with discussion). Journal of Royal Statistical Society Series B 69: 507564.