European Journal of Operational Research 249 (2016) 457–464

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

The stability of survival model parameter estimates for predicting the probability of default: Empirical evidence over the credit crisis

Mindy Leow, Jonathan Crook*

Credit Research Centre, University of Edinburgh Business School, 29 Buccleuch Place, Edinburgh EH8 9JS, Scotland, United Kingdom

ARTICLE INFO

Article history:
Received 18 February 2014
Accepted 1 September 2014
Available online 18 October 2014

Keywords:
Forecasting
Robustness and sensitivity analysis
Macroeconomic variables
Structural change
Probability forecasting

ABSTRACT

Using a large portfolio of credit card loans observed between 2002 and 2011 provided by a major UK bank, we investigate the stability of the parameter estimates of discrete survival models, especially since the start of the credit crisis of 2008. Two survival models are developed for accounts that were accepted before and since the crisis. We find that the two sets of parameter estimates are statistically different from each other. By applying the estimated parameters onto a common test set, we also show that they give different predictions of probabilities of default. The changes in the predicted probability distributions are then investigated. We theorise them to be due to the quality of the cohort accepted under different economic conditions, or due to the drastically different economic conditions that was seen in the UK economy, or a combination of both. We test for each effect.

1. Introduction

The application of survival analysis models onto credit-related problems is not new (for example, see Banasik, Crook, & Thomas, 1999; Pennington-Cross, 2010) and is welcomed for its ability to take into account factors that are inherent in the modelling of credit risk and the prediction of credit events, where regression methods are unable to. First, survival models are able to account for censoring, which allows for a realistic and practical model to be developed. Second, they are able to incorporate time-dependent variables with ease, which will allow the inclusion of time-dependent account-specific covariates as well as time-dependent macroeconomic variables in credit models. When this is combined with simulation, a plausible platform for stress testing is created, as proposed by Rodriguez and Trucharte (2007), Leow, Mues, and Thomas (2011) and Bellotti and Crook (2013, 2014). Third, and most crucially, survival models are able to generate probabilities of how likely an event is to occur over time, conditional on the event not having occurred before, and this provides a dynamic framework for the prediction of credit events (e.g. default or customer churn of credit loans, repossession or early-prepayment for mortgage loans). Because the likelihood of the credit event occurring over time can be estimated, the corresponding losses (Mcdonald, Matuszyk, & Thomas, 2010) or profits (Ma, Crook, & Ansell, 2010) can also be predicted. In terms of how well survival models predict, there has been some work done specifically to compare its prediction to that of regression models: Stepanova and Thomas (2002) looked at the model performances in the prediction of early prepayment and default of personal loans; Bellotti and Crook (2009) looked at model performances in the prediction of default of credit card loans. Both papers found that survival models are able to predict better than static regression models.

This work does not attempt to revisit the advantages of survival models over their regression counterparts – that much has been established in the literature over different retail products. The work here differs from the existing literature in two ways. First, we have a rich source of credit card loan data that goes from 2002 to 2011, and so encompasses the credit crisis from 2008, which is not commonly available. Macroeconomic indicators over time will show a large difference in values, and it would be interesting to explore how these large and unexpected changes would affect default models and their predictions. Second, we investigate the stability of survival model parameter estimates before and after the credit crisis. Using a portfolio of active credit cards observed between January 2002 and March 2011, we investigate whether parameter estimates change over the crisis period, and whether the inclusion of time-varying covariates representing the economy are able to adequately account for changes to debtors' propensity to default. By separately and independently estimating a survival model for periods before and since the start of the credit crisis, i.e. 2002–2007 and 2008–2011 respectively, we use the Chow test (more details in Section 4.1) to check for statistical differences between the two sets of parameter estimates. To illustrate how the two sets of parameter estimates are different, we apply each survival model developed onto a common test set to get the average predicted probabilities over the (duration) time of the loan.

* Corresponding author. Tel.: +44 131 650 3802.
E-mail addresses: mindy.leow@ed.ac.uk (M. Leow), j.crook@ed.ac.uk (J. Crook).

http://dx.doi.org/10.1016/j.ejor.2014.09.005

During the course of this work, population drift, and how it might affect parameter estimates, is also considered as a related issue, due to the differing types of debtors securing credit accounts before and during the credit crisis. However, because of the large variations in macroeconomic conditions that was seen in our period of interest, it is also possible that changes in distributions of probabilities are due to the changes in these macroeconomic variables. We investigate the effects of either by selecting two cohorts, representing a set of accounts accepted during a non-downturn period and a downturn period, and estimating a survival model for each period. We then create test sets based on each training set, by holding constant either the cohort quality or the macroeconomic conditions, and compare the distribution of predicted probabilities to see how the distributions shift due to changes in cohort quality or economic conditions. We find macroeconomic conditions do affect probabilities of default, and could affect different groups of debtors in different ways.

2. Methodology

We use data gathered at regular, discrete monthly points in time, and the default event is recorded in a particular month with reference to the month the account was open. Therefore we estimate the survival models in discrete time. Another advantage of discrete time rather than continuous time survival models is a much lower computational time in model estimation. This is important because we deal with a large dataset.

Let $P_{i\tau}$ be the probability that an individual account i goes into default at duration time (of loan) τ, given that default has not happened up to time $\tau - 1$, and the final model developed is given in Eq. (1).

$$\log\left(\frac{P_{i\tau}}{1-P_{i\tau}}\right) = \alpha_\tau + \beta_1\mathbf{X}_i + \beta_2\mathbf{Y}_{i\tau-3} + \beta_3(\mathbf{Z}_{\tau-3} - \mathbf{Z}_{\tau-15}) + \beta_4\mathbf{X}_i(\mathbf{Z}_{\tau-3} - \mathbf{Z}_{\tau-15}) \tag{1}$$

where α_τ represents the effect of time on the odds of default; $\mathbf{X}_i$ is a vector which represents the time-independent, account-dependent covariates, i.e. application variables; $\mathbf{Y}_{i\tau-3}$ is a vector which represents time- and account-dependent covariates, i.e. behavioural variables, lagged 3 months; $\mathbf{Z}_{\tau-3} - \mathbf{Z}_{\tau-15}$ is a vector which represents time-dependent, account-independent covariates, i.e. macroeconomic variables, at 12th difference and lagged 3 months; and $\mathbf{X}_i\left(\mathbf{Z}_{\tau-3} - \mathbf{Z}_{\tau-15}\right)$ is a vector which represents interaction terms between selected application variables and macroeconomic variables at 12th difference and lagged 3 months.

In this regression model, the dependence of the hazard on time, α_t, is specified as $\alpha_1\tau + \alpha_2\tau^2 + \alpha_3 \ln\tau + \alpha_4(\ln\tau)^2$. By doing so, we allow the relationship between the effect of time and the odds of default to be very flexible with an added advantage of allowing for prediction beyond the maximum duration time that is observed in the training set.

A number of model variations were considered in the course of this work, mainly experimenting with the way the macroeconomic variables were included in the model. Lags of between 3 and 12 months were considered, and to address the possible correlation between macroeconomic variables, both levels and 12th differences, lagged or otherwise, of each macroeconomic variable were examined.

3. Data

The data is supplied by a major UK bank and is a random sample of credit cards that were issued in the UK between 2002 and 2010. It consists of almost 538,000 unique credit card accounts and each account is tracked monthly up to March 2011, or until the time the credit card account is closed, whichever is earlier. Common application variables are available: type of employment, length of time the debtor has been with the bank, income at application and age at application, among others. Because each account is updated monthly, we also have behavioural variables, including repayment amount, credit limit and outstanding balance, from which further behavioural indicators could be inferred, for example, how frequently the account misses payment(s) over its entire history. Any behavioural variables included in the model are lagged by 3 months.

Table 1
Dataset splits.

Dataset	Acceptance period	Observation period
Training set I	January 2002 to August 2007	May 2002 to December 2007
Training set II	January 2008 to July 2010	May 2008 to March 2011
Combined/“test”	January 2002 to July 2010	May 2002 to March 2011

Although default information is available from the dataset, it is not consistently defined across the entire dataset. Therefore, a monthly minimum repayment amount is defined and is used to define arrears and default. This minimum repayment amount is 2.5 percent of the previous month’s outstanding balance or £5, whichever is higher, unless the account is in credit, in which case the minimum repayment amount is £0, or the account has an outstanding balance of less than £5, in which case the minimum repayment amount would be the full outstanding amount. An account is said to be in arrears if it does not make the minimum payment. A default is said to occur if and when an account goes 3 months in arrears (not necessarily consecutive). Note that this definition of default is not the conventional “three consecutive months of missed payment”, but is acceptable as financial institutions are not bound to this definition of default (Basel Committee on Banking Supervision, 2004 , paragraphs 452–456). As the work here only focuses on the default event, we do not specify the transitions between states of arrears in the preceding months; further details can be found in Leow and Crook (2014).

3.1. Training and test set splits

The dataset is used in a number of ways here. In order to accommodate the lagged behavioural covariates, only accounts that are observed for longer than 3 months since each was opened are included.

First, the dataset is split into two training sets (see Table 1). The first consists of accounts that started between January 2002 and December 2007 inclusive, with an observation period up to December 2007, i.e. any remaining active accounts are censored in December 2007. The second consists of accounts that started between January 2008 and July 2010 inclusive, with an observation period up to March 2011, i.e. accounts are censored in March 2011, if the account has not been closed earlier. Note that the two training sets are completely separate. The creation of these two training sets represent portfolios of loans that were accepted before and during the credit crisis, since we expect bank policies and acceptance decisions to change slightly over the years, with distinguished differences before and since the start of the credit crisis.

Due to the split of the training sets, it is not sensible to try and reduce the length of either training set further to get a test set. In order to get an indication of how similar (or different) the models of each training set would predict, we apply the respective models onto the entire dataset, i.e. combining training sets I and II, as a test set. Doing so would mean that a common test set is used without any further loss to both training sets in terms of observations and observation period.

3.2. Macroeconomic variables

The macroeconomic variables considered are given in Table 2. The main source of macroeconomic variables is the Office of National