

Credit Scoring With Macroeconomic Variables Using Survival Analysis

Tony Bellotti and Jonathan Crook
Credit Research Centre
Management School and Economics
University of Edinburgh

7 May 2007

Abstract

Survival analysis can be applied to build models for time of default on debt. In this paper we report an application of survival analysis to model default on a large data set of credit card accounts. We show that survival analysis is competitive for prediction of default in comparison with logistic regression. We explore the hypothesis that probability of default is affected by general conditions in the economy over time. These macroeconomic variables cannot readily be included in logistic regression models. However, survival analysis provides a framework for their inclusion as time-varying covariates. Various macroeconomic variables, such as interest rate and unemployment index, are included in the survival model as time-varying covariates. We show that inclusion of these indicators improves model fit and affects probability of default and provides a statistically significant improvement in predictions of default on an independent test set.

Keywords: credit scoring, survival analysis; risk; banking.

JEL: C41, D14, G21

Introduction

The use of statistical methods for credit scoring and prediction of default on credit card accounts is now well-known. In particular, logistic regression has become a standard method for this task (Thomas et al 2002). There has recently been interest in using survival analysis for credit scoring. This allows us to model not just if a borrower will default, but when. The advantages of this method are that (i) survival models naturally match the loan default process, (ii) it gives a clearer approach to assessing the likely profitability of an applicant and (iii) survival estimates will provide a forecast as a function of time (Banasik et al 1999). Survival analysis has been applied in many financial contexts including explaining financial product purchases (Tang et al. 2007), behavioural scoring on credit customers (Stepanova &

Thomas 2001), predicting default on personal loans (Stepanova & Thomas 2002) and the development of generic score cards for retail cards (Andreeva 2006).

In this paper we test the hypothesis that probability of default (PD) is affected by general economic conditions that are measured by macroeconomic variables such as bank interest rates, unemployment index, house price and so on. For example, we expect that rises in interest rates may increase the risk of an individual failing to make payments due to increased payment demands on loans and mortgages as well as outstanding credit card debt. We do this using survival analysis since it has the additional advantage that such macroeconomic time series data can naturally be incorporated into a survival model as time-varying covariates (TVCs) (Banasik et al 1999). This cannot be done so easily with the usual regression or logistic regression models. We conduct experiments to test the affect of macroeconomic variables on the PD of credit card account holders. We are interested in assessing macroeconomic variables in terms of explanation and prediction of default. Survival models with TVCs are constructed and contrasted with standard logistic regression models to determine any uplift in predictive performance. We show that the inclusion of macroeconomic variables gives a statistically significant explanatory model of the data and a statistically significant uplift in predictive performance over not including macroeconomic variables.

Data

Sample application and monthly performance data were used for a single UK credit card product provided by a UK bank. This sample spanned a period of credit card accounts opened from 1997 to mid-2005. Accounts opened between 1997 and 2001 were used as a training data set, and those opened between 2002 and 2005 were used as a test data set. Each data set contained over 100,000 accounts with application variables such as income, age, housing and employment status along with a bureau score taken at the time of application.

For this experiment, an account is defined as having *defaulted* if it goes 3 months down or more within the first 12 months. An account that defaults is referred to as a

bad case and a non-defaulting account is referred to as a *good* case. For this data set, using this definition, the proportion of bad cases in the data was small.

Macroeconomic Variables

Several macroeconomic variables are used and are described in Table 1. These macroeconomic variables were selected as the most likely to affect default. Table 1 shows the effect we expect each variable to have on risk of default. A positive value means that as the value of the macroeconomic variable rises, this is linked to a rise in risk of default. Conversely a negative expectation means that an increase in the value is linked to a decrease in risk. So an increase on interest rates, unemployment and house prices is expected to place further stress on the population which will translate as higher risk. However, increases in earnings, the FTSE index and production are indicators of an improving economy providing conditions for reduced risk of default.

TABLE 1 HERE

A rise in consumer confidence is expected to generally lead to greater relative risk since people will be more likely to consume and borrow more making repayment more difficult. We expect that interest rates will have the largest influence on risk of default, since this change has a more direct affect on repayments.

Methods

We describe model training, model selection and model assessment methods below.

Model Training

Training data is modelled using a Cox proportional hazard (PH) survival model to model the time of default of each case. It is contrasted with logistic regression (LR) which is a standard model in credit scoring. Both types of models are described below. A Cox PH model without macroeconomic variables is also built to determine whether any uplift in performance is due to the use of the Cox PH model or the inclusion of macroeconomic variables. Since the data is skew, in terms of numbers of good to bad cases, we give greater weight to bad cases for training in proportion to numbers of bad to good cases in the training data. This is possible for both Cox PH

and LR models since both use maximum likelihood estimation for which bad cases can be included in the likelihood function multiple times.

Cox Proportional Hazard Model with Time-varying Covariates

Survival analysis is used to study time to *failure* of some population. This is called the *survival time*. Survival analysis is able to facilitate the inclusion of observations that have not failed. These are treated as *censored* data and an observation time can be given for censored cases indicating the last time they were observed. In the context of consumer credit, the population comprises individuals applying for credit in the form of loans or credit cards. When a consumer defaults on a loan or credit card payment then this is a failure event. Survival time is measured from the date the account was opened. If a consumer never defaults during the lifetime of their account then they are censored and observation time is the period of time the account was open or, if the account was never closed, the time from when the account was opened to the date of data collection.

A common means to analyze survival data is through the *hazard function* which gives the rate of change of probability of failure at a time t :

$$h(t) = \lim_{\delta t \rightarrow 0} \left(\frac{P(t \leq T < t + \delta t \mid T \geq t)}{\delta t} \right) \quad (1)$$

where T is a random variable associated with survival time. The probability of survival at time t can be given in terms of the hazard function:

$$S(t) = P(T \geq t) = \exp\left(-\int_0^t h(u) du\right). \quad (2)$$

This is the probability of survival up to time t (Collett 1994, Section 1.3). For credit data, this is the probability that an account has not defaulted by some time t after the account has been opened, ie 1-PD at time t .

A series of n observations $i=1$ to n is given in survival analysis in terms of observation times t_i and indicators c_i where $c_i=0$ for a censored observation and $c_i=1$ for a failure event, in which case t_i is the survival time. In addition, each observation will include a vector of covariates that may be associated with survival time. Some of these may be time-varying so, in general, they are given as functions of time, $\mathbf{x}_i(t)$. Application data is fixed with respect to time. However, macroeconomic variables

change over time and the value of the covariate is given as the value of the macroeconomic variable at the time of failure.

Several models of the hazard function are available, but in this paper we use the Cox PH model since it allows for the inclusion of macroeconomic variables as time-varying covariates (TVCs). This model is semiparametric, depending partly on a vector of coefficients $\boldsymbol{\beta}$ that are linear multiples of the covariates and a nonparametric baseline hazard function h_0 dependent on time but not the covariates. With TVCs, the Cox PH model is

$$h(t, \mathbf{x}(t), \boldsymbol{\beta}) = h_0(t) \exp(\boldsymbol{\beta} \cdot \mathbf{x}(t)) \quad (3)$$

which gives hazard at time t for observation \mathbf{x} , given parameters $\boldsymbol{\beta}$. This model yields a partial likelihood function on the training observations,

$$l_p(\boldsymbol{\beta}) = \prod_{i=1}^n \left(\frac{\exp(\boldsymbol{\beta} \cdot \mathbf{x}_i(t_{(i)}))}{\sum_{j \in R(t_{(i)})} \exp(\boldsymbol{\beta} \cdot \mathbf{x}_j(t_{(i)}))} \right)^{c_i} \quad (4)$$

where $t_{(i)}$ are ordered survival times and the risk set $R(t) = \{j : t_{(j)} \geq t\}$. This allows the use of maximum likelihood estimation to estimate $\boldsymbol{\beta}$ without needing to know the baseline hazard (Hosmer & Lemeshow 1999, Section 7.3). However, in order to use the model for estimation of survival probabilities, the baseline hazard is needed and can be estimated based on the parameter estimates $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ given by maximum likelihood estimation.

Logistic Regression

LR is a standard parametric technique used for credit scoring. It is used to model the log-odds of an event given a vector of covariates \mathbf{x} :

$$\log\left(\frac{p}{1-p}\right) = \mathbf{w} \cdot \mathbf{x} \quad (5)$$

where p is probability of the event and \mathbf{w} is a linear combination of weights on \mathbf{x} . Given a set of n training observations \mathbf{x}_1 to \mathbf{x}_n , the weight vector can be estimated using maximum likelihood estimation. For credit scoring, the event is default, therefore p is the PD (Thomas et al. 2002, Section 4.5).

Model Selection

We expected that the inclusion of interactions between application and macroeconomic variables may lead to better models since some categories of credit consumers would be more prone to changes in economic conditions than others. The following model selection method is employed to determine which interactions to include, based on the strategies described by Collett (1994, Section 3.6). Each macroeconomic variable is interacted with an application variable and added to the basic model. The uplift of model fit is then measured using the log-likelihood ratio (LLR) derived from the maximum likelihood procedure used to estimate the model. For each macroeconomic variable, the interaction giving the lowest p-value for its LLR is included in the optimal macroeconomic Cox PH model. Note that due to the large size of the training set, processing time to fit each model was long. This meant constraining the model selection phase and, in particular, it was judged that forward selection or backward elimination methods would be too time consuming to use.

Assessment

We assess our optimal model in terms of both its explanatory power on the training data and its predictive power on the independent test set.

Explanatory Model

The Cox PH model is assessed as an explanatory model by reporting its fit to the training data with and without macroeconomic variables using LLR. The significance of each coefficient in the model is determined using a Wald statistic derived from the maximum likelihood estimation. The Wald statistic follows a chi-square statistic, so a p-value can be computed for the null hypothesis that the coefficient value is zero (Hosmer & Lemeshow 1999, Section 3.3).

When a covariate x interacts with one or more covariates y_1, \dots, y_n , it is difficult to immediately determine the effect of x on the PD. However, it is possible to determine the marginal effect on log-hazard, γ_x of x , conditional on the interaction terms, as

$$\gamma_x = \beta_x + \sum_{i=1}^n \beta_{xy_i} y_i \quad (6)$$

where β_x and β_{xy_i} are coefficient estimates for x and each interaction xy_i for $i=1$ to n , respectively (Brambor et al 2005). In this paper we report a single figure for marginal effect by substituting the mean values of each interaction term y_1, \dots, y_n in Equation (6), thus providing a value of marginal effect of x for the mean observation. Table 2 lists prior expectations of the effects of each macroeconomic variable and the sign of the observed marginal effect can be tested against this expectation. In assessing the model, we would require that most match.

TABLE 2 HERE

The *importance* of a macroeconomic variable can be measured by the magnitude of the standardized marginal effect; ie the absolute value of the marginal effect multiplied by the standard deviation of the macroeconomic variable over the period of time of the data. This will give an approximate indication of the relative importance of each macroeconomic variable in the model.

Predictive Performance

To determine its usefulness as a credit scoring system, the Cox PH model is tested as a predictor of default. Predictions are made using survival probabilities for different survival times computed using the Cox PH model. Given a cut-off threshold, the survival probabilities are used as scores to predict good or bad cases. That is, if a case has a survival probability at 12 months that is greater than the cut-off then it is predicted as good (ie it is predicted as surviving default), otherwise it is predicted as a bad case. Predictions are made with LR in a similar way using a cut-off on PDs computed by the LR model.

A cost function is used to determine the value of a prediction:

- A correctly classified case has a cost of 0.
- A good case wrongly predicted as bad incurs a cost of 1.
- A bad case wrongly predicted as good incurs a cost of 20.

This skew in the cost of error reflects the fact that the cost of accepting bad debt is much higher than the cost of rejecting a good debt (Thomas et al 2002, Section 7.2). A cost of 20 is chosen since the low default rate in the data for this experiment means that a cost of 10 or lower would be so low that the most cost effective policy would simply be to accept all credit card applications whilst, on the other hand, a cost much greater than 20 is unlikely to reflect a realistic ratio of costs between decisions made based on wrongly predicted good and bad cases. Nevertheless, to demonstrate robustness, relative costs of 15 and 25 are also reported. We have chosen not to use receiver operating characteristic (ROC) curves to assess the models since they are insensitive to the relative costs that we can expect between errors on good and bad cases in consumer credit and, therefore, can possibly give rise to misleading conclusions (Hand 2005).

For each model, a cut-off threshold is computed for prediction which minimizes the total cost of errors on the training set. This cut-off is then applied to make predictions on the test set. Therefore, the predictions made on the test set are completely independent of the training data. However, the cut-off computed in this way is unlikely to be optimal for the test set and there will be a degree of fluctuation between the computed cut-off and the cost optimal cut-off for the test set. This will affect the relative performance of the models. In order to determine that improvement in performance is mostly due to the model, rather than a fluctuation in the cost effectiveness of the cut-off on the test set, the analysis is repeated with cut-offs computed so as to minimize total cost on the test set. This is likely to introduce a bias, but it does, however, allow us to discount fluctuations in the cut-off term as a cause of performance gain. If a particular model performs well with both cut-offs derived from training and test sets, then it shows that the model is both an unbiased good predictor and that the results are not due to fluctuations in the optimality of the cut-off threshold.

Assessment is made on the independent test set. The *mean cost* per observation is computed on the test set for each model as the sum of costs of error for all cases in the test set. Models resulting in a lower mean cost have performed better. To see how relative performance between models changes over time, the costs will be broken down over the period of the test set by year and quarter. The significance of any

improvement in performance of one model over another is measured using a paired *t*-test on the sequence of relative predictive costs between two models on the sequence of independent test cases (Witten & Frank 2005, Section 5.5). *Sensitivity* and *specificity* will also be reported for the optimal Cox PH with macroeconomic variables. These are the proportions of good cases in the test set predicted as good and bad cases predicted as bad, respectively. These figures allow us to contrast with results using other credit models to ensure our model's behaviour is typical (Baesens et al 2003).

Results

All the models fit the training data well and significantly as is shown in Table 2. The last row of this table shows that the inclusion of macroeconomic variables into the Cox PH model is highly significant in improving model fit.

TABLE 2 HERE

Several of the macroeconomic variables and their interactions included in the optimal Cox PH model proved to be important explanatory variables. These are shown in Table 3. The interactions were selected using the automated model selection method described in Section 2.2. The table reveals that most of the macroeconomic variables and their interactions are significant with p-values below a 0.01 significance level, as calculated using the Wald statistic.

TABLE 3 HERE

Table 4 shows the marginal effect for each macroeconomic variable, taking the interaction terms into consideration. They are calculated using Equation (6) with figures for coefficient estimates given in Table 3. The prior expected sign for each of the marginal effects, taken from Table 1, is also shown. These are correct for all the macroeconomic variables except the production index. In particular, the coefficients are positive for IR and Unemp, indicating a marginal increase in hazard (risk of default) with increases in bank interest rates and levels of unemployment. This is

what we would expect since higher interest rates mean generally higher repayments on credit and higher levels of unemployment mean less economic stability. Conversely, hazard decreases with increases in the FTSE index and levels of real earnings which is what we would expect since these are indicators of ability to repay.

TABLE 4 HERE

Table 4 also shows the relative importance (magnitude of standardized margin effect) of each macroeconomic variable. They are also shown in Figure 1 which indicates that interest rate is by far the most important macroeconomic variable influencing default risk as we would expect.

FIGURE 1 HERE

Table 5 shows prediction results on the test set for each model when the cut-off is computed using the training set. Results on the training set are also shown for contrast with the results on the test set. Table 5 clearly reveals that survival analysis improves performance in terms of reduced mean cost and that this is largely due to the inclusion of macroeconomic variables. This is true with a range of costs on bad errors. Significance tests given in Table 6 demonstrate that the improvement in performance is significant and that this is largely due to the inclusion of macroeconomic variables. The sensitivity and specificity using the Cox PH model with macroeconomic variables on the test data set, with cost on bad cases=20, are 97.1% and 13.8% respectively. These figures, with a much higher sensitivity in relation to specificity, are typical of credit data (see eg Baesens et al 2003) which demonstrates that our model is behaving typically as a credit model.

TABLE 5 HERE

TABLE 6 HERE

Table 7 shows test results when the experiments are repeated with the cut-off computed using the test set to yield optimal performance. Again, these results reveal an improvement in performance when macroeconomic variables are included,

indicating that the results are not fundamentally due to the method of computing the cut-off. Table 8 shows that the performance uplift is significant and due mainly to the inclusion of macroeconomic variables.

TABLE 7 HERE

TABLE 8 HERE

Figure 2 shows the mean cost difference between models over time. This figure shows that the inclusion of macroeconomic variables consistently gives better performance over time. The Cox PH model with macroeconomic variables outperforms LR in all periods except one in quarter 3 in 2002. The figure also shows that there is a general improvement in prediction over time using macroeconomic variables, in relation to LR.

FIGURE 2 HERE

Discussion and Conclusions

These results demonstrate that survival analysis is competitive in comparison with logistic regression as a credit scoring method for prediction. The inclusion of macroeconomic variables gives a statistically significant improvement in predictive performance. We show that model fit improves significantly and that the direction of the marginal effect on log-hazard rate of most macroeconomic variables is as we would expect. Additionally, Figure 1 indicates that interest rate is the more important macroeconomic variable for estimation of risk of default. This is also as expected. Figure 2 suggests that the survival model with macroeconomic variables is robust over time with a general trend of improvement in performance in relation to the logistic regression model.

In practice, this model can be used for credit scoring by incorporating forecasts of macroeconomic conditions into the assessment of credit card applications. Based on equation (2), the survival probability can be estimated by integrating the hazard rate

incorporating estimates of the macroeconomic time series across the period that default is to be considered. For our experiments this was 12 months. This method of estimation also makes this model suitable for stress testing by including macroeconomic conditions that simulate a depressed or booming economy. This makes it valuable for the implementation of the requirements of the Basel II Accord (eg see Basel II paragraph 415).

Future lines of research will focus on further application of these methods to other credit card and fixed loan products. Also, although the analysis of the explanatory model gives an understanding of how each macroeconomic variable contributes to modelling the data, further extensive experimental work is required to determine the affect of each of the microeconomic variables on the *prediction* of PD.

Acknowledgements

We would like to thank Professors David Hand, Lyn Thomas and other members of the Quantitative Financial Risk Management Centre for discussion of survival analysis and credit scoring during the course of this research. We are grateful for funding through an EPSRC grant EP/D505380/1.

References

Andreeva G (2006). European generic scoring models using survival analysis. *J Opl Res Soc* **57**(10): 1180-1187

Banasik J, Crook JN, Thomas LC (1999). Not if but when will borrowers default. *J Opl Res Soc* **50**: 1185-1190

Baesens B, van Gestel T, Viaene S, Stepanova M, Suykens J and Vanthienen J (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *J Opl Res* **54**: 1082-1088.

Basel II: International Convergence of Capital Measurement and Capital Standards (2006) at www.bis.org/publ/bcbsca.htm

Brambor T, Clark WR, Golder M (2005). Understanding interaction models: improving empirical analyses. *Political Analysis* **14**: 63-82

Collett D (1994). *Modelling Survival Data in Medical Research*. Chapman & Hall.

Hand DJ (2005) Good practice in retail credit scorecard assessment. *J Opl Res Soc* **56**: 1109-1117.

Hosmer Jr. DW and Lemeshow S (1999). *Applied Survival Analysis: regression modelling of time to event data*. Wiley.

Stepanova M and Thomas LC (2001). PHAB scores: proportional hazards analysis behavioural scores. *J Opl Res Soc* **52**: 1007-1016.

Stepanova M and Thomas LC (2002). Survival analysis for personal loan data. *Opl Res* **50**: 277-289.

Tang L, Thomas LC, Thomas S, Bozzetto J-F (2007). It's the economy stupid: modelling financial product purchases. *International Journal of Bank Marketing*. Vol.25, issue 1, pp.22-38.

Thomas LC, Edelman DB and Crook JN (2002). *Credit Scoring and its Applications*. SIAM Monographs on Mathematical Modeling and Computation. SIAM: Philadelphia, USA.

Witten IH and Frank E (2005). *Data Mining*. 2nd ed. Elsevier.

Figure 1. Importance of each macroeconomic variable in Cox PH model.

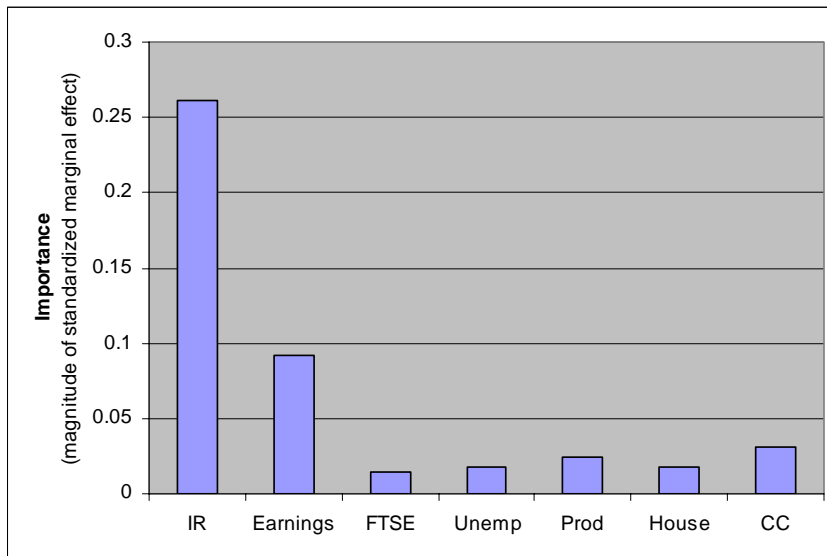
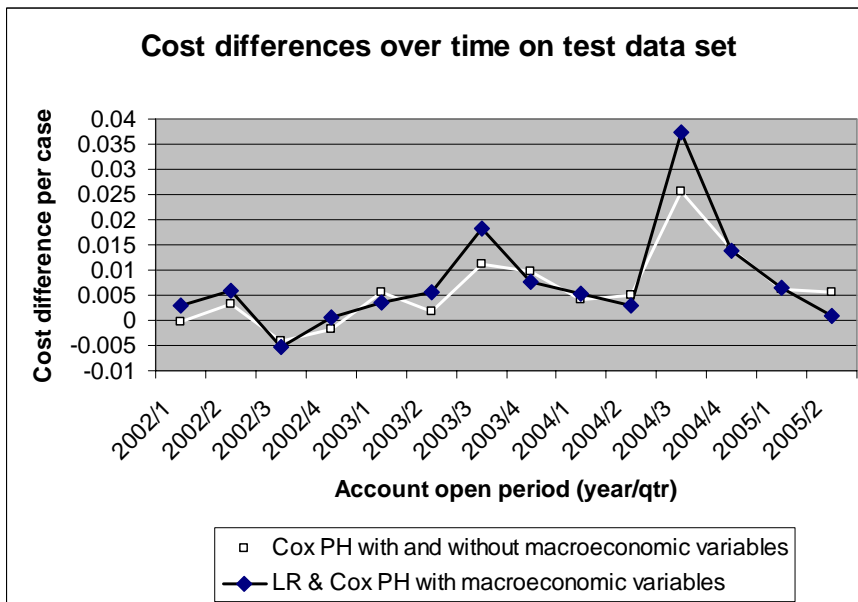


Figure 2. Cost differences between models over time.



Note: cost on bad cases = 20.

Table 1. Macroeconomic variables

<i>Code</i>	<i>Macroeconomic variable</i>	<i>Data source</i>	<i>Expected effect on default risk</i>
IR	Interest rates: Selected UK Retail Banks Base Rate.	ONS	+ve
Earnings	Ratio of UK earnings including bonuses and retail price index on all items, not seasonally adjusted.	ONS	-ve
FTSE	FTSE all-share index.	Publicly available	-ve
Unemp	Unemployment index for males unemployed for 6 to 12 months, seasonally adjusted.	ONS	+ve
Prod	Index of all UK production, not seasonally adjusted.	ONS	-ve
House	House price index.	Nationwide building society	+ve
CC	UK consumer confidence index, not seasonally adjusted.	ONS	+ve

ONS = Office of National Statistics

Table 2. Model fit statistics

<i>Model</i>	<i>LLR</i>	<i>Degrees of freedom</i>	<i>P-value</i>
LR	18396	22	<.0001
Cox PH without TVCs	19387	22	<.0001
Cox PH with macroeconomic variables	21166	47	<.0001
Difference between Cox PH with and without macroeconomic variables	1779	25	<.0001

Table 3. Cox PH model coefficient estimates.

<i>Macroeconomic variables</i>	<i>Coefficient estimate</i>	<i>Chi-sq value</i>	<i>P-value</i>
IR	0.133	373	<.0001
Earnings	-7.08	65	<.0001
FTSE	0.000216	19	<.0001
Unemp	0.0044	9	0.003
Prod	-0.0491	74	<.0001
House	0.821	47	<.0001
CC	0.0331	46	<.0001
<i>Selected interactions</i>	<i>Coefficient estimate</i>	<i>Chi-sq value</i>	<i>P-value</i>
IR * income	-0.00000174	512	<.0001
IR * unknown income (y/n)	-0.107	0	0.9976
Earnings * home owner (y/n)	3.25	13	0.0004
Earnings * private tenant (y/n)	8.59	60	<.0001
Earnings * home council (y/n)	6.63	21	<.0001
FTSE * home owner (y/n)	-0.00025	23	<.0001
FTSE * private tenant (y/n)	-0.00045	50	<.0001
FTSE * home council (y/n)	-0.00029	13	0.0004
Unemp * employed (y/n)	-0.00349	5	0.0191
Unemp * self-employed (y/n)	-0.00687	18	<.0001
Unemp * unemployed (y/n)	0.0955	71	<.0001
Prod * bureau score	0.0000582	81	<.0001
House * home owner (y/n)	-0.721	39	<.0001
House * private tenant (y/n)	-0.839	34	<.0001
House * home council (y/n)	-0.282	2	0.1204
CC * employed (y/n)	-0.0217	19	<.0001
CC * self-employed (y/n)	-0.0457	72	<.0001
CC * unemployed (y/n)	0.0789	1	0.2416

Notes: 1. (y/n) = yes/no indicator variable (yes=1, no=0).

2. Many variables which were included in the models are not reported in the Table for confidentiality reasons.

Table 4. Marginal effects of macroeconomic variables.

<i>Macroeconomic variable</i>	<i>Marginal effect</i>	<i>Expected sign</i>	<i>Standard deviation</i>	<i>Importance</i>
IR	0.34	+ ✓	0.767	0.261
Earnings	-3.56	- ✓	0.0257	0.0917
FTSE	-0.0000347	- ✓	405	0.0141
Unemp	0.0017	+ ✓	10.5	0.0179
Prod	0.00465	- ✗	5.29	0.0246
House	0.136	+ ✓	0.130	0.0177
CC	0.0108	+ ✓	2.90	0.0314

Table 5. Prediction results.

<i>Cost on bad case</i>	<i>Model</i>	<i>Data set</i>	<i>Mean cost</i>
15	LR	Training	0.2135
		Test	0.2364
	Cox PH without TVCs	Training	0.2133
		Test	0.2364
	Cox PH with macroeconomic variables	Training	0.2125
		Test	0.2362
20	LR	Training	0.2824
		Test	0.3083
	Cox PH without TVCs	Training	0.2816
		Test	0.3077
	Cox PH with macroeconomic variables	Training	0.2817
		Test	0.3028
25	LR	Training	0.3500
		Test	0.3728
	Cox PH without TVCs	Training	0.3486
		Test	0.3532
	Cox PH with macroeconomic variables	Training	0.3417
		Test	0.3505

Note: cut-off is computed on the training set.

Table 6. Significance of uplift on test set due to macroeconomic variables.

<i>Costs on models compared</i>	<i>Mean cost difference</i>	<i>N</i>	<i>t</i>	<i>p-value</i>
LR and Cox PH with macroeconomic variables	0.0055	Over 100000	3.89	<.0001
Cox PH with and without macroeconomic TVCs	0.0049	Over 100000	3.74	0.0002

Note: cost on bad cases = 20.

Table 7. Prediction results.

<i>Cost on bad case</i>	<i>Model</i>	<i>Data set</i>	<i>Mean cost</i>
15	LR	Test	0.2348
	Cox PH without TVCs	Test	0.2343
	Cox PH with macroeconomic variables	Test	0.2329
20	LR	Test	0.2985
	Cox PH without TVCs	Test	0.2978
	Cox PH with macroeconomic variables	Test	0.2932
25	LR	Test	0.3524
	Cox PH without TVCs	Test	0.3510
	Cox PH with macroeconomic variables	Test	0.3465

Note: cut-off is computed on the test set.

Table 8. Significance of uplift on test set due to macroeconomic variables.

<i>Costs on models compared</i>	<i>Mean cost difference</i>	<i>N</i>	<i>t</i>	<i>p-value</i>
LR and Cox PH with macroeconomic variables	0.0053	Over 100000	3.45	0.0006
Cox PH with and without macroeconomic TVCs	0.0046	Over 100000	2.84	0.0045

Note: cost on bad cases = 20.

Figure 1. Importance of each macroeconomic variable in Cox PH model.

Figure 2. Cost differences between models over time.

Table 1. Macroeconomic variables

Table 2. Model fit statistics

Table 3. Cox PH model coefficient estimates.

Table 4. Marginal effects of macroeconomic variables.

Table 5. Prediction results.

Table 6. Significance of uplift on test set due to macroeconomic variables.

Table 7. Prediction results.

Table 8. Significance of uplift on test set due to macroeconomic variables.