Sensitivity of stress testing metrics to estimation risk, account behaviour and volatility for credit defaults

Citation for published version:
Djeundje, VB & Crook, J 2022, 'Sensitivity of stress testing metrics to estimation risk, account behaviour and volatility for credit defaults', *Journal of the Operational Research Society*. https://doi.org/10.1080/01605682.2022.2115413

Digital Object Identifier (DOI):
10.1080/01605682.2022.2115413

Published In:
Journal of the Operational Research Society

Journal of the Operational Research Society

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjor20

Sensitivity of stress testing metrics to estimation risk, account behaviour and volatility for credit defaults

Viani Biatat Djeundje & Jonathan Crook

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
https://doi.org/10.1080/01605682.2022.2115413

THE OPERATIONAL RESEARCH SOCIETY
Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Sensitivity of stress testing metrics to estimation risk, account behaviour and volatility for credit defaults

Viani Biatat Djeundje and Jonathan Crook

[a]Credit Research Centre, University of Edinburgh Business School, Edinburgh, UK

ABSTRACT
One approach to stress testing the amount of capital required by a bank for credit risk is to use parameterised account level models with credit application characteristics, behavioural characteristics and macroeconomic factors as predictors. The standard methodology underestimates the amount of capital required because it fails to include uncertainty over the model parameters, over the future trajectory of behavioural variables and over volatility. We provide a methodology for estimating the magnitudes of these additional losses and so a methodology to gain a more accurate estimate of the amount of capital required.

ARTICLE HISTORY
Received 7 June 2021
Accepted 9 August 2022

KEYWORDS
Credit scoring; stress testing; banks; model risk; risk capital

1. Introduction

Every major bank undertakes an annual stress test to indicate whether it has enough capital to protect depositors in the event of a rare but plausible severe negative shock to the value of its assets. These tests are required under the Basel Accords (Pillars I and II) and are implemented by most western banking supervisors including the Federal Reserve, the European Banking Authority and the Bank of England. Major banks also stress test their portfolios to examine the adequacy of their computed economic capital under their internal capital adequacy assessment process (ICAAP). For regulatory capital the national Banking Supervisor supplies a small number of macroeconomic scenarios to a bank and asks it to predict both the amount of capital it will have under these scenarios in a $2-5$ year horizon and its risk weighted assets (RWA) (see BCBS (2006, 2010)). The ratio of the former to the latter (the capital adequacy ratio) must exceed a figure given by the regulator. Typically, statistical methods are used to make the predictions but recently, machine learning methods have been proposed (see Petropoulos et al., 2020, Brummelhuis and Luo (2019), Gogas et al. (2018) and Kupiec (2018)). Specifically, for credit risk, there are at least two broad methodologies that are followed to compute the RWA and so capital required. For regulatory capital, the bank is asked to compute the amount of capital it must hold according to the Vasicek formula (see, for example, Siemsen and Vilsmeier: 2018). The amount is essentially the amount of unexpected losses defined as the difference between the 99.9th percentile of the distribution of expected losses, the Value-at-Risk (VaR) and the mean of that distribution. An alternative methodology is to simulate a large number of scenarios, for example of macroeconomic values for each macroeconomic variable and compute the expected loss distribution and so the VaR from that distribution.

Both methodologies for predicting the VaR required for a portfolio of loan accounts have at least three weaknesses. First, they assume that the parameterised models that are used hold for the population; there is no uncertainty over the parameters. Second, there is no uncertainty over the future values of the behavioural variables which play a vital role in credit scoring models. Third, there is no residual error in the prediction. Clearly omission of these sources of uncertainty results in underestimation of the VaR for a loan portfolio and so of the amount of capital required. This could mean that depositors appear to be adequately protected when in fact they are not. Our article suggests ways to make the scenario approach to stress testing that is used by financial regulators more accurate in terms of the amount of capital needed by banks to protect against plausible but extreme adverse events.

In this article, we contribute to the literature in three ways. First, we give a methodology to assess the sensitivity of the VaR of a loss distribution for a credit card portfolio to the uncertainty over the future values of (a) behavioural variables in a credit scoring model used as inputs to the loss

CONTACT Jonathan Crook j.crook@ed.ac.uk Credit Research Centre, University of Edinburgh Business School, 29 Bucceleuch Place, Edinburgh EH8 9JS, UK

distribution, (b) the parameters of the credit scoring model and (c) residual values of the scoring model. Second we show the relative sizes of these amounts of uncertainty for a portfolio of credit card accounts. Third we show the sensitivity of the amount of capital required relative to the mean that is implied by each type of risk in terms of the VaR and expected shortfall values. We find that the underestimation of the amount of capital required is substantial and therefore that regulators may need to consider seeking estimates of these values from banks in annual stress tests.

There is a large literature on sensitivity analysis (see Borgonovo and Plischke (2016) for a review of methods and Pesenti et al. for an alternative approach). We are interested in measuring the sensitivity of VaR to each of the risks we have outlined. Considering each of the specific sources of risk, and starting with that associated with future values of behavioural variables, we are unaware of any papers that evaluate the implications of uncertainty over their future values in a credit scoring model for metrics of the predicted loss distribution.

In contrast, there is a substantial literature on model parameter estimation risk. Two streams of literature are most closely related to this article.[1] One stream looks at methods to reduce model uncertainty. They include Bayesian model averaging (BMA) (Henry & Kok, 2013; Siemsen & Vilsmeier, 2018) and BMA and ridge regression (Hofmarcher et al. (2014)). The second stream of literature considers methods for evaluating the implications of different sources of prediction uncertainty. Misina and Tessier (2008) consider a model relating industry sector default rates to macroeconomic variables. Using a Cholesky decomposition to retain the correlations between the macroeconomic variables with simulated values of the macroeconomic variables and error terms, they illustrate the differences between the loss distribution if the model is linear compared with non-linear, thus arguing uncertainty over the correct form can have major implications for the VaR. Jacobs et al. (2015) proposed Bayesian regression to take into account parameter uncertainty. But, their model took aggregate banking sector charge-off rates as the measure of loss rather than using account level data and included only three macroeconomic variables. Bignozzi and Tsanakas (2016) consider model uncertainty, parameter uncertainty and residual risk. They show that a Bayesian approach with a non-informative prior removes residual estimation risk if the model selected is correct. Using different Bayesian approaches, for example choose the model with the highest posterior weight, they argue the amount of residual estimation risk indicates the amount of model risk. Wang et al. (2020) estimated Bayesian models of PD distributions at account level using application, behavioural and macroeconomic variables and compared the implied VaR with that from a frequentist model. Garcia-Cespedes and Moreno (2016) suggest uncertainty over probability of default (PD) estimation can be computed from the variance of the estimates and they illustrate using estimates of the PD in different corporate credit ratings. They find that the randomness of PD, rather than the fixed value in the Basel Accord, increases the probabilities of extreme losses but the VaR increases by much less than the uncertainty over the PD. A further paper that considers uncertainty over behavioural model parameters is by Bakoben et al. (2020) who propose a method to cluster accounts by vector autoregression (VAR) parameters of behavioural equations taking into account uncertainty as to their values in the clustering algorithm. They then estimate cross sectional scoring models using cluster membership or the degree of cluster membership uncertainty as the covariate, but they do not consider the implications of parameter uncertainty for the future expected loss distribution or VaR and their application is relatively simple.

The closest paper to ours is Gross and Poblacion (2019). They decompose prediction uncertainty about required Common Equity Tier 1 (CET1), at the level of the bank, using data for 75 banks across 16 countries into model uncertainty (34%), coefficient uncertainty (26%) and residual uncertainty (40%). However Gross and Poblacion consider only bank level data rather than account level data and do not consider a stress test of a credit scoring model. They also do not consider the implications of stressing behavioural variables which are important at account level but are not normally included at bank level.

None of the papers in the literature give a methodology for measuring the implications for the VaR of a portfolio of the three types of risk in a credit scoring model that we are interested in here: behavioural variable value risk, parameter uncertainty and residual uncertainty. Whilst there are several methods for estimating parameter uncertainty they have not been applied to credit scoring models to predict the amount of capital necessary to cover this risk.[2] Importantly, as far as we are aware, there is no paper that decomposes prediction uncertainty for a credit scoring model into different sources and shows the contribution of each to VaR. That is the aim of our article.

Section 2 describes the baseline hazard model that forms the credit scoring model we will be assuming. Section 3 explains the methodology: the use of the Poisson-Binomial distribution to measure volatility risk, the use of multiple generated trajectories using account level vector auto-regressions to gain a measure of the risk of not knowing future values of the behavioural variables, and the distribution of the maximum likelihood estimator of the parameters to measure the risk of mis-estimation. Section 4 shows an implementation for a credit card portfolio and Section 5 concludes.

2. Dynamic models for credit risk data

Let us consider a portfolio of n credit card accounts, and denote by $q_{i,\tau}$ the default probability for account i at time τ (from the opening date), assuming that this account is still active just prior to this time. In this article, we assume that time is measured in months. The probabilities $q_{i,\tau}$ are driven by a number of factors including application variables, behavioural variables and macroeconomic variables. A standard way to quantify this dependence is via a discrete-time survival model (also refereed to as a dynamic model) as follows

$$g(q_{i,\tau}) = h_{0,\tau} + \boldsymbol{U}_i\ \boldsymbol{\alpha} + \boldsymbol{Z}_{i,\tau}\ \boldsymbol{\beta}, \quad i = 1, \ \cdots, \ n\cdot \tag{1}$$

In this expression, $\boldsymbol{U}_i$ is a one-row matrix of time-independent variables associated with account i (these are often refereed to as application variables), and $\boldsymbol{Z}_{i,\bullet}$ is the one-row matrix of dynamic or time-dependent variables on account i. These time-dependent variables often include variables measuring the behaviour of the accounts (behavioural variables), as well as macro-economic variables. Interactions between variables can also be appended to the $\boldsymbol{U}_i$'s and $\boldsymbol{Z}_{i,\bullet}$'s. To complete the definition, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are unknown coefficients to be estimated, $h_{0,\bullet}$ is a baseline function and g is the logit link function,.

In practice, the baseline function is often smoothed and this allows one to reduce irregular fluctuations from one month to the next. Various smoothing methods can be used. Typical examples include the use of polynomial-type functions (Bellotti & Crook, 2013) as well as spline basis functions (Djeundje & Crook, 2018); our preferred method is to use B-splines basis functions (De Boor, 1978). The foundation and attractiveness of this approach are discussed by Eilers and Marx (1996); more recent exposures of the benefit of this approach in the credit risk context can be found in Djeundje and Crook (2018, 2019a).

Using B-splines, the baseline function can be expressed as $h_{0,\tau} = \mathcal{B}(\tau)\boldsymbol{a}$, where $\mathcal{B}(\tau)$ is a one-row matrix of B-splines at time τ, and $\boldsymbol{a}$ is a vector of unknown coefficients to be estimated. In practice, using too many B-splines leads to over-fitting. Conversely, using too few splines results in poor fit. An attractive way of solving this dilemma is to allow a rich set of B-splines and then apply a penalty on adjacent elements of the coefficient vector $\boldsymbol{a}$ to achieve smoothing (O'Sullivan, 1986; Eilers & Marx, 1996). We prefer this method in this article.

Within credit risk, dynamic models of the form (1) have been investigated in the recent literature, starting with sub-models that include application or behavioural variables only, through to larger models involving macroeconomic variables (Djeundje & Crook, 2019a, 2019b). Various extensions have also been considered and implemented successfully including varying coefficients models (Djeundje & Crook, 2019a) and multi-state intensity models, some with random effects to account for heterogeneity (Lando & Skødeberg, 2002; Djeundje & Crook, 2018).

In this article, we show and describe how the framework of discrete survival models can be used to assess and quantify the impact of three sources of uncertainty when stress testing credit defaults.

3. Mis-estimation, volatility. and behavioural risks

Consider a portfolio of credit card accounts over multiple years. We denote by t_o the current calendar time. Thus, we assume that data is observed only up to $t_o\cdot$ These data include application, behavioural and macroeconomic variables, as well as the survival times of the accounts that have already defaulted. We are interested in constructing and investigating the distribution of the loss random variable over a set future time horizon of length δ; for example, one month, six months, one year, etc.

Consider an account i still active at current calendar time t_o. The loss associated with this account over the future horizon $(t_o, \ t_o + \delta]$ can be expressed as

$$\mathrm{EAD}_{i,(0,\delta]} \times \mathrm{LGD}_{i,(0,\delta]} \times y_{i,(0,\delta]} \tag{2}$$

where $\mathrm{EAD}_{i,(0,\delta]}$ and $\mathrm{LGD}_{i,(0,\delta]}$ are the exposure and loss given default associated with account i during the future horizon $(t_o, \ t_o + \delta]$, and $y_{i,(0,\delta]}$ is an indicator variable taking value 1 if account i defaults during the future horizon $(t_o, \ t_o + \delta]$, and 0 otherwise.

Our prime interest in this work is to assess the impact of a number of sources of uncertainty involved in the estimation of the risk of default. Thus, if one assumes a 100% LGD and constant EAD for all accounts as in Bellotti and Crook (2013), the aggregated loss (relative to the constant EAD) associated with the future time horizon $(t_o, \ t_o + \delta]$ is obtained as

$$\mathcal{L}_{(0,\delta]} = \sum_{i \in \mathcal{R}_o} y_{i,(0,\delta]}, \tag{3}$$

where $\mathcal{R}_o$ represents the set of accounts still active at the beginning of the future horizon (i.e., the current calendar time t_o).

If we can generate a large sample from the entire distribution of the aggregated losses, then economic capital can be derived as VaR less the mean (or ES less the mean). For example, if $\boldsymbol{L} = \{l_1, \ldots, l_N\}$ is a reasonable sample generated from the loss distribution, the $\alpha\%$-level VaR relative to the expected loss, which we denote by VaR_α, can be estimated by

$$\mathrm{VaR}_\alpha = \frac{\alpha^{\text{th}} \text{ percentile of } \boldsymbol{L}}{\text{mean of } \boldsymbol{L}} \tag{4}$$

Before we proceed, it is imperative to first identify the main components or contributors of the loss distribution. Let us denote by $p_{i,(0,\delta]}$ the probability that account i defaults during the future time horizon $(t_o,\ t_o+\delta]$, given that this account is still active at the opening time t_o. That is,

$$p_{i,(0,\delta]} = 1 - \prod_{k=1}^{\tau}(1 - \tilde{q}_{i,t_o+k}) \tag{5}$$

where $\tilde{q}_{i,t_o+k}$ are the conditional monthly default probabilities for account i at future calendar time $t_o + k$. The expected value and variance of the aggregated loss are given by

$$\mu_{(0,\delta]} = \sum_{i \in \mathcal{R}_o} p_{i,(0,\delta]} \quad \text{and}$$
$$\sigma^2_{(0,\delta]} = \sum_{i \in \mathcal{R}_o} p_{i,(0,\delta]}(1 - p_{i,(0,\delta]}) \tag{6}$$

Hence, the structure of the loss distribution is driven by the default probabilities $\tilde{q}_{i,t_o+k}$ over the future horizon $(t_o,\ t_o+\delta]$, in conjunction with the sample volatility or noise due to the composition of the portfolio being analysed. Note that the true values of these default probabilities $\tilde{q}_{i,t_o+k}$ are unknown themselves, but they can be predicted by applying the fitted parameters from model (1) to the corresponding variables over the future horizon. Hence, a number of components contributing to the distribution of the credit default loss can be identified, including *model risk, behavioural risk, macroeconomic risk, mis-estimation risk, volatility risk*, etc.

We focus on the measurement and comparison of behavioural, mis-estimation and volatility risks only. By behavioural risk we mean the risk of not knowing future values of behavioural variables. By mis-estimation risk we mean the risk of not knowing the process generating the population values of the parameters. By volatility risk we mean the risk of not knowing if an account will actually default even if we can predict the probability of default perfectly.

In particular in Section 3.1, we show how to quantify the impact of volatility via the Poison-Binomial distribution. In Section 3.2, we use the distribution of the maximum likelihood estimator of the parameters to measure mis-estimation. In Section 3.3, we measure behavioural risk through account-level vector auto-regressions. Section 4 then illustrates, in terms of stress metrics, the magnitude of different combinations of these three types of risk within a global sensitivity analysis framework (Borgonovo & Plischke, 2016).

3.1. Volatility risk and quantification

The aggregated loss $\mathcal{L}_{(0,\delta]}$ is a random variable consisting of a sum of Bernoulli random variables with not-all-equal individual success probabilities, $p_{i,(0,\delta]}$, $i = 1, \ldots, n$. Thus, conditional on the values of the cumulated probabilities $p_{i,(0,\delta]}$, the aggregated loss variable $\mathcal{L}_{(0,\delta]}$ follows the so-called Poisson-Binomial (PB) distribution which we denote by

$$\mathcal{L}_{(0,\delta]} \sim \mathcal{PB}\{p_{i,(0,\delta]},\ i \in \mathcal{R}_o\} \tag{7}$$

Hence, for a credible risk management strategy, capital must be held against the fact that each portfolio is made up of many individuals' accounts, irrespective of whether the true values of the default probabilities for these accounts are known or not.

We propose to quantify the contribution of the portfolio-volatility on stress metrics (such as VaR or ES) based on the PB distribution (7). Specifically, we simulate prospective losses using (7) conditional on the individual probabilities $p_{i,(0,\delta]}$ and then use these simulated losses to estimate the stress metrics. More details on this are given in given in Sections 4.4 and 4.5.

The PB distribution itself is not new. It was first considered by (Neyman, 1939) and an early application was by (Barrett, 2007) in the context of criminal and civil courts. But more recent investigations of the properties of this distribution can be found in Daskalakis et al. (2015), Hong (2013), or Chen and Liu (1997). In particular, Daskalakis et al. (2015) explore the sample complexity of learning from PB distributions, while the other two papers explore different methods of computing and sampling for PB distributions. In this work, we sample from any given PB distribution based on the Fourier transform of its characteristic function as in Hong (2013).

3.2. Mis-estimation risk and quantification

In practice, for each account still active at current calendar time t_o, its conditional monthly default probabilities over the future horizon $(t_o, t_o+\delta]$ are unknown. They must be predicted using the fitted model. Hence uncertainty exists over the portfolio's underlying default rates, since these can only be estimated to a degree of confidence linked to the size and richness of the data. Part of this uncertainty is reflected in the estimation process of the model parameters $\boldsymbol{a}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (where $\boldsymbol{a}$ represents the vector of parameters used to model the baseline function $h_{0,\bullet}$).

For notational convenience, let us denote by $\boldsymbol{\theta}$ the joint vector of all parameters. That is,

$$\boldsymbol{\theta} := [\boldsymbol{a}^T, \boldsymbol{\alpha}^T, \boldsymbol{\beta}^T]^T, \tag{8}$$

Also, let $\hat{\boldsymbol{\theta}}$ denote the maximum likelihood estimator of $\boldsymbol{\theta}$ and $\mathcal{H}(\hat{\boldsymbol{\theta}})$ denotes the Hessian matrix evaluated at the maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$. It is well known from the theory of MLE that $\hat{\boldsymbol{\theta}}$ is efficient and asymptotically normal:

$$\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \ -\mathcal{H}(\hat{\boldsymbol{\theta}})^{-1}) \tag{9}$$

We can use expression (9) with MLE plugged in, to generate plausible alternative values of the parameters of the model, which can be used in turn to calculate alternative plausible default probabilities and to quantify the amount of capital that should be held against the risk of mis-estimation. More details on this are provided in Sections 4.3 and 4.5. A similar approach was used in the context of longevity risk insurance by Richards (2016) to quantify the cost of mis-estimation for capital adequacy and solvency.

3.3. Behavioural risk and quantification

Accounts are assumed to be observed from opening date up to current calendar time t_o only. Thus, the values of the behavioural variables beyond t_o are unknown. In order to generate plausible values of default probabilities $p_{i,(0,\delta]}$, we first need to generate plausible monthly values of the underlying behavioural variables (for each account still active at time t_o) over the future horizon $(t_o, t_o + \delta]$.

One possibility is to consider each behavioural variable in isolation and then sample plausible future values based on past observed values, separately for each account. However, behavioural variables are potentially correlated with one another. Thus, it is unrealistic to generate plausible future values of each behavioural variable in isolation.

In this work, we propose to generate plausible future values of the behavioural variables through account-level vector-auto-regressive processes. Hence, for a given account i still active at the calendar time t_o, we consider models of the form

$$\boldsymbol{Z}_{i,\tau} = \boldsymbol{\alpha}_{o,i} + \sum_{k=1}^{p} \boldsymbol{\Theta}_{i,k}\ \boldsymbol{Z}_{i,\tau-k} + \boldsymbol{\varepsilon}_{i,\tau}, \quad \boldsymbol{\varepsilon}_{i,\tau} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_i), \tag{10}$$

where $\boldsymbol{Z}_{i,\tau}$ represents the vector of behavioural processes associated with account i at duration time τ, $\boldsymbol{\alpha}_{o,i}$ is the vector of intercepts, p is the lag, $\boldsymbol{\Theta}_{i,k}$ is the matrix of coefficients associated with lagged values of these processes, $\boldsymbol{\varepsilon}_{i,\tau}$ is the vector of error terms, and $\boldsymbol{\Sigma}_i$ is the covariance matrix $\boldsymbol{\Sigma}_i$.

In particular, if $p = 1$ and elements within $\boldsymbol{\Theta}_{i,1}$ are set to the identity matrix, then Equation (10) simplifies to a multivariate random walk with drift vector $\boldsymbol{\alpha}_{o,i}$, in which case the correlation between the behavioural processes are measured through the correlation matrix $\boldsymbol{\Sigma}_i$ alone. Similarly, if the $\boldsymbol{\Theta}_{i,k}$ are diagonal, then Equation (10) simplifies to a set of correlated or uncorrelated auto-regressive processes, depending of the structure imposed on $\boldsymbol{\Sigma}_i$. But in general, Equation (10) allows each behavioural process to affect the prediction of its counterparts through the coefficients matrices $\boldsymbol{\Theta}_{i,k}$ and covariance matrix $\boldsymbol{\Sigma}_i$.

Models of the form of (10) can be fitted using observed data up to calendar time t_o. In this article, these models are fitted separately for each account via the method of least squares. The fitted parameters are then used to simulate plausible values of the behavioural processes for each account over the future horizon $(t_o, t_o + \tau]$. These simulated values are used subsequently in conjunction with the regression coefficients estimated from model (1) to generate prospective default probabilities and to construct the loss distribution. Details about the exact forms of models (10) implemented in this article are presented in Section 4 below.

4. Data, implementation and results

4.1. Data and base survival model

The data that motivated this work is a sample of about 50,000 credit card accounts from a major bank in the UK. The accounts in the sample were opened from 2002 through to 2011.

This dataset comprises a number of application variables as well as behavioural variables; the list of variables used in this analysis is summarised in the first column of Table 1. In particular, four behavioural variables are used in this investigation, including the Credit limit (denoted by CL), the Repayment amount (denoted by PAY), the percentage of time spent with one outstanding payment (denoted by DEL), and the Proportion of the credit limit that is drawn (denote by PDR). An illustration of three of these behavioural variables for a typical account is shown in Figure 1.

We define an account as being in default if and when it became three months in arrears. The three missed payments are not required to have been missed consecutively. To ensure a consistent definition over time we computed a minimum payment each month using constant parameters over time. This definition is not the same as that used in practice by the data provider.

Table 1. Fitted coefficients for the base survival model. Interaction terms have been omitted. The variable *Number of cards* measures the number of cards that the applicant already has at the time of application. Names of some variables cannot be revealed for commercial confidentiality reasons.

Variables' type	Variables' name/categories	Coefficient	S.E.	p-val
Application variables	*Employment: employed*	0		
	Employment: self-employed	−0.029	0.028	0.294
	Employment: retired or unemployed	0.002	0.060	0.971
	Employment: students	0.352	0.036	0.000
	Employment: not given	−0.153	0.028	0.000
	Number of cards: "0 card"	0		
	Number of cards: "1 card"	−0.048	0.022	0.031
	Number of cards: "2 to 5 cards"	0.146	0.023	0.000
	Number of cards: " >5 cards"	0.307	0.077	0.000
	Variable X: group B	−0.011	0.029	0.704
	Variable X: group C	0.068	0.032	0.032
	Variable X: group D	−0.042	0.032	0.184
	Variable X: group E	0.019	0.035	0.590
	Age at application: $\leq$ 22 *year old*	0		
	Age at application: (22 − 27)	−0.070	0.033	0.035
	Age at application: (27 − 32)	−0.083	0.037	0.026
	Age at application: (32 − 37)	−0.146	0.040	0.000
	Age at application: (37 − 42)	−0.125	0.042	0.003
	Age at application: (42 − 47)	−0.177	0.044	0.000
	Age at application: (47 − 52)	−0.221	0.048	0.000
	Age at application: (52 − 57)	−0.327	0.054	0.000
	Age at application: (57 − 62)	−0.388	0.063	0.000
	Age at application: $\geq$ 62 *year old*	−0.562	0.069	0.000
	Adjusted credit limit (CL)	0.586	0.012	0.000
Behavioural variables	*Adjusted delinquency index (DEL)*	1.757	0.034	0.000
	Adjusted repayment amount (PAY)	−0.227	0.004	0.000
	Adjusted proportion drawn (PDR)	4.263	0.028	0.000

For illustration, we set the current calendar date t_o to December 2010. Thus, the base survival model (1) is fitted using data available from 2002 through to December 2010. The accounts still active at the end of December 2010 form the stress-testing set. We can use this stress-testing set to quantify the impact of different sources of uncertainty, going forward. Various future horizons can be considered; the illustrations in this article correspond to a twelve months future horizon (i.e., $\delta = 12$).

The starting point is to fit the base survival model (1). The fitted coefficients are shown in Table 1. The coefficients have the expected signs and are consistent with the literature. Thus, holding other variables constant, the higher the credit limit, the greater the percentage of duration time the account has been one payment behind and the higher the proportion of the limit that is drawn the higher the PD. Also the higher is the repayment amount the lower the PD. Older account holders are seen to have lower PDs and those with more cards of types C and D have a higher PD. More discussion of the magnitude and significance of most of these coefficients can be found in Djeundje and Crook (2018, 2019a). Our aim in this article is the quantification of uncertainty.

In the next few sections, we shall focus on the assessment of the three sources of uncertainty described in Section 3. We use four scenarios to investigate and compare the impact of these sources over the future time-period $(\texttt{Dec2010} - \texttt{Dec2011}]$. As we shall see, each source can be switched on/off when estimating the aggregated loss distribution.

4.2. Scenario1: Behavioural risk only

In this first scenario, we consider behavioural risk alone. To quantify its impact, we generate $N = 5000$ plausible trajectories for each behavioural variable and for each account (in the stress-testing set) over the future twelve months (i.e., January to December 2011). These generated behavioural trajectories are then used in conjunction with the MLE of the regression coefficients from the base survival model to estimate plausible monthly PDs (over the future horizon) and subsequently N candidates for the cumulative probabilities $p_{i,(Dec2010, Dec2011]}$ for each account. The application variables were set at their observed values in the data. A sample of N plausible aggregated losses can then be derived using Equation (6), and used to construct the empirical loss distribution, as well as to estimate the VaR according to Formula (4).

To generate future trajectories of the behavioural variables, we consider three alternative variants of vector autoregressive models for the behavioural processes.

First, we consider multivariate random walks with drift terms, which we outline as follows

$$\begin{cases} \mathrm{CL}_{i,\tau} = \alpha_{i,1} + \mathrm{CL}_{i,\tau-1} + \varepsilon_{i,\tau,1} \\ \mathrm{PAY}_{i,\tau} = \alpha_{i,2} + \mathrm{PAY}_{i,\tau-1} + \varepsilon_{i,\tau,2} \\ \mathrm{DEL}_{i,\tau} = \alpha_{i,3} + \mathrm{DEL}_{i,\tau-1} + \varepsilon_{i,\tau,3} \\ \mathrm{PDR}_{i,\tau} = \alpha_{i,4} + \mathrm{PDR}_{i,\tau-1} + \varepsilon_{i,\tau,4}. \end{cases} \tag{11}$$

This is a very simplified version of model (10), where $\boldsymbol{Z}_{i,\tau} = (\mathrm{CL}_{i,\tau}, \mathrm{PAY}_{i,\tau}, \mathrm{DEL}_{i,\tau}, \mathrm{PDR}_{i,\tau})^T$, $\boldsymbol{\alpha} = (\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3}, \alpha_{i,4})^T$, $\boldsymbol{\Theta}_{i,1}$ is the identity matrix,

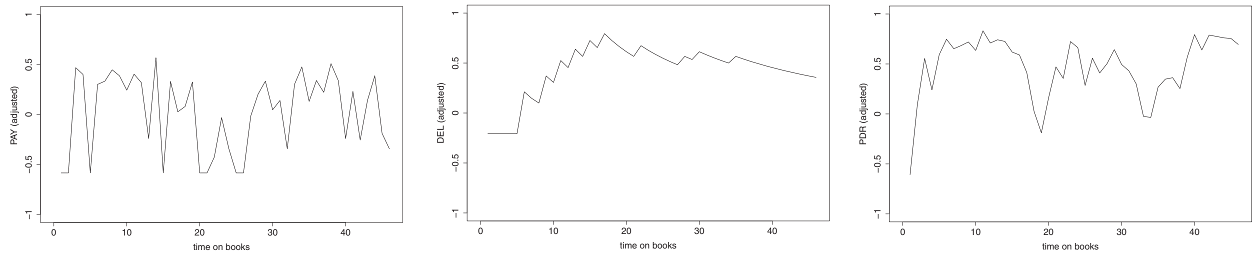

Figure 1. Monthly behavioural data for a typical account from opening date through to December 2010.

and $\boldsymbol{\varepsilon}_{i,\tau} = (\varepsilon_{i,\tau,1},\ \varepsilon_{i,\tau,2},\ \varepsilon_{i,\tau,3},\ \varepsilon_{i,\tau,4})$. In this reduced model, the correlation between the behavioral variables is captured by the (4×4) covariance matrices $\boldsymbol{\Sigma}_{i,\tau} = covar(\boldsymbol{\varepsilon}_{i,\tau})$.

For each account, the drift vector $\boldsymbol{\alpha}$ can be estimated using the method of least squares, and an estimate of the covariance matrix $\boldsymbol{\Sigma}_{i,\tau}$ derived from the residuals. For each account we simulate by choosing values of $\boldsymbol{\varepsilon}_{i,\tau} = (\varepsilon_{i,\tau,1},\ \varepsilon_{i,\tau,2},\ \varepsilon_{i,\tau,3},\ \varepsilon_{i,\tau,4})$ randomly to predict each behavioural variable one month ahead. These predicted values are used as inputs to generate future predictions. We follow this simulation method in all of the following models.

Second, we consider correlated autoregressive processes for the behavioural processes as follows

$$\begin{cases} \mathrm{CL}_{i,\tau} &= \alpha_{i,1} + \beta_{i,1}\ \mathrm{CL}_{i,\tau-1} + \varepsilon_{i,\tau,1} \\ \mathrm{PAY}_{i,\tau} &= \alpha_{i,2} + \beta_{i,2}\ \mathrm{PAY}_{i,\tau-1} + \varepsilon_{i,\tau,2} \\ \mathrm{DEL}_{i,\tau} &= \alpha_{i,3} + \beta_{i,3}\ \mathrm{DEL}_{i,\tau-1} + \varepsilon_{i,\tau,3} \\ \mathrm{PDR}_{i,\tau} &= \alpha_{i,4} + \beta_{i,4}\ \mathrm{PDR}_{i,\tau-1} + \varepsilon_{i,\tau,4} \end{cases} \tag{12}$$

In this case, the matrices of coefficients $\boldsymbol{\Theta}_{i,1}$ are diagonal with unknown parameters to be estimated on the diagonal. The interaction between the individual behavioural processes are captured by the diagonal elements of $\boldsymbol{\Theta}_{i,1}$ and the covariance matrix $\boldsymbol{V}_i$. An illustration of simulated paths by applying this process to a typical account is shown in Figure 2.

Third, we consider a full vector autoregressive model of order one. That is

$$\begin{cases} \mathrm{CL}_{i,\tau} &= \alpha_{i,1} + \beta_{i,1,1}\ \mathrm{CL}_{i,\tau-1} + \beta_{i,1,2}\ \mathrm{PAY}_{i,\tau-1} + \beta_{i,1,3}\ \mathrm{DEL}_{i,\tau-1} + \beta_{i,1,4}\ \mathrm{PDR}_{i,\tau-1} + \varepsilon_{i,\tau,1} \\ \mathrm{PAY}_{i,\tau} &= \alpha_{i,2} + \beta_{i,2,1}\ \mathrm{CL}_{i,\tau-1} + \beta_{i,2,2}\ \mathrm{PAY}_{i,\tau-1} + \beta_{i,2,3}\ \mathrm{DEL}_{i,\tau-1} + \beta_{i,2,4}\ \mathrm{PDR}_{i,\tau-1} + \varepsilon_{i,\tau,2} \\ \mathrm{DEL}_{i,\tau} &= \alpha_{i,3} + \beta_{i,3,1}\ \mathrm{CL}_{i,\tau-1} + \beta_{i,3,2}\ \mathrm{PAY}_{i,\tau-1} + \beta_{i,3,3}\ \mathrm{DEL}_{i,\tau-1} + \beta_{i,3,4}\ \mathrm{PDR}_{i,\tau-1} + \varepsilon_{i,\tau,3} \\ \mathrm{PDR}_{i,\tau} &= \alpha_{i,4} + \beta_{i,4,1}\ \mathrm{CL}_{i,\tau-1} + \beta_{i,4,2}\ \mathrm{PAY}_{i,\tau-1} + \beta_{i,4,3}\ \mathrm{DEL}_{i,\tau-1} + \beta_{i,4,4}\ \mathrm{PDR}_{i,\tau-1} + \varepsilon_{i,\tau,4} \end{cases} \tag{13}$$

In this case, $\boldsymbol{\Theta}_{i,1}$ is a full 4×4 matrix filled with the β's, and the structure of the interactions between behavioural processes is more complex and flexible compared to models (11) and (12).

We quantify the size of the behavioural risk in terms of the VaR and ES relative to the mean of the predicted losses. The latter are the aggregated values of those losses predicted by the base hazard model, (1), where the observed values of the application variables and the predicted values of the behavioural variables are substituted in.

Among the three model structures (11), (12), and (13) for the behavioural processes, the multivariate random walks are the easiest to implement. However, this structure can struggle to capture certain important features seen in the data. In contrast, the full vector autoregressive process is more flexible and can adjust to capture more complex associations between the behavioural processes. However, it can become computationally intensive as the number of accounts increases. High order vector autoregressive processes were considered but were found to be too complex for the vast majority of accounts, especially those with a small number of observations or low variability.

An illustration of the aggregated loss (corresponding to a 12 months horizon) subject to behavioural risk alone is shown in Figure 3. To create this graphic, plausible trajectories of the underlying behavioural processes over the future horizon were jointly simulated based on the correlated autoregressive processes in Equation (12). Similar distributions can be constructed based on correlated random walk models (11) or the full vector autoregressive models (13).

A comparison of the VaR and ES relative to the mean when using these three processes to quantify behavioural risk is shown in Table 2. This highlights the sensitivity of risk measures and economic capital with respect to the choice of the behavioural process models. In particular, modelling behavioural

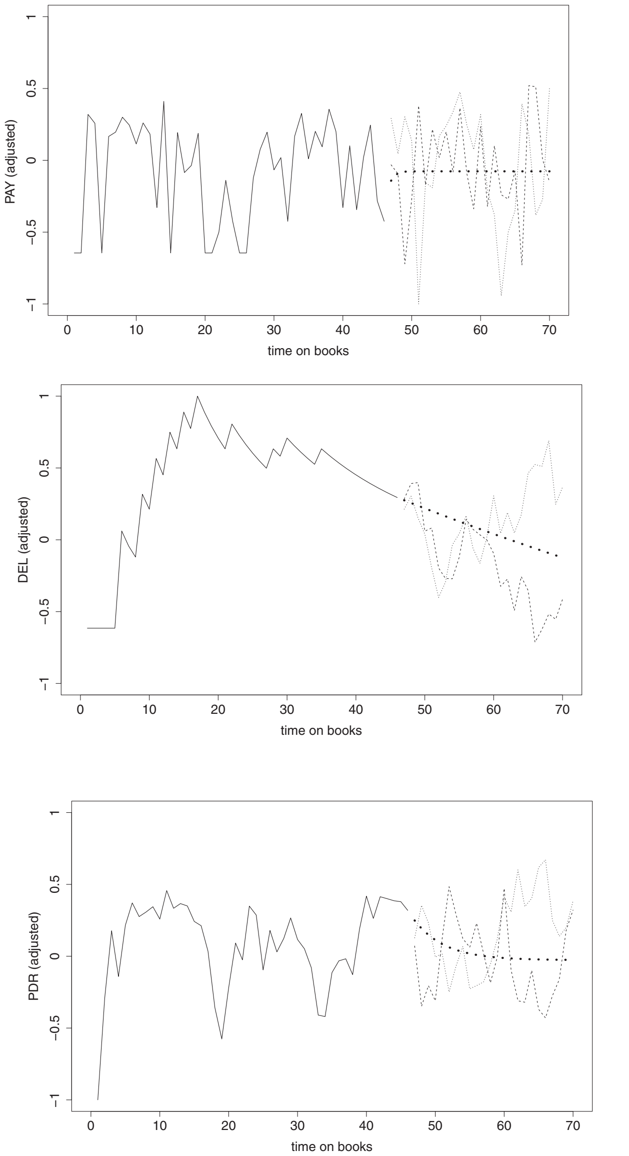

Figure 2. Two simulated future paths for a typical account using correlated autoregressive processes (12). The dotted line is the central/expected forecast.

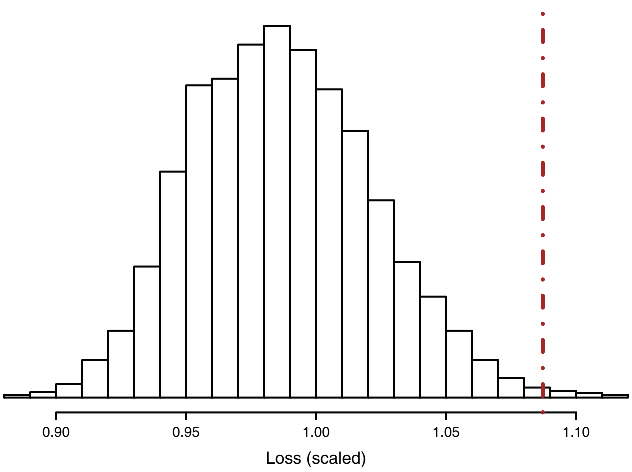

Figure 3. Distribution of the aggregated loss (relative to the mean) subject to behavioural risk alone over twelve months ($\delta = 12$); the dashed line represents the 99.5^{th} percentile. Plausible values of the underlying behavioural processes over the future horizon were jointly simulated based on correlated auto-regressive processes; see Equation (12).

Table 2. Assessment of behavioural risk in terms of VaR and ES over a 12-months horizon relative to the mean. The $VaR_{99.5}$ values are estimated based on a sample of 5000 losses, and the the loss samples were computed based on simulated values of the behavioural processes, in conjunction with the MLE of the vector of regression parameters $\boldsymbol{\theta}$. The underlying simulated behavioural processes over the 12 months horizon were generated using three different forms of vector auto-regressive processes; see Equations (11)(12) and (13).

Model for behavioural processes:	$VaR_{99.5}$	$ES_{99.5}$
Multivariate random walk	113.6%	115.9%
Correlated auto-regressive processes	108.3%	109.3%
First order vector auto-regressive process	108.7%	110.1%

patterns using the simplest multivariate random walk (11) leads to more uncertainty compared to the other two models. This is expected since the second and third processes (12) and (13) are extensions of (11), and so would be expected to model the data more accurately. What is especially interesting is that the VaR and ES of the autocorrelated process (12) and the VAR model (13) are very similar; there is little benefit from using a VAR model (13).

4.3. Scenario2: Mis-estimation risk only

Under this scenario, we consider the impact of mis-estimation risk alone as follows.

(i) Upon fitting the base survival model (1), the Hessian matrix is estimated and used to simulate N alternative plausible vectors for the regression parameters $\boldsymbol{\theta}$ via distribution (9).
(ii) The behavioural processes are jointly forecasted over twelve months using correlated auto-regressive models (12) at account level.
(iii) Each simulated vector of regression parameters from (i) is applied (to the application variables and central forecasts of the behavioural variable from (ii)) yielding a sample of N monthly default probabilities over 12 months for each account. Finally, these monthly probabilities are used to compute the cumulated probabilities $p_{i,(Jan2011,Dec2011]}$ and to generate a sample of N potential loses.

An illustration of the negative inverse of the Hessian matrix (i.e., the covariance matrix of $\hat{\boldsymbol{\theta}}$) is shown in Figure 4. We find the values of VaR and ES required for mis-estimation risk are 6.5% and 7.3% of the expected values, respectively. These are a little less than those due to behavioural risk.

4.4. Scenario3: volatility risk only

Under this scenario, we consider the impact of volatility risk alone. To generate the cumulated

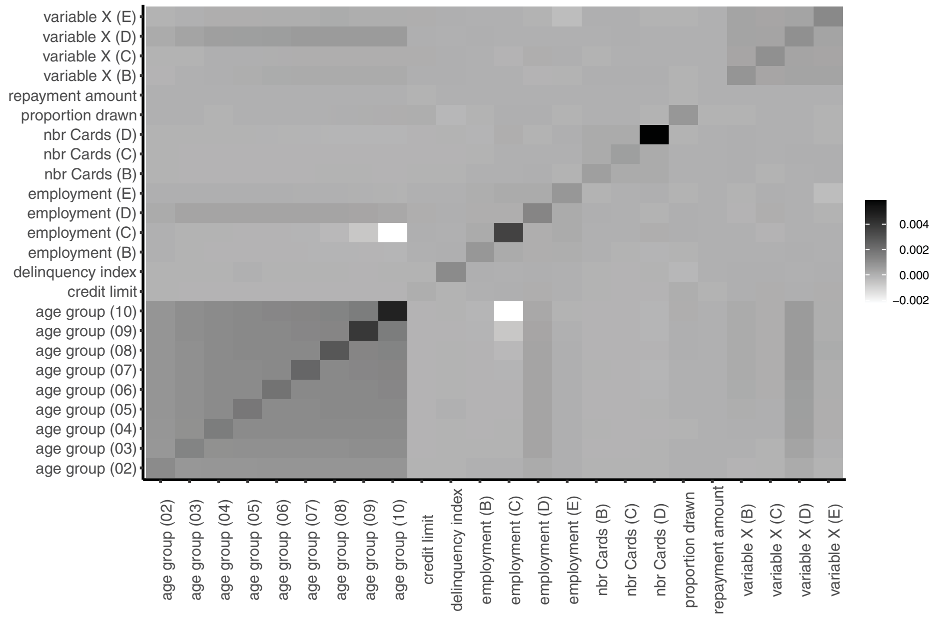

Figure 4. Covariance matrix of the fitted regression coefficients from the survival model.

Table 3. Assessment of three sources of uncertainty in terms of VaR and ES over a 12-months horizon. The $VaR_{99.5}$ are estimated based on a sample of 5000 losses, and the $ES_{99.5}$ are estimated as the average of simulated losses above $VaR_{99.5}$. In this table, the underlying projected behavioural processes were carried out using correlated auto-regressive models separately for each account; we use this model form for illustration because an early investigation (e.g., Table 2) showed that, for our dataset, this model form produces stress metrics that are similar to those from full VAR(1) models.

Source of uncertainty	$VaR_{99.5}$	$ES_{99.5}$
Behavioral only	108.3%	109.3%
Mis-estimation only	106.5%	107.3%
Volatility only	121.8%	125.9%
Behavioral & mis-estimation & volatility	129.6%	133.9%

predicted probabilities $p_{i,(Jan2011,Dec2011]}$ and hence the potential losses, we proceed as follows:

- The MLE estimate $\hat{\theta}$ of the vector of regression parameters is obtained by fitting the base survival model (1).
- The behavioural processes are jointly forecasted over a 12-months horizon using vector auto-regressive models, separately for each account.
- The MLE $\hat{\theta}$ is applied (unto the application variables and forecasted behavioural variables) yielding (a) predicted values of the monthly default probabilities over 12 months and (b) the cumulated probabilities $p_{i,(Jan2011,Dec2011]}$.
- Potential losses are then simulated based on these cumulated probabilities within the Poisson-Binomial distribution (7).

A quantification of the volatility risk in terms of VaR and ES is presented in the third row of Table 3. This suggests that the volatility risk is much higher than either that associated with behavioural or misestimation risk.

4.5. Scenario4: volatility, mis-estimation and behavioural risks

In this scenario, all three sources of uncertainty (i.e., volatility, mis-estimation and behavioural risks) are switched on, through the following steps.

(i) Upon fitting the base survival model (1), the normal distribution (9) is used to simulate a plausible alternative vector of regression parameters.
(ii) For each account, upon fitting the vector auto-regressive model (12) to its behavioural data, a path is simulated of each behavioural variable over the twelve month horizon.
(iii) For each account, the simulated vector generated in (*i*) is applied to the application variables and the simulated behavioural paths generated in (*ii*). This yields a potential path of monthly default probabilities over twelve months, and subsequently the cumulated probabilities $p_{i,(Jan2011,Dec2011]}$.
(iv) These cumulated probabilities are then used to generate a potential loss through the Poisson-Binomial distibution (7).
(v) Steps (*i*)-(*iv*) above are repeated N times, yielding a sample of N potential losses.

A summary of the resulting aggregated risk is shown in the bottom row of Table 3 in terms of VaR and ES. Notice that the VaR and ES resulting from all three sources of risk together is not the

Table 4. Assessment of three sources of uncertainty by employment group.

Source of uncertainty	Employment class	$VaR_{99.5}$	$ES_{99.5}$
Behavioural risk only	Employed	110.0%	111.2%
	Self-employed	104.2%	104.6%
	Retired, unemployed	102.4%	102.8%
	Students	103.1%	103.5%
	Not given	103.1%	103.5%
Mis-estimation only	Employed	106.2%	107.1%
	Self-employed	101.6%	101.7%
	Retired, unemployed	102.5%	102.8%
	Students	100.8%	100.9%
	Not given	103.5%	104.2%
Volatility only	Employed	126.8%	130.2%
	Self-employed	111.8%	104.0%
	Retired, unemployed	107.2%	108.9%
	Students	107.5%	109.2%
	Not given	125.3%	129.0%
Behavioral & mis-estimation & volatility	Employed	134.1%	138.1%
	Self-employed	115.1%	117.7%
	Retired, unemployed	109.5%	111.4%
	Students	116.8%	119.1%
	Not given	130.2%	135.3%

sum of the values for each individual risk component because the risk sources are not independent.

4.6. Risk assessment by employment class

So far, the assessment of uncertainty has been presented at portfolio level. In practice, a portfolio is often made up of different business segments, each segment having a different composition and can cause different risk to the lender. To illustrate the differences between applicant segments we undertake an assessment of the risk separately for each of three levels of employment. We follow the same methodology as above. A summary is shown in Table 4.

Table 4 shows that for all five segments the capital required for volatility risk is much higher than for the other two risks. But the segments differ noticeably in the amount of each type of risk they are associated with. The employed segment has considerably higher behavioural and mis-estimation risk than the other segments and the employed segment and that for which there is no data available have higher volatility risk than those in the other risk segments. Overall, the retired or unemployed segment has the lowest volatility and behavioural risk and the lowest overall risk from these three sources. If all other sources of risk were the same for all segments these results suggest that the amount of capital that needs to be held for the employed segment and the segment where no details are available is larger than that for those retired or unemployed or self employed or students and so these former segments would yield, ceteris paribus, a lower return on equity when lent to.

5. Discussion and conclusions

Many regulators require banks to adopt model risk management principles. For example the UK's Prudential Regulation Authority (PRA) states that banks must engage in "appropriate testing of models to take into account potential limitations, assess their robustness and stability over time, and across a variety of economic and market conditions" (Prudential Regulation Authority, 2018, pp. 8). The PRA does not specify exactly how uncertainty over the future values of behavioural variables, uncertainty over model parameters and uncertainty over the occurrence of default given the predicted PD, should be analysed. In this article, we have developed a framework for the quantification of the impact of these three sources of uncertainty in the prediction of the VaR for a credit portfolio and so the amount of capital a bank is required to hold for that portfolio. There are, of course, many other sources of uncertainty in the computation of required capital, for example the state of the macroeconomy. However, in this article, we are concentrating on different types of modelling risk, conditional on the state of the macroeconomy.

The sources of model risk we are considering are typically omitted from the calculation of the VaR of a retail credit portfolio. In this case the behavioural variables take on fixed values as do the PD model parameters, and so does the predicted PD for each account. The aggregated loss across all accounts (given by Equation 3) takes on a single value; conditional on the macroeconomy, there is no additional uncertainty. In this case the VaR in Equation (4) has the value 1. Comparing the values of the VaR under each scenario from Table 3 with the value of 100% we can see the increase in VaR and so capital if the three sources of uncertainty are incorporated rather than ignored. Thus, in the portfolio we considered, if all three sources of uncertainty are incorporated, the value of capital to be held will increase by up to 30% relative to if they are omitted.

Our article contributes to prudential policy by offering a methodology to incorporate all three sources of risk in the VaR calculation for a portfolio of credit products. Given the potential under-capitalisation if this is not done, this is an important message to lenders and to regulators. We offer a method that regulators might consider advancing to lenders to protect depositors and others and the methodology might be wisely adopted by lenders when computing their economic capital. Parts of the methodology may also be incorporated into the the prediction of expected cash flows for loans for which risk has increased under IFRS9 (and similarly CICL). Under IFRS9 (International Accounting Standards Board, 2013) expected cash flows from loans may be computed using PD probabilities in each future month for which a loan is outstanding and these can be gained from a survival model. The Standard requires that a range of possible outcomes is assessed and their probabilities of occurrence. A survival model of PD may include behavioural variables (as we have shown). So, the range of PDs may be gained following the methods in our article: by simulation of survival model parameters, simulation of behavioural variables values and of the occurrence of defaults given the PD. Of course, other sources of uncertainty would also be included such as possible changes in macroeconomic variables.

We have concentrated on uncertainty associated with PD modelling. A limitation of our approach is that we have not considered uncertainty over the future states of the macroeconomy—we have concentrated on model risk. A second limitation is that the residuals from the VAR that is used to model the behavioural variables may not include enough values or range of values to represent the population distributions. This means the analyst should use as long a time series of the behavioural variables as possible. A further limitation is that aggregate loss is the product of PD, LGD and EAD (Equation 2) and we have concentrated only on the former. However, a similar approach may be applied to these other two terms since there would be uncertainty over the estimation model parameters, and uncertainty as to the future values of behavioural variables which might enter both models. In principle, uncertainties over these variables should be included in the estimation of the VaR.

On the other hand, our method has a number of strengths. One of these is that one can follow it to gain the VaR at a time horizon of the analyst's choice. This follows from our use of a survival model for PD prediction and a VAR to product the future values of the behavioural variables. This means that the method could inform managerial decisions over different time horizons: for example one year, as in Basel or five years. Different stakeholders may be interested in these predictions over different horizons: certain institutions may be more interested in a five year horizon and some depositors only in a one year horizon.

Notes

1. A further group of papers considers worst outcomes; that is, theoretical upper bounds on VaR from model risk (Cont, 2006; Embrechts et al., 2013; Puccetti & Ruschendorf, 2012; Bernard et al., 2017; Bernard 2017; and Glasserman & Xu, 2014.
2. In the context of credit risk, Jacobs et al. (2015) and Gross and Poblacion (2019) give methods for estimating parameter uncertainty, but they do not do so for a credit scoring model and hence not for capital prediction for a specific credit portfolio. Wang et al. (2020) do give a method for estimating parameter uncertainty but do not decompose predicted loss into the three types of risk we are interested in.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bakoben, M., Bellotti, T., & Adams, A. (2020). Identification of credit risk based on cluster analysis of account behaviours. *Journal of the Operational Research Society*, *71*(5), 775–783. https://doi.org/10.1080/01605682.2019.1582586

Barrett, B. E. (2007). Detecting Bias in Jury Selection. *The American Statistician*, *61*(4), 296–301. https://doi.org/10.1198/000313007X243629

Basel Committee on Banking Supervision. (2006). International Convergence of capital measurement and capital standards: a revised framework. Bank for International Settlements

Basel Committee on Banking Supervision. (2010). Basel III: A global regulatory framework for more resilient banks and banking systems. Bank for International Settlements

Bellotti, T., & Crook, J. (2013). Forecasting and stress testing credit card default using dynamic models. *International Journal of Forecasting*, *29*(4), 563–574. https://doi.org/10.1016/j.ijforecast.2013.04.003

Bernard, C., Ruschendorf, L., Vanduffel, S., & Yao, J. (2017). How robust is the value-at-risk of credit risk portfolios? *The European Journal of Finance*, *23*(6), 507–534. https://doi.org/10.1080/1351847X.2015.1104370

Bignozzi, V., & Tsanakas, A. (2016). Model uncertainty in risk capital measurement. *The Journal of Risk*, *18*(3), 1–24. https://doi.org/10.21314/J0R.2016.326

De Boor, C. (1978). *A practical guide to splines*. Springer.

Borgonovo, E., & Plischke, E. (2016). Sensitivity analysis: A review of recent advances. *European Journal of Operational Research*, *248*(3), 869–887. https://doi.org/10.1016/j.ejor.2015.06.032

Brummelhuis, R., Luo, Z. (2019). Interest margin forecasting and capital adequacy stress testing by machine

learning techniques. SSRN. https://ssrn.com/abstract=3282408 or https://doi.org/10.2139/ssrn.3282408.
Chen, S. X., & Liu, J. S. (1997). Statistical applications of the Poisson-binomial and conditional Bernoulli distributions. *Statistica Sinica*, *7*, 875–892.
Cont, R. (2006). Model uncertainty and irts impact on the pricing of derivative instruments. *Mathematical Finance*, *16*(3), 519–547. https://doi.org/10.1111/j.1467-9965.2006.00281.x
Daskalakis, C., Diakonikolas, I., & Servedio, R. A. (2015). Learning Poisson binomial distributions. *Algorithmica*, *72*(1), 316–357. https://doi.org/10.1007/s00453-015-9971-3
Djeundje, V. A. B., & Crook, J. (2018). Incorporating heterogeneity and macroeconomic variables into multi-state delinquency models for credit cards. *European Journal of Operational Research*, *271*(2), 697–709. https://doi.org/10.1016/j.ejor.2018.05.040
Djeundje, V. A. B., & Crook, J. (2019a). Dynamic survival models with time varying coefficients for credit risks. *European Journal of Operational Research*, *275*(1), 319–333. https://doi.org/10.1016/j.ejor.2018.11.029
Djeundje, V. A. B., & Crook, J. (2019b). Identifying hidden patterns in credit risk survival data using generalised additive models. *European Journal of Operational Research*, *277*(1), 366–376. https://doi.org/10.1016/j.ejor.2019.02.006
Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, *11*(2), 89–121. https://doi.org/10.1214/ss/1038425655
Embrechts, P., Puccetti, G., & Ruschendorf, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking & Finance*, *37*(8), 2750–2764. https://doi.org/10.1016/j.jbankfin.2013.03.014
Garcia-Cespedes, R., & Moreno, M. (2016). Probability of default uncertainty in the Vasicek credit risk framework. Available from *SSRM 2761013*
Glasserman, P., & Xu, X. (2014). Robust risk measurement and model risk. *Journal of Financial Stability*, *9*, 139–149.
Gogas, P., Papadimitriou, T., & Agrapetidou, A. (2018). Forecasting bank failures and stress testing: A machine learning approach. *International Journal of Forecasting*, *34*(3), 440–455. https://doi.org/10.1016/j.ijforecast.2018.01.009
Gross, M., & Poblacion, J. (2019). Implications of model uncertainty for bank stress testing. *Journal of Financial Services Research*, *55*(1), 31–58. https://doi.org/10.1007/s10693-017-0275-4
Henry J. and Kok C. (Eds). 2013. A macro stress testing framework for assessing systemic risks in the banking sector. European Central Bank Occasional paper, No 152.
Hofmarcher, P., Kerbl, S., Grün, B., Sigmund, M., & Hornik, K. (2014). Model uncertainty and aggregate default probabilities: New evidence from Austria. *Applied Economics*, *46*(8), 871–879. https://doi.org/10.1080/00036846.2013.859378
Hong, Y. (2013). On computing the distribution function for the Poisson binomial distribution. *Computational Statistics & Data Analysis*, *59*, 41–51. https://doi.org/10.1016/j.csda.2012.10.006
International Accounting Standards Board. (2013). Exposure draft ED/2013/3 financial instruments: Expected credit losses. http://www.ifrs.org/Current-Projects/IASB-Projects/Financial-Instruments-A-Replacement-of-IAS-39-Financial-Instruments-Recognitio/Impairment/Exposure-Draft-March-2013/Comment-letters/Documents/ED-Financial-Instruments-Expected-Credit-Losses-March-2013.pdf
Jacobs, M., Karagozoglu, A. K., & Sensenbrenner, F. J. (2015). Stress testing and model validation: Application of the Bayesian approach to a credit risk portfolio. *The Journal of Risk Model Validation*, *9*(3), 41–70. https://doi.org/10.21314/JRMV.2015.140
Kupiec, P. (2018). On the accuracy of alternative approaches for calibrating bank stress test models. *Journal of Financial Stability*, *38*, 132–146. https://doi.org/10.1016/j.jfs.2018.08.001
Lando, D., & Skødeberg, T. M. (2002). Analysing rating transitions and rating drift with continuous observations. *Journal of Banking & Finance*, *26*(2 − 3), 423–444. https://doi.org/10.1016/S0378-4266(01)00228-X
Misina, M., & Tessier, D. (2008). Non-linearities, model uncertainty and macros stress testing. *Bank of Canada Working Paper*, 2008–2030. https://ssrn.com/abstract=2185664.
Neyman, J. (1939). On a New Class of "Contagious" Distributions, Applicable in Entomology and Bacteriology. *The Annals of Mathematical Statistics*, *10*(1), 35–57. https://doi.org/10.1214/aoms/1177732245
O'Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems. *Statistical Science*, *1*, 502–518.
Petropoulos, A., Siakoulis, V., Panousis, K. P., Christophides, T., & Chatzis, S. (2020). A deep learning approach to dynamic balance sheet stress testing. arXiv: 2009.11075v1.
Prudential Regulation Authority. (2018). Model risk management principles for stress testing. Supervisory Statement SS3/18.
Puccetti, G., & Ruschendorf, L. (2012). Comparison of sharp bounds on the distribution of a function of dependent risks. *Journal of Computational and Applied Mathematics*, *236*(7), 1833–1840. https://doi.org/10.1016/j.cam.2011.10.015
Richards, S. J. (2016). Mis-estimation risk: Measurement and impact. *British Actuarial Journal*, *21*(3), 429–457. https://doi.org/10.1017/S1357321716000040
Siemsen, T., Vilsmeier, J. (2018). On the quest for robustness: about model risk, randomness and discretion in credit risk stress tests. Deutsche Bundesbank, No 31/2018.
Wang, Z., Crook, J., & Andreeva, G. (2020). Reducing estimation risk using Bayesian posterior distribution approach: Application to stress testing mortgage loan default. *European Journal of Operational Research*, *287*(2), 725–738. https://doi.org/10.1016/j.ejor.2020.04.020