European Journal of Operational Research 255 (2016) 649–658

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Interfaces with Other Disciplines

Incorporating lifecycle and environment in loan-level forecasts and stress tests

Joseph L. Breeden*

Prescient Models LLC, 300 Catron St., Suite B, Santa Fe, NM 87501, USA

ARTICLE INFO

Article history:
Received 22 February 2014
Accepted 4 June 2016
Available online 11 June 2016

Keywords:
Forecasting
Risk
Banking
Time series
Age-Period-Cohort models

ABSTRACT

The new FASB current expected credit loss (CECL) proposal, IASB's IFRS 9, and regulatory stress testing all require that the industry move toward forecasting probabilities of future events, rather than simply rank-ordering loans. Even more importantly, effective loan pricing requires this same forward-looking, loan-level forecasting.

We created a loan-level version of Age-Period-Cohort (APC) models suitable for forecasting individual loan performance at a point-in-time or for the loan's lifetime. The APC literature explains that any model of loan performance must make either an explicit or implicit assumption around the embedded model specification error between age of the loan, vintage origination date, and performance date. We have made this assumption explicit and implemented a technique using augmented macroeconomic history to stabilize the analysis.

The preceding steps provide robust estimates of lifecycle and environmental impacts. We then use a Generalized Linear Model (GLM) with a population odds offset for each age/time combination derived from the lifecycle and environment functions in order to estimate origination and behavior scores. Analyzing a small US auto loan portfolio, we demonstrate that this model is robust out-of-sample and out-of-time for predicting both rank-ordering and probabilities by inserting the odds offset appropriate for the environment being modeled.

In addition to producing loan-level forecasts and stress tests, the scores produced have higher rank-order performance out-of-sample and out-of-time than standard scores. The scores prove to be robust years into the future with no measurable degradation in performance because of the stabilizing effect of the offset factor during model construction.

1. Introduction

Credit scores were originally developed to aid loan origination. Applicants would be screened by estimating a score from available application data. The score served as an impartial criterion for assessing risk. The "cut-off" score was the threshold below which riskier applicants were denied loans.

The transition to risk-based pricing meant that a wider range of applicants could be accepted by modifying the loan terms to match the riskiness of the applicant. Risk-based pricing came in response to competitive pressure among lenders. As margins contracted, lenders needed to better target their pricing.

As lenders became more reliant on scores, greater effort was put into improving score accuracy. Improved estimation techniques, the transition to logit and probit models, and the use of bureau attributes all enhanced the ability of the scores to assess risk.

The scorecard uses these factors from a dataset, $\boldsymbol{x}$, to predict the probability of "good", $p(G|\boldsymbol{x})$. The odds of a loan being good are (Thomas, 2009)

$$o(G|\boldsymbol{x}) = \frac{p(G|\boldsymbol{x})}{p(B|\boldsymbol{x})} = \frac{p_G}{p_B}\frac{p(\boldsymbol{x}|G)}{p(\boldsymbol{x}|B)} \equiv o_{Pop} \times I(\boldsymbol{x}), \boldsymbol{x} \in X \tag{1}$$

or

$$\log(o(G|\boldsymbol{x})) \equiv \log(o_{Pop}) + \log(I(\boldsymbol{x})) \tag{2}$$

where p_G and p_B are the unconditional odds of good or bad, $\frac{p_G}{p_B}$ is the population odds o_{Pop}, and $I(\boldsymbol{x}) = \frac{p(\boldsymbol{x}|G)}{p(\boldsymbol{x}|B)}$ is the information odds. We could also say that o_{Pop} captures systematic effects for the portfolio and $I(\boldsymbol{x})$ captures the idiosyncratic effects for an individual loan.

* Tel.: +1 505 490 6042.
E-mail address: breeden@prescientmodels.com

http://dx.doi.org/10.1016/j.ejor.2016.06.008

When we create a credit score, we expect the information odds to be reasonably robust out-of-sample. The population odds, however, are dependent upon the macroeconomic conditions prevailing during the in-sample period. In future time periods, the population odds should change due to factors not captured in the credit score. For this reason, credit scores are used as risk ranking tools out-of-time, not as predictors of $o(G|\boldsymbol{x}_{oot})$ where $\boldsymbol{x}_{oot}$ are the loan attributes out-of-time. (Out-of-time refers to data from time periods not included in the training sample.)

Many practitioners expand the attributes $\boldsymbol{x}$ to include macroeconomic factors and the age of the loan in an attempt to predict the population odds as well, but with mixed results. As explained in the Age-Period-Cohort (APC) literature (Glenn, 2005; Mason & Fienberg, 1985) and applied to the context of credit risk modeling (Breeden & Thomas, 2016), a model specification error is embedded in the dynamics of retail lending. Traditional credit scoring attributes are measured in the loan origination month, also known as the vintage date, v. Macroeconomic data is measured with calendar date t. Lifecycle functions as in survival models (Cox & Oakes, 1984; Efron, 2002; Hosmer & Lemeshow, 1999; Therneau & Grambsch, 2000) are measured versus the age of the loan a. However, $a = t - v$, leading to a linear specification error if factors measured along all three dimensions are included in the model simultaneously and without constraint. In cases where some of these dimensions are excluded, as with traditional credit scores that rely solely on information from the origination (vintage) date, a unique solution is obtained, but at the cost of being unable to predict probabilities in future time periods.

The APC literature proves that no general solution exists for this specification error, with the implication that we can never be certain of the linear trends in lifecycle, macroeconomic, or credit risk functions. Instead, domain-specific solutions are recommended by incorporating constraints suitable to the specific situation being modeled. An inability to be certain of linear trends in time would mean that we could not reliably predict the population odds in future time periods, bringing the effectiveness of any stress test model into question.

Breeden and Thomas (2016) described a constraint that would appear to be reasonable for retail lending. Namely, that after fitting macroeconomic factors to the environment function, the slope of that macroeconomic impact should be zero when extrapolated backward across multiple macroeconomic cycles. For any model that includes macroeconomic factors, this is equivalent to creating an environmental index from a subset of the model as $\hat{E} = f(E_i(t), c_i)$ where the E_i are the individual macroeconomic factors and c_i are the estimated coefficients for their inclusion in the credit risk model. Then create the constraint that $\text{slope}(\hat{E}(t)) = 0$ when $\hat{E}(t)$ is extrapolated backward over multiple decades of macroeconomic history.

Over any short time frame of less than one economic cycle, the environment will certainly not show zero trend, but the trend is also unlikely to be identically zero even over longer periods. Nonetheless, the assumption of zero environmental trend over many cycles is consistent with the assumption that a through-the-cycle average can be defined for macroeconomic impacts, i.e. that a through-the-cycle probability of default (PD) exists in the sense that the population odds can be a constant when measured for a reference portfolio across multiple economic cycles. Therefore, although the zero environmental trend assumption is a good starting point, it must be tested and potentially corrected if it is not precisely true.

Using the technique of Breeden and Thomas to control the model specification error, we demonstrate a method of creating credit scores that estimates both the population odds and information odds in-sample, and provides for extrapolating the population odds out-of-sample so that loan-level probabilities are forecasted. This paper represents the first time that APC algorithms have been used for account-level forecasting. Incorporating the retrending process of Breeden and Thomas means that the score created will be stable through economic cycles without risk of cross-correlation between scoring factors and macroeconomic factors.

The newly adopted accounting standards of IFRS9 (International Financial Reporting Standards Foundation, 2014) and CECL (Financial Accounting Standards Board, 2012) require just such features. On the one hand, they are both based upon lifetime loss estimation from any point in the age of the loan. That necessitates inclusion of a lifecycle such as is found in survival models or APC models. However, other requirements within the guidelines for IFRS9 and CECL suggest that account-level models are to be preferred. The Cox proportional hazards algorithm (Cox, 1992) in principle applies to such problems, but the paper by Breeden and Thomas demonstrates that a risk of linear trend ambiguity exists in all credit risk models, and subsequent work by Breeden, Bellotti, Yablonski, and Leonova (2016) demonstrates this problem explicitly for Cox PH. Therefore, the authors believe that a new approach is needed that has the scoring and hazard function attributes of Cox PH, but with explicit control of the linear trend so that no estimation confusion occurs between the scoring and macroeconomic factors.

2. Modeling approach

Although a single-stage approach is in principle possible using a constrained optimization, we followed a sequential analysis using simpler algorithms. The following steps were performed in the analysis.

1. Decompose loan-level performance data with an Age-Vintage-Time (AVT)
2. Fit the time function to macroeconomic data
3. Retrend the age and vintage functions
4. Fit a credit score with age and time offsets

2.1. Age-Vintage-Time decomposition

The first step is to estimate the lifecycle as a function of age of the loan, the vintage quality as a function of vintage, and the environment as a function of time (calendar date). This analysis should be performed on the longest history available, preferably longer than the two to three years of history that is typical of credit scores.

For the Age-Vintage-Time (AVT) decomposition, we can use standard Age-Period-Cohort (APC) implementations to analyze vintage-aggregate time series with a logit transformation, or use an equivalent loan-level implementation of APC. Each vintage will be measured each month to create an appropriate rate. For example, to predict the default rate, default accounts and active accounts would be reported each month. The APC algorithm would estimate

$$r(a, v, t) = \frac{\text{Defaults}(t)}{\text{Active.Accounts}(t-1)} = \frac{1}{1 + e^{-(F(a)+G(v)+H(t))}} \tag{3}$$

where $r(a, v, t)$ is the default rate, $F(a)$ is the lifecycle with age, $G(v)$ is the credit quality by vintage, and $H(t)$ is the environment function over time. With standard implementations of the APC algorithm, all three functions are estimated via splines with the analyst specifying the number of spline nodes. Although $F(a)$ is usually given fewer nodes on the assumption of a relatively smooth lifecycle function, $G(v)$ and $H(t)$ should have as many nodes as the data will support in order to capture sudden changes in the portfolio composition or macroeconomic environment respectively. Alternatively, standard APC implementations usually support nonparametric estimation of any of the three functions.