→ data → information → profit

A. BACKGROUND AND SCOPE

A.1. Goals of Paper

This paper defines a road map for scorecard changes arising from payment holidays. It focuses on ongoing behavioral scores, used to support operational decisions, IFRS9 provision calculations and Basel capital calculations. Like any good road map, it sets out alternative routes. It is not a recipe book set of instructions.

A.2. Structure

- Section A defines the problem and clarifies terminology and scope.
- Section B gives a three-step procedure to create a scorecard which is appropriate to customers currently on payment holidays.

A.3. Payment Holidays

During the Covid-19 crisis, many creditors have offered payment holidays (PH) to their borrowers. In many cases, this has been done at the direction of governments and regulators. The effect is to suspend the customer's payments for a period of time without recording these as missed payments.

Payment holidays are also known as deferrals and moratoria.

The goal of a payment holiday is to reduce outgoings for borrowers who have lost income. They are based on the assumption that the customer will return to a normal level of income in a relatively short period of time and can then resume normal payments.

A.4. Variations on a Theme

The definition of the payment holiday differs from market to market. The main differences are:

- In some cases, the payment holidays have been granted automatically to all borrowers. But in most, the customer must request the holiday. Hence, there are customers with continuing payments and others with no payments due.
- Lenders may have discretion to allow or refuse a payment holiday. Often, this is done through clear policy rules (e.g. available to all customers who are currently up to date or one cycle in arrears). But in some cases, it is subject to the discretion of the lender, usually implemented through a score (or PD) cut-off.
- The duration of the payment holiday is most frequently three months. However, many are being extended, often for a further three months.
- In most cases, interest continues to accrue during the payment holiday. Thus, the lender continues to recognize income. But in some cases this is not reflected on the reported balance on the account.
- The payment holidays have been accompanied by a variety of commitments that the customer's "credit rating" will not suffer directly from using a payment holiday. The wording of these guarantees varies widely from one country to another and from one portfolio to another.

A.5. Credit Bureau Reporting

Payment holidays are not treated as delinquency episodes. In most cases, the credit will be reported to bureaus each with the same status as in the month immediately prior to the holiday. Sometimes, the presence of a payment holiday can be detected, either as a specific status recorded at the bureau, or by observing a static balance from month to month.

A.6. Impact on Credit Scores

In most cases, monthly behavior scores (which underlie Basel IRB and IFRS 9 calculations) do not change as a result of payment holidays. Either the borrower's score is frozen at the level observed before the lookdown, or the score is recalculated each month, but without any penalty for missed payments.

A.7. Sources

The paper draws heavily from material covered in the Scoreplus training courses, notably Building Better Scorecards (BBS). To a lesser degree, it uses ideas from:

- Basel Models and Validation (BMV)

- Portfolio Management Analytics (PMA).

A.8. Health Warning

The recommendations in this paper are based on experience and reflection. But they have not been tested on empirical data. In several cases, alternative approaches are suggested.

It is quite possible that there are hidden snags or complications that have not been anticipated. Scoreplus would welcome suggestions, questions and debate about the ideas in this paper.

A.9. Exclusion from Scope: Broader Economic Impact on Scores

The Covid-19 crisis has led to an exacerbation of risk on almost credit portfolios. Losses for 2020 and 2021 will be much higher than those seen in recent years. Therefore, score models must be adjusted to allow for this change in point in time "expected loss". Some of this change can be observed in payment patterns observed since the onset of the crisis. But much of the deterioration will depend on the economic trajectory of employment and incomes over the next two years.

This issue is broader than payment holidays. This paper does not cover it. It will need considerable analysis to the end of 2020. But the final outcome will not be based on solely on data analysis or on economic models. It will require considered management judgment - and the certainty of getting it wrong, to a greater or lesser degree.

B. CALIBRATING SCORES DURING PAYMENT HOLIDAYS

B.1. Overall Idea

The three steps in this phase are discussed in turn:

- Identify customers who are statistically identical apart from taking payment holidays. These are called "twins". Some will have taken a payment holiday (PH) and some not (NH). This can be done through an *ad hoc* scorecard, the H-score.
- For each PH customer, calculate the "current" scores (or PDs - probability of default) for NH "twin" customers. The twin scores - and a judgmental adjustment for the difference between PH and NH customers - are used to assign a score (the S-score) to each PH customer. This serves as a surrogate of what would have been the current score for this customer if the payment holiday had not been offered.
- Build a model - in the simplest case by linear regression - for the S-score based on available history for the PH cases at the current point. This is denoted by T-score.

The T-score is on the same scale as the S-score and can be used to assign a score for PH customers which is compatible with the S-scores on the NH population.

B.2. Identifying Twins

B.2.1. H-score

The customers who opt for payment holidays (PH) are not typical of the rest of the portfolio (NH). The goal of this step is to understand the differences - who opted for a payment holiday?

A binary logistic regression is built to estimate the propensity to take a payment holiday, called the H-score. The specifications are:

- Population: All customers eligible for a payment holiday observed at the point when payment holidays were first offered.
- Predictors: Behavioral score applied at the observation data and any other characteristics which would have been available at that date. Also, any post-holiday data that is available on a uniform basis for both PH and NH customers should be used (see note below).
- Outcome: PH = 1, NH = 0.
- Method: Mother-child scorecard, using the current behavioral score as a mother score (see BBS, Session 11, slides 1133-1134).

The H-score should be a sufficient statistic for customers' propensity to take a payment holiday - i.e. given the H-score, there is no further information available to indicate who is likely to take a payment holiday. However, knowing that a customer took a payment holiday gives the lender extra information, which can be used to drive future estimates of behavior, notably of risk behavior.

B.2.2. Characteristics for H-score - Pre-Payment Holiday

The score model should be relatively straightforward. In addition to the mother score, any other information available before the payment holiday is given should be considered. Candidate characteristics for the model include:

- Stable Income Index (see Scoreplus paper "Stable Income Index", Helen McNab and Gerard Scallan, June 2020)
- Recent delinquency
- Any recently acquired credit (from credit bureau)
- Size of payment due (if possible in relation to income or some surrogate)
- Anything else that matches the business understanding of portfolio managers.

B.2.3. Characteristics for H-score - Since Payment Holiday

In addition to this pre-holiday data, the model should also use predictors which are generated during the payment holiday on a consistent basis for PH and NH customers. Relevant predictors are likely to be:

- recent delinquency on other accounts (from credit bureau)
- trend in current account turnover since start of holiday period (surrogate for income reduction)
- evidence of payment holiday with other credit grantors
- recent usage for revolving credit products.

Anything where the presence of the payment holiday will not directly influence the value of a characteristic is a valid predictor. So all data derived from account usage (as opposed to payments) can be used. The rationale for using the holiday data is discussed in B.5.1 below.

B.2.4. Identifying Twins - fixed methods

- The simplest method of identifying twins is to find a fixed number, k, of NH customers with H-scores closest to that of customer c:

  $$\text{Twins}(c) = \{d \in NH \mid \|H(d) - H(c)\| \le \varepsilon_c\}$$

  where ε_c is selected to give exactly k twins. k can range from 1 to around 30.
- An alternative is to define the Twins by a fixed value of ε, which does not vary from case to case. Some PH customers will now have lots of twins while others may have none.

B.2.5. Identifying Twins - randomization

It is also possible to randomize the selection of (a single) twin:

- For every PH customer c, with score H_c, calculate the standard deviation of H_c, sd_c. See BBS, session 5, slides 520-522. This measures the random noise in the H-score.
- Assign a weight to each NH customer based on their density in a normal distribution with mean H_c, and variance sd_c^2.
- Sum all the weights and normalize by dividing the weights by this value.
- Randomly select one twin from the PH customers by according to the normalized weights.

Customers with "sure" H-scores will have concentrated distributions and have a twin with an H-score very close to H_c. Those with a lot of uncertainty will have wide distributions, with relatively heavy weights on NH customers far removed from H_c.

B.2.6. Nearest Neighbor Method

Instead of the H-score, a K Nearest Neighbor (KNN) algorithm can identify the NH twins for each customer c. But the metric used to drive the KNN algorithm needs definition. The KNN replaces both the H-score and the twinning process.

B.3. Assigning the S-score

B.3.1. Overview

The goal is to estimate what score a PH customer would have now ("current score" or S-score), had the payment holiday not been offered. There are two sub-steps, each discussed in more detail in the following paragraphs:

- Assign a score S_0 to customer c based on the current scores of the twins. In the simplest case, this can be done by randomly choosing one twin and assigning its score to customer c.
- Apply a (judgmental) offset, J, as a penalty to S_0. This reflects that taking a payment holiday is evidence of higher vulnerability and therefore of higher risk. So the PH customer c is a worse risk than the NH twins. The S-score is given by

$$S = S_0 - J$$

B.3.2. *Combining Twins to Calculate the S-score for customer c*

There are three methods possible to assign an S-score to a PH customer, c:

- Mean: $S_{c0} = \text{mean}\{S_d | d \in \text{Twins(c)}\}$
- Median: $S_{c0} = \text{median}\{S_d | d \in \text{Twins(c)}\}$
- Randomized: $S_c^0 = \text{random}\{S_d | d \in \text{Twins(c)}\}$ - one element is chosen at random from the set Twins(c).

The randomization method can also be applied directly through the random twin method described in B.2.5.

The mean and median methods are likely to reduce the amount of random variation in the distribution of the S^0 scores. This will make the next step look artificially precise. Randomization is also more robust in where the H-scores are extreme - in other words, where the propensity to take a payment holiday is strongly polarized and the Gini coefficient of the H-score is strong. For this reason, the randomization method is preferred.

The “current score”, S_d, for each NH twin is the behavioral score used in the current month. This may have been modified since the beginning of the crisis based on data or on judgment. Its statistical reliability does not enter into the use of the algorithm - but obviously is critical for the use of the scores.

B.3.3. *Adjusting the S-score for Payment Holiday*

The S_{c0}-score is a risk score. It represents the risk of NH twins of the PH customer c. But there is an important difference. The NH customers had the choice of taking a payment holiday and chose not to; the PH customer c took the holiday. This extra piece of information gives some extra insight into future risk: at the same S^0 score, the PH customer is considerably riskier than the NH customers.

Log-odds scores work well to quantify risk differences from other information. Thus, it is reasonable that the extra risk from the payment holiday be quantified by a constant offset on the score, labeled J. This will be applied to all the PH customers.

When payment holidays are first granted, there is no data to quantify this risk offset J. It is a judgmental parameter. The following factors guide the corporate judgment:

- Taking a payment holiday is less serious than missing a payment. So J should be less than the score penalty which would result from missing a payment on the S-score model.
- If the scorecard uses external bureau data each month, the penalty arising from learning that the customer has had a delinquency with another creditor can be used as the J-adjustment.
- In the absence of this data, a penalty corresponding to multiplying odds by ⅔ to ½ seems a reasonable point for discussion. On a scorecard with 20 points to double odds, this would give a penalty of -12 to -20 points.

B.3.4. *Data calibration of J adjustment*

Once some payment holidays have come to an end (say in month m), it is possible to calibrate J based on data. This is done by running a logistic regression:

- Population: All customers due to make a payment in month m - both those exiting payment holidays in this month and those who were not on payment holidays, the NH customers.
- Outcome: Schedules payment made = 1; No payment (or partial payment) = 0;
- Predictors: S-score and an indicator variable for exiting payment holidays (and intercept); the S-score incorporates the original judgmental estimate of J.
- The coefficient of the indicator variable shows the correction to be applied to the judgmental J.
- See B.5.2 below for further discussion.

B.3.5. Averaging when PH penalties are not permitted

In some cases, there is a regulatory prohibition on adverse treatment of customers who have taken payment holidays. In this case, J should still be estimated. But a weighted adjustment is applied to all customers, PH and NH alike. The adjustment is given by:

$$J' = \|PH\| \times J / (\|PH\| + \|NH\|)$$

In other words, every customer's score - both PH and NH - is reduced by the average penalty, to take into account the incidence of payment holidays on portfolio risk.

Once J has been estimated, the estimated S-score for each PH customer is given by:

$$S = S_0 - J$$

Each PH customer has now been assigned a score on the same scale as the S-score which applies to NH customers. This is the best estimate of their risk, based on the available evidence. But it is not calculated through a "standard" score calculation - the weighted sum of variables calculated from the customer's own behavior.

B.4. Estimating a Modified Scorecard

B.4.1. Linear Regression for the T-score

The goal of this section is to obtain a score that can be calculated from the behavioral characteristics of the PH customers and which matches as closely as possible the S-scores assigned in the previous section. This is called the T-score.

The simplest method to obtain the T-score model to run a linear regression:

- Outcome variable: S_c
- Population: PH customers measured at some month during the payment holiday
- Predictors: any behavioral characteristics available at the point where the S-scores are computed.

Shrinkage regression methods, such as Lasso, Ridge or Least Angle Regression can also be used. These give more "economical" models (with fewer parameters) and are less sensitive to outliers.

B.4.2. Logistic Regression for the T-score

For analysts with experience in scoring (and appropriate software), it is more natural to "parcel" the PH customers:

- Convert the S-score into a probability of bad, PD_c, using the assumed score-odds relationship for the NH population.
- For each customer c, generate a random number (uniform between 0 and 1), U_c.
- If $U_c < PD_c$, label customer c as a "bad"; otherwise label customer c as a "good"

A logistic regression is then run with outcome variable "good" = 1, "bad" = 0, and using the same population and predictors as for the linear regression.

B.4.3. Other Estimation Methods

Alternatively, the T-score can be estimated using any other discriminant method, such as random forests or neural networks, estimating either the S-score or the randomized parceled output 1 or 0.

B.4.4. Predictive Characteristics

- The last pre-holiday score is likely to be the strongest predictor and should be used as a mother score for the T-scorecard.
- The H-score should also be used, either as a continuous or discretized predictor. Alternatively, the constituent characteristics could be used, to avoid collinearity with the "regular" behavioral score.
- Months since entry to the payment holiday is also likely to be predictive.
- The SII (Stable Income Index) is likely to be an important predictor.
- External bureau information will give important insight into the behavior of the customer on other credits.
- Current account behavior is key in understanding what has happened in the payment holiday period.

B.5. Design Issues

B.5.1. Information During Payment Holiday

If the H-score is restricted to "at the time" data, then the identification of "twins" will not match post-holiday behavior patterns. For instance, a customer whose income has gone to zero during the payment holiday will not have a different S-score from someone whose income has been untouched during that time. Thus, the S-score will not take into account changes in risk that occur during the payment holiday.

The T-score mimics the S-score. So if there is a problem with the S-score, the T-model will not take into account this information either.

To avoid this problem, the H-score must use post-holiday data which is not directly dependent on whether the customer is in a payment holiday or not. Therefore, customers whose income has gone down significantly can be differentiated from those whose incomes have remained steady during the crisis. But defining which characteristics are eligible for the H-score requires judgment and portfolio expertise.

Limitations on the depth of this data may limit the effectiveness of the T-score as a whole. It is important that this score reflects the details of customer behavior (other than payments) during the payment holiday.

This is probably the most important limitation on the method suggested here.

B.5.2. Short-term and Full Outcomes

The usual outcome on a behavioral score is at least 12 months. But it is not possible to wait for a year before correcting the weights on judgmental parameters, such are the J-offset. Short term adjustments are necessary.

The earliest indicator of actual performance is missing the first payment due at the expiry of a payment holiday. However, this correction is likely to be too harsh. Some customers exiting payment holidays may miss payments for administrative reasons. Once a longer history is available (say three months after the end of the payment holiday), the exercise should be repeated to obtain a more reliable estimate based on delinquency (or not) three months after the end of the holiday.

Gerard Scallan
27 July 2020
gerard.scallan@scoreplus.com