Chapter 6
Probabilities of Default (PD): Discrete-Time Hazard Models

INTRODUCTION

In the previous chapter, we saw that credit scores are indicators of the credit risk of borrowers. Default probabilities or probabilities of default (PDs) are in essence credit scores that are standardized likelihood measures with a range between zero and one, whereby zero implies that an event is impossible to occur and one implies certainty. Realistic models generally assign PDs between zero and 30 percent to loans.

The PD is the most scrutinized parameter in credit risk analytics and subject to minimum standards imposed by prudential regulators. For example, banks are required to include and exclude specific risk factors. Furthermore, minimum floors such as three basis points are imposed (often to have a nonzero PD value for low default portfolios, which is a subject that we explore later). Banks are also required to validate their PD estimates with rigorous tests, which we discuss in the validation chapter.

Default Events

A PD describes the likelihood of a default event. Banks observe whether borrowers default, and generally indicate this with a default indicator:

$$D_{it} = \begin{cases} 1 & \text{borrower } i \text{ defaults at time } t \\ 0 & \text{otherwise} \end{cases}$$

with $i = 1, \ldots, I$ and $t = 1, \ldots, T$. We assume that the default event is random and use an uppercase letter D as the random variable and a lowercase letter d as its realization. A default event may be defined by any of the following events:

- Payment delinquency of a number of days or more; popular thresholds are 30, 60, and 90 days
- Bankruptcy of the borrower
- Collateral owned by a bank (e.g., real estate owned after an unsuccessful sale at a foreclosure auction)
- Foreclosure of loan
- Short sale of loan
- Loss/write-down amount

- Involuntary liquidation
- Debt modification as a positive interest, expense, or principal forgiveness

Alternative default definitions are possible, and one example may be the Basel definition, which is based on a payment delinquency of 90 days or more. In later sections, we discuss examples where banks observe other outcomes such as payoff, winsorizing, or competing default states (e.g., delinquency, liquidation, or receivership).

Conditional and Unconditional Default

Banks track the performance of loans over time and include dynamic information in their measurements. Time-varying information may include information about the borrower, loan, and collateral (idiosyncratic information), and macroeconomic conditions (systematic information). Common assessment periods may be a day, a month, a quarter, or a year. The choice of time period may depend on the availability of information and the needs of stakeholders such as depositors, investors, prudential regulators, or shareholders. The assessments of credit risk support various functional bank areas. Some areas (e.g., capital allocation and risk reporting under Basel) require risk estimates for the next period, whereas others (e.g., loan loss provisioning under IFRS 9) cover multiple periods, often the lifetime of financial instruments.

Exhibit 6.1 shows the conditionality of default events for two periods. Defaults are generally terminating events, and a default event in one period is therefore conditional on survival (i.e., nondefault) in prior periods.

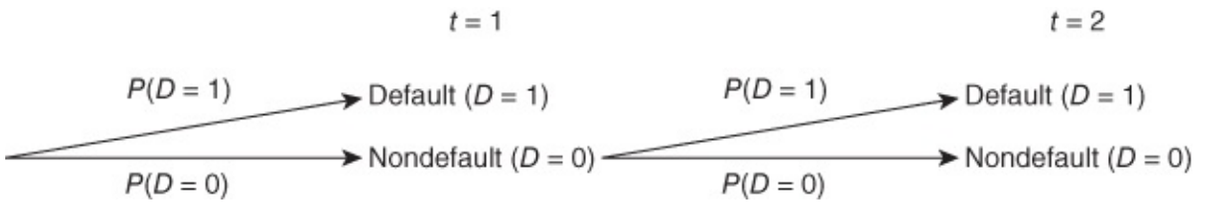

Exhibit 6.1 Conditionality of Default Events

Banks measure the probability of default (PD) for both counterparties and financial instruments. As mentioned, the terminology of default probability is generally used to refer to the one-period default probability. Other terms used are the conditional default probability (conditional on survival) and default intensity.

Furthermore, multiyear default probabilities may be computed as unconditional default probabilities and applied in multiyear evaluations, as is done in the context of computing expected present values or loss values for financial instruments. The unconditional probability of default is often used to measure the likelihood of default from the perspective of the loan's origination. The conditional probability of default measures the likelihood of default conditional on survival and is often used to measure the risk after origination. The conditional and unconditional default probabilities are identical for the first period.

Assuming that every borrower has the same conditional probability of default $PD_{t-1,t}$ at any

period $t-1,t$, we can omit the borrower index i. The unconditional probability of default $UPD_{t-1,t}$ can then be computed as follows:

$$\begin{aligned} UPD_{t_1,t_2} &= S(t_1) - S(t_2) \\ &= \prod_{t=1}^{t_1}(1 - PD_{t-1,t}) - \prod_{t=1}^{t_2}(1 - PD_{t-1,t}) \\ &= \prod_{t=1}^{t_1}(1 - PD_{t-1,t}) - \prod_{t=1}^{t_1}(1 - PD_{t_1,t}) \prod_{t=t_1}^{t_2}(1 - PD_{t-1,t}) \\ &= \prod_{t=1}^{t_1}(1 - PD_{t-1,t}) \left(1 - \prod_{t=t_1}^{t_2}(1 - PD_{t-1,t}) \right) \\ &= S(t_1) PD_{t_1,t_2} \end{aligned}$$

$S(t)$ is the cumulative survival probability to time t (i.e., no default by time t).

Real-World versus Risk-Neutral

Default probabilities are called real-world default probabilities if they are modeled for real-world default realization. The majority of a bank's credit risk exposures relate to illiquid loans (in particular mortgage loans) for which a bank estimates real-world default probabilities.

For marketable exposures, risk-neutral probabilities of default may be derived from observed market prices (e.g., share prices or credit default swap spreads). This approach is very popular in the finance literature but limited to a small fraction of total credit exposures, namely public and/or large counterparties for which equity, debt, or derivatives markets exist. Moreover, risk-neutral probabilities can be quite different from real-world probabilities due to premiums for risk aversion and are therefore of only limited value for risk management.

Basel Requirements

In order to determine regulatory capital requirements, banks often assign rating classes to borrowers and then compute default probabilities for these rating classes. For corporates, central governments, and central banks, the default rating should exclusively reflect the risk of obligor default. At least seven ratings should be available for the nondefaulted obligors and one rating for the defaulters. For retail exposures, a credit rating needs to reflect both obligor and transaction risk. Hence, credit product characteristics and collateralization also need to be taken into account. An excessive concentration of obligors in a rating should be avoided. Instead, ratings should be as homogeneous as possible in terms of default risk. No minimum number of ratings for retail exposures is suggested.

For corporates, central governments, central banks, and retail exposures, a PD needs to be provided per rating. Note that depending upon geographical location, ratings may also be referred to as pools, segments, grades, classes, or clusters. A defaulter is defined as an obligor who is unlikely to pay his or her obligations or is past due more than 90 days. In the United States this has been set to either 120 or 180 days depending on the exposure class. In the

United Kingdom, both 90 and 180 days are used. A lower bound has been set to the PD of three basis points. For obligors in default, it is obvious that the PD equals 100 percent. To calculate the PD, at least five years of historical data should be used, although not all data should receive equal weight in case older data is less relevant. The PD of a particular rating can then be estimated as the long-run average of the one-year default rates.

Parameter Estimation

It is common to assume that default events are driven by an unobservable data-generating process (DGP). The data-generating process is unknown and much research has focused on understanding its key components. Many different models have been proposed in the literature, and a popular one is the Merton (1974) model for corporate borrowers. In this model, default occurs if the market value of the assets (or the return) falls below the market value of the outstanding debt. Exhibit 6.2 shows the derivation of a one-year default probability from such a structural model.

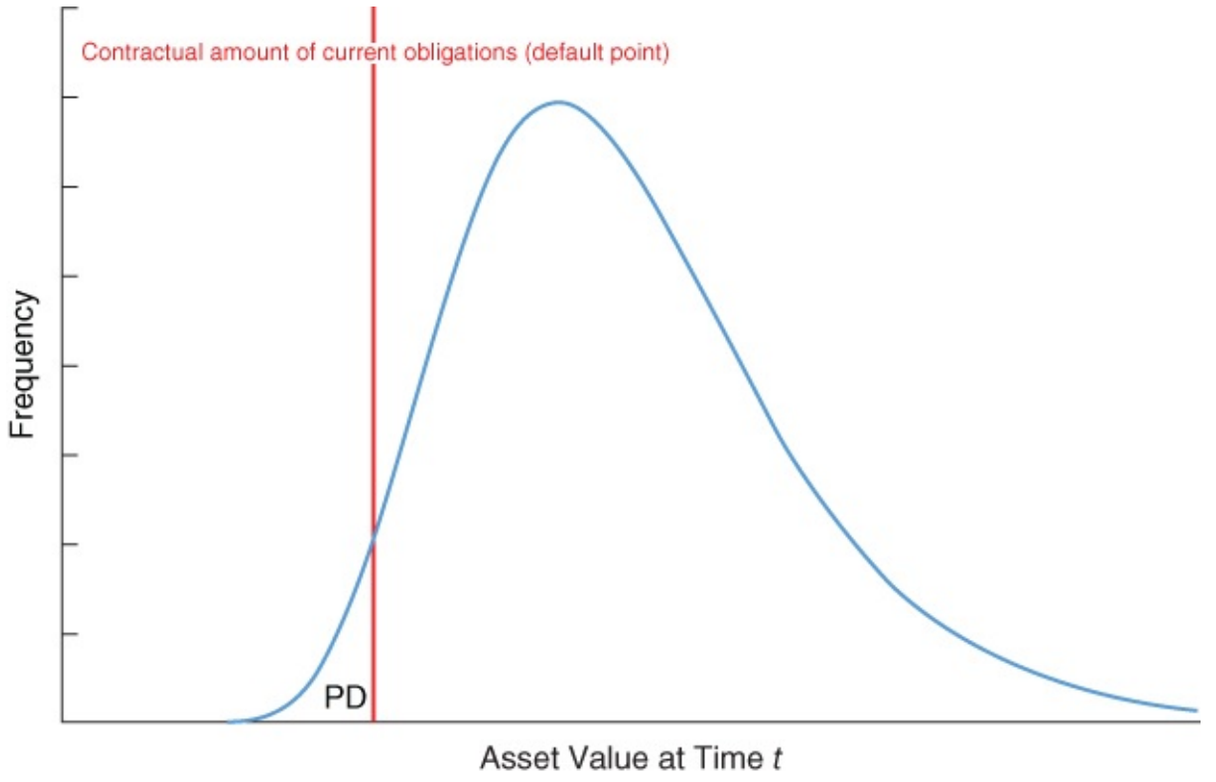

Exhibit 6.2 Merton Model

The asset value is often assumed to follow a lognormal distribution, and the asset return a normal distribution. We define $\log(x)$ as the natural logarithm of x and $\exp(x)$ as e raised to the power of x throughout this chapter. The standardized asset return A_{it} of borrower i in time period t can be modeled as a latent process. A default event occurs if the asset return A_{it} falls below a threshold γ_{it}. The probability of default (PD) then becomes:

$$PD_{it} = P(D_{it} = 1) = P(A_{it} < \gamma_{it}) = \Phi(\gamma_{it})$$

with Φ the cumulative density function of the standard normal distribution. This simple credit

model results in a probit model with the linear predictor γ_{it}. Similar models may be constructed for other distributions and other risk segments such as consumer loans. The latter include a consumer's asset value or credit score and a threshold, which if passed results in a default event. For example, Hamerle, Liebig, and Scheule (2006) derive a logit regression model using an asset value return/threshold model. We will extend this model by a decomposition of the asset value returns into systematic and idiosyncratic components in our default correlation and credit portfolio risk chapter.

The observable outcome is the default event. In the banking industry, regression models have been introduced that link the observable default/nondefault outcome to information that is available at the time of the risk assessment. Credit risk modelers approximate the structural model (i.e., the comparison of asset value and debt value, also known as the distance to default) by using observable information such as the earnings, debt, liquidity of the creditor, and so on. Parameters explain the sensitivities for the impact of observable information on observableoutcomes, and a link function (e.g., a linear combination or a nonlinear link) is assumed.

Econometric methods for parameter estimation are discussed later. More specifically, we will introduce both discrete-time and continuous-time (survival) models. Discrete-time models explain the default event within a certain time period, while survival models estimate the time to default. There are close linkages between these methodologies.

Popular examples of such estimation techniques are: the unconditional mean of the default indicators for single or multiple periods, regression techniques that condition the default probabilities on observable variables (e.g., linear regressions for averages of default indicators), and nonlinear regressions for the default indicator itself. These approaches are generally based on maximum likelihood estimation techniques that maximize a theoretical likelihood for the realized dependent variable given the observed information variables and estimated parameters.

Estimated parameters are subject to parameter uncertainty. Most estimation procedures report estimates for the parameters, their standard deviation (standard errors) and covariance/correlation matrices. These are typically based on the assumption of a normal distribution with the parameter estimate being the estimated mean and the standard error being the estimated standard deviation. We will follow up on this kind of model risk in our stress testing chapter.

DISCRETE-TIME HAZARD MODELS

Probabilities of default may be modeled with discrete-time hazard models. In order to estimate such models in SAS, the data should be arranged in panel form. Note that the word “panel” needs to be interpreted with care, as discrete-time hazard models are estimated by maximizing the product over the observational likelihoods (see later discussion). This implies that discrete-time hazard models treat the observations as conditionally (i.e., given the covariates) independent. For loan exposures, every loan is observed in periodic intervals to either default,

payoff, or end of the observation period.

Every row in a panel data set represents the observation of one loan in a single period. The columns represent variables, which include information on the propensity of a borrower default, including borrower-specific variables, macroeconomic variables, origination variables, and collateral variables as well as the random outcome variable D_{it} with realization d_{it}, which is zero if a borrower does not default and one if a borrower does default.

The code and output in Exhibit 6.3 show the last three observations of three loans from our mortgage data set.

id	orig_ time	time	default_ time	payoff_ time	FICO_ orig_ time	LTV_ orig_ time	LTV_time	gdp_time	uer_time
46	19	27	0	0	581	80.0	67.5913	2.36172	4.4
46	19	28	0	0	581	80.0	68.2919	1.22917	4.6
46	19	29	1	0	581	80.0	68.8752	1.69297	4.5
47	19	25	0	0	600	80.0	66.7938	2.89914	4.7
47	19	26	0	0	600	80.0	66.9609	2.15136	4.7
47	19	27	0	1	600	80.0	67.5853	2.36172	4.4
56	-15	58	0	0	664	52.5	17.3599	2.86859	6.2
56	-15	59	0	0	664	52.5	17.2625	2.44365	5.7
56	-15	60	0	0	664	52.5	16.8980	2.83636	5.7

Exhibit 6.3 Panel Data

In our mortgage data set, the loans in the data are numbered consecutively (id = 1 to $50{,}000$) and the observation periods (time = 1 to 60) are numbered consecutively from the first observation period. Loans may have originated prior to the first observation period and may carry negative values in such cases. Examples of loan-specific variables are the FICO score at origination and the loan-to-value (LTV) ratio at origination. The FICO mortgage score is a credit score with values between 300 and 850. LTV is the ratio of the outstanding loan amount to the collateral value, and banks traditionally extend loans in the region of 80 percent. The unemployment rate is time-varying and has the same value for all loans in a given observation period.

The loans result in different outcomes over time. Loan 46 defaults in period $time = 29$, which is indicated by the binary variable default_time. Loan 47 is paid off in period $time = 27$, which is indicated by the binary variable payoff_time. Loan 56 survives until the end of our observation period (i.e., $time = 60$), and we will later refer to this as a right-censored observation (see Chapter 7).

Discrete-time hazard regression models establish a link between the probability of the binary default variable taking on a particular value (e.g., default) and the observable information through nonlinear link functions F: $P(D_{it} = 1) = F(lp_{it-1})$. Note that the models are generally formulated in terms of expectation due to the binary character of the dependent variable. Furthermore, lp_{it-1} is a linear predictor formulated in terms of $lp_{it-1} = \boldsymbol{\beta}' \boldsymbol{x}_{it-1}$, that is a linear

combination of observable information and parameters that are estimated, as discussed in what follows.

Note that the probability of default is time-varying and that the explanatory variables need to be known at the beginning of the observation period (hence the index $it-1$) to enable the models to provide out-of-time econometric forecasts (see fitting and forecasting section). Examples of papers that apply discrete-time hazard models to estimate borrower-specific default probabilities are Hamerle et al. (2006), Crook, Edelman, and Thomas (2007), Crook and Bellotti (2010), and Leow and Mues (2012).

In the following section, we discuss four discrete-time hazard models: the linear model, the logit model, the probit model, and the complementary log-log (cloglog) model. In the industry, the logit and probit models are the most popular.

Linear Model

Linear models are generally estimated with an ordinary least squares (OLS) technique in which the distances between the predicted outcome and realized outcome are squared and summed over all observations, and the parameters arechosen such that this deviation measure is minimized. The model equation is:

$$D_{it} = \boldsymbol{\beta}' \boldsymbol{x}_{it-1} + \epsilon_{it}$$

with a vector of appropriate covariates $\boldsymbol{x}_{it-1}$, the corresponding sensitivities $\boldsymbol{\beta}$ and ϵ_{it}, which is normally distributed with mean zero and standard deviation σ. The models can be estimated in SAS with PROC REG:

```
PROC REG DATA=data.mortgage;
MODEL default_time = FICO_orig_time
LTV_orig_time gdp_time;
RUN;
```

We do not show all output tables for the various procedures but limit our presentation to the ODS tables of interest throughout the book. Exhibit 6.4 shows the parameter estimates and *R*-squared measures (adjusted and nonadjusted) for model accuracy.

The REG Procedure
Model: MODEL1
Dependent Variable: default_time

Root MSE	0.15349	*R*-Square	0.0083
Dependent Mean	0.02435	Adj *R*-Sq	0.0083
Coeff Var	630.33859		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	0.07957	0.00259	30.71	<.0001
FICO_orig_time	1	−0.00011544	0.00000275	−42.02	<.0001
LTV_orig_time	1	0.00038077	0.00001944	19.59	<.0001
gdp_time	1	−0.00545	0.00009914	−55.02	<.0001

Exhibit 6.4 Linear Model

The estimate for σ is given by root mean squared error (MSE). No additional link function is applied in a linear model, which implies that the parameter estimates can be interpreted in terms of the dependent variable. For example, the impact of a change in the FICO score can be assessed by multiplying the change by the parameter estimate. We have set up the data such that the covariates refer to the time stamp, and the dependent variable (here default_time) refers to the period that follows the observation of the covariates. In other words, the dependent variable and the covariates are time lagged ($t-1$). Various time lags may be chosen for all or a set of selected covariates.

The sign of the parameters indicates the directional impact on the default probability. The FICO score and GDP growth have negative signs, which implies that the default probability decreases with a higher FICO score or GDP growth. The LTV ratio has a positive sign, which implies that the default probability increases with a higher LTV ratio.

The linear model is not popular for estimating borrower/loan-level default probabilities, as the predicted probabilities of default are not constrained within the range of the defined default probabilities of zero and one. This has the caveat that in some instances probabilities of default that are negative or above one may result. These values are unreasonable and often cause problems in follow-up applications such as the computation of regulatory capital or loan pricing.

Note that the model fit as indicated by the *R*-squared is low. The low *R*-squared is common for binary models and a reflection of the data-generating process: The realization of default events is based on a random binary variable and the default probability as event likelihood. The random experiment explains the low fit, and better results are achieved when default rates (rather than default events) are modeled. As a matter of fact, linear models and extensions thereof are very popular to model default rates for risk segments and time periods. The reason for this popularity is that the estimated parameters can now be interpreted in terms of the

default probability, and advanced econometric modeling techniques, such as state space models, require a linear link.

Furthermore, a nonlinear transformation for the dependent variable default rate is often chosen. Popular transformation functions are inverse cumulative density functions such as the logit and probit functions. The transformation in essence increases the range of values to the range of possible outcomes (i.e., from minus infinity to plus infinity) to match the range of possible values for the sum of the linear predictor and residual terms.

Nonlinear Models: Probit, Logit, and Cloglog Models

Specifying the Link Function

In order to constrain the dependent variable to the defined range of zero to one, nonlinear link functions can be applied. The nonlinear link functions are generally inverse cumulative distribution functions of well-known distributions. The functions that we discuss in this and the following subsections are the probit, the logit, and the complementary log-log (cloglog) functions. These functions map the linear predictor into a default probability, which is shown in Exhibit 6.5, where the linear predictor is shown on the x-axis, and the default probability, which is bounded between zero and one, is shown on the y-axis.

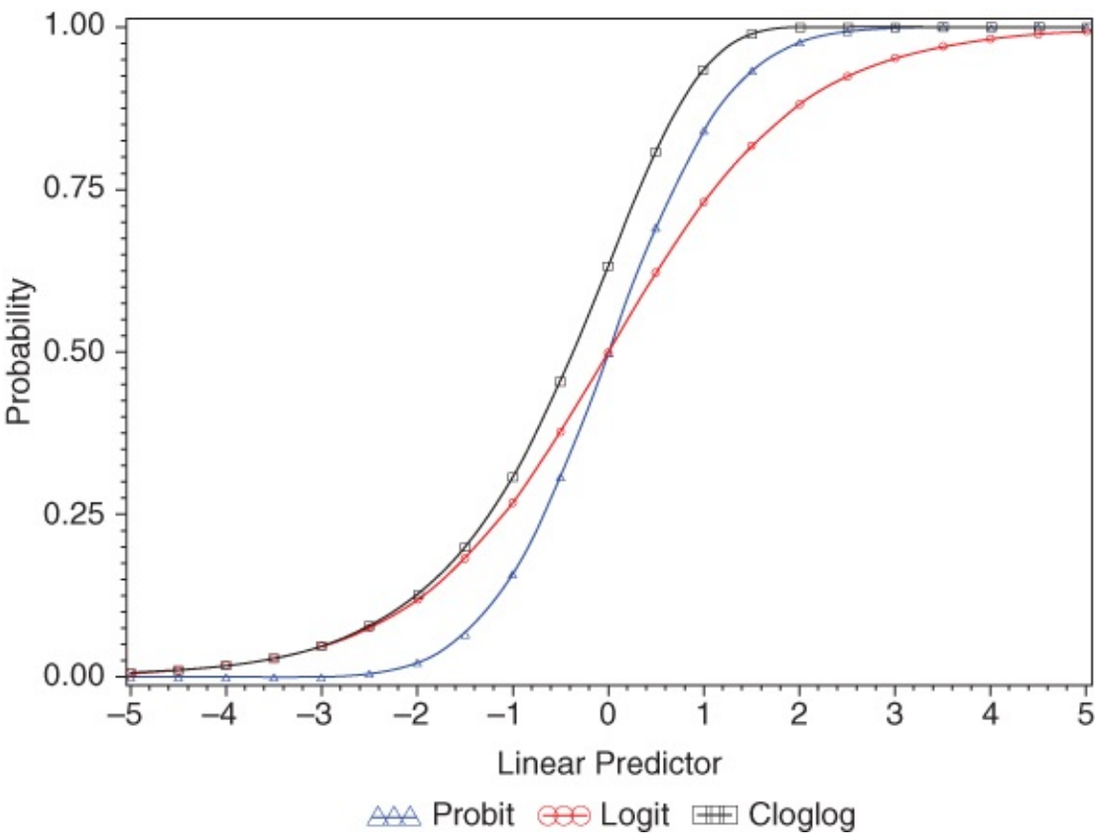

Exhibit 6.5 Nonlinear Link Functions

Creating Charts in SAS

Most credit risk analysts either export the data from SAS to other programs, in particular

Excel, or directly draw the graphs in SAS. The graph in Exhibit 6.5 can be generated by the following code:

```
DATA graph;
DO x=-5 TO 5 BY 0.5;
logit=cdf('LOGISTIC',x);
probit=cdf('NORMAL',x);
cloglog=1-EXP(-EXP(x));
OUTPUT;
END;
RUN;
ODS GRAPHICS ON;
AXIS1 ORDER=(-5 to 5 BY 1) LABEL=('Linear predictor');
AXIS2 ORDER=(0 to 1 BY 0.25) LABEL=('Probability');
SYMBOL1 INTERPOL=SPLINE WIDTH=2 VALUE=TRIANGLE C=BLUE;
SYMBOL2 INTERPOL=SPLINE WIDTH=2 VALUE=CIRCLE C=RED;
SYMBOL3 INTERPOL=SPLINE WIDTH=2 VALUE=SQUARE C=BLACK;
LEGEND1 LABEL=NONE SHAPE=SYMBOL(4,2)
POSITION=(BOTTOM OUTSIDE);
PROC GPLOT DATA=test; PLOT logit*x probit*x cloglog*x
/ OVERLAY LEGEND=LEGEND1 HAXIS=AXIS1 VAXIS=AXIS2;
RUN;
ODS GRAPHICS OFF;
```

The shapes of the three link functions are comparable, yet distinct. Research has shown that different link functions result in different parameter estimates but the predicted values, here the probabilities of default, are comparable (see, e.g., Hamerle et al. 2006). Hence, practitioners may choose any of the presented link functions.

Probit Model

In what follows, we compare three models with different link functions: the logit model, the probit model, and the cloglog model. As we have already introduced the logit model in the credit scoring chapter, we now focus on the probit model, which can be estimated using PROC LOGISTIC or PROC PROBIT. We prefer PROC LOGISTIC as it has exactly the same syntax for all link functions and the link function can easily be changed for robustness checks.

As said, we generally find that the link function is of minor importance but different functions have important applications. For example, the logit link function has useful properties if the resulting mean PDs have to be calibrated to a different level. The probit link function is particularly useful for estimating parameters that are in line with the internal ratings based (IRB) models under the Basel regulations, as these assume the probit link.

The probit model equation is:

$$P(D_{it} = 1 | \boldsymbol{X}_{it-1}) = \Phi(\boldsymbol{\beta}' \boldsymbol{x}_{it-1})$$

with $\Phi(.)$ the cumulative density function of the standard normal distribution, $\boldsymbol{\beta}$ a vector of sensitivity parameters, and $\boldsymbol{x}_{it-1}$ a vector of time-lagged (with regard to the observable default event) covariates (i.e., risk factors). The following code estimates a logit model:

```
PROC LOGISTIC DATA=data.mortgage DESCENDING;
MODEL default_time =FICO_orig_time LTV_time gdp_time
/ LINK=PROBIT RSQUARE;
OUTPUT OUT=probabilities PREDICTED=PD_time;
RUN;
```

The PROC LOGISTIC statement invokes the procedure that estimates the probit, logit, as well as a range of other models. By default, SAS sorts the dependent variable from low to high and models the first category. The option DESCENDING reverses the default order and specifies that the probability of $y_{it} = 1$ is modeled.

The OUTPUT statement in combination with the PREDICTED= statement requests that the default probabilities are calculated for the estimation sample and stored in a separate variable in the input data set and also saved in the OUT= location. Alternatively, a STORE statement may be added that requests that the procedure parameter estimates are stored in the file specified by the name associated with the OUT= statement. This file cannot be modified but can be used for estimation and prediction of default probabilities, which we explain in detail later in the book. The contents of the item store can be processed with the PROC PLM procedure, which we will explain in more detail later in this chapter.

The Model Information section of the output (Exhibit 6.6) lists the reference data set, the dependent variable, and the number of response categories. Our dependent variable is default-time, which is binary and coded 0 for nondefault and 1 for default (other codes such as "Yes"/"No" are also possible). We use the suffix_time to indicate that the variable is time-varying. We use the suffix_orig_time to indicate that the variable is collected at loan origination and hence is constant over time. Furthermore, the number of observations and a breakdown by category of the dependent variable are provided. These are important descriptive statistics that are often included in risk reports.

The LOGISTIC Procedure

Model Information	
Data Set	DATA.MORTGAGE
Response Variable	default_time
Number of Response Levels	2
Model	Binary probit
Optimization Technique	Fisher's scoring

Number of Observations Read	622489
Number of Observations Used	622219

Response Profile		
Ordered Value	default_time	Total Frequency
1	1	15153
2	0	607066

note	Probability modeled is default_time=1.

Model Convergence Status
Convergence criterion (GCONV=1E−8) satisfied.

Exhibit 6.6 Probit Model

Maximum Likelihood Estimation

In SAS, the Fisher's scoring method is used as a default method of iteratively estimating the regression parameters via maximization of the log-likelihood. The likelihood is defined as follows:

$$L(\boldsymbol{\beta}, \boldsymbol{x}_{it-1}) = \prod_{i=1}^{I} \prod_{t=1}^{T} \left((P(D_{it} = 1)^{d_{it}} \cdot (1 - P(D_{it} = 1))^{(1-d_{it})} \right)$$

with $P(d_{it} = 1) = \Phi(\boldsymbol{\beta}' \boldsymbol{x}_{it-1})$. The likelihood function in essence multiplies all event probabilities by the probability for a default event of $P(d_{it} = 1)$ and the probability for a nondefault event of $1 - P(d_{it} = 1)$. The likelihood is then transformed by the monotone natural logarithm (i.e., log-likelihood):

$$\log\ L(\boldsymbol{\beta}, \boldsymbol{x}_{it-1}) = \sum_{i=1}^{I} \sum_{t=1}^{T} (d_{it} \log\ (P(D_{it} = 1)) + (1 - d_{it}) \log\ (1 - P(D_{it} = 1)))$$

The SAS algorithm then maximizes the natural logarithm, as the latter maintains monotonicity and results in the same parameter estimates with the merit that the sum of log-likelihoods is

computationally simpler than the product of likelihoods.

Note that other optimization techniques may result in different standard errors. The Number of Observations Read and Number of Observations Used are the number of observations in the data set and the number of observations used in the analysis. The latter may be lower if values for the dependent or independent variables can not be processed, or if they are missing.

The Response Profile shows that the probability of a default event is modeled as we have applied the DESCENDING command, which sorts the dependent variable default_time in descending order (high to low). This implies that a positive parameter estimate (note that the logistic link function is monotone) increases the default probability. SAS models by default the probability of the lowest category, here the probability of a nondefault event, which is coded as zero. SAS confirms this in the output file with the sentence "Probability modeled is […]."

The next output section (Exhibit 6.7) includes the likelihood-based performance measures.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	142525.57	134965.28
SC	142536.91	135010.64
−2 Log L	142523.57	134957.28

R-Square	0.0121	Max-rescaled *R*-Square	0.0590

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7566.2941	3	<.0001
Score	7861.2257	3	<.0001
Wald	6854.6897	3	<.0001

Exhibit 6.7 Probit Model (cont.)

The model fit statistics are measures for model fit based on −2 times the log-likelihood (−2LogL). Both the Akaike information criterion (AIC) and the Schwartz criterion (SC) are based on −2 times the log-likelihood and include a penalty for the number of estimated parameters. A lower AIC/SC/−2LogL indicates a better fit, which holds for nested models. These measures are absolute measures, which depend on the sample size. In other words, these measures are not suitable to compare models based on different sample sizes, which may be a result of the availability of dependent and independent variables (see the previous comment on Number of Observations Used).

We have also requested the generalized R-squared measure (i.e., the likelihood-based pseudo R-squared measure and its rescaled variant) for the model by using the RSQUARE option after the model statement. The measure is a relative performance measure, as it includes a

comparison with a noninformative model that assigns all default observations the same average default rate and is defined between zero and one. A model with a higher R-squared value dominates models with lower R-squared values. Note that relative performance measures share the same critique as absolute performance measures and should be used with great care when models are estimated using different sample sizes (see Hamerle, Rauhmeier, and Rösch 2003, and the validation chapter).

The Global Null Hypothesis test that is included tests whether all parameter estimates are jointly equal to zero. As we primarily deal with large data, this hypothesis is generally rejected, implying that at least one explanatory variable is significant.

The next output section (Exhibit 6.8) includes the actual parameter estimates and hence the meat of our model analysis.

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	−1.0091	0.0354	814.0574	<.0001
FICO_orig_time	1	−0.00242	0.000050	2366.9776	<.0001
LTV_time	1	0.00781	0.000158	2447.5807	<.0001
gdp_time	1	−0.0496	0.00167	883.8005	<.0001

Exhibit 6.8 Probit Model (cont.)

The column Estimate shows the parameter estimates and mean of the normal parameter distribution. Standard Error is the estimated standard deviation, and Wald Chi-Square is the test statistic for a test that verifies whether the parameter is equal to zero (i.e., the explanatory variable has no impact on the default behavior). A higher value generally indicates greater significance. The information is also included in the p-value, which is labeled Pr $>$ ChiSq. An explanatory variable is considered to be significant if the p-value is less than or equal to a chosen significance level (often termed α), which effectively limits the type 1 error rate.

The sensitivities have to be interpreted in terms of the linear predictor. The sensitivity of the FICO score is -0.00242 and implies that the linear predictor and hence the PD decreases with the FICO score. A FICO score which is 100 points greater results in a linear predictor that is -0.242 lower ($= -0.00242 * 100$). The interpretation for the LTV ratio and the GDP growth follows analogously.

The last output table (Exhibit 6.9) includes statistics that are of great interest to credit risk analysts.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	69.5	Somers' D	0.422
Percent Discordant	27.3	Gamma	0.436
Percent Tied	3.2	Tau-a	0.020
Pairs	9198871098	c	0.711

Exhibit 6.9 Probit Model (cont.)

Somers' D is the accuracy ratio (AR) and c is the area under the ROC curve (AUROC). Both measures can be transformed into one another: AR = 2 * AUROC − 1. The accuracy ratio is sometimes also referred to as the Gini coefficient. We will discuss these measures in more detail in the validation chapter.

Estimation of Default Probabilities

The PROC LOGISTIC has also generated a data set called probabilities with the OUTPUT command. The command adds an additional column (named PD_time) to the input data set. In other words, a default probability has been estimated for every observation based on the

- Model assumptions
- Estimated model parameters
- Historical data that was used to estimate the parameters

Calibration of Probit Models

An important consideration in estimating default probabilities is their calibration to the default rates. Discrete-time hazard models by definition are calibrated (see later analysis): The mean of the estimated default probabilities matches the default rate of the estimation sample. The intercept parameter captures any baseline risk that is not attributable to the explanatory variables.

Calibration of default rates is a “hard” requirement in building Basel-compliant credit risk models and models that do not meet this standard are not approved for capital allocation, loan loss provisioning, and stress testing. Furthermore, banks are unlikely to consider such models for internal economic purposes such as loan pricing. Note that because of this, the industry has hesitated to embrace risk-neutral credit risk models based on share prices. We test the calibration of the logistic model by comparing the mean estimated default probabilities with the default rate in the sample.

We run PROC MEANS for the default indicator and default probabilities and find that the mean of the in-sample PD estimates matches the default rate. Furthermore, the dispersion of the estimated default probabilities (i.e., the difference between the maximum and minimum estimates) is an indication of the ability of a PD model to predict default. This ability is highest if PDs with a value of zero are assigned to nondefault events and PDs with a value of one are assigned to default events. Note that this measure is also data dependent and the discriminatory power increases with the degree of discrimination in the data-generating process. We have

estimated the in-sample default probabilities with the OUTPUT statement, which has evaluated the model equation and estimated parameters (indicated by a hat) as follows:

$$\widehat{P}(D_{it} = 1|\boldsymbol{X}_{it-1}) = \Phi(\widehat{\beta}_0 + \widehat{\beta}_1 * \text{FICO\_orig\_time} + \widehat{\beta}_2 * \text{LTV\_time} + \widehat{\beta}_3 * \text{gdp\_time})$$

with $\widehat{\beta}_0$ to $\widehat{\beta}_3$ the estimated parameters. A PROC MEANS provides the mean for the default indicators and the estimated PDs (Exhibit 6.10).

The MEANS Procedure

Variable	Mean	Variable	Mean
default_time	0.0243506	PD_time	0.0242548

Exhibit 6.10 Calibration of Probit Models: Comparison of Default Indicators and Estimated Default Probabilities

```
PROC MEANS DATA=probabilities MEAN NOLABELS;
VAR default_time PD_time;
RUN;
```

The calibration is clear, as the mean of the default event almost matches the mean of the estimated PD. The minor difference that we observe is due to the estimation algorithm that iteratively maximizes the likelihood and stops if a target function indicates a low model improvement.

We include a more detailed discussion of the estimation and forecasting of default probabilities using the data set generated by the STORE command and PROC PLM later in the book.

Logit Model

As mentioned in the credit scoring chapter, a logit model is the default option of PROC LOGISTIC and hence does not require the explicit specification of a link function:

```
PROC LOGISTIC DATA=data.mortgage DESCENDING;
MODEL default_time = FICO_orig_time LTV_time gdp_time;
RUN;
```

The parameter estimates have the same signs as the probit model but a different magnitude (see Exhibit 6.11).

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	−1.6523	0.0832	394.8522	<.0001
FICO_orig_time	1	−0.00540	0.000116	2156.7605	<.0001
LTV_time	1	0.0184	0.000372	2446.3140	<.0001
gdp_time	1	−0.1094	0.00375	851.1896	<.0001

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	69.4	Somers' D	0.423
Percent Discordant	27.1	Gamma	0.438
Percent Tied	3.5	Tau-a	0.020
Pairs	9198871098	c	0.711

Exhibit 6.11 Logit Model

The model fit as indicated by the accuracy ratio (Somers' D) and AUROC (c) is comparable to the probit model.

Cloglog Model

The following code estimates a cloglog model:

```
PROC LOGISTIC DATA=data.mortgage DESCENDING;
MODEL default_time = FICO_orig_time
LTV_time gdp_time/ LINK=CLOGLOG;
RUN;
```

Again, the parameter estimates have the same signs as the probit and logit models, but different magnitudes (see Exhibit 6.12). We will show that the combination of link function, parameter estimates, and covariates will result in very similar estimates for the probability of default, regardless of the choice of link function. The model fit as indicated by the accuracy ratio (Somers' D) and AUROC (c) is lower than the one by the probit and logit models.

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	−0.8041	0.0761	111.6063	<.0001
FICO_orig_time	1	−0.00535	0.000112	2268.0798	<.0001
LTV_time	1	0.00885	0.000283	980.3850	<.0001
gdp_time	1	−0.1415	0.00355	1591.2030	<.0001

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	67.7	Somers' D	0.396
Percent Discordant	28.1	Gamma	0.413
Percent Tied	4.2	Tau-a	0.019
Pairs	9198871098	c	0.698

Exhibit 6.12 Cloglog Model

Applications

Qualitative Information

Another important feature is the CLASS statement that enables the coding of categorical variables. It is common to control for origination periods (also known as vintages) and other qualitative variables by means of appropriate coding. Popular approaches include reference/dummy coding and effect coding. For both approaches, $K-1$ new variables are created if the categorical variable has K categories.

Reference Coding

In reference coding, the respective variable is coded one if the category is given, and zero otherwise:

$$C_k = \begin{cases} 1 & \text{category } k \text{ is given} \\ 0 & \text{otherwise} \end{cases}$$

with $k = 1, \cdots K-1$. The interpretation of the parameter estimates is in terms of the linear predictor where category k is given relative to the linear predictor where the reference category K is given.

Effect Coding

In effect coding, the respective variable is coded one if the category is given and minus one if the reference category is given:

$$C_k = \begin{cases} 1 & \text{category } k \text{ is given} \\ -1 & \text{reference category } K \text{ is given} \\ 0 & \text{otherwise} \end{cases}$$

with $k = 1,\ldots K-1$.

Controlling for Categorical Information in SAS

A popular categorical variable is the origination year, which is also known as the vintage. The data set has many vintages with different observation counts. To simplify the analysis, we restrict the number of categories to vintages with sufficient observation counts by generating a new variable called orig_time2:

```
data mortgage;
SET data.mortgage;
orig_time2=orig_time;
IF orig_time NOT IN (20,21,22,23,24,25) THEN orig_time2 =0;
RUN;
```

We then include orig_time2 and specify dummy coding (PARAM=REFERENCE, effect coding would be PARAM=EFFECT):

```
PROC LOGISTIC DATA=mortgage DESCENDING;
CLASS orig_time2/PARAM=REFERENCE;
MODEL default_time = FICO_orig_time
LTV_time gdp_time orig_time2 / LINK=PROBIT;
RUN;
```

The parameter estimates are reported in Exhibit 6.13. Note that the accuracy ratio has slightly increased as we add more meaningful variables to the model. We encourage the reader to identify and add more risk factors.

The LOGISTIC Procedure

Class Level Information							
Class	Value	Design Variables					
orig_time2	0	1	0	0	0	0	0
	20	0	1	0	0	0	0
	21	0	0	1	0	0	0
	22	0	0	0	1	0	0
	23	0	0	0	0	1	0
	24	0	0	0	0	0	1
	25	0	0	0	0	0	0

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	−0.9670	0.0379	649.8974	<.0001
FICO_orig_time		1	−0.00241	0.000050	2329.3365	<.0001
LTV_time		1	0.00770	0.000162	2260.7985	<.0001
gdp_time		1	−0.0495	0.00167	877.8470	<.0001
orig_time2	0	1	−0.0365	0.0116	9.9937	0.0016
orig_time2	20	1	−0.1176	0.0213	30.4655	<.0001
orig_time2	21	1	−0.0528	0.0185	8.1113	0.0044
orig_time2	22	1	−0.0646	0.0169	14.6329	0.0001
orig_time2	23	1	−0.0101	0.0171	0.3475	0.5556
orig_time2	24	1	−0.0353	0.0164	4.6079	0.0318

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	69.5	Somers' D	0.423
Percent Discordant	27.3	Gamma	0.437
Percent Tied	3.2	Tau-a	0.020
Pairs	9198871098	c	0.711

Exhibit 6.13 Probit Model with Categorical Covariates

Through-the-Cycle (TTC) versus Point-in-Time (PIT)

We now analyze default probabilities generated by different modeling methodologies. Through-the-cycle (TTC) models generally abstract from the state of the overall economy by excluding macroeconomic risk drivers. Point-in-time (PIT) models explicitly control for the state of the economy. In reality, the distinction is not so easy, as many borrower-specific (i.e., idiosyncratic) risk factors are time-varying and often correlated with the economy. Various alternative definitions may be found in the literature for TTC and PIT models that are not

considered here.

For a first comparison, we estimate two models: a TTC model and a PIT model.

The TTC model is based on application data (i.e., information that is observable at loan origination). The model is a logit model based on the FICO score at origination and LTV ratio at origination:

```
PROC LOGISTIC DATA=data.mortgage DESCENDING;
MODEL default_time = FICO_orig_time
LTV_orig_time / LINK=probit;
OUTPUT OUT=probabilities PREDICTED=PD_TTC_time;
RUN;
```

The PIT model is based on application data and time-varying information. It is a logit model based on the FICO score at origination and the LTV ratio at observation time, as well as the following macroeconomic indicators: (1) GDP growth rate, (2) unemployment rate, and (3) house price index.

```
PROC LOGISTIC DATA=probabilities DESCENDING;
MODEL default_time = FICO_orig_time
LTV_time gdp_time uer_time hpi_time / LINK=PROBIT;
OUTPUT OUT=probabilities2 PREDICTED=PD_PIT_time;
RUN;
```

Exhibit 6.14 compares the default rate and the mean estimated default probability for the TTC model and the PIT model over time. The chart is also known as an in-sample real-fit diagram.

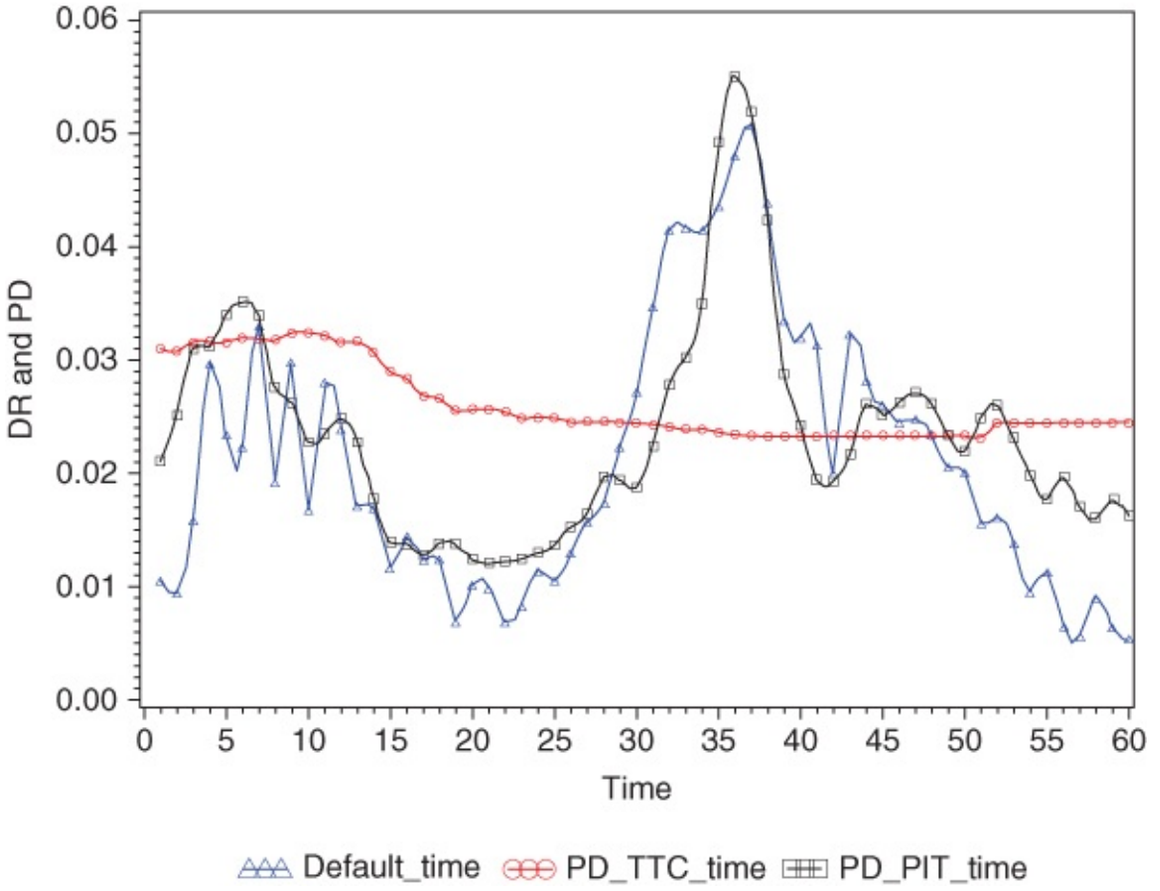

Exhibit 6.14 Real-fit diagram for the TTC Probit Model and the PIT Probit Model

```
DATA means;
SET means;
LABEL PD_TTC_time="PD_TTC_time";
LABEL PD_PIT_time="PD_PIT_time";
RUN;
ODS GRAPHICS ON;
AXIS1 ORDER=(0 to 60 by 5) LABEL=('Time');
AXIS2 order=(0 to 0.06 by 0.01) LABEL=('DR and PD');
SYMBOL1 INTERPOL=SPLINE WIDTH=2 VALUE=TRIANGLE C=BLUE;
SYMBOL2 INTERPOL=SPLINE WIDTH=2 VALUE=CIRCLE C=RED;
SYMBOL3 INTERPOL=SPLINE WIDTH=2 VALUE=SQUARE C=BLACK;
LEGEND1 LABEL=NONE SHAPE=SYMBOL(4,2) POSITION=(bottom outside);
PROC GPLOT DATA=means;
PLOT (default_time PD_TTC_time PD_PIT_time)*time
/ OVERLAY HAXIS=AXIS1 VAXIS=AXIS2 LEGEND=LEGEND1;
RUN;
ODS GRAPHICS OFF;
```

The TTC model produces default probabilities that are mostly time-invariant. The only sources for time variation are changes in the loan population. For example, if the fraction of low-risk borrowers increases over time, then the model-implied mean default probability decreases; if the fraction of low-risk borrowers decreases, then the default probability increases. A hybrid TTC model may include time-varying borrower information that is correlated with the economy and may show a moderate time variation.

The PIT model includes the time-varying LTV ratio and a number of macroeconomic factors. Its mean default probability is closer to the realized default rate. From an econometric viewpoint, PIT models are closer to the data-generating process and hence more accurate. For examples in the literature, see Rösch and Scheule (2010) and Rösch and Scheule (2004).

Controversy Regarding Procyclical Risk Measures

The use of macroeconomic variables is sometimes of concern to prudential regulators as their use leads to time-varying risk measures and hence time-varying loan loss provisioning and regulatory capital requirements. Banks often find it challenging to raise capital during economic downturns, and the use of PIT models leads to procyclical capital requirements that may further exacerbate economic downturns. To avoid such situations, regulators may prohibit the use of macroeconomic indicators, mandate forward-looking (i.e., anticipating) risk measures, or offset the impact of procyclical capital requirements through other actions (e.g., countercyclical capital buffers).

Note that PIT models are more accurate than TTC models as they control for important (time-varying) covariates and should be considered for the measurement of economic risks.

Estimation of Rating Migration Probabilities

It is common in banking to estimate rating migration probabilities (often displayed in rating migration matrices) to simplify rating analytics. As with the case of default probabilities, multiyear rating migration probabilities can be computed by means of matrix multiplication.

We first define the rating classes for the current period $k = 1, 2, \ldots K$ and the past period $k^* = 1, 2, \ldots K$. The last rating class K may be defined as the terminating default event, in which case the number of past rating classes reduces to $k^* = 1, 2, \ldots K - 1$ as only borrowers that did not default in the prior period can be observed.

We prepare the data set by categorizing the FICO score to derive rating classes. Various categorization techniques are discussed in the model validation chapter. We generate three rating classes and assume that the default event constitutes a fourth rating class:

```
PROC SORT DATA=data.mortgage OUT=mortgage;
BY id time;
RUN;
DATA mortgage;
SET mortgage;
IF FICO_orig_time>350 AND FICO_orig_time<=500 THEN rating=1;
IF FICO_orig_time>500 AND FICO_orig_time<=650 THEN rating=2;
IF FICO_orig_time>650 AND FICO_orig_time<850 THEN rating=3;
lagid=LAG(id);
lagrating=LAG(rating);
RUN;
DATA mortgage;
SET mortgage;
IF id NE lagid THEN lagrating=.;
IF default_time=1 THEN rating=4;
RUN;
```

The example should be read with care. Rating migration models do not generally provide for more accurate models than a loan-level PD model. PD models have an infinite number of possible values, while the aggregation of PDs or any metric score to rating classes implies a loss of information. However, we acknowledge that situations may exist within particular risk applications of a bank where ratings classes and rating migration probabilities provide added value.

Rating Matrix from Observed Migration Frequencies

A rating matrix tabulates the relative frequencies or estimated probabilities of migrating from a rating at the beginning of the period (generally indicated by the first column) to a rating including the terminal default state at the end of the period (generally indicated by the first row). It can be easily calculated using an actuarial method by PROC FREQ:

```
PROC FREQ DATA=mortgage(WHERE=(rating NE . AND lagrating NE .));
TABLES lagrating*rating /NOCOL NOPERCENT NOCUM;
RUN;
```

We cross-tabulate the rating at the beginning of the period and the rating at the end of the period. We are interested in the conditional rating migration probabilities (i.e., the rating migration probabilities conditional on a particular rating at the beginning of the period). Hence, we suppress the display of column percentages, percentages, and cumulative frequencies using the NOCOL, NOPERCENT, and NOCUM options. The resulting table shows the absolute and relative frequencies for the rating migrations (see Exhibit 6.15). Relative frequencies are conditional on the rating at the beginning of the period. As a result, all relative frequencies for a given rating at the beginning of the period add up to one.

The FREQ Procedure

Table of Lagrating by Rating

Lagrating / Rating Frequency Row Pct	1	2	3	4	Total
1	2731 96.06	0 0.00	0 0.00	112 3.94	2843
2	5 0.00	193070 96.50	15 0.01	6978 3.49	200068
3	3 0.00	50 0.01	362080 97.97	7445 2.01	369578
Total	2739	193120	362095	14535	572489

Exhibit 6.15 Rating Migration Matrix Based on Observed Migration Frequencies

As default is a terminating state, ratings cannot be observed at the end of a period if the loan has defaulted at the beginning of the period. Therefore, the number of rows of the rating migration matrix (i.e., the rating at the beginning of the observation period) is one less than the number of columns (i.e., the rating at the end of the observation period).

Cumulative Probit Model

Discrete-time hazard models may be extended by cumulative probit models (also known as ordered probit models) to accommodate the probability of migrating from one state to another. Comparable extensions exist for continuous-time models such as a competing hazard model. We discuss these in the exposure at default (EAD) modeling chapter. We model the probability of migrating to rating R_{it} conditional on the realized rating of the previous period r_{it-1}:

$$\begin{cases} P(R_{it} = 1 | r_{it-1} = k^*) = \Phi(\theta_1 + \boldsymbol{\beta}' \boldsymbol{X}_{it-1}) \\ \ldots \\ P(R_{it} = k | r_{it-1} = k^*) = \Phi(\theta_k + \boldsymbol{\beta} \boldsymbol{X}_{it-1}) - \Phi(\theta_{k-1} + \boldsymbol{\beta} \boldsymbol{X}_{it-1}) \\ \ldots \\ P(R_{it} = K | r_{it-1} = k^*) = 1 - \Phi(\theta_{K-1} + \boldsymbol{\beta} \boldsymbol{X}_{it-1}) \end{cases}$$

The parameters are determined by estimating one model for every starting rating class and maximizing the natural logarithm of the following likelihood:

$$L = \prod_{k=1}^{K} \prod_{i=1}^{I} \prod_{t=1}^{T} P(R_{it} = k | r_{it-1} = k^*)^{I(R_{it}=k)}$$

The indicator variable $I(R_{it} = k)$ equals one if R_{it} is equal to k and zero otherwise. We include the rating at the beginning of the period as a categorical control variable using PROC LOGISTIC with the CLASS statement.

```
PROC LOGISTIC DATA = mortgage;
CLASS lagrating(REF='3');
MODEL rating = lagrating /LINK=PROBIT;
OUTPUT OUT=probabilities PREDICTED=PROB_CUM;
RUN;
```

The code generates an output data set that includes the cumulative probability for rating migration from the rating at the beginning of the period to the ratings one to three. Note that the cumulative probability for the fourth rating class is per definition one.

The parameter estimates include the threshold estimates and the parameter estimate for the first and second class (see Exhibit 6.16).

The LOGISTIC Procedure

Model Information	
Data Set	WORK.MORTGAGE
Response Variable	Rating
Number of Response Levels	4
Model	Cumulative probit
Optimization Technique	Fisher's scoring

Class Level Information			
Class	Value	Design Variables	
Lagrating	1	1	0
	2	0	1
	3	−1	−1

c

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	−2.7377	0.0151	32904.2195	<.0001
Intercept	2	1	1.5819	0.0111	20178.6211	<.0001
Intercept	3	1	4.8703	0.0118	171445.072	<.0001
Lagrating	1	1	3.5768	0.0216	27410.2541	<.0001
Lagrating	2	1	−0.2444	0.0109	500.0602	<.0001

Exhibit 6.16 Cumulative Probit Model for Rating Migration Probabilities

The rating migration matrix can be derived from the cumulative probabilities generated by the OUTPUT command in the data set probabilities in a number of steps. First, we add rows for the rating class 4, which was not included in the output data set, as it is the reference category in the model:

```
DATA rating4;
INPUT lagrating _LEVEL_;
DATALINES;
1 4
2 4
3 4
;
RUN;
DATA probabilities2;
SET probabilities rating4;
RUN;
```

Second, we create a new data set that includes one observation per unique combination of lagrating and rating:

```
PROC SORT DATA=probabilities2 OUT=probabilities3(WHERE=(lagrating NE .))
NODUPKEY;
BY lagrating _LEVEL_;
RUN;
```

Third, we compute the cumulative migration probabilities for all ratings at the beginning of the observation and the marginal migration probabilities for all ratings at the beginning of the observation:

```
DATA probabilities3(KEEP =lagrating rating probability);
SET probabilities3;
lagprob_cum=LAG(prob_cum);
IF _LEVEL_=1 THEN probability=prob_cum;
IF _LEVEL_=2 THEN probability=prob_cum-lagprob_cum;
IF _LEVEL_=3 THEN probability=prob_cum-lagprob_cum;
IF _LEVEL_=4 THEN probability=1-lagprob_cum;
rating=_LEVEL_;
RUN;
```

Fourth, we transpose the data set to match the format of a rating migration matrix:

```
PROC TRANSPOSE DATA=probabilities3 OUT=probabilities4(DROP=_NAME_)
PREFIX=rating;
BY lagrating;
ID rating;
RUN;
```

The resulting data set, which includes the rating matrix, is presented in Exhibit 6.17.

Obs	Lagrating	rating1	rating2	rating3	rating4
1	1	0.79930	0.20070	0.00000	0.000000
2	2	0.00143	0.90804	0.09052	0.000002
3	3	0.00000	0.04001	0.89795	0.062039

Exhibit 6.17 Rating Migration Matrix

This rating migration matrix is different from the empirical one presented earlier, as we have made some simplifying assumptions with regard to the parameter estimates. More specifically, we have assumed that the differences in rating migration probabilities are identical for all ratings at the beginning of the period.

Other ways to control for the rating at the beginning of a period exist. For example, a credit risk modeler may include interactions between the rating at the beginning of an observation period and other control variables, or, alternatively, estimate separate equations for every rating class at the beginning of a period. In this extension, all parameters are conditional on $k^* = 1, 2, \ldots K-1$. For example, the rating migration model can be extended by including other control variables such as macroeconomic effects. Both the resulting rating migration probabilities and rating migration matrix then become time-varying. The following model conditions on GDP growth:

```
PROC LOGISTIC DATA = mortgage;
CLASS lagrating(REF='3');
MODEL rating = lagrating gdp_time/LINK=PROBIT;
RUN;
```

The parameter estimates include the threshold estimates, the parameter estimate for the first and second classes, and GDP growth (see Exhibit 6.18).

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	−2.8034	0.0152	34027.7024	<.0001
Intercept	2	1	1.5275	0.0112	18574.9967	<.0001
Intercept	3	1	4.8284	0.0118	167505.368	<.0001
Lagrating	1	1	3.5811	0.0216	27435.3449	<.0001
Lagrating	2	1	−0.2448	0.0109	501.3024	<.0001
gdp_time		1	0.0430	0.00103	1728.8677	<.0001

Exhibit 6.18 Cumulative Probit Model for Rating Migration Probabilities with Time-Varying Covariates

Another interesting extension could be the inclusion of higher-order lags into the analysis.

WHICH MODEL SHOULD I CHOOSE?

Choice of Link Functions, Distributions, and Optimization Algorithms

Multiple distributional choices are often available within a model class. As a general rule, the empirical distribution should match the assumed distribution. However, it can be hard to determine which model is best due to data limitations. An important driver in credit risk analytics is the macroeconomy. Time series data are generally limited due to the evolution of data storage facilities (often from 2000 onward) and typically also include structural breaks such as changes in economies, financial markets, institutions, instruments, and borrowers. This implies that distributions and parameters may change over time and that the out-of-time validation of choices is limited.

Furthermore, different models lead to different estimates. However, the final estimation outputs are always calibrated on data histories, and the economic content is often comparable across model types (see Hamerle et al. 2006, for one example). In other words, these choices typically lead to very similar conclusions, as the aim to bring default probabilities close to default indicators with a scientific method is common to all approaches.

Variable Selection

The development of economically founded models that follow the thinking of borrowers, lenders, and the general economy provides strong improvements for model quality. The choice of variables is crucial, portfolio dependent, and often subject to a first-stage analysis that is executed prior to the actual model building. In granular retail portfolios (e.g., mortgage portfolios that consist of a large number of exposures with small to moderate exposure amounts), macro-economic variables such as house prices or bank lending standards may dominate idiosyncratic (i.e., borrower-specific) variables. In smaller portfolios such as corporate lending, borrower-specific variables may dominate macroeconomic variables.

Adding explanatory variables to increase the model fit (e.g., via a forward selection technique) is potentially dangerous, as one may risk overfitting by providing excellent fits that cannot be replicated in an out-of-sample context. Hence, statistical validation techniques should be applied in an out-of-sample context (see also thevalidation chapter).

An important consideration is the selection of good predictive variables as inputs to the model. It also means that only variables that can sufficiently discriminate between default and nondefault should be considered.

There are various statistical techniques to test discriminatory power of individual independent variables such as means comparison in different groups, information value, and accuracy ratio.

SAS provides a procedure to statistically test the difference between two means of independent groups (default and nondefault). For example:

```
PROC TTEST DATA=data.mortgage;
CLASS default_time;
VAR FICO_orig_time;
RUN;
```

The output is based on two methods: the Satterthwaite approximation, which does not assume equal variance between the default and nondefault samples, and the pooled method, which assumes equal variance in both samples (see Exhibit 6.19).

The TTEST Procedure

Variable: FICO_orig_time

Method	Variances	DF	*t* Value	Pr > |*t*|
Pooled	Equal	622487	42.87	<.0001
Satterthwaite	Unequal	16027	45.52	<.0001

Exhibit 6.19 T-test for FICO_orig_time by default_time

It can be seen that both tests indicate a difference in means for defaults and nondefaults. Furthermore, PROC HPBIN in SAS is useful for binning and computing weight-of-evidence values and the corresponding information value (see the credit scoring chapter). An example of this is:

```
PROC HPBIN DATA=data.mortgage NUMBIN=5;
INPUT FICO_orig_time;
```

```
ODS OUTPUT MAPPING=Mapping;
RUN;
PROC HPBIN DATA=data.mortgage WOE BINS_META=Mapping;
TARGET default_time/LEVEL=BINARY;
RUN;
```

Remember, as discussed in the credit scoring chapter, the weight of evidence (WOE) measures the relative risk of each variable bin (or category). The information value is the weighted sum of the weight of evidence measures across the bins and indicates the predictive strength of the variable. The weights of evidence and information value that are obtained are shown in Exhibit 6.20.

The HPBIN Procedure

Weight of Evidence								
Variable	Binned Variable	Range	Nonevent Count	Nonevent Rate	Event Count	Event Rate	Weight of Evidence	Information Value
FICO_orig_time	[FICO...]		0	0	0	0	0	0
		FICO_orig_time<488	1533	0.95217391	77	0.04782609	−0.6993690	0.00178735
		488<=FICO_orig_time<576	59952	0.96482024	2186	0.03517976	−0.3790746	0.01724805
		576<=FICO_orig_time<664	196955	0.96836128	6435	0.03163872	−0.2693221	0.02699478
		664<=FICO_orig_time<752	251238	0.97860025	5494	0.02139975	0.13219861	0.00677210
		752<=FICO_orig_time	97653	0.99020473	966	0.00979527	0.92546632	0.08982732

Variable Information Value	
Variable	Information Value
FICO_orig_time	0.14262960

Exhibit 6.20 Weights-of-Evidence and Information Value for FICO_orig_time with Regard to default_time

As discussed in the credit scoring chapter, SAS is able to help selecting variables with the SELECTION=BACKWARD, SELECTION=FORWARD, and SELECTION=STEPWISE statements. These techniques vary the number of variables included by adding variables that are significant and dropping variables that are insignificant. See the credit scoring chapter for more details. While these techniques provide a first way to identify reasonable variable sets, variable selection in credit risk models has to take into account particular economic aspects, as purely machine-based selections are often not robust in out-of-sample forecasting exercises.

Other Helpful SAS Features

Other useful features in SAS that you may be interested in exploring include:

- The ROC statement within PROC LOGISTIC allows receiver operating characteristic evaluation (see the validation chapter for more details).
- PROC SURVEYLOGISTIC provides for clustered standard errors. The parameter estimates are comparable with PROC LOGISTIC, but standard errors are now computed by assuming dependence between observations of the same cluster unit. Clustering variables can be defined by the CLUSTER command.

- PROC DISCRIM estimates a discriminant criterion to classify each observation into default or nondefault. Early scoring functions (e.g., the Altman z-score) were developed using discriminant analysis.

FITTING AND FORECASTING

Sampling strategy: training and validation sample

For large data sets, it is common to create cross-sectional random samples to reduce the data volume. Within a random sample, it is common to stratify the sample further into a training sample with regard to the cross-section and time series dimension and borrower, and a validation sample (for an example, see Lee, Rösch, and Scheule 2016).

With regard to the cross section, one may differentiate between in-sample validation if the borrowers of the training sample are analyzed, and out-of-sample validation if different borrowers than those in the training sample are analyzed.

With regard to the time series dimension, one may differentiate between in-time (if the same time periods of the training sample are analyzed), and out-of-time (if different time periods than those in the training sample are analyzed).

Exhibit 6.21 shows the various combinations of these classifications and resulting data samples.

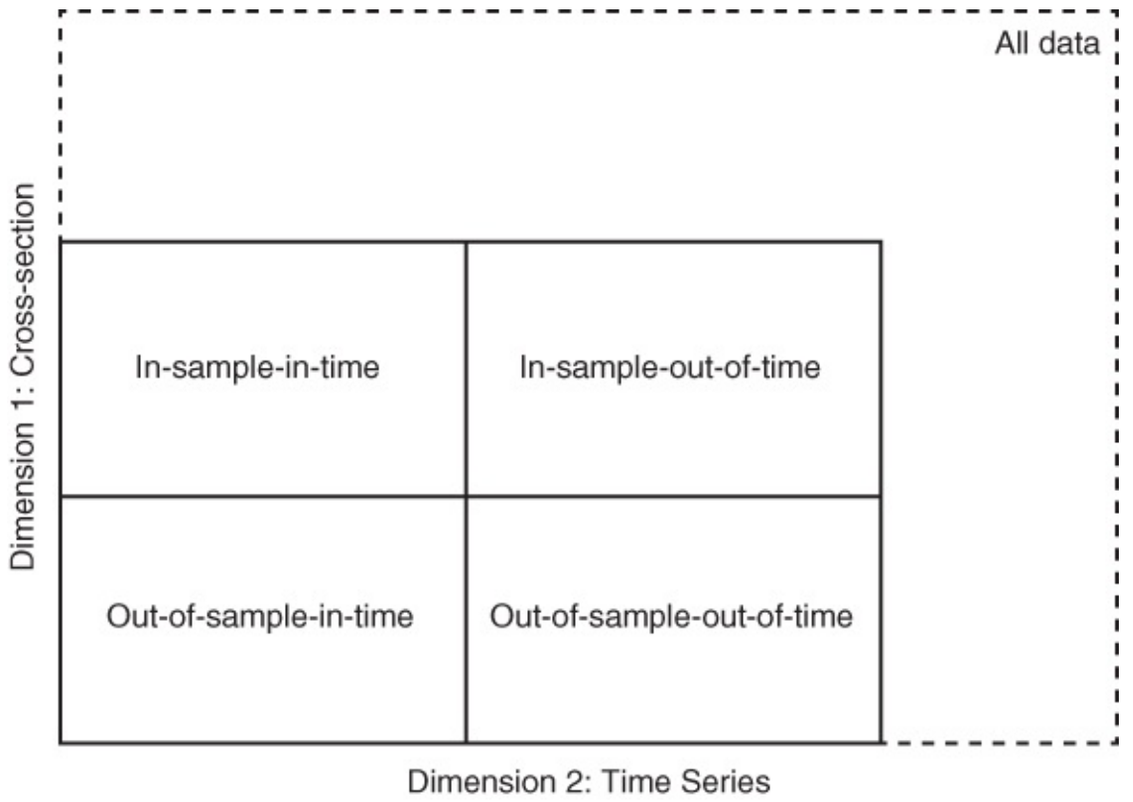

Exhibit 6.21 Data Sampling Strategies

In a rigorous model validation, validation samples are generally out-of-sample and out-of-time

with regard to the training sample. The data sampling strategy may be simpler if data is limited.

Please note that it is possible that the training sample is not representative with regard to the population. Bayesian models are able to include prior assumptions with regard to the population and hence correct for this bias. These methods are discussed in the chapter on Bayesian methods.

Generation of Samples

We now illustrate the generation of a training sample. Here, we do not distinguish between in-time and out-of-time and use all available data (i.e., the four tiles in the above chart are equal to all data). PROC SURVEYSELECT generates random samples. By default, the output data set includes only those observations selected for the sample. If OUTALL is specified, then a selection indicator named Selected is added to the input data set. Furthermore, it is common in panel data sets to select by subjects (here borrowers/loans), which we achieve with the SAMPLINGUNIT command:

```
PROC SURVEYSELECT DATA=data.mortgage SAMPRATE=0.8 OUTALL SEED=12345
OUT=mortgage;
SAMPLINGUNIT id;
RUN;
```

The command SEED=12345 fixes the random experiment so that the random draw results in the same outcome if executed another time. We note that the observed mortgage data experiences an increase in default risk at approximately time period 25. Hence, one may build a model that is estimated precrisis and apply the estimated parameters (out-of-sample) to periods during the crisis to assess whether the model is capable of predicting credit risk outcomes in severe economic downturns. Lee et al. (2016) perform such an analysis based on a larger number of risk factors.

In-Sample and Out-of-Sample Validation

We can now estimate a logit model using PROC LOGISTIC for the training sample:

```
PROC LOGISTIC DATA=mortgage(WHERE=(selected=1)) DESCENDING;
MODEL default_time = FICO_orig_time
LTV_time gdp_time;
STORE OUT=model_logistic;
RUN;
```

Next, we calculate the probabilities for the training and validation sample using PROC PLM. The procedure allows the post-modeling estimation of default probabilities. In case of out-of-time subsets, we actually forecast the default probabilities, as all covariates are time lagged with regard to the default indicator.

```
PROC PLM SOURCE=model_logistic;
SCORE DATA=mortgage OUT=mortgage2;
RUN;
```

As PROC PLM calculates the linear predictor, we need to convert the calculated linear predictor to a default probability using the logit link function; note that we would use the cumulative density function of the standard normal distribution for a probit model by replacing the line PD_time=... by PD_time= PROBNORM(predicted):

```
DATA mortgage2;
SET mortgage2;
PD_time=EXP(predicted)/(1+EXP(predicted));
RUN;
```

We then analyze the default rates and mean default probabilities in-sample and out-of-sample:

```
PROC SORT DATA=mortgage2;
BY selected time;
RUN;
PROC MEANS DATA=mortgage2;
BY selected time;
OUTPUT OUT=means MEAN(default_time PD_time)=default_time PD_time;
RUN;
```

Exhibits 6.22 and 6.23 show the real-fit diagrams (produced by PROC GPLOT), which include the default rate (DR) and the mean estimated default probabilities (PD) for the in-sample and out-of-sample.

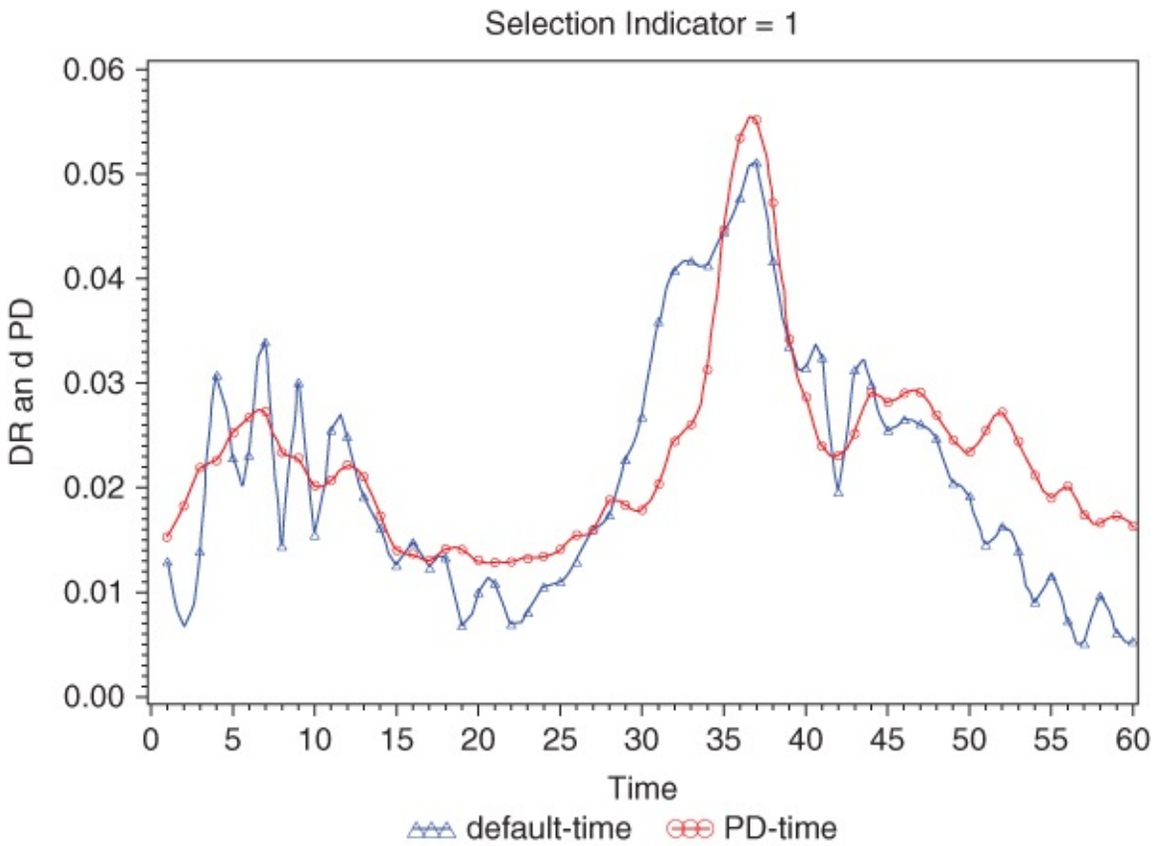

Exhibit 6.22 Real-Fit Diagram for In-Sample

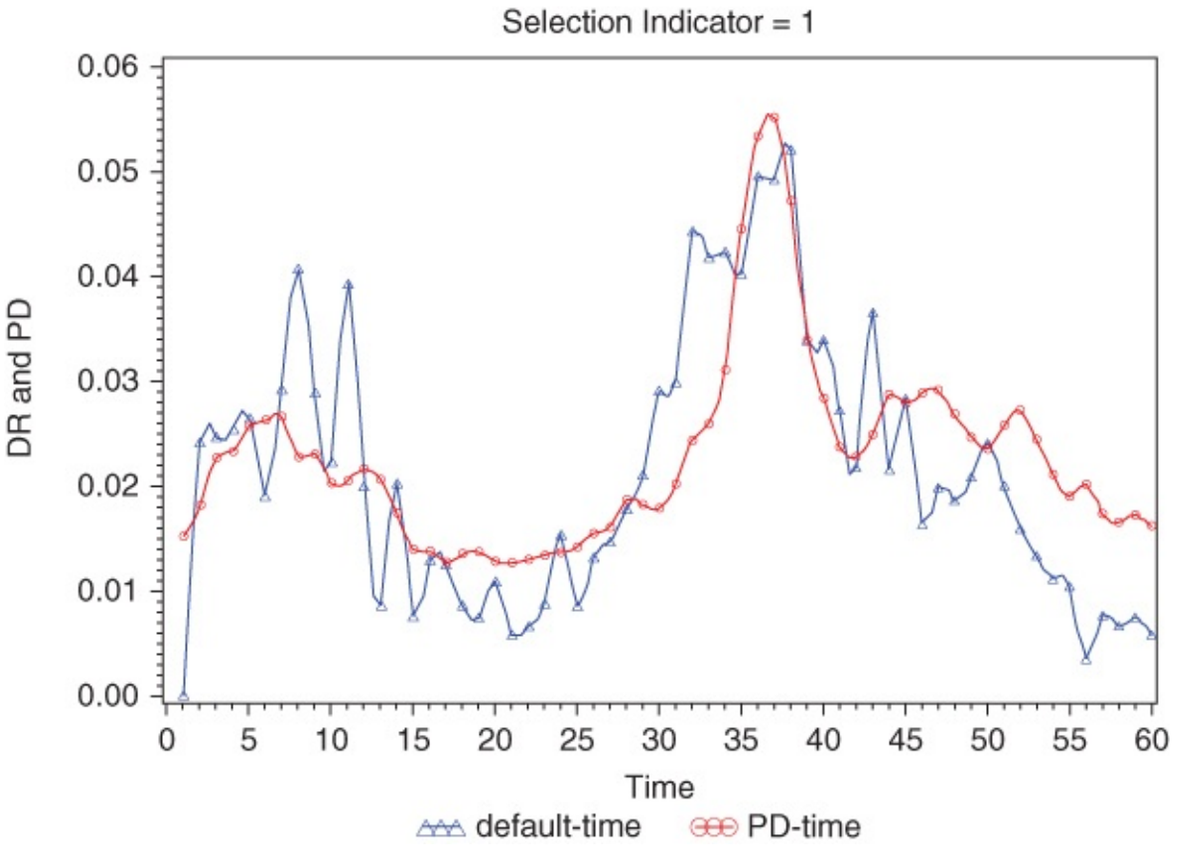

Exhibit 6.23 Real-Fit Diagram for Out-of-Sample

```
ODS GRAPHICS ON;
AXIS1 ORDER=(0 TO 60 BY 5) LABEL=('Time');
AXIS2 ORDER=(0 TO 0.06 BY 0.01) LABEL=('DR and PD');
SYMBOL1 INTERPOL=SPLINE WIDTH=2 VALUE=TRIANGLE C=BLUE;
SYMBOL2 INTERPOL=SPLINE WIDTH=2 VALUE=CIRCLE C=RED;
LEGEND1 LABEL=NONE SHAPE=SYMBOL(4,2) POSITION=(BOTTOM OUTSIDE);
PROC GPLOT DATA=means;
  PLOT (default_time PD_time)*time / OVERLAY HAXIS=AXIS1 VAXIS=AXIS2
            LEGEND=LEGEND1;
BY selected;
RUN;
ODS GRAPHICS OFF;
```

Generally speaking, in-sample estimated default probabilities dominate out-of-sample estimates and in-time default probabilities dominate out-of-time estimates.

FORMATION OF RATING CLASSES

Model performance measurement often requires a large number of observations, and banks form rating classes, allocate observations to rating classes, and compute risk and risk model validation metrics for these rating classes. Observations (i.e., borrowers at a given time period) are assumed to be homogeneous for a given rating class. Rating classes not only are formed for model validation but may also match expectations of regulators (i.e., part of prudential regulation), or serve various risk applications of a bank, as ratings often simplify

complicated measurement challenges.

Rating classes are generally created by (1) observing classes or (2) categorizing meaningful metric variables or metric credit scores. Classes may be observable if borrowers are rated by external credit rating agencies or similar service providers. External ratings are often available for only the largest corporates within an economy. Various options are available to categorize scores into rating classes, and three popular examples are:

1. Set class intervals to have equal sizes.
2. Set class intervals so that classes are expected to have equal numbers of observations.
3. Set class intervals so that classes are expected to have equal numbers of default observations.

Other categorization techniques are also possible. For example, a mapping to an external rating agency scale can be adopted, or an exponential heuristic method, which says, for example, that the default rate should double from one rating to the next. In practice (based on our experience), ratings can be normally classified from the PD. There are a few points to consider when developing a rating system:

- Distribution of ratings: The distribution of ratings is often expected to follow a skewed bell curve. This implies that a large proportion of borrowers/loans concentrates in the medium-risk ratings, low-risk ratings are at one tail, and high-risk ratings at the other tail.
- Monotonicity of default rate curve: The default rate for the ratings should be monotonically increasing or decreasing over the ratings.
- Dispersion: The default rate difference between the lowest and the highest rating should be large (e.g., 20 to 30 percent).
- Master scale for ratings: Banks often develop multiple PD models for different products or segments within the whole credit portfolio. It would be handy if the ratings were aligned in terms of risk levels across risk segments. This can be accomplished by the definition of a master scale as a corporate-wide standard.
- Aligning with external rating agencies: Banks often assign internal credit ratings in line with the default rate of external rating agencies.

We now implement these three options in SAS to demonstrate the formation of rating classes from a credit score—here the FICO score.

Rating approach 1: Set class intervals to have equal sizes

In this application, we assume 10 rating classes (classes zero to nine), define lower and upper bounds of the rating classes with regard to the FICO score, and assign the observations/borrowers to the rating classes based on the observed FICO scores. The new variable FICO_orig_time_rank1 collects the rating classes for this approach.

```
DATA rank1;
SET data.mortgage;
```

```
IF FICO_orig_time>=300 AND FICO_orig_time<=400 THEN FICO_orig_time_rank1=0;
IF FICO_orig_time>400 AND FICO_orig_time<=450 THEN FICO_orig_time_rank1=1;
IF FICO_orig_time>450 AND FICO_orig_time<=500 THEN FICO_orig_time_rank1=2;
IF FICO_orig_time>500 AND FICO_orig_time<=550 THEN FICO_orig_time_rank1=3;
IF FICO_orig_time>550 AND FICO_orig_time<=600 THEN FICO_orig_time_rank1=4;
IF FICO_orig_time>600 AND FICO_orig_time<=650 THEN FICO_orig_time_rank1=5;
IF FICO_orig_time>650 AND FICO_orig_time<=700 THEN FICO_orig_time_rank1=6;
IF FICO_orig_time>700 AND FICO_orig_time<=750 THEN FICO_orig_time_rank1=7;
IF FICO_orig_time>750 AND FICO_orig_time<=800 THEN FICO_orig_time_rank1=8;
IF FICO_orig_time>800 AND FICO_orig_time<=850 THEN FICO_orig_time_rank1=9;
RUN;
```

Note that the FICO score is defined between 300 and 850 in this data set.

Rating approach 2: Set class intervals so that classes are expected to have equal numbers of observations

In this application, we use PROC RANK to assign the observations into 10 rating classes with an equal number of observations. The GROUPS command defines the number K of rating classes and assigns the $1/K$ observations with the lowest FICO score to class 0, $1/K$ observations with the second lowest FICO score to class 1, and so on. The new variable FICO_orig_time_rank2 collects the rating classes for this approach.

```
PROC RANK DATA=rank1 OUT=rank2 GROUPS=10;
VAR FICO_orig_time;
RANKS FICO_orig_time_rank2;
RUN;
```

Rating approach 3: Set class intervals so that classes are expected to have equal numbers of default observations

In this application, we use PROC RANK to assign the observations into 10 rating classes with an equal number of observations but stratify the ranking by the default indicator default_time. The new variable FICO_orig_time_rank3 collects the rating classes for this approach.

```
PROC SORT DATA=rank2;
BY default_time;
RUN;
PROC RANK DATA=rank2 OUT=rank3 GROUPS=10;
VAR FICO_orig_time;
RANKS FICO_orig_time_rank3;
BY default_time;
RUN;
```

We then sort according to the FICO scores and overwrite the ranking for stratum for nondefaults by the ranking of the previous default event.

```
PROC SORT DATA=rank3;
BY FICO_orig_time descending default_time;
RUN;
DATA rank3(DROP=temp);
SET rank3;
```

```
BY FICO_orig_time descending default_time;
IF first.default_time AND default_time = 1 THEN temp =
FICO_orig_time_rank3;
ELSE IF default_time = 0 THEN DO;
FICO_orig_time_rank3 = temp;
END;
RETAIN temp;
RUN;
DATA rank3;
SET rank3;
IF FICO_orig_time_rank3 = . THEN FICO_orig_time_rank3 = 0;
RUN;
```

In this example, we created 10 rating classes. Various options within PROC RANK (option ties) are available to specify how tied values should be ranked and include the lowest, mean, or highest rank value of tied observations. Tied observations are of a lower importance if the number of ties is small relative to the total number of observations. An indication for this might be a mortgage loan book with a continuous score.

Exhibit 6.24 shows the relative frequencies of observations per rating class. We have generated the data points by computing the number of observations, the default rate, and the mean FICO score by rating class for the three approaches. We obtain three data sets and assign different names to them (_1, _2, and _3).

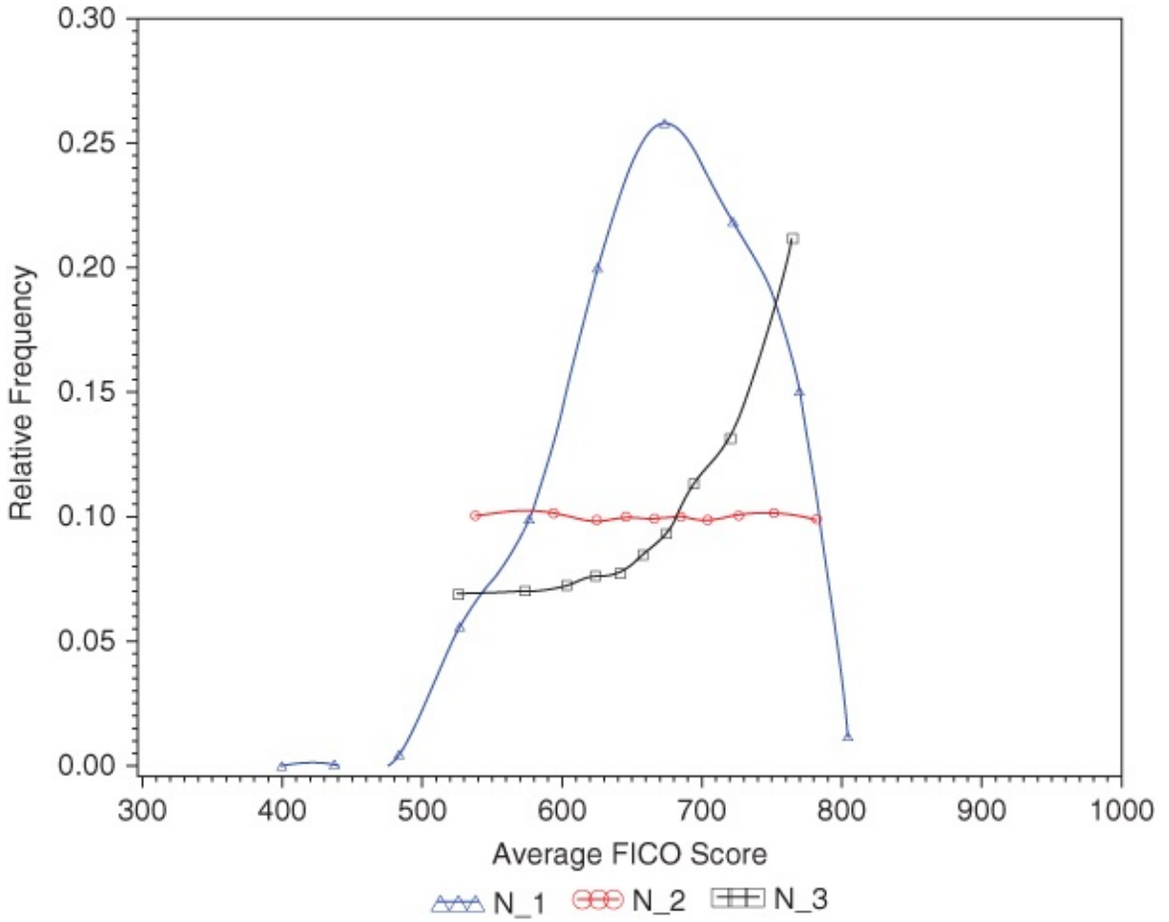

Exhibit 6.24 Relative Frequencies of Observations per Rating Class

```
PROC SORT DATA=rank3;
BY FICO_orig_time_rank1;
RUN;
PROC MEANS DATA=rank3;
BY FICO_orig_time_rank1;
OUTPUT OUT=orank1 N(default_time)=N_1 MEAN(default_time)=DR_1
MEAN(FICO_orig_time)=FICO_1;
RUN;
PROC SORT DATA=rank3;
BY FICO_orig_time_rank2;
RUN;
PROC MEANS DATA=rank3;
BY FICO_orig_time_rank2;
OUTPUT out=orank2 N(default_time)=N_2 MEAN(default_time)=DR_2
MEAN(FICO_orig_time)=FICO_2;
RUN;
PROC SORT DATA=rank3;
BY FICO_orig_time_rank3;
RUN;
PROC MEANS DATA=rank3;
BY FICO_orig_time_rank3;
OUTPUT OUT=orank3 N(default_time)=N_3 MEAN(default_time)=DR_3
MEAN(FICO_orig_time)=FICO_3;
run;
```

We then rename the rating class variable to a joint name and merge the three data sets:

```
DATA orank1(KEEP= FICO_orig_time_rank FICO_1 N_1 DR_1);
SET orank1;
RENAME FICO_orig_time_rank1=FICO_orig_time_rank;
N_1=N_1/622489;
RUN;
DATA orank2(KEEP= FICO_orig_time_rank FICO_2 N_2 DR_2);
SET orank2;
RENAME FICO_orig_time_rank2=FICO_orig_time_rank;
N_2=N_2/622489;
RUN;
DATA orank3(KEEP= FICO_orig_time_rank FICO_3 N_3 DR_3);
SET orank3;
RENAME FICO_orig_time_rank3=FICO_orig_time_rank;
N_3=N_3/622489;
RUN;
DATA orank;
MERGE orank1 orank2 orank3;
BY FICO_orig_time_rank;
RUN;
```

We use PROC GPLOT to plot the relative frequencies of observations per rating class:

```
ODS GRAPHICS ON;
AXIS1 ORDER=(300 TO 1000 BY 100) LABEL=('Average FICO score');
AXIS2 ORDER=(0 TO 0.3 BY 0.05) LABEL=('Relative frequency');
SYMBOL1 INTERPOL=SPLINE WIDTH=2 VALUE=TRIANGLE C=BLUE;
SYMBOL2 INTERPOL=SPLINE WIDTH=2 VALUE=CIRCLE C=RED;
```

```
SYMBOL3 INTERPOL=SPLINE WIDTH=2 VALUE=SQUARE C=BLACK;
LEGEND1 LABEL=NONE SHAPE=SYMBOL(4,2) POSITION=(BOTTOM OUTSIDE);
PROC GPLOT DATA=orank; PLOT N_1*FICO_1 N_2*FICO_2 N_3*FICO_3 /
OVERLAY LEGEND=LEGEND1 HAXIS=AXIS1 VAXIS=AXIS2;
RUN;
ODS GRAPHICS OFF;
```

The number of observations follows a hump-shape distribution if rating class intervals are set by fixed intervals of equal length. The number of observations is per definition equally distributed for the second approach, and increases for equal numbers of default events per rating classes, as the credit risk of borrowers and hence frequency of default events is lower.

We use PROC GPLOT to plot the default rate per rating class (see Exhibit 6.25):

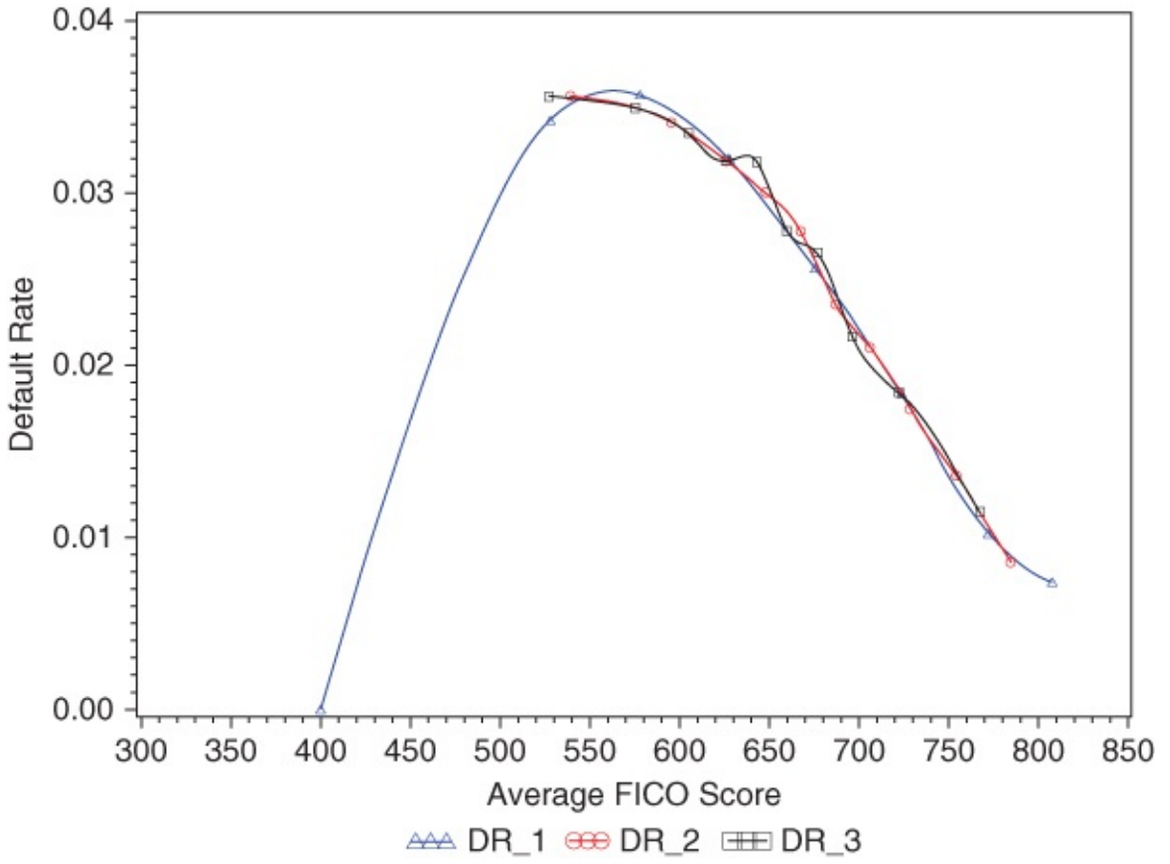

Exhibit 6.25 Default Rate per Rating Class

```
ods graphics on;
AXIS1 ORDER=(300 TO 850 BY 50) LABEL=('Average FICO score');
AXIS2 ORDER=(0 TO 0.04 BY 0.01) LABEL=('Default rate');
SYMBOL1 INTERPOL=SPLINE WIDTH=2 VALUE=TRIANGLE C=BLUE;
SYMBOL2 INTERPOL=SPLINE WIDTH=2 VALUE=CIRCLE C=RED;
SYMBOL3 INTERPOL=SPLINE WIDTH=2 VALUE=SQUARE C=BLACK;
LEGEND1 LABEL=NONE SHAPE=SYMBOL(4,2) POSITION=(BOTTOM OUTSIDE);
PROC GPLOT DATA=orank; PLOT DR_1*FICO_1 DR_2*FICO_2 DR_3*FICO_3 /
OVERLAY LEGEND=LEGEND1 HAXIS=AXIS1 VAXIS=AXIS2;
RUN;
ODS GRAPHICS OFF;
```

Generally speaking, the default rate decreases from low to high rating class. In the example,

this is only the case for Options 2 and 3, but not for Option 1. Note that this property is based on our definition of risk ordering, which is consistent with the FICO score (low value: high risk to high value: low risk). External rating agencies like Fitch and Moody's have chosen letters to assign to rating classes (AAA/Aaa: low risk to C: high risk), +/−, (+: lower risk to −: higher risk), or numbers 1/2/3 (1: lower risk to 3: higher risk). Others may use traffic light approaches (red: high risk, yellow: medium risk, and green: low risk).

Trade-Off Effect

The first approach is more likely to result in a nonmonotone decline of the default rate by rating class. This is due to the random nature of default events. Lower rating classes have fewer observations and the default rate is volatile. Hence, there are trade-offs between these options that relate to the nature of the data set. Generally speaking, the number of default observations is small relative to the total sample size.

Option 1 may be preferable as it provides for an equally spaced coverage over the score range if the number of default events is sufficiently large. However, Option 2 or 3 may become suitable if default events are limited, as both approaches (Option 3 to a greater extent than Option 2) focus on ranges where reasonable numbers of default events are observed. Note that the first and last rating classes may encompass a large range of score values and are limited in terms of their informational value and interpretation. Therefore, we prefer in most of our empirical applications to use the third option.

PRACTICE QUESTIONS

1. Estimate a probit model for the default probability, and interpret the parameters with regard to the PD. Include the following risk factors: FICO_orig_time, LTV_time, REtype_CO_orig_time, REtype_PU_orig_time, and REtype_SF_ orig_time. Use data set mortgage.
2. Categorize the current LTV ratio into 10 rating classes, and estimate the rating migration probabilities. Use data set mortgage.
3. Categorize the current LTV ratio and include the effect-coded LTV ratio next to the FICO score into a logit model for the PD. How do you interpret the parameter estimates for the LTV ratio, and what may be the advantages and disadvantages of including a metric variable in categories(relative to a stand-alone inclusion)? Use data set mortgage.
4. Estimate a probit model and a logit model based on the interest rate and FICO score at origination. Compare the parameter estimates. Estimate the minimum, maximum, mean, and median for the resulting PDs, and compare the moments. Which model dominates in your opinion? Use data set mortgage.
5. Compute a proxy for house price appreciation and include it next to the current LTV ratio in a probit model for PDs. Estimate the model and interpret the parameter estimates. Why would the house price appreciation be relevant next to the current LTV ratio that includes

the price of the house collateral?

6. Show an example of a TTC model and a PIT model for PDs. What are the implications for both from a regulatory perspective?

REFERENCES

Crook, J., and T. Bellotti. 2010. "Time Varying and Dynamic Models for Default Risk in Consumer Loans," *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173 (2): 283–305.

Crook, J. N., D. B. Edelman, and L. C., Thomas. 2007. "Recent Developments in Consumer Credit Risk Assessment," *European Journal of Operational Research* 183 (3): 1447–1465.

Hamerle, A., T. Liebig, and H. Scheule. 2006. "Forecasting Credit Event Frequency—Empirical Evidence for West German Firms." *Journal of Risk* 9: 75–98.

Hamerle, A., R. Rauhmeier, and D. Rösch. 2003. "Uses and Misuses of Measures for Credit Rating Accuracy." Available at SSRN 2354877

Lee, Y., D. Rösch, and H. Scheule. 2016. "Accuracy of Mortgage Portfolio Risk Forecasts during Financial Crises." *European Journal of Operational Research* 249 (2): 440–456.

Leow, M., and C. Mues. 2012. "Predicting Loss Given Default (LGD) for Residential Mortgage Loans: A Two-Stage Model and Empirical Evidence for UK Bank Data." *International Journal of Forecasting* 28 (1): 183–195.

Merton, R. C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* 29: 449–470.

Rösch, D., and H. Scheule. 2004. "Forecasting Retail Portfolio Credit Risk." *Journal of Risk Finance* 5 (2): 16–32.

Rösch, D., and H. Scheule. 2010. "Downturn Credit Portfolio Risk, Regulatory Capital and Prudential Incentives." *International Review of Finance* 10 (2), 185–207.