Segmentation, Probability of Default and Basel II Capital Measures for Credit Card Portfolios

Draft: Aug 13, 2007

Dennis Ash
Federal Reserve Bank of Philadelphia

Shannon M. Kelly
HSBC*

William W. Lang
Federal Reserve Bank of Philadelphia

William Nayda
Capital One Corporation

Haining Yin
Capital One Corporation

*Work completed while at Federal Reserve Bank of Philadelphia

Introduction

This paper examines alternative methods for differentiating the likelihood of default among credit card borrowers and how these alternative methods affect capital requirements for the portfolio when using a variant of the Vasicek (2002) asymptotic single risk factor (ASRF) model to construct a value-at-risk (VAR) capital measure.

The results in our paper have important implications for banking practice. First, the general modeling approach discussed in the paper is an approach commonly used in the banking industry (see Risk Management Association, 2003). Moreover, this modeling framework is embodied in the proposed Basel II regulatory framework for bank capital.[1]

In the ASRF model, accurate estimates of the tail of the portfolio distribution require differentiating the portfolio into separate "homogeneous" risk buckets. Credit card risk managers often refer to these differentiated risk buckets as "segments." While there are some important differences between credit card and commercial loan portfolios, the segmentation process for the consumer portfolio is in many ways analogous to the loan rating process commonly used for estimating loss distributions of a commercial lending portfolio (see Treacy and Carey, 1998).

We show that the estimates of VARs obtained from the ASRF model are inversely related to the degree of homogeneity of the portfolio segments. Other things equal, a finer differentiation of default risk among borrowers produces a more accurate and lower estimated VAR.

---

[1] See Basel Committee on Banking Supervision, 2004.

Using data from a sample of credit card loans over the period 1999-2004, we examine the importance of various customer attributes in differentiating the likelihood of borrower default in credit card portfolios. The attributes we examine are customer credit score at the time the loan was originated, updated or refreshed credit scores, and delinquency status. We estimate the one-year horizon probability of default (PD) for various alternative segmentation approaches based on these attributes. We also examine the role of loan age or "seasoning" in predicting credit card defaults. Our PD estimates are then entered into the portfolio loss model to estimate economic capital measures using different segmentation schemes. This economic capital measure can be thought of as a summary indicator of the importance of improving segmentation for measuring economic capital.

The next section of the paper outlines a simple theoretical model of consumer default and the corresponding distribution of portfolio losses. The following section then discusses segmentation criteria for credit card portfolios. Section III of the paper describes the data and analysis design. Section IV discusses PD results using our alternative segmentation criteria. Section V discusses the affects of loan age on PD. Section VI analyzes the capital measures resulting from alternative approaches to segmentation.

I. A Model of Consumer Defaults

In this section, we detail a consumer credit version of the ASRF model. The model is a special case of a Merton (1974) options-based structural model of default.

Consider a consumer borrower j with net worth (measured in natural logs) at time t of $w_{j,t}$ and that net worth follows a standard geometric random walk model with drift:

$$w_{j,t+1} = w_{j,t} + \mu_t + v_{j,t} \; ; with \;\; v_{j,t} \sim N(0, \sigma^2_{vj}) \tag{1}$$

In the model, a period is equivalent to the forecast time horizon. Most portfolio risk models use a one-year time horizon and our empirical estimates will be based on a one-year horizon. Default is assumed to occur at time t+1 when the borrower's net worth falls below some threshold value:

$$w_{j,\,t+1} < \theta_{j,\,t+1} \tag{2}$$

The threshold value might differ across individuals depending on a number of observable factors such as marital status, debt burden, state laws, and employment status. Individuals might also differ by nonobservable traits such as their attitude toward default or reputation costs associated with default.

Subtracting $w_{j,\,t}$ from both sides of equation (2):

$$g_{j,\,t+1} = \mu_t + \nu_{j,\,t} < m_{j,\,t+1} \tag{3}$$

where $g_{j,\,t+1} = w_{j,\,t+1} - w_{j,\,t}$ and $m_{j,t+1} = \theta_{j,\,t+1} - w_{j,t}$

The shock to borrower j's wealth is assumed to be driven by a stochastic systematic factor and an idiosyncratic factor. Both the systematic and idiosyncratic random variable are assumed to be independent and identically distributed. Without loss of generality, we assume that these shocks are standard normal random variables.

$$v_{j,t} = \sqrt{\rho}Y_t + \sqrt{1-\rho}\varepsilon_{jt} \text{ and } Y_t \sim \mathrm{N}(0,1) \text{ and } \varepsilon_{jt} \sim N(0,1) \tag{4}$$

where Y_t is the systematic factor and ε_{jt} is the borrower-specific idiosyncratic shock and ρ is the correlation of borrower wealth with the systematic risk factor.

Equations (3) and (4) imply:

$$\pi_{j,\,t+1} = \Phi(m_{j,\,t+1} - \mu_j) \tag{5}$$

where $\pi_{j,t+1}$ is the unconditional probability of default of borrower j at time t+1, and Φ is the standard normal cumulative density function. Alternatively:

$$\Phi^{-1}(\pi_{j,\,t+1}) = m_{j,\,t+1} - \mu_j \tag{6}$$

The state of default conditional on Y_t can now be written as:

$$\varepsilon_{j,\,t} < [\Phi^{-1}(\pi_{j,\,t+1}) - \sqrt{\rho} Y_t] / \sqrt{1-\rho} \tag{7}$$

This implies a probability of default for borrower j conditional on the realization of Y as:

$$\pi_{j,\,t+1}(Y_t) = \Phi([\Phi^{-1}(\pi_{j,\,t+1}) - \sqrt{\rho} Y_t] / \sqrt{1-\rho}) \tag{8}$$

Now consider a portfolio of loans to *n* borrowers with homogeneous (type *j*) risk characteristics. For simplicity, we assume that all loans are of equal size.[2] Let $D^p_{j,\,t+1}(Y_t)$ be the default frequency of the portfolio. Then as $n \to \infty$:

[2] Allowing for different size exposures is straightforward as long as we assume that there are a large number of borrowers and individual loans represent a very small share of the total loan portfolio.

$$D^{p}_{j,\, t+1}(Y_t) \cong \pi_{j,\, t+1}(Y_t) = \Phi\left[\frac{\Phi^{-1}(\pi_{j,\, t+1}) - \sqrt{\rho}\, Y_t}{\sqrt{1-\rho}}\right] \tag{9}$$

Equation (9) can be used for calculating the VAR for any chosen threshold. The threshold chosen corresponds to some target level of solvency. The Basel II proposal sets this threshold at the .001 probability of failure. Let ν be the chosen threshold probability of failure. Since by assumption idiosyncratic risk is diversified away, the ν upper tail of the distribution of $D^{p}_{j,\, t+1}(Y_t)$ is equivalent to the lower ν tail of Y_t. Let $\underline{Y_t}$ be the ν lower tail of the distribution of Y_t. Then,

$$D^{p}_{j,\, t+1}(\underline{Y_t}) \cong \Phi\left[\frac{\Phi^{-1}(\pi_{j,\, t+1}) - \sqrt{\rho}\, \underline{Y_t}}{\sqrt{1-\rho}}\right] = \Phi\left[\frac{\Phi^{-1}(\pi_{j,\, t+1}) - \sqrt{\rho}\, \Phi^{-1}(\nu)}{\sqrt{1-\rho}}\right] \tag{10}$$

Subtracting expected losses from equation (10) provides an estimate for capital required for a homogeneous portfolio of credit card loans. What if the loan portfolio consists of borrowers with different default probabilities? Suppose a lender is able to discriminate between *S* discrete types of borrowers with different default probabilities but a common correlation with the systematic risk factor. The default rates between borrower types are correlated to the systematic factor, but the idiosyncratic shocks for all borrowers are assumed to be uncorrelated. The lender is assumed to have a large number of borrowers of each type, with α_j representing the share of type j borrowers in the portfolio. The ν upper tail of the default frequency distribution for a portfolio with *S* heterogeneous borrower types is then:

$$\sum_{j=1}^{S} \alpha_j D^{p}_{j,\, t+1}(\underline{Y_t}) \cong \sum_{j=1}^{S} \alpha_j \Phi\left[\frac{\Phi^{-1}(\pi_{j,\, t+1}) - \sqrt{\rho}\, \Phi^{-1}(\nu)}{\sqrt{1-\rho}}\right] \; ; \; \sum_{j=1}^{S} \alpha_j = 1 \tag{11}$$

Suppose that lenders differ in their ability to distinguish borrower types, with some lenders being able to distinguish a smaller number of borrower types or some lenders making more errors in identifying the borrower's type. In either case, the effect will be to group borrowers with different default frequencies and estimate the model as if they are the same type. As discussed in Laurent (2004), the VAR measure in equation (11) is concave for most of the relevant range of PDs. From Jensen's inequality, it is easy to demonstrate that treating loans with different PDs as a single group results in overestimating the upper tail of the default frequency distribution. Since aggregating across different borrower types does not bias estimates of expected loss, more accurately distinguishing borrower types lowers the estimated capital for the portfolio. Thus, we expect, and our empirical analysis below confirms, that more accurate segmentation of a portfolio will produce lower estimates of economic capital using the ASRF model.

II. Segmentation Criteria for Credit Card Portfolios

The model discussed in the previous section requires that loans be grouped into homogeneous risk classes to accurately measure the tail of the loss distribution and the resulting capital requirement for the portfolio. For credit card portfolios, this grouping or segmentation process is typically done by assigning loans to risk groups where the groups are defined by proxy borrower risk characteristics reflecting borrowers' ability and intent to make payments. PD is then estimated using the realized default experience of individual segments. In this section, we discuss risk factors in credit card portfolios that are typically analyzed by risk managers of credit card lenders.

Credit Scores and Delinquency Status

Unlike large commercial lenders, credit card lenders typically devote limited resources to analyzing the idiosyncratic risk of an individual borrower or individual loan. Rather than relying on direct analysis and monitoring of the idiosyncratic characteristics of individual consumers, large credit card lenders rely heavily on statistical models of borrower performance based on standardized data for credit approval decisions, risk-based pricing, determining credit limits, and setting collection strategies.

The primary statistical tool for making credit card lending decisions is credit scoring. The existence of extensive credit bureau data in the U.S. allows lenders to use a wealth of readily available data on individuals to estimate "bureau" scoring models. In addition to using standardized bureau scoring models, banks often buy or develop customized scoring models tailored to a bank's own client population. Many companies with a sufficiently large consumer lending portfolio also employ "application" scoring models that allow an institution to incorporate additional information collected during the loan application process.

In addition to scoring customers at the time of application, scoring is used in a dynamic way for managing accounts and for performing internal bank analytics. Large credit card lenders typically obtain updated or refreshed bureau scores on a monthly basis. Some lenders also estimate "behavioral" credit scores that combine updated credit bureau information with information on the borrower's performance on accounts with the bank. These updated scores are then used for a wide variety of purposes, including credit line changes, repricing, and collection strategies.

One commonly used modeling technique for estimating credit scores is to build a standard logistic regression model using credit bureau data and possibly bank-specific data. The discrete outcome typically modeled is the probability of a borrower ever becoming seriously delinquent over some fixed time horizon. Seriously delinquent is typically defined as more than

60 days past due or more than 90 days past due. The time horizon window can vary from six months to two years.

Logistic and other scoring models can be used to generate an estimate of the probability of a “bad” outcome. However, while the output of a standard scoring model produces a “probability,” the probability estimate of the standard scoring model is generally not used as the PD estimate in internal economic capital models and would not produce the PD required by regulators under Basel II.

One reason is that the “bad” outcome used in the scoring model is usually some measure of serious delinquency (e.g., 90 days or more past due) rather than a probability of default or loss. At U.S. banks, credit card loans are typically charged off when they are 180 days or more past due or when the borrower declares bankruptcy. Moreover, the PD estimate entering the ASRF model is an unconditional PD representing an average default rate over a long time period for borrowers within a risk class. The probability of default generated by credit scoring is typically a short-run conditional estimate.

While most credit scoring models do not produce a PD that is appropriate as a direct input into the ASRF loss distribution model, the rank ordering properties of scoring models play a critical role in the segmentation process, since they are strong indicators of risk type. A common practice among banks for estimating PD is to estimate the relationship between *ex post* one-year default rates and credit scores, controlling for other risk factors. Our methodology for incorporating scores into PD estimation follows this approach. In addition, we examine the importance of updating credit scores in estimating PD and segmenting the portfolio.

Delinquency status is another critical risk factor for predicting loss for a credit card portfolio. When using a one-year horizon for estimating future losses, a large percentage of those losses will come from borrowers who are currently delinquent. In our estimates below, we examine the relationship between delinquency status and PD as well as the importance of incorporating credit scores into the segmentation process controlling for delinquency status.

Refreshed Credit Scores

Some banks with sophisticated risk management systems obtain refreshed credit scores as well as updated information on other risk factors. Clearly, updating this information provides a better measure of an individual borrower's risk type. Thus, segmentation of borrowers into homogeneous risk types will be more accurate if banks reallocate loans into segments based on this updated information.

Some credit card lenders use an alternative risk measurement approach that fixes loans into segments at the time of origination. Individual loans remain within a single segment for the entire life of the loan. Performance is then tracked for these "fixed" segments, and risk parameter estimates can be estimated based on the historical performance of segments with similar characteristics at origination. "Vintage analysis" is a variant of the fixed segment approach where all loans in the segment are originated during a common time period (e.g., all loans in the segment are originated in the same month). Vintage analysis is commonly used at banks for long-term projections of loss and profitability. In the analysis below we compare what effect using refreshed credit scores has on estimating PD when allocating loans to segments versus using a fixed segmentation approach based on scores at origination.

Loan Age

For many loan types with long effective maturities, loan age is a predictive factor for default. In credit card portfolios this effect of loan age on defaults is referred to as "seasoning," and plots of portfolio default rates against time on books are called seasoning curves. The typical seasoning curve has an upward slope that peaks typically between 18 months to three years and then either flattens out or declines. Vintage curves are a variant of seasoning curves where each seasoning curve represents portfolio performance of loans originated in the same time period. Credit card lenders compare the seasoning curves for different vintages to determine relative performance across portfolios originated at different time periods and use extrapolation methods for predicting future default rates.

Note that the upward slope typically observed in a standard seasoning curve for credit card loans does not necessarily imply that age is a material predictor of default, controlling for other risk factors. Standard seasoning or vintage curve analysis uses fixed segmentation methods where exposures do not migrate across segments as their risk profile changes. Thus, in the standard seasoning or vintage analysis, loans remain in the same segment even if the risk of the loan changes over time.

Thus, the shape of a standard seasoning curve does not represent a pure marginal age effect, but rather incorporates both loan age effects and credit quality transitions.[3] To illustrate this point, consider the following simple example. Assume that available information allows a lender to classify all non-defaulted borrowers into two groups: good (G) or poor (P). Further assume that type G borrowers have a 1% probability of default and type P borrowers have a 5%

[3] It also incorporates the effects of economic state variables. That is, the dynamic path of vintage performance will be affected by the dynamic path of the economy.

probability of default and that these probabilities are independent of loan age. Now further assume that there is a 50% probability that a type G borrower (excluding defaulted loans) will become a type B borrower.

Consider the seasoning curve for a portfolio of type G borrowers assuming that actual defaults equal expected defaults. In the first year, the default frequency will be 1%. In the second year, on average, half of the remaining portfolio will be type P borrowers and the portfolio will have an expected default rate of 3% (.5x1% + .5x5%). Thus, the seasoning curve will slope upward even though loan age is not a risk factor after controlling for credit quality. In our empirical analysis, we examine the importance of loan age and whether it continues to be a material risk factor after controlling for updated measures of credit quality.

III. Data Description and Design of the Analysis

Analysis Population

The empirical data used in this research are monthly observations of credit card accounts originated in 1999 and 2000 from selected business lines at Capital One. For the purposes of this analysis, the loans have been limited to non-rewards and non-affinity accounts, thereby ruling out product-specific factors that could otherwise distort the results. This analysis focuses on two discreet portfolio segments, low risk and high risk, that reflect the highest and lowest quality credits in the portfolio.

Within each portfolio, accounts are further grouped by cohort, i.e., the quarter in which credit card accounts were originated. To examine loan performance over a five-year window, we have eight quarterly cohorts originated over the period 1999 to 2000 and observe their performance from origination until September 2004. The method used to determine sample sizes for the data is discussed in the Appendix.

Observation Point and Outcome Window

For any group of accounts with certain common characteristics (i.e., accounts originated in the same cohort from the same business line), monthly snapshots are taken over an extended time period, usually three to five years, depending on the history of available data. To examine the seasoning effects of PD and capital, we arrange these monthly observations on the basis of account ages rather than calendar months. Cohorts are defined by the quarter of origination and age. For example, accounts originated in January 1999 and observed in January 2001 correspond to account age of 24 months, while accounts originated in March 1999 reach the age of 24 months in March 2001. In this analysis, both accounts would be included in a single observation point: cohort Q1 1999 at the age of 24 months.

The outcome window represents the 12-month window immediately following the observation point. That is, each observation point t corresponds to the observation window of [t+1, t+13]. Following the earlier example, cohort Q1 1999 at the age of 24 months corresponds to the outcome window of 25 to 37 months of age. Since the data run as late as September 2004, we have a minimum of 33 observation points/outcome windows for the youngest accounts originated in December 2000.

Chart 1 illustrates the observations and the outcome windows for quarterly cohorts of which PDs are observed quarterly. Each shaded window corresponds to an outcome period starting at an observation point, at which point the historical one-year PD is assigned. When multiple cohorts are grouped together, observations with the same age are analyzed in aggregate to examine seasoning effects. The exact framework used in our analysis allows for more granularity by taking monthly observations of monthly originations grouped into quarterly cohorts. Later in the analysis we note that functional relationships and correlations are not significantly different between cohorts; therefore, for this paper we display graphical results from all cohorts only in aggregate.

Analysis Variables

For any combination of cohort and business line at any observation point, the following measures are calculated and analyzed, either at the entire cohort level or, more relevant to this paper, at the segment level defined by different criteria. The segmentation criteria used in the analysis are described in detail in the next section.

- PD (probability of default) = % of accounts charged off during the 12-month outcome window.

  PD is measured as the number of accounts that default over a 12-month window divided by the total number of open accounts at the observation point.

- EAD (exposure at default) = the expected dollar losses for the segment if all accounts in the segment default within the 12-month outcome window.

  EAD is estimated using the actual historical increase in balances up to the period of default. The multiplication of a cohort's PD and EAD equals the expected gross dollar loss rate for the cohort.

- LGD = expected net losses as a percent EAD after accounting for recoveries.

  LGD is measured using broad industry averages for recovery rates.

IV. Segmentation Analysis

This section examines the estimation of PD using the segmentation criteria discussed above. For each combination of cohort and portfolio at any observation point, the population is divided into segments using various combinations of origination credit score, refreshed credit score, and delinquency status.

Our segmentation factors are:

- ORIGSC = credit bureau FICO score assigned to a borrower at the time of credit application

  For each account, FICO scores are typically available from all three of the major credit bureaus - Equifax, Experian, and TransUnion. To obtain ORIGSC, we use a cascaded score, i.e., use credit bureau A's score if available; otherwise use credit bureau B's score if available; and finally use score from credit bureau C if neither of the other two is available.

- REFRSC = the updated credit bureau score assigned on a monthly basis.

- DELINQ = the delinquency state of an account that has not yet defaulted, updated on a monthly basis.

  DELINQ can take on one of the following discrete values: current, 30-59 days past due, 60-89 days past due, 90-119 days past due, 120-149 days past due, 150-179 days past due. Since loans 180 days or more past due are required to be charged off, these are defaulted accounts.

- AGE = the number of statements since the account was originated.

Other variables used in the analysis are:

BALANCE = outstanding dollar balances for an account.

LINE = dollar value of an account's credit line.

REFRSC, DELINQ, AGE, BALANCE, and LINE are updated at each observation point.

Regression Model

Our prior discussion of segmentation is primarily drawn from current industry risk management practice. To confirm the importance of the role of these risk factors, we conduct some preliminary multivariate regression analyses. We estimate logistic models predicting PD. The control variables used are ORIGSC or REFRSC, AGE, DELINQ, BALANCE, and LINE.

The regressions are run separately for the low risk and high risk business portfolios. All eight quarterly cohorts are used in the regression. With two portfolios (low risk and high risk) and two credit scores (ORIGSC and REFRSC), we estimate four models of the following form:

$$\log\frac{PD}{1-PD} = \alpha_0 + \alpha_1 * AGE + \alpha_2 * SC + \alpha_3 * DELINQ + \alpha_4 * BALANCE + \alpha_5 * LINE + \varepsilon \quad (11)$$

Table 1 and 2 report the maximum likelihood estimates for the four regressions. As expected, DELINQ has a strong positive coefficient, while the score variables are negatively associated with PD. In addition, BALANCE is positively related to PD, while the coefficients on LINE are negative. Note that since we are controlling for outstanding balances, a marginal increase in LINE implies a marginal decrease in credit line utilization rate.

AGE is a much stronger positive predictor of PD in the low risk portfolio when using ORIGSC rather than REFRSC. This result suggests that much, though not all, of the relationship between loan age and default for the low risk portfolio is due to credit quality deterioration that is captured in the updated scores. Note that AGE remains a strong predictor for the high risk portfolio even when controlling for REFRSC, but the coefficient on AGE is negative. This suggests that within a portfolio of newly acquired high risk credit card accounts, poorer quality borrowers default quickly leaving a higher quality pool of residual borrowers.

Analysis of Segmentation Criteria: Credit Scores and Delinquency Status

We examine the importance of updating credit scores and delinquency status for risk segmentation for the low risk and high risk samples. Figure 1 plots PD against REFRSC and ORIGSC. The charts report PDs as an index rather than using the actual estimated PD values to avoid revealing proprietary information. Figure 1 shows the expected negative relationship between credit scores and PD. Not surprisingly, there is a steeper negative slope when using refreshed PDs, indicating that REFRSC provides important additional information about borrower quality. Note that REFRSC is particularly important in differentiating credit risk for the lower score bands in the high risk portfolio. These results indicate that borrowers who default typically go through a period of declining credit performance and migrate to lower credit scores prior to entering their year of default. However, note that some borrowers do default directly out of the high REFRSC buckets, indicating that some borrowers move to the default state relatively quickly without going through a protracted period of credit problems.

Figure 2 plots PD rates by DELINQ for the low risk and high risk portfolios. As would be expected, there is a monotonic positive relationship between the degree of delinquency and PD for both portfolios.

For the most part, Figures 1 and 2 display graphically what is already well known: credit scores provide information on borrower risk, more recent credit scores are better indicators of risk than stale credit scores, and delinquency status is a strong indicator of credit risk. It is less clear how well updated credit scores differentiate risk controlling for DELINQ. While it is reasonable to expect that updated credit scores will differentiate risk among current borrowers, it is less clear whether credit scores will help in predicting PD among delinquent borrowers. In particular, a credit card lender generally becomes aware of a delinquency prior to its incorporation into a bureau credit score. In other words, DELINQ provides updated information not necessarily reflected in REFRSC. In addition, the credit score is built to predict the likelihood of a borrower becoming seriously delinquent and a different model might be relevant for predicting the transition from seriously delinquent to default.

Figure 3 displays results from segmenting the portfolios by REFRSC and DELINQ for both portfolios. REFRSC continues to rank risk for the current bucket and for lower stages of delinquency. This can be seen by the negative slope for these categories. However, at later stages of delinquency, PD is no longer monotonically declining in REFRSC. These results suggest that REFRSC and DELINQ are jointly important segmentation factors for the current bucket and earlier stage delinquencies, while REFRSC may not be a useful segmentation factor for later stage delinquencies.

V. Seasoning Effects

Figure 4 shows seasoning curves that plot PD against AGE for the low risk portfolio and the high risk portfolio. The low risk seasoning curve is monotonically increasing up to 33 months with a concave shape. This is in sharp contrast to the high risk portfolio, which has a sharp upward slope that peaks after seven months and then is downward sloping. This is consistent with the evidence in the preceding section indicating that higher quality loans often go through a period of decline prior to entering their year of default. It also suggests that newly originated high risk portfolios contain a distinct subpopulation of borrowers that move quickly to default.

The difference in seasoning patterns for the low risk vs. high risk portfolio suggests that seasoning patterns differ depending on portfolio credit quality. To investigate this issue further, we examine seasoning curves controlling for either ORIGSC or REFRSC.

Figure 5 displays seasoning curves controlling for ORIGSC. For portfolios in bands with ORIGSC of 720 or less, the seasoning curves display an upward slope that peaks in the 18-20 month range and then flattens out. For higher score bands, the seasoning curves are monotonically increasing out to 33 months and the curves are concave. These seasoning curves for highly scored borrowers have the same general seasoning pattern as the seasoning curve of the pooled low risk portfolio, indicating that the seasoning peak in default rates for portfolios with very high credit scores at origination is at least 33 months.

The results for the high risk portfolio indicate that ORIGSC rank orders borrower performance for high risk credits. However, there are some distinct differences in the seasoning pattern of the high risk portfolio as compared to the low risk portfolio. All of the high risk portfolio credit bands display a common seasoning pattern, with PDs rising very rapidly, peaking at seven months, and then falling. Interestingly, the PDs for all of the score bands converge to a very narrow range at the end of 33 months. That is, conditional on surviving for nearly three years, ORIGSC adds little information on PD.

As discussed earlier, seasoning curves based on segmentation criteria that reflect credit quality at origination implicitly include the effects of changes in credit quality prior to default. Does loan age continue to be a factor when we separate out the portfolio by updated credit quality information? To examine this question, Figure 6 plots seasoning curves by REFRSC within the low risk and high risk portfolios.

As expected, REFRSC provides better risk separation compared to ORIGSC. For the low risk portfolio, there is a substantial upward slope to the seasoning curve for REFRSC of 740 or below. These curves either peak at around two years or are monotonically increasing up to 33 months. The seasoning curves for accounts with REFRSC above 740 are flat.

For the high risk portfolio, seasoning curves controlling for REFRSC show the familiar seven-month peak in the seasoning curve for accounts with scores of 580 or less. For these score bands, PDs fall for a few months after the peak and then quickly flatten out. Note that for higher score bands the seasoning curves are flat and then have a slight downward slope that eventually flattens out.

Figure 7 displays seasoning curves for the two portfolios controlling for DELINQ, which is an alternative updated measure of credit quality. For the low risk portfolio, conditional on the loan being current, AGE does not appear to have an independent effect on PD after controlling for delinquency. There remains a monotonically upward seasoning curve for earlier stage delinquencies. However, for delinquencies beyond 60 DPD the seasoning curves become essentially flat except for very young loans.

For the high risk portfolio shown in the second graph, seasoning curves display a downward slope for the first few months the loans are on the books and then are either flat or downward sloping. For the high risk portfolio, loans that become delinquent at an early age have a very high PD, but loans that become delinquent after having been on the books for some time are less likely to default. This evidence, combined with the previous seasoning curves for the high risk portfolio, indicates that there is a subpopulation within the newly booked high risk accounts that moves quickly to delinquency status and subsequently default. Once these accounts are purged from the sample, the remaining accounts display improved credit quality.

VI. Regulatory Capital and Segmentation

In this section, we calculate Basel II regulatory capital requirements associated with the alternative segmentation schemes discussed above. As discussed previously, capital requirements in the model fall if the segmentation system produces better risk separation within the portfolio. For regulatory purposes, this has the advantage that banks have an incentive to improve their ability to differentiate risk. Moreover, the effect on capital requirements can be interpreted as one measure of the relative quality of risk separation when using alternative segmentation systems.

The PD used for our calculations are the same as those we have used in our graphs: PD is averaged over our entire sample, broken down by the drivers of the particular segmentation. For example, in a score-only segmentation PD is averaged for all the data cells within a particular score band within a portfolio. These averages are account weighted. For a score by delinquency segmentation, PD is averaged over the score band by delinquency by portfolio combination. LGD is also averaged. Here the averages are taken over each of the portfolios, high risk or low risk. The weighting is by the number of defaults in the data cell. EAD is the sum of the actual EADs in each segment.

The proposed Basel II capital regulations require an upward adjustment to PD where seasoning effects are material. This reflects the view that for some types of credits, significant credit deterioration does not show up in defaults when the credits are young and that capital should cover these types of losses. Adjusting PD upward for seasoning is an issue in our data only when AGE is a segmenting factor. For young accounts of a particular age, we adjusted PD upward by averaging over all data cells from that unseasoned age up to the average maturity of that portfolio. Where the average maturity exceeds the age of our sample, we extrapolated. We used a flat extrapolation from the last age or average of the last few ages for high risk and a linear extrapolation for low risk segments. These should be conservative, since delinquencies are declining or are increasing at a falling rate as seen in our figures.

The PDs, LGDs, and EADs using various segmentation approaches are then entered into the current Basel II risk-weight formulas for credit card exposures to calculate regulatory capital for each portfolio. These regulatory capital calculations for different segmentation methods are indexed relative to the capital requirements with no segmentation to avoid revealing proprietary information. The results of these calculations for the low risk and high risk portfolios are shown in Figure 8.

There is a relatively small reduction in regulatory capital requirements for the low risk portfolio when segmenting by ORIGSC, and this reduction is negligible for the high risk portfolio. These relatively small effects are in part due to the limitations of ORIGSC for differentiating risk and in part due to the relative homogeneity at origination of each of these portfolios. There is a substantial drop in capital requirements for both portfolios when using REFRSC in the segmentation system. There is an additional drop in capital requirements when segmenting by ORIGSC and DELINQ. Note that this drop is more substantial for the lower credit quality portfolio. There is a further drop when using REFRSC and DELINQ, but here the drop is more substantial for the higher credit quality portfolio.

These results point to the significant improvements in differentiating risk by moving to a segmentation method that includes scores and delinquency status as segmenting criteria. For higher credit quality portfolios, our data suggest that it is important to combine updated credit scores with delinquency status. For lower credit quality portfolios that segment by DELINQ, our evidence suggests that there are relatively small gains from using REFRSC rather than ORIGSC in the segmentation system.

Our results for seasoning suggest negligible changes when applying seasoning adjustments compared to a segmentation system that includes delinquency and score factors. This may not be surprising, since our previous analysis showed that AGE was not an important risk factor for lower quality credits after controlling for delinquency status, and lower quality credits typically account for a large share of the capital calculation. However, since our previous results showed that AGE remains a risk factor for higher quality credits after controlling for scores and delinquency, this negligible effect of seasoning adjustments on the capital calculation depends on the composition of the specific credit card portfolio.

To summarize, our calculations indicate substantial capital relief incentives for improved risk differentiation when segmenting the credit card portfolio. They also indicate the significant gains in risk differentiation from using updated credit quality information, since the bulk of the regulatory capital relief occurs when moving to a segmentation system that provides updated risk measures such as refreshed score and delinquency status.

VII. Summary

This paper examines methods for differentiating the likelihood of default among credit card borrowers and how these alternative methods affect capital requirements for the portfolio in an ASRF model. Using proprietary data on two large credit card portfolios with different average credit quality characteristics, we demonstrate the importance of updating credit quality information in differentiating risk among borrowers and for calculating tail risk of a credit card portfolio. In particular, we find important benefits to controlling for both delinquency status and updated credit scores, and these benefits are greater for relatively high credit card portfolios.

Finally, we find that while credit card defaults are strongly correlated with loan age, this correlation is significantly weak after controlling for updated credit quality, particularly for lower credit quality loans. Moreover, after segmenting the portfolio using updated credit quality measures, adjusting PD for loan-age effects had little impact on estimated portfolio capital requirements.

Bibliography:

Basel Committee on Banking Supervision (2004), “International Convergence of Capital Measurement and Capital Standards: A Revised Framework,” available at www.bis.org.

Lang, William, W. and Anthony M. Santomero (2004), "Risk Quantification of Retail Credit: Current Practices and Future Challenges," in Monetary Integration, Markets and Regulation Research in Banking and Finance, Volume 4, 1–15, Elsevier Ltd.

Laurent, Marie-Paule (2004), "Asset Return Correlation in Basel II: Implications for Credit Risk Management," Working Papers CEB 04-017.RS, Université Libre de Bruxelles, Solvay Business School, Centre Emile Bernheim (CEB).

Merton, Robert C. (1974). “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates,” *Journal of Finance* 29, 449-470.

Risk Management Association Report, “Retail Credit Economic Capital Estimation – Best Practices,” February 2003.

Treacy, William F., and Mark S. Carey (1998). “Credit Risk Rating at Large U.S. Banks,” *Federal Reserve Bulletin* 84, 897-921.

Vasicek, Oldrich (2002). “Loan Portfolio Value,” *Risk* 15, 160-162.

Charts, Tables and Figures

Chart 1

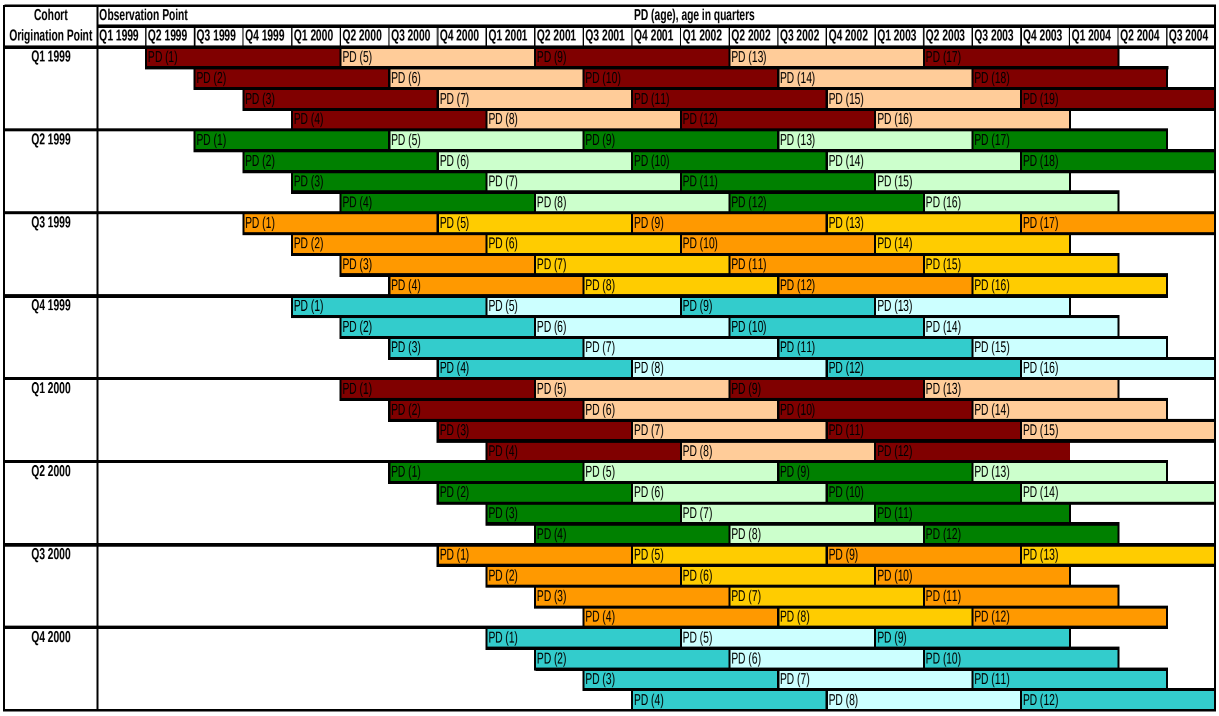


Table 1a: Maximum Likelihood Estimates of Logistic Regressions

Origination Score Model

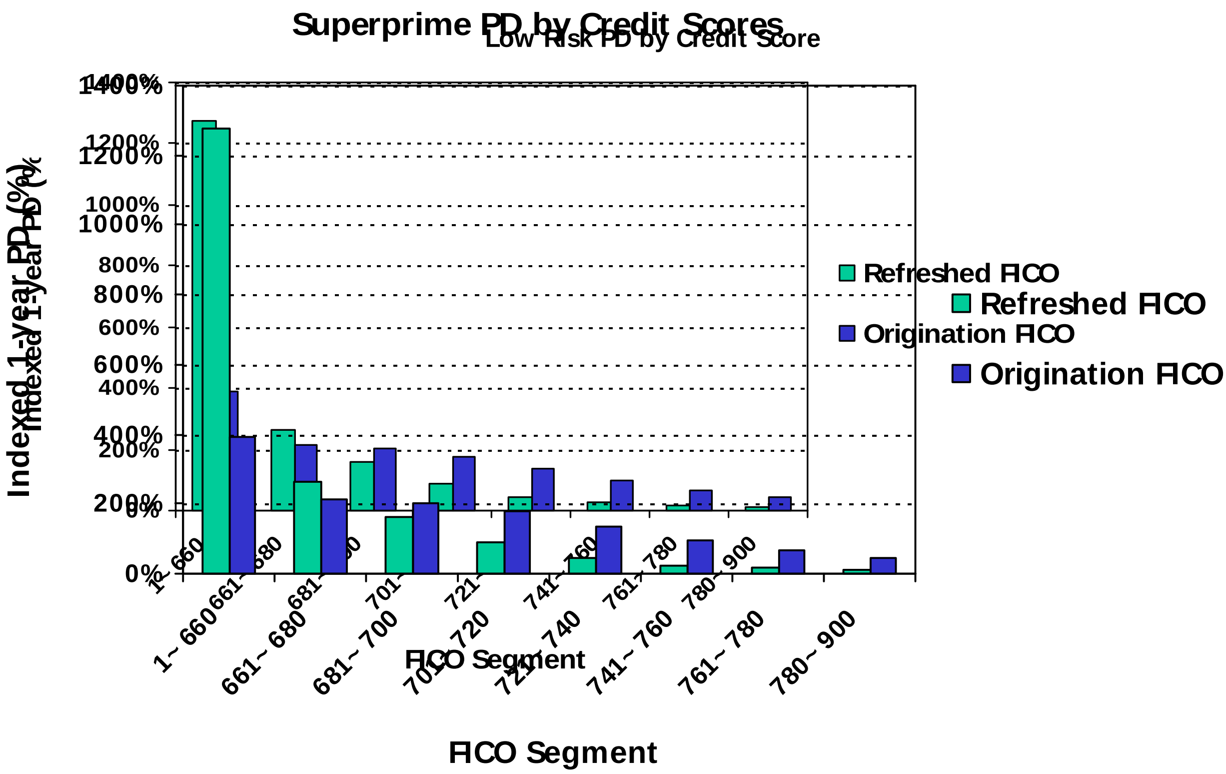
Model	Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Odds Ratio
Low Risk	INTERCEPT	-1.9695	0.0583	1,141	<.0001	
	AGE	0.0203	0.000171	14,088	<.0001	1.021
	ORIGINATION SCORE	-0.0043	0.000083	2,712	<.0001	0.996
	DELINQ	1.3908	0.00336	171,813	<.0001	4.018
	BALANCE	0.000217	1.42E-06	23,307	<.0001	1
	LINE	-0.00011	1.32E-06	6,697	<.0001	1
High Risk	INTERCEPT	-1.0923	0.0032	116,824	<.0001	
	AGE	-0.0215	0.000035	376,252	<.0001	0.979
	ORIGINATION SCORE	-0.00182	6.07E-06	90,275	<.0001	0.998
	DELINQ	0.8784	0.000422	4,329,955	<.0001	2.407
	BALANCE	0.00136	3.62E-06	142,069	<.0001	1
	LINE	-0.0009	3.17E-06	80,544	<.0001	1

Table 1b: Maximum Likelihood Estimates of Logistic Regressions

Refreshed Score Model

Model	Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Odds Ratio
Low Risk	INTERCEPT	7.6245	0.0294	67,288	<.0001	
	AGE	0.00632	0.000171	1,362	<.0001	1.006
	REFRSC	-0.0178	0.000048	140,817	<.0001	0.982
	DELINQ	0.8594	2.68E-03	102,871	<.0001	2.362
	BALANCE	0.000168	1.26E-06	17,659	<.0001	1
	LINE	-0.00008	1.17E-06	4,618	<.0001	1
High Risk	INTERCEPT	3.4976	0.00421	691,523	<.0001	
	AGE	-0.01	0.000038	68,791	<.0001	0.99
	REFRSC	-0.00931	8.08E-06	1,326,018	<.0001	0.991
	DELINQ	0.7177	0.000396	3,286,676	<.0001	2.05
	BALANCE	0.00172	4.13E-06	173,542	<.0001	1.002
	LINE	-0.0019	3.61E-06	276,916	<.0001	0.998

Figure 1


Superprime PD by Credit Scores
Low Risk PD by Credit Score
Indexed 1-year PD (%)
1400%
1200%
1000%
800%
600%
400%
200%
0%
Refreshed FICO
Origination FICO
1~660
661~680
681~700
701~720
721~740
741~760
761~780
780~900
FICO Segment

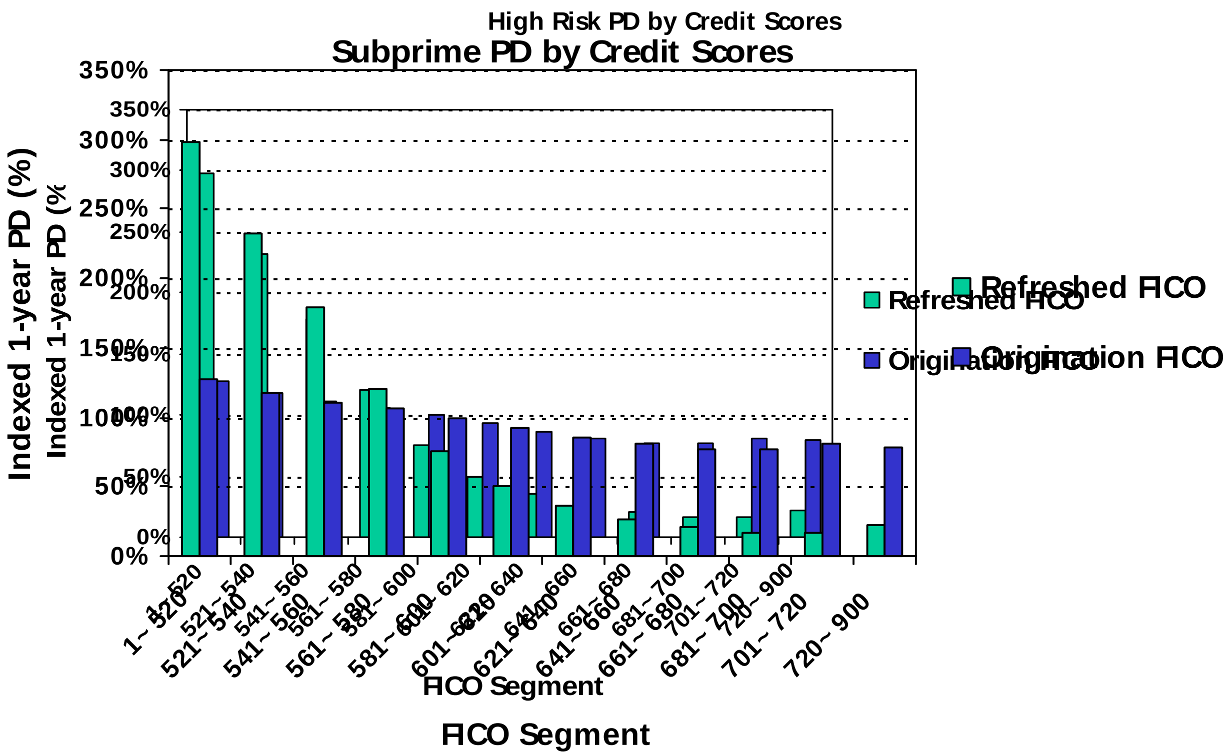
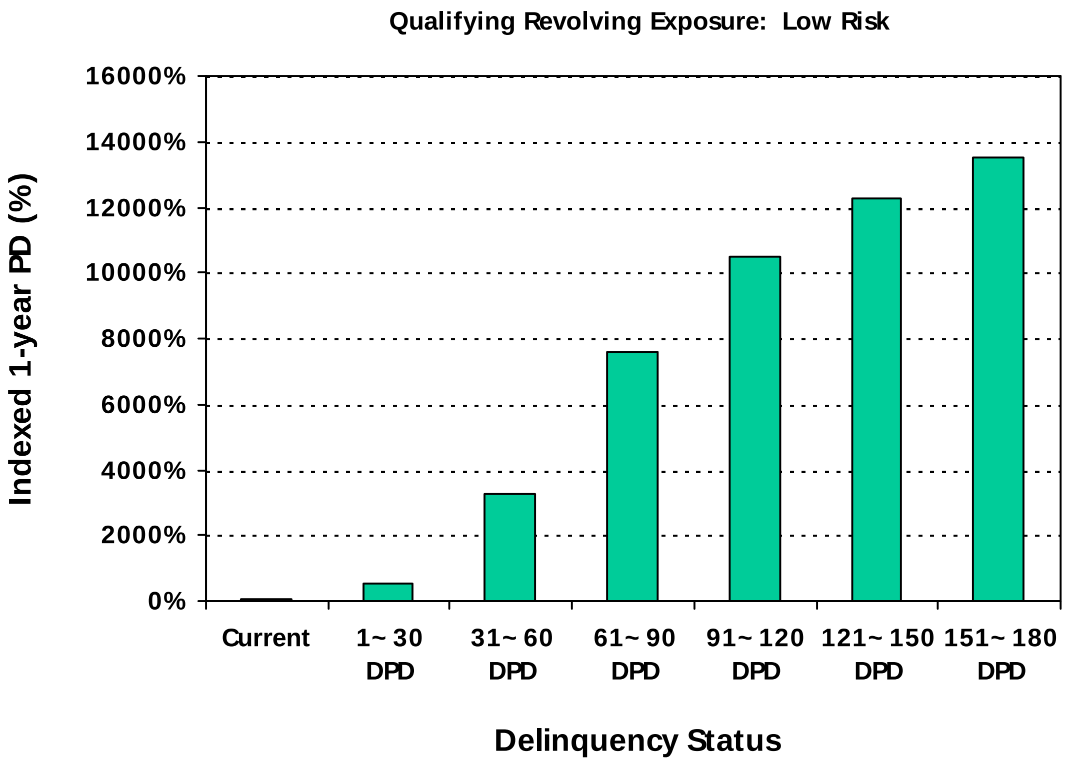

High Risk PD by Credit Scores
Subprime PD by Credit Scores
Indexed 1-year PD (%)
350%
300%
250%
200%
150%
100%
50%
0%
Refreshed FICO
Origination FICO
1~520
521~540
541~560
561~580
581~600
601~620
621~640
641~660
661~680
681~700
701~720
720~900
FICO Segment

Figure 2


Qualifying Revolving Exposure: Low Risk
Indexed 1-year PD (%)
16000%
14000%
12000%
10000%
8000%
6000%
4000%
2000%
0%
Current
1~ 30 DPD
31~ 60 DPD
61~ 90 DPD
91~ 120 DPD
121~ 150 DPD
151~ 180 DPD
Delinquency Status

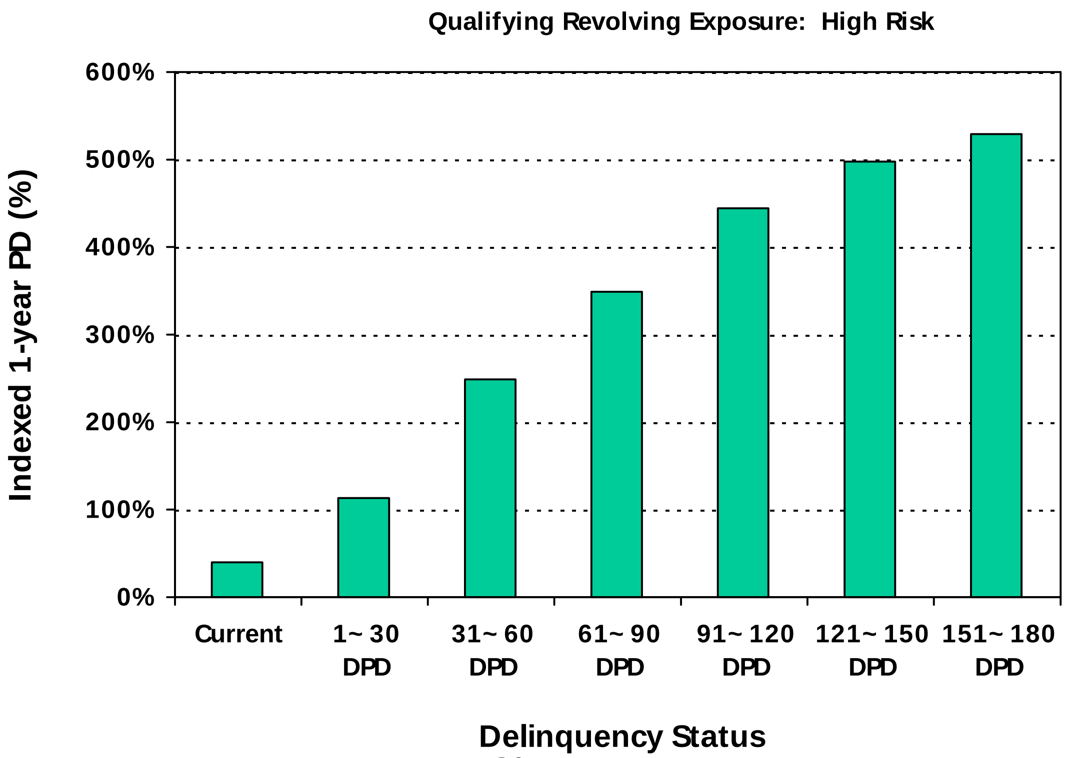
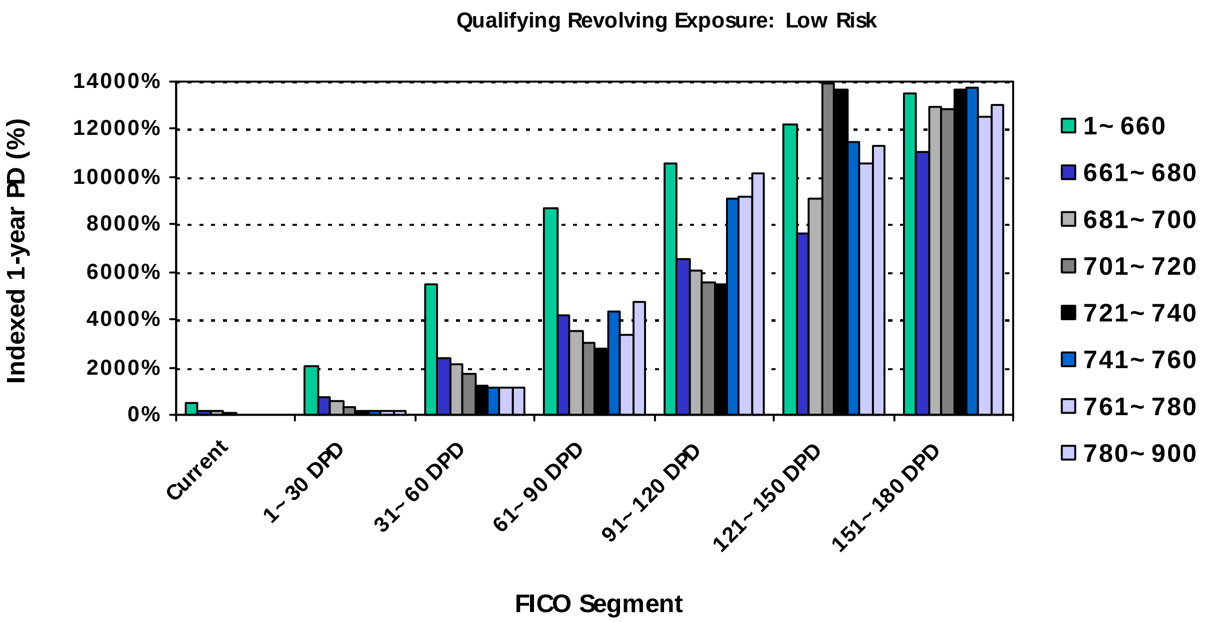

Qualifying Revolving Exposure: High Risk
Indexed 1-year PD (%)
600%
500%
400%
300%
200%
100%
0%
Current
1~ 30 DPD
31~ 60 DPD
61~ 90 DPD
91~ 120 DPD
121~ 150 DPD
151~ 180 DPD
Delinquency Status

Figure 3

Low Risk Indexed PD by Refreshed Score and Delinquency



High Risk Indexed PD by Refreshed Score and Delinquency

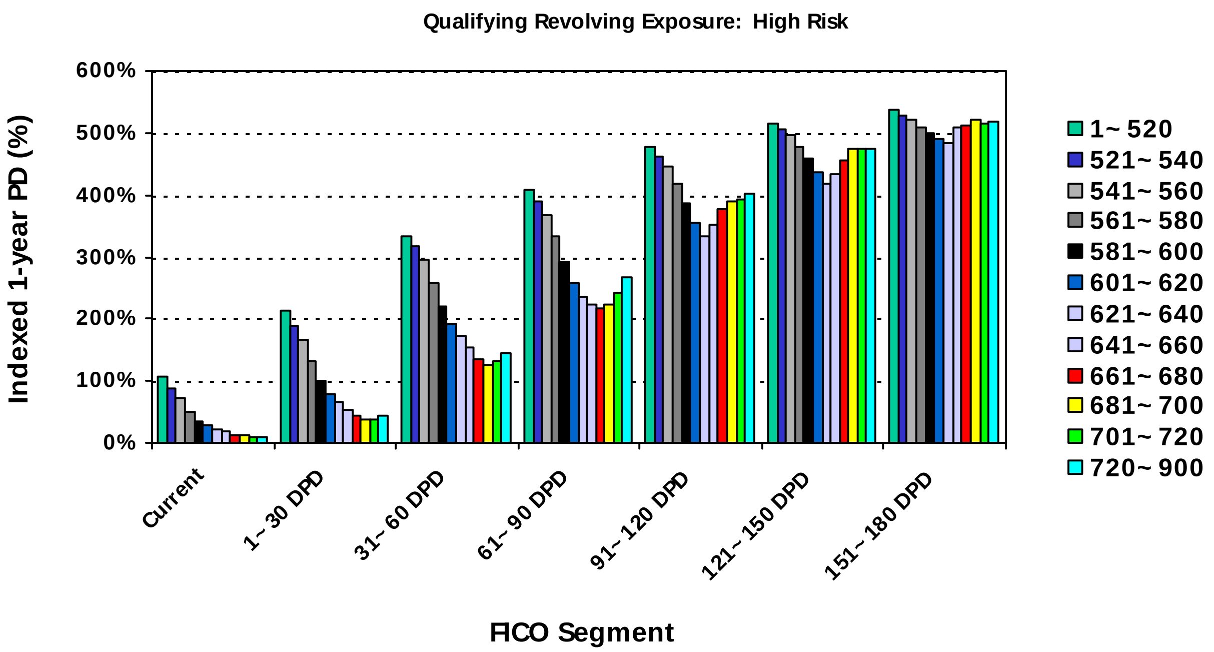
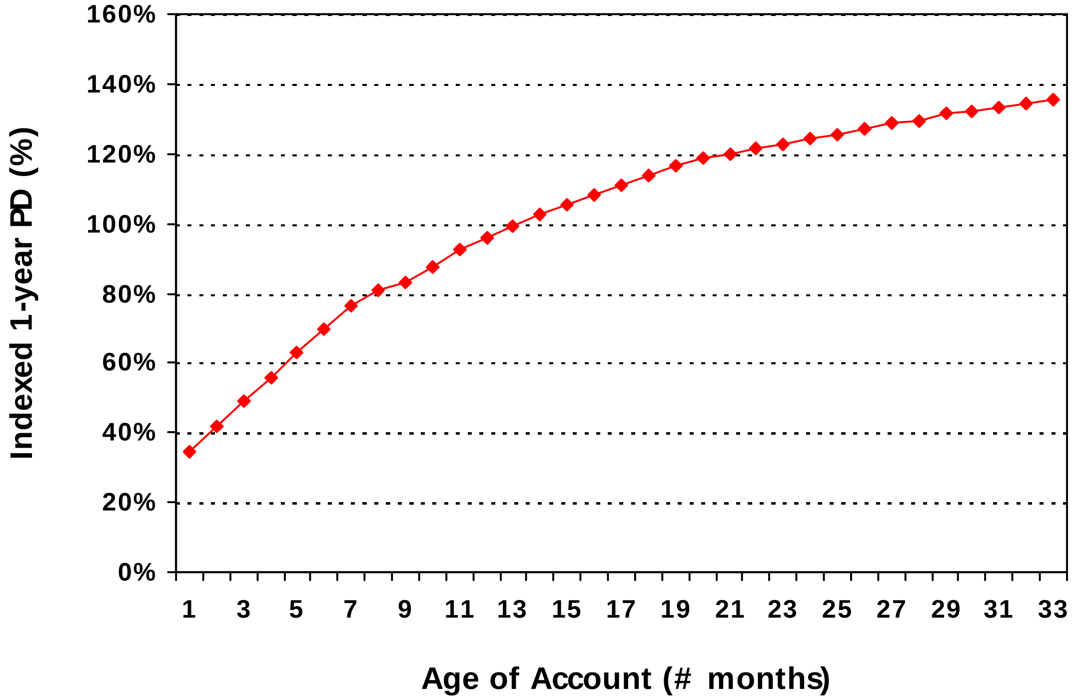


Figure 4

Low Risk Indexed PD by Age



High Risk Indexed PD by Age

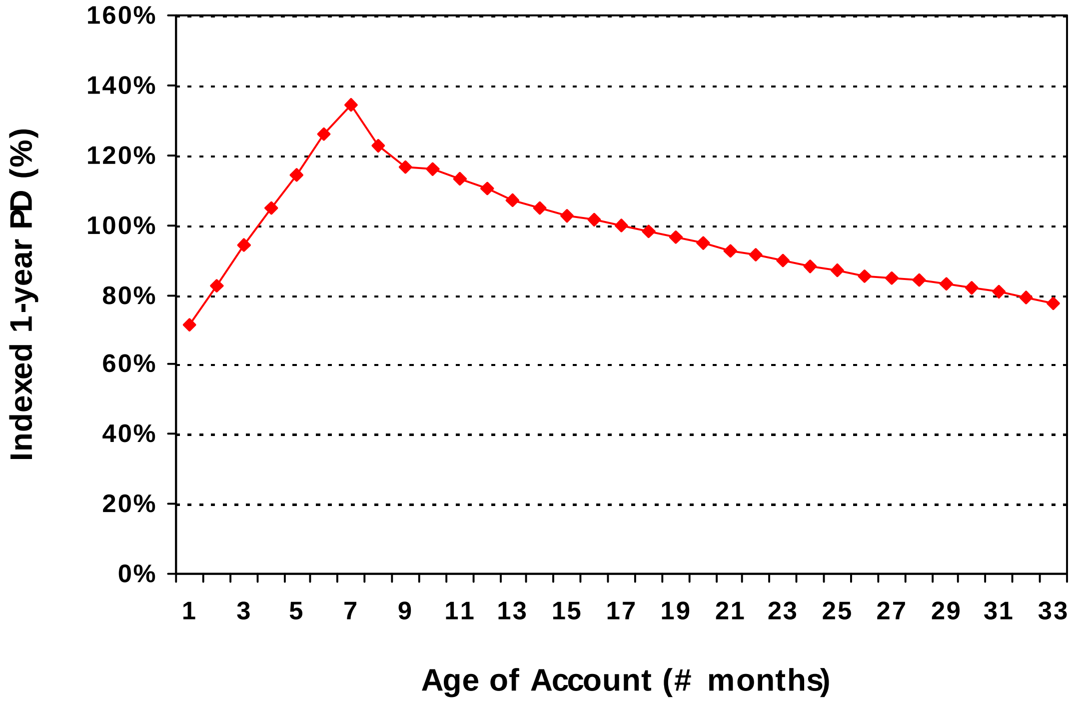
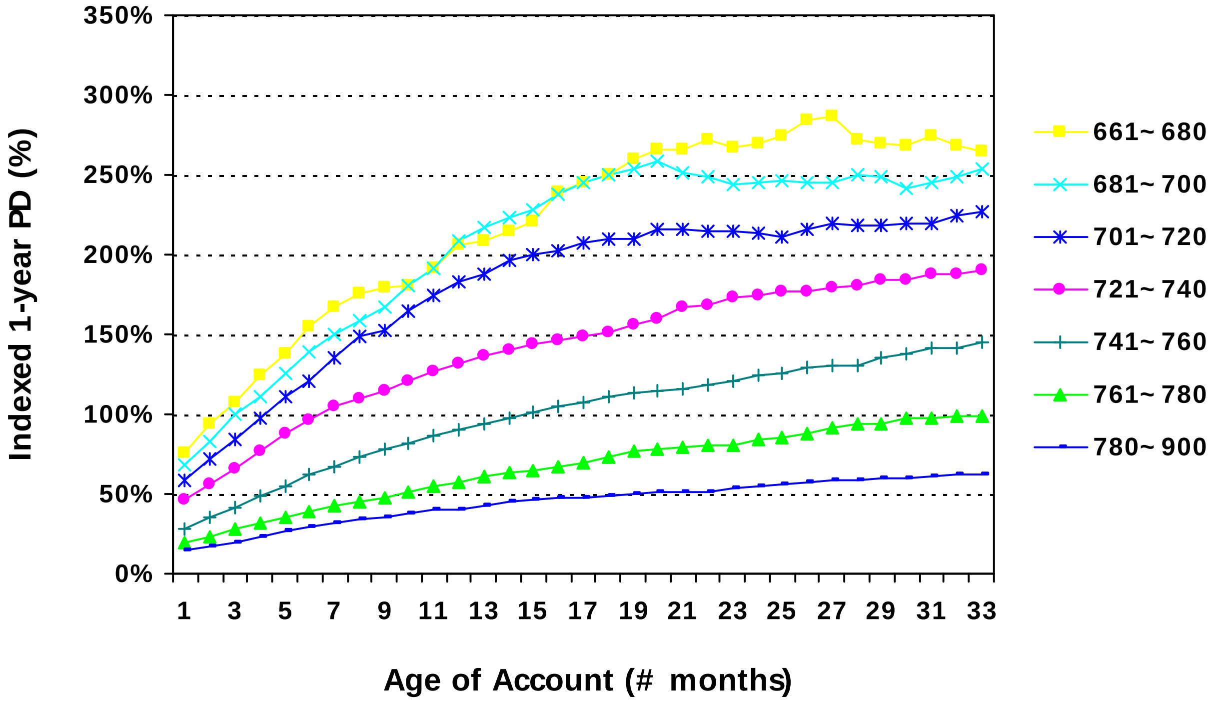


Figure 5

Low Risk Indexed PD by Origination Score and Age



High Risk Indexed PD by Origination Score and Age

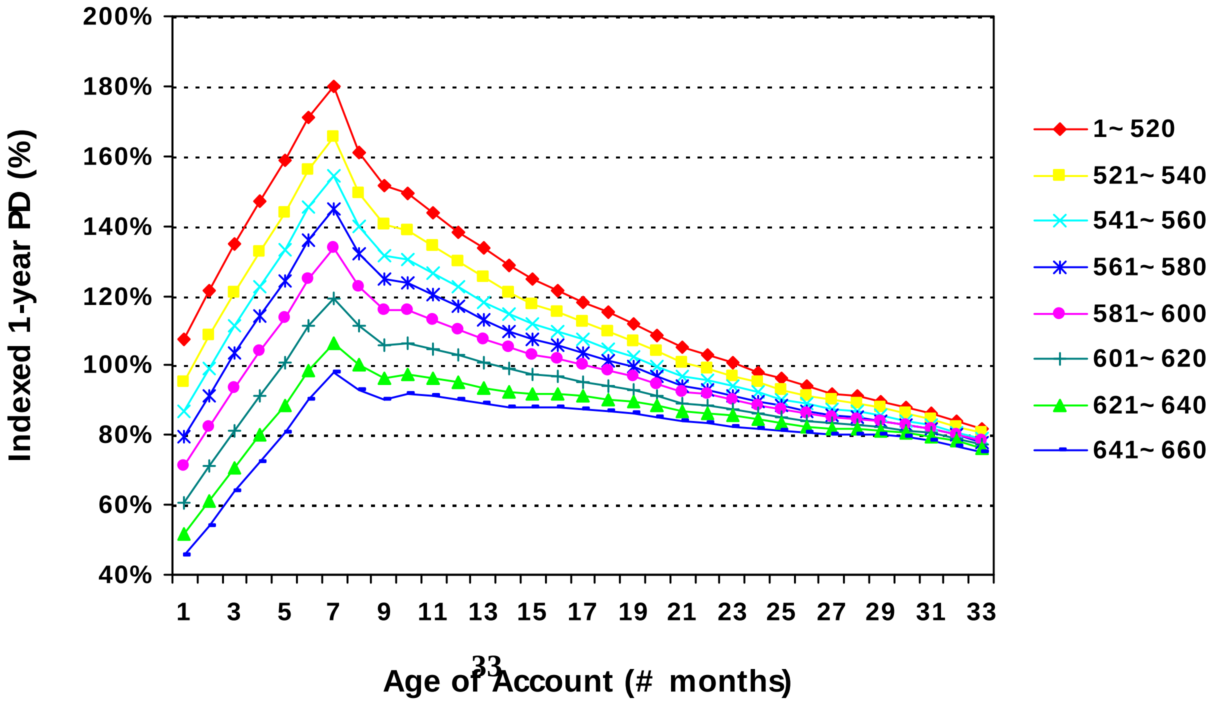
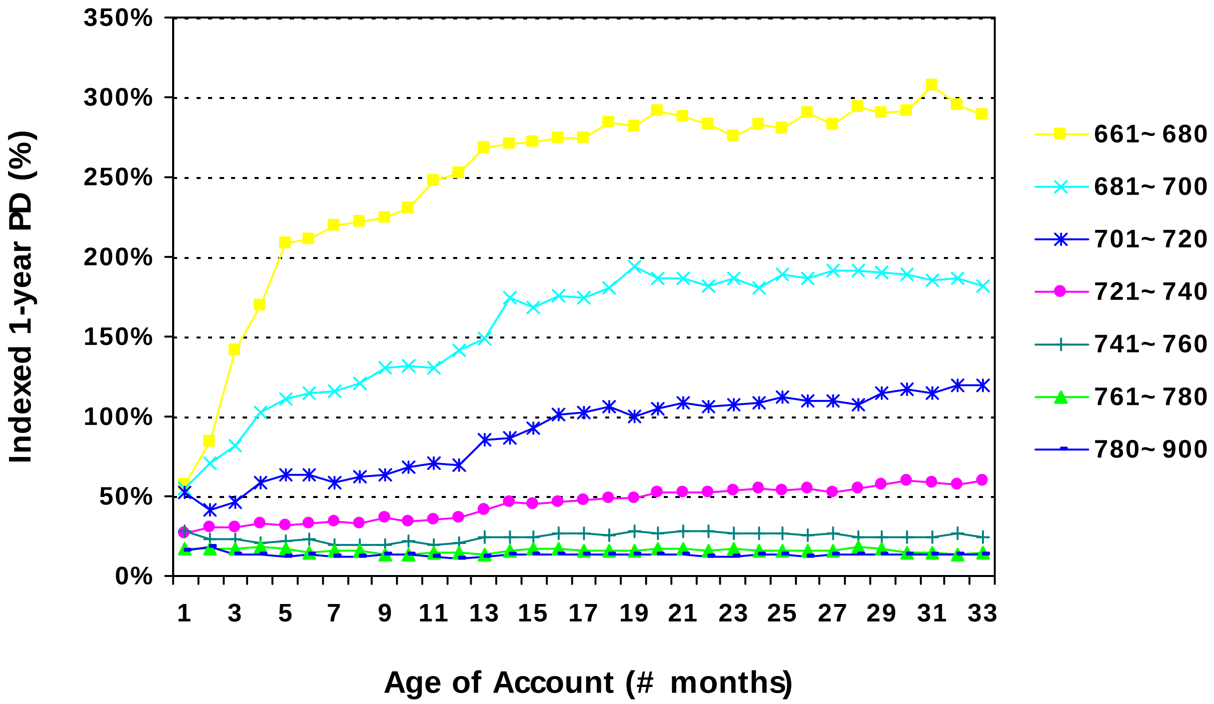


Figure 6

Low Risk PD by Refreshed Score and Age



High Risk PD by Refreshed Score and Age

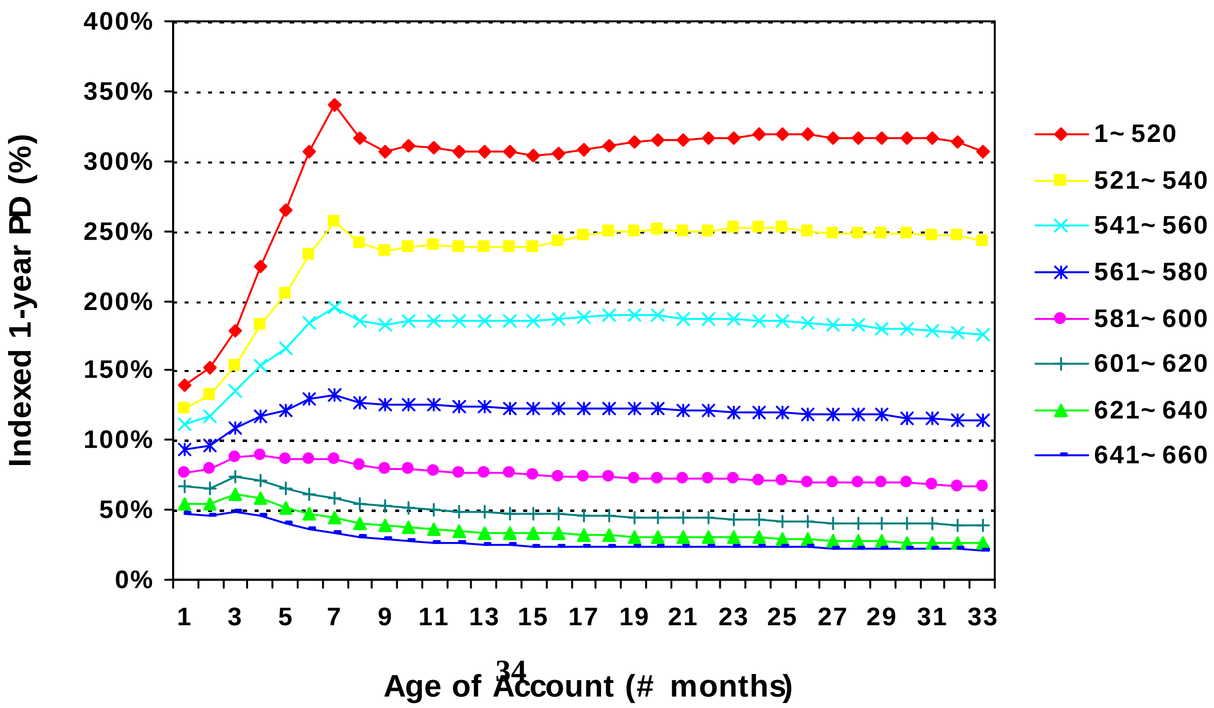
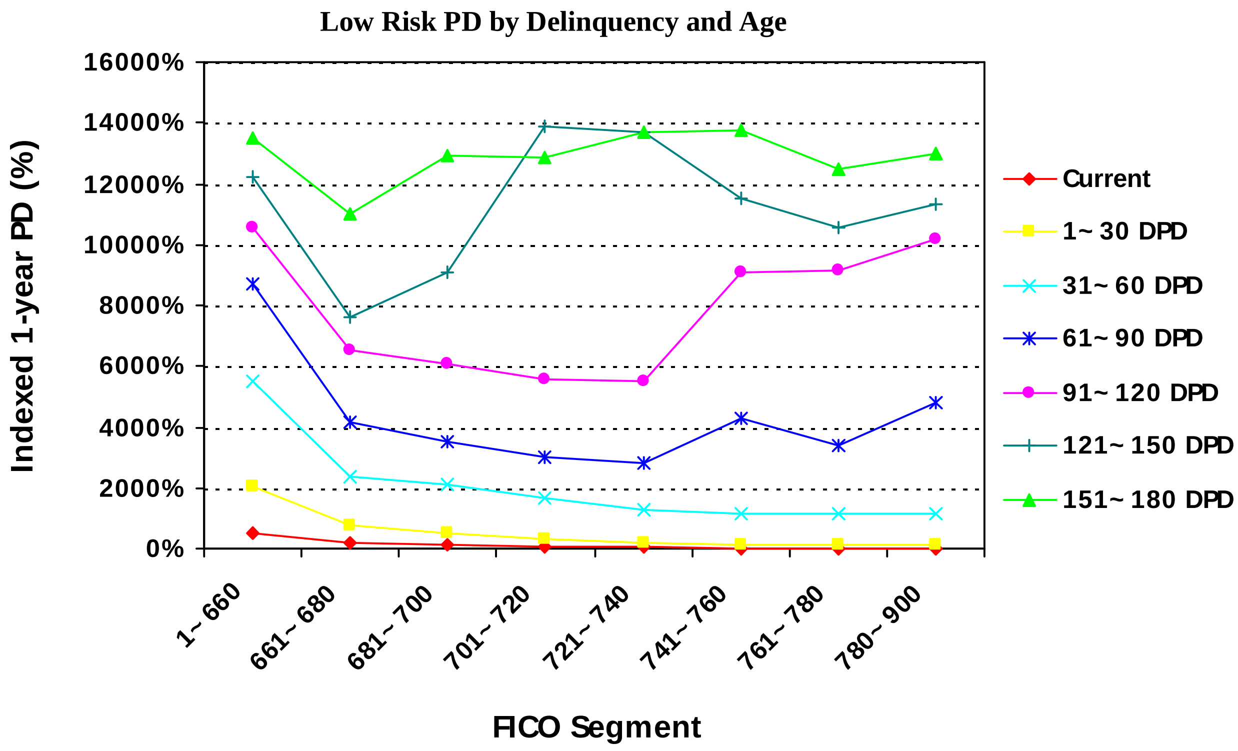


Figure 7


Low Risk PD by Delinquency and Age
16000%
14000%
12000%
10000%
8000%
6000%
4000%
2000%
0%
Indexed 1-year PD (%)
1~660
661~680
681~700
701~720
721~740
741~760
761~780
780~900
FICO Segment
Current
1~30 DPD
31~60 DPD
61~90 DPD
91~120 DPD
121~150 DPD
151~180 DPD

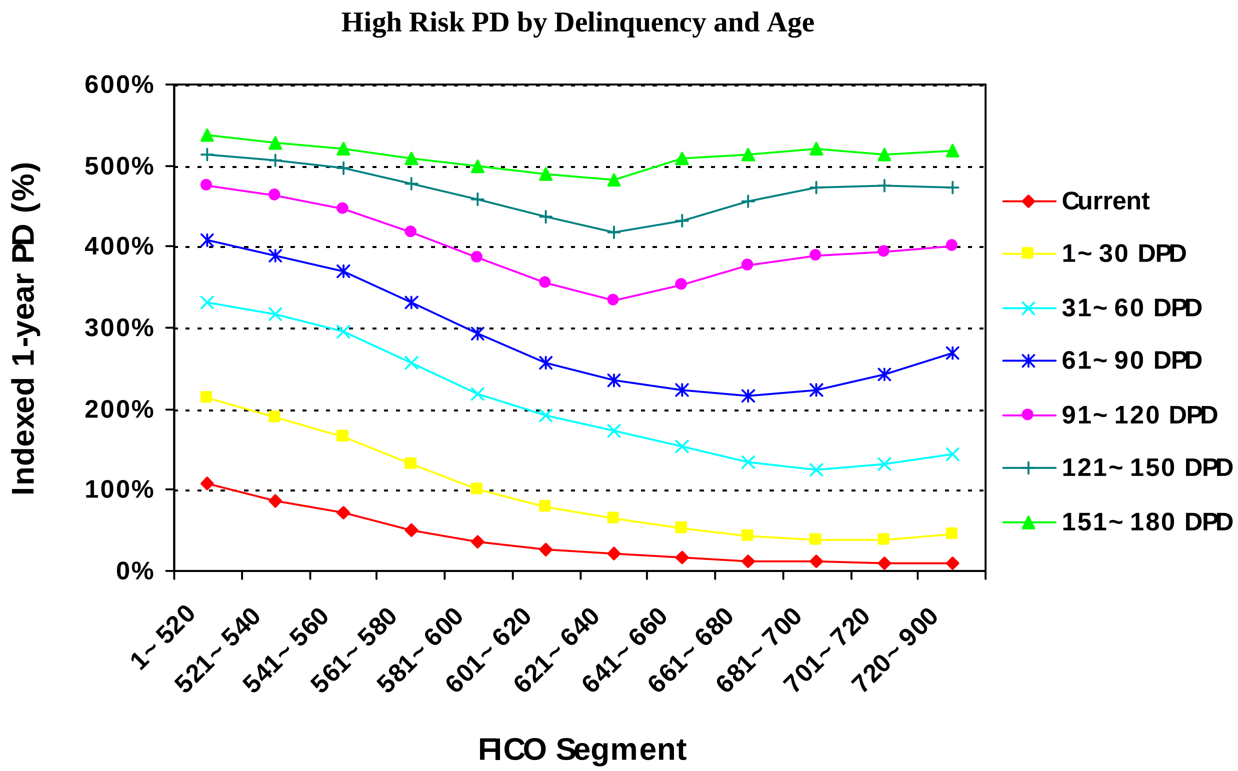
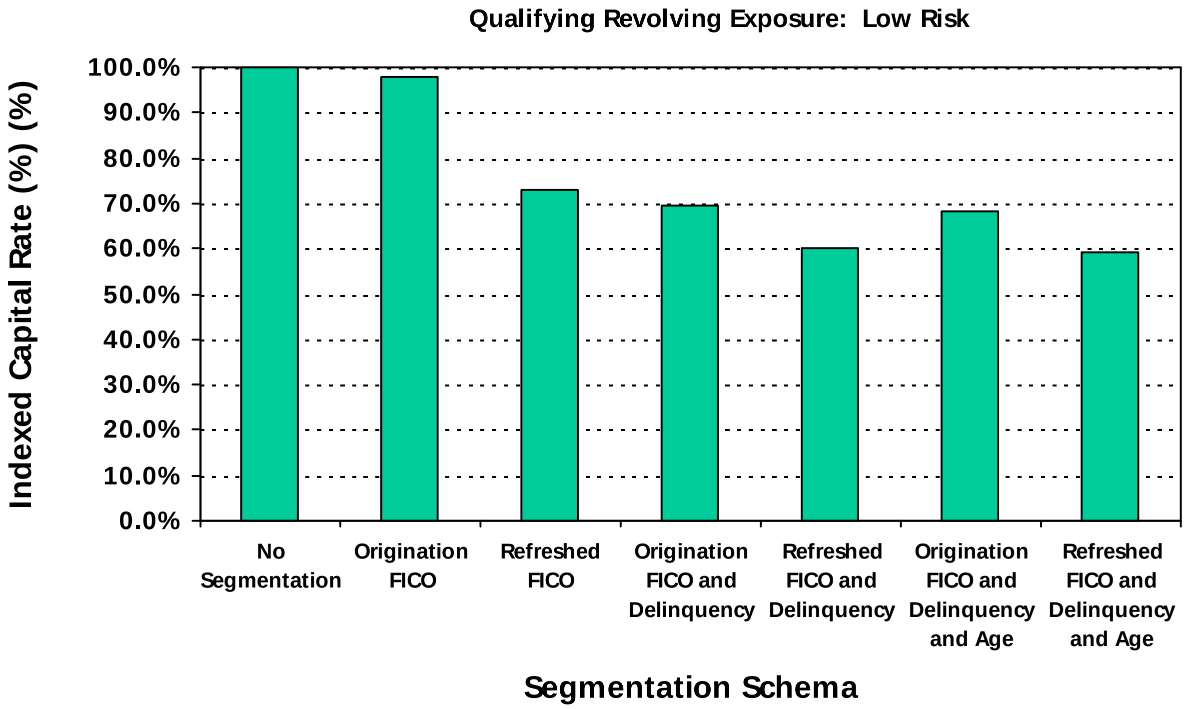
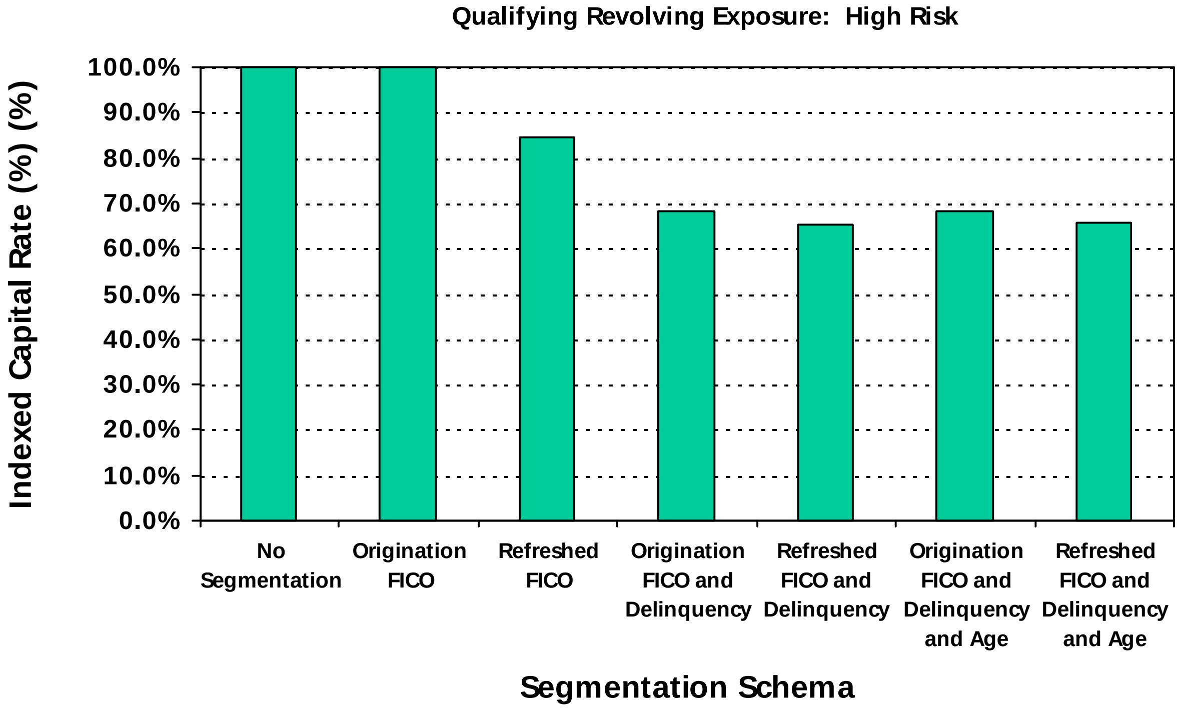

High Risk PD by Delinquency and Age
600%
500%
400%
300%
200%
100%
0%
Indexed 1-year PD (%)
1~520
521~540
541~560
561~580
581~600
601~620
621~640
641~660
661~680
681~700
701~720
720~900
FICO Segment
Current
1~30 DPD
31~60 DPD
61~90 DPD
91~120 DPD
121~150 DPD
151~180 DPD

Figure 8

Capital Based on Long Term PD (Low Risk)



Capital Based on Long Term PD (High Risk)



Appendix: Sample Size Determinations

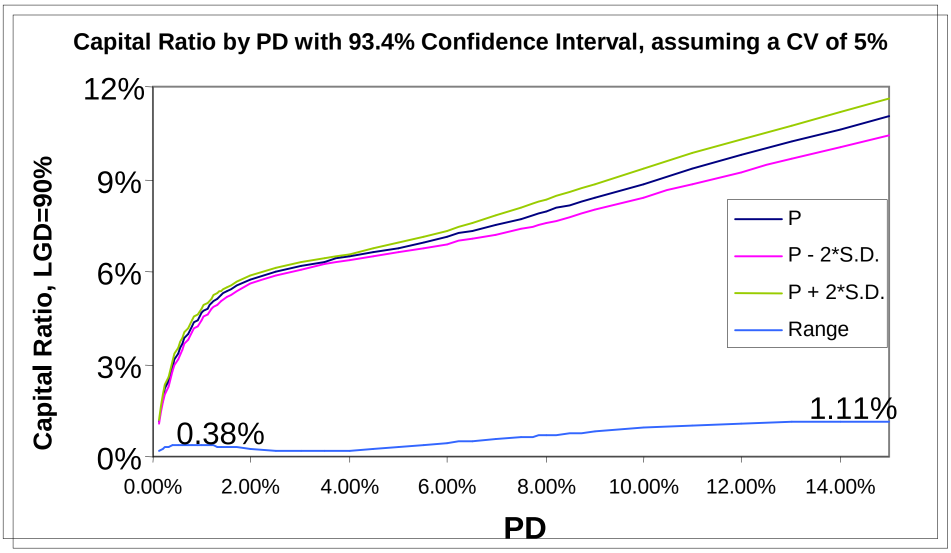
Accounts were pulled within each band of ORIGSC with a sample size sufficient to produce reliable estimates of the PD over the lifetime of each cohort. This required an initial analysis of the expected lifetime PD per score band to determine the required sample sizes. A good benchmark was to use the coefficient of variation:

CV = standard deviation/mean = sqrt(1-p)/sqrt(pn), where n is the sample size.

By using the CV rather than the standard deviation, the level of variance is within a certain percentage of the average PD. This is especially important for very low PD segments, since the effect on capital of small changes in the PD is the greatest in the low-PD segments. Sample sizes for one-year PDs, with a CV of 10% are as follows.

PD	N = (1-PD)/(PD*CV^2)	Standard Deviation = sqrt(PD*(1-PD)/N)	PD - 2*S.D.	PD + 2*S.D.	K-IRB PD	K-IRB PD - 2*S.D.	K-IRB PD + 2*S.D.
0.10%	399,600	0.005%	0.09%	0.11%	1.14%	1.06%	1.23%
0.25%	159,600	0.013%	0.23%	0.28%	2.18%	2.03%	2.32%
0.50%	79,600	0.025%	0.45%	0.55%	3.33%	3.14%	3.51%
0.75%	52,933	0.038%	0.68%	0.83%	4.10%	3.90%	4.28%
1.00%	39,600	0.050%	0.90%	1.10%	4.65%	4.45%	4.83%
1.25%	31,600	0.063%	1.13%	1.38%	5.05%	4.87%	5.21%
1.50%	26,267	0.075%	1.35%	1.65%	5.35%	5.18%	5.50%
1.75%	22,457	0.088%	1.58%	1.93%	5.58%	5.43%	5.71%
2.25%	17,378	0.113%	2.03%	2.48%	5.90%	5.78%	6.00%
2.75%	14,145	0.138%	2.48%	3.03%	6.11%	6.00%	6.20%
3.25%	11,908	0.163%	2.93%	3.58%	6.27%	6.17%	6.36%
3.75%	10,267	0.188%	3.38%	4.13%	6.41%	6.31%	6.51%

The upper and lower bounds on the PD above (PD – 2xS.D. and PD + 2xS.D., respectively) provide a 93.4% confidence interval, using the normal approximation to the binomial. This translates into a confidence interval on the capital ratio, using the Basel II capital function for credit cards.



However, for larger PDs the standard deviation and width of the resulting confidence interval is beyond an acceptable level. For higher PD segments the sample size was fixed at approximately 10,000 accounts.

PD	N	Standard Deviation = sqrt(PD*(1-PD)/N)	PD - 2*S.D.	PD + 2*S.D.	K-IRB		
					PD	PD - 2*S.D.	PD + 2*S.D.
4%	10,000	0.196%	3.61%	4.39%	6.48%	6.37%	6.59%
5%	10,000	0.218%	4.56%	5.44%	6.77%	6.64%	6.91%
6%	10,000	0.237%	5.53%	6.47%	7.11%	6.94%	7.30%
7%	10,000	0.255%	6.49%	7.51%	7.51%	7.30%	7.73%
8%	10,000	0.271%	7.46%	8.54%	7.95%	7.71%	8.20%
9%	10,000	0.286%	8.43%	9.57%	8.41%	8.14%	8.67%
10%	10,000	0.300%	9.40%	10.60%	8.87%	8.59%	9.15%
15%	10,000	0.357%	14.29%	15.71%	11.03%	10.75%	11.31%
20%	10,000	0.400%	19.20%	20.80%	12.73%	12.49%	12.95%

Sample size reductions in the segments due to defaults, attrition, and score migration were also considered; this is especially true for defaults in the higher PD segments. The total sample size in each segment within each cohort was increased so that the segment (using either REFRSC or ORIGSC) maintained the minimum sample size throughout the sample time frame.

To reduce the sample required, accounts that default by the end of the observation period (three years) were sampled at 100% with only a 10% sample of accounts that do not default by the end of the period. The under-sampling of non-defaulted accounts by random sampling leads to increased variance and a biased estimate. However, taking every 10th account does not lead to the problem of increased variance, unless there is some cyclic nature in the order of the accounts. Over the interim, the number of accounts that default within each one-year time horizon and the total number of non-defaulted accounts (from the sample times 10) at the beginning of the year were used to produce unbiased, consistent estimates of the one-year PD.