WORKING PAPERS

RESEARCH DEPARTMENT

WORKING PAPER NO. 10-17

CAN BANKS CIRCUMVENT MINIMUM CAPITAL REQUIREMENTS? THE CASE OF MORTGAGE PORTFOLIOS UNDER BASEL II

Christopher Henderson
Federal Reserve Bank of Philadelphia

Julapa Jagtiani
Federal Reserve Bank of Philadelphia

May 19, 2010

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Can Banks Circumvent Minimum Capital Requirements? The Case of Mortgage Portfolios under Basel II

Christopher Henderson
Retail Risk Officer, Federal Reserve Bank of Philadelphia
Ten Independence Mall, Philadelphia, PA 19106
Phone: 215-574-4139
E-mail: Chris.Henderson@phil.frb.org

Julapa Jagtiani
Special Advisor, Federal Reserve Bank of Philadelphia
Ten Independence Mall, Philadelphia, PA 19106
Phone: 215-574-7284
E-mail: Julapa.Jagtiani@phil.frb.org

Draft: May 19, 2010

Abstract

The recent mortgage crisis has resulted in several bank failures as the number of mortgage defaults increased. The current Basel I capital framework does not require banks to hold sufficient amounts of capital to support their mortgage lending activities. The new Basel II capital rules are intended to correct this problem. However, Basel II models could become too complex and too costly to implement, often resulting in a trade-off between complexity and model accuracy. In addition, the variation of the model, particularly how mortgage portfolios are segmented, could have a significant impact on the default and loss estimated and, thus, could affect the amount of capital that banks are required to hold. This paper finds that the calculated Basel II capital varies considerably across the default prediction model and segmentation schemes, thus providing banks with an incentive to choose an approach that results in the least required capital for them. We also find that a more granular segmentation model produces smaller required capital, regardless of the economic environment. In addition, while borrowers' credit risk factors are consistently superior, economic factors have also played a role in mortgage default during the financial crisis.

Keywords: Basel II, Bank Capital, Bank Regulation, Mortgage Default, Retail Credit Risk
JEL Classification Codes: G28, G21, G38, L51

The authors thank Joanne Chow, Vidya Nayak and Ali Cannoni for their research assistance. Special thanks to Joanne Chow for her dedication on this project. Please direct correspondence to Julapa Jagtiani, Federal Reserve Bank of Philadelphia, Supervision, Regulation and Credit Dept., Ten Independence Mall, Philadelphia, PA 19106, e-mail: Julapa.Jagtiani@phil.frb.org. The opinions in this paper are the authors' opinions and not necessarily those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

Can Banks Circumvent Minimum Capital Requirements? The Case of Mortgage Portfolios under Basel II

Christopher Henderson and Julapa Jagtiani

1. Introduction

The recent U.S. mortgage crisis has caused turbulence in financial and payment systems around the globe. The dramatic growth in subprime mortgages, in conjunction with the decline in lending standards in the early and mid 2000s, put the entire U.S. banking industry at great risk. A number of financial institutions have been closed, and those considered systemically important were rescued by the federal government during the financial crisis of 2007 and the ensuing recession. It has been evident that banks were not holding sufficient amounts of capital as a cushion to support their risky mortgage activities, causing them to become insolvent as the number of mortgage defaults increased as house prices declined and job losses mounted.[1]

Banking regulators recognized the potential threat of widespread capital inadequacy in the midst of the crisis as well as its potentially disastrous effects on the U.S. financial system and launched a program (the Supervisory Capital Assessment Program, or SCAP) for the largest U.S. bank holding companies (BHCs). SCAP was designed to challenge firms' loss assumptions and to build a capital buffer sufficient to withstand losses in the event the recession became more severe. Ten of the 19 BHCs involved in SCAP were required to raise the additional capital buffer within six months of the SCAP assessment; and one year later, none of these 19 BHCs has had an additional need for capital support,

[1] Lang and Jagtiani (2010) argue that such high concentration in mortgage-related securities at large financial institutions prior to August 2007 violated the basic principles of modern risk management and that forecasted tail risk was ignored at these institutions.

suggesting that the capital assessment resulting from SCAP was successful.[2] This one regulatory intervention was a timely response, among many others, to stave off a financial and perhaps a global meltdown. In the absence of a SCAP exercise, what regulatory infrastructure is currently in place that will force capital to adjust to heightened credit risk and prevent government bailouts? Regulators and financial analysts view the updated risk-based capital rules (Basel II) as a starting place to provide such a flexible, risk-sensitive capital framework.

The current Basel I capital requirement does not take into account the risk of mortgages, but rather it considers all mortgage activities to be in the same risk category regardless of the creditworthiness of the borrowers, the counterparties, or the level of systemic risk in the economy. The new Basel II capital requirement is intended to require large and complex banks (i.e., core banks) to better measure their risk so that they can manage it better and set aside an appropriate amount of capital to support the risk that they take on. [3]

Under Basel II, mortgage activities would be treated differently depending on the risk associated with the loans, which is a big improvement over the Basel I framework. Core banks, including JPMC, Citigroup, and Bank of America, would be required to use their own internal models to estimate the risk of default and the potential losses expected to occur due to their lending and other activities in a stress environment, in order to estimate the amount of required capital. Specifically, these banks would be required to establish appropriate models to estimate the various risk parameters and to transform these risk

[2] For more details on the SCAP program, see Hirtle, Schuermann, and Stiroh (2009).
[3] In the U.S., all large banks with assets of more than $250 billion or with foreign exchange activities of more than $10 billion (so-called core banks) would be required to implement the most sophisticated approach to risk measurement, the advanced internal ratings-based (A-IRB) approach, for the purposes of calculating capital requirements.

parameters into a risk-based capital amount that they need to hold. For a given "soundness standard" (99.9th percentile), the bank's calculated level of capital should be sufficient to cover unexpected losses up to the soundness standard to keep the bank solvent in the face of actual losses of that magnitude.[4]

Core banks have been working on developing their internal models for compliance with Basel II for many years, but none have begun operating under the final rule as of the first quarter of 2010. It is important to note that the Basel II guidelines are not intended to be prescriptive but rather to allow flexibility for banks to develop models that are appropriate to the organization's risk profile and unique business model. However, some variations of the modeling approaches could have a significant impact on the calculated required capital, holding fixed the risk that banks take, and, thus, could result in a level of required capital that is not correct and may be insufficient to cover unexpected losses.

Given the flexibility allowed under the Basel II guidelines, banks have responded to the new rules by employing different and sometimes competing practices to develop a segmentation scheme that reflects current risk management concerns to generate homogeneous pools of risk. Basel II rules only require that banks group their retail exposures into segments with homogeneous risk characteristics. These segments are necessary to estimate key risk parameters, such as probability of default (PD), loss given default (LGD), and exposure at default (EAD), and to determine risk-weighted assets (RWA) and required capital estimates. Do alternative segmentation methods result in important differences in required minimum capital for the same level of overall portfolio risk?

[4] This requirement is also applied to other retail portfolios, such as credit cards, auto, consumer loans, student loans, and small business loans under ($100,000).

We focus on the required capital for mortgage activities. As for other retail products, the Basel II framework allows a high degree of flexibility for mortgage segmentation in terms of criteria used and the number of segments and sub-segments. A well constructed segmentation system should have a sufficient number of loans in each segment in order for the risk parameters to be estimated with statistical significance and to be stable over time. There are no standard methods for grouping retail credits into risk buckets (segments) for Basel II purposes. Focusing on mortgage modeling, where risks are measured at the segment level, we examine whether banks would have the flexibility to choose a segmentation approach that allows them to reduce the capital requirements under the Basel II framework.[5]

In addition, previous studies have argued that risk measurement under credit risk models tends to underestimate the true risk during extremely stressful economic conditions. The Basel II capital calculation for mortgage lending could substantially understate prudential capital adequacy due to the skewed nature of returns from mortgage lending, where collateral values decline dramatically during an extreme systemic crisis and tail risk is substantially underestimated; see Dimou, Lawrence, and Milne (2009) and Lehnert (2009). We, therefore, explore the capital impact with different loss scenarios both before and during the 2007 financial crisis. Our results demonstrate that the choice of PD segmentation could have a significant impact across all loss scenarios and economic environments – both during the boom period prior to the financial crisis (2000-2006) and during the global financial crisis that started in 2007 (2007-2009). Our findings serve to

---

[5] Unlike wholesale products, retail products are much smaller in account size; thus, it is more efficient for the capital requirement calculation to be performed on the pool of retail loan accounts (segments), rather than at the account level.

fill the literature gap on the subject and to provide policy implications with respect to the flexibility allowed under the Basel II capital calculation.

2. Literature Review

Given the complexity of the A-IRB approach of Basel II for retail products, especially mortgages, and the potential importance of portfolio segmentation in measuring a bank's risk exposure in its mortgage lending activities, it is surprising to see a sparse literature on this subject. Berlin and Mester (2004) and Allen, DeLong, and Saunders (2004) provide an overview of issues in credit risk modeling for retail portfolios. Dimou, Lawrence, and Milne (2005) focus on the tail risk of mortgage lending by examining the appropriateness of the Basel II framework for computing capital adequacy for mortgage products. Owing to the extreme skewness of returns from mortgage products, where loan losses are usually small but would rise sharply in the event of a severe financial crisis, Basel II capital calculation tends to understate tail risk in an extreme systemic crisis and, thus, could substantially underestimate the appropriate amount of prudential capital adequacy.

Ambrose, Lacour-Little, and Sanders (2005) examine the potential motivation for mortgage loan securitization, where banks may choose to exploit the asymmetric information advantage to securitize risky loans and sell them to the public market or where they could retain risky loans for themselves while securitizing and selling good loans to the public. They find that based on data up to 2000, securitized mortgage loans are less risky (lower expected PD and lower probability of prepayment) than those mortgage loans banks retain in their portfolios, thus, supporting the capital arbitrage motivation argument. Again, the results are consistent with an argument that the Basel capital framework would tend to underestimate the amount of capital required to support those unsecuritized (more risky) mortgages that remain on the books. Krainer and Laderman (2009) also find consistent results based on more recent data up to October 2008; that is, securitized mortgage loans default at a faster rate than retained mortgage loans. In addition, Calem and Henderson (2010) find that securitized subprime mortgages during 2005-2006 were more risky than those retained on the book.

Cowan and Cowan (2004) find another potential reason for the Basel capital framework to provide inadequate capitalization for residential subprime mortgages, since defaults on these mortgages tend to be highly correlated (more defaults as the internal credit ratings decline). Unlike other residential loan portfolios where default correlation is generally low, the subprime (low credit quality) portfolios are subject to high default correlation. The coincidental movements in defaults among subprime borrowers are likely to be triggered by common underlying factors, such as an increase in unemployment – thus, potential capital inadequacy to support subprime lending under the Basel II capital rules.

A few studies have examined issues related to retail portfolio segmentation and the capital implications. In general, for products with a very large customer base such as mortgages, banks have the option to go with a very fine or a rough segmentation approach. Previous studies have claimed that finer levels (more granularity) of segmentation are associated with lower required capital estimates under the Basel II framework. Lang and Santomero (2004) point out conceptually that banks have incentive to estimate their PDs at a more granular level, since the Basel II formula produces a lower required capital number. This is because capital factors are concave in PD for a given LGD under the Basel II one-factor model, and because PDs are generally estimated at a more granular level than the LGD estimates.

Laurent (2004) elaborates further that since the capital is concave for most of the relevant range of PDs, using less granular segmentation (i.e., treating loans with different PDs as a single group) would result in overestimating the upper tail of the default frequency distribution. And since aggregating across different borrower types does not bias estimates of expected loss, more accurately distinguishing borrower types (more granular segmentation by PDs) would lower the estimated capital for the portfolio. Their analysis of the various segmentation criteria, using auto lease financing from European financial institutions in 2002, indicates that segmentation on the basis of scoring reduces the capital requirement by 30 basis points, a finding consistent with the argument that banks could reduce their capital requirements through their choice of segmentation. We provide a more in-depth empirical support in this paper, focusing on mortgage portfolios.

Ash, Kelly, Lang, Nayda, and Yin (2007) estimate the one-year PD for various alternative segmentation approaches for credit card loans. They find significant variation in the capital number based on the various set of risk factors that separate the sample into a low-risk segment and a high-risk segment. Their most granular segmentation criteria, which include refreshed score and delinquency (compared with their other criteria that include only one factor or a combination with origination score), produce the lowest capital numbers. Their results imply that there are capital relief incentives for improved risk segmentation, based on proprietary data from Capital One for credit card accounts originated in 1999 and 2000, with a performance period up to September 2004. We use more recent data (2000-2009) and provide a more in-depth analysis in this paper by exploring segmentation alternatives that contain up to 31 terminal nodes and that include additional important risk factors for segmentation criteria.

Kaltofen, Paul, and Stein (2006) use auto loan data from 2000 to 2002 to demonstrate that the overall required capital under the Basel II framework would be lower if PDs were estimated at a finer level of segmentation. Using a CHAID (chi-square automatic interaction detection algorithm) to develop the segmentation approaches, they conclude that the bank's ability to differentiate between sound and risky loans with greater precision through the formation of homogeneous segments would reduce the calculated amount of required capital.[6]

3. The Data

Our primary source of data is the loan level monthly mortgage data from the McDash LPS database, which covers the majority of all mortgage loans. We take a 5 percent random sample of observations and exclude those loans that were originated before 2000 or have missing FICO scores at origination. Our sample consists of approximately 2.46 million mortgage loans, originated during the period 2000-2008, a total of approximately 75.4 million monthly observations.

The LPS monthly data are then divided into nine cohorts, with the cohort observation dates being December 31 of each year (2000-2008) and the 12 months following each cohort date are the performance period for the cohort. For example, the performance period for the December 31, 2000 cohort is the period January 1 to December 31, 2001. The first six cohorts are defined as pre-crisis cohorts, consisting of cohort

[6] Note that Basel II requires that banks treat the loans in each segment in the same manner as in their internal risk management process.

observation dates December 31, 2000… and December 31, 2005 – with the cohort observation period from January 1, 2001 to December 31, 2006. The last three cohorts are defined as crisis cohorts, consisting of cohort observation dates December 31, 2006… and December 31, 2008 – with the cohort observation period from January 1, 2007 to December 31, 2009. The performance period data are used to define default within 12 months after the observation date. Once the default has been defined, the analysis includes only the year-end (cohort-level) data; this brings the observation number down to 2.43 million observations for the pre-crisis period and 3.69 million observations for the crisis period.

The McDash LPS loan level data are then merged with the quarterly credit bureau data from the Equifax database for the same time period, using year-end data for the period 2000-2009. The primary purpose of merging these two data sets is to obtain additional information about the loans and the borrowers, which is not available from the LPS database, specifically, information about second liens (or "piggybacks") and credit card utilization. The Equifax data contain customer information, including all information about first mortgages, second mortgages, and all credit card loans. Following the merging approach used in Elul, Souleles, Chomsisengphet, Glennon, and Hunt (2010), we merge the LPS and Equifax data based on the following characteristics of first mortgage loans: open date, initial balance, and ZIP code.

Unlike in Elul et al. (2010), we exclude all first mortgage loans that are associated with customers who have more than one first mortgage. We do this to make sure that we can match all the second mortgage loans of the same customer with the correct first mortgage loan and ensure that they belong to the same property (for the purpose of calculating the combined loan-to-value ratio). The merging of our cohort level (year-end) LPS data with the Equifax data results in the final loan observations of 211,061 for the pre-crisis period and 329,854 for the crisis period. Our economic data include (state-level) home-price index data from the Federal Housing Finance Agency (FHFA), formerly the Office of Federal Housing Enterprise Oversight (OFHEO), and the number of initial unemployment claim applications collected from the Haver Analytics database. The HPI is a weighted, repeat-sales index based on mortgage transactions on single-family properties (purchased or securitized by Fannie Mae or Fannie Mac) and within the conforming amount limits.

Table 1 presents a summary description of the combined data set for the pre-crisis and crisis periods. The statistics presented for the pre-crisis period are calculated based on loans originated during the period 2000-2005 only. However, the data for the crisis period include loans that were originated prior to the crisis as well, that is, loans originated during the period 2000-2008, with the default events occurring during the period 2007-2009. We divide the loans into three different product categories: prime (origination FICO at least 710), alt-A (origination FICO at least 620 or origination FICO at least 710 but with limited documentation), and subprime (origination FICO less than 620).[7] All statistics related to loan characteristics *at origination* include only loans that were originated during the period 2000-2005 (pre-crisis) and 2007-2009 (crisis). It is obvious from the summary statistics in Table 1 and the plot of vintage curves in Figure 1 that both lending standards and loan quality had deteriorated over the sample period, resulting in an increasing default rate

[7] Loans with an origination FICO score of at least 710, but with missing information on whether it is low-doc or no-doc, are considered prime mortgages; we assume that these loans are not low-doc or no-doc unless such status is clearly indicated in the McDash LPS database. Note that the average FICO score among prime borrowers was 710 in 2004 and declined to 706 in 2007 (see Amronin and Paulson, 2009). We keep the prime cut-off at 710 throughout the sample period in this study.

from the pre-crisis to the crisis periods. The pre-recession vintages performed much better than the post-crisis vintages, where cumulative gross losses for the 2006 vintage (including prime, alt-A, and subprime) had already reached 33 percent on average by 2009. Previous studies claim that the skewed nature of the banks' returns in mortgage lending (stable or good returns occur in a good economy and sharply decline during the crisis) has caused the Basel II framework to underestimate the PDs at the tail (during a systemic financial crisis). In this paper, we intend to shed more light on this issue by separating our analysis into the pre-crisis boom period (2000-2006) and the financial crisis period (2007-2009).

4. Mortgage Default Model

Logistic regressions are used to examine important factors that determine mortgage defaults, as defined in equation (1) below.

Log(P/1-P) = F (Idiosyncratic Risk factors, Economic Factors) ----- (1)

where P is the probability that the loan would default (default is defined as at least 60 days past due) within the next 12 months following the observation month.[8] In order to estimate this probability of default (PD) model, we divide the monthly data for each loan into cohorts, with December 31 of each year being the observation date and the 12 months following being the outcome period. The variable P takes a value of 1 if the loan defaulted during the following 12-month outcome period, and zero otherwise.

The idiosyncratic variables include the origination FICO score, origination debt-to-income (DTI) ratio, effective combined loan-to-value ratio (aggregated across first lien and second liens and marked-to-market), aggregated credit card utilization (across all cards

[8] Note that another definition of default (e.g., at least 90 days past due) was also considered; the results are similar but are not reported here.

and all joint borrowers), spot delinquency (30+ days past due as of observation date), product type (prime, alt-A, subprime), vintages (months on book), and the geographic indicators. The economic variables include unemployment (state level) and housing price appreciation index (state level).[9]

The results of the logistic regression analysis, based on equation (1), are reported in Tables 2A (Pre-Crisis) and 2B (Crisis), where Model (1) is the least granular model and Model (5) is the most granular. Model (1) includes only product type (prime, alt-A, or subprime) and geographic region. In model (2), loan age as measured by months on book is included as well. Model (3) incorporates economic factors (home price index and unemployment) at the state level into the analysis; thus, there is no need to keep the geographic dummy variables that were included in Models (1) and (2). Model (4) includes five important credit risk factors: FICO score at origination (*FICO*), debt-to-income ratio at origination (*DTI*), combined card utilization ratio (*Utilization*), effective combined loan-to-value ratio (*EC_LTV*), and spot delinquency, which is defined as 30+ days past due as of the cohort observation date (*Delinquency*).[10] The variable *Utilization* is the customer's combined utilization ratio for all credit cards, and it is calculated as the ratio of the combined balances (from all credit cards and all joint borrowers) to the combined credit lines (from all credit cards and joint borrowers). The *EC_LTV* is the ratio of all mortgage loan balances (first and second liens) to the appraised value of the property as of

[9] Other economic factors were included in our preliminary analysis but are less significant or insignificant, for example GDP (state level), bankruptcy filing (state level), unemployment claim (state level), Treasury spread (10-year Treasury yield vs. 3-month Treasury yield) , and credit risk spread (AAA bond yield vs. BBB bond yield).

[10] We include origination FICO scores from the McDash LPS database, rather than the updated credit scores from Equifax, because the updated credit scores are highly correlated with other credit risk factors that have already been included in the model (i.e., the spot delinquency and the current card utilization rate).

origination, adjusted for the changes in home price since the loan origination date to the cohort observation date.

The results overall demonstrate the relationship between mortgage defaults and the various idiosyncratic risks associated with the loans, borrowers, product types, and economic factors. The borrower's risk characteristics in Model (4) seem to be the most powerful of all idiosyncratic risk factors, as they fit the PD model the best (in terms of predictive ability of default) during the pre-crisis period, with 95.5 percent concordant. Adding the economic factors as in Model (5) does not improve the model's predictive ability. One of the reasons may be that our credit risk factors are measured so well and we have already incorporated economic conditions (particularly the home price index) into the credit risk measure through our EC_LTV variable.

Our earlier analysis (in the previous version of this paper) did not include the card utilization ratio and the previous loan-to-value variable (LTV) did not incorporate the second lien information into the LTV measure, and we found that economic factors added predictive value then. Unlike in pre-crisis period, for the crisis period in Table 2B, Model (5) performs slightly better than Model (4) – with 94.1 percent concordant compared with 93.9 percent. The economic factors do provide some additional information not captured by the credit risk factors during the financial crisis. In addition, from Tables 2A and 2B, the variable *HPI* is not significant in Model (5) during the pre-crisis period but becomes highly significant (at the 1 percent level) during the financial crisis period. This is consistent with the findings in Mayer, Pence, and Sherlund (2009) that the house-price decline during the mortgage crisis eroded home equity, resulting in higher defaults among prime borrowers.

The factors included in each of these models are then used for PD segmentation: five different segmentation schemes from Model (1) being least granular to Model (5) being the most granular. The first two models may be considered judgment-based segmentation models, since they leverage business expertise in devising the segmentation schemes, supported by judgment used in normal business practices. The last three models are considered more granular and are primarily statistics-based, using decision-tree methods to determine key risk drivers that differentiate risk within a retail portfolio.

The objective of the PD segmentation is to build the right number of segments with homogeneous loans, where the number of loans within each segment is neither too small nor too large. There tend to be issues with statistical significance when the segments contain few homogeneous observations; however, segments with large concentrations might also be an issue, since they call into question the ability of the segmentation scheme to identify homogeneous pools of risk. In addition, the risk ranking order is another important criterion in the PD segmentation, so that the segments can be arranged from high risk to low risk in such a way that adjacent segments do not share the same risk characteristics and do not have the same average PD.[11] We explore the various alternative segmentation schemes (with varying degrees of granularity) that meet these segmentation objectives and demonstrate that different qualified segmentation approaches could result in significantly different required capital and that banks may have incentives to choose a segmentation scheme that helps reduce their regulatory capital burden – capital avoidance.

5. <u>Portfolio Segmentation – Using a CHAID Algorithm</u>

[11] This cardinal ranking of segments is similar to credit bureau score bands that map risk levels as measured by average default or delinquency rates to non-overlapping score ranges.

Basel II requires banks to classify each of its retail exposures into one of three retail subcategories: residential mortgage exposures, qualifying retail exposures (QRE), and other retail exposures. By relying on an assumption that a bank's portfolio is infinitely granular (large in size and geographically diverse), the bank must further classify exposures into segments with similar risk characteristics. Once the portfolio has been segmented to a point at which each adjacent segment is statistically different, the segments should be rank ordered by some well defined measure of risk (such as average PD) and should not overlap. In this paper, we leverage decision-tree methods, using a chi-square automatic interaction detector (CHAID), which is one of the oldest tree classification methods originally proposed by Kass (1980). It allows for a wide choice of tree algorithms based on the chi-square test.

CHAID is also commonly used in the industry for the purpose of classifying data used in credit risk models. It involves formulating a set of rules that generate a split from a parent node to a child node based on the maximum similarity statistic within and between the nodes, to determine how records from the parent node are to be distributed across the child nodes. The CHAID tree diagram allows for multiple ways to split the observations into many categories (segments) based on categorical predictors with many classes. It is a useful method of summarizing data and can show major natural divisions of the observations (e.g., mortgage loans) by various defining variables.

In constructing the tree, for which the dependent variable is categorical in nature (e.g., default or not default), CHAID relies on the chi-square test to determine the next best split at each step. The first step in CHAID is to prepare the predictors by creating categorical predictors out of any continuous predictors by dividing the respective continuous distributions into a number of categories with an approximately equal number of observations. By grouping values of the independent variable into pre-determined intervals, a tree branch is formed by iteratively grouping each interval into finer intervals called categories (or segments). After each split is found, the algorithm attempts to improve the significance by taking single dependent variable values from segments to which they were assigned and grouping them with other categories. If improvement is achieved, the split that produced it becomes the candidate. This process is repeated until no further splits are possible.[12]

The segmentation results, based on different alternative groups of predictors, are presented in Tables 3A (Pre-Crisis) and 3B (Crisis), where total number of nodes, number of terminal nodes, and the depth of the tree are reported. The ranking order of the predicting factors and the estimated average PD for each model are also presented. The ranking of importance of risk factors are the same for both crisis and pre-crisis (boom) periods. The most important mortgage default predictors are obviously the indicator that loans have already been past due (at least 30 days past due) as of the observation date. Other important factors are the borrower's FICO score at origination, the effective combined loan-to-value (aggregated across all first and second liens and adjusted with the home price index from origination to observation date), and interestingly, the aggregate credit card utilization across all cards and all joint borrowers. Economic factors are not important (after including idiosyncratic credit risk factors) during the boom period but become highly significant during the financial crisis period.

[12] For more detail on the CHAID algorithm used in this paper, see KnowledgeSEEKER (2009).

6. <u>Does the Required Capital Vary Across Segmentation Approaches?</u>

Once the segmentation process is completed, the various Basel II risk parameters (PD, LGD, EAD) are calculated for each segment, and the required capital is calculated for the entire mortgage portfolio. Our analysis demonstrates how the Basel II parameters, the calculated risk-weighted assets and the required Basel II regulatory capital vary across the segmentation approaches, across the method for calculating averages, and between the pre-crisis and the crisis periods.

Probability of Default (PD): The average PD for each node (segment) is calculated as the ratio of defaulted accounts to total number of accounts in the node, where default is defined as being at least 60 days past due during the next 12 months following the cohort date (observation date). Based on the PD of each node, we calculated average PDs for the entire mortgage portfolio using three different approaches: 1) simple average across nodes, 2) balance weighted average PD, where the ratio of loan balance for each node to total loan balance for the portfolio is used as the weight, and 3) account-weighted average PD, where the ratio of number of accounts for each node to total number of accounts in the portfolio is used as the weight. The results are presented at the bottom of Tables 3A and 3B. Average PDs vary greatly across the calculation method (simple average vs. balance weight vs. account weight) and vary across the segmentation approaches from Model (1) to Model (5). The account-weight PDs do not vary across models because default is calculated based on number of accounts. The simple average method tends to over-estimate the PD during good times, since there are a smaller number (or amount) of loans in bad segments but could under-estimate the PDs during bad times. The balance-weight method tends to produce the smallest estimated PDs regardless of model granularity or economic environment. Our results suggest that a bank's choice of PD estimation weighting method could significantly affect the amount of required capital.

Loss Given Default (LGD) and Expected LGD (ELGD): The parameter LGD is allowed to vary in our analysis due to the hypothetical nature of the portfolios. Under the Basel II framework, LGD is the bank's empirically based best estimate of the economic loss, per dollar of exposure at default (EAD) that the bank would expect to incur if the exposures in the segment were to default within a one-year horizon during economic downturn conditions.[13] Up until late 2006, many core banks did not have economic downturn conditions (stress data) in their historical reference data systems. As a result, they could measure LGD using a linear supervisory mapping function proposed by federal regulators to convert ELGD into LGD for risk-based capital purposes.[14] Although banks are not required to use the supervisory mapping function, the rule permits its use in the event that data issues create serious problems in LGD estimation.

In this study, we use the same segmentation schemes for both PD and LGD. The LGD is calculated based on foreclosed accounts in the PD segments, and it is equal to the ratio of loss amount to account balance prior to foreclosure. We provide several ranges for LGD based on industry practices to establish various sensitivities for LGD in order to test the robustness of the capital estimates for different loss severities that could be observed in actual bank portfolios. That is, we estimate LGD from three different assumptions about the recovery cost (40, 60, and 80 percent). The loss amount is equal to account balance minus the current value of the property (appraised value at origination adjusted by HPI)

---

[13] See page 69,402 of the Federal Register, Vol. 72, Friday, December 7, 2007. The rule also permits long-run dollar-weighted average economic loss estimates. LGD must always be positive. Expected LGD or ELGD differs from LGD only by the fact that it is measured over a mix of economic conditions, including economic downturn conditions. In other words, LGD is a stress concept and ELGD is a" through the cycle" concept.

[14] The linear supervisory mapping function is LGD=0.08+0.92xELGD.

plus the recovery cost. Following the Basel II final rules for mortgage LGD, the calculated LGD is floored at 10 percent and capped at 100 percent (at the loan level).

Basel II Capital (K): K is calculated using the Basel II regulatory capital function, where the asset correlation (R) is equal to 0.15 as defined by the Basel II final rules for mortgages. K is calculated for each segment based on the segment's PD, LGD, and R, with three different recovery cost assumptions. The calculated Ks (for the entire mortgage portfolio) are reported in Table 4 – top panel for pre-crisis and bottom panel for crisis -- for the five different segmentation approaches (Model (1) to Model (5)) and across three different approaches for averaging across segments (simple average vs. balance weight vs. account weight). Again, the calculated K varies across the segmentation approaches and the methods of calculating average K. Using a simple average approach, the K value increases with the model granularity; thus, the least granular model produces less required capital. In contrast, a balance-weight or account-weight approach produces smaller K values as the model's granularity increases from Model (1) to Model (5), regardless of the economic environment. Our results suggest that, under the Basel II risk-based capital framework, banks may have incentives to choose a calculation method that still meets the final rules yet produces the least amount of capital requirements for them.[15]

Risk-Weighted Assets (RWA) and Required Capital: From the calculated K for each node, we also calculated the RWA for each segment (node), and it is equal to K * 12.5 * EAD, where the exposure at default (EAD) is the dollar amount of loan balance prior to foreclosure. The portfolio's RWA is equal to the combined RWA across all segments, and the required tier I capital is equal to 8 percent of RWA, reported in Table 5 for the five

[15] This is consistent with Hasan, Siddique, and Sun (2009), who find that the market expectation of capital can be quite different from the regulatory required capital.

segmentation models (top panel for pre-crisis and bottom panel for crisis). The results reported in Table 5 demonstrate that the required capital declines as the granularity of the segmentation scheme increases (consistent with the average K calculation using balance-weight method since it is EAD weighted here) – moving from Model (1) to Model (50) would result in a capital reduction of 40 to 48 percent depending on the LGD assumptions. Note that the RWA here is calculated at the segment level and added across all segments. The RWA number could be significantly different if banks choose to calculate their portfolio's RWA based on an average K for the portfolio (rather than segment) as the bank's choice of calculation for the portfolio's average K would directly impact the RWA and the required capital. Overall, our results indicate potential incentives for banks to choose a segmentation approach and RWA calculation method that help reduce the required capital for them.

7. <u>Are More Granular Segmentation Schemes More Stable?</u>

K-S and ROC Charts: In evaluating and comparing the quality of the various segmentation schemes, Figures 2A and 2B present the Kolmogorov-Smirnov (K-S) chart and the Receiver-Operator Characteristics (ROC) chart, respectively, for Models (1) to (5) for the pre-crisis period. The K-S plot in Figure 2A shows that the K-S statistics range from 0.5701 in Model (1) to 0.8184 in Model (5), suggesting a better segmentation outcome for more granular models (in terms of homogeneity within segments and risk ranking across segments). The ROC curve is a graphical representation of the trade-off between the false negative and the false positive. The larger the area under the curve, the better it is for the segmentation scheme in producing segments with homogeneous loans. Figure 2B plots true positive against false positive and shows that Models (4) and (5), which include idiosyncratic credit risk factors, are superior to the less granular segmentation schemes in Model (1) to Model (3). Figures 3A and 3B present similar plots for Model (1) to Model (5) during the crisis period. Again, the plots demonstrate improved segmentation schemes with greater granularity.

Population Stability Index (PSI): This is a measure of stability of the segmentation over time, comparing the proportion of the population (mortgage loans) in each year that falls into each segment. We calculate PSI as indicated in equation (2) below:

$$PSI = \sum_i (F_{i,t} - F_{i,t+1}) * Ln(F_{i,t}/F_{i,t+1}) \quad \text{-------------------------} \ (2)$$

where $F_{i,t}$ is the proportion of loans in Cohort (t) that falls in segment (node) i and $F_{i,t+1}$ is the proportion of loans in Cohort (t+1) that falls in segment i. A PSI index of 0.10 or less indicates no real population shift or a stable segmentation, and an index greater than 0.25 indicates a definite population shift and that the segmentation is not stable. An index between 0.10 and 0.25 suggests some shift in population. Table 6 reports the PSI for both the pre-crisis and crisis periods. The results indicate that Models (3) and (2), in which PD segmentation schemes are based on economic factors, geographic indicators, and loan age only, are the least stable ones, during both the crisis and pre-crisis periods. The most granular segmentation schemes in Model (4) and Model (5) are not only more stable (with smaller PSI) but also produce a smaller estimated required capital for the bank (based on RWA calculated at the segment level, as shown in Table 5).

8. Conclusions and Policy Implications

The recent mortgage crisis has resulted in several bank failures as the number of mortgage defaults increased. The current Basel I capital framework does not require banks to hold sufficient amounts of capital to support their mortgage lending activities, since all mortgages are treated as if they were equal in terms of riskiness, regardless of the borrower's credit risk or whether they are no-doc mortgages. The new Basel II capital rules are intended to correct this problem. However, Basel II models could become too complex and too costly to implement, often resulting in a trade-off between complexity and model accuracy. More importantly, the variations in the models, particularly how mortgage portfolios are segmented, could potentially have a significant impact on the default and loss estimated, thus, significantly affecting the amount of capital that banks are required to hold.

This paper examines the implications from the various mortgage segmentation approaches. Our analysis, which compares five different models with varying degrees of granularity during the pre-crisis and crisis periods, demonstrates that there is a large difference across the segmentation approaches in terms of estimated default probability and the calculated required capital. More granular segmentation models (including idiosyncratic credit risk as well as a home-price index) produce smaller required capital (if calculated at the segment level and sum across segments in the portfolio) and they also appear to be more stable than the less granular models. Our results suggest that banks may be faced with incentives to choose an approach that results in the least required capital for them.

Our default prediction models also demonstrate the importance of the borrower's credit risk factors -- which include not only the borrower's FICO score but also the aggregate utilization ratio across all credit cards held by the borrower and joint borrowers – in determining mortgage default probability. A less granular mortgage default prediction and/or segmentation model – e.g., one that incorporates only economic factors, product types, and/or loan age into the analysis -- would tend to be less accurate and subject to population shift across segments over time. While economic factors may play a little more significant role during the crisis period than during the pre-crisis (boom) period, we find that the quality of both the mortgage default prediction and the PD segmentation models can consistently be enhanced considerably by including the following important credit risk factors: spot delinquency, FICO score, effective combined loan-to-value (aggregated across all first and second liens and adjusted with the home-price index), and aggregated utilization across all credit cards and joint borrowers. The origination debt-to-income ratio does not contribute much after including the other credit risk factors in the model.

Overall, we find that banks have the flexibility to choose an approach for their capital calculation that meets the risk-based Basel II capital framework and yet produces the least amount of capital for them, thus resulting in potential under-capitalization. For the purpose of Basel II qualification review, it is important that banks provide analytical support for important modeling assumptions to demonstrate that the segmentation system reflects risk drivers that are commonly found in the bank risk management process and that the segmentation system provides accurate, reliable, and consistent estimates of all the Basel II risk parameters (PD, LGD, and EAD). Furthermore, given the ability of the banks to leverage their own internally developed risk parameters to calculate credit risk capital requirement, a formal periodical review (post-qualification for Basel II) would be necessary to ensure that the segmentation system is reviewed and updated appropriately to reflect a rapidly changing environment and economic impacts. As a result, Pillar II (Supervisory Review) could play a critical role in detecting the bank's incentive to balance the segmentation-capital benefit tradeoff (i.e. to invest in more granular segmentation systems for the benefit of reducing regulatory capital requirement).[16]

Our results not only fill the gap in the literature, but they also serve as a benchmark and as useful supervisory tools for reviewing mortgage credit risk under the new Basel II risk-based capital requirement framework. It is important that banks choose an approach that is most appropriate for their organization and business model to derive the most accurate risk measures and required capital for their retail portfolios.

[16] The new risk-based Basel II capital framework is built on three main Pillars (principal): Pillar 1 (minimum required capital), Pillar 2 (supervisory review), and Pillar 3 (market discipline and disclosure).

<u>References:</u>

Allen, L., G. DeLong, and A. Saunders (2004) "Issues in the Credit Risk Modeling of Retail Markets," *Journal of Banking and Finance* 28, 727-752.

Ambrose, B., M. Lacour-Little, and A. Sanders (2005) "Does Regulatory Capital Arbitrage, Reputation, or Asymmetric Information Drive Securitization?" *Journal of Financial Services Research*, 28, Number 1-3, 113-133.

Amronin, G. and A. Paulson (2009) "Comparing Patterns of Default Among Prime and Subprime Mortgages," *Economic Perspectives*, Federal Reserve Bank of Chicago, Quarter 2, 18-37.

Ash, D., S. Kelly, W. Lang, W. Nayda, and H. Yin (2007) "Segmentation, Probability of Default and Basel II Capital Measures for Credit Card Portfolios," Unpublished Working Paper, Federal Reserve Bank of Philadelphia, August.

Berlin, M. and L. Mester (2004) "Retail Credit Risk Management and Measurement: An Introduction to the Special Issue," *Journal of Banking and Finance* 28, 721-725.

Calem, P. and C. Henderson (2010) "Cream-Skimming in Subprime Mortgage Securitizations: Which Subprime Mortgage Loans Were Sold by Depository Institutions Prior to the Crisis of 2007?" Federal Reserve Bank of Philadelphia, Working Paper No. 10-8, March.

Cowan, A. and C. Cowan (2004) "Default Correlation: An Empirical Investigation of a Subprime Lender," *Journal of Banking and Finance* 28, 753-771.

Dimou, P., C. Lawrence, and A. Milne (2005) "Skewness of Returns, Capital Adequacy, and Mortgage Lending," Journal of Financial Services Research, 28, Number 1-3, 135-161.

Elul, R., N. Souleles, S. Chomsisengphet, D. Glennon, and R. Hunt (2010) "What 'Triggers' Mortgage Default?" Federal Reserve Bank of Philadelphia, Working Paper No. 10-13, April.

Hasan, I., A. Siddique, and X. Sun (2009) "Capital Adequacy Revisited: What Banks Report Versus What Markets Say," Working Paper, Rensselaer Polytechnic Institute, December.

Hirtle, B., T. Schuermann, and K. Stiroh (2009) "Macroprudential Supervision of Financial Institutions: Lessons from the SCAP," Federal Reserve Bank of New York Staff Report No. 409, November.

Kaltofen, D., S. Paul, and S. Stein (2006) "Retail Loans & Basel II: Using Portfolio Segmentation to Reduce Capital Requirements," European Credit Research Institute (ECRI) Research Report No. 8, August.

Krainer, J. and L. Laderman (2009) "Mortgage Loan Securitization and Relative Loan Performance," Federal Reserve Bank of San Francisco, Working Paper, September.

Kaltofen, D., S. Paul, and S. Stein (2006) "Retail Loans & Basel II: Using Portfolio Segmentation to Reduce Capital Requirements," European Credit Research Institute (ECRI) Research Report No. 8, August.

Kass, G.V. (1980) "An Exploratory Technique for Investigating Large Quantities of Categorical Data," *Journal of Applied Statistics* 29, 119-127.

Knowledge Seeker (2009) "Mathematical Formulations (Technical Notes)," Angoss Software Corporation.

Lang, W. and J. Jagtiani (2010) "The Mortgage and Financial Crises: The Role of Credit Risk Management and Corporate Governance," University of Pennsylvania, Wharton Financial Institution Center, Working Paper No. 10-12 (February), forthcoming in *Atlantic Economic Journal.*

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

Laurent, M. (2004) "Asset Return Correlation in Basel II: Implications for Credit Risk management," Working paper CEB 04-17, Solvay Business School, University Libre de Bruxelles.

Lehnert, A. (2009) "The Outlook for House Prices and Foreclosures," Working Paper, Federal Reserve Board, presented at the Federal Reserve Bank of Atlanta, July.

Mayer, C., K. Pence, and S. Sherlund (2009) "The Rise in Mortgage Defaults," *Journal of Economic Perspectives*, Volume 12, No. 1, Winter, 27-50.

Table 1
Summary Statistics of the Sample

Pre-crisis data include loans originated in 2000-2005, with performance period 1/1/2001-12/31/2006. Crisis data include loans originated in 2000-2008, with performance period 1/1/2007-12/31/2009. Factors measured at origination are calculated based on loans that were originated in 2000-2005 (Pre-Crisis) and 2007-2008 (Crisis).

Variables	Pre-Crisis (2000-2006)	Crisis (2007-2009)
Average FICO Score at Origination	719	708
Average Debt-to-Income at Origination (%)	36.85	36.7
Average Loan-to-Value (LTV) at Origination (%)	68.00	73.80
Average Combined LTV (%)	67.09	68.53
Average Effective Combined LTV (%)	62.75	65.39
Average Utilization Rate (%)	29.26	31.51
Number of All Loans in the Sample	211,061	329,854
Average Loan Balance at Origination ($)	179,044	204,800
Average Loan Age (Months-On-Book)	19.82	37.43
Number of Prime Loans in the Sample (%)	55.29	44.41
$ Amount of Prime Loans in the Sample (%)	55.70	48.76
Number of Alt-A Loans in the Sample (%)	37.62	44.16
$ Amount of Alt-A Loans in the Sample (%)	38.81	43.18
Number of Subprime Loans in the sample (%)	7.09	11.43
$ Amount of Subprime Loans in the sample (%)	5.49	8.06
Number of No-Doc Loans in the Sample (%)	4.74	9.62
$ Amount of No-Doc Loans in the Sample (%)	6.73	10.55
Number of Jumbo Loans in the Sample (%)	6.35	5.83
$ Amount of Jumbo Loans in the Sample (%)	19.00	16.41
Average Default (%)	0.82	3.00
Number of Foreclosure (%)	0.21	0.93
Average Home Price Index (HPI)	387.66	426.74
Number of Unemployment Claims	75,595	87,143
Number of Mortgage Loans from McDash LPS	2.43 Mill	3.69 Mill
Number of Observations from Equifax	251.6 Mill	119 Mill
Final Number of Mortgage Loan Observations (Include only customers with 1 first mortgage)	211,061	329,854

Table 2A
Logistic Analysis: Pre-Crisis Period (2000-2006)

The analysis is based on equation (1). Dependent variable is probability of default (60+ dpd) in the next 12 months. Model (1) is the least granular model, including only product type and geographic factors. Model (2) includes months on book (MOB) as well. Model (3) includes economic variables, product type, and MOB. Model (4) includes idiosyncratic risk factors. Model (5) is the most granular, including all but regional factors, since the regional impact is already incorporated into the analysis through the regional HPI and unemployment. P-values are reported in parentheses. The ***, **, and * denote the 1%, 5%, and 10% significance, respectively.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Intercept	-4.77***	-5.57***	-6.72***	-0.86	-0.58
	(0.0001)	(0.0001)	(0.0001)	(0.1298)	(0.6139)
D_Alt-A	2.19***	2.19***	2.20***		0.27
	(0.0001)	(0.0001)	(0.0001)		(0.1915)
D_Subprime	4.23***	4.23***	4.22***		0.06
	(0.0001)	(0.0001)	(0.0001)		(0.8578)
D_East	-2.31**	-2.33*			
	(0.0359)	(0.0634)			
D_Midwest	-1.81	-1.85			
	(0.1010)	(0.1407)			
D_Southwest	-1.96*	-2.01			
	(0.0750)	(0.1091)			
D_West	-2.80**	-2.77**			
	(0.0110)	(0.0272)			
MOB		0.04***	0.04***		0.0219***
		(0.0001)	(0.0001)		(0.0001)
HPI			-0.003***		-0.0004
			(0.0001)		(0.3093)
Unemployment			-1.75E-6***		-1.87E-6*
			(0.0015)		(0.0912)
FICO Score				-0.01***	-0.01***
				(0.0001)	(0.0001)
DTI Ratio				0.003	0.004*
				(0.1983)	(0.0641)
Utilization				0.01***	0.01***
				(0.0001)	(0.0001)
EC_LTV				0.02***	0.02***
				(0.0001)	(0.0001)
Delinquency				3.57***	3.48***
				(0.0001)	(0.0001)
% Concordant	82.0	85.6	85.7	95.5	95.2
% Discordant	10.0	10.0	9.9	2.3	2.3

Table 2B
Logistic Analysis: Crisis Period (2007-2009)

The analysis is based on equation (1). Dependent variable is probability of default (60+ dpd) in the next 12 months. Model (1) is the least granular model, including only product type and geographic factors. Model (2) includes months on book (MOB) as well. Model (3) includes economic variables, product type, and MOB. Model (4) includes idiosyncratic risk factors. Model (5) is the most granular, including all but regional factors, since the regional impact is already incorporated into the analysis through the regional HPI and unemployment. P-values are reported in parentheses. The ***, **, and * denote the 1%, 5%, and 10% significance, respectively.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Intercept	-5.77***	-5.71***	-4.57***	-1.97***	-1.88***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
D_Alt-A	1.84***	1.83***	1.82***		0.12**
	(0.0001)	(0.0001)	(0.0001)		(0.0476)
D_Subprime	3.44***	3.42***	3.43***		0.16
	(0.0001)	(0.0001)	(0.0001)		(0.1229)
D_East	0.35	0.39			
	(0.6249)	(0.5892)			
D_Midwest	0.54	0.58			
	(0.4568)	(0.4220)			
D_Southwest	0.77	0.81			
	(0.2862)	(0.2650)			
D_West	0.71	0.74			
	(0.3269)	(0.3057)			
MOB		-0.003***	-0.01***		0.001
		(0.0001)	(0.0001)		(0.2646)
HPI			-0.002***		-0.001***
			(0.0001)		(0.0007)
Unemployment			4.36E-6***		3.7E-6***
			(0.0001)		(0.0001)
FICO Score				-0.008***	-0.01***
				(0.0001)	(0.0001)
DTI Ratio				0.004***	0.004***
				(0.0001)	(0.0001)
Utilization				0.01***	0.01***
				(0.0001)	(0.0001)
EC_LTV				0.03***	0.02***
				(0.0001)	(0.0001)
Delinquency				2.99***	2.99***
				(0.0001)	(0.0001)
% Concordant	75.3	77.3	80.2	93.9	94.1
% Discordant	16.2	17.9	17.5	5.4	5.1

Table 3A
Segmentation Schemes – Characteristics and Estimated PD
Pre-Crisis Period (2000-2006)

The maximum depth of tree is limited to no more than 5 and the segmentation process is set to stop splitting when a node contains less than 500 loan observations.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Total Number of Nodes	12	37	48	44	49
Number of Terminal Nodes	8	22	30	27	31
Maximum Depth of Tree	3	5	5	5	5
Ranking:					
• Loan Category (Prime, Alt-A, Subprime)	1	1	1	--	3
• Region (East, Midwest, Southwest, West)	2	3	--	--	--
• Months On Book	--	2	2	--	6
• HPI (State Level)	--	--	3	--	7
• Unemployment Claims (State Level)	--	--	4	--	8
• FICO Score (at origination)	--	--	--	2	2
• Debt-to-Income Ratio (at origination)	--	--	--	5	9
• Combined Card Utilization	--	--	--	3	4
• Combined LTV (HPI adjusted)	--	--	--	4	5
• Spot Delinquency (At least 30 days past due)	--	--	--	1	1
Average PD for Mortgage Portfolio:					
• Simple Average	2.63%	2.72%	3.42%	4.93%	7.84%
• Balance Weight	0.68%	0.65%	0.63%	0.65%	0.63%
• Account Weight	0.83%	0.83%	0.83%	0.83%	0.83%

Table 3B
Segmentation Schemes – Characteristics and Estimated PD
Crisis Period (2007-2009)

The maximum depth of tree is limited to no more than 5 and the segmentation process is set to stop splitting when a node contains less than 500 loan observations.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Total Number of Nodes	12	47	108	82	94
Number of Terminal Nodes	8	29	68	52	59
Maximum Depth of Tree	3	5	5	5	5
Ranking:					
• Loan Category (Prime, Alt-A, Subprime)	1	1	1	--	3
• Region (East, Midwest, Southwest, West)	2	3	--	--	--
• Months On Book	--	2	2	--	6
• HPI (State Level)	--	--	3	--	7
• Unemployment Claims (State Level)	--	--	4	--	8
• FICO Score (at origination)	--	--	--	2	2
• Debt-to-Income Ratio (at origination)	--	--	--	5	9
• Combined Card Utilization	--	--	--	3	4
• Combined LTV (HPI adjusted)	--	--	--	4	5
• Spot Delinquency (At least 30 days past due)	--	--	--	1	1
Average PD for Mortgage Portfolio:					
• Simple Average	7.01%	6.93%	6.48%	8.51%	11.26%
• Balance Weight	2.80%	2.88%	2.96%	2.96%	2.96%
• Account Weight	3.00%	3.00%	3.00%	3.00%	3.00%

Table 4
Basel II Capital (K) – By Segmentation Schemes and LGD Assumption

Capital (K) is calculated based on the Basel II formula for mortgages, where the asset correlation is 0.15. The LGDs used are based on three different assumptions for recovery cost (40%, 60%, and 80%).

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Pre-Crisis Average PD:					
PD_Simple Average (%)	2.63	2.72	3.42	4.93	7.84
PD_Balance Weight (%)	0.68	0.65	0.63	0.65	0.63
PD_Account Weight (%)	0.83	0.83	0.83	0.83	0.83
Pre-Crisis K_Basel II:					
K_Simple Average_40 (%)	2.54	2.88	3.26	2.46	3.93
K_Simple Average_60 (%)	6.42	6.73	7.72	5.59	8.79
K_Simple Average_80 (%)	10.32	10.66	12.33	9.30	14.04
K_Balance Weight_40 (%)	0.87	0.87	0.89	0.65	0.53
K_Balance Weight_60 (%)	2.45	2.28	2.26	1.59	1.30
K_Balance Weight_80 (%)	4.05	3.74	3.70	2.58	2.11
K_Account Weight_40 (%)	1.01	1.03	1.10	0.77	0.64
K_Account Weight_60 (%)	2.77	2.66	2.70	1.86	1.57
K_Account Weight_80 (%)	4.54	4.35	4.35	2.99	2.54
Crisis Average PD:					
PD_Simple Average (%)	7.01	6.93	6.48	8.51	11.26
PD_Balance Weight (%)	2.80	2.88	2.96	2.96	2.96
PD_Account Weight (%)	3.00	3.00	3.00	3.00	3.00
Crisis K_Basel II:					
K_Simple Average_40 (%)	7.29	7.27	6.98	4.89	5.64
K_Simple Average_60 (%)	13.24	12.94	11.89	8.47	9.51
K_Simple Average_80 (%)	19.20	18.64	16.94	13.49	14.87
K_Balance Weight_40 (%)	4.77	4.67	4.41	2.76	2.77
K_Balance Weight_60 (%)	8.36	8.15	7.77	4.71	4.65
K_Balance Weight_80 (%)	11.94	11.68	11.29	6.78	6.78
K_Account Weight_40 (%)	4.78	4.60	4.35	2.77	2.83
K_Account Weight_60 (%)	8.46	8.12	7.71	4.74	4.72
K_Account Weight_80 (%)	12.14	11.69	11.23	6.86	6.97

Table 5
Basel II Capital Required for the Mortgage Portfolio
By Segmentation Schemes and LGD Assumption

The dollar amount of required capital reported below is calculated based on the Basel II capital (K) for each node, the LGD for each node, and the EAD for each node (equal to loan balance). The $ amount of capital calculated for each of the nodes is then added together to derive the $ capital required for the mortgage portfolio. The node LGD is calculated based on three different recovery cost assumptions (40%, 60%, and 80%) and LGD is floored at 10% and capped at 100% (following the Basel II final rules).

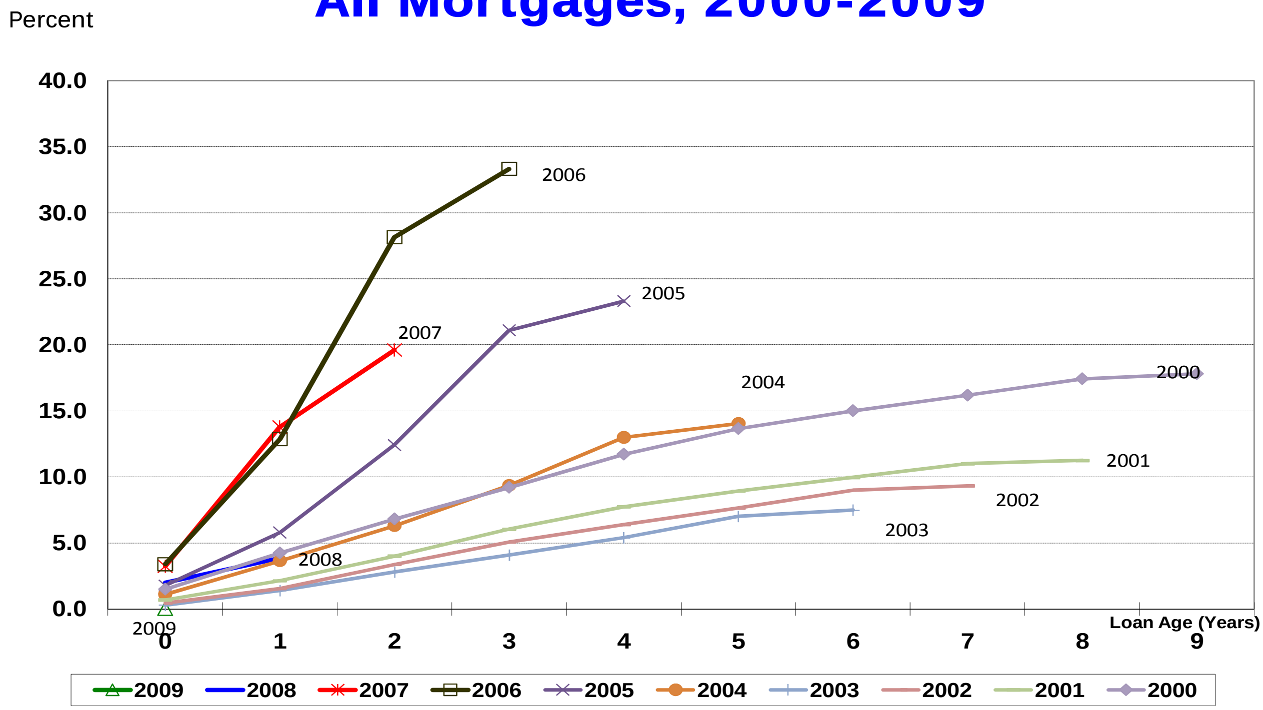
	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Pre-Crisis: Capital Required for the Mortgage Portfolio ($Mil):					
40 % Recovery Cost	218	217	221	163	131
60 % Recovery Cost	613	569	564	398	324
80 % Recovery Cost	1,010	935	924	644	527
Crisis: Capital Required for the Mortgage Portfolio ($Mil):					
40 % Recovery Cost	2,082	2,036	1,923	1,205	1,207
60 % Recovery Cost	3,645	3,555	3,389	2,055	2,027
80 % Recovery Cost	5,207	5,094	4,924	2,956	2,955

Table 6
Stability of the PD Segmentation – Population Stability Index (PSI)
By Segmentation Schemes

The PSI calculation is based on the proportion of loans in each cohort that falls into each node, using equation (2). PSI < 0.10 indicates no real change from one cohort to the next. The ** denotes PSI between 0.10 and 0.25 (suggesting some shift), and *** denotes PSI > 0.25 (indicating a definite change in population from one cohort to the next).

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Pre-Crisis: Population Stability Index					
Cohort 2000-2001	0.076	0.115**	0.145**	0.100	0.114**
Cohort 2001-2002	0.008	0.130**	0.135**	0.036	0.096
Cohort 2002-2003	0.017	0.078	0.149**	0.036	0.055
Cohort 2003-2004	0.008	0.470***	0.555***	0.020	0.227**
Cohort 2004-2005	0.011	0.349***	0.405***	0.031	0.042
Pre-Crisis: Population Stability Index					
Cohort 2006-2007	0.002	0.195**	0.222**	0.025	0.030
Cohort 2007-2008	0.000	0.068	0.524***	0.054	0.077

Figure 1: Vintage Curve
All Mortgages, 2000-2009



1

Figure 2A: Pre-Crisis (2000-2006): K-S Chart Comparison (Models 1-5)

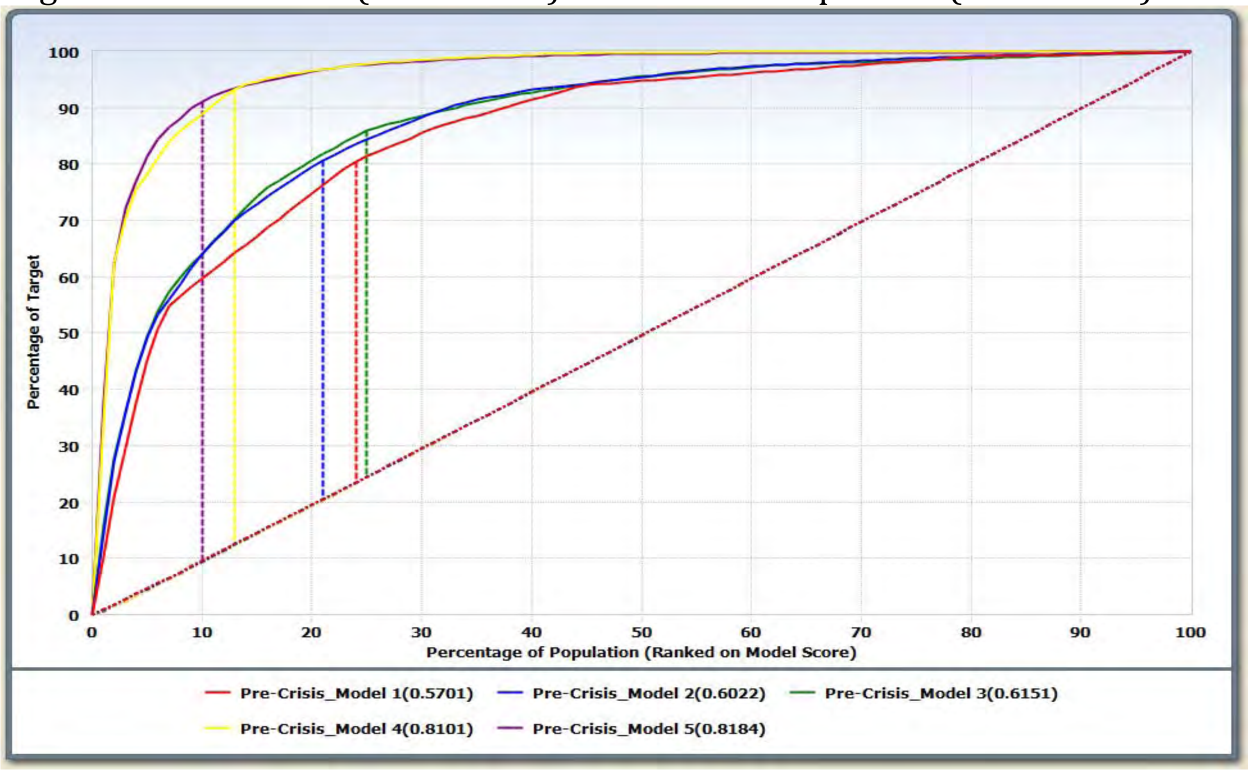


Figure 2B: Pre-Crisis (2000-2006): ROC Chart Comparison (Models 1-5)

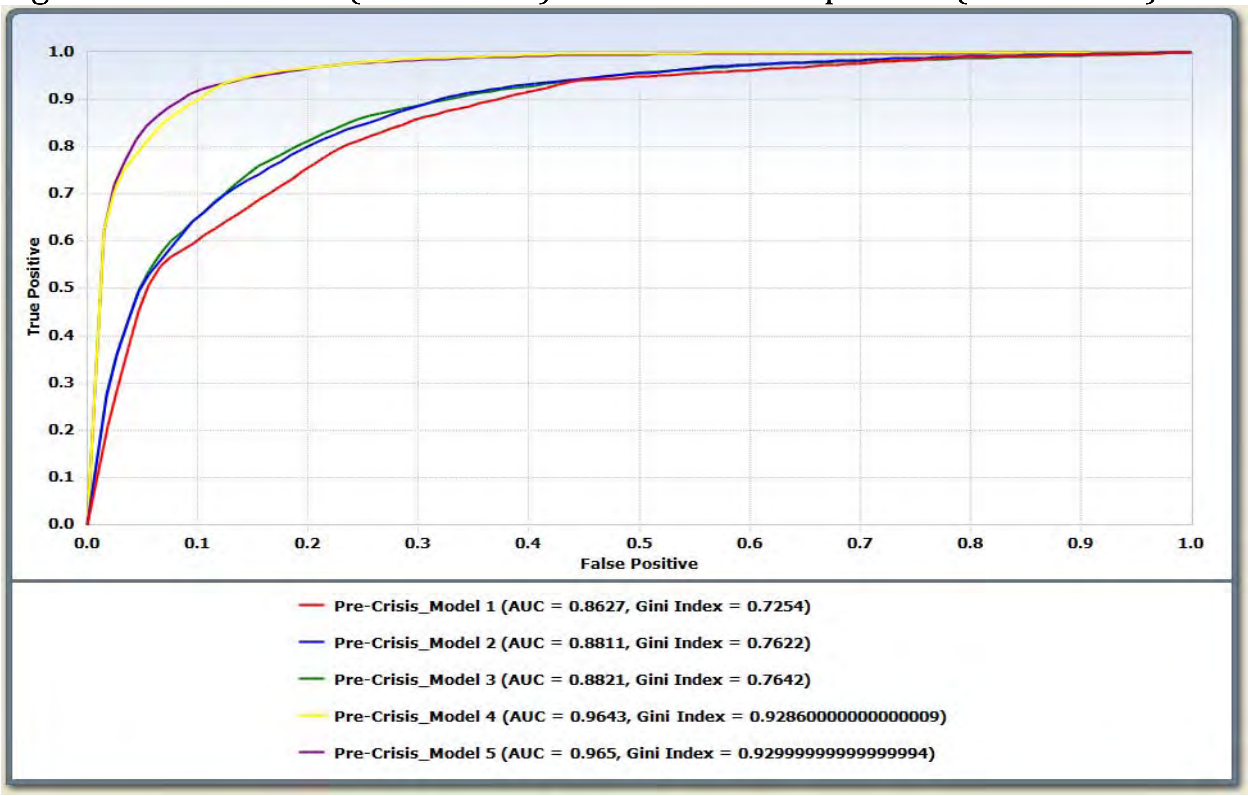


Figure 3A: Crisis Period (2007-2009): K-S Chart Comparison (Models 1-5)

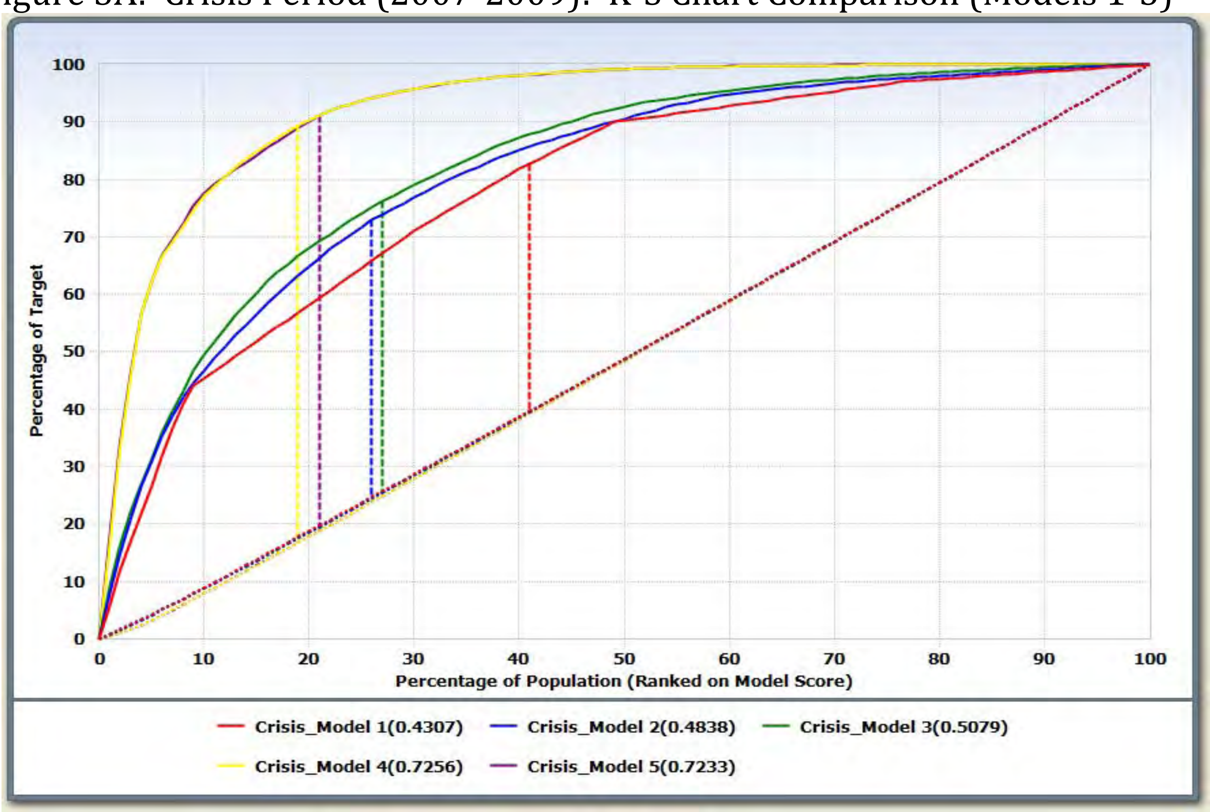


Figure 3B: Crisis Period (2007-2009): ROC Chart Comparison (Models 1-5)