WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 16-05
CONSUMER RISK APPETITE, THE CREDIT CYCLE, AND THE HOUSING BUBBLE

Joseph L. Breeden
Prescient Models LLC

José J. Canals-Cerdá
Supervision, Regulation, and Credit
Federal Reserve Bank of Philadelphia

February 2016

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Consumer Risk Appetite, the Credit Cycle, and the Housing Bubble

Joseph L. Breeden and José J. Canals-Cerdá*

February 2016

Abstract

We explore the role of consumer risk appetite in the initiation of credit cycles and as an early trigger of the U.S. mortgage crisis. We analyze a panel data set of mortgages originated between the years 2000 and 2009 and follow their performance up to 2014. After controlling for all the usual observable effects, we show that a strong residual vintage effect remains. This vintage effect correlates well with consumer mortgage demand, as measured by the Federal Reserve Board's Senior Loan Officer Opinion Survey, and correlates well to changes in mortgage pricing at the time the loan was originated. Our findings are consistent with an economic environment in which the incentives of low-risk consumers to obtain a mortgage decrease when the cost of obtaining a loan rises. As a result, mortgage originators generate mortgages from a pool of consumers with changing risk profiles over the credit cycle. The unobservable component of the shift in credit risk, relative to the usual underwriting criteria, may be thought of as macroeconomic adverse selection.

Keywords: credit risk, credit cycle, mortgages, lending standards, financial crisis

JEL Classifications: G20, G21, G32

* Contact information: José J. Canals-Cerdá, Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106; 215-574-4127, fax: 215-574-4146; e-mail: jose.canals-cerda@phil.frb.org. Corresponding author: Joseph L. Breeden, Prescient Models LLC, 300 Catron St., Suite B, Santa Fe, NM 87501; 505-670-7670; e-mail: breeden@prescientmodels.com. We thank Sharon Tang for outstanding research support and Amy Sill for outstanding logistic support. We are particularly grateful to William W. Lang for his support and assistance on this project. We would also like to gratefully acknowledge the assistance of the Editorial Services team at the Federal Reserve Bank of Philadelphia. Any remaining errors or omissions are our own. The views expressed in the paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers.

1. Introduction

In our experience developing models for forecasting and stress testing portfolio credit risk through the U.S. mortgage crisis, we have often observed the suboptimal performance of standard underwriting measures, insufficient to explain observed variations in credit quality. In this paper, we explore the possible causes of this unexplained variation and conjecture that consumer risk appetite may be a root cause. We refer to this effect as *macroeconomic adverse selection* to emphasize that loans exhibit anomalous credit risk because of the consumers' perception of macroeconomic conditions.

Changes in default risk that cannot be observed via standard credit scores and are suspected of being caused by consumer behavior are generally referred to as adverse selection. The macroeconomic adverse selection mechanism that we consider in this paper relates to anomalous credit risk associated with consumers' perception of macroeconomic conditions. In this regard, the aim of our paper is similar to that of other studies by Breeden, Thomas, and McDonald (2008) and Calem, Cannon, and Nakamura (2011), which will be discussed in some detail later in this section.

In contrast, in the standard example, adverse selection can impact a specific lender when it fails to respond to precautionary product or pricing changes made by its peers. Through the lender's inaction, consumers with lower credit risk are drawn to other lenders, leaving only the riskier borrowers for the unresponsive lender. In this scenario, the credit risk faced by the lender for the originated pool of loans can be much worse than what one could expect using traditional measures of credit quality, such as borrower credit scores. In terms of nomenclature, we have chosen to relabel this standard form of adverse selection as *competitive adverse selection* to differentiate it from the *macroeconomic adverse selection* mechanism that is the subject of analysis in this paper.

When the real estate bubble burst in the U.S. and across several European countries, it precipitated a deep financial crisis accompanied by an unsettling sovereign crisis in Europe. Understanding the mechanisms that led to the creation of the real estate bubble can prove extremely helpful, particularly for the purpose of implementing appropriate policies to minimize the risks of asset bubbles in the future. Recognizing this, the analysis of the leading factors contributing to the real estate bubble has generated a growing body of research. In the following, we review some of the most plausible proposed explanations and highlight our contribution to this literature.

Researchers in the empirical macroeconomics field have pointed at the simultaneity of rising asset values and current account deficits in the U.S. as well as other countries affected by real estate bubbles.[1] Their analysis suggests that current account deficits need to be accompanied by mispricing risk and falling lending standards to generate bubbles. In a similar vein, some economists have pointed out that the unusually low interest rates in the years before the crisis may have exacerbated the housing boom and bust (Taylor, 2014). Other authors, however, are critical of that view. Bernanke (2010) argues that monetary policy during that period was close to his preferred Taylor rule and was appropriate, given deflationary concerns at the time. Furthermore, significant increases in house prices preceded the period of accommodative monetary policy. In addition, cross-country analysis does not support the view that monetary policy played a fundamental role in the housing bubble. MacGee (2010) points out that Canada followed a monetary policy similar to that of the U.S. but did not suffer from a housing bubble.

Existing empirical microeconomics research points at mispricing risk and falling lending standards as fundamental catalysts of the crisis. In particular, researchers have considered the impact of investors in the mortgage market, either through direct purchases of houses or through the purchase of mortgage-backed securities. Haughwout, Lee, Tracy, and van der Klaauw (2011) point to the increasing role played by investors during the bubble years. Specifically, they document that investors were responsible for almost half of purchase mortgage originations at the peak of the market bubble. Investors were also associated with higher rates of default after the bubble burst.

Several authors have argued that securitized loans were originated using lower lending standards than loans held in bank portfolios. Elul (2015) calculates that, after controlling for observable risk factors, loans that are privately securitized have a 20 percent higher rate of becoming delinquent. His finding is consistent with research by Keys, Mukherjee, Seru, and Vig (2010). These authors point out that the securitization framework can reduce lenders' incentives to monitor lending standards (see also Nadauld and Sherlund, 2013). Securitization also may have contributed to lowering lending standards more broadly through its effects in a competitive market. Levitin, Pavlov, and Wachter (2009) argue that not only was private-label securitization a contributor to the crisis, but it was in fact the root of the crisis.[2] Ruckes (2004) describes theoretically a mechanism of transmission of low screening activity

[1] See Adam, Kuang, and Marcet (2011); Bergin (2011); Gete (2014); and In't Veld, Kollmann, Pataracchia, Ratto, and Roeger (2014).

[2] Levitin and Wachter (2013) argue that securitization was also responsible for the commercial real estate bubble in the U.S.

resulting from intense price competition among lenders.[3] Foote, Gerardi, and Willen (2012) take a contrarian view and argue that investment decisions made during the bubble years were rational and logical given investors beliefs about future house prices at the time.

Several authors have focused their attention on the way lending standards were lowered during the years before the bubble burst. Dell'Ariccia, Igan, and Laeven (2012), using mortgage origination information from the Home Mortgage Disclosure Act (HMDA), document the lowering of lending standards, particularly in areas that experienced faster credit demand growth. Palmer (2014), using data from privately securitized subprime mortgages, points out that mortgages originated in the two years before the cycle were about three times more likely to default within a three-year period than mortgages that originated around 2003. He argues that one-third of the increase in defaults can be attributed to changing borrower and loan characteristics, while the remaining two-thirds can be attributed to the price cycle.

Previous studies of the U.S. mortgage crisis have suggested that factors beyond those visible to the lenders had a strong impact on credit quality. Breeden (2011) analyzed a 15-year data set of mortgage performance employing a dual-time dynamics approach (Breeden et al., 2008) and found that dramatic cycles in credit quality occurred three times during the observation period, even after segmenting by product type, credit score, and loan-to-value. Further, they found that these cycles correlate to macroeconomic factors, such as changes in housing prices and mortgage interest rates. Similarly, Calem et al. (2011) used a combination of competing risk models and panel regression to show that riskier households tended to borrow more on their home equity loans when the expected unemployment risk increased.

In this paper, we quantify the impact of macroeconomic adverse selection on a data set of first-lien, installment, fixed-rate, conventional mortgages. We intentionally avoid option adjustable-rate mortgages (option ARM) and negative amortizing products to focus specifically on the question of the impact of macroeconomic adverse selection effects in this core mortgage product. We create a complete loan-level probability of default (PD) model that includes all of the standard predictive factors (loss timing versus age, also known as the life cycle; credit risk scoring attributes, such as FICO score, loan-to-value, etc.; and macroeconomic drivers, such as unemployment and house prices), and, using this framework, we demonstrate that a strong vintage-based effect persists beyond these observables.

[3] See Berlin (2009) for a survey of alternative theories of the bank lending cycle.

Additionally, we demonstrate that this residual credit risk is highly correlated to consumer mortgage demand based on the Federal Reserve Board's (FRB) Senior Loan Officer Opinion Survey (SLOOS) and changes in mortgage pricing at the time of loan origination. To our knowledge, this is the first paper to affirm correlation of credit risk to consumer demand and macroeconomic factors for residential mortgages, after controlling for all available scoring attributes.

In the next section, we present the data and provide descriptive statistics for some of the key variables in our sample. Section 3 contains the empirical methodology, and Section 4 presents the empirical model results. Section 5 concludes the paper.

2. Data and Descriptive Analysis

We analyze mortgage industry data from McDash Analytics, Residential Mortgage Servicing Database. This database mainly is composed of the servicing portfolios of the largest residential mortgage servicers in the U.S. and covers about two-thirds of installment-type loans in the residential mortgage servicing market. The database includes mortgages from Fannie Mae, Freddie Mac, Ginnie Mae, and private securitized portfolios as well as banks' portfolios. The original data set contains monthly loan performance data from mortgages originated dating from 1992. The data include a broad range of loan attributes from the underwriting process (such as product type, documentation type, loan purpose, property type, and zip code), borrower characteristics (such as credit score, debt-to-income ratio, and owner occupancy), and dynamic loan-level attributes (such as delinquency status, loan balance, current interest rate, and investor type).

Our sample of the mortgage industry data includes the full performance history of a randomly selected sample of loans in the McDash Residential Mortgage Servicing Database. Much has been written about how negative amortizing loans and second liens caused exceptionally high loss rates. To focus our analysis on the question of macroeconomic adverse selection, we restricted our analysis to fixed-term, fixed-rate, first-lien mortgages. We also restrict the sample of analysis to loan performance data from 2000 through 2014 on mortgages that originated from 2000 through 2009. We focus on modeling loan delinquency status of 60 to 89 days past due (DPD), as the later delinquency data were significantly thinner in the sample.

Table 1 lists the primary risk drivers used in our statistical analysis of credit risk. Relevant variables include loan-specific characteristics, such as term, documentation, loan-to-value (LTV) (defined as the percentage of the loan amount to the appraisal value at origination), loan purpose, loan source, and occupancy.

Table 1: Variable Definitions

Risk Score	*Borrower's FICO credit score*
Risk Score Dummies	*By ranges: 250–539, 540–619, 620–659, 660–699, 700–739, 740–779, 780–819, and 820+*
Debt-to-Income (DTI)	*Ratio of loan*
Jumbo	*Dummy variable for jumbo loan type*
Private Mortgage Insurance (PMI)	*PMI dummy*
Term	*Loan term (in months)*
Term Dummies	*Term dummies for ranges: up to 120, 120–180, 180–240, 240–360, 360+*
Documentation	*Loan documentation type; or unknown if type is not known*
Full	*Full documentation*
Low	*Low documentation*
No	*No documentation*
Loan-to-Value (LTV)	*Ratio of loan balance to current home value (i.e., at each observation point in time)*
LTV Dummies	*By ranges: 0–0.75, 0.75–0.80, 0.80–0.85, 0.85–0.90, 0.90–0.95, 0.95–1.00, 1.00+*
Loan Purpose	*Purpose of the loan; or unknown if type is not known*
New	*New loan*
Refinance	*Refinance loan*
Other	*Other (home improvement, debt consolidation, etc.)*
Loan Source	*Loan origination source; or unknown if type is not known*
Retail	*New loan originated by client organization*
Wholesale	*Wholesale origination*
Correspondent	*Correspondent and flow/co-issue loans*
Transfer	*Servicing rights purchased or transferred*
Other	*Other loan source*
Occupancy	*Occupancy type*
Owner	*Owner-occupied*
Nonowner	*Nonowner-occupied*
Other/Unknown	*Other occupancy type*
Vintage Year Dummies	*Dummy variables specific to the origination date*

Other borrower-specific characteristics include FICO scores at origination and debt-to-income (DTI) ratios at origination. In addition, we update the LTV variable over time using a repeated-sale house price index (the Federal Housing Finance Agency's (FHFA) house price index (HPI)) as well as the loan delinquency state at each point in time. Several variables included in our model specifications are represented as dummy variables, reflecting nonoverlapping ranges across the overall variable range. This approach allows us to estimate the potential nonlinear impact of particular variables without having to rely on specific functional form assumptions.

Table 2 presents descriptive statistics across origination vintages for the representative sample used in our analysis.

Table 2: Descriptive Statistics at Origination by Vintage for Fixed-Term and Rate, First-Lien Mortgages

	2000–03	2004	2005	2006	2007	2008	2009
Risk Score (mean)	715	710	710	704	705	715	738
Score in [250,540)	0.8	1.09	1.37	1.81	1.47	0.67	0.15
Score in [540,660)	13.01	15.78	15.73	18.8	19.68	17.99	10.72
Score in [660,700)	12.12	14.61	14.67	15.49	15.52	14.1	11.97
Score in [700,740)	15.25	16.81	16.57	16.51	16.84	16.44	15.93
Score in [740,900)	30.52	30.52	30.96	29.32	30.95	37.13	51.02
Jumbo	3.59	4.38	5.52	3.9	3.61	1.98	3.3
Term in months (mean)	305	323	335	340	343	338	333
Term up to 180 m.	28.24	21.64	12.51	8.26	7.04	10.63	12.31
Term 360+ m.	66.03	71.93	82.01	87.41	89.25	86.32	84.32
Documentation							
Full	25.54	25.97	28.03	30.35	39.26	48.72	54.01
Low	6.07	6.71	7.21	7.57	8.87	5.95	4.93
No Documentation	1.56	3.13	3.73	5.09	4.74	5.79	3.35
Unknown	66.83	64.19	61.03	56.98	47.12	39.54	37.71
Loan-to-Value (mean)	71.8	70.5	68.1	66.2	69.8	76.3	75.1
LTV in [0,0.75)	44.34	43.39	42.17	37.68	35.42	34.96	40.54
LTV in [0.75, 0.90)	31.05	34.26	37.32	38.28	35.07	29.83	26.65
LTV in [0.90,1.00)	17.15	15.65	13.44	14.94	19.37	29.1	24.77
LTV in [1.00+)	7.47	6.7	7.08	9.1	10.15	6.11	8.04
Loan Purpose							
New	30.37	36.63	39.03	42.37	39.6	36.49	30
Refinance	4.33	6.88	16.3	16.09	17.74	15.52	15.85
Other	50.16	39.2	27.26	23.7	23.88	26.25	39.43
Unknown	15.14	17.3	17.41	17.84	18.78	21.75	14.73
Loan Source							
Branch	39.17	37.17	34.69	33.72	37.72	41.55	46.05
Correspondent	22.85	24.93	25.04	25.79	26.69	32.55	36.29
Transfer	16.84	16.85	15.87	14.8	10.01	7.65	4.96
Other	12.63	14.97	16.77	17.36	20.08	16.2	11.51
Unknown	8.51	6.08	7.63	8.33	5.5	2.06	1.18
Occupancy							
Owner	91.66	89.32	83.72	83.06	86.19	87.92	92.67
Nonowner	5.94	7.28	7.79	8.69	8.09	5.6	2.68
Other/Unknown	2.4	3.4	8.49	8.25	5.72	6.48	4.65

Data source: McDash Analytics, Residential Mortgage Servicing Database

Observed changes in loan characteristics at origination are consistent with our expectation. We first observe a decrease in origination FICO scores across years up to 2007 and a reversal in this trend after that. However, changes in FICO scores are not dramatic. The percentage of originated loans with full documentation increased significantly during the crisis years, although this variable has a significant proportion of noncategorized loans. As expected, we also observe a decrease in nonowner-occupied loans during the crisis years. Overall, while we observe changes in the average characteristics of loans originated over the years, these changes are by no means dramatic. Thus, loan origination characteristics in the segment of the market composed of the fixed-term, fixed-rate, first-lien mortgages considered in our study remained relatively stable across the years and across observable risk dimensions.

3. The Modeling Approach

We follow the lives of the loans in our sample from origination to the time a loan is paid off or defaults. Our primary test for macroeconomic adverse selection is to create a loan-level model that includes all available observable factors and vintage fixed effects. The vintage fixed effects are intended to allow us to quantify the magnitude of adverse selection through time, if any. It will be important to compensate for life cycle, as a function of months-on-books,[4] and for changes in the macroeconomic environment that can contribute to higher losses across vintages. The comparison of estimation results from models with and without vintage effects will assist us in ascertaining the presence and relevance of a residual component that cannot be explained by standard scoring factors.

The odds of a loan defaulting can be represented as a combination of the average population odds of default (i.e., the average performance across all loans) and the idiosyncratic odds (i.e., divergence of an individual loan from the mean of the population) (Thomas, 2009),

$$\textit{log-odds of default}(a,v,t,i) = \log\big(\textit{population odds}\,(a,v,t)\big) + \log\big(\textit{idiosyncratic odds}\,(i)\big),$$

where *a* denotes the age of a loan (or months on book), *v* denotes the loan's vintage by origination date, *t* denotes the calendar date, and *i* denotes a loan-specific identifier.

Attempting to simultaneously estimate both the population odds and the idiosyncratic odds can lead to instability because of potential colinearity of macroeconomic and scoring factors when modeled on short timescales relative to the economic cycle. Therefore, we first create a *model of the population odds of default* as a function of months-on-books, vintage origination date, and calendar date. The

[4] Also known as the loss timing, seasoning, or credit loss hazard function. All of these refer to the changing probability of loss as a function of the age of the loan (months-on-books).

population odds are used as a fixed input to a panel data model such that the idiosyncratic odds are measured relative to the calendar date and age-varying population mean.

The two-stage approach of creating the population odds model and then the idiosyncratic odds model allows us to make explicit assumptions and tests around the linear trend specification error present in any model that includes age, vintage, and time effects. We can solve this in the population odds forecast before computing the idiosyncratic odds, so that the results will be robust. In the following subsections, we describe our approach to modeling population odds and idiosyncratic odds.

A. Modeling Population Odds

When modeling population odds, we are focused on drivers affecting all loans rather than idiosyncratic effects. The most important factors for modeling default rate are the life cycle and environmental effects.

The life cycle captures the fact that newly underwritten loans have much lower default rates than loans that are a few years old. Further, very old loans will be seasoned and are low risk. The precise shape of this life cycle function will depend upon the specific product and is usually measured nonparametrically, as in survival models. The life cycle function is also referred to as a hazard function or loss timing function.

Environmental impacts are traditionally thought of as the macroeconomic environment experienced by all active loans. Changes in unemployment and house prices are the primary drivers of mortgage defaults. However, other portfolio management drivers may be present. Because we are conducting an industrywide study, these drivers would have to be industrywide portfolio management trends, which do occur. By using the approach in which an environmental function is estimated directly from the data, we do not need to explicitly include macroeconomic factors in the model. In this way, we will capture the net effect of both macroeconomic drivers and portfolio management trends.

Any model that includes factors related to age of the loan, calendar date, and vintage will have a linear specification error because of the simple relationship, $a = t - v$, where a is age, t is time, and v is vintage (Breeden and Thomas, 2016). This specification error is well explained in the age-period-cohort (APC) literature (Mason and Fienberg, 1985; Glenn, 2005), and no general solution exists. In cases in which some of these dimensions are excluded, as with traditional credit scores that rely solely on information from the origination (vintage) date, a unique solution is obtained but at the cost of being unable to predict probabilities in future time periods.

To model the population odds, a Bayesian APC model was used to analyze data for the 60 to 89 DPD rate (Schmid and Held, 2007). Each rate was decomposed into a life cycle function with age of the account, *F(a);* vintage quality, *G(v);* and environment function with time*, H(t)*. Specifically,

$$\log\left(\frac{p(a,v,t)}{1-p(a,v,t)}\right) = F(a) + G(v) + H(t), \qquad \textbf{Equation 1}$$

where a logistic link function has been chosen because the probability of default follows a binomial distribution. This formulation does not consider any idiosyncratic variation; it just captures the mean of the distribution through age, vintage, and time.

A Bayesian APC algorithm was chosen because it provides a nonparametric estimate of the three functions. Relative to an initial mean-zero prior for each function, the values of the functions are adjusted to optimally predict the in-sample performance. A detailed description of the Bayesian APC algorithm is given in the Appendix.

In some of the later analyses, the data were segmented so we could study various effects. For segmented data, we can choose to segment any or all of the previously defined functions. For example, segmenting the environment function, *H(t)*, at the state level allows us to estimate it separately by state. Using this approach, we are able to include variations caused by the local economic environment in our estimates. Similarly, we will use segmentation to explore differences in the vintage function across segments. Note that in all the segments tested, the life cycle function was unchanged across the segments.

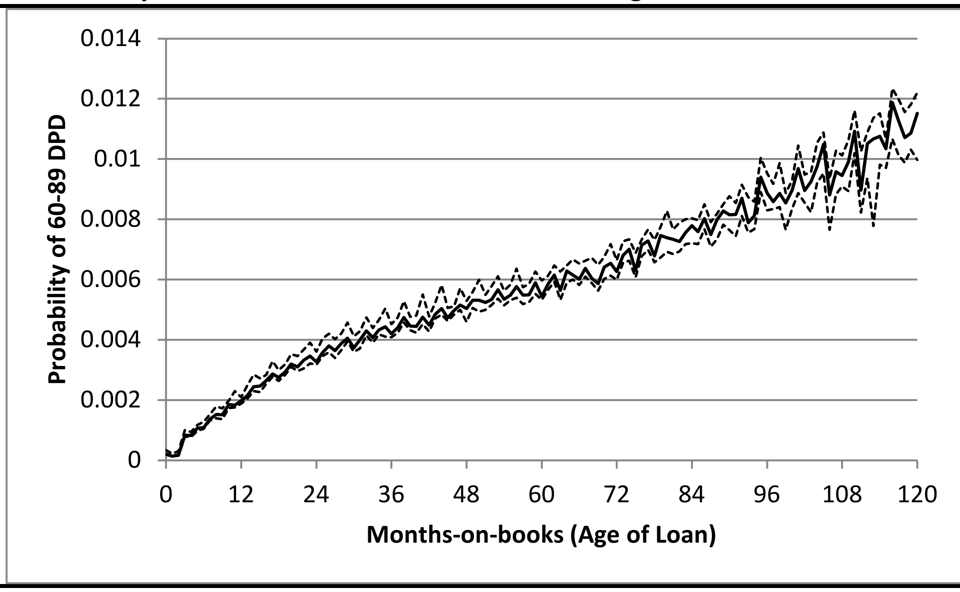
To test variations in the population odds by segment, we will create a set of models, as listed in Table 3.

Table 3: High-Level Model Description

Model Design	Life Cycle	Vintage	Environment
Primary model	Single function	Single function	By state
Score segmentation	By score	By score	By state
Purpose segmentation	By purpose	By purpose	By state

When the APC algorithm is applied to create the primary model, the algorithm provides point estimates for each value of the life cycle along with 5% and 95% confidence intervals (Figure 1). This represents the expected probability of delinquency for the entire sample. By estimating via the APC algorithm, it is normalized for portfolio variations in credit quality and environment, but it is conceptually equivalent to a hazard function.

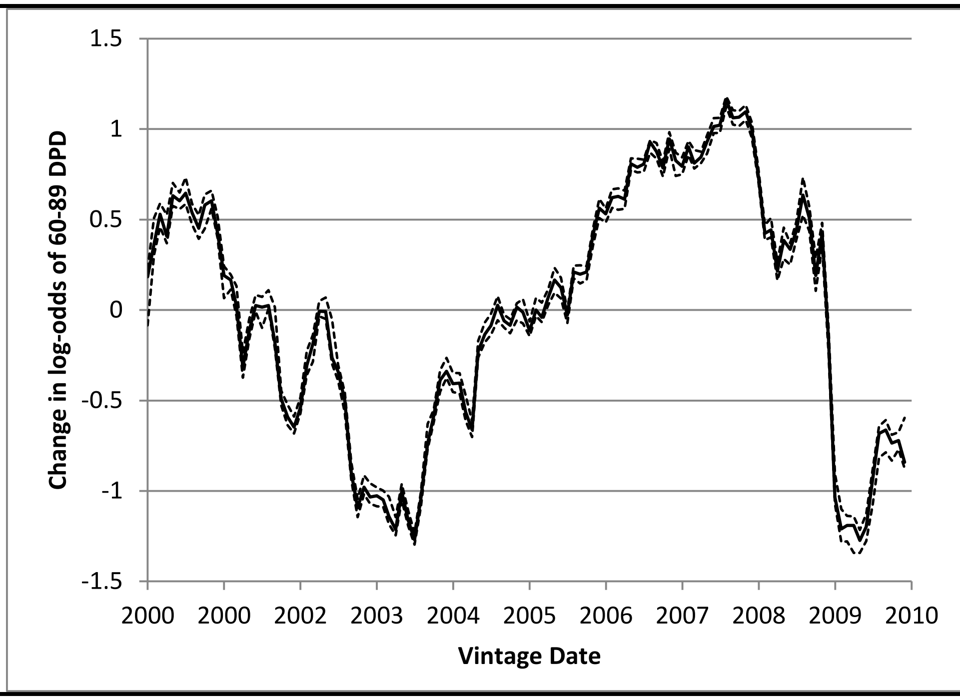
Figure 1: Life Cycle Function Estimated from the APC Algorithm for the Full Data Set



Note: This figure represents the expected probability of delinquency for the entire sample along with 5% and 95% confidence intervals. Results derived using McDash Analytics, Residential Mortgage Servicing Database.

Figure 2 shows the credit risk function obtained from the APC algorithm. It shows that loans originated in 2002 through 2004 had lower-than-average log-odds of delinquency, whereas loans originated in 2006 through 2008 had significantly higher-than-average log-odds of delinquency. The rest of this article will focus on testing the possible causes of this credit cycle.

Figure 2: Credit Risk Function Estimated from the APC Algorithm for the Full Data Set

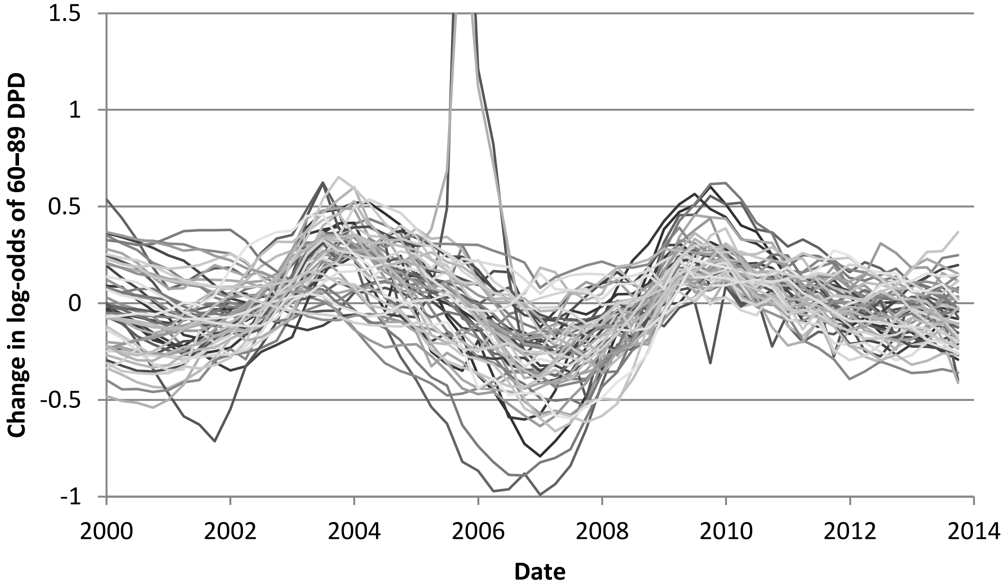


Note: Results derived using McDash Analytics, Residential Mortgage Servicing Database.

To measure the environment function, we segmented by state. As seen in Figure 3, in a summary across all risk bands, the states were highly correlated through the aftermath of the 2001 and 2009 recessions. In Figure 4, the large outliers in 2005 were Louisiana and Mississippi following Hurricane Katrina. In the final analysis, we segmented the environment function by both state and risk bands.

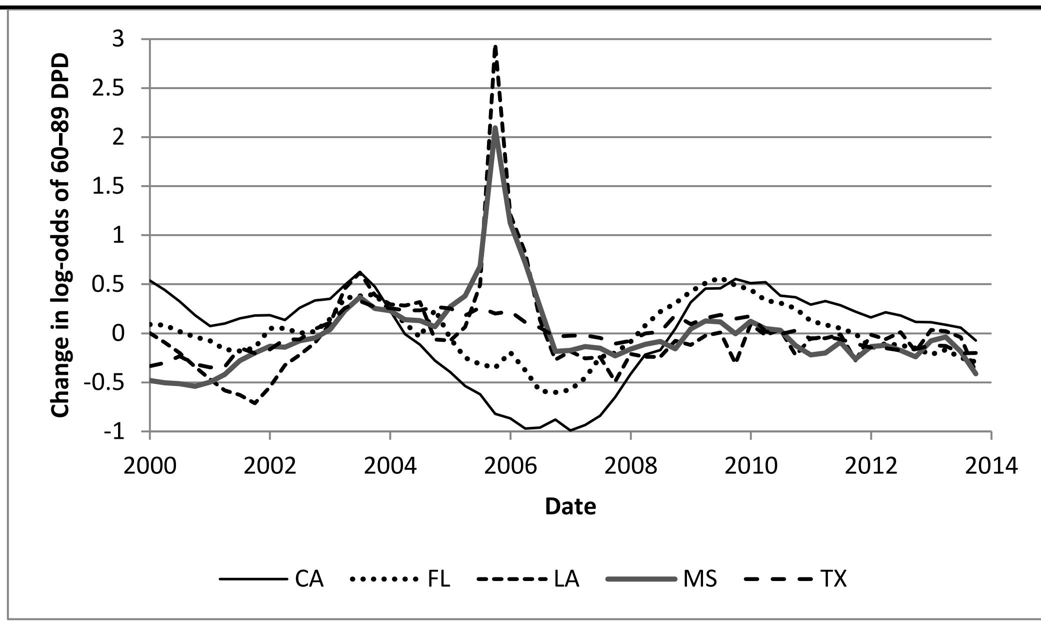
The environment function shows the change in log-odds of delinquency for all loans active on a given calendar date. The life cycle serves as the baseline against which the change is computed, so loans of different ages will be adjusted relative to their life cycle estimates.

Figure 3: Environmental Function by State Segments Estimated for 60 to 89 DPD



Note: Results derived using McDash Analytics, Residential Mortgage Servicing Database.

Figure 4: Environmental Function for Selected State Segments Estimated for 60 to 89 DPD



Note: California and Florida show the biggest swings through the recessions, but Louisiana and Mississippi show the greater impacts from Hurricane Katrina. Results derived using McDash Analytics, Residential Mortgage Servicing Database.

The results shown for life cycle, credit quality, and environment form a complete portfolio model in themselves but without causal explanation.

No macroeconomic model is needed for the macroeconomic adverse selection study. The environment function from the Bayesian APC algorithm will remove the maximum amount of temporal variability from the signal, most of which should be driven by the economy, but effects such as those from Hurricane Katrina are also obvious in the data. By using the environment function, any deviation as a function of calendar date will be removed, regardless of cause. That said, we created a panel data model of the environment functions measured by state segmentation. We built a single model simultaneously predicting the environment functions for all states, but we included fixed effects for states to allow for level shifts between them. The purpose of the panel data modeling of the environment functions with macroeconomic factors was to test for the necessity of a secular trend. If adding a c_t term was statistically significant, where c_t is an estimated constant for specific calendar date t, this would indicate that the environment functions are nonstationary with respect to macroeconomic effects. We designed the Bayesian APC to produce stationary environment functions, but the actual constraint we want is that the residuals be stationary when modeled against macroeconomic data. By showing that no time component is necessary, we can accept the decomposition as stable with respect to our design goals.

B. Modeling Idiosyncratic Odds

The final step in our analysis is to create loan-level models using first origination and then refreshed FICO and LTV attributes. The goal is to model the idiosyncratic odds separately from the population odds estimated via the APC algorithm. To create a score that incorporates the systematic effects (population odds) caused by life cycle and environmental impacts, we include the population odds as fixed offsets to a generalized linear model (GLM).[5] This has the effect of adjusting the log-odds on the left by the population odds as reflected by *F(a)* and *H(t)*:

$$\log\left(\frac{p_i(a,v,t,i)}{1-p_i(a,v,t,i)}\right) = c_0 + offset\big(F(a) + H(t)\big) + \sum_{j=1}^{n_s} c_j\, x_{ij} + \sum_{v=1}^{n_v} g_v,$$

Equation 2

[5] In the language of GLM, “fixed offsets” are factors that have a coefficient identically equal to 1.

where x_{ij} are the values of scoring attribute *j* for account *i*, c_j are the corresponding coefficients, and n_s is the number of scoring attributes. Again, $p_i(a,v,t)$ is the probability of a loan being 60 to 89 DPD.

The vintage function, *G(v)*, is not included in the offset (population odds) because we want to explicitly test how much of the population odds shift by vintage can be explained by population shifts in the scoring factors. Therefore, rather than include *G(v)* for the overall vintage function, we include fixed effects (dummy variables) for the vintages g_v to capture the residual vintage performance not explained by the scoring variables.

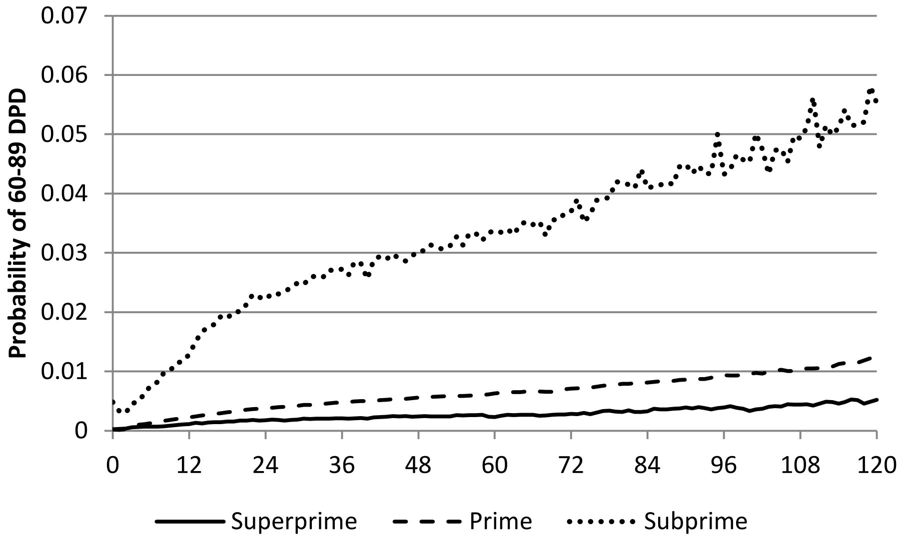
The method described here is broadly equivalent to a discrete time survival model with the added nuance of carefully controlling for the linear trend ambiguity. With the *F(a)* and *H(t)* functions as fixed offsets, the linear trend cannot change by inclusion of scoring factors. As soon as one function in *a*, *v*, or *t* is held fixed, the other two will be uniquely determined, as explained in the APC literature.

4. APC Model Results

To estimate the population odds, we estimated the Bayesian APC algorithm with 60 to 89 DPD as our target metric. The life cycle functions were segmented by subprime, prime, and super prime. In general, it also may be advisable to segment the life cycles by loan term. In our data, the loan terms were primarily for 10, 15, 20, and 30 years, and even with our large panel, we could not distinguish differences in the life cycles with this additional level of segmentation.

The y axis of the life cycle graph (Figure 5) is the expected average monthly delinquency rate averaged across the full time range.

Figure 5: Life Cycle Average Expected Monthly Delinquency Rate

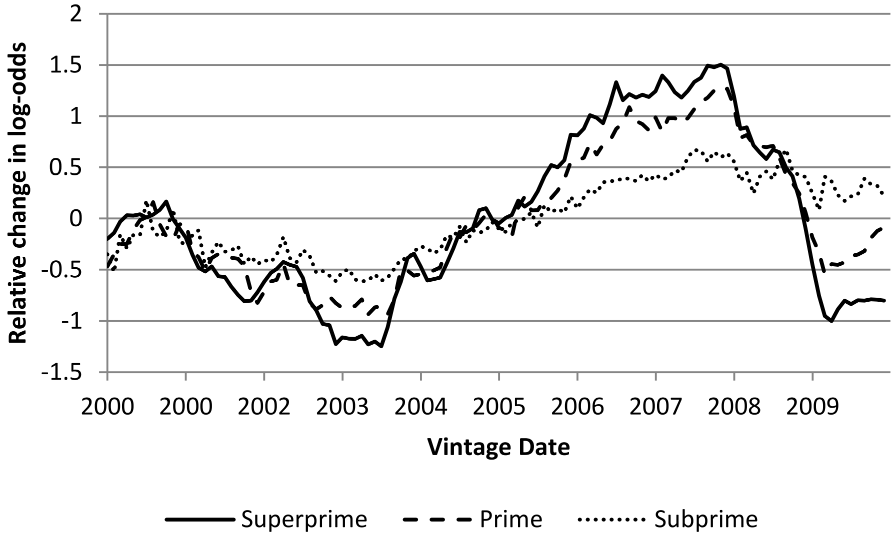


Note: The y axis represents the monthly conditional probability of default. Results derived using McDash Analytics, Residential Mortgage Servicing Database.

Figure 6 measures credit risk across vintages. Credit risk is measured as a relative scaling of the log-odds. The values shown represent the relative risk of a given vintage for the entire life of those loans. We observe that subprime loans have a smaller dynamic range than prime and superprime loans. Thus, on a relative basis, subprime loans tend to be less sensitive to the economic cycle in terms of underwriting (credit quality versus vintage) and the environment functions versus calendar date. This is a well-established result. However, in terms of total numbers of delinquent loans, the subprime segment will see the most growth for risky loans.[6]

[6] These findings regarding the performance of subprime versus prime loans over the life cycle are not specific to mortgages. Specifically, Canals-Cerdá and Kerr (2015) report similar findings in credit cards.

Figure 6: Relative Credit Risk by Vintage



Note: Results derived using McDash Analytics, Residential Mortgage Servicing Database.

A. Idiosyncratic Odds Results

The estimated credit risk function by vintage date from the APC algorithm captures both the known changes due to observable shifts in underwriting and possible unobserved effects for which we are searching. To distinguish between these two effects, we specify a loan-level probability model where the life cycle versus age by risk band and the environment function versus date by state are used as fixed offsets (Equation 2). In addition to these inputs, we also include the typical scoring attributes listed in Table 1. In particular, we estimate models with and without quarterly vintage effects and separately for subprime, prime, and super prime-segments.[7]

Applying this method to predicting the probability of being 60 to 89 DPD for the first-lien mortgage data provides the scoring results reported in Table 4. The table provides the GLM output for the full sample for all parameters except the vintage effects (graphed later), where the life cycle function and environment function by state are included in the model as fixed offsets. The coefficients shown are in line with industry intuition.

[7] Tables of parameter estimates are available from the authors.

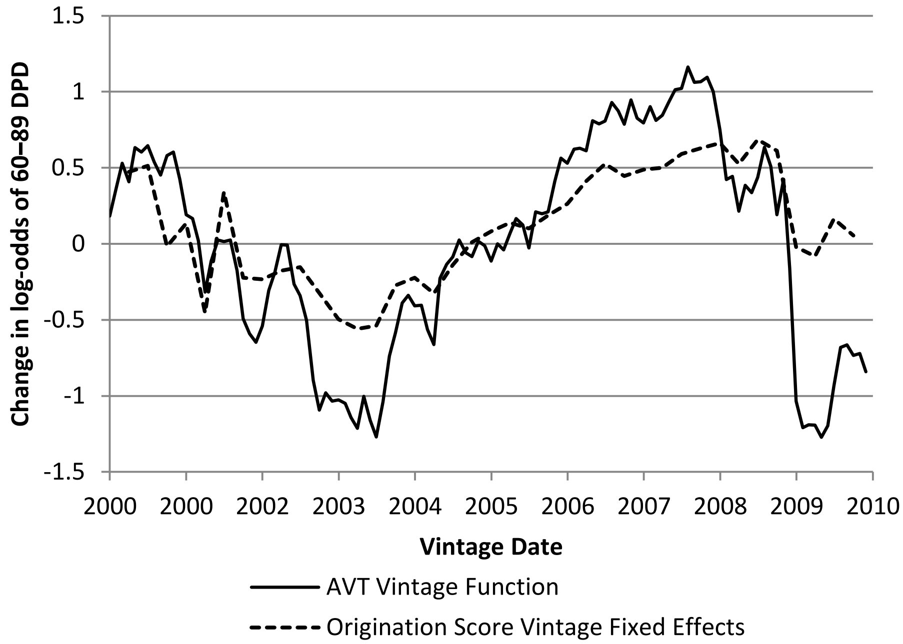
Table 4: Output Coefficients from the GLM Analysis of Mortgage Delinquency

Variables	Coef.	t-val	Variables (*cont.*)	Coef.	t-val
Intercept	2.268	67.83	**Source Channel**		
Jumbo loan	-0.128	-16.57	Retail	control	
Documentation			Wholesale	0.217	63.27
Full documentation	control		Correspondent	0.100	30.25
Low documentation	0.103	25.48	Transfer	0.240	38.63
No documentation	-0.030	-5.16	Other	0.437	32.19
Documentation unknown	0.135	40.61	**Occupancy**		
FICO at Origination			Owner	control	
up to 540	control		Nonowner	-0.109	-17.75
540 to 580	-0.188	-17.33	Other or unknown	-0.187	-17.84
580 to 620	-0.435	-44.66	**PMI**		
620 to 660	-0.807	-85.86	No	control	
660 to 700	-1.373	-145.22	Yes	0.119	32.06
700 to 740	-1.956	-202.66	Unknown	0.170	37.92
740 to 780	-2.671	-264.46	**Term**		
780 to 820	-3.380	-283.96	0 to 120	control	
820+	-3.623	-52.71	120 to 180	0.144	10.39
Loan-to-Value			180 to 240	0.407	27.38
0 to 0.75	control		240 to 360	0.593	44.05
0.75 to 0.8	0.157	40.09	360+	0.640	36.25
0.8 to 0.85	0.221	46.80	**Purpose**		
0.85 to 0.9	0.247	42.01	Purchase	control	
0.9 to 0.95	0.262	42.53	Refinance	-0.001	-0.41
0.95 to 1	0.305	46.79	purposeU	-0.462	-64.29
1 to 1.13	0.285	36.81	purposeZ	0.090	5.08
DTI	0.007	73.35			

Note: The model specification also includes quarterly vintage dummies that are not reported in this table. Results derived using McDash Analytics, Residential Mortgage Servicing Database.

Even though we included all available scoring factors, the fixed effect in vintage is still significant. When the APC decomposition is compared with having fixed effects in the scores, the major variation is still present (Figure 7). However, by including the scoring factors, the dynamic range for the vintage fixed effects is less pronounced than for the original credit risk function by vintage. In addition, the transition in 2009 is less dramatic. Both measures are normalized for life cycle and environment, so this suggests that only half of the variation in credit risk observed with APC is explainable by observable underwriting changes.

Figure 7: APC Vintage Function Versus Origination Score Vintage Effects

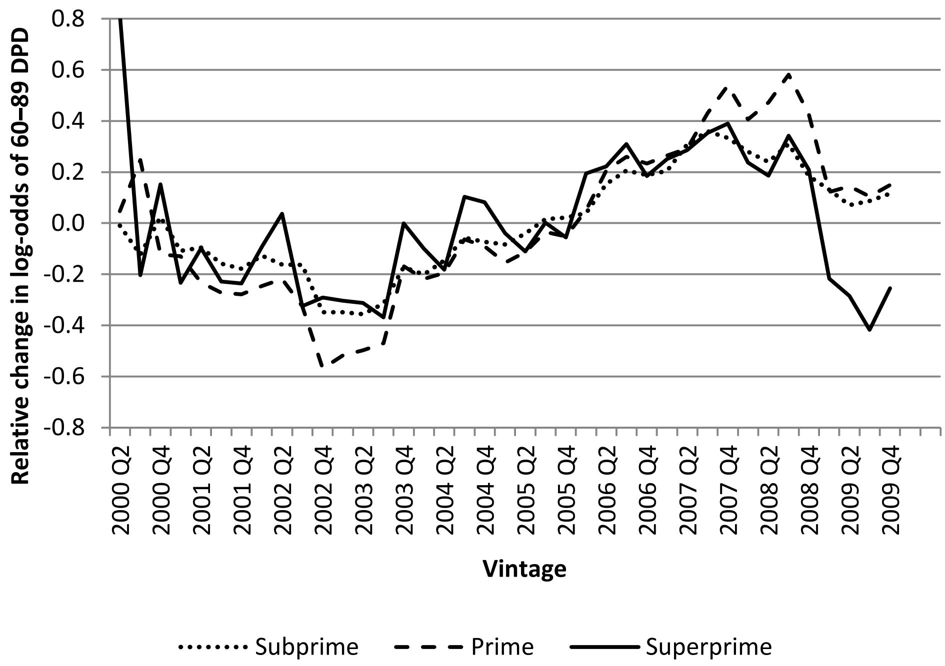


Note: Results derived using McDash Analytics, Residential Mortgage Servicing Database.

To test the result seen in Figure 7, the analysis was rerun segmented by score band, with entirely separate models built for each segment. Figure 8 shows the vintage fixed effect functions extracted from these models (Equation 2). The results are nearly identical, with the exception of the most recent history where the superprime function improves even more than the others.

Compare Figure 8, which resulted from the model with scoring factors, with Figure 6, which resulted purely from the APC analysis. The range of variation is less in Figure 8 than in Figure 6, but the vintage effects are much more aligned after adjustment for scoring factors. Thus, the inclusion of scoring factors does not eliminate the structure observed in Figure 6. Rather, the scoring factors clarify the unobserved vintage effects.

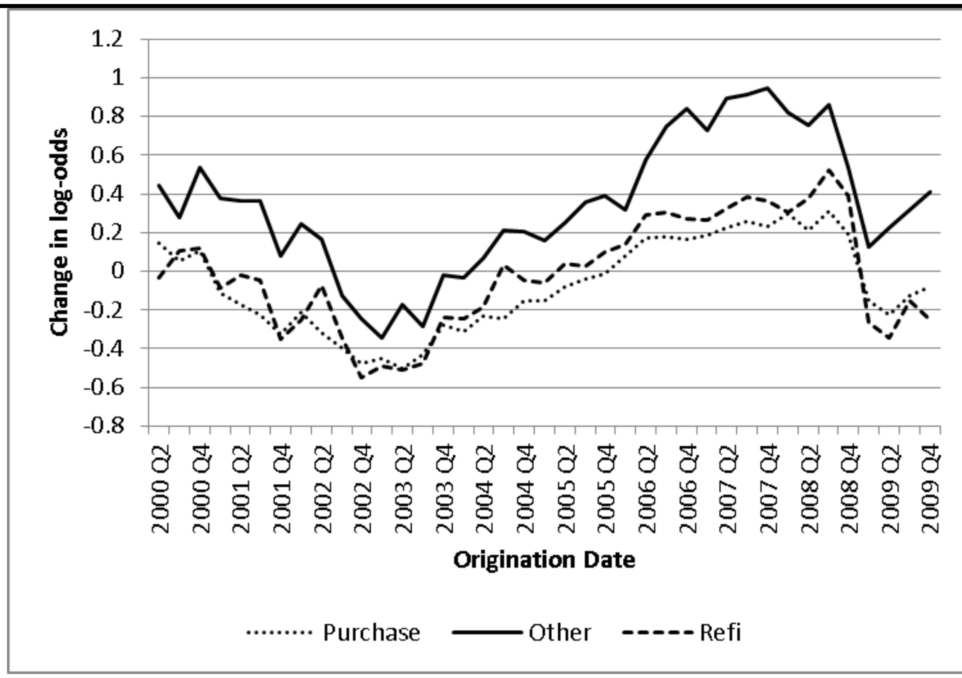
Figure 8: Relative Credit Risk by Vintage After Controlling for Scoring Attributes



Note: Results derived using McDash Analytics, Residential Mortgage Servicing Database.

The analysis by risk band in Figures 7 and 8 included loan purpose as a factor in the overall score. Rather than just including a scalar, we also tested to see if each loan-purpose segment exhibited the same dynamism with vintage. Figure 9 shows the separate estimates for the vintage fixed effect functions when segmented by loan purpose. We observe that "other" as a loan purpose segment is more risky across all vintages relative to purchase or refinance, but it is just a level shift equivalent to the scaling observed in the original model of Table 4. Thus, we conclude that the same credit risk cycle is present across loan purpose segments as well as risk bands.

Figure 9: Comparison of Vintage Fixed Effects by Purpose for Obtaining a Mortgage



Note: Results derived using McDash Analytics, Residential Mortgage Servicing Database.

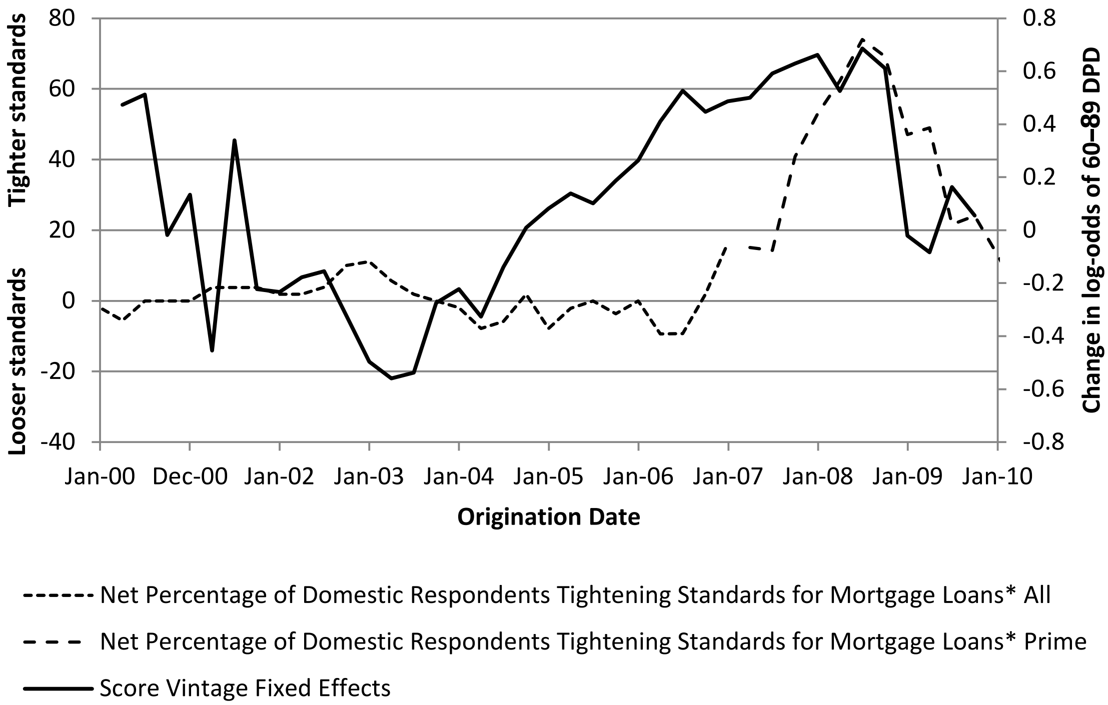
Any way we segment the data, we continue to observe that typical scoring factors do not capture all of the variation in credit risk by vintage. Vintage fixed effects (dummies) add significantly to the analysis and show that risk rose steadily from a low in early 2003 to a peak in 2007.

B. Comparison with Senior Loan Officer Opinion Survey

Given the similarity between risk bands in Figure 8, we continue now with a single credit risk function for all mortgages derived the same way as before but without segmentation. When looking for possible ways to explain the variation in credit risk after adjusting for available observed factors, we consider the FRB's SLOOS (Federal Reserve Board, 2014). This survey asks questions of senior loan officers regarding loan origination for several loan types. Before 2007, a single question was asked regarding mortgage underwriting. After 2007, this was separated into three questions for subprime, prime, and superprime mortgage origination. To create a continuous history, we computed the survey average after 2007.

Figure 10 shows the history with SLOOS (dashed line) for loosening and tightening of underwriting standards (left y axis). The solid line is the vintage fixed effect for first-lien, fixed-rate installment loans. These two lines should be anticorrelated. As underwriting standards are tightened, credit risk should decrease. In fact, a small positive correlation of 0.41 ± 0.38 is observed between these two measures.

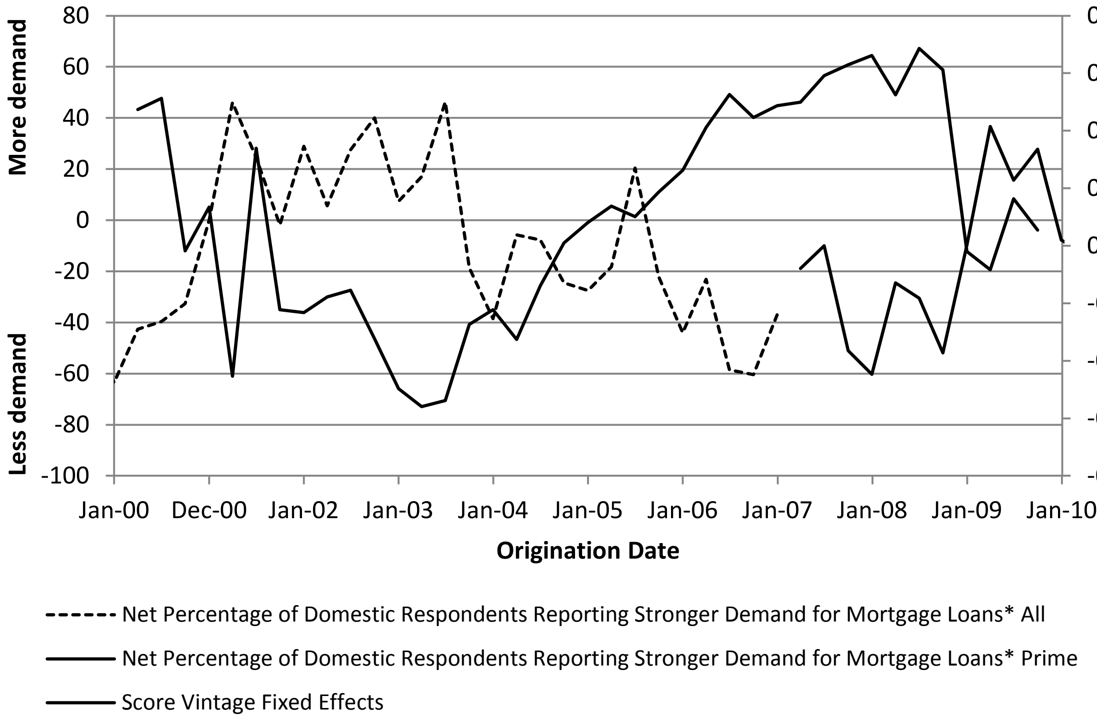
Figure 10: SLOOS Reported Underwriting Standards Versus First-Lien Vintage Effects



Note: Results derived using SLOOS and the McDash Analytics, Residential Mortgage Servicing Database.

The same survey asks the same senior loan officers a related question about consumer demand for loans. Figure 11 compares this measure of consumer demand (left y axis), dashed line, with the same measure of credit risk. Again, we computed the average demand index for data after 2007. Unlike the previous graph, this one shows significant anticorrelation of -0.69 ± 0.30. When consumer demand is high, credit risk is low, or when consumer demand is low, credit risk is high.

Figure 11: SLOOS Reported Mortgage Demand Versus First-Lien Vintage Effects



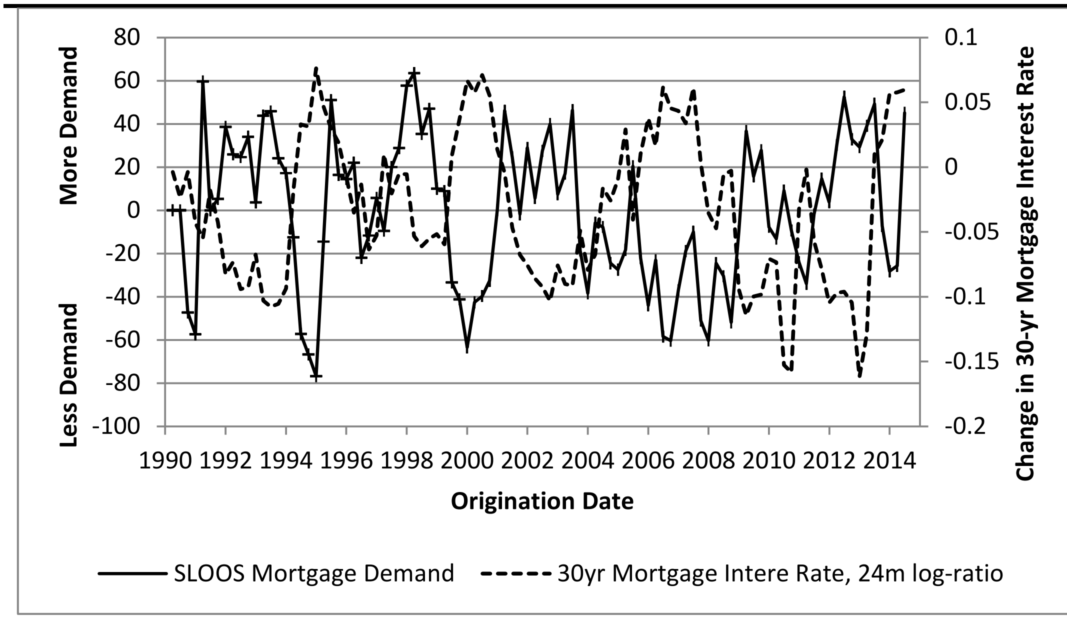
Note: Results derived using SLOOS and the McDash Analytics, Residential Mortgage Servicing Database.

The irony in these graphs is that the same senior loan officers answered both questions and, therefore, have all of the information shown here available to them; yet, their expectations on credit risk do not align with portfolio realities.

C. Comparison with Economic Drivers

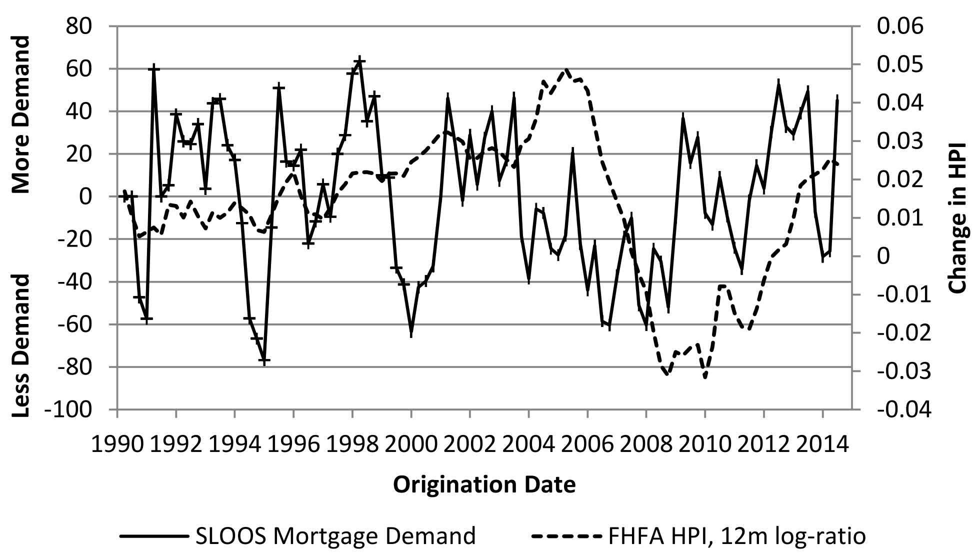
Because consumer demand changes significantly through time, we want to understand what might cause these changes. Figures 12 and 13 compare the SLOOS mortgage demand index with the change in 30-year mortgage rates and the change in HPI. The interest rate story is clear. We found that the optimal relationship was to the change over a 24-month horizon with a correlation of −0.56. The interpretation is that when interest rates have experienced a significant decline over an extended period of time, consumer demand rises, and conversely for rising rates.

Figure 12: SLOOS Mortgage Demand Index Versus the Change in 30-Year Mortgage Rates



Note: Data source SLOOS

Figure 13: SLOOS Mortgage Demand Index Against the Change in the HPI

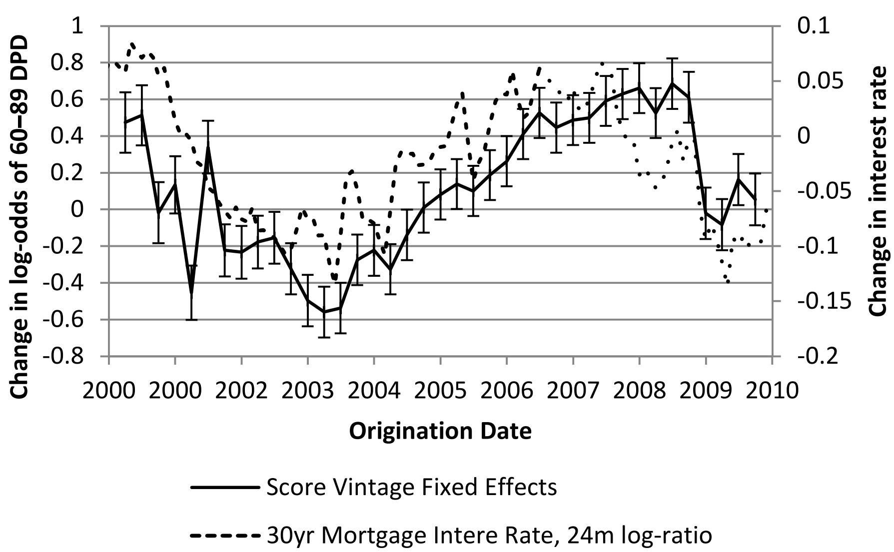


Note: Data source SLOOS

Figure 13 shows the relationship between mortgage demand and changes in the HPI. In a regression including change in the 30-year interest rate and change in HPI, we find that both are significant, and there is a positive relationship between demand and HPI. However, we really only have a single event in HPI against which to model. The results would be more reliable if we could conduct the analysis by geographic region, but demand is only available as a national measure.

Overall, the relationship between demand and interest rates is stronger and more intuitive. In Figure 14, we compare the vintage fixed effects for 60 to 89 DPD directly with the 24-month change in the 30-year mortgage interest rate without the intermediate measure of mortgage demand. Again, the relationship is clear.

Figure 14: Change in Log-odds of Default Versus Change in Interest Rate



Note: Results derived using McDash Analytics, Residential Mortgage Servicing Database.

Our best interpretation of these results is that consumer risk appetite changes based on economic conditions. Credit risk for the loan is a function of economic conditions at the time the loan was originated as well as later on if conditions worsen during the life of the loan.

As an industry, we tend to assume that credit risk is driven primarily by underwriting. However, it would be more accurate to say that underwriting is the process of selecting the best borrowers among those who are interested in getting loans. Thus, the primary drivers of risk may actually be the conditions that change the consumers' perspectives on their financial risks. Therefore, consumer risk appetite determines the pool of interested borrowers. Underwriting selects from those. When interest rates are falling, homes are more affordable and naturally conservative consumers come to the market to borrow. When interest rates rise, demand from conservative consumers dries up, and we are left with those consumers who are riskier in ways not always visible to bureau scores and other typical underwriting factors.

From the perspective of understanding the housing bubble, this suggests that the financially conservative consumers withdrew from the market in 2005 just as the problems with poor underwriting were taking hold. The pool of interested borrowers had a high proportion of risky consumers, and lenders went deeply into that pool. A disaster was in the making.

5. Conclusion

Although many explanations have been offered for the U.S. mortgage crisis, our research advocates that shifts in consumer risk appetite were a major contributing factor. In our approach, we used an age-period-cohort model to capture trends in the population odds. The age and period functions were then included in a generalized linear model of delinquency, which also included all available scoring factors. The original cohort function was, thereby, replaced with the scoring factors and a series of fixed effects to capture any residual structure. Although we had normalized for product life cycles, macroeconomic conditions by state, and all available scoring factors, the remaining vintage fixed effects were both significant and persistent through multiple segments.

The residual vintage fixed effects demonstrate a strong credit risk cycle, but correlations to external information suggest possible causes. Using the Senior Loan Officer Opinion Survey (SLOOS), we found that self-reported changes in underwriting standards did not correlate to the vintage fixed effects. This is reasonable because those changes in underwriting might already be captured in the scoring factors incorporated into the model. Surprisingly, the changes in consumers' mortgage demand reported by SLOOS correlated strongly to the vintage fixed effects, suggesting that periods of high demand correspond to low-risk vintages and periods of low demand correspond to high-risk vintages.

Further investigation of the SLOOS-reported changes in demand showed that both demand and the vintage fixed effects correlate strongly to long-term changes in interest rates. This suggests that declining interest rates drive increased demand from a broad spectrum of consumers, including the important low-risk borrowers. When interest rates are rising, the low-risk consumers no longer want mortgages, so the resulting vintages are lower in volume by much higher in risk.

Modern risk management relies heavily on statistical models. Often models are estimated over a short time horizon that does not cover a full cycle or a cycle with a sufficiently severe downturn period. Our paper emphasizes the importance of estimating models over a full cycle whenever possible. It is also of vital importance to pay special attention to the credit cycle when conducting model validation and for the analysis of model risk in particular. Recent regulatory guidance on model risk from the Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency (2011) highlights the increased awareness of regulatory agencies on this subject. Furthermore, the Basel II framework and the Comprehensive Capital Analysis and Review framework emphasize the use of models for effective supervision and surveillance. Our analysis stresses the importance of accounting for the credit cycle as an important element of model development, implementation, validation, and control. It is of particular importance for bank supervision to improve its understanding of the credit cycle as a catalyst of credit bubbles and its effects on the procyclicality of capital.

6. References

Adam K., P. Kuang, and A. Marcet (2011). "House Price Booms and the Current Account," NBER Working Paper 17224, National Bureau of Economic Research.

Bergin, P. (2011). "Asset Price Booms and Current Account Deficits," *Federal Reserve Bank of San Francisco Economic Letter*, December 5.

Berlin, M. (2009). "Bank Credit Standards," Federal Reserve Bank of Philadelphia *Business Review*, Q2:1–10.

Bernanke, B.S. (2010). "Monetary Policy and the Housing Bubble," speech given at the Annual Meeting of the American Economic Association, Atlanta, GA, available at www.federalreserve.gov/newsevents/speech/bernanke20100103a.htm.

Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency (2011). "Supervisory Guidance on Model Risk Management." OCC Bulletin 2011-12 and FRB SR Letter 11-7, available at www.federalreserve.gov/bankinforeg/srletters/sr1107a1.pdf.

Breeden, J.L. (2011). "Macroeconomic Adverse Selection: How Consumer Demand Drives Credit Quality," Proceedings of the Credit Scoring and Credit Control XII Conference, Edinburgh, Scotland.

Breeden, J.L., and L.C. Thomas (2016). "Solutions to Specification Errors in Stress Testing Models," *Journal of the Operations Research Society*, 67(2).

Breeden, J.L., L.C. Thomas, and J. McDonald III (2008). "Stress Testing Retail Loan Portfolios with Dual-Time Dynamics," *Journal of Risk Model Validation*, 2(2):43–62.

Calem, P., M. Cannon, and L.I. Nakamura (2011). "Credit Cycle and Adverse Selection Effects in Consumer Credit Markets — Evidence from the HELOC Market," Federal Reserve Bank of Philadelphia Working Paper 11-13.

Canals-Cerdá, J., and S. Kerr (2015). "Forecasting Credit Card Portfolio Losses in the Great Recession: A Study in Model Risk," *Journal of Credit Risk,* 11(1):29–57.

Dell'Ariccia, G., D. Igan, and L. Laeven (2012). "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market." *Journal of Money, Credit and Banking*, 44(2–3):367–384.

Elul, R. (2015). "Securitization and Mortgage Default," *Journal of Financial Services Research*, forthcoming.

Federal Reserve Board (2014). Senior Loan Officer Opinion Survey on Bank Lending Practices, available at www.federalreserve.gov/BoardDocs/snloansurvey.

Foote, C., K.S. Gerardi, and P.S. Willen (2012). "Why Did So Many People Make So Many *Ex Post* Bad Decisions? The Causes of the Foreclosure Crisis." Federal Reserve Bank of Boston Public Policy Discussion Paper 12-2.

Gete, P. (2014). "Housing Markets and Current Account Dynamics," unpublished manuscript, Georgetown University, available at http://faculty.georgetown.edu/pg252/H&CAbyGete.pdf.

Glenn, N.D. (2005). *Cohort Analysis*, 2nd ed. London: Sage.

Haughwout, A., D. Lee, J. Tracy, and W. van der Klaauw (2011). "Real Estate Investors, the Leverage Cycle, and the Housing Market Crisis," Federal Reserve Bank of New York Staff Report 514.

In't Veld, J., R. Kollmann, B. Pataracchia, M. Ratto, and W. Roeger (2014). "International Capital Flows and the Boom-Bust Cycle in Spain," Federal Reserve Bank of Dallas, Globalization and Monetary Policy Institute Working Paper 181.

Keys, B.J., T. Mukherjee, A. Seru, and V. Vig (2010). "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans," *Quarterly Journal of Economics*, 125(1):307–362.

Levitin, A., A.D. Pavlov, and S.M. Wachter (2009). “Securitization: Cause or Remedy of the Financial Crisis?” University of Pennsylvania Law School Research Paper 09-31, Institute for Law and Economics.

Levitin, A.J., and S.M. Wachter (2013). “The Commercial Real Estate Bubble,” *Harvard Business Law Review*, 3: 83–118.

MacGee, J. (2010). “Not Here? Housing Market Policy and the Risk of a Housing Bust,” E-brief, C.D. Howe Institute, August 1.

Mason, W.M., and S. Fienberg (1985). *Cohort Analysis in Social Research: Beyond the Identification Problem*. New York: Springer.

Nadauld, T.D., and S.M. Sherlund (2013). “The Impact of Securitization on the Expansion of Subprime Credit," *Journal of Financial Economics*, 107(2):454–476.

Palmer, C. (2014). “Why Did So Many Subprime Borrowers Default During the Crisis: Loose Credit or Plummeting Prices?” University of California, Berkeley Working Paper.

Ruckes, M. (2004). “Bank Competition and Credit Standards,” *Review of Financial Studies*, 17(4):1073–1102.

Schmid, V.J., and L. Held (2007). “Bayesian Age-Period-Cohort Modeling and Prediction — BAMP,” *Journal of Statistical Software*, 21(8):1–15.

Taylor, J.B. (2014). “The Role of Policy in the Great Recession and the Weak Recovery.” *American Economic Review*, 104(5):61–66.

Thomas, L.C. (2009). *Consumer Credit Models: Pricing, Profit, and Portfolios*. Oxford: Oxford University Press.

Appendix: Bayesian Age-Period-Cohort (APC)

The Bayesian APC implementation assumes that the functions *F(a)*, *G(v)*, and *H(t)* are values of random variables for a, v, and t, respectively. To estimate these functions from real data, we need to calculate the joint density for *F*, *G*, and *H*.

In the following, let $i = 1, \ldots, I$ denote the index of the age group, $j = 1, \ldots, J$ the index of the period (time), and $k = 1, \ldots, K$ the index of the vintage. The vintage index can be explicitly computed from age group and period index $v_k = t_j - a_i$, assuming that all variables are on the same time scale. We address the identifiability problem by imposing the following restrictions: $\sum_i F'(a_i) = \sum_k G(v_k) = \sum_j H(t_j) = 0$, where $F(a) = F'(a) + \mu$ and μ is the intercept.

The Bayesian APC algorithm uses a Bayesian hierarchical approach. Specifically, random walk (RW) priors of different orders are used for parameters *F*, *G,* and *H*. The random walk of first order (RW1) prior assumes a constant trend over the time scale, whereas the random walk of second order (RW2) prior assumes a linear time trend.

Next, we describe random walk priors for F', while the other parameters can be treated analogously. For random walks of first order, we assume

$$p\left(F'(a_1)\right) \propto \text{const.}$$

$F'(a_i)|F'(a_{i-1}) \sim N\left(F'(a_{i-1})\right), k_a^{-1})$, for i=2,…, I

with k_a^{-1} a precision parameter. This means that $F'(a_1)$ has a uniform distribution, $F'(a_i)$ has conditional Gaussian distribution with mean $F'(a_{i-1})$ and variation k_a^{-1}.

For random walks of second order, we assume

$$p\left(F'(a_1)\right) = p\left(F'(a_2)\right) \propto \text{const.}$$

$F'(a_i)|F'(a_{i-1}), F'(a_{i-2}) \sim N\left(2F'(a_{i-1}) - F'(a_{i-2})\right), k_a^{-1})$, for i=3,…I.

The joint distribution of $\mathbf{F}' = \left(F'(a_1),\ldots,F'(a_I)\right)$ is

$$p(\mathbf{F}') \propto k_a^{rank(\mathbf{R})/2} \exp\left(-\frac{k_a}{2}\mathbf{F}'^t\mathbf{R}\mathbf{F}'\right)$$

with the precision matrix **R** of rank $rank(\mathbf{R})$. For RW1, $I \times I$ matrix **R** is the following

$$\mathbf{R} = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

with $rank(\mathbf{R}) = I - 1$. For RW2,

$$\mathbf{R} = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & \cdots \\ -2 & 5 & -4 & 1 & 0 & 0 & 0 & \cdots \\ 1 & -4 & 6 & -4 & 1 & 0 & 0 & \cdots \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & \cdots \\ 0 & 0 & 1 & -4 & 6 & -4 & 1 & \cdots \\ 0 & 0 & 0 & 1 & -4 & 6 & -4 & \cdots \\ 0 & 0 & 0 & 0 & 1 & -4 & 6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

with $rank(\mathbf{R}) = I - 2$.

For the intercept μ, we use a flat prior

$$p(\mu) \propto \text{const.}$$

The joint prior distribution is

$$p(\mathbf{F},\mathbf{G},\mathbf{H}) = p(\mu)\,p(\mathbf{F}')\,p(\mathbf{G})\,p(\mathbf{H}).$$

The precision parameters k_a, k_v, and k_t are smoothing parameters and will be estimated simultaneously. A gamma distribution is used for the precision parameters. The full conditional of the APC parameter vectors are nonstandard distributions. Therefore, the algorithm uses a Metropolis–Hastings algorithm to sample from the posterior.

For the final functions, *F*, *G*, and *H*, we add the intercept term to *F* and linear terms to *F* and *G*, so that the life cycle function is calibrated to historic probabilities and *G* and *H* act as adjustments to that. *H* is assumed to have no net trend over the span of the training data, which is tested later by allowing for the possibility of including a time trend in the macroeconomic fit.