Vintage Effects in Loan Performance Models

Andrew Haughwout, Joseph Tracy and Wilbert van der Klaauw

Conference on "Banking, credit and macroprudential policy: what can we learn from micro data." Central Bank of Ireland, December 3-4, 2018

The views presented here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Dallas or New York, or the Federal Reserve System

Background

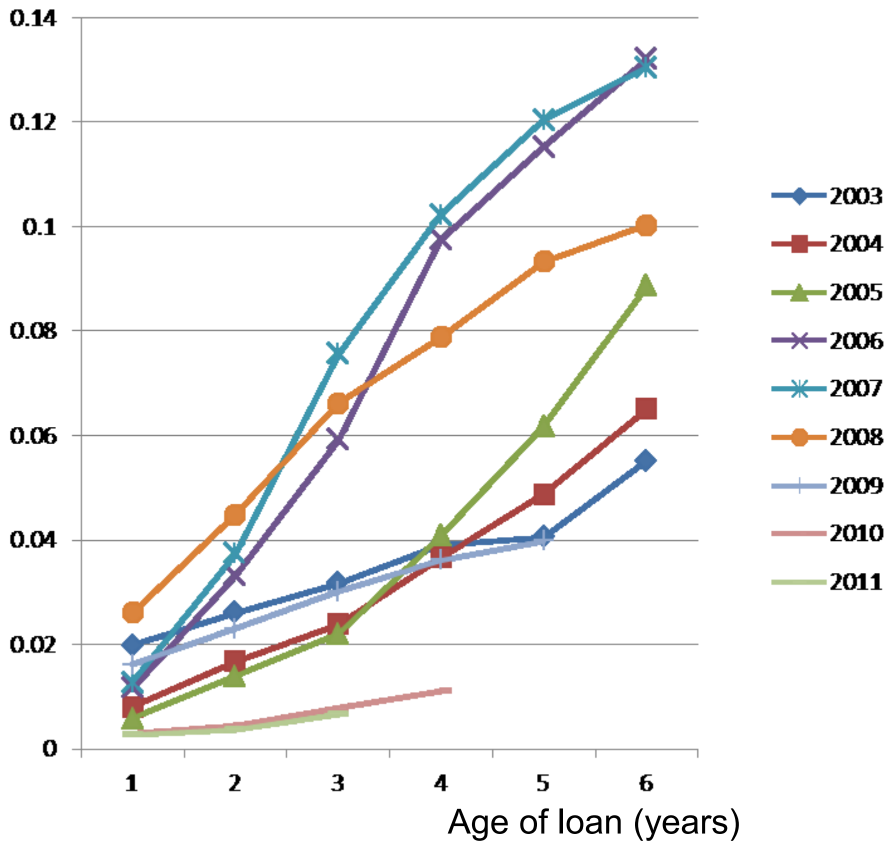
- Access to loan level mortgage servicing data has greatly enhanced empirical modeling of mortgage default and prepayment
  - Key application has been for stress testing bank portfolios

- Commonly find loan performance to co-vary strongly with a mortgage's origination period
  - Even after conditioning on a wide array of observed mortgage, borrower, property characteristics at origination, and on changes in the borrower's economic environment
  - Described as "vintage effects" (origination-period fixed-effects)
  - Their importance is documented by Demyanyk and van Hemert (2009) for subprime loans

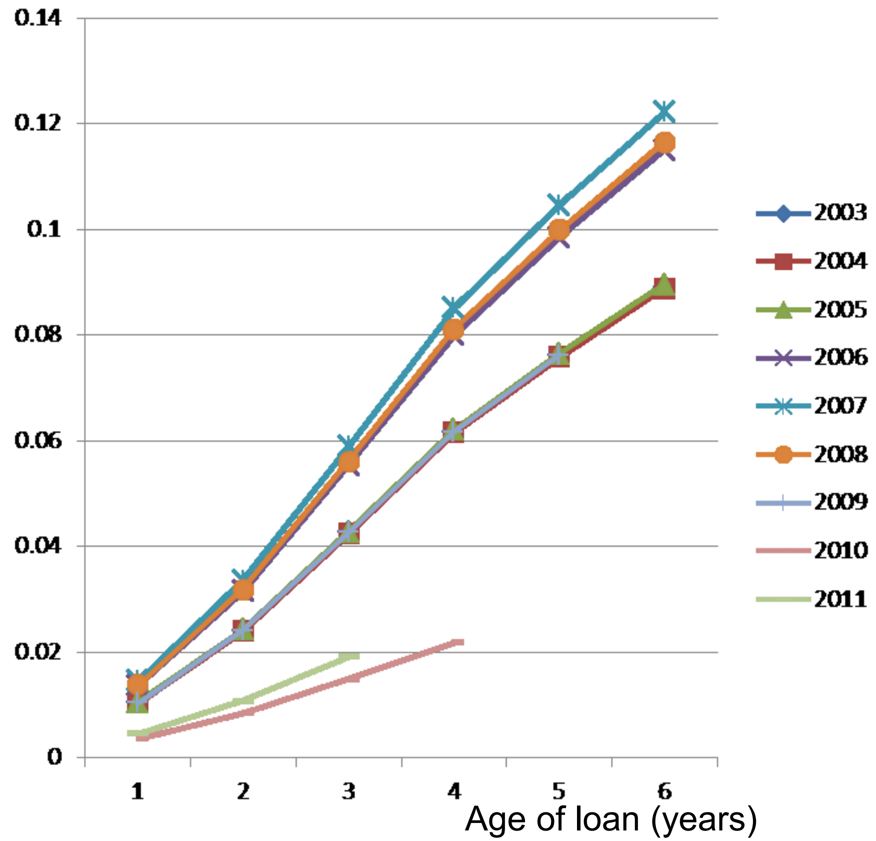
- Vintage effects also present when analyzing 30yr FRM portfolio loans originated between 2002-2011

Even after including many controls, find 2006-2008 vintages underperform, 2010-11 vintages outperform

(a) Actual Cumulative **Default** Rates by Vintage Year (default=90+ dpd)



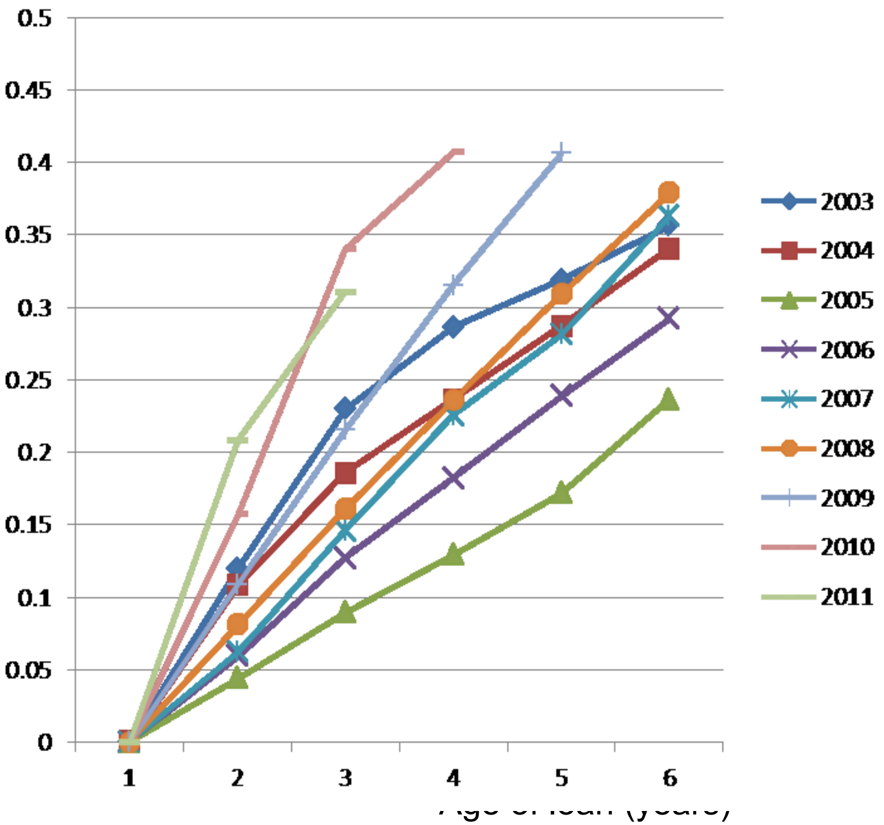
(b) Adjusted Cumulative **Default** Rates by Vintage Year



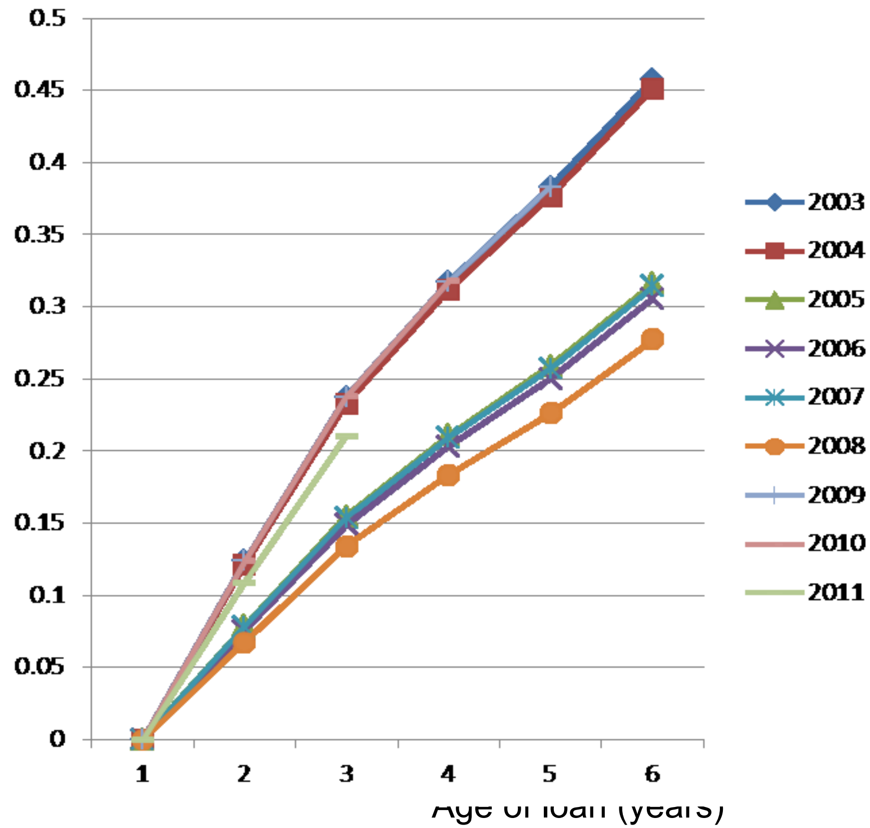
Adjusted rates account for differences in loan characteristics and economic conditions, following Demyanyk and Van Hemert (RSF 2009)

Similarly large unexplained differences in loan pre-payment rates across vintages (lower for 2006-2008)

(a) Actual Cumulative **Prepayment** Rates by Vintage Year



(b) Adjusted Cumulative **Prepayment** Rates by Vintage Year



Adjusted rates account for differences in loan characteristics and economic conditions, following Demyanyk and Van Hemert (RSF 2009)

Sources of vintage effects

Why do vintage effects show up in mortgage performance?

They capture variation in:

- Extent to which the origination LTV correctly reflects borrower equity in the property (piggyback seconds or inflated appraisals)
- Calculation of the borrower's front- and back-end DTI – what measure of income is used?
- Share of first-time borrowers across origination years – their credit scores do not reflect experience managing a mortgage
- Extent to which key elements of the loan file are documented by the borrower and verified by the underwriter
- Extent of misreporting of "occupancy status" on loan application – misreported investors ended up defaulting at high rates (Haughwout et al 2011)

Clearly vintage effects are important, but how best to model them?

- Goals of paper:
  - Explore suitability of standard fixed-effects approach to capturing vintage effects in loan default and prepayment models
  - Propose an alternative approach
  - Illustrate impact on stress-test type performance projections

Modeling vintage effects

- Standard approach: vintage fixed effects
  - Assumes that all mortgages of vintage share same unobserved traits
  - Important shortcoming – approach cannot easily capture dynamics in a vintage's effect over time
  - This is particularly problematic for out-of-sample forecasting of loan performance
  - Need marginal vintage effects, but average vintage effects are estimated

- Alternative approach: model vintage effects as differences in (endog. time-varying) distribution of unobserved characteristics
  - Each mortgage draws from distribution of unobserved heterogeneity determined by underwriting process at time of origination
  - Over time there is dynamic selection among surviving mortgages

Illustration of these two approaches using LPS data

- LPS mortgage selection criteria:

  - Portfolio loans originated between Jan 2002 and Dec 2011, with performance observed until July 2014
  - Purchase loans
  - 30yr FRM
  - Excludes loans with missing time-invariant or time-varying covariates

- Sample of 20,368 loans, with 336,645 quarterly observations

Standard vintage fixed-effects approach

- Conventional competing risk proportional hazards model with 2 exits (e.g. Deng, Quigley and van Order 2000):
  - **default** = mortgage reaches 90-days delinquent for the first time
  - **prepayment**

  When loan servicing is transferred, or duration is right-censored at end of observation period, treat as a random censoring

Probability that mortgage *i* **originated in year $\boldsymbol{\tau_i}$** defaults (k=1) or prepays (k=2) at duration *t* and calendar time τ_i+t:

$$h^k[t, x_i(t,\tau_i), \tau_i] = \bar{h}^k(t) e^{x_i(t,\tau_i)'\beta_k + \boldsymbol{\delta_{\tau_i k}}},$$

where $\bar{h}^k(t)$ is the baseline hazard, $x_i(t,\tau_i)$ is a vector of time-invariant and calendar-time-dependent characteristics and $\boldsymbol{\delta_{\tau_i k}}$ the exit-k specific fixed effect associated with origination vintage τ_i

Conventional Competing Risks Model Estimates

Estimate	Default	Pre-payment
Origination Balance ($10,000)	−0.011 (0.022)	**0.051** (0.002)
Jumbo Loan	0.069 (0.134)	**0.339** (0.036)
Single Family Residence	0.035 (0.063)	−0.038 (0.027)
Non Full Documentation	**0.725** (0.094)	0.015 (0.044)
Unknown Documentation	**−0.460** (0.101)	**−0.162** (0.053)
Time varying covariates		
LTV	**0.020** (0.001)	**−0.019** (0.001)
Unemployment rate	**0.062** (0.010)	**−0.012** (0.006)
Refinance incentive	−0.086 (0.060)	**0.722** (0.030)
Vintage effects		
2004	−0.004 (0.158)	−0.020 (0.059)
2005	0.005 (0.151)	**−0.477** (0.065)
2006	0.269 (0.148)	**−0.518** (0.066)
2007	**0.333** (0.151)	**−0.484** (0.067)
2008	0.283 (0.167)	**−0.633** (0.078)
2009	−0.188 (0.210)	−0.055 (0.079)
2010	**−1.061** (0.303)	0.002 (0.070)
2011	**−0.815** (0.297)	−0.136 (0.070)
Other controls: Spline in FICO, spline in DTI		
Number Exits	1718	6682

Notes: LPS data. Omitted vintages are 2002-2003 years. Number of loans and loan-quarter observations are 20,368 and 336,645, respectively. Flexible quarterly baseline hazards. Standard errors are given in parentheses.

Conventional Competing Risks Model Estimates

Estimate	Default	Pre-payment
Origination Balance ($10,000)	−0.011 (0.022)	**0.051** (0.002)
Jumbo Loan	0.069 (0.134)	**0.339** (0.036)
Single Family Residence	0.035 (0.063)	−0.038 (0.027)
Non Full Documentation	**0.725** (0.094)	0.015 (0.044)
Unknown Documentation	−0.460 (0.101)	**−0.162** (0.053)
Time varying covariates		
LTV	**0.020** (0.001)	**−0.019** (0.001)
Unemployment rate	**0.062** (0.010)	**−0.012** (0.006)
Refinance incentive	−0.086 (0.060)	**0.722** (0.030)
Vintage effects		
2004	−0.004 (0.158)	−0.020 (0.059)
2005	0.005 (0.151)	**−0.477** (0.065)
2006	0.269 (0.148)	**−0.518** (0.066)
2007	**0.333** (0.151)	**−0.484** (0.067)
2008	0.283 (0.167)	**−0.633** (0.078)
2009	−0.188 (0.210)	−0.055 (0.079)
2010	**−1.061** (0.303)	0.002 (0.070)
2011	**−0.815** (0.297)	−0.136 (0.070)
Other controls: Spline in FICO, spline in DTI		
Number Exits	1718	6682

Notes: LPS data. Omitted vintages are 2002-2003 years. Number of loans and loan-quarter observations are 20,368 and 336,645, respectively. Flexible quarterly baseline hazards. Standard errors are given in parentheses.

Conventional Competing Risks Model Estimates

Estimate	Default	Pre-payment
Origination Balance ($10,000)	−0.011 (0.022)	**0.051** (0.002)
Jumbo Loan	0.069 (0.134)	**0.339** (0.036)
Single Family Residence	0.035 (0.063)	−0.038 (0.027)
Non Full Documentation	0.725 (0.094)	0.015 (0.044)
Unknown Documentation	**−0.460** (0.101)	**−0.162** (0.053)
Time varying covariates		
LTV	**0.020** (0.001)	**−0.019** (0.001)
Unemployment rate	**0.062** (0.010)	**−0.012** (0.006)
Refinance incentive	−0.086 (0.060)	**0.722** (0.030)
Vintage effects		
2004	−0.004 (0.158)	−0.020 (0.059)
2005	0.005 (0.151)	**−0.477** (0.065)
2006	0.269 (0.148)	**−0.518** (0.066)
2007	**0.333** (0.151)	**−0.484** (0.067)
2008	0.283 (0.167)	**−0.633** (0.078)
2009	−0.188 (0.210)	−0.055 (0.079)
2010	**−1.061** (0.303)	0.002 (0.070)
2011	**−0.815** (0.297)	−0.136 (0.070)
Other controls: Spline in FICO, spline in DTI		
Number Exits	1718	6682

Notes: LPS data. Omitted vintages are 2002-2003 years. Number of loans and loan-quarter observations are 20,368 and 336,645, respectively. Flexible quarterly baseline hazards. Standard errors are given in parentheses.

Conventional Competing Risks Model Estimates

Estimate	Default	Pre-payment
Origination Balance ($10,000)	−0.011 (0.022)	**0.051** (0.002)
Jumbo Loan	0.069 (0.134)	**0.339** (0.036)
Single Family Residence	0.035 (0.063)	−0.038 (0.027)
Non Full Documentation	**0.725** (0.094)	0.015 (0.044)
Unknown Documentation	**−0.460** (0.101)	**−0.162** (0.053)
Time varying covariates		
LTV	**0.020** (0.001)	**−0.019** (0.001)
Unemployment rate	**0.062** (0.010)	**−0.012** (0.006)
Refinance incentive	−0.086 (0.060)	**0.722** (0.030)
Vintage effects		
2004	−0.004 (0.158)	−0.020 (0.059)
2005	0.005 (0.151)	**−0.477** (0.065)
2006	**0.269** (0.148)	**−0.518** (0.066)
2007	**0.333** (0.151)	**−0.484** (0.067)
2008	**0.283** (0.167)	**−0.633** (0.078)
2009	−0.188 (0.210)	−0.055 (0.079)
2010	**−1.061** (0.303)	0.002 (0.070)
2011	**−0.815** (0.297)	−0.136 (0.070)
Other controls: Spline in FICO, spline in DTI		
Number Exits	1718	6682

Notes: LPS data. Omitted vintages are 2002-2003 years. Number of loans and loan-quarter observations are 20,368 and 336,645, respectively. Flexible quarterly baseline hazards. Standard errors are given in parentheses.

Dynamics of vintage effects

- Standard vintage fixed-effects approach assumes vintage effects are uniform across loans belonging to same vintage and **time-invariant**

- However, in practice we see dynamics over time in a vintage's effect

- For example, if over time the riskiest mortgages default and drop out, the (unobserved) quality of the remaining mortgages in the pool may improve. Similarly, can have negative selection due to nonrandom prepayment over time of mortgages with better (unobserved) quality

- Any evidence of such dynamics in vintage effects? To illustrate, estimate conventional model with sample that updates as additional performance years and vintages are added over time

Convergence of average vintage effects

Table 2. Proportional Hazard Vintage Effect Estimates

a. Vintage Effects on **Default** ($e^{\delta_{\tau_i k}}$)

	Last vintage included (+1yr extra data)							
Vintage	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2011***
2004	0.670	0.871	1.007	0.993	0.937	0.955	0.958	0.996
2005	0.460	0.681	0.747	0.842	0.898	0.928	0.925	1.005
2006		1.335	1.568	1.477	1.312	1.249	1.180	1.309
2007			**1.532**	**1.486**	**1.419**	**1.284**	**1.247**	**1.395**
2008				1.786	1.502	1.234	1.132	1.327
2009					1.419	0.946	0.726	0.828
2010						0.229	0.291	0.346
2011							0.338	0.443

b. Vintage Effects on **Prepayment**

	Last vintage included (+1yr extra data)							
Vintage	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2011***
2004	0.895	0.907	0.947	0.984	1.022	1.022	0.994	0.980
2005	0.319	0.391	0.455	0.502	0.533	0.575	0.599	0.620
2006		0.530	0.544	0.518	0.551	0.583	0.610	0.596
2007			**0.434**	**0.427**	**0.503**	**0.556**	**0.590**	**0.617**
2008				0.464	0.471	0.510	0.529	0.531
2009					1.011	0.960	0.839	0.947
2010						1.001	1.010	1.002
2011							0.864	0.873

Convergence of average vintage effects

Table 2. Proportional Hazard Vintage Effect Estimates

a. Vintage Effects on **Default** ($e^{\delta_{\tau_i k}}$)

	Last vintage included (+1yr extra data)							
Vintage	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2011***
2004	**0.670**	0.871	1.007	0.993	0.937	0.955	0.958	0.996
2005	**0.460**	0.681	0.747	0.842	0.898	0.928	0.925	1.005
2006		1.335	1.568	1.477	1.312	1.249	1.180	1.309
2007			**1.532**	**1.486**	**1.419**	**1.284**	**1.247**	**1.395**
2008				1.786	1.502	1.234	1.132	1.327
2009					1.419	0.946	0.726	0.828
2010						0.229	0.291	0.346
2011							0.338	0.443

Initially 2004-2005 vintages outperform earlier vintages

b. Vintage Effects on **Prepayment**

	Last vintage included (+1yr extra data)							
Vintage	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2011***
2004	0.895	0.907	0.947	0.984	1.022	1.022	0.994	0.980
2005	0.319	0.391	0.455	0.502	0.533	0.575	0.599	0.620
2006		0.530	0.544	0.518	0.551	0.583	0.610	0.596
2007			**0.434**	**0.427**	**0.503**	**0.556**	**0.590**	**0.617**
2008				0.464	0.471	0.510	0.529	0.531
2009					1.011	0.960	0.839	0.947
2010						1.001	1.010	1.002
2011							0.864	0.873

Convergence of average vintage effects

Table 2. Proportional Hazard Vintage Effect Estimates

a. Vintage Effects on **Default** ($e^{\delta_{\tau_i k}}$)

	Last vintage included (+1yr extra data)							
Vintage	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2011***
2004	0.670	0.871	1.007	0.993	0.937	0.955	0.958	**0.996**
2005	0.460	0.681	0.747	0.842	0.898	0.928	0.925	**1.005**
2006		1.335	1.568	1.477	1.312	1.249	1.180	1.309
2007			**1.532**	**1.486**	**1.419**	**1.284**	**1.247**	**1.395**
2008				1.786	1.502	1.234	1.132	1.327
2009					1.419	0.946	0.726	0.828
2010						0.229	0.291	0.346
2011							0.338	0.443

By 2001 no evidence of outperforming

b. Vintage Effects on **Prepayment**

	Last vintage included (+1yr extra data)							
Vintage	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2011***
2004	0.895	0.907	0.947	0.984	1.022	1.022	0.994	0.980
2005	0.319	0.391	0.455	0.502	0.533	0.575	0.599	0.620
2006		0.530	0.544	0.518	0.551	0.583	0.610	0.596
2007			**0.434**	**0.427**	**0.503**	**0.556**	**0.590**	**0.617**
2008				0.464	0.471	0.510	0.529	0.531
2009					1.011	0.960	0.839	0.947
2010						1.001	1.010	1.002
2011							0.864	0.873

Convergence of average vintage effects

Table 2. Proportional Hazard Vintage Effect Estimates

a. Vintage Effects on **Default** ($e^{\delta_{\tau_i k}}$)

	Last vintage included (+1yr extra data)							
Vintage	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2011***
2004	0.670	0.871	1.007	0.993	0.937	0.955	0.958	0.996
2005	0.460	0.681	0.747	0.842	0.898	0.928	0.925	1.005
2006		1.335	1.568	1.477	1.312	1.249	1.180	1.309
2007			1.532	1.486	1.419	1.284	1.247	1.395
2008				1.786	1.502	1.234	1.132	1.327
2009					1.419	0.946	0.726	0.828
2010						0.229	0.291	0.346
2011							0.338	0.443

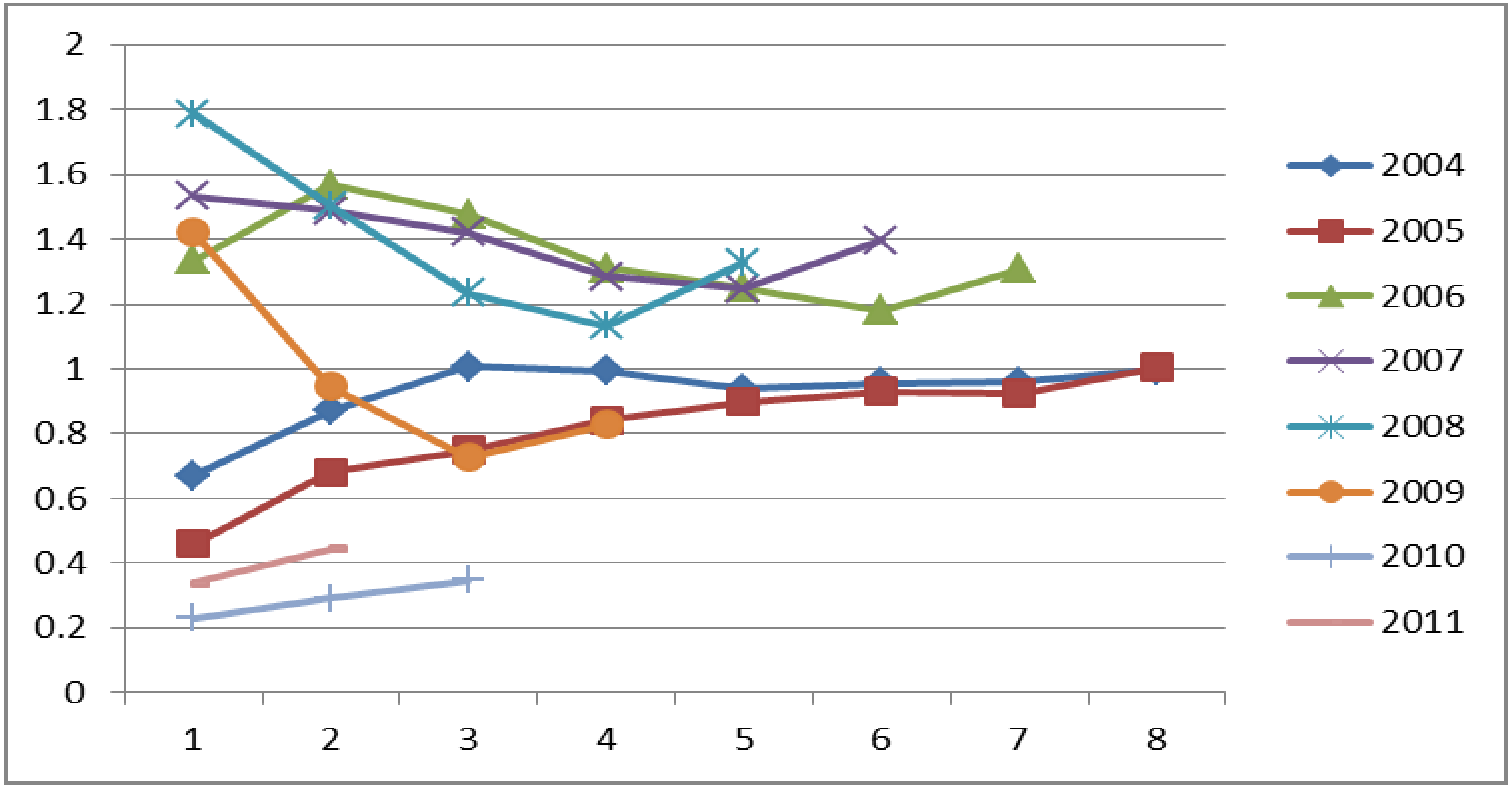
b. Vintage Effects on **Prepayment**

	Last vintage included (+1yr extra data)							
Vintage	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2011***
2004	**0.895**	0.907	0.947	0.984	1.022	1.022	0.994	**0.980**
2005	**0.319**	0.391	0.455	0.502	0.533	0.575	0.599	**0.620**
2006		0.530	0.544	0.518	0.551	0.583	0.610	0.596
2007			0.434	0.427	0.503	0.556	0.590	0.617
2008				0.464	0.471	0.510	0.529	0.531
2009					1.011	0.960	0.839	0.947
2010						1.001	1.010	1.002
2011							0.864	0.873

Similar pattern for prepayment vintage effects

Convergence in estimated vintage effects over time – consistent with dynamic selection

Vintage Effects on Default by Years Since Origination

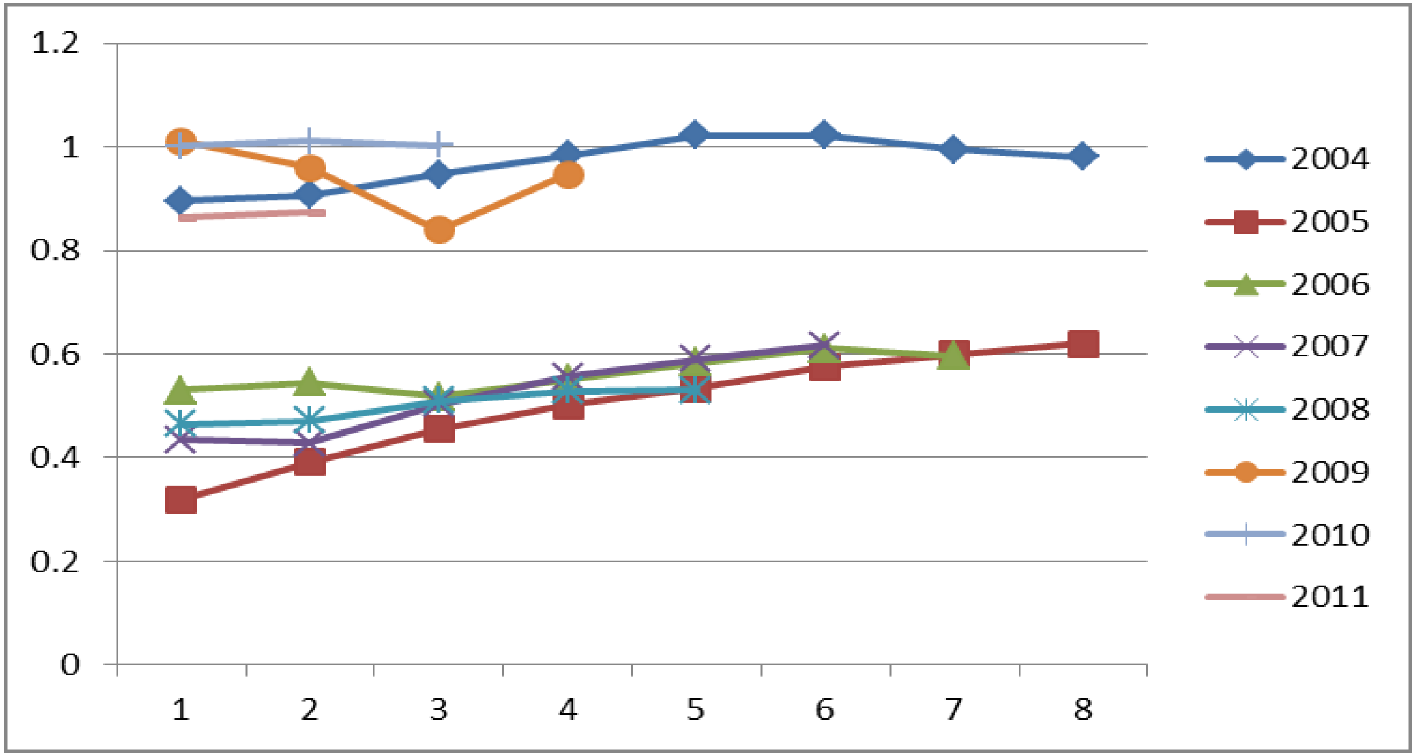


Years since origination year

Convergence towards performance of 2002-2003 vintage (reference)

Also slow convergence in estimated vintage effects for prepayment

Vintage Effects on Prepayment by Years Since Origination



Years since origination year

Gradual convergence towards performance of 2002-2003 vintage (reference)

Modeling with vintage fixed-effects

- Conventional model unable to capture dynamics in vintage effects
- Highly relevant for stress testing
  - Need to consider quality of loans remaining in portfolio today
  - What matters is how marginal vintage effects evolve into the future, not average realized vintage effects since origination

Alternative approach: model vintage effects as changes in unobserved heterogeneity distribution

Dependent Competing Risks Model with transition rates

$$h^k[t, x_i(t,\tau_i), \tau_i] = \bar{h}^k(t)e^{x_i(t,\tau_i)'\beta^k + \boldsymbol{v_i^k}}, \qquad k = 1,2,$$

- Where we specify the joint distribution for (v^1, v^2) as a flexible discrete multinomial distribution with J point of support (Heckman and Singer,1984)

- In doing so, we allow for *dependence* between $v^1$and v^2 and specify their bivariate distribution as a discrete multinomial distribution with J mass points $\mu^j = (\mu_1^j, \mu_2^j), j = 1, \dots, J$ with probabilities Pr(v^1=μ_1^j, v^2=μ_2^j) =π_j, $j = 1, \dots, J$.

- We then model vintage effects by allowing the mixture distribution to vary across vintages. More specifically, we allow the probabilities $\pi_j(\tau_i)$ associated with the mass points to vary across vintages.

Advantages of approach

- More natural way of modeling vintage effects – as capturing differences in the distribution of unobserved loan/borrower characteristics (e.g. as determined by underwriting process at time of origination)

- By explicitly accounting for unobserved heterogeneity – able to avoid estimation biases from ignoring such heterogeneity

- Captures dynamic selection among surviving mortgages – allows marginal vintage effect to differ from average vintage effect
  - Captures endogenous evolution of marginal vintage effect
  - Updated distribution reflects initial underwriting, as well as history of exposure to macroeconomic conditions including past incentives to refinance

Advantages of approach (continued)

- Note that model does **<u>not</u>** necessarily imply convergence across vintages in default risk (conditional on observed loan characteristics)

  - The competing risks setting can generate non-monotonic changes in vintage effects over time. e.g . a decline in mortgage interest rates could lead to increased refinancing (prepayment) – generating negative selection when lower quality loans ineligible for refinancing remain in the sample

  - Dynamics of vintage effects also depends on correlation between unobserved heterogeneity in default and prepayment risks (correlation between v^1 and v^2)

Estimation

- Observed duration and exit type can be interpreted as realizations of random variables T and D defined as $T = \min_{k=1,..,K} T^k$ and $\mathrm{D} = \operatorname{argmin}_{k=1,...K} T^k$ where each independent random variable T^k, k=1,..,K is a latent duration until exit type k in absence of other types of exit risks

- Given discrete bivariate distribution of $v^1$and v^2 the marginal likelihood function is:

$$L(\bar{h}^k, \beta^k, k = 1,..,K; (\mu_j, \pi_j), j = 1, \dots, J) =$$

$$\prod_{i=1}^{N} \sum_{j=1}^{J} \pi_j \ \Pr(T^1 \geq t_i, T^2 \geq t_i | \mu_j)^{c_i} \prod_{k=1}^{2} \Pr(t_i \leq T^k < t_i + 1, T^j > T^k | \mu_j)^{I(d_i=k).(1-c_i)}$$

with $j \neq k$. c_i is the censoring indicator with $c_i = 0$ for a complete uncensored spell and $c_i = 1$ if the duration is right censored at t_i

Estimation

- Observed duration and exit type can be interpreted as realizations of random variables T and D defined as $T = \min_{k=1,..,K} T^k$ and $\mathrm{D} = \operatorname{argmin}_{k=1,...K} T^k$ where each independent random variable T^k, k=1,..,K is a latent duration until exit type k in absence of other types of exit risks

- Given discrete bivariate distribution of $v^1$and v^2 the marginal likelihood function is:

$$L(\bar{h}^k, \beta^k, k = 1,..,K; (\mu_j, \pi_j), j = 1, ..., J) =$$

$$\prod_{i=1}^{N} \sum_{j=1}^{J} \pi_j \ \Pr(T^1 \geq t_i, T^2 \geq t_i | \mu_j)^{c_i} \prod_{k=1}^{2} \Pr(t_i \leq T^k < t_i + 1, T^j > T^k | \mu_j)^{I(d_i = k).(1-c_i)}$$

$$= \Pr(T^1 \geq t_i) \star \Pr(T^2 \geq t_i) \quad where$$

$$\Pr(T^j \geq t_i) = \exp(-\sum_{s=1}^{t_i} e^{x_i(s-1,\tau_i)'\beta^j + v_i^j} \gamma_j(s)) \text{ with } \gamma_j(s) = \int_{s-1}^{s} \bar{h}^j(u) du$$

with $j \neq k$. c_i is the censoring indicator with $c_i = 0$ for a complete uncensored spell and $c_i = 1$ if the duration is right censored at t_i

Estimation

- Observed duration and exit type can be interpreted as realizations of random variables T and D defined as $T = \min_{k=1,..,K} T^k$ and $\mathrm{D} = \operatorname{argmin}_{k=1,...K} T^k$ where each independent random variable T^k, k=1,..,K is a latent duration until exit type k in absence of other types of exit risks

- Given discrete bivariate distribution of $v^1$and v^2 the marginal likelihood function is:

$$L\left(\bar{h}^k, \beta^k, k = 1,..,K; \left(\mu_j, \pi_j\right), j = 1, \dots, J\right) =$$

$$\prod_{i=1}^{N} \sum_{j=1}^{J} \pi_j \ \Pr(T^1 \geq t_i, T^2 \geq t_i | \mu_j)^{c_i} \prod_{k=1}^{2} \Pr(t_i \leq T^k < t_i + 1, T^j > T^k | \mu_j)^{I(d_i=k).(1-c_i)}$$

$$= [\ 1 - \exp(-e^{x_i(t_i,\tau_i)'\beta^k + v_i^k} \gamma_k(t_i +$$

Estimation

- Observed duration and exit type can be interpreted as realizations of random variables T and D defined as $T = \min_{k=1,..,K} T^k$ and $\mathrm{D} = \operatorname{argmin}_{k=1,\dots,K} T^k$ where each independent random variable T^k, k=1,..,K is a latent duration until exit type k in absence of other types of exit risks

- Given discrete bivariate distribution of $v^1$and v^2 the marginal likelihood function is:

$$L(\bar{h}^k, \beta^k, k = 1, .., K; (\mu_j, \pi_j), j = 1, \dots, J) =$$

$$\prod_{i=1}^{N} \sum_{j=1}^{J} \pi_j \ \Pr(T^1 \geq t_i, T^2 \geq t_i | \mu_j)^{c_i} \prod_{k=1}^{2} \Pr(t_i \leq T^k < t_i + 1, T^j > T^k | \mu_j)^{I(d_i=k).(1-c_i)}$$

with $j \neq k$. c_i is the censoring indicator with $c_i = 0$ for a complete uncensored spell and $c_i = 1$ if the duration is right censored at t_i

Dependent competing risks model estimates – vintage effects through unobserved heterogeneity

Estimate		Default	Pre-payment
Origination Balance ($10,000)		−0.030 (0.027)	0.060 (0.003)
Jumbo Loan		0.305 (0.172)	0.397 (0.042)
Single Family Residence		0.037 (0.083)	−0.072 (0.032)
Non Full Documentation		**1.092** (0.132)	0.117 (0.050)
Unknown Documentation		−0.685 (0.141)	−0.283 (0.060)
Time varying covariates			
LTV		**0.025** (0.001)	**−0.022** (0.001)
Unemployment rate		**0.087** (0.013)	**−0.016** (0.006)
Refinance incentive		0.019 (0.058)	**0.715** (0.029)
Mass point (μ_j)		3.509 (0.148)	1.511 (0.088)
	$\pi_1=\Pr(\mu_1, \mu_2)$	$\pi_2=\Pr(\mu_1,0)$	$\pi_3=\Pr(0, \mu_2)$
2003	0.229	0.003	0.364
2004	0.242	0.001	0.334
2005	0.151	0.082	0.130
2006	0.348	0.003	0.008
2007	0.351	0.011	0.057
2008	0.307	0.017	0.013
2009	0.200	0.002	0.465
2010	0.067	0.002	0.632
2011	0.110	0.003	0.472

Other controls:

Spline in FICO, spline in DTI

Dependent competing risks model estimates – vintage effects through unobserved heterogeneity

Estimate		Default	Pre-payment
Origination Balance ($10,000)		−0.030 (0.027)	0.060 (0.003)
Jumbo Loan		0.305 (0.172)	0.397 (0.042)
Single Family Residence		0.037 (0.083)	−0.072 (0.032)
Non Full Documentation		1.092 (0.132)	0.117 (0.050)
Unknown Documentation		−0.685 (0.141)	−0.283 (0.060)
Time varying covariates			
LTV		0.025 (0.001)	−0.022 (0.001)
Unemployment rate		0.087 (0.013)	−0.016 (0.006)
Refinance incentive		0.019 (0.058)	0.715 (0.029)
Mass point (μ_j)		**3.509** (0.148)	**1.511** (0.088)
	$\pi_1=\Pr(\mu_1, \mu_2)$	$\pi_2=\Pr(\mu_1,0)$	$\pi_3=\Pr(0, \mu_2)$
2003	0.229	0.003	0.364
2004	0.242	0.001	0.334
2005	0.151	0.082	0.130
2006	**0.348**	0.003	**0.008**
2007	**0.351**	0.011	**0.057**
2008	**0.307**	0.017	**0.013**
2009	0.200	0.002	0.465
2010	**0.067**	0.002	0.632
2011	**0.110**	0.003	0.472

Other controls:
Spline in FICO, spline in DTI

Type proportions and correlation varies across vintages

Dependent versus independent risks specification

- LR test rejects independence ($p<0.001$)
- Estimates of covariate coefficients very similar
- But estimates of mixture distribution quite different

	Independent risks			Dependent risks			
Mass points							
default (μ_1)	3.262 (0.154)			3.509 (0.148)			
prepayment(μ_2)	1.155 (0.118)			1.511 (0.088)			
	$\pi_1=\Pr(\mu_1,\mu_2)$	$\pi_2=\Pr(\mu_1,0)$	$\pi_3=\Pr(0,\mu_2)$	$\pi_1=\Pr(\mu_1,\mu_2)$	$\pi_2=\Pr(\mu_1,0)$	$\pi_3=\Pr(0,\mu_2)$	ρ
2003	0.100	0.128	0.340	0.229	0.003	0.364	0.44
2004	0.095	0.137	0.315	0.242	0.001	0.334	0.48
2005	0.004	0.232	0.014	0.151	0.082	0.130	0.45
2006	0.014	0.302	0.030	0.348	0.003	0.008	0.97
2007	0.025	0.305	0.049	0.351	0.011	0.057	0.86
2008	0.001	0.310	0.003	0.307	0.017	0.013	0.93
2009	0.074	0.116	0.317	0.200	0.002	0.465	0.35
2010	0.025	0.031	0.422	0.067	0.002	0.632	0.16
2011	0.032	0.068	0.291	0.110	0.003	0.472	0.28

Model produces marginal vintage effects that change over the duration of the loan

Table 5. Implied Vintage Effects

Average Marginal Vintage Effects on Default

$$\sum_{k=1}^{K} \Pr(v = \mu^k | T^1 > t, T^2 > t) e^v$$

Duration (quarters)

Vintage	**1**	**9**	**17**	**25**	**33**	**41**
2003	8.54	7.54	6.34	5.16	3.85	2.55
2004	8.90	7.96	6.74	5.13	3.64	2.61
2005	8.54	7.73	6.36	4.91	3.75	NA
2006	12.38	10.71	7.40	4.64	2.92	NA
2007	12.74	10.65	7.41	4.37	NA	NA
2008	11.51	9.24	6.33	4.19	NA	NA
2009	7.54	6.62	5.25	NA	NA	NA
2010	3.22	2.92	2.41	NA	NA	NA
2011	4.65	4.03	NA	NA	NA	NA

- *Notes*: Average marginal vintage effects computed for loans still ongoing and current at each duration.
- Based on estimated dependent competing risks model with unobserved heterogeneity, presented in previous table.

Source: LPS

Comparison of performance projections

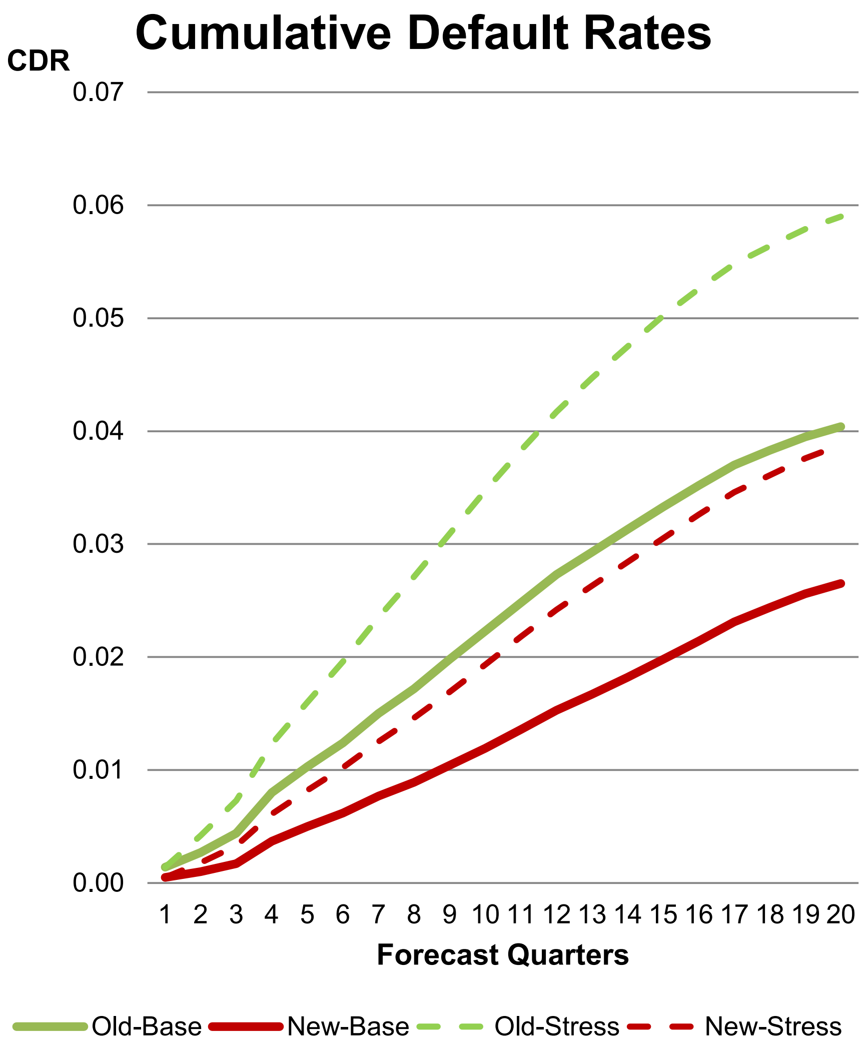
- Conduct series of "stress test" exercises
  - Start with stock of surviving loans at end of 2014
  - Forecast loan performance 5 years into the future
  - Assume value of baseline hazard is constant beyond its within-sample longest-duration level estimated (44 quarters)

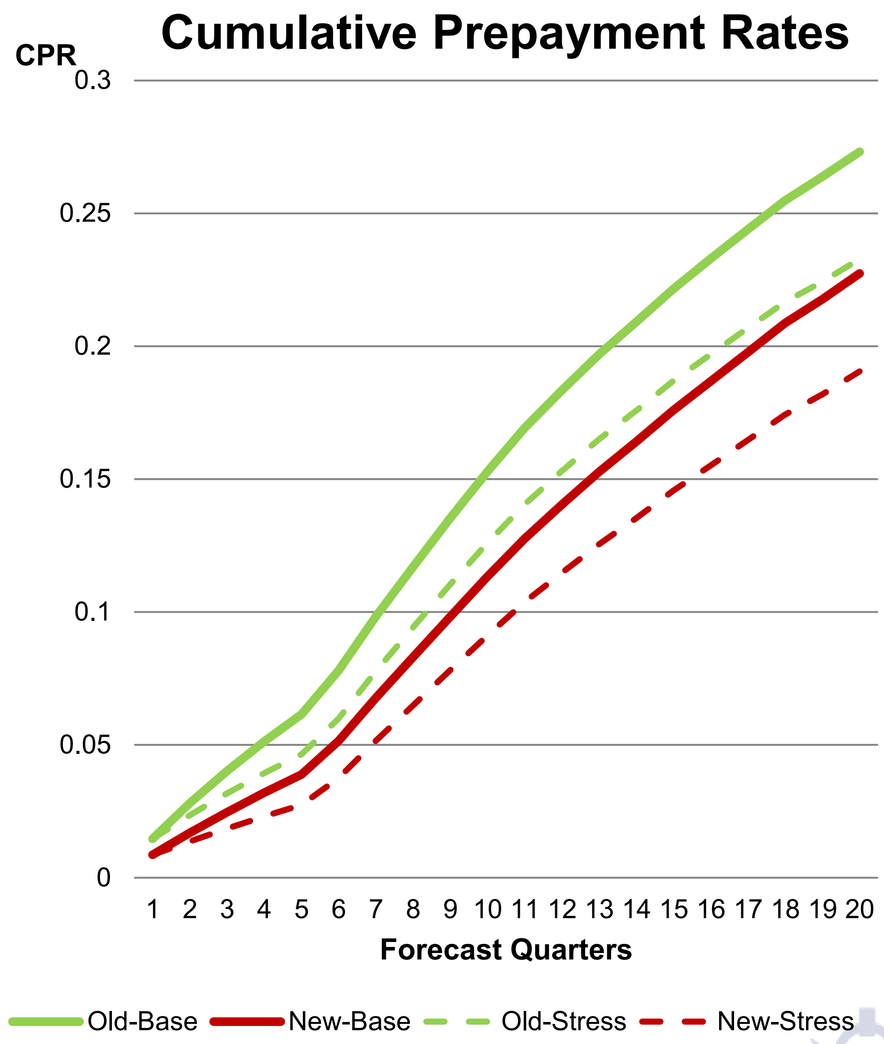
Baseline out-of-sample projection: assume that all time-varying covariates (such as LTV and unemployment rate) remain at their 2014 observed values

Alternative "stress test" scenario: assume a sudden increase beyond 2014 in the local unemployment rate to 10 percent, and a 30 percent increase in LTV (to capture a sharp drop in local home prices)

Comparison of performance projections

Old = conventional model **New = unobserved-heterogeneity model**





Conclusions – next steps

- Alternative model generates lower baseline projections for both default and repayment
- Alternative model shows smaller negative impact of adverse scenario than conventional model

Next steps

- Further investigate non-monotonicity in convergence of average treatment effects – explained by new model?
- Include additional vintages and performance years
- Include non-portfolio loans
- Perform within and out-of-sample performance tests
- Include more mass points
- Explore incorporating unobserved heterogeneity in *sensitivity* of default to covariates (random coefficients)

Reference Slides

Independent competing risk model estimates

Estimate		Default	Pre-payment
Origination Balance ($10,000)		−0.021 (0.028)	0.052 (0.002)
Jumbo Loan		0.082 (0.173)	0.351 (0.039)
Single Family Residence		0.082 (0.082)	−0.048 (0.032)
Non Full Documentation		1.043 (0.133)	0.015 (0.044)
Unknown Documentation		−0.598 (0.141)	−0.162 (0.053)
Time varying covariates			
LTV		0.026 (0.001)	−0.020 (0.001)
Unemployment rate		0.090 (0.013)	−0.013 (0.006)
Refinance incentive		−0.045 (0.063)	0.703 (0.028)
Mass point (μ_j)		3.262 (0.154)	1.155 (0.118)
	$\pi_1=Pr(\mu_1, \mu_2)$	$\pi_2=Pr(\mu_1,0)$	$\pi_3=Pr(0, \mu_2)$
2003	0.100	0.128	0.340
2004	0.095	0.137	0.315
2005	0.004	0.232	0.014
2006	0.014	0.302	0.030
2007	0.025	0.305	0.049
2008	0.001	0.310	0.003
2009	0.074	0.116	0.317
2010	0.025	0.031	0.422
2011	0.032	0.068	0.291

Other controls:
Spline in FICO, spline in DTI

Does dependence between risks matter?