Vintage Effects in Loan Default Models

by Haughwout, Tracy, and van der Klaauw

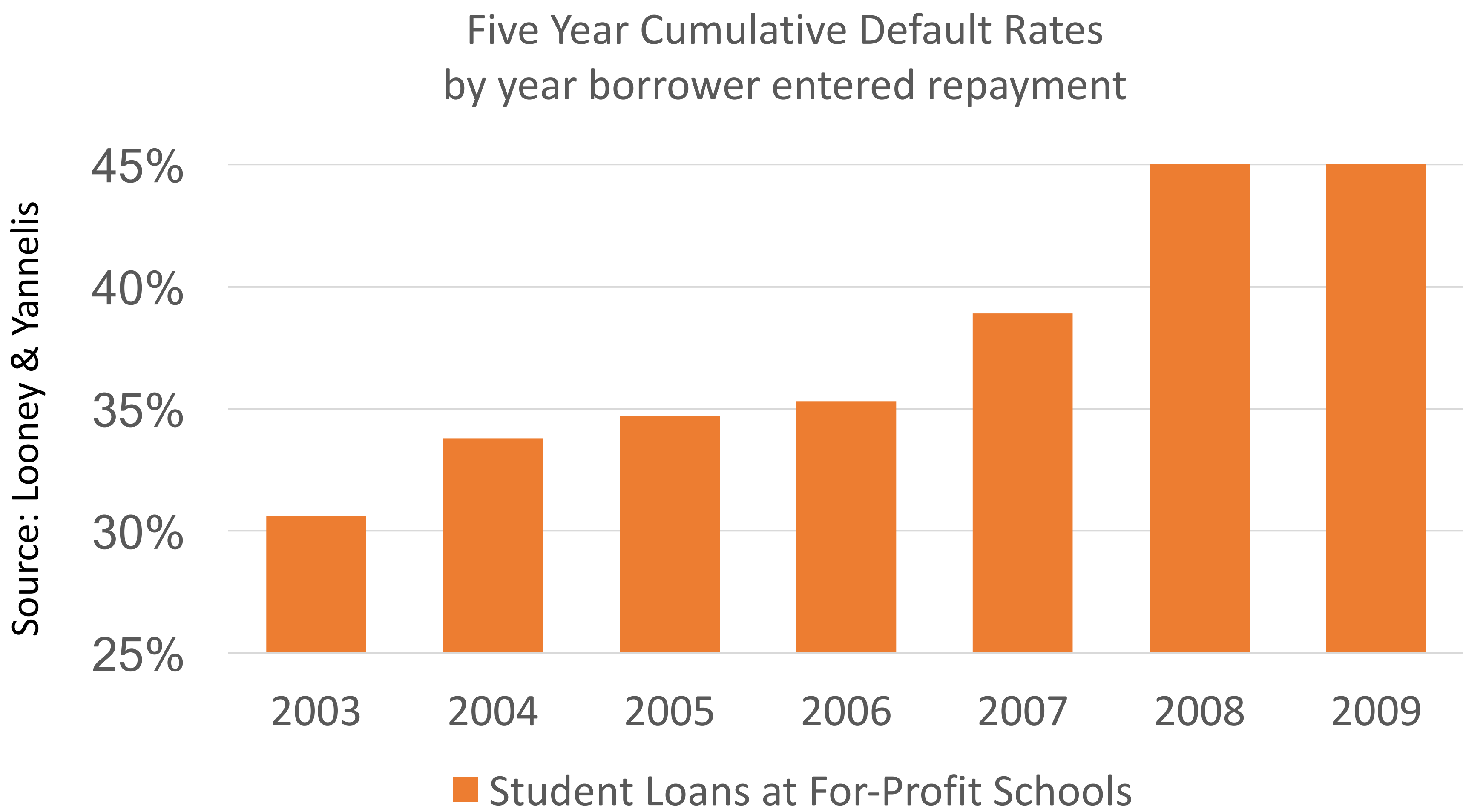
Discussion by Christopher Palmer
MIT Sloan

2017 Stress Testing Research Conference
Federal Reserve Bank of Boston

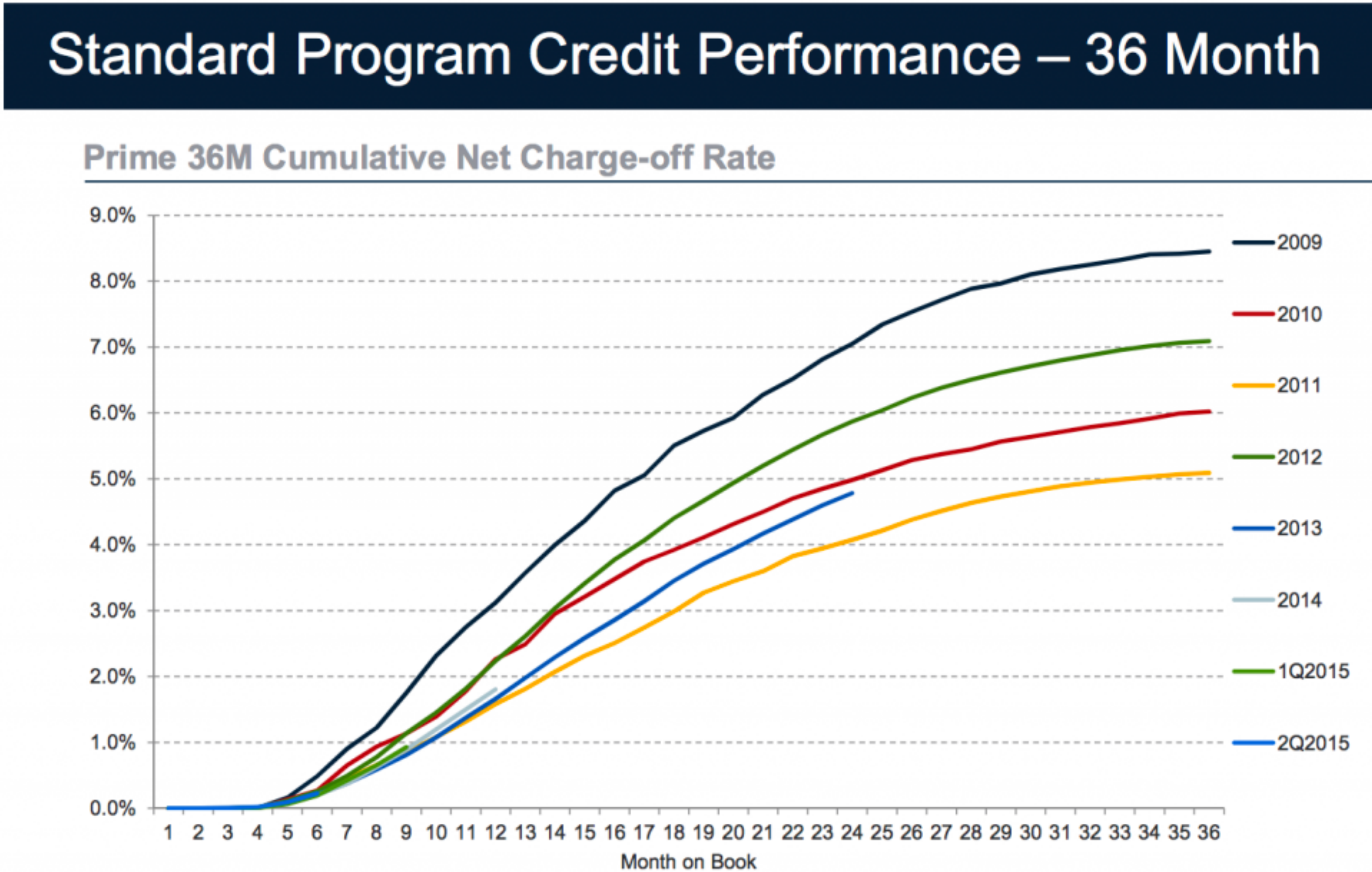
Vintage Effects: Etymology

- Why is vintage a sufficient statistic for wine?
- Idea is that only conditions at bottling matter
- Thereafter, all vintages face same controlled environment
- No interaction between cellar conditions and vintage
- Some usefulness, but analogy breaks down
- Origination regimes drift but…
- Credit markets are not wine cellars
- Clearest example: interaction between equity and house price declines → “cohort” effects

Student Loan Vintage Effects

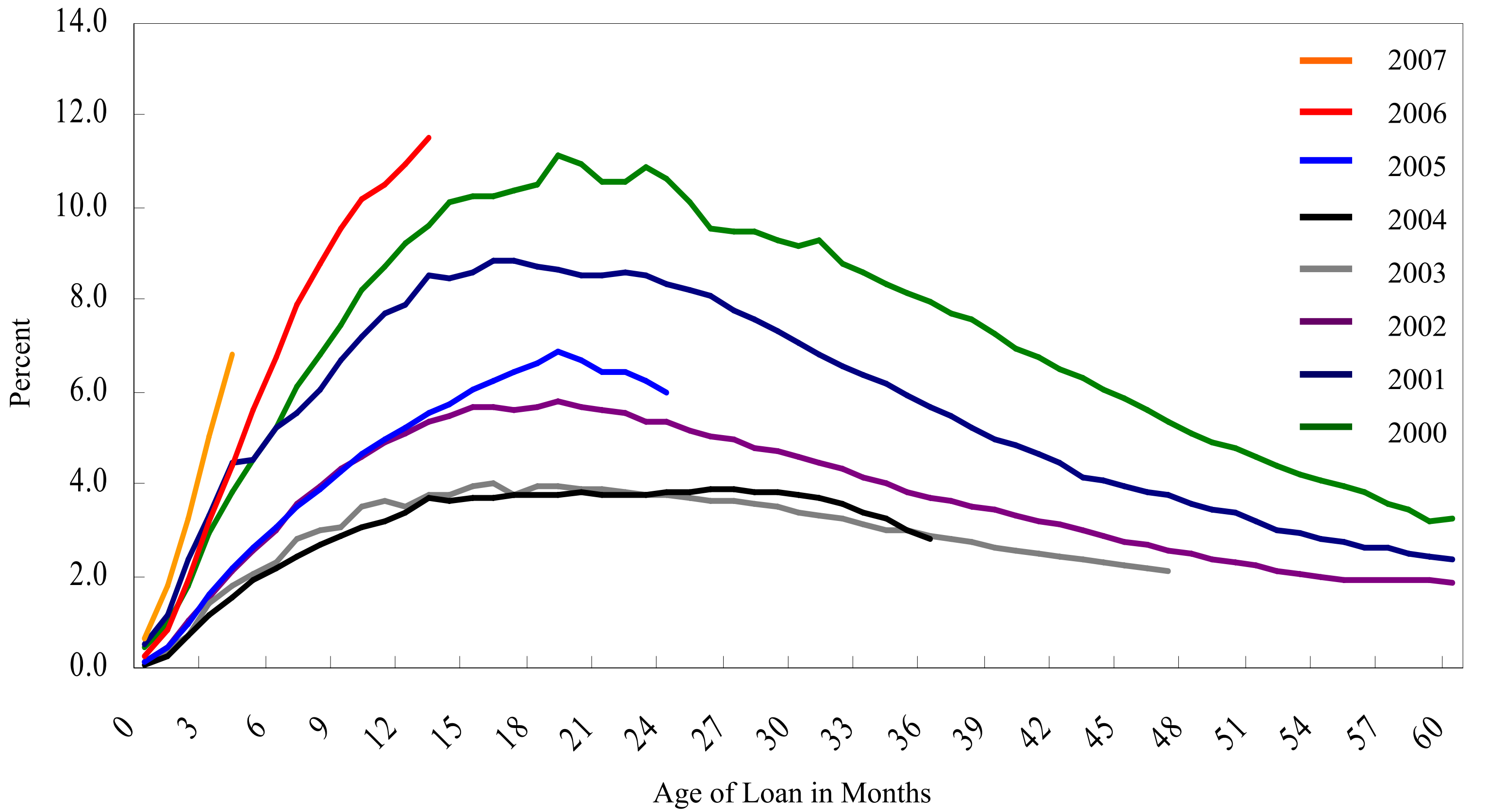


Lending Club Vintage Effects



Subprime Mortgage Vintages

Figure 14: Recent Subprime Vintages Have Performed Poorly
Percent of Loans 60+ Days Delinquent by Year of Origination



Source: Moody's Investors Service

Relevance for Stress Testing

- Want to kick the tires on a portfolio of loans

$$Loss_t = PD_t \times LGD_t \times EAD_t + \xi_t$$

- How can we model changing environment if there are vintage effects?
- Usual strategy: use whatever data we have to estimate fixed vintage effects
- But that doesn't allow for dynamics, making out-of-sample predictions (i.e. stress testing!) tenuous
- Fundamental problem with RF out-of-sample

This paper

- Estimate mixture model with vintage-specific weights in loan hazards
- There are *J* types of mortgages
- Estimated, but we know there is type heterogeneity
  - Woodheads (Campbell et al., 2016)
  - Fastidious (Aiello, 2016)
  - Liars (Haughwout et al., 2011)
- Mass point levels common across vintages but
- Different origination years have different type mix
- Allow default/prepayment types to be correlated

Wine intuition

- Suppose 1979 vintage had 2 (unobservable) types

1. "juvenile delinquents" that burst bottles by 1983
2. perfect wine

- Conventional model would predict in 1983 that 1979 was a uniformly terrible year

- Authors' mixture model would recognize that residual bottles are a great bet

Mortgage Intuition

- Can’t judge a portfolio to be particularly exposed if it’s already “burned out” of all the risky types
  - And vice-versa via prepayment
- Data allows us to estimate share of good and bad types in each vintage with the data we do have
- Allows us to postulate residual distribution of types
- Should improve out-of-sample prediction under alternative scenarios

1. Does risk dependence matter?

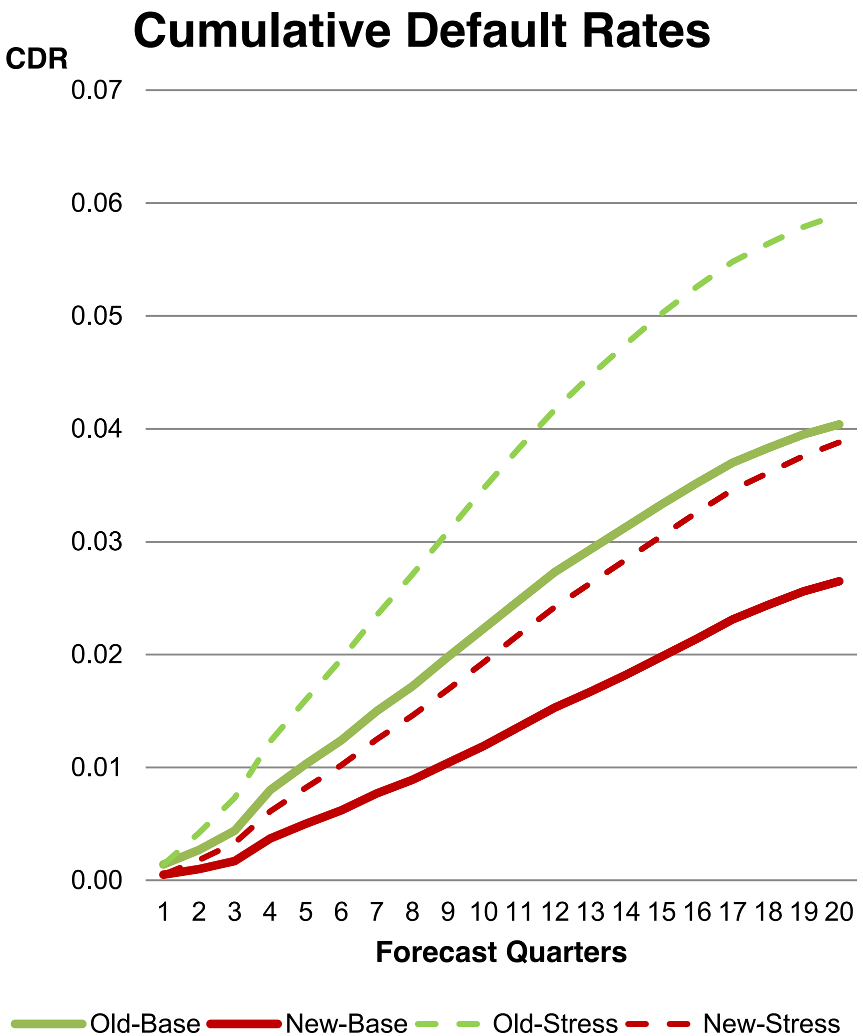
- Given that “estimate of covariate coefficients similar” between independent/dependent risks specs…

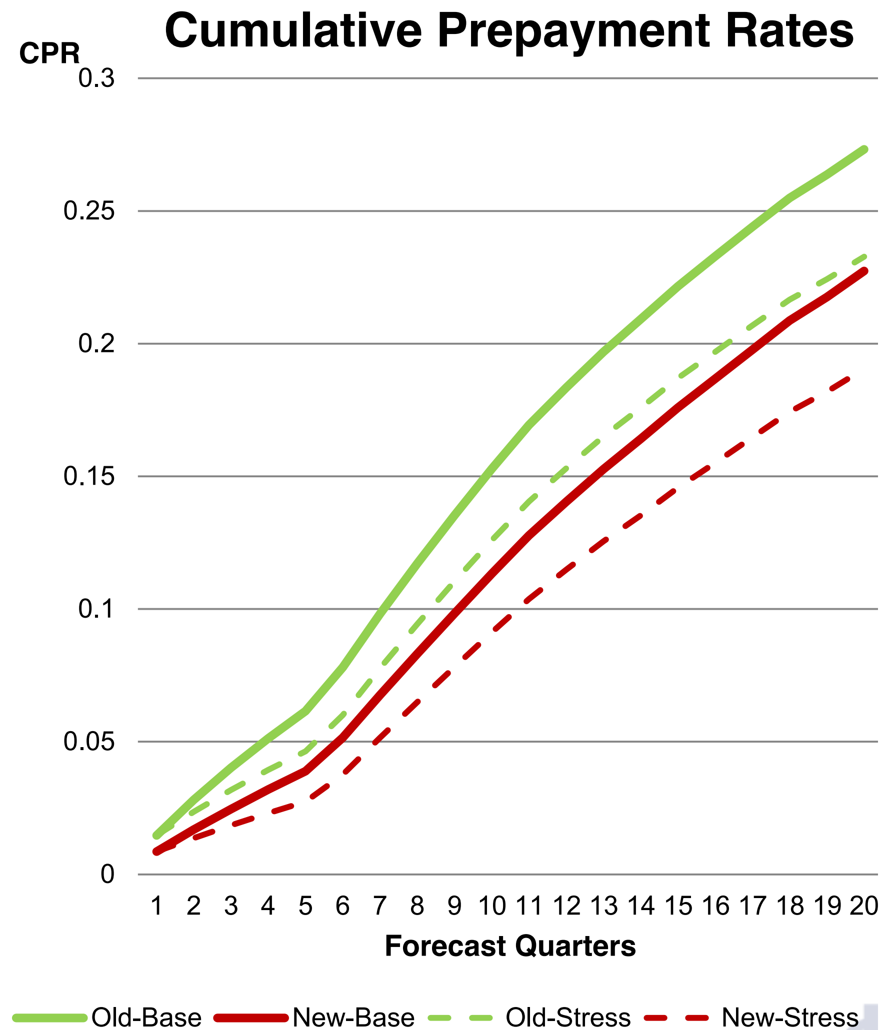
1. Should we care about accounting for dependence?
   - Stress testing: care about exposure to *covariates*, mass points less relevant

2. Why difference in effect of stress?
   - Is there?
   - Baseline model accuracy is empirical question.

Is difference significant?

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

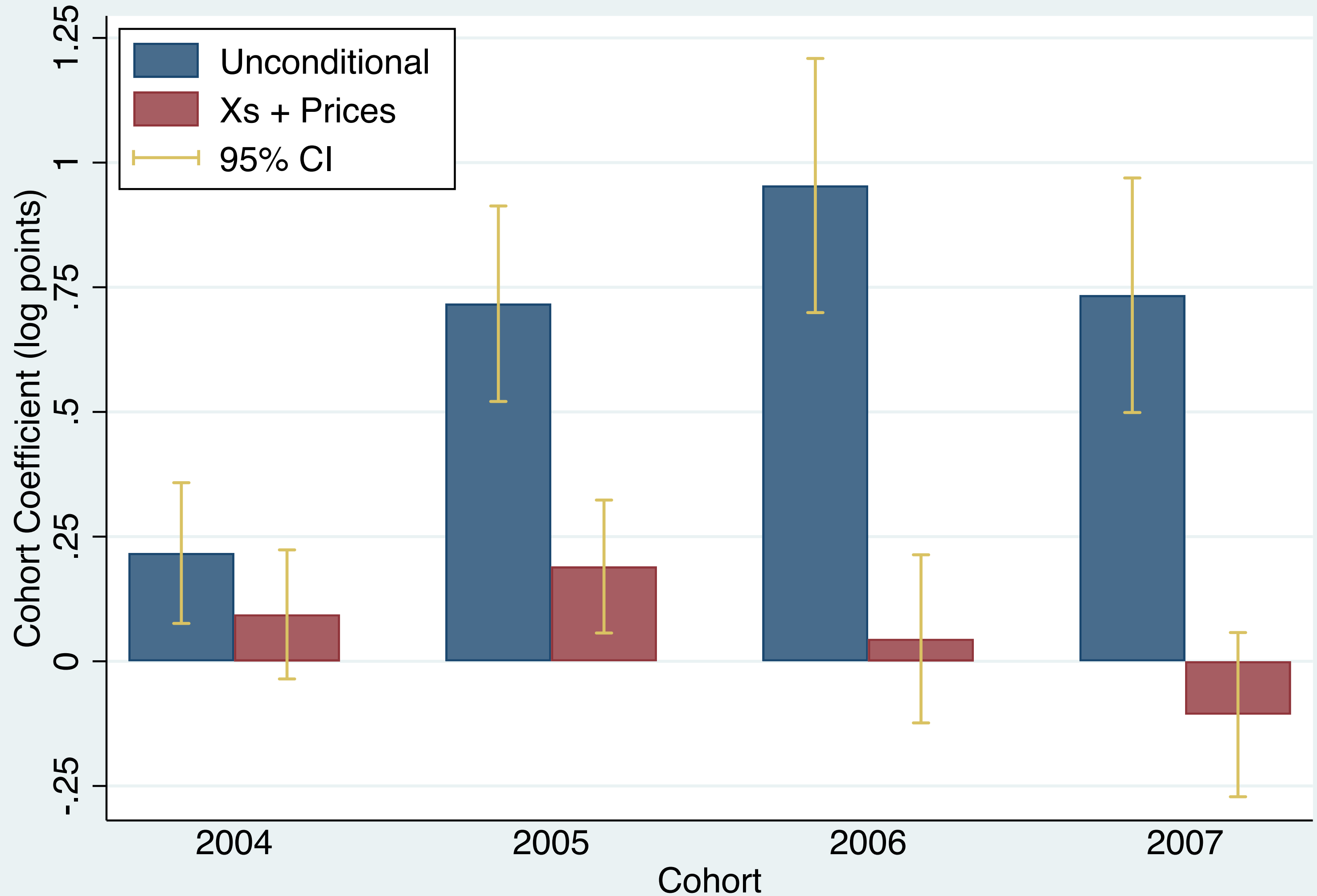




2. Alternative goal: drive $\delta_{\boldsymbol{\tau}k} \rightarrow 0$

- Why not find a spec that **fully** explains vintages?
- If key problems of FEs are dynamic selection, out-of-sample prediction, marginal vs. average, why not use rich enough covariates that explain away entirety of vintage effects?
- Ideal for stress testing!
- Authors cite Demyanyk and van Hemert (2011) as evidence that residual vintage effects cannot be explained with covariates.
- But hazard spec in Palmer (2015) does just this.

Subprime Vintage Effects



Source: Palmer (2015)

3. Simulate out of sample

- Table 5 full of NAs
- Virtue of this method is allows simulation of those
- Could back-test by estimating in 2012 and simulate using 2014 realized covariates
- Compare to 2014 actual realizations key test

Average Marginal Vintage Effects on Default[1]

$$\sum_{k=1}^{K} \Pr(v = \mu^k | T^1 > dur, T^2 > dur) e^{v_1}$$

	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

Little things

- Notation confusing: *j* sometimes indexes unobserved heterogeneity types and an exit type in the same equation.
- Still requires extrapolating out-of-sample (baseline hazard, for example—virtue of parametric λ_0)
- At what level is unemployment measured?
- What level of HPI used to impute time-varying LTV?
- More flexible LTV function? Spline?
- Include other controls like ΔHPI

Conclusion

- Authors develop a very intuitive mixture model
- Captures vintage heterogeneity in a way that captures dynamics while still doubling down on vintage concept
- Could be especially important when evaluating legacy pools
- Why not capture vintage effects fully with Xs?
- Need more evidence that this matters: back-testing simulations