Unobserved Heterogeneity and Its Effect on Mortgage Default and Prepayment Options

Min Hwang, Raphael Kuznetsovski

George Washington University
The US mortgage market is huge and it is dominated by long-term fixed rate loans

Source: Moody's Economy.com
Borrowers always have options to terminate the mortgage contract at any time

- Voluntary Prepayment
- Default
- Contractual Payoff (Full Amortization)

Please do not quote without authors' explicit consent
The theoretical underpinning of rational option exercising has been discussed since late 1980s-early 1990s

• Option pricing models of mortgage consumer behavior
  – Borrowers are rational
    • Compare the value of (discounted) future payments against outstanding mortgage balance
    • Compare the value of collateral against outstanding mortgage balance
  – Both prepayment and default options must be considered jointly:
    • Cannot accurately value the mortgage contract without taking both options into consideration

  – Epperson *et al.* (1985)
  – and others
Exercising options is far from costless

Costs of Prepayment:
- Monetary Costs of New Mortgage
- Lost Productivity
- Inconvenience /Difficulty to Complete Paperwork

Costs of Default:
- Ruined Credit History
- Litigation Costs
- Collections

Please do not quote without authors' explicit consent
Previous studies tried introducing costs heterogeneity into mortgage valuation models

- Stanton (1995)
  - Structural option-based prepayment model
  - Incorporates three exogenous components
    - Heterogeneous transaction costs
    - Random discrete time intervals at which prepayment decisions are evaluated
    - Random process of forced prepayments (“housing turnover”)

- However,
  - Default options were not considered in Stanton’s model
  - Estimated heterogeneity of transaction costs was a distribution with mean value of 41% of the loan balance
    - Too high to be plausible
Deng, Quigley, Van Order (DQV, 2000 – *Econometrica*)

- Reduced-form early termination model
- Introduced unobserved heterogeneity under competing risk survival framework to mortgages
  - Two hazards – *prepayment* and *default* – are dependent competing risks
  - Estimated jointly (using MLE)
  - Unobserved heterogeneity was measured through discretely distributed mass point mixed hazard
  - Considered cases of m=2 and m=3
- Identified borrowers with high propensity to prepay
- Still, DQV and other models left many questions about suboptimal option exercising unanswered

Please do not quote without authors' explicit consent
Why do borrowers fail to take advantage of refinance opportunities?

- Laziness?
- Lack of Education?
- Large Personal Transaction Costs?
- Deterioration of Borrower Credit-worthiness?
- Property Valuation Measurement Error?
- Over-leveraged?

“Rational”

Very Difficult to Control

More Difficult to Control

Can Control (given the data)
Whatever the cause, an approach incorporating unobserved heterogeneity is important to model real behavior.
Joint competing risks hazard mortgage termination model with unobserved heterogeneity can be estimated using MLE

- The unconditional survivor function:

\[
S(t_p, t_d | X) = \int \exp \left( -\eta_p \sum_{k=1}^{t_p} e^{(\gamma_{pk} + \beta_p X)} - \eta_d \sum_{k=1}^{t_d} e^{(\gamma_{dk} + \beta_d X)} \right) dG(\eta_p, \eta_d)
\]

- Log-likelihood function:

\[
\ln(L) = \sum_{j=1}^{N} \left[ \delta_j^p \ln \left( \int \{ S(t-1, t-1 | X_j, \eta_p, \eta_d) - S(t-1, t | X_j, \eta_p, \eta_d) \} dG(\eta_p, \eta_d) \right) \\
- A(t | X_j, \eta_p, \eta_d) dG(\eta_p, \eta_d) \\
+ \delta_j^d \ln \left( \int \{ S(t-1, t-1 | X_j, \eta_p, \eta_d) - S(t, t-1 | X_j, \eta_p, \eta_d) \} dG(\eta_p, \eta_d) \right) \\
- A(t | X_j, \eta_p, \eta_d) dG(\eta_p, \eta_d) \\
+ \delta_j^s \ln \left( \int \{ S(t-1, t-1 | X_j, \eta_p, \eta_d) \} dG(\eta_p, \eta_d) \right) \right]
\]

- Computationally burdensome for practical purposes
  - Trick is to assume discrete mass-point distribution for \((\eta_p, \eta_d)\)
We had access to rich loan-level data on which to empirically test new JCRH models

- First-lien mortgages originated from 1999 through 2008
  - Large national mortgage lender
  - Mix of prime and Alt-A type FRM30 products
The observation period covers full business cycle, several refi booms, and the Great Recession.
Such rich loan-level data on mortgage performance allowed expanding research in new directions

<table>
<thead>
<tr>
<th>Borrower-Level</th>
<th>Loan-level (Static)</th>
<th>Loan-level (Time-varying)</th>
<th>Macro-economic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origination FICO</td>
<td>Loan-to-Value</td>
<td>Updated balance</td>
<td>Unemployment rate (MSA-level)</td>
</tr>
<tr>
<td>Borrower Age</td>
<td>Debt-to-income</td>
<td>Current equity</td>
<td>HPI (MSA-level)</td>
</tr>
<tr>
<td>Family Status</td>
<td>Loan term</td>
<td>Refinance incentive</td>
<td>Underwriting index (US-level)</td>
</tr>
<tr>
<td>Employment</td>
<td>Loan purpose</td>
<td>Updated FICO</td>
<td></td>
</tr>
<tr>
<td>Years at Current Residence</td>
<td>Documentation level</td>
<td>Bankruptcy indicator</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Occupancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Property location</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Number of borrowers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Loan note rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Points paid</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Some of the covariates have been absent from prior research on JCRH framework
Results: Adding asymmetric information helps reduce unobserved heterogeneity

• Some borrowers did not refinance because they were no longer considered credit-worthy
  – For every 50 points drop in FICO:
    • Prob(prepayment) decreases by 17%
    • At the same time, Prob(default) more than doubles

• Borrower’s age, family status, current residency track also found to explain mortgage terminations

• For model with 2 mass points, the distance between “fast” prepayers and “slow” prepayers shrinks by 15% (3% in DQV)

Please do not quote without authors' explicit consent
Results: By paying origination points borrowers send a signal that they plan to stay in the house for a long time

- With each extra point paid, prepayment rate drops by about 14%
Results: Increasing the number of mass points provides new insights about possible unobserved heterogeneity distribution

- The higher the $m$, the more the distribution of $\eta_p$ resembles “humped-shaped” distribution (normal? lognormal?)
- However, distribution of $\eta_d$ remains difficult to parameterize
- Highlights the challenges one might face when trying to impose parametric assumptions on $(\eta_p, \eta_d)$

Please do not quote without authors’ explicit consent
Results: Estimated correlation between prepayment and default options

- ... is positive and economically significant

<table>
<thead>
<tr>
<th>Number of mass points</th>
<th>Estimated correlation between $\eta_p$ and $\eta_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.43362</td>
</tr>
<tr>
<td>4</td>
<td>0.47813</td>
</tr>
<tr>
<td>5</td>
<td>0.42446</td>
</tr>
</tbody>
</table>
Competing risks survival model is a good candidate for PD model under Basel II Advanced-IRB approach

• PD is probability of exposure default in the next 12 months

• Competing risks survival approach has several advantages over simpler Basel PD models (12 mos. cumulative logit, regression tree models):
  – Explicitly accounts not only for default but also for prepayment
  – Accounts for correlation between default and prepayment
  – Allows explicit alignment between loss forecasting, (internal) economic capital, and regulatory capital (Basel II) frameworks (Basel II “use test”)

Please do not quote without authors’ explicit consent
However, not accounting for unobserved heterogeneity can lead to mis-estimation of baseline hazard.

- Unobserved heterogeneity can bias the duration dependence downward
  - Bias exists even if the unobserved heterogeneity is uncorrelated with observed variables
  - If unobserved risk factors are correlated with fixed covariates included in the model, there could be spurious time-covariate interactions
  - Obtaining better model specification by including additional covariates can help mitigate the problem

Please do not quote without authors' explicit consent
Distribution of empirical 12-mos PD rate is skewed and has heavy right tail

Mean=41bps   Median=33bps   Std=24bps
Results: Increasing number of mass points (groups) could significantly impact Basel capital ratios

\[
\Phi\left(\frac{\Phi^{-1}(PD) + \Phi^{-1}(0.999)\sqrt{\rho(PD)}}{\sqrt{1 - \rho(PD)}}\right) \cdot LGD \quad - \quad PD \cdot LGD
\]
Conclusions

• Additional borrower-specific and time-varying information helps reduce unobserved heterogeneity among mortgage holders
• Suggested a way for how the distribution of unobserved heterogeneity can be discretely approximated by increasing number of mass points in the joint competing risk hazard framework
• Found positive correlation between prepayment and default mortgage options
• Proper estimation of unobserved heterogeneity could impact calculation of minimum regulatory capital