Refinancing Frictions, Mortgage Pricing and Redistribution

David Berger (Duke)
Konstantin Milbradt (Northwestern University)
Fabrice Tourre (Copenhagen Business School)
Joseph Vavra (University of Chicago)

Federal Reserve Bank of Philadelphia, May 24, 2023

What do US households know about mortgages?

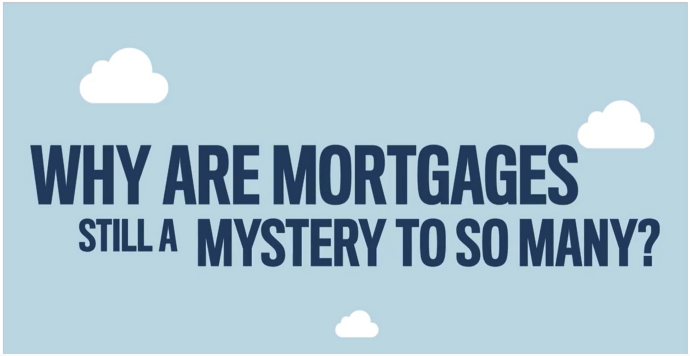
SECTIONS

NEW YORK POST

More than half of Americans don't know what a mortgage actually is: poll

By Brittany Miller

August 11, 2022 | 4:10pm | Updated



Research questions and objectives

- Growing body of work showing that
  - households refinance fixed-rate mortgages sub-optimally
  - large cross-sectional differences in degree of sub-optimality

Research questions and objectives

- Growing body of work showing that
  - households refinance fixed-rate mortgages sub-optimally
  - large cross-sectional differences in degree of sub-optimality

- Mortgage institutional settings limit ability for lenders to price-discriminate
  - US (originators' use of TBA market, legal barriers)
  - Denmark (covered bond market funding model)

Research questions and objectives

- Growing body of work showing that
  - households refinance fixed-rate mortgages sub-optimally
  - large cross-sectional differences in degree of sub-optimality

- Mortgage institutional settings limit ability for lenders to price-discriminate
  - US (originators' use of TBA market, legal barriers)
  - Denmark (covered bond market funding model)

- This paper: consequences for equilibrium mortgage pricing
  - Cross-subsidies from mortgage market pooling
  - Distributional effects of
    - new contract designs: auto-refinancing mortgages
    - recent origination trends: rise of non-bank and fintech lenders

Related Literature

- Mistakes in households' refinancing decision, cross-sectional heterogeneity
  - Campbell (2006), Agarwal, Driscoll, Laibson (2013); Keys, Pope, Pope (2016), Agarwal, Rosen, Yao (2016), Gerardi, Willen, Zhang (2020), Anderson, Campbell, Nielsen, Ramadorai (2020), Amromin, Huang, Sialm, Zhong (2018)
- Mortgages and redistribution
  - Fisher, Gavazza, Liu, Ramadorai, Tripathy (2021), Zhang (2022)
- MBS pricing
  - Stanton (1995), McConnell, Singh (1994), Stanton, Wallace (1998)
- Mortgage refinancing and macroeconomic outcomes
  - Berger, Milbradt, Tourre, Vavra (2021), Guren, Krishnamurthy, McQuade (2021), Eichenbaum, Rebelow, Wang (2022)
- Inattention and redistribution in other fields
  - Gabaix, Laibson (2006), Gottlieb, Smetters (2021)

Preview of results

- Tractable mortgage pricing model capturing cross-sectional differences in attention
  - Analytic characterization of household behavior in special cases
  - Existence/uniqueness/comparative statics of MPE in various settings
  - Influence (or lack thereof) of fixed upfront costs for household behavior and MPE

Preview of results

- Tractable mortgage pricing model capturing cross-sectional differences in attention
  - Analytic characterization of household behavior in special cases
  - Existence/uniqueness/comparative statics of MPE in various settings
  - Influence (or lack thereof) of fixed upfront costs for household behavior and MPE
- Document substantial heterogeneity in attention rates in US borrower population

Preview of results

- Tractable mortgage pricing model capturing cross-sectional differences in attention
  - Analytic characterization of household behavior in special cases
  - Existence/uniqueness/comparative statics of MPE in various settings
  - Influence (or lack thereof) of fixed upfront costs for household behavior and MPE
- Document substantial heterogeneity in attention rates in US borrower population
- Calibration suggests significant cross-subsidies between different household types:
  - fastest HHs pay **60bps** p.a. less than avg. HH, and **150bps** less than in separating eqm.
  - slowest HHs pay **30bps** p.a. more than avg. HH, and **60bps** more than in separating eqm.
  - substantial redistribution via *equilibrium* effects

Preview of results

- Tractable mortgage pricing model capturing cross-sectional differences in attention
  - Analytic characterization of household behavior in special cases
  - Existence/uniqueness/comparative statics of MPE in various settings
  - Influence (or lack thereof) of fixed upfront costs for household behavior and MPE
- Document substantial heterogeneity in attention rates in US borrower population
- Calibration suggests significant cross-subsidies between different household types:
  - fastest HHs pay **60bps** p.a. less than avg. HH, and **150bps** less than in separating eqm.
  - slowest HHs pay **30bps** p.a. more than avg. HH, and **60bps** more than in separating eqm.
  - substantial redistribution via *equilibrium* effects
- Policy interventions, contractual designs, origination trends
  - auto-refi mortgages:
    - increase in mortgage rates ($\sim$ **100bps** p.a.);
    - **16%** of borrowers forced to select smaller homes/higher down-payments;
    - decreases avg. coupons for slow borrowers, but less than in PEQ
  - rise of non-bank lending: **35bps** p.a. increase in equilibrium mortgage rates

Bigger picture

- Modeling framework to study environments with
  - economic agents making dynamic discrete choices about entering long term (non-state contingent) contract subject to (ex-ante) heterogeneous frictions
  - other side of the market that is competitive but cannot price-discriminate

Bigger picture

- Modeling framework to study environments with
  - economic agents making dynamic discrete choices about entering long term (non-state contingent) contract subject to (ex-ante) heterogeneous frictions
  - other side of the market that is competitive but cannot price-discriminate

- Potential applications
  - Labor market
  - Business loan credit markets

Road map

Mortgage refinancing decisions

Markov perfect equilibrium with homogeneous households

Heterogeneous Households

Quantitative implications

Conclusion

Appendices

Mortgage refinancing in partial equilibrium

Households

- continuous time
- continuum of measure 1, risk-neutral, discount rate ρ
- \$1 notional fixed rate mortgage, coupon c_t
- *non-strategic* prepayment at rate ν (move, amortization, divorce)
- option to refinance at any time
- frictions
  - make decisions at discrete points in times arriving with **attention rate** χ
  - upon refinancing, pay **upfront closing cost** ψ

Mortgage rates $m_t = m(x_t)$

- x_t (diffusion and driving process) latent state for term structure model $r(x_t)$
- mortgage rate $m(x_t)$ taken as given by households (for now, exogenous)
- at origination or refinancing time τ, $dc_\tau = m(x_\tau) - c_{\tau-}$

Optimal household behavior in partial equilibrium

Household life-time cost function

$$V(x, c) := \inf_{a_t \in \{0,1\}} \mathbb{E}_{x,c}\left[\int_0^{+\infty} e^{-\rho t}\left(c_t^{(a)} dt + a_t \psi dN_t^{(\chi)}\right)\right],$$

$$\text{s.t.} \qquad dc_t^{(a)} = \left(m(x_t) - c_{t-}^{(a)}\right)\left(a_t dN_t^{(\chi)} + dN_t^{(\nu)}\right)$$

Properties of cost function

- $V(x, c)$ smooth in x, increasing in current coupon c
- refinance when **rate gap** $c - m(x) \geq \theta(x)$
- optimal rate gap threshold $\theta(x)$ satisfies

$$V(x, m(x)) + \psi = V(x, m(x) + \theta(x))$$

Optimal household behavior: special cases with analytical solutions

No upfront closing costs

- When $\psi = 0$, optimal rate gap threshold $\theta(x) = 0$

Mortgage rates following random walk

- assume $m_t = \sigma B_t$
- setup of Agarwal, Driscoll and Laibson 2013 ("ADL"), but with inattention
- optimal rate gap threshold $\theta(x) = \theta$, solution to non-linear equation
- closed form Taylor series approximation $\hat{\theta}$
- θ is increasing in attention rate χ
  - conditional on paying attention, households refinance more aggressively
  - limited attention reduces importance of fixed costs for refi decisions
- [illustration]

Road map

Mortgage refinancing decisions

Markov perfect equilibrium with homogeneous households

Heterogeneous Households

Quantitative implications

Conclusion

Appendices

Mortgage rates in equilibrium

Mortgage investors

- capital elastically supplied by risk-neutral investors ($r_t = r(x_t)$)
- wedge f between coupon paid by households and cash-flow received by investors
  - G-fees, servicing fees
- mortgage origination costs $\psi + \pi$ incurred by lenders recouped via
  - **upfront closing costs** ψ paid by households
  - **gain on sale** π realized in secondary MBS market
- mortgage price $P(x, c)$

$$P(x,c) := \mathbb{E}_x \left[\int_0^{\tau_\theta} e^{-\int_0^t r(x_s)ds} (c - f)\, dt + e^{-\int_0^{\tau_\theta} r(x_s)ds} \right]$$

Equilibrium "break-even" condition at origination

$$P(x, m(x)) = 1 + \pi$$

Equilibrium concept: Markov perfect equilibrium ("MPE")

Mortgage rates with homogeneous households

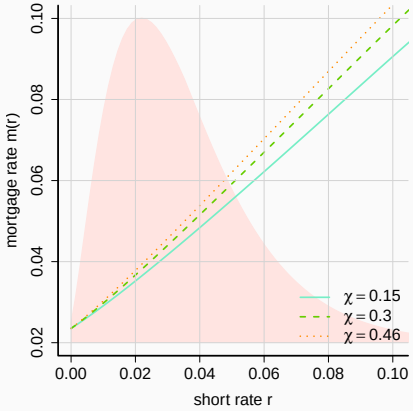
Special case: no upfront closing costs borne by households ($\psi = 0$)

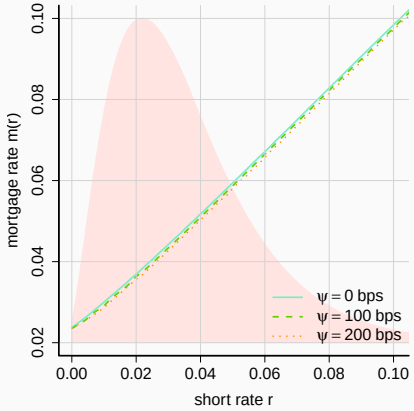
- there exists a unique MPE (subject to technical conditions).
- $m(x)$ increasing in attention rate χ:
  - with higher χ, households exercise prepayment option more "optimally"
  - since investors are short the option, they react by charging more for it

General case: (arbitrary ψ)

- need numerical methods to compute MPE
- MPE not very sensitive to level of upfront closing costs ψ
- intuition:
  - upfront closing costs get large coupon changes upon refi wrong
  - inattention gets small coupon changes upon refi wrong

Upfront closing costs less important for mortgage pricing than inattention





- χ from 50% to 150% of estimated avg attention: $\mathbb{E}[m(x_t)]$ increases 46bps
- ψ from 0% to 200% of empirical avg closing costs: $\mathbb{E}[m(x_t)]$ decreases 14bps

Assume no upfront closing costs

Summarizing results related to upfront closing costs:

1. Inattention significantly dampens refinancing barrier θ (PEQ result)
2. 80% of origination costs financed by lenders via higher rates (Zhang (2022))
3. Fixed costs have limited GE effects in homogeneous model (GE result)

$\Rightarrow$ **going forward, no upfront closing costs born by households:** $\psi = 0, \pi > 0$

In environment with ex-ante HH heterogeneity, it will

- allow us to derive additional theoretical results
- simplify substantially numerical equilibrium calculations

Road map

Mortgage refinancing decisions

Markov perfect equilibrium with homogeneous households

Heterogeneous Households

Quantitative implications

Conclusion

Appendices

Heterogeneity and Pooling

Households are heterogeneous:

- Use CRISM data to estimate heterogeneity in attention intensity (more on this later)
- Find substantial heterogeneity (consistent with other papers)

Hard for conforming mortgage originators in US to price discriminate:

- originate and distribute model and the TBA market [institutional details]
- Fair Credit Lending Act (illegal to price some covariates, even unintentionally)
- regression of mortgage coupon on (a) origination time, (b) FICO and (c) LTV: R^2 of 95%

Mortgage rates with heterogeneous households in GE

Assumption

- cross-sectional distribution $H(\chi)$, and investors cannot screen on χ

Household problem

- without upfront closing costs ($\psi = 0$), refi when paying attention and $m(S_t) \leq c_t$

Mortgage pricing: infinite dimensional problem

- state variable $S_t = (x_t, f_t)$ with $f_t(c, \chi)$ joint density over coupons c and types χ
- *shadow* price $P(S, c; \chi)$ conditional on type χ
- *pool* price $\bar{P}_\phi(S, c) = \mathbb{E}^\phi [P(S_t, c; \chi)]$ for arbitrary pool with attention distribution ϕ

Origination attention distribution $g_t(\chi) = \frac{\int_c \left(\nu + \chi \mathbb{1}_{\{c > m_t\}}\right) f_t(c,\chi) dc}{\int_\chi \int_c \left(\nu + \chi \mathbb{1}_{\{c > m_t\}}\right) f_t(c,\chi) dc d\chi}$

Equilibrium "break-even" condition in Pooling MPE $\bar{P}_{G_t}(S_t, m(S_t)) = 1 + \pi$

Simplifications

Pooling MPE

- Infinite dimensional problem

Simplification: Approximate Pooling MPE

- "bounded rationality" assumption — lenders:
  - know that distribution of refinancers G_t differs from that in the population H, but
  - do not account for the fact that distribution of refinancers depends on the full history of rates
  - price as if cross-sectional origination density is
    - a constant $G(\chi)$, corresponding to erg. avg. origination distribution, or
    - a state-dependent $G(\chi|x)$, equal to conditional erg. avg. origination distribution
- lenders/investors break-even *on average*

Properties of approximate pooling MPE

Existence/uniqueness

- when $r(x_t)$ is one-factor model, there exists a unique (monotone) Approximate Pooling MPE if the "candidate" $m(x)$ is monotone

Properties of approximate pooling MPE

Existence/uniqueness

- when $r(x_t)$ is one-factor model, there exists a unique (monotone) Approximate Pooling MPE if the "candidate" $m(x)$ is monotone

How does cross-sectional heterogeneity in χ affect equilibrium mortgage rates?

- mortgage pool price $\bar{P}_G$ satisfies

$$\bar{P}_G(x,c) = P(x,c;\bar{\chi}_G) - \mathbb{E}_x\left[\int_0^T e^{-\left(\int_0^t r(x_s)ds\right)} \mathbb{1}_{\{m(x_t)\leq c\}} \mathbb{C}\mathrm{ov}^G\left(\chi, P(x_t,c;\chi)\right) dt\right]$$

  - mortgages priced as if homogeneous pool with attention $\bar{\chi}_G$...
  - ... with covariance adjustment that is negative if shadow price declining in χ in expectation
  - ... in which case $m(x;G) \leq m(x;\bar{\chi}_G)$

Properties of approximate pooling MPE

Existence/uniqueness

- when $r(x_t)$ is one-factor model, there exists a unique (monotone) Approximate Pooling MPE if the "candidate" $m(x)$ is monotone

How does cross-sectional heterogeneity in χ affect equilibrium mortgage rates?

- mortgage pool price $\bar{P}_G$ satisfies

$$\bar{P}_G(x,c) = P(x,c;\bar{\chi}_G) - \mathbb{E}_x\left[\int_0^\tau e^{-\left(\int_0^t r(x_s)ds\right)} \mathbb{1}_{\{m(x_t)\leq c\}} \mathbb{C}\mathrm{ov}^G\left(\chi, P(x_t,c;\chi)\right) dt\right]$$

  - mortgages priced as if homogeneous pool with attention $\bar{\chi}_G$...
  - ... with covariance adjustment that is negative if shadow price declining in χ in expectation
  - ... in which case $m(x;G) \leq m(x;\bar{\chi}_G)$
- if $r(x_t)$ is one-factor model, lowest rate $m(x)$ is invariant to origination distribution $G \Rightarrow$ heterogeneity mostly affects mortgage rates when $r(x)$ is high.

Road map

Mortgage refinancing decisions

Markov perfect equilibrium with homogeneous households

Heterogeneous Households

Quantitative implications

Conclusion

Appendices

Measuring inattention in the data

CRISM

- monthly data set following 232k individuals from 2005m6 to 2017m12
- unit of observation: month-household
- observables: monthly refi, cash-out, purchase, FICO, loan balance, LTV (constructed via local price indeces), mortgage rate gap

SFLP (Fannie Mae)

- monthly data set following 250k mortgages from 2000m1 to 2021m12
- unit of observation: month-mortgage
- observables: monthly prepayment, loan balance, mortgage rate gap, initial FICO, initial LTV

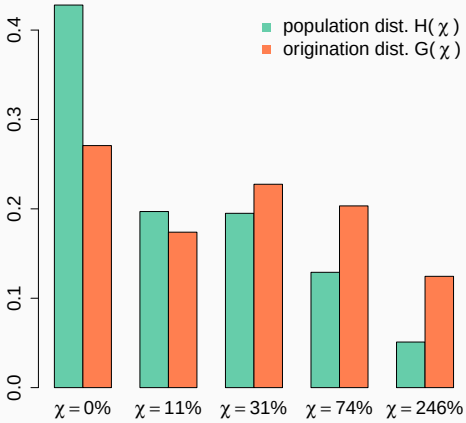
Estimating attention distribution H

Clustering algorithm

- Assume that there is N distinct attention groups in the population
- MLE delivers
  - non-strategic prepay rate ν
  - attention rates $\{\chi_i\}_{i \leq N}$
  - allocation of each household into one group

Group	$\hat{\chi}_i$ (% per year)	$p(\hat{\chi}_i)$ (% per month)	$H(\hat{\chi}_i)$ (%)
base	5.2	0.4	NA
1	0	0.0	40.0
2	11.0	0.9	18.6
3	28.6	2.4	20.3
4	67.2	5.5	15.2
5	240.8	16.8	6.0
avg	30.4	2.4	100.0

Origination distribution $G(\chi)$ [conditional distribution $G(\chi|x)$]



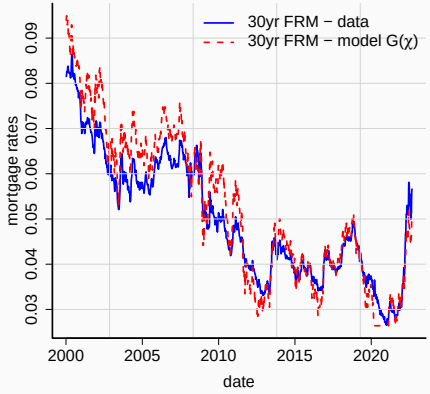
- Average attention rate in the population $\mathbb{E}^H[\chi] = 30\%$
- Average attention rate for newly originated mortgages $\mathbb{E}^G[\chi] = 55\%$ [derivation]

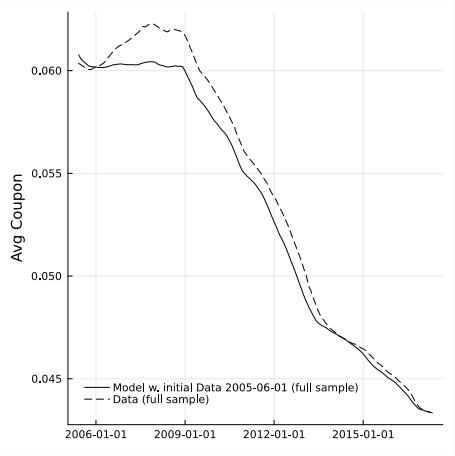
Calibration/estimation of other model parameters

- non-strategic prepay rate (from clustering algo) + amortization rate $\rightarrow \nu$
- CIR model of short rates (MLE) $\rightarrow (\mu, \kappa, \sigma)$
- G-fee as wedge between mortgage coupon and investor cash flows $\rightarrow f$
- mortgage origination costs $\psi + \pi$ from Zhang (2022)
- upfront closing costs rolled into higher rates $\rightarrow \psi = 0$

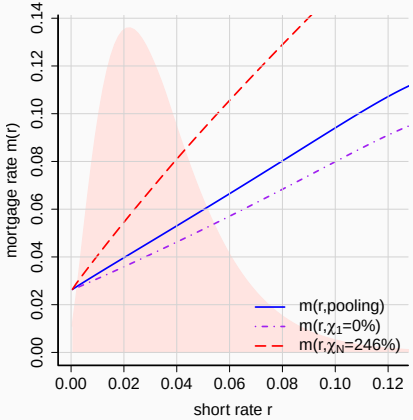
Parameter	Value	Interpretation
μ	3.5%	long run mean short rate
κ	0.13	speed of mean reversion
σ	6%	volatility
ν	10.7%	non-strategic prepay + amortization rate
f	0.45%	G-fees
π	3.7%	gain on sale
ψ	0%	upfront closing costs (rolled into higher rates)

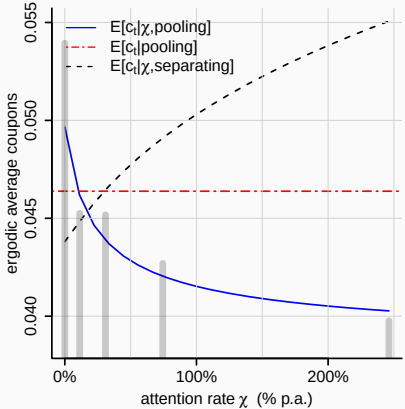
Model validation





Equilibrium mortgage rates and redistribution





Redistribution: PE and GE

- Both PE and GE forces important but with significant heterogeneity
  - Most attentive HHs pay 61 bps p.a. less than avg. HH, and 148 bps less than in separating eqm
  - Least attentive HHs pay 33 bps p.a. more than avg. HH, and 59 bps more than in separating eqm

Group	PE (bps p.a.)	GE (bps p.a.)	total (bps p.a.)
1 (least attentive)	+33	+26	+59
2	-2	+15	+13
3	-25	-1	-26
4	-43	-27	-70
5 (most attentive)	-61	-87	-148

Reducing frictions via contract design: automatic refinancing

Households' refinancing mistakes

- Existing literature: welfare gains achievable via automatically refinancing mortgages
- Conclusion often does not factor equilibrium response

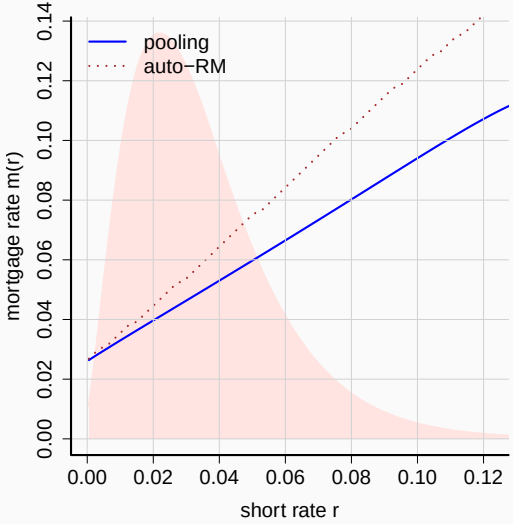
Automatically refinancing mortgages ("Auto-RM")

- Smart contract: no origination costs upon "rate reset", only upon new mortgage
- equivalent to homogeneous MPE with $\chi = +\infty$ (when no gain on sale π)

Properties of Auto-RM contract

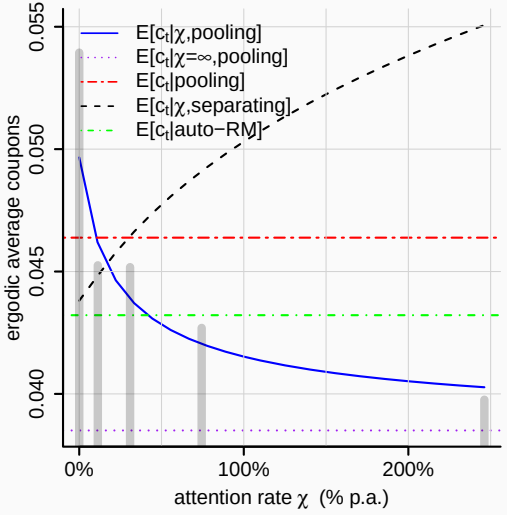
- Auto-RM more costly than ARM: $m(x; \infty) > r(x) + f$ for all latent state x
- "unravelling": if Auto-RM made available, all households eventually migrate to it

Automatically refinancing mortgages: pricing



- Auto-RM leads to significantly higher mortgage rates (**100bps** on average), even though it "saves" on closing costs
- Cross-subsidies disappear
- $\sim$ **16%** of borrowers forced to select smaller homes/smaller initial mortgage balance due to DTI constraints [illustration]

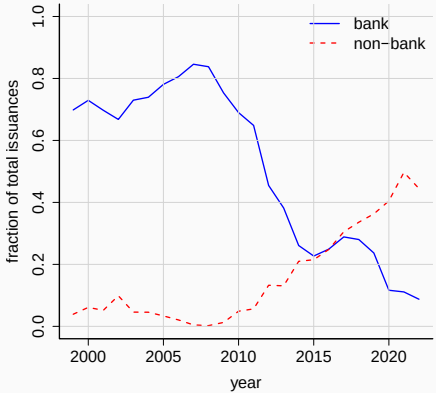
Automatically refinancing mortgages: winners and losers

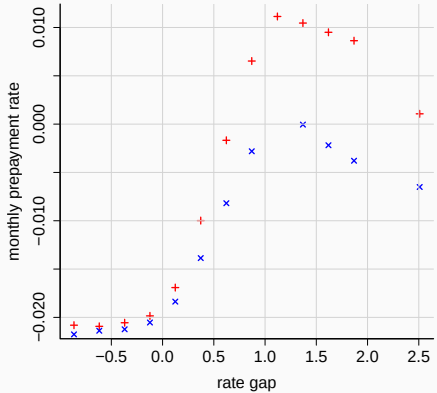


- PE: keep mortgage rates constant, remove all refinancing frictions
  - everyone gains
  - least attentive the most
- GE: adjust mortgage rates to equilibrium auto-RM level
  - least attentive 3 groups gain
  - most attentive 2 groups lose

Group	PE (bps p.a.)	total (bps p.a.)
1 (least attentive)	-112	-64
2	-77	-29
3	-54	-7
4	-35	+12
5 (most attentive)	-18	+29

The rise of non-bank mortgage lenders





- Effective attention rates: $\hat{\chi}_{non-bank} - \hat{\chi}_{bank} \approx 100\text{bps/month}$
- Counterfactual: $\mathbb{E}\left[m_t^{\text{non-bank}}\right] - \mathbb{E}\left[m_t^{\text{bank}}\right] \approx 35bps$

Road map

Mortgage refinancing decisions

Markov perfect equilibrium with homogeneous households

Heterogeneous Households

Quantitative implications

Conclusion

Appendices

Conclusion

- Tractable mortgage pricing model capturing cross-sectional differences in attention
  - Novel results for household behavior and equilibrium
  - Systematic study of inattention friction onto equilibrium outcomes
- Measure distribution of attention in large panel of US borrowers
- Substantial cross-subsidies arising from pooling MPE
  - regressive cross-subsidies
  - direction opposite that from credit guarantee scheme
- Policy analysis and model counterfactuals
  - Financial literacy programs will hurt those that are “untreated”
  - Automatically refinancing mortgages with large equilibrium effects
  - Recent rise of non-bank lending has caused increase in mortgage interest rates
- Other economic application of general modeling framework [Harris-Holmstrom]

Road map

Mortgage refinancing decisions

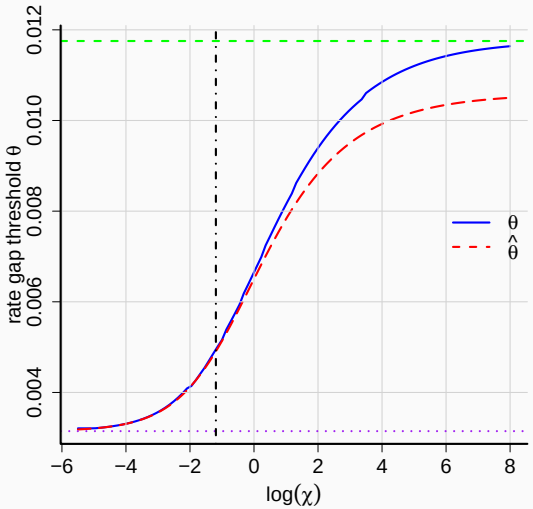
Markov perfect equilibrium with homogeneous households

Heterogeneous Households

Quantitative implications

Conclusion

Appendices

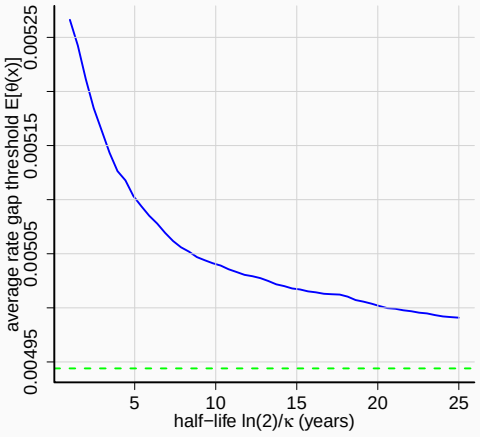


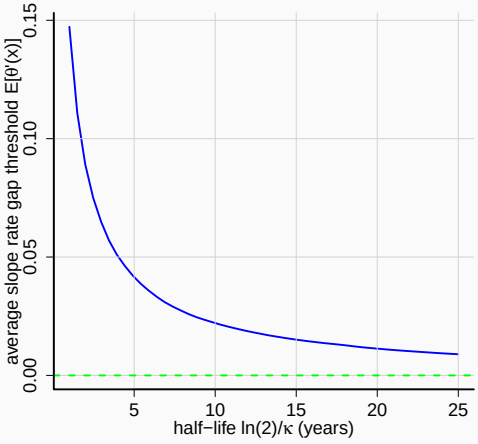
- blue solid (red dash) line: rate gap threshold θ (approximation $\hat{\theta}$)
- data: $\chi \approx 30\%$
- **Key insight:** Upfront costs become less important in the presence of inattention
- Agarwal et al (2016), Fuster et al (2019) revisited:
  - are households refinancing too early?
  - are households refinancing optimally, subject to attention friction?

Mean reverting rates: state dependent policies

[Back]

- mortgage rate follows OU process: $dm_t = \kappa(\bar{m} - m_t)dt + \sigma dB_t$
- optimal rate gap threshold now state-dependent: $\theta(x)$





Institutional details – conforming mortgage market

[Back]

Origination and securitization process

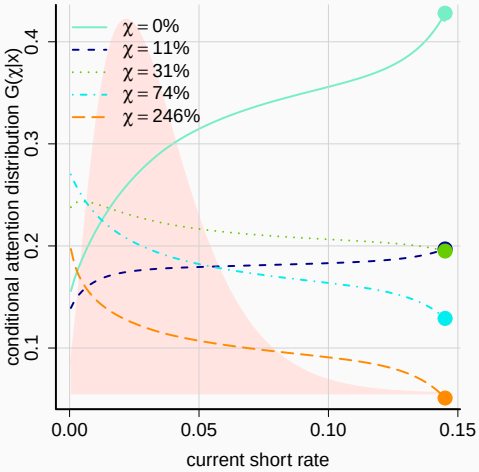
- Originator extends mortgage to borrower
- Originator then hedges
  - rate risk via the to-be-announced ("TBA") market
  - fallout risk via swaptions or other derivatives
- Once a pool of loans is assembled
  - loans delivered to the agency (Fannie or Freddie) vs. MBS security
  - originator delivers MBS into TBA contract
  - originator either retains or sells mortgage servicing right ("MSR")

TBA contract

- effectively, a forward contract
- term (e.g. 30 years), issuer (e.g. Fannie), coupon (e.g. 2.5%, in 50bps increments), settlement month, notional, price
- no specification of exact pool to be delivered $\Rightarrow$ cheapest-to-deliver
- SIFMA "good delivery" guidelines

Conditional origination distribution $G(\chi|x)$

[Back]



Origination distribution G

Intuition

- Suppose everyone's refinancing option is *in the money*, then high χ over-represented

$$g_t(\chi) = \frac{\nu + \chi}{\nu + \bar{\chi}_H} h(\chi)$$

- Suppose everyone's refinancing option is *out of the money*, then "equal" representation

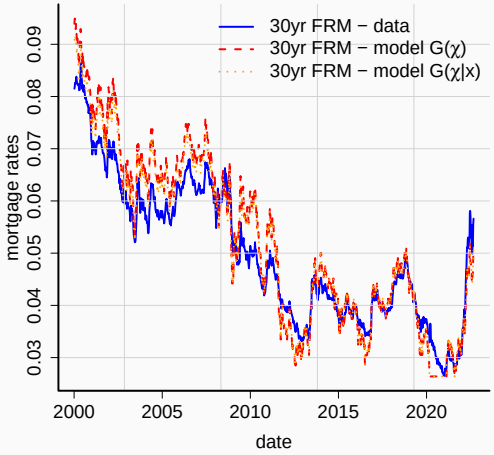
$$g_t(\chi) = h(\chi)$$

Derivation of G

- derive ergodic joint distribution $f_\infty(x, c|\chi)$ given conjecture of price function $m(x, G)$
- derive ergodic conditional (unconditional) origination distribution $G(\chi|x)$ ($G(\chi)$)
- iterate on pricing function $m(x, G) \implies$ in 1D, no iteration needed

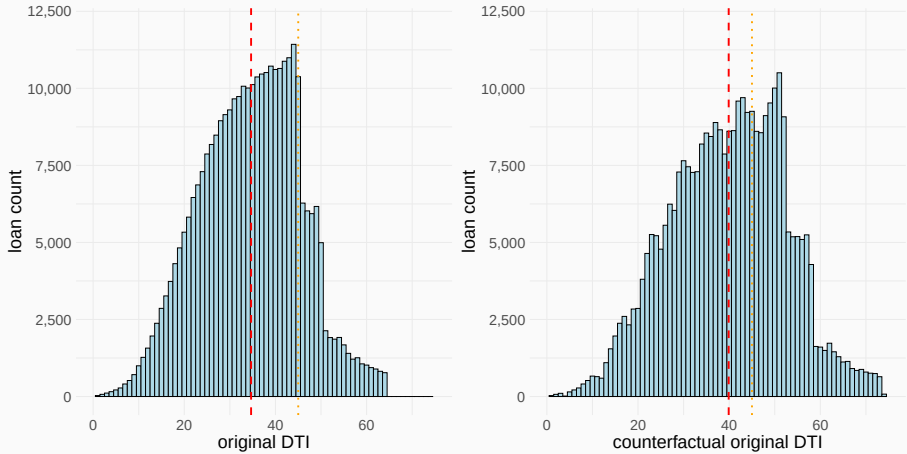
Pricing under conditional origination distribution $G(\chi|x)$

[Back]



DTI AND COUNTERFACTUAL DTI AT ORIGINATION

[BACK]



General framework – labor market application

[Back]

Workers — value function $V(x, w)$

- risk-averse and cannot save
- productivity $x_{i,t}$ (Itô process, idiosyncratic shocks), fixed wage rate w_{it}
- outside offers at rate χ ("loyalty"), cross-sectional distribution $H(\chi)$

Firm — match value $\Pi(x, w; \chi)$ conditional on worker type

- offers competitive wage s.t.
  1. no discrimination (i.e. offer not contingent on χ)
  2. maximal insurance given workers outside offer (i.e. fixed wage contract)
  3. breaks even given distribution of job applicant types $G(\chi)$
- worker-firm pairs separate at unconditional rate ν

MPE — pooling equilibrium wage rate for new workers $\mathcal{W}(x)$

- cross-subsidies between "loyal" and "on-the-job hunting" workers
- at separation, wage losses even without firm-specific human capital