International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(20s) 2026 pp. 1627-1634
https://doi.org/10.70917/ijcisim-2026-6023
Article

Beyond Static Prepayment Curves: Machine Learning and Simulation for Prepayment and Liquidity Risk in Government-Guaranteed Loan Pools

Ankur Mehra

Independent Researcher, New Jersey, United States
Corresponding author: reachankurm@gmail.com

Abstract: However, for the guaranteed portion of a loan, a borrower default is an early return of principal to the investor, making static prepayment curves structurally unfit to price SBA 7(a) guaranteed portions and Ginnie Mae pools of FHA and VA loans. Replacing them with machine learning is necessary, but not sufficient. It shows that loan level prediction contributes to valuation under an accelerated simulation structure, a governance layer for validators to contest. Our three contributions are: (i) a taxonomy of prepayment mechanisms for guaranteed pools, (ii) a comparison mechanism for benchmark static, structural, gradient boosted, deep neural network and survival models via five criteria, and (iii) a reference architecture coupling prepayment prediction with servicer liquidity stress. The timing of the repurchase, and the funding capacity of the servicer, should be modeled jointly, since their effects are related in a way that makes separate modeling misleading.
Keywords: prepayment modeling; government guaranteed loans; machine learning; Monte Carlo simulation; liquidity risk; model risk governance

1. Introduction

Consequently, to the investor holding the guaranteed portion of the loan, default looks like a prepayment rather than a loss. For example, in the event of a default by a borrower on a U.S. SBA 7(a) loan, the guaranteed share is repurchased and the principal is advanced to the lender. The same is true for an FHA or VA loan in a Ginnie Mae pool that becomes seriously delinquent, with the issuer having the option to repurchase the loan out of the Ginnie Mae pool. Credit risk is still present, but is called something else. Any valuation engine which relies on interest rate motivation for cash flow timing is using the process incorrectly.

Compounding the issue is the structure of the servicing industry. As a majority of Ginnie Mae servicing is now performed by non-bank companies, their share of Ginnie Mae loans grew from 34 percent in 2014 to 83 percent in 2023 (Financial Stability Oversight Council [FSOC], 2024). Non-banks pay investors ahead of missed payments and buy out loans from their balance sheets, with the 2020 wave of forbearances showing how quickly servicing obligations can amass (Cherry et al., 2021). The buyout timing, in other words, depends, in part, on the servicer's capacity to pay.

This is a more granular analysis than the normal fixed-income analysis in ML and models at the loan level capture the characteristics of prepayment not observed on a static curve. Even for a held-for-investment book, the daily mark-to-market is not a benchmarking exercise: it has to be stress-tested against thousands of rate paths,

Copyright: © 2026 by the authors

explained to an independent validator and reconciled against the liquidity available to the servicer. Without any of those three layers, a more advanced model adds little in the way of value to the pool.

There are three contributions to this.

1. A taxonomy. A classification of prepayment mechanisms that occur specifically because of government guarantees (Section 3).
2. An evaluation. An ordinal evaluation of five families of models based on their production value, rather than out-of-sample accuracy (Section 4.1).
3. An architecture. A reference design coupling together a competing-risk prepayment model and a servicer liquidity overlay with accelerated Monte Carlo pricing, all within supervisory model-risk controls (Sections 4.4 and 4.5).

2. Background and Related Work

2.1 The representative-borrower inheritance

The two traditions that have shaped the industry's perception of prepayment both assume a representative borrower: the econometric tradition estimates the prepayment hazard from observed pool behavior and passes the fitted hazard into valuation (Schwartz & Torous, 1989), while from this structural view as an embedded option, transaction costs have been used to explain why most mortgages are not refinanced when it is optimal to do so (Stanton, 1995). Arising from this inheritance are the conventions for speed, like the PSA benchmark and CPR assumptions, which function for large pools of homogenously behaving loans at the agency level; the guaranteed pools are rarely similar to each other. An SBA pool can include restaurant franchises, medical practices, different percentages of guarantees, and servicers of very different financial strength.

2.2 What Loan-Level Learning Reveals

Machine learning studies challenge the representative-borrower view on empirical grounds. Prepayment has the strongest nonlinear effects of any loan event the authors model, across networks trained on origination and performance records for over 120 million mortgages (Sadhwani et al., 2021). Loan level gradient boosted classifiers when aggregated upward are as good or better than the modular prepayment models long in use for agency securities (Schultz and Fabozzi, 2021). Liu and Liang (2024)'s predicted performance window goes back to origination, finding different behavior between bank, non-bank and fintech lenders, which matters where there are pools with mixed servicers.

Competing risks fits may be appropriate for guaranteed collateral. Blumenstock et al. (2022) show deep survival networks outperform Cox and Fine-Gray specifications when default and prepayment compete for the same loan. In the guaranteed pool case, these two outcomes are distinct events leading to the same cash flow outcome, and a model that distinguishes between these will explain the underlying mechanism better. Scalable tree-boosting implementations are a key enabler of the economics of building loan tape scale models (Chen & Guestrin, 2016).

Enthusiasm needs a counterweight. Loan-level models inherit the biases present in their training data and the world when they were trained. Research into explainability can help: SHapley Additive exPlanations (SHAP) decompose predictions into feature contributions (Lundberg & Lee, 2017). Tree-based methods (Lundberg et al., 2020) extend to global model monitoring (see data attribution methods and interpretability).

2.3 The simulation bottleneck

However, valuing the pool with embedded prepayment options requires path simulation, and the loan-level model must be evaluated at each time step along each simulated path. Early-exercise valuation is tractable because the continuation values can be regressed on simulated state variables (Longstaff and Schwartz, 2001). Their method remains the auditable reference point. Kahale (2020) proves that multilevel estimators reduce the variance of path-

dependent payoffs. Alternatively, Becker et al. (2020) compute the neural optimal stopping, while Kunsági-Máté et al. (2023) show hybrid neural-weighted Monte Carlo retains the same accuracy at a reduced cost. None of these studies have discussed loan pools directly (this is an inference that the article makes).

2.4 Liquidity where guarantees operate

Liquidity research provides the missing link. Cherry et al. (2021) document the sheer scale of pandemic forbearance, the majority of which was mandated by the government. Kim et al. (2024) find that less liquid non-bank servicers provided less forbearance, suggesting that funding capacity matters for borrower outcomes. Pence (2022) uses this view of the mortgage crisis to directly compare 2008 to 2020 and finds that government heightened market stability. In small business, Bachas et al. (2021) show that lending volume responds to guarantee generosity, and Lee and Lee (2023) show that 7(a) pricing responds weakly to local default risk.

Taken together, these literatures describe one system from three separate vantage points. Nobody has yet drawn the whole picture.

3. Materials and Methods

The study is a synthesis of peer-review evidence and reasoning for building valuation and analytics platforms for loan portfolios, as well as fixed-income analytics for government-guaranteed small-business lending. The study does not include proprietary data. This places restrictions on what is measured, but makes every claim traceable to published sources.

Each model family is scored according to five production criteria:

1. predictive accuracy when loans in a pool differ widely,

2. interpretability to an independent validator;

3. compute the cost per full valuation run, meaning for each loan across every simulated path;

4. data requirements;

5. alignment with supervisory model risk guidance (Board of Governors of the Federal Reserve System [FRB] & Office of the Comptroller of the Currency [OCC], 2011).

Since these scores are ordinal (low, medium, high) and depend on the weight of the cited evidence rather than a pooled statistical estimate, where the studies disagreed, the more conservative score was kept.

One constraint deserves direct statement. Loan-level prepayment histories of the SBA guaranteed portions are not published with anything like the granularity of agency mortgage datasets. This means ML is mostly supported by evidence on residential mortgages, and generalizing from that to small-business guaranteed pools is an inferred generalization, untested by the paper's results. The first taxonomy lumped issuer buyouts together with guarantee purchases. They had to be separated once servicer funding entered the picture, since the buyouts would depend on the cash position of the servicer.

Table 1. Prepayment mechanisms in government-guaranteed pools

Mechanism	**Trigger**	**Investor experience**	**Depends on servicer liquidity?**
Voluntary refinancing	Rate incentive; borrower credit improvement	Standard prepayment	No
Guarantee purchase after default	Borrower default; guarantor pays holder	Early return of principal at par	Weakly

Issuer buyout of delinquent loans	Sustained delinquency in Ginnie Mae pools	Early return of principal	Yes, strongly
Streamline refinancing	Government-program refinance channels	Rate-driven prepayment wave	No
Curtailment or business sale	Excess cash; sale of the borrowing firm	Partial or full early return	No

4. Results and Discussion

4.1 Model families under production constraints

None of the model families have all the merits in Table 2, for example, static curves are cheap, transparent and easy to govern but ignore the variation across loans that is the distinguishing characteristic of guaranteed collateral. Structural models econometrically discipline and are difficult to fit to heterogeneous small-business populations. It finds gradient boosting achieves good predictive performance at moderate cost, and SHAP makes its outputs interpretable (Lundberg et al., 2020). Deep networks capture the richest nonlinearities (Sadhwani et al., 2021), and carry the heaviest validation burden. Deep survival models are structurally most similar because they separate voluntary prepayment from default-driven repurchase (Blumenstock et al., 2022).

The most plausible reading remains a survival model or gradient-boosted competing-risk model as the defensible primary engine and a static curve as governance benchmark running in parallel, not to be discarded.

Table 2. Evaluation of prepayment model families on five production criteria

Family	**Accuracy with diverse loans**	**Interpretability**	**Compute cost per run**	**Data needs**	**Governance fit**
Static curve (PSA/CPR)	Low	High	Low	Low	High
Structural (rational exercise)	Medium	Medium	Medium	Medium	Medium
Gradient boosting	High	Medium (with SHAP)	Low–Medium	High	Medium
Deep neural network	High	Low	Medium–High	Very high	Low–Medium
Deep survival / competing risk	High	Medium	Medium	High	Medium–High

4.2 Location of computation

The cost is relatively high, since the guaranteed pool must be evaluated at every monthly step along every simulated rate path for every loan under the prepayment model. Thus, the number of calls grows multiplicatively along all three dimensions. Daily mark-to-market repeats the full run every business day, but accuracy cannot be computed before the morning profit-and-loss cut-off, and thus does not reach the books.

Table 3 shows the acceleration options classified according to what they change. For computing path-dependent payoffs, multilevel estimators can get the variance down by a factor of order m, while leaving the pricing approach the same (Kahale, 2020). Neural stand-in models (Becker et al., 2020; Kunsági-Máté et al., 2023) replace the inner valuation loops, and share their own validation file and matrices. Elastic cloud compute, including serverless functions and on-demand data science clusters, replaces the vertical scaling of computation time while keeping the mathematics unchanged. But it is the cheapest gain from a governance perspective, and platform teams will fully exhaust it before accepting neural stand-ins.

Table 3. Simulation acceleration options for loan-level valuation

Approach	Source of speed	Accuracy risk	Auditability	Validation burden
Least-squares Monte Carlo	Baseline regression on paths	Low	High	Baseline
Multilevel Monte Carlo	Variance reduction across levels	Low	High	Low
Neural stand-in model	Replaces inner valuation	Medium	Medium	High
Elastic cloud compute	Parallel path execution	None	High	Minimal

4.3 Why liquidity belongs inside the prepayment model

This subsection describes the article's finding that Ginnie Mae servicers must pay principal and interest on defaults, then buy out the delinquent loans from pools (FSOC, 2024). During the pandemic, over 70 million borrowers with $2.3 trillion in loans obtained forbearance (Cherry et al., 2021). Servicers with less buffer and slow replacement rates were less likely to grant forbearance, indicating a greater effect of servicers' funding capacity (Kim et al., 2024).

From the investor's perspective, each buyout is a prepayment. From the servicer's perspective, it is an outflow of cash. If the timing of one is the timing of the other, a prepayment model that predicts buyouts from borrower delinquency implicitly assumes every servicer can fund every buyout as-scheduled. Concentration, though, makes that assumption fragile: when non-banks service most of the market (Figure 1), a squeeze in funding can shift buyout timing across much of the market at once, and at speeds that no borrower-level feature can predict or explain.

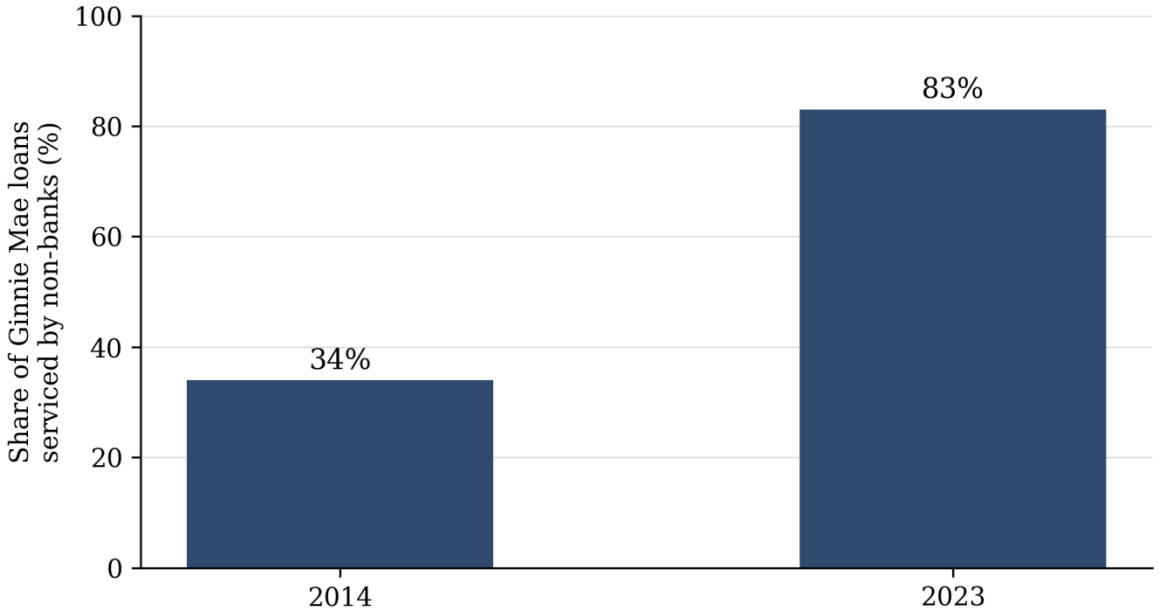


Figure 1. Non-bank share of Ginnie Mae loan servicing, 2014 and 2023. *Source: FSOC (2024).*

A candidate for a servicer funding-stress indicator could be built without privileged information. The candidate inputs would rely on information that is imperfectly observable: the share of the servicer's Ginnie Mae portfolio that was already delinquent, the reliance on secured credit facilities to fund advances, and any available financial disclosure (FSOC, 2024). None of these are sufficient to predict funding stress by themselves, but together as an ordinal stress score mapped to each loan via its servicer identifier, they give the overlay something to condition on.

Where a servicer cannot be scored, the overlay assumes buyouts are delayed under stressed paths (this is intentionally conservative as late buyouts extend the cash-flow profile that an investor has to fund).

4.4 A reference architecture

Figure 2 shows the six layers:

1. Ingestion. Servicer loan tapes pass loan quality validation checks and are written into a versioned feature store, such that any past valuation can be replayed exactly.

2. Competing-risk prediction. A competing-risk model estimates the chance of voluntary prepayment, and the chance of repurchase because of default, for each loan/period.

3. Scenario generation. Stress paths of rates and macroeconomic variables are generated for the liquidity analysis.

4. Liquidity overlay. The timing of the predicted buyout is adjusted based upon the servicer funding-stress indicator from Section 4.3.

5. Pricing. An accelerated Monte Carlo layer where values, option-adjusted spreads (OAS), and scenario-based profit and loss are computed. It uses multilevel estimators first and only uses a neural stand-in when validation allows.

6. Governance. A monitor layer stores SHAP attributions per run, tracks feature drift, and compares actual vs. predicted speeds grouped by the Table 1 mechanisms.

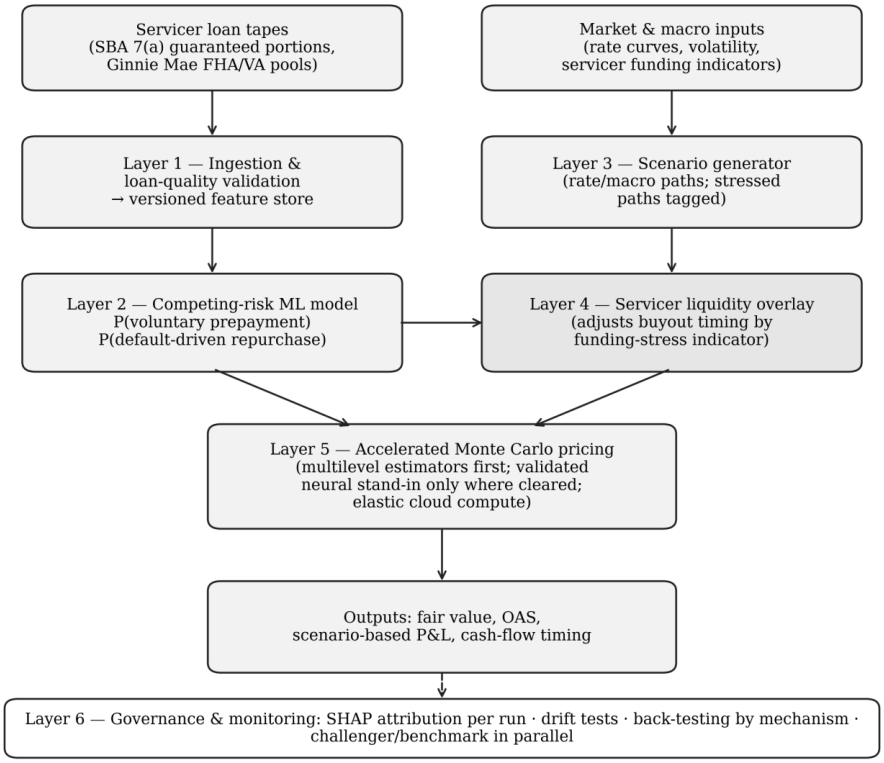


Figure 2. Reference architecture for coupled prepayment–liquidity valuation of government-guaranteed pools.

The layers are designed with weak coupling. Replay is typically the most underappreciated property. For example, if the feature store and scenario sets are versioned together, then any historical mark can be reconstructed

exactly, and a validator who is asking for the last quarter's valuation will get the numbers the desk actually saw, not a reconstruction. In the loose coupling case a parallel run. With the new prepayment model, the full book can be priced alongside the incumbent for an agreed period, with differences mechanism by mechanism before cut-over. That is, how production valuation platforms are usually migrated, and it turns model replacement from a single decision into a monitored transition.

4.5 Governance as a property of architecture

Supervisory requirements for conceptual soundness, continuing monitoring and outcomes analysis, and annual review (FRB & OCC, 2011) become operational for complex credit models with explainable ML (Bussmann et al., 2021). Bücker et al. (2022) insist that transparency and auditability must be designed in, not documented after the fact; the architecture treats governance the same way (Table 4). Explanations are recorded at run time rather than reconstructed for an examiner months later.

Table 4. Supervisory model-risk elements mapped to architectural controls

SR 11-7 element	**Architectural control**
Conceptual soundness	Documented rationale for features; structural benchmark comparison
Ongoing monitoring	Drift tests on features and residuals; alerting thresholds
Outcomes analysis	Back-testing of predicted vs realized speeds, by mechanism
Effective challenge	Challenger model and static-curve benchmark run in parallel
Explainability	SHAP attribution persisted with every valuation run
Reproducibility	Versioned feature store and scenario sets

4.6 Where the argument is exposed

There are three objections to this.

- Regime change. Once learned, models trained during periods of low rates may fail under sudden increases. Though this architecture reduces the problem, it cannot completely remove it. Training windows should be, if possible, at least one cycle long. Scenario paths that leave the support of the training data can be marked at run time, and the static and structural benchmarks can be strongly relied upon until speeds are realized to confirm or disprove the ML prediction.
- Fairness. Loan-level models can encode distributional effects that go beyond market pricing. For guaranteed-pool models, the owners are responsible for watching for these.
- The argument for the liquidity overlay was based largely on mortgage evidence. It is unclear whether the SBA guarantee purchases have similar sensitivity to servicer funding needs, and the magnitude may be weaker.

5. Conclusion

The guarantee is what breaks apart the static curve; by transforming the default into early principal return, government guarantees fold both credit risk and servicer liquidity into the single number a valuation engine most needs: when the cash comes back. A better prediction model addresses one piece of that number, another a simulation layer that could do that same work overnight. The others (the liquidity overlay, the run-time governance) are kind of a single thing that is specified at the level of architecture with trade-offs documented explicitly. Yet, the question is less which prepayment model does best out of sample, but rather which does least badly. It is whether a held-for-investment book of guaranteed loans can be marked daily, explained to a validator, and trusted when servicers come under stress.

Conflict of Interest

The author has no conflicts of interest to disclose.

Funding

This research did not receive any outside funding.

REFERENCES

1. Bachas, N., Kim, O. S., & Yannelis, C. (2021). Loan guarantees and credit supply. *Journal of Financial Economics, 139*(3), 872–894. https://doi.org/10.1016/j.jfineco.2020.08.008
2. Becker, S., Cheridito, P., & Jentzen, A. (2020). Pricing and hedging American-style options with deep learning. *Journal of Risk and Financial Management, 13*(7), Article 158. https://doi.org/10.3390/jrfm13070158
3. Blumenstock, G., Lessmann, S., & Seow, H.-V. (2022). Deep learning for survival and competing risk modelling. *Journal of the Operational Research Society, 73*(1), 26–38. https://doi.org/10.1080/01605682.2020.1838960
4. Board of Governors of the Federal Reserve System & Office of the Comptroller of the Currency. (2011). *Supervisory guidance on model risk management* (SR Letter 11-7). https://www.federalreserve.gov/frrs/guidance/supervisory-guidance-on-model-risk-management.htm
5. Bücker, M., Szepannek, G., Gosiewska, A., & Biecek, P. (2022). Transparency, auditability, and explainability of machine learning models in credit scoring. *Journal of the Operational Research Society, 73*(1), 70–90. https://doi.org/10.1080/01605682.2021.1922098
6. Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics, 57*(1), 203–216. https://doi.org/10.1007/s10614-020-10042-0
7. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785
8. Cherry, S., Jiang, E., Matvos, G., Piskorski, T., & Seru, A. (2021). Government and private household debt relief during COVID-19. *Brookings Papers on Economic Activity, 2021*(2), 141–221. https://doi.org/10.1353/eca.2022.0002
9. Financial Stability Oversight Council. (2024). *Report on nonbank mortgage servicing 2024*. U.S. Department of the Treasury. https://home.treasury.gov/system/files/261/FSOC-2024-Nonbank-Mortgage-Servicing-Report.pdf
10. Kahale, N. (2020). General multilevel Monte Carlo methods for pricing discretely monitored Asian options. *European Journal of Operational Research, 287*(2), 739–748. https://doi.org/10.1016/j.ejor.2020.04.022
11. Kim, Y. S., Lee, D., Scharlemann, T., & Vickery, J. (2024). Intermediation frictions in debt relief: Evidence from CARES Act forbearance. *Journal of Financial Economics, 158*, Article 103873. https://doi.org/10.1016/j.jfineco.2024.103873
12. Kunsági-Máté, S., Fáth, G., Csabai, I., & Molnár-Sáska, G. (2023). Deep weighted Monte Carlo: A hybrid option pricing framework using neural networks. *Quantitative Finance, 23*(4), 615–629. https://doi.org/10.1080/14697688.2023.2181206
13. Lee, C. K., & Lee, M. (2023). Regional redistribution through SBA guaranteed loan programs. *Journal of Corporate Finance, 78*, Article 102344. https://doi.org/10.1016/j.jcorpfin.2022.102344
14. Liu, Z., & Liang, H. (2024). The prediction of mortgage prepayment risks in the early stages of loan origination: A machine learning approach. *The Journal of Risk, 27*(2), 14–43. https://doi.org/10.21314/JOR.2024.017
15. Longstaff, F. A., & Schwartz, E. S. (2001). Valuing American options by simulation: A simple least-squares approach. *The Review of Financial Studies, 14*(1), 113–147. https://doi.org/10.1093/rfs/14.1.113
16. Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence, 2*(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9
17. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30* (NIPS 2017). https://papers.nips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
18. Pence, K. (2022). Liquidity in the mortgage market: How does the COVID-19 crisis compare with the global financial crisis? *Real Estate Economics, 50*(6), 1405–1424. https://doi.org/10.1111/1540-6229.12389
19. Sadhwani, A., Giesecke, K., & Sirignano, J. (2021). Deep learning for mortgage risk. *Journal of Financial Econometrics, 19*(2), 313–368. https://doi.org/10.1093/jjfinec/nbaa025
20. Schultz, G. M., & Fabozzi, F. J. (2021). Rise of the machines: Application of machine learning to mortgage prepayment modeling. *The Journal of Fixed Income, 31*(3), 6–19. https://doi.org/10.3905/jfi.2021.1.123
21. Schwartz, E. S., & Torous, W. N. (1989). Prepayment and the valuation of mortgage-backed securities. *The Journal of Finance, 44*(2), 375–392. https://doi.org/10.1111/j.1540-6261.1989.tb05062.x
22. Stanton, R. (1995). Rational prepayment and the valuation of mortgage-backed securities. *The Review of Financial Studies, 8*(3), 677–708. https://doi.org/10.1093/rfs/8.3.677