Second Opinions: Human Review Beyond Algorithmic Credit Pricing

Henry Woo[1], Kyoungwon (KY) Kim[2], Dean Ryu[3]

[1]Harvard Business School, [2]University of Florida, [3]ITAM Business School

Abstract

We examine human intervention in AI-assisted lending using data from LendingClub, where a subset of borrowers is randomly selected for verification, enabling human review of algorithmic decisions. Exploiting this institutional variation and restricting attention to borrowers with identical observable credit profiles, we identify the causal effect of human intervention. We find that human verification leads to significantly higher interest rates, even among observationally identical borrowers. Moreover, human-induced interest rate adjustments predict ex-post default beyond the AI prediction, reflecting superior screening of latent borrower risk rather than higher interest rates driving borrower outcomes. We prove these findings as arising from human screeners processing the same borrower information differently and potentially applying conservative adjustments because of agency concerns or uncertainty about AI recommendations. Our results highlight the role of heterogeneous information processing in human-AI decision-making.

Introduction

Backgrounds and Research Question

- Algorithmic decisions are frequently subject to human review.
- This raises a central question: when humans review AI-generated decisions, do they contribute incremental predictive information or simply alter algorithmic recommendations?
- Even when humans and AI observe the same borrower information, they may process that information differently, leading to either improved risk assessment or systematic distortions.

Empirical Challenge and Solution

- A key empirical challenge is that human intervention is typically endogenous, making it difficult to distinguish informational advantages from selection effects.
- This paper addresses this challenge by exploiting an institutional feature of LendingClub (LC), a large U.S. peer-to-peer lending platform. The platform also randomly selects a subset of borrowers for verification, even when additional review is not required. This institutional feature generates plausibly exogenous variation in post-AI human review.

Findings

- Human verification leads to significantly higher interest rates, even among borrowers with identical credit profiles.
- Human-induced interest rate adjustments contain incremental predictive information about subsequent default beyond the AI recommendation.

Data & Empirical Strategy

Quasi-random variation with a multivariate matching design

- Using the LC loan application data (2007-2020), we construct cohorts of borrowers with identical observable credit profiles.
- It allows us to compare otherwise identical borrowers who differ only in whether they receive post-AI human review (verification).

Human Screening Captures Residual Borrower Risk

- IV (2016 LC funding shock)-based 2SLS estimates indicate that human-induced interest rate adjustments primarily reflect "predictive selection (capturing borrower risk beyond the AI)" rather than a "causal effect" of higher interest rates on default.

Rationalized Mechanism: Loss-Minimization Model of Human Screening

- Human screeners optimally adjust AI-assigned interest rates by minimizing an asymmetric loss.
- Upward adjustments reflect both informational corrections and precautionary conservatism.

Results

- *Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.*
- *Robustness checks are conducted separately for the high- and low-credit-profile subsamples.*

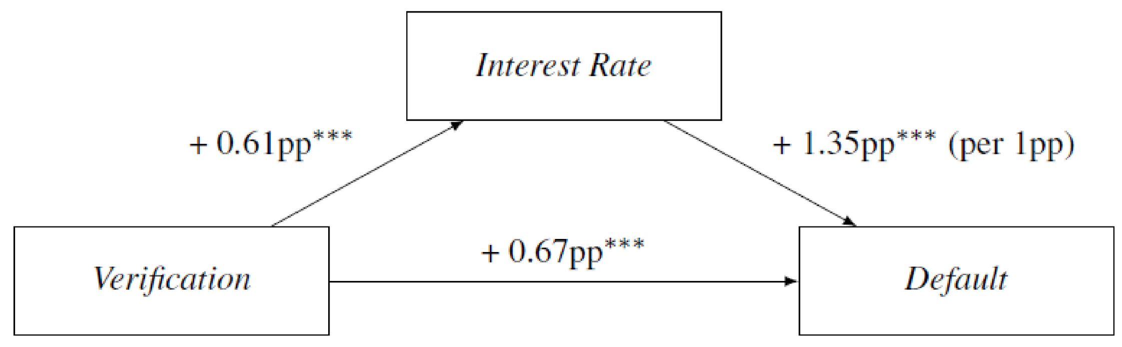
Table 1. Human Verification and Interest Rates

Dep. Variable:	Interest Rate				
	(1)	(2)	(3)	(4)	(5)
Human Verification	0.010***	0.005***	0.006***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Observations	334,900	334,900	334,900	334,899	334,899
R-squared	0.012	0.377	0.399	0.400	0.651
Controls	No	Yes	Yes	Yes	Yes
Year-Quarter FE	No	No	Yes	Yes	Yes
State FE	No	No	No	Yes	Yes
Cohort FE	No	No	No	No	Yes

Table 2. Default Risk: AI-Assigned Rates vs. Human Adjustments (Predictive Content)

Dep. Variable:	Default					
	(1)	(2)	(3)	(4)	(5)	(6)
AI-Assigned Interest Rate	1.296***	1.139***	1.316***	-0.154*	0.633***	0.552***
	(0.089)	(0.081)	(0.087)	(0.091)	(0.082)	(0.082)
Human Interest Rate Adjustment			0.631***	1.409***		
			(0.034)	(0.058)		
Upward Adjusted					0.050***	
					(0.003)	
Downward Adjusted					-0.026***	
					(0.004)	
Upward Adjusted (50bp)						0.055***
						(0.003)
Downward Adjusted (50bp)						-0.034***
						(0.004)
Observations	216,047	166,798	216,047	166,798	166,798	166,798
R-squared	0.073	0.464	0.077	0.470	0.466	0.466
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	No	Yes	Yes	Yes

Figure 1. Verification Effect Decomposition: Direct and Interest-Rate Channels



Verification → Default	
Direct effect	+ 0.67*** (0.22)
Indirect effect	+ 0.83*** (0.06)
Total effect	+ 1.50*** (0.21)

Discussions & Conclusion

Interpretations

- Human screeners improve AI-assisted risk assessment by processing the same borrower information differently from the algorithm.
- Human intervention serves as an informational correction to the algorithmic recommendation while potentially incorporating conservative adjustments arising from agency concerns, ambiguity about the AI recommendation, or algorithm aversion.

Contributions

- We provide causal evidence on human intervention in AI-assisted decision-making by exploiting quasi-random verification together with multivariate exact matching, allowing us to isolate the effect of post-AI human review
- We contribute to the literature on human–algorithm interaction by showing that human intervention can improve prediction beyond the AI recommendation even when humans rely on the same underlying information, highlighting heterogeneous information processing as a source of complementary predictions (Athey et al., 2020).
- We contribute to the literature on FinTech lending by showing that human-induced interest rate adjustments primarily reflect superior screening of latent borrower risk rather than a causal pricing effect (Morse, 2015; Iyer et al., 2016; Liberti & Petersen, 2019).

References

Athey, S. C., Bryan, K. A., and Gans, J. S. The allocation of decision authority to human and artificial intelligence. In *American Economic Association Papers and Proceedings, 110*, 80–84, 2020.
Baron, R. M. and Kenny, D. A. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6):1173–1182, 1986.
Berg, T., Burg, V., Gombovi ´c, A., and Puri, M. On the rise of fintechs: Credit scoring using digital footprints. *Review of Financial Studies, 33*(7):2845–2897, 2020.
Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., and Walther, A. Predictably unequal? the effects of machine learning on credit markets. *Journal of Finance, 77*(1):5–47, 2022.
Imai, K., Keele, L., and Yamamoto, T. Identification, inference, and sensitivity analysis for causal mediation effects. *Statistical Science, 25*(1):51–71, 2010.
Iyer, R., Khwaja, A. I., Luttmer, E. F., and Shue, K. Screening peers softly: Inferring the quality of small borrowers. *Management Science, 62*(6):1554–1577, 2016.
Liberti, J. M. and Petersen, M. A. Information: Hard and soft. *Review of Corporate Finance Studies, 8*(1):1–41, 2019.
MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. Mediation analysis. *Annual Review of Psychology, 58*(1):593–614, 2007.
Morse, A. Peer-to-peer crowdfunding: Information and the potential for disruption in consumer lending. *Annual Review of Financial Economics, 7*(1):463–482, 2015.

Acknowledgements

We thank Kwangwon Ahn, Jaehoon Hahn, Ki Hyun, Jihyun (Jason) Kim, and Taeyoung Park for helpful comments. Henry Woo gratefully acknowledges support and encouragement from Paul Gompers, Harry Mamaysky, and Marco Sammon, as well as financial support from the Hyundai Motor Chung Mong-Koo Foundation and the ICML-TAIGR 2026 Travel Grant.