Second Opinions: Human Review Beyond Algorithmic Credit Pricing

Henry Woo [1] **Kyongwon (KY) Kim** [2] **Dean Ryu** [3]

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

> "*A single human bias might lead to a poor investment or a misguided decision. However, an AI bias, magnified across millions of interactions, has the potential to subtly influence markets.*"
> — Forbes, August 24, 2025

Artificial intelligence (AI) is playing an increasingly central role in decision-making in financial markets. Algorithmic models are now widely used to evaluate borrower risk and determine loan interest rates, reshaping how credit is allocated (Berg et al., 2020; Fuster et al., 2022). Yet, in most real-world applications, AI does not operate autonomously. Instead, 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.

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, a large U.S. peer-to-peer lending platform. LendingClub uses proprietary algorithms to assign credit grades and interest rates while also verifying borrower information. Although verification is typically triggered by suspicious applications or missing documentation, 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.

We combine this quasi-random variation with a multivariate exact matching design to identify the causal effect of human intervention. Using loan application data from 2007 through the third quarter of 2020, we construct cohorts of borrowers with identical observable credit profiles. Within these matched cohorts, some borrowers are verified while others are not, allowing us to compare otherwise identical borrowers who differ only in whether they receive post-AI human review.

Our analysis yields two main findings. First, human verification leads to significantly higher interest rates, even among borrowers with identical credit profiles. Second, human-induced interest rate adjustments contain incremental predictive information about subsequent default beyond the AI recommendation. Instrumental-variable estimates exploiting the exogenous contraction in loan funding following the 2016 LendingClub governance scandal show that this relationship reflects superior screening of latent borrower risk rather than a causal effect of higher interest rates on default probability. Taken together, these findings suggest that human screeners improve AI-assisted risk assessment

[1]Harvard Business School, MA, USA [2]University of Florida, FL, USA [3]ITAM Business School, Mexico City, Mexico. Correspondence to: Henry Woo <hwoo@hbs.edu>.

Second Workshop on Technical AI Governance Research (TAIGR) @ ICML 2026, Seoul, South Korea. 2026. Copyright 2026 by the author(s).

by processing the same borrower information differently from the algorithm.

To interpret these findings, we develop a stylized model in which both the AI system and the human screener observe the same borrower information but process it differently. 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. The model rationalizes why human verification can simultaneously increase assigned interest rates and improve the prediction of subsequent default.

We contribute to several strands of the literature. First, 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. Second, 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). Third, 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).

Our findings highlight an important design challenge for hybrid human–AI systems. The value of human oversight need not arise from access to additional information, but can instead emerge from alternative processing of the same information. Designing effective human–AI systems therefore requires identifying when human judgment provides genuinely incremental predictive value and when it primarily reflects conservative adjustments or other behavioral distortions.

Data and Identification

Data

We use loan-level data from LendingClub covering the period from the first quarter of 2007 through the third quarter of 2020. The sample consists of 2,712,230 accepted loan applications. For each loan, the dataset provides detailed information on borrower characteristics, including loan terms, verification status, and realized loan performance.

A key variable in our setting is *verification*, which captures whether a borrower undergoes additional document screening by LendingClub. We define an indicator equal to one if a borrower undergoes verification (verified borrower) and zero otherwise (non-verified borrower). Verification reflects the involvement of human screeners who review borrower information and may adjust the algorithmically assigned loan terms. To mitigate the influence of extreme values, we winsorize all continuous variables at the 1st and 99th percentiles. Summary statistics are reported in Table A.

Multivariate cohort matching

To identify the effect of human intervention following AI-based decisions, we focus on borrowers with identical observable credit profiles who differ only in whether they receive post-AI human verification. Specifically, we contrast (i) non-verified borrowers whose interest rates are determined solely by the algorithm with (ii) verified borrowers who receive additional human review after the algorithmic decision.

We construct matched cohorts using a multivariate exact matching procedure. Borrowers are first grouped by loan issuance year. Within each year, continuous variables are discretized into economically meaningful bins, while categorical variables retain their original categories. Matching is then performed across all borrower characteristics simultaneously so that verified and non-verified borrowers within the same cohort share identical observable credit profiles. Detailed matching criteria are reported in Table B.

Our analysis is restricted to matched cohorts containing at least one verified and one non-verified borrower. The resulting matched sample consists of 334,900 borrowers, including 200,189 verified borrowers and 134,711 non-verified borrowers. Table C reports covariate balance before and after matching. Following matching, standardized differences for all borrower characteristics are small, indicating that verified and non-verified borrowers are highly comparable on observable dimensions.

By construction, our matching procedure compares borrowers within the same region of the observable covariate space and avoids support extrapolation inherent in propensity-score-based approaches. Within the matched sample, remaining variation in verification is plausibly driven by LendingClub's institutional verification procedures rather than observable borrower characteristics, allowing us to interpret the estimated effects as causal.

Empirical Strategy

Selection regression test after matching

Prior to matching, borrower characteristics strongly predict the likelihood of verification. Because our matching procedure groups borrowers into identical bins of all covariates, these characteristics should no longer predict verification in the matched sample. To verify that matching removes observable selection, we estimate regressions of verification status on the original borrower characteristics while controlling for matched-cohort fixed effects. Standard errors are

double-clustered at the issuance year and matched cohort levels:

$$verification_i = \alpha_0 + \boldsymbol{\delta_0}'\mathbf{X}_i + \gamma_{c(i)} + \varepsilon_{i0} \quad (1)$$

where *verification*$_i$ is an indicator equal to one if borrower i undergoes verification; $\mathbf{X}_i$ is a vector of borrower characteristics measured at the time of application; $\boldsymbol{\delta_0}$ denotes the coefficients on $\mathbf{X}_i$; $\gamma_{c(i)}$ represents matched-cohort fixed effects; and ε_{i0} is an error term.

Insignificant coefficients on the borrower characteristics indicate that, conditional on the matched cohorts, verification is no longer systematically related to observable credit profiles. These results support our interpretation that verification is plausibly quasi-random within the matched sample.

Verification treatment effect on interest rates and default
We first estimate whether post-AI human verification affects assigned interest rates:

$$\textit{interest rate}_i = \alpha_1 + \beta_1 verification_i + \boldsymbol{\delta_1}^\top \mathbf{X}_i + \gamma_{c(i)} + \varepsilon_{i1}, \quad (2)$$

where *interest rate*$_i$ denotes borrower i's assigned interest rate. Standard errors are double-clustered at the issuance year and matched cohort levels. We also estimate specifications excluding $\mathbf{X}_i$ because matched-cohort fixed effects absorb most variation in observable borrower characteristics.

Next, among verified borrowers, we measure the human-induced interest rate adjustment relative to comparable non-verified borrowers within the same matched cohort. Verified borrowers whose assigned interest rates exceed the cohort average for non-verified borrowers are classified as *upward adjusted*, whereas those receiving lower interest rates are classified as *downward adjusted*.

To examine whether human adjustments contain incremental predictive information beyond the AI recommendation, we estimate

$$default_i = \alpha_2 + \beta_2 verification_i + \boldsymbol{\theta_2}^\top \mathbf{Z}_i + \boldsymbol{\delta_2}^\top \mathbf{X}_i + \gamma_{c(i)} + \varepsilon_{i2}, \quad (3)$$

where $default_i$ equals one if borrower i subsequently defaults or is charged off, and zero if the loan is fully repaid. The vector $\mathbf{Z}_i$ alternatively includes the assigned interest rate, the human-induced interest rate adjustment, or indicators for upward- and downward-adjusted borrowers. Standard errors are clustered at the issuance year level.

Statistically significant coefficients on the human adjustment variables indicate that human intervention contains incremental predictive information beyond the AI recommendation. However, predictive relationships alone cannot distinguish whether human screeners identify latent borrower risk (selection) or whether higher human-induced interest rates themselves increase default (causal pricing). We therefore separately examine these competing mechanisms using an instrumental-variable strategy.

Decomposing effects via mediation analysis
Because interest rates are assigned during the verification process, they are post-treatment variables rather than predetermined controls. Conditioning directly on interest rates may therefore obscure the pathways through which verification is associated with subsequent default.

To characterize these pathways, we adopt the mediation framework of Baron & Kenny (1986) and subsequent developments by MacKinnon et al. (2007) and Imai et al. (2010). Specifically, we estimate

$$default_i = \alpha_{21} + \beta_{21} verification_i + \theta_{21} interest\ rate_i + \boldsymbol{\delta_{21}}^\top \mathbf{X}_i + \gamma_{c(i)} + \varepsilon_{i21}, \quad (4)$$

which corresponds to Equation 3 with $\mathbf{Z}_i$ containing only the assigned interest rate. Under the standard linear mediation assumptions of Imai et al. (2010), the coefficient on verification captures the direct effect, while the product $\beta_1 \times \theta_{21}$ measures the indirect effect operating through interest rates. The total effect corresponds to the coefficient on verification in Equation 3 when the mediator is omitted.

Selection versus causal pricing
Although human-induced interest rate adjustments may predict subsequent default, this relationship admits two distinct interpretations. Human screeners may assign higher interest rates because they identify borrowers with greater latent default risk (selection), or higher interest rates may themselves increase subsequent default (causal).

To distinguish between these mechanisms, we estimate two-stage least squares specifications using the exogenous contraction in loan funding following the 2016 LendingClub governance scandal as an instrument for human-induced interest rate adjustments. This strategy isolates the causal effect of higher interest rates on subsequent default from the predictive content of human screening, allowing us to distinguish superior risk selection from causal pricing effects.

Results

Multivariate cohort matching and selection regressions
Table D reports the selection regressions before and after matching. In the full sample, borrower characteristics strongly predict verification status, indicating that verification is systematically related to observable borrower characteristics. After implementing the multivariate exact matching procedure and conditioning on matched-cohort fixed effects, these relationships become economically small and

statistically insignificant. Thus, within matched cohorts of borrowers with identical observable credit profiles, verification is no longer explained by observable borrower characteristics.

These results provide strong support for our identification strategy. Conditional on matched cohorts, remaining variation in verification is plausibly quasi-random, allowing us to interpret differences between verified and non-verified borrowers as the causal effect of post-AI human intervention.

Human intervention: interest rates and default

Table 1 (see Table E for details) shows that, within matched cohorts, human verification significantly increases assigned interest rates by approximately 50–60 basis points. Even among borrowers with identical observable credit profiles, post-AI human review systematically results in higher interest rates, indicating that human screeners do not simply reaffirm the algorithmic recommendation but instead revise it upward.

Table 2 (see Table F for details) examines whether human-induced interest rate adjustments contain incremental predictive information beyond the algorithmic recommendation. The AI-assigned interest rate strongly predicts subsequent default, confirming that the algorithm captures economically meaningful variation in ex ante borrower risk. More importantly, human-induced interest rate adjustments remain significantly associated with subsequent default even after controlling for the AI-assigned interest rate. Likewise, borrowers whose interest rates are adjusted upward exhibit significantly higher default probabilities, whereas those receiving downward adjustments exhibit significantly lower default probabilities. Similar patterns emerge when focusing on adjustments exceeding 50 basis points. Together, these results indicate that human-induced adjustments contain incremental predictive information beyond the algorithmic recommendation.

To further characterize the relationship between verification, interest rates, and subsequent default, we decompose the association between verification and default into direct and indirect components using a standard mediation framework, illustrated in Figure 1 and reported in Table G. Verification is positively associated with assigned interest rates, and assigned interest rates are positively associated with subsequent default, generating a positive indirect association through interest rates. Verification also remains positively associated with default after accounting for the interest-rate channel, implying that the overall association between verification and subsequent default is positive.

These findings suggest that post-AI human intervention contributes incremental predictive information beyond the algorithmic recommendation. However, the positive relationship between human-induced interest rate adjustments and subsequent default does not, by itself, distinguish whether human screeners identify borrowers with greater latent default risk or whether higher interest rates themselves increase subsequent default. We investigate these competing explanations in the following section using an instrumental-variable strategy.

Selection and Causal Effects of Human-Induced Interest Rates

Human-induced interest rate adjustments are predictive of subsequent default, but this relationship admits two distinct interpretations. One possibility is *selection*: human screeners identify borrowers with higher latent default risk and assign higher interest rates accordingly. The alternative is a *causal pricing effect*: higher interest rates increase borrowers' repayment burden and thereby raise default probabilities.

In Table H, our evidence strongly favors the selection interpretation. Human-induced interest rate adjustments contain incremental predictive information beyond the AI recommendation, indicating that human screeners identify residual borrower risk not fully reflected in the algorithmic assessment. Exploiting the exogenous funding shock following the 2016 LendingClub governance scandal as an instrument for human-induced interest rate adjustments, two-stage least squares estimates provide little evidence that higher interest rates themselves increase default. We therefore separately examine the predictive content of human adjustments and the causal effect of interest rates on borrower default.

A Stylized Model of Human-Induced Interest Rate Increases

We develop a simple model to interpret the paper's two main empirical findings: human verification raises AI-assigned interest rates, and human-induced interest-rate adjustments predict ex-post default. The model assumes that both the AI system and the human verifier observe the same borrower characteristics. Human verification therefore does not introduce additional borrower information but instead reflects an alternative interpretation of the same observable information. We first model human intervention as an informational correction to the AI recommendation and then allow for an additional conservative adjustment arising from agency concerns, ambiguity about the AI recommendation, or algorithm aversion. The complete derivation is provided in Appendix A.

Let F_i denote the observed borrower characteristics and p_i the true default probability of borrower i. The AI system maps borrower information into an estimated default probability, $\widehat{p}_i^A = \mathbb{E}[p_i \mid F_i]$, and assigns an algorithmic interest rate $r_i^A = r(\widehat{p}_i^A)$. Because the algorithmic recommendation

Table 1. Human Verification and Interest Rates

This table reports OLS estimates of the effect of human verification on loan interest rates. The dependent variable is the annualized interest rate assigned at origination. *Human Verification* is an indicator equal to one if the borrower underwent additional human screening. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

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. Predictive Content of Human Adjustment

This table examines the incremental predictive information of human rate adjustments beyond the algorithmic baseline for loan default. See Table F in the Appendix for details. The dependent variable equals one if the loan is classified as charged off (clearly non-performing) or default (overdue by more than 120 days), and zero if it is fully paid. *AI-Assigned Interest Rate* is the OLS-predicted rate a non-verified borrower would receive given the same observable characteristics. *Human Interest Rate Adjustment* is the deviation of the final assigned rate from the AI-assigned rate (zero for non-verified borrowers). The 50bp variants require the deviation to exceed 50 basis points in absolute value. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

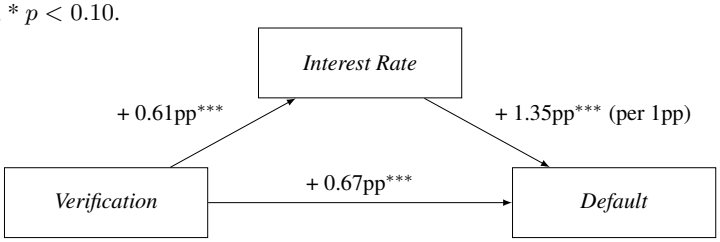
Dep. Variable:	Default					
	(1)	(2)	(3)	(4)	(5)	(6)
AI-Assigned Interest Rate	1.296***	1.139***	1.316***	-0.154*	0.633***	0.552***
	(0.089)	(0.081)	(0.087)	(0.091)	(0.082)	(0.082)
Human Interest Rate Adjustment			0.631***	1.409***		
			(0.034)	(0.058)		
Upward Adjusted					0.050***	
					(0.003)	
Downward 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

need not coincide with the efficient benchmark, we allow $r_i^* = r_i^A + u_i$, where u_i denotes residual borrower risk not fully reflected in the AI recommendation.

A randomly assigned human verifier observes the same borrower information together with the AI recommendation and forms an informational correction based on an alternative interpretation of the observed characteristics. The human adjustment is $h_i = s_i + b_i$, where s_i denotes the informational correction and b_i captures a potential conservative adjustment arising from agency concerns, ambiguity about the AI recommendation, or algorithm aversion. The final assigned interest rate is therefore $r_i^H = r_i^A + h_i$. When $b_i = 0$, human intervention serves purely as an informational correction. Human adjustments therefore identify residual borrower risk not fully incorporated into the AI rec-

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

Notes: This figure visualizes the mediation analysis reported in Table G, decomposing the effect of verification on default into direct and indirect paths through interest rates. Coefficients are expressed in percentage points (pp), and standard errors (in parentheses) are double-clustered by issue year and matched cohort. See Figure A in the Appendix for details. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.



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

ommendation and are informative about subsequent default outcomes. More generally, when $b_i > 0$, human verification raises assigned interest rates beyond the informational correction while preserving the predictive content of s_i. Consequently, the model rationalizes the paper's two central empirical findings: $\mathbb{E}[h_i \mid \text{Verification}_i = 1] > 0$ and $\text{Cov}(h, \text{Default}) > 0$.

Identification Checks via Credit Quality

Low-credit-quality borrowers: selection concerns
A potential concern is that, among borrowers with weaker credit profiles, verification may lead not only to higher assigned interest rates but also to outright loan rejection. If verification disproportionately removes the riskiest borrowers from the observed sample, the estimated effect of human intervention on interest rates may understate the true effect because the most extreme cases are no longer observed.

To examine this possibility, we re-estimate our baseline specifications using only low-credit-quality borrowers (Grades D–G), for whom such selection concerns are most relevant. Table I shows that verification continues to increase assigned interest rates by magnitudes comparable to those in the full sample. Likewise, Table J shows that human-induced interest rate adjustments remain positively associated with subsequent default. Although some indicator specifications become less precisely estimated owing to the substantially smaller matched sample, the overall pattern is consistent with the baseline results: borrowers receiving larger upward human adjustments exhibit higher subsequent default rates, and continuous measures of human adjustments remain strongly predictive of default.

Importantly, any rejection of the riskiest verified borrowers would attenuate the estimated treatment effect because these borrowers would be excluded from the observed sample. In this sense, our estimates should be interpreted as conservative. The persistence of both the pricing effect and the predictive content of human adjustments suggests that our main findings are unlikely to be driven by sample selection arising from loan rejection among low-credit-quality borrowers.

High-credit-quality borrowers: verification endogeneity
A related concern is that verification may still reflect unobserved borrower characteristics within matched cohorts. Although our matching procedure removes observable differences, some verified borrowers may nevertheless be selected because they appear suspicious based on information unavailable to the researcher. To address this concern, we focus on the highest-credit-quality borrowers (Grade A), for whom verification is unlikely to be triggered by underlying credit risk and therefore more plausibly reflects LendingClub's random verification procedures.

Table K shows that verification continues to increase assigned interest rates in this subsample, although the magnitude is smaller than in the full sample, consistent with more limited scope for discretionary adjustments among borrowers with already strong credit profiles. More importantly, Table L shows that human-induced interest rate adjustments continue to contain incremental predictive information beyond the AI recommendation. While some coefficient estimates are attenuated because of the substantially lower default rate among Grade A borrowers, the continuous human adjustment remains significantly associated with subsequent default, and upward adjustments continue to predict higher default risk. These patterns indicate that the predictive content of human intervention is not driven by suspicious borrowers disproportionately entering the verification process.

Taken together, these subsample analyses strengthen our identification strategy. The low-credit-quality analysis suggests that our estimates are unlikely to be driven by sample selection arising from loan rejection, while the high-credit-quality analysis alleviates concerns that verification reflects residual endogeneity within matched cohorts. Across both settings, human verification continues to increase assigned interest rates, and human-induced interest rate adjustments continue to contain incremental predictive information be-

yond the algorithmic recommendation.

Conclusion

This paper studies how human intervention affects AI-assisted lending decisions. Exploiting quasi-random verification at LendingClub together with multivariate exact matching, we identify the causal effect of post-AI human review among borrowers with identical observable credit profiles. We find that human verification systematically increases assigned interest rates and that human-induced interest rate adjustments contain incremental predictive information about subsequent default beyond the AI recommendation.

Our evidence indicates that this predictive content reflects superior screening rather than a causal effect of higher interest rates on default probability. Instrumental-variable estimates exploiting the exogenous contraction in loan funding following the 2016 LendingClub governance scandal show that human-induced interest rate adjustments have no statistically significant causal effect on subsequent default. Instead, human screeners appear to identify residual borrower risk that is not fully reflected in the algorithmic recommendation.

To interpret these findings, we develop a stylized model in which humans and AI observe the same borrower information but process it differently. Human intervention therefore reflects an informational correction to the algorithmic recommendation and can additionally incorporate conservative adjustments arising from agency concerns or uncertainty about AI recommendations. This framework rationalizes why human verification both raises interest rates and improves the prediction of subsequent default.

Our findings contribute to the growing literature on human–AI decision-making by showing that the value of human oversight need not arise from access to additional information. Instead, complementary predictions can emerge because humans and algorithms process the same information differently. More broadly, our results highlight heterogeneous information processing as a distinct source of value in hybrid human–AI decision systems.

These findings also have important implications for the design of AI-assisted decision systems. Rather than viewing human review solely as a safeguard against algorithmic errors, institutions should identify when human judgment provides incremental predictive value beyond algorithmic recommendations and structure decision processes to preserve those complementary signals while limiting unnecessary conservative adjustments. More broadly, our results suggest that the effectiveness of hybrid human–AI systems depends less on whether decisions are made by humans or algorithms alone than on how institutions combine their distinct modes of information processing.

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.

Impact Statement

Human intervention in AI-assisted decision systems can provide incremental predictive information beyond algorithmic recommendations, even when humans and AI rely on the same underlying information. Our findings highlight the importance of understanding how heterogeneous information processing shapes decision quality and the design of hybrid human–AI systems in high-stakes domains. Appendix B discusses the broader policy implications.

LLM Usage Statement: LLMs were used for language editing and programming assistance. All analyses were conducted and verified by the authors.

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*, volume 110, pp. 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., Gombović, A., and Puri, M. On the rise of fintechs: Credit scoring using digital footprints. *The 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. *The 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.

A. Appendix: Mechanism Proof through Stylized Model

Introduction

The main empirical result of the paper is that post-AI human verification raises AI-assigned interest rates conditional on observable credit profiles. In our empirical results, verified borrowers receive higher interest rates than non-verified borrowers in the same matched cohort. At the same time, the human-induced adjustment to AI-assigned interest rates contains predictive information about ex-post default: upward adjustments are associated with higher default probabilities, while downward adjustments are associated with lower default probabilities.

We develop a stylized model that rationalizes these empirical findings. The model focuses on how existing borrower information is interpreted. Specifically, the model assumes that the algorithm and the human screener observe the same set of reported borrower characteristics. Human verification does not introduce new borrower characteristics; rather, it allows the screener to assess the credibility of the reported information, which may lead to a different assessment of borrower risk. In particular, the algorithm provides a scalable and systematic mapping from borrower characteristics to interest rates. The human screener, however, may disagree with this mapping, perceive that certain characteristics are underweighted, or place greater weight on characteristics viewed as particularly informative about borrower risk.

The baseline model treats the human adjustment as an informational correction based on the human screener's alternative processing of the same borrower information. This captures the idea that human screeners may identify residual risk that is not fully reflected in the algorithmic interest rate. We then extend the model by allowing for an additional distortion or markup arising from agency concerns, ambiguity aversion, algorithm aversion, salience, or defensive screening. This extension provides a way to ask whether human intervention is merely informative, or whether humans are over-conservative relative to the efficient benchmark.

Human verification as informational correction

Consider a borrower i with true default probability p_i. The platform observes a vector of borrower information F_i, such as FICO score, income, debt-to-income ratio, loan amount, credit history, employment length, loan purpose, and other observable borrower characteristics. The algorithm uses this information to form an estimate of default risk,

$$\widehat{p}_i^A = \mathbb{E}[p_i \mid F_i],$$

and assigns an algorithmic interest rate

$$r_i^A = r(\widehat{p}_i^A).$$

The algorithmic rate is informative because it summarizes the platform's prediction based on borrower characteristics. However, the algorithmic mapping from F_i to default risk need not be complete or perfectly calibrated. Let r_i^* denote the efficient benchmark rate that would be assigned if the available information were processed optimally. We allow the algorithmic rate to differ from this benchmark:

$$r_i^* = r_i^A + u_i,$$

where u_i captures the residual gap between the algorithmic rate and the efficient benchmark. If $u_i > 0$, the efficient rate is higher than the algorithmic rate, so the algorithmic rate is too lenient relative to the benchmark. If $u_i < 0$, the efficient rate is lower than the algorithmic rate, so the algorithmic rate is too conservative.

This residual can arise for several reasons. The algorithm may not fully exploit all information relevant for predicting default risk. Alternatively, even if the algorithm processes the available information correctly given its objective, the objective or instruction embedded in the algorithm may itself be misspecified by human designers. In this sense, the residual need not reflect a mechanical failure of the machine; it may also reflect uncertainty about whether the algorithm has been designed, calibrated, or instructed to map borrower information into risk in the correct way.

A subset of post-AI loans is then randomly assigned to a human screener for verification. The human observes the same borrower information F_i and the algorithmic rate r_i^A. The human does not necessarily take the algorithmic rate at face value. Instead, the human forms a subjective correction to the algorithmic rate based on a different interpretation of the same information:

$$s_i = g_H(F_i, r_i^A).$$

The term s_i captures the informational component of human intervention. It reflects the human screener's attempt to correct perceived residual risk in the algorithmic rate. For example, the human may place different weight on combinations of

borrower characteristics, such as high income but weak credit history, low FICO score but low debt-to-income ratio, or large loan amount with short employment tenure.

In the baseline model, the human adjustment (rate) is this informational correction:

$$h_i = s_i.$$

The final human-adjusted rate is therefore

$$r_i^H = r_i^A + h_i = r_i^A + s_i.$$

If the human correction successfully captures residual risk not fully reflected in the algorithmic rate, then h_i should be informative about ex-post default. In particular, upward human adjustments should be associated with higher default probabilities, while downward human adjustments should be associated with lower default probabilities.

The efficient benchmark is a useful theoretical reference point. If the human correction exactly equals the residual component of risk, then

$$s_i = u_i,$$

and the human-adjusted rate satisfies

$$r_i^H = r_i^A + s_i = r_i^A + u_i = r_i^*.$$

In this ideal setting, human intervention fully corrects the gap between the algorithmic rate and the efficient benchmark.

In practice, however, neither r_i^* nor u_i is observed. Therefore, the empirical analysis cannot directly test whether $s_i = u_i$. Instead, it can test a weaker implication of the model: whether human adjustments are informative about residual default risk not fully reflected in the algorithmic rate. If so, human adjustments should predict subsequent default outcomes. However, predictive power alone does not imply that the magnitude of the human adjustment is optimally calibrated or that it coincides with the efficient benchmark.

Empirical implications and interpretation

The model provides a simple framework for interpreting the two main empirical findings. The first finding is that human verification further raises interest rates initially assigned by AI, on average. The second finding is that human-induced interest-rate adjustments contain predictive information about ex-post default.

We begin with the average effect of verification on interest rates. Empirically,

$$\mathbb{E}[h \mid \text{Verification} = 1] > 0.$$

In the baseline model,

$$h_i = s_i.$$

A positive average adjustment can arise when verified loans are disproportionately cases in which human screeners perceive residual risk that is not fully reflected in the algorithmic rate. In such cases, the human correction increases the assigned interest rate relative to the algorithmic benchmark.

The second empirical finding concerns the informational content of human intervention. Since we assume the human adjustment is equal to the informational correction, the human adjustment reflects the screener's assessment of residual risk not fully incorporated into the algorithmic rate. If human screeners identify such residual risk, then human adjustments should predict subsequent default outcomes. Formally,

$$\mathrm{Cov}(h, \text{Default}) > 0.$$

Borrowers receiving larger upward adjustments should exhibit higher subsequent default rates, while borrowers receiving downward adjustments should exhibit lower subsequent default rates.

These findings suggest that human intervention may both alter prices and incorporate information about borrower risk that is not fully reflected in the algorithmic rate. However, the baseline model treats human intervention as a purely informational correction and therefore abstracts from the possibility that human adjustments may also contain systematic distortions. We next extend the model to allow for this possibility.

Human over-conservatism as an extension

The baseline model treats human intervention as an informative correction. However, human screeners may not simply implement the efficient correction. They may face asymmetric accountability, career concerns, ambiguity about the algorithmic recommendation, or a tendency to overweight salient negative information. To capture the possibility that humans are over-conservative, we extend the model by allowing the actual human adjustment to include an additional distortion term, denoted by b_i.

The total human adjustment is now

$$h_i = s_i + b_i.$$

The final human-adjusted rate is

$$r_i^H = r_i^A + h_i = r_i^A + s_i + b_i.$$

The term s_i is the informational correction. The term b_i is the distortion component.

The wedge between the actual human-adjusted rate and the efficient benchmark is

$$r_i^H - r_i^* = b_i$$

when

$$s_i = u_i.$$

Therefore, human intervention is over-conservative when

$$b_i > 0.$$

In this case, human screeners correctly identify residual default risk but raise interest rates by more than what is justified by the efficient correction.

The extended model also provides a natural interpretation of the positive average effect of verification on interest rates observed in the data. Empirically,

$$\mathbb{E}[h \mid \text{Verification} = 1] = \mathbb{E}[s + b \mid \text{Verification} = 1] > 0.$$

This condition can hold regardless of the sign of $\mathbb{E}[s_i \mid \text{Verification}_i = 1]$, provided that the average bias component is sufficiently positive. In this sense, the observed increase in rates following verification may be driven by bias rather than by information alone. Because the human adjustment now contains both an informational component and a distortion component, the observed increase in interest rates may reflect not only corrections of perceived residual risk but also systematic upward markups arising from human conservatism. Agency concerns, ambiguity aversion, algorithm aversion, salience, and defensive screening all generate forces that push human adjustments upward relative to the efficient benchmark.

At the same time, the extended model preserves the informational content of human intervention. Since the human adjustment continues to embed the informational component s_i, human adjustments may remain predictive of subsequent default outcomes:

$$\mathrm{Cov}(h, \text{Default}) = \mathrm{Cov}(s + b, \text{Default}) > 0.$$

The informational component improves risk assessment, while the distortion component pushes interest rates above the efficient benchmark. Thus, the extended model can simultaneously rationalize why human verification raises interest rates on average and why human-induced interest-rate changes remain predictive of ex-post default.

A microfoundation for the distortion term

The previous subsection showed that a positive distortion component can help explain why human verification raises interest rates on average. We now provide a simple agency-based microfoundation for such distortions.

The human screener is responsible for reviewing the loan file. If the screener leaves the algorithmic rate unchanged or assigns a relatively low rate and the borrower later defaults, the screener may be blamed for being too lenient. By contrast, if the screener raises the interest rate, the private downside is smaller. The borrower may face harsher terms, and the platform may lose some surplus, but the screener is less exposed to blame.

To capture this asymmetric accountability, suppose the human screener chooses the distortion term b_i after forming the informational correction s_i. The screener's private loss is

$$L_H(b_i) = \frac{1}{2}b_i^2 - \lambda q_i b_i.$$

The first term captures the cost of deviating from the efficient correction. Large distortions may require justification, attract scrutiny, or create procedural costs.

The second term captures the private benefit of defensive conservatism. The parameter $\lambda > 0$ measures the strength of the screener's career concern, ambiguity aversion, or blame-avoidance motive. The variable $q_i \geq 0$ captures the perceived blame exposure associated with borrower i. Blame exposure may be higher when the borrower has an ambiguous profile, when the loan amount is large, when recent platform defaults are salient, or when macroeconomic and platform-level default conditions make errors more costly.

The screener chooses b_i to minimize private loss:

$$\min_{b_i} \left\{ \frac{1}{2}b_i^2 - \lambda q_i b_i \right\}.$$

The first-order condition is

$$b_i - \lambda q_i = 0.$$

Thus, the privately optimal distortion is

$$b_i^* = \lambda q_i.$$

Substituting this expression into the human-adjusted rate gives

$$r_i^H = r_i^A + s_i + \lambda q_i.$$

This expression summarizes the main mechanism in the extended model. Human verification contains an informational component, s_i, and a defensive markup, λq_i.

The informational component reflects the screener's assessment of residual risk not fully incorporated into the algorithmic rate and therefore helps explain why human-induced interest-rate adjustments predict subsequent default outcomes. The defensive markup reflects agency concerns and blame avoidance and therefore helps explain why human verification raises interest rates on average.

Consequently, the model can simultaneously rationalize the two central empirical findings of the paper: verified borrowers receive higher interest rates on average, and human-induced interest-rate adjustments contain predictive information about ex-post default.

AI uncertainty and algorithm aversion

The distortion component can also arise from uncertainty about the algorithmic recommendation itself. Whereas the previous subsection emphasized agency concerns and defensive screening, this interpretation focuses on uncertainty surrounding the AI system.

The algorithm provides a fast and scalable estimate of credit risk, but the human screener may not observe the full algorithmic model or may not understand why the algorithm assigned a particular interest rate. This lack of transparency can make the algorithmic recommendation appear uncertain or incomplete. In such cases, human screeners may be reluctant to rely exclusively on the algorithmic rate and may instead apply an additional cautionary markup.

Let σ_i^A denote the perceived uncertainty associated with the algorithmic rate. This uncertainty may be high when the borrower profile is unusual, when similar past borrowers are sparse, when borrower characteristics send conflicting signals, or when the AI-assigned rate deviates substantially from a simple rule-of-thumb benchmark.

In this case, the updated distortion component can be written as

$$b_i^* = \lambda q_i + \phi \sigma_i^A,$$

where $\phi > 0$ captures the human screener's aversion to algorithmic uncertainty.

The final human-adjusted rate becomes

$$r_i^H = r_i^A + s_i + \lambda q_i + \phi \sigma_i^A.$$

This formulation highlights two distinct sources of upward adjustment. The first source, λq_i, reflects agency concerns and defensive screening. The second source, $\phi \sigma_i^A$, reflects uncertainty about the reliability or interpretability of the algorithmic recommendation.

When the algorithmic recommendation is perceived as more uncertain or less explainable, human screeners may require a larger cautionary markup. As a result, human intervention may raise interest rates even when the informational correction itself is modest.

At the same time, the informational component s_i remains embedded in the human adjustment. Consequently, human-induced interest-rate changes may continue to contain predictive information about ex-post default outcomes even when part of the adjustment reflects uncertainty about the AI system.

The agency-based and uncertainty-based interpretations suggest that human intervention may contain both informational and distortive components. The informational component helps explain why human adjustments predict subsequent default outcomes, while the distortive component helps explain why verification raises interest rates on average. This framework provides a unified interpretation of the paper's two central empirical findings.

B. Appendix: Policy Implications

Our findings have important implications for the design of hybrid human–AI decision systems. A common view is that human oversight is valuable primarily because humans can contribute information unavailable to algorithms. Our evidence suggests a different mechanism. Even when human screeners rely on essentially the same observable borrower information, they generate interest-rate adjustments that contain incremental predictive information about subsequent default. This implies that the value of human oversight need not arise from access to additional information, but can instead arise from alternative processing of existing information.

This distinction has important implications for the allocation of decision authority in AI-assisted systems. Rather than viewing algorithmic and human decisions as substitutes, platforms should recognize that they may extract different predictive signals from the same underlying data. Accordingly, the objective of hybrid system design should not be to maximize the frequency of human overrides, nor to eliminate them altogether, but to identify the settings in which human judgment consistently contributes incremental predictive value beyond the algorithmic benchmark.

The results also suggest that human intervention should be evaluated differently from conventional post-decision review. Because the predictive content of human adjustments reflects superior screening rather than the causal effect of higher interest rates on borrower outcomes, human overrides should be interpreted as additional risk signals rather than merely discretionary pricing decisions. In practice, this suggests that platforms may benefit from incorporating human adjustments into subsequent model updating and algorithm retraining, allowing the AI system to learn systematically from human judgments that prove informative ex post while reducing reliance on purely discretionary intervention over time.

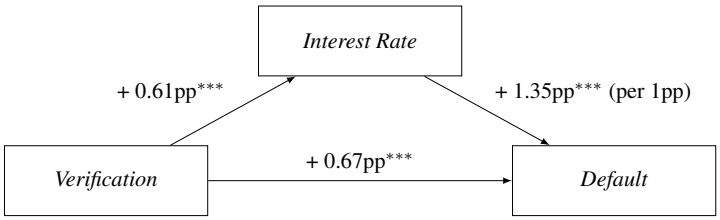
Our findings point to a shift in how human oversight should be incorporated into AI governance. Existing policy discussions often emphasize human review as a safeguard against algorithmic errors. Our evidence instead suggests that the primary benefit of human involvement may lie in generating complementary interpretations of the same information. Designing effective hybrid systems therefore requires identifying when human judgments contain genuine incremental predictive information and when they instead reflect behavioral biases, such as excessive conservatism or algorithm aversion. Distinguishing between these two channels is essential for allocating decision authority efficiently in high-stakes financial decisions.

Finally, these findings have broader implications for AI-assisted decision-making beyond credit markets. Many high-stakes applications—including insurance underwriting, hiring, medical diagnosis, and judicial decision-making—combine algorithmic recommendations with discretionary human review. In such settings, the central policy question is not whether humans or algorithms should make decisions, but how institutions should combine multiple modes of information processing to improve predictive performance while limiting systematic distortions. Our results suggest that the greatest gains from hybrid intelligence may come not from expanding human discretion, but from designing institutions that identify, preserve, and systematically learn from the aspects of human judgment that add predictive value.

C. Appendix: Figures

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

This figure visualizes the mediation analysis reported in Table G. All values are expressed in percentage points (pp), and standard errors (in parentheses) are double-clustered by issue year-quarter and matched cohort. All estimates use the default-observed subsample (N= 166,798). The *Interest Rate* → *Default* effect is equal to the *Interest Rate* coefficient (+ 1.35pp per 1pp) in Column (4), indicating that a 1 percentage point increase in the interest rate changes the default probability by + 1.35 percentage points. The direct effect is the *Verification* coefficient in Column (4). The indirect effect (*Verification* → *Interest Rate* → *Default*) is computed as the product of the *Verification* coefficient (*Verification* → *Interest Rate*) in Column (1) and the *Interest Rate* coefficient (*Interest Rate* → *Default*) in Column (4), with its standard error obtained via the delta method assuming zero covariance between the two coefficients. The total effect (*Verification* → *Default*) is equal to the *Verification* coefficient in Column (3). *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.



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

D. Appendix: Tables

Table A. Summary Statistics

This table reports summary statistics for the main variables used in the analysis. *Interest Rate* is the annualized loan interest rate assigned at origination, and *Verification* is an indicator equal to one if the borrower underwent additional human screening. *Default* equals one if the loan is classified as charged off or default, and zero otherwise. *Annual Income* and *Loan Amount* are reported in U.S. dollars. *DTI* is the debt-to-income ratio, *Revolving Utilization* is the ratio of revolving balance to credit limit, and *Credit History* is measured in years since the earliest credit line. The sample consists of 2,712,230 accepted loan applications from LendingClub over 2007Q2–2020Q3.

Variable	Mean	Std. Dev.	Median	Min	Max
Interest Rate (%)	0.13	0.05	0.12	0.05	0.31
Verification	0.64	0.48	1	0	1
FICO Score	701	33	692	627	848
Annual Income ($)	83,792	49,201	71,798	20,246	300,917
Debt-to-Income Ratio (%)	0.18	0.08	0.18	0.02	0.38
Loan Amount ($)	15,884	9,765	13,371	601	43,133
Revolving Credit Utilization (%)	0.49	0.25	0.49	0.01	0.98
Credit History (Years)	16.28	7.52	15	4	40
Employment Length (Years)	5.86	3.73	6	0	10
Number of Open Accounts	11.67	5.51	11	3	30
Credit Inquiries (Last 6 Months)	0.55	0.81	0	0	3
Delinquencies (Past 2 Years)	0.28	0.70	0	0	4
Public Records	0.17	0.42	0	0	2
60-Month Term	0.29	0.45	0	0	1
Loan Default	0.20	0.40	0	0	1

LendingClub classifies loan outcomes into several categories, including “fully paid,” “charged off” (clearly non-performing), “default” (loans overdue by more than 120 days that are generally classified as non-performing loans), and ongoing repayment statuses, including “delayed” (or overdue). Importantly, interest rates are assigned at origination and are observed for all loans regardless of subsequent performance. In our analysis, we define a *default* indicator that equals one if a loan is classified as either “charged off” or “default,” both of which correspond to states in which repayment is effectively not completed. It takes the value of zero when the loan status is “fully paid.”

Table B. Matching Rule

This table reports the matching tolerances applied to each covariate in the multivariate exact matching procedure. Cohorts are defined as groups of borrowers sharing identical values across all discretized characteristics. The matched sample is restricted to cohorts containing at least one verified and one non-verified borrower, yielding a final sample of 334,900 matched borrowers.

Variable	**Matching Tolerance**
FICO Score	± 5 points
Log(Annual Income)	± 0.3 Std. Dev.
Debt-to-Income Ratio	± 5 percentage points
Log(Loan Amount)	± 0.1 Std. Dev.
Revolving Credit Utilization	± 1.5 percentage points
Credit History	± 2 years
Employment Length (Years)	± 1 years
Number of Open Accounts	± 2
Credit Inquiries (Last 6 Months)	Exact match (0, 1, 2, 3)
Delinquencies (Past 2 Years)	Exact match (0, 1, 2, 3, 4)
Public Records	Exact match (0, 1, 2)

We construct cohorts using a multivariate caliper-based exact matching procedure. Each non-verified borrower serves as the anchor observation of a cohort. Verified borrowers are assigned to the same cohort if they satisfy all matching criteria reported in Table B. Specifically, matching requires exact agreement on recent credit inquiries (past six months), delinquencies (past two years), and public records, while all remaining continuous covariates must fall within the corresponding caliper thresholds. Matching is performed without imposing one-to-one restrictions, so each cohort may contain one non-verified borrower and multiple verified borrowers. Observations that do not belong to a cohort containing at least one verified and one non-verified borrower are excluded from the matched sample.

To improve computational efficiency without affecting the matching results, we partition the data into bins defined by the exact-match variables and coarse partitions of FICO score, annual income, and loan amount. Candidate matches are searched only within adjacent partitions that fully encompass the specified calipers, after which all matching conditions are re-evaluated using the original continuous variables. This partitioning strategy is purely computational and does not alter the set of matched observations. The final matched sample contains 334,900 borrowers.

Table C. Covariate Balance: Pre- and Post-Matching

This table reports mean covariate values for non-verified (NV) and verified (V) borrowers before and after multivariate exact matching. The pre-matching sample consists of 2,712,230 loan applications, and the matched sample comprises 334,900 borrowers (134,711 NV and 200,189 V). *Difference* reports the V−NV mean difference; statistical significance is based on Welch's two-sample *t*-tests with unequal variances. *SMD* denotes the absolute standardized mean difference; values below 0.10 are conventionally regarded as well-balanced. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Pre-Matching				Post-Matching			
	NV	V	Difference	SMD	NV	V	Difference	SMD
Observations	977,269	1,734,961			134,711	200,189		
Interest Rate	0.118	0.137	0.019***	0.414	0.121	0.131	0.010***	0.228
FICO Score	708.606	697.400	-11.206***	0.337	700.094	697.920	-2.174***	0.081
Log(Annual Income)	11.161	11.210	0.049***	0.094	11.134	11.130	-0.004***	0.011
Debt-to-Income Ratio	0.179	0.185	0.005***	0.064	0.180	0.182	0.002***	0.029
Log(Loan Amount)	9.315	9.534	0.220***	0.313	9.559	9.600	0.041***	0.082
Revolving Credit Utilization	0.452	0.515	0.062***	0.254	0.548	0.568	0.020***	0.097
Credit History	16.500	16.162	-0.339***	0.045	13.578	13.374	-0.204***	0.041
Employment Length	5.776	5.913	0.137***	0.037	6.183	6.420	0.237***	0.063
Number of Open Accounts	11.612	11.702	0.090***	0.016	10.219	10.065	-0.153***	0.041
Credit Inquiries (Last 6 Months)	0.476	0.591	0.115***	0.145	0.200	0.161	-0.039***	0.089
Delinquencies (Past 2 Years)	0.245	0.292	0.048***	0.069	0.053	0.040	-0.013***	0.052
Public Records	0.137	0.183	0.046***	0.111	0.046	0.036	-0.010***	0.051
60-Month Term	0.231	0.316	0.085***	0.192	0.297	0.354	0.057***	0.122

Table D. Randomness Check: Verification Status and Borrower Characteristics

This table reports OLS estimates from regressions of verification status on borrower characteristics. The dependent variable is an indicator equal to one if the borrower underwent human verification. Column (1) uses the full pre-matching sample of 2,712,230 loan applications; Column (2) uses the matched sample of 334,900 borrowers and absorbs matched-cohort fixed effects. In Column (2), Home Ownership and Loan Purpose are exact-match variables and are therefore fully absorbed by the matched cohort fixed effect. Standard errors in parentheses are clustered by issue year-quarter in Column (1) and double-clustered by issue year-quarter and matched cohort in Column (2), and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Verification	
	(1)	(2)
Sample:	Pre-Matching	Post-Matching
FICO Score	-0.002***	-0.005***
	(0.000)	(0.000)
Log(Annual Income)	-0.003	-0.100***
	(0.010)	(0.021)
Debt-to-Income Ratio	0.182***	0.542***
	(0.038)	(0.061)
Log(Loan Amount)	0.117***	0.376***
	(0.012)	(0.061)
Revolving Credit Utilization	0.037***	0.298
	(0.010)	(0.188)
Credit History	-0.002***	-0.008***
	(0.000)	(0.001)
Number of Open Accounts	-0.002***	-0.005***
	(0.000)	(0.001)
Employment Length	0.000	0.007*
	(0.001)	(0.004)
Credit Inquiries (Last 6 Months)	0.034***	
	(0.002)	
Delinquencies (Past 2 Years)	0.002	
	(0.002)	
Public Records	0.021***	
	(0.006)	
Observations	2,534,152	334,899
R-squared	0.088	0.104
F-statistic	205.601	49.459
Controls	Yes	Yes
Year-Quarter FE	Yes	Yes
State FE	Yes	Yes
Cohort FE	No	Yes

Table E. Human Verification and Interest Rates

This table reports OLS estimates of the effect of human verification on loan interest rates. The dependent variable is the annualized interest rate assigned at origination. *Human Verification* is an indicator equal to one if the borrower underwent additional human screening. Column (1) includes no fixed effects or controls; Column (2) adds controls and home ownership and loan purpose fixed effects; Column (3) adds year-quarter fixed effects; Column (4) adds state fixed effects; Column (5) further absorbs matched-cohort fixed effects. The sample consists of 334,900 matched borrowers. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. All columns also include home ownership and loan purpose fixed effects, which are absorbed by the matched cohort fixed effect in Column (5) since home ownership and loan purpose are exact-match variables. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

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 F. Default Risk: AI-Assigned Rates versus Human Adjustments

Panel A. Effect of Verification on Default Risk Beyond AI Prediction

This panel examines whether human verification affects loan default. The dependent variable equals one if the loan is classified as charged off (clearly non-performing) or default (overdue by more than 120 days), and zero if it is fully paid; loans in ongoing repayment status are excluded from the analysis. *Human Verification* is an indicator equal to one if the borrower underwent additional human screening. *AI-Assigned Interest Rate* is the OLS-predicted rate a non-verified borrower would receive given the same observable characteristics. Column (1) includes no controls or fixed effects; Column (2) adds controls and home ownership and loan purpose fixed effects; Column (3) additionally absorbs year-quarter and state fixed effects; Column (4) further absorbs matched-cohort fixed effects. Columns (5)–(6) add *AI-Assigned Interest Rate* alongside *Human Verification*, without and with matched-cohort fixed effects, respectively. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. All columns also include home ownership and loan purpose fixed effects, which are absorbed by the matched cohort fixed effect in Columns (4) and (6) since home ownership and loan purpose are exact-match variables. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Default					
	(1)	(2)	(3)	(4)	(5)	(6)
Human Verification	0.033***	0.013***	0.014***	0.015***	0.015***	0.016***
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
AI-Assigned Interest Rate					1.303***	1.148***
					(0.086)	(0.080)
Observations	216,048	216,048	216,047	166,798	216,047	166,798
R-squared	0.002	0.057	0.069	0.462	0.073	0.464
Controls	No	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	No	No	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	No	Yes	No	Yes

Table F. Default Risk: AI-Assigned Rates versus Human Adjustments (continued)

Panel B. Predictive Content of Human Adjustment

This panel examines the predictive content of human rate adjustments for loan default. The dependent variable is the same as in Panel A. *AI-Assigned Interest Rate* is the OLS-predicted rate a non-verified borrower would receive given the same observable characteristics. *Human Interest Rate Adjustment* is the deviation of the final assigned rate from the AI-assigned rate (zero for non-verified borrowers). *Upward Adjusted* (*Downward Adjusted*) equals one for verified borrowers whose assigned rate exceeds (falls below) the cohort non-verified mean; the 50bp variants require the deviation to exceed 50 basis points in absolute value. The reference group in all columns is non-verified borrowers. All columns include controls and home ownership, loan purpose, year-quarter, and state fixed effects; Columns (2), (4), (5), and (6) further absorb matched-cohort fixed effects. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

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

Table F examines whether human-induced interest rate adjustments contain incremental predictive information beyond the algorithmic baseline. Columns (1) and (2) show that the AI-assigned interest rate strongly predicts subsequent default, confirming that the algorithm captures economically meaningful variation in ex ante borrower risk. Although the coefficient is attenuated after introducing matched-cohort fixed effects in Columns (3) and (4), this attenuation is expected because comparisons are restricted to borrowers with identical observable credit profiles, leaving only within-cohort variation.

The key result is that human-induced interest rate adjustments remain strongly predictive of default. In Columns (3) and (4), larger positive adjustments are associated with higher subsequent default probabilities even after conditioning on the AI-assigned rate and matched-cohort fixed effects. Likewise, borrowers whose interest rates are adjusted upward exhibit significantly higher default rates, whereas those receiving downward adjustments exhibit significantly lower default rates, with similar patterns for the 50-basis-point threshold in Columns (5) and (6). These findings indicate that human intervention contains incremental predictive information beyond the algorithmic recommendation.

One interpretation is that human screeners extract additional information from the same observable borrower characteristics by processing them differently from the algorithm. Importantly, this predictive relationship should not be interpreted as evidence that higher interest rates themselves increase default. As shown in Table H, the predictive content of human adjustments is primarily consistent with superior screening of latent borrower risk rather than a causal pricing effect.

More broadly, these results suggest that human judgment can complement algorithmic predictions even when both rely on the same underlying information. Rather than merely overriding algorithmic recommendations, human intervention appears to identify residual risk that is not fully reflected in the AI-assigned interest rate.

Table G. Mediation: The Interest Rate Channel from Verification to Default

This table decomposes the effect of verification on default into direct and indirect paths. The dependent variable is the annualized interest rate in Column (1), the interest rate change in Column (2), and an indicator for default in Columns (3)–(5). *Human Verification* is an indicator equal to one if the borrower underwent additional human screening, and *Interest Rate* is the annualized rate assigned at origination. All columns are estimated on the default-observed subsample ($N = 166{,}798$) and absorb matched-cohort fixed effects. Column (1) estimates the effect of verification on interest rate (Path a); Column (2) estimates the total effect of verification on default; Column (3) jointly estimates the direct effect of verification and Path b, the effect of interest rate on default adjusted for verification. The indirect effect equals the product of the *Verification* coefficient in Column (1) and the *Interest Rate* coefficient in Column (3), with its standard error obtained via the delta method. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. Home ownership and loan purpose fixed effects are absorbed by the matched cohort fixed effect since they are exact-match variables. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Interest Rate	Interest Rate Change	Default	Default	Default
	(1)	(2)	(3)	(4)	(5)
Human Verification	0.006***	0.006***	0.015***	0.007***	0.007***
	(0.000)	(0.000)	(0.002)	(0.002)	(0.002)
Interest Rate				1.352***	
				(0.049)	
Human Interest Rate Adjustment					1.352***
					(0.049)
Observations	166,798	166,798	166,798	166,798	166,798
R-squared	0.696	0.689	0.462	0.470	0.470
Controls	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes

Table H. Causal Effect of Human Adjustment on Default

This table reports two-stage least squares (2SLS) estimates of the effect of interest rates on loan default. The instrument is the LendingClub funding shock in May 2016. Specifically, the excluded instrument is *Post*, an indicator equal to one for loans originated after the funding shock and zero for loans originated during the one-year period preceding the event. The dependent variable is an indicator for default in all columns. The reported coefficients are from the second-stage regressions only. Column (1) instruments the assigned interest rate. Columns (2)–(5) instrument the human interest rate adjustment, defined as the difference between the final assigned interest rate and the AI-assigned interest rate. The AI-assigned interest rate is included as a control variable in Columns (3)–(5). Column (4) additionally includes grade fixed effects, and Column (5) instead includes matched-cohort fixed effects. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. Standard errors are reported in parentheses. The first-stage F-statistics are reported at the bottom of the table. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Default				
	(1)	(2)	(3)	(4)	(5)
Interest Rate	0.696***				
	(0.240)				
Human Interest Rate Adjustment		1.062***	-0.571	0.080	-1.327
		(0.368)	(0.405)	(0.301)	(1.459)
AI-Assigned Interest Rate			1.319***	0.146	2.336
			(0.125)	(0.218)	(1.393)
Observations	92,464	92,464	92,464	92,464	37,870
R-squared	0.076	0.068	0.055	0.018	0.003
F-stat (1st Stage)	117.06	88.80	69.81	136.95	54.76
Controls	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Grade FE	No	No	No	Yes	No
Cohort FE	No	No	No	No	Yes

This Table reports the second-stage estimates from the 2SLS analysis. We exploit the LendingClub funding shock in May 2016 as an exogenous source of variation in interest rates. The first-stage results (not reported) show that the funding shock strongly predicts both the assigned interest rate and the human interest rate adjustment, with first-stage F-statistics well above conventional thresholds.

The IV estimates indicate that exogenous increases in interest rates do not significantly affect subsequent default once the AI-assigned interest rate is held constant. In particular, the coefficients on the instrumented human interest rate adjustment in Columns (3)–(5) are economically small and statistically insignificant. This finding suggests that the positive correlation between human interest rate adjustments and default in the baseline OLS results is not primarily driven by a causal pricing effect. Instead, it is consistent with human underwriters identifying borrowers with higher latent default risk and adjusting interest rates accordingly. In other words, the predictive value of human intervention arises mainly from superior risk selection rather than from changes in borrowers' repayment behavior induced by higher interest rates.

Table I. Verification Effect on Interest Rates (Low Credit Quality Subsample)

This table reestimates Table E on the low credit quality subsample (Grades D, E, F, and G). The dependent variable is the annualized interest rate, and *Human Verification* is an indicator equal to one if the borrower underwent additional human screening. Column (1) includes no fixed effects or controls; Column (2) adds controls and home ownership and loan purpose fixed effects; Column (3) adds year-quarter fixed effects; Column (4) adds state fixed effects; Column (5) further absorbs matched-cohort fixed effects. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. All columns also include home ownership and loan purpose fixed effects, which are absorbed by the matched cohort fixed effect in Column (5) since these are exact-match variables. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Interest Rate				
	(1)	(2)	(3)	(4)	(5)
Huamn Verification	0.002*	0.001	0.005***	0.005***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	58,135	58,135	58,133	58,133	23,317
R-squared	0.001	0.036	0.239	0.241	0.586
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 J. Default Risk: AI-Assigned Rates versus Human Adjustments (Low Credit Quality)

Panel A. Effect of Verification on Default Risk Beyond AI Prediction

This table replicates Table F on the low credit quality subsample (Grades D, E, F, and G) with observed default outcomes. The dependent variable is an indicator for default. *Human Verification* is an indicator equal to one if the borrower underwent additional human screening. *AI-Assigned Interest Rate* is the OLS-predicted rate a non-verified borrower would receive given the same observable characteristics. Column (1) includes no controls or fixed effects; Column (2) adds controls and home ownership and loan purpose fixed effects; Column (3) additionally absorbs year-quarter and state fixed effects; Column (4) further absorbs matched-cohort fixed effects. Columns (5)–(6) add *AI-Assigned Interest Rate* alongside *Human Verification*, without and with matched-cohort fixed effects, respectively. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. All columns also include home ownership and loan purpose fixed effects, which are absorbed by the matched cohort fixed effect in Columns (4) and (6) since home ownership and loan purpose are exact-match variables. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Default					
	(1)	(2)	(3)	(4)	(5)	(6)
Human Verification	0.032***	0.014*	0.016***	0.025**	0.039***	0.053***
	(0.008)	(0.007)	(0.006)	(0.011)	(0.008)	(0.018)
AI-Assigned Interest Rate					0.459***	0.618*
					(0.162)	(0.325)
Observations	41,348	41,348	41,346	13,275	41,346	13,275
R-squared	0.001	0.026	0.045	0.506	0.046	0.506
Controls	No	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	No	No	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	No	Yes	No	Yes

Table J. Default Risk: AI-Assigned Rates versus Human Adjustments (Low Credit Quality) (continued)

Panel B. Predictive Content of Human Adjustment

This table replicates Table F on the low credit quality subsample (Grades D, E, F, and G) with observed default outcomes. The dependent variable is the same as in Panel A. *AI-Assigned Interest Rate* is the OLS-predicted rate a non-verified borrower would receive given the same observable characteristics. *Human Interest Rate Adjustment* is the deviation of the final assigned rate from the AI-assigned rate (zero for non-verified borrowers). *Upward Adjusted* (*Downward Adjusted*) equals one for verified borrowers whose assigned rate exceeds (falls below) the cohort non-verified mean; the 50bp variants require the deviation to exceed 50 basis points in absolute value. The reference group in all columns is non-verified borrowers. All columns include controls and home ownership, loan purpose, year-quarter, and state fixed effects; Columns (2), (4), (5), and (6) further absorb matched-cohort fixed effects. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Default					
	(1)	(2)	(3)	(4)	(5)	(6)
AI-Assigned Interest Rate	-0.056	-0.169	0.296**	-0.351*	0.356	0.078
	(0.109)	(0.199)	(0.115)	(0.204)	(0.331)	(0.320)
Human Interest Rate Adjustment			0.412***	1.260***		
			(0.042)	(0.181)		
Upward Adjusted					0.065***	
					(0.017)	
Downward Adjusted					0.015	
					(0.023)	
Upward Adjusted (50bp)						0.058***
						(0.017)
Downward Adjusted (50bp)						-0.010
						(0.024)
Observations	41,346	13,275	41,346	13,275	13,275	13,275
R-squared	0.045	0.506	0.047	0.509	0.507	0.507
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

Table K. Verification Effect on Interest Rates (High Credit Quality Subsample)

This table reestimates Table E on the high credit quality subsample (Grade A). The dependent variable is the annualized interest rate, and *Human Verification* is an indicator equal to one if the borrower underwent additional human screening. Column (1) includes no fixed effects or controls; Column (2) adds controls and home ownership and loan purpose fixed effects; Column (3) adds year-quarter fixed effects; Column (4) adds state fixed effects; Column (5) further absorbs matched-cohort fixed effects. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. All columns also include home ownership and loan purpose fixed effects, which are absorbed by the matched cohort fixed effect in Column (5) since these are exact-match variables. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Interest Rate				
	(1)	(2)	(3)	(4)	(5)
Human Verification	0.001	0.001	0.001***	0.001***	0.001***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	75,410	75,410	75,410	75,410	34,551
R-squared	0.001	0.121	0.369	0.369	0.683
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 L. Default Risk: AI-Assigned Rates versus Human Adjustments (High Credit Quality)

Panel A. Effect of Verification on Default Risk Beyond AI Prediction
This table replicates Table F on the high credit quality subsample (Grade A) with observed default outcomes. The dependent variable is an indicator for default. *Human Verification* is an indicator equal to one if the borrower underwent additional human screening. *AI-Assigned Interest Rate* is the OLS-predicted rate a non-verified borrower would receive given the same observable characteristics. Column (1) includes no controls or fixed effects; Column (2) adds controls and home ownership and loan purpose fixed effects; Column (3) additionally absorbs year-quarter and state fixed effects; Column (4) further absorbs matched-cohort fixed effects. Columns (5)–(6) add *AI-Assigned Interest Rate* alongside *Human Verification*, without and with matched-cohort fixed effects, respectively. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. All columns also include home ownership and loan purpose fixed effects, which are absorbed by the matched cohort fixed effect in Columns (4) and (6) since home ownership and loan purpose are exact-match variables. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Default					
	(1)	(2)	(3)	(4)	(5)	(6)
Human Verification	0.002	-0.000	0.002	0.004	-0.005	0.003
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.006)
AI-Assigned Interest Rate					0.300	0.053
					(0.200)	(0.256)
Observations	40,354	40,354	40,354	12,194	40,354	12,194
R-squared	0.000	0.016	0.025	0.496	0.026	0.496
Controls	No	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	No	No	Yes	Yes	Yes	Yes
State FE	No	No	Yes	Yes	Yes	Yes
Cohort FE	No	No	No	Yes	No	Yes

Table L. Default Risk: AI-Assigned Rates versus Human Adjustments (High Credit Quality) (continued)

Panel B. Predictive Content of Human Adjustment

This table replicates Table F on the high credit quality subsample (Grade A) with observed default outcomes. The dependent variable is the same as in Panel A. *AI-Assigned Interest Rate* is the OLS-predicted rate a non-verified borrower would receive given the same observable characteristics. *Human Interest Rate Adjustment* is the deviation of the final assigned rate from the AI-assigned rate (zero for non-verified borrowers). *Upward Adjusted* (*Downward Adjusted*) equals one for verified borrowers whose assigned rate exceeds (falls below) the cohort non-verified mean; the 50bp variants require the deviation to exceed 50 basis points in absolute value. The reference group in all columns is non-verified borrowers. All columns include controls and home ownership, loan purpose, year-quarter, and state fixed effects; Columns (2), (4), (5), and (6) further absorb matched-cohort fixed effects. Control variables include FICO score, log annual income, DTI, log loan amount, revolving utilization, credit history, open accounts, credit inquiries, delinquencies, public records, employment length, and loan term. Standard errors in parentheses are double-clustered by issue year-quarter and matched cohort, and *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dep. Variable:	Default					
	(1)	(2)	(3)	(4)	(5)	(6)
AI-Assigned Interest Rate	0.173	0.145	0.284*	-0.014	-0.069	-0.027
	(0.128)	(0.207)	(0.149)	(0.223)	(0.270)	(0.260)
Human Interest Rate Adjustment			0.120**	0.994***		
			(0.058)	(0.352)		
Upward Adjusted					0.014*	
					(0.008)	
Downward Adjusted					-0.004	
					(0.007)	
Upward Adjusted (50bp)						0.018**
						(0.009)
Downward Adjusted (50bp)						-0.016**
						(0.008)
Observations	40,354	12,194	40,354	12,194	12,194	12,194
R-squared	0.025	0.496	0.026	0.497	0.496	0.497
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