JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 60, No. 4, June 2025, pp. 1727–1759

doi:10.1017/S0022109024000565

Attention Constraints and Financial Inclusion

Bo Huang
Remin University School of Finance
bohuang@ruc.edu.cn

Jiacui Li
University of Utah David Eccles School of Business
jiacui.li@eccles.utah.edu (corresponding author)

Tse-Chun Lin
Zhejiang University School of Economics, and University of Hong Kong Faculty of Business and Economics
tsechunlin@hku.hk

Mingzhu Tai
University of Hong Kong Faculty of Business and Economics
taimzh@hku.hk

Zoey Yiyuan Zhou
Hong Kong University of Science and Technology Academy of Interdisciplinary Studies Division of Environment and Sustainability
zoeyzhou@ust.hk

Abstract

We show that attention constraints on decision-makers create barriers to financial inclusion. Using administrative data on retail loan-screening processes, we find that attention-constrained loan officers exert less effort reviewing applicants of lower socioeconomic status (SES) and reject them more frequently. More importantly, when externally imposed increases in loan officers' workloads tighten attention constraints, loan officers are even more prone to quickly reject low-SES applicants but quickly accept very high-SES applicants without careful review. Such selective attention allocation further widens the approval rate gap between high- and low-SES applicants—a unique prediction of this attention-based mechanism.

I. Introduction

Having access to basic financial services is crucial to one's well-being in contemporary society.[1] Even in the United States, however, nearly one fifth of

The authors appreciate feedback from Sumit Agarwal, Patrick Bayer, Tobias Berg, Douglas Bernheim, Phillip Bond, Patrick Button, Lynn Cornell-Price, Shaun Davies, Daniel Ferreira, Matthew Gentzkow, Troup Howard, Tingyan Jia, Timothy McQuade, Filip Matejka, Jonathan Parker, Wenlan Qian, and Alminas Zaldokas as well as participants at the AEA Discrimination and Disparities in Credit and Housing Markets Session, Colorado Leeds, Consumer Financial Protection Bureau Research Conference (6th), Econometric Society meetings (Asia+Australia+China+Africa), European Economic Association – European meeting of the Econometric Society Meeting, Great Bay Area Finance Conference, NBER Household Finance Working Group Meeting, China Financial Research Conference, Online Seminar on the Economics of Discrimination and Disparities, Utah Behavioral Lab, Utah Eccles, and Washington Foster. Hyun Joong Kim provided excellent research assistance. Authors are listed alphabetically and have contributed equally to this work.

[1]It has been found that financial inclusion plays an important role in determining human capital investment (e.g., Stein and Yannelis (2020)), wealth accumulation (e.g., Célerier and Matray (2019)), long-term financial health (e.g., Brown, Cookson, and Heimer (2019)), etc.

adults remain unbanked or underbanked, and there exist significant barriers to financial inclusion for those of lower socioeconomic status (SES).[2] In this study, we show that *attention constraints* on key decision-makers, such as loan officers, can further restrict financial inclusion of and resource allocation to low-SES borrowers even when many are qualified for financial access.

How do attention constraints impact inclusion? We use a simple model to illustrate a mechanism that might explain this relationship. If loan officers had infinite time, they would review every application carefully and make informed decisions based on borrower credit quality. If loan officers are attention-constrained, however, they may choose to ration their attention based on easily observed signals, such as labels that indicate borrowers' SES status. Our model predicts that, in the *base case*, when loan officers face tighter attention constraints, they will allocate disproportionately less attention to low-SES borrowers, leading to "rash rejections," even when a significant portion of those borrowers are qualified and should otherwise be approved if adequate attention were paid to their applications. Moreover, in the *special case* where some borrowers enjoy extremely high SES status, they may even experience the opposite fate and be "rashly accepted" without careful screening. As a result, tighter loan officer attention constraints will widen the inclusion gap between high- and low-SES applicants. We want to emphasize that such loan officer behavior can be fully rational and constrained optimal from the perspective of the lender. This behavior does, however, create distributional consequences for financial inclusion on low-SES borrowers, which is what we focus on in this paper.

We face two challenges in studying the impact of attention on inclusion empirically. First, it is difficult to measure attention allocation, as noted by Gabaix (2019): "measuring attention is … a hard task—we still have only a limited number of papers that measure attention in field settings." Second, we need variations in attention constraints that are orthogonal to candidate fundamentals. Using administrative data on the screening processes associated with approximately 146,000 retail loans in one of the largest national banks in China, our paper overcomes both difficulties. First, as we observe accurate timestamps in the decision-making process, we can track the amount of time that loan officers spend reviewing each application—a direct measure of attention allocation.

To overcome the second challenge, we utilize a unique institutional feature of our sample: The bank allocates applications across loan officers using an algorithm that induces variations in loan officer workloads. Such variations are orthogonal to borrower creditworthiness and loan-officer behaviors, enabling us to identify the consequences of attention constraints. To measure loan officer attention constraints, we begin by exploiting variations in loan officer *busyness*, which is defined as the number of applications processed by an officer on a given day. We then introduce 2 instrumental variables to capture variations in loan officer busyness that are orthogonal to loan officer behaviors and borrower credit quality.

[2]Source: Report on the Economic Well-Being of U.S. Households in 2020. According to the report, almost half of all families with annual incomes below $50,000 experienced credit denials or could not obtain sufficient credit. Across all income levels, on average about one third reported experiencing difficulty obtaining credit.

When loan officers are busy, they consider certain salient socioeconomic labels as simple signals to guide their attention allocation to each borrower. In each application form, the front page reports a number of indicators related to an applicant's SES, such as whether the applicant is a local resident (rather than a migrant), a public employee (i.e., employed by a government agency or a state-owned firm), a worker with stable long-term employment and income, and/or a homeowner. Based on these labels, applicants are sorted into high- and low-SES groups.

Our attention-driven mechanism suggests that, when loan officers are time-constrained, they allocate disproportionately less time to reviewing loan applications from low-SES applicants, which results in disproportionately lower financial inclusion for those borrowers. Our empirical findings support this main hypothesis. We first plot these patterns in Figure 1 as an exploratory illustration. Graphs A and B show that, when officers are busier, their disproportionate reduction in attention results in a wider attention gap between the high- and low-SES groups. Graphs C and D further show a sharper decline in approval rates for low-SES applicants, suggesting that some are rejected rashly by busy loan officers. This is consistent with the *base case* predictions associated with this attention-driven mechanism. Also, the approval rate for high-SES applicants seems to increase slightly when loan officers are busier, which would be consistent with the *special case* where some extremely high-SES applications are "rashly approved."

We then formally test this attention-driven mechanism using difference-in-differences regression analyses, in which we control for officer-month-year, week, bank branch, and loan-type fixed effects as well as a comprehensive list of applicant characteristics. The results are similar to those plotted in Figure 1. When loan officer busyness varies from the bottom to the top decile, the review time they spend on low-SES applicants declines by 53% (52%), which is larger than the 38% (37%) decline observed for high-SES applicants. More importantly, the approval rate for the low-SES group drops by 45% (39%) relative to the average levels, while approval rates for the high-SES groups were not significantly affected and in some cases even increased slightly.

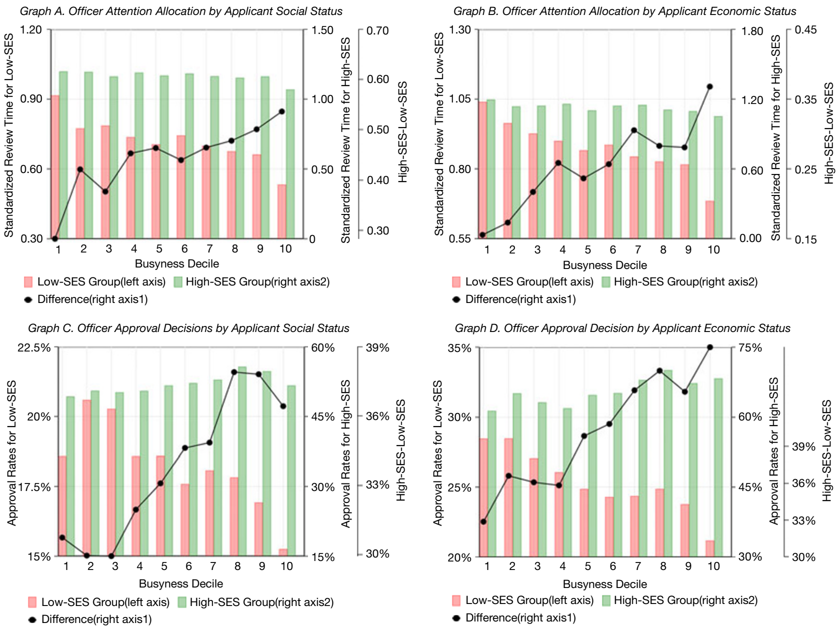
Two empirical concerns might arise in using *realized* loan officer busyness to measure attention constraints. First, loan officers may have leeway to work faster or more slowly, so realized busyness may reflect an endogenous choice rather than an external constraint. To address this concern, we instrument the busyness measure by the number of applications *assigned* to officers. Because applications are assigned by a central dispatcher algorithm over which officers have no influence, the process generates externally imposed variations in loan officers' attention constraints. Furthermore, conditioning on week, branch, loan type, and loan officer-month-year fixed effects in all our specifications, we are effectively utilizing the idiosyncratic assignment variation that is unrelated to loan officer preferences or systematic shifts in risk-management criteria over time.

The second empirical concern is that loan officer busyness may be correlated with unobserved borrower quality. We should note that our results can be biased only if loan officer busyness and work assignments are *negatively correlated* with the credit quality of low-SES borrowers but *uncorrelated or even positively*

FIGURE 1

Attention Allocation and Approval Decisions by Loan Officer Constraints

Figure 1 exhibits results indicating how loan officer attention allocation and approval decisions for high- and low-SES applicants vary with officer attention constraints. As explained in Section III.C, we use the possession (or not) of various labels to classify applicants into high- and low-SES groups based on social status (Graphs A and C) or economic status (Graphs B and D). For all 4 graphs, we sort the sample into deciles by officer attention constraints measured by *busyness*, which is defined as the number of applications processed per day. Graphs A and B plot the average officer attention allocation, measured as the standardized review time for each loan in the screening process, by busyness decile. Graphs C and D plot the average loan approval rate by busyness decile. The measurement of standardized review time is explained in Section IV.A. Each red (green) bar graphs the average for the low- (high-) SES applicant groups. The black line plots the differences between the two groups.



correlated with the credit quality of high-SES borrowers, which is very unlikely in practice. Nevertheless, to alleviate this concern, we first verify that the assignment-instrumented busyness measure is orthogonal to a comprehensive list of credit-worthiness metrics. We then further address this concern by constructing another leave-one-out (LOO) Bartik-type instrument. The idea is as follows: if a loan officer's attention constraint is tightened following an idiosyncratic spike in application volume from Province A, her decision-making on applications from Province B can also be affected, even if there is no change in either the quantity or quality of applications from Province B. In this sense, by utilizing variations in officer busyness that are driven by assignments from *other* provinces and directly controlling for busyness driven by assignments from the *local* province, we capture variations in a loan officer's attention constraints that are orthogonal to a particular application she is screening. We find that, when using the abovementioned instrumental variables, our main results are qualitatively and quantitatively robust.

This paper's main contribution to the literature lies in showing that decision-maker attention constraints can have consequences for financial inclusion. This attention-driven mechanism suggests that, as long as attention-constrained decision-makers have small differential priors across SES groups, whether driven by prejudice or statistics, their prior-based disproportionate attention allocation can widen the inclusion gap between applicants with differing SES labels.

Regarding policy implications, our attention-based mechanism suggests that policies and technologies that relax decision-maker attention constraints may promote financial inclusion for borrowers from low-SES backgrounds. For instance, recent developments in financial technologies ("Fintech"), which use automated underwriting algorithms (and thus are subject to low or no attention constraints) to assist in screening borrowers, may improve financial access. Meanwhile, this use of technology could also change the way that soft and hard information are used in the screening process.[3,4] By the same logic, the introduction of credit scores such as the Fair Isaac Corporation (FICO) scores, which were not available in China during our sample period, may also help improve credit access for low-SES applicants. In addition, to the extent that loan officer specialization in screening applicants from specific backgrounds can make information processing more efficient, this may also improve outcomes for low-SES applicants. Finally, taking attention constraints more seriously also generates additional insights into optimal workload allocation. For instance, in our setting, while each loan officer may be acting optimally given their attention constraints, the bank is likely behaving suboptimally in distributing workloads unevenly across officers (see Appendix A.2).

There are two empirical limitations that we cannot fully address, but we believe that they do not invalidate our conclusions. First, we do not observe ex post default outcomes for our sample borrowers and thus cannot determine how loan officers' attention-driven behavior affects bank loan losses. As previously noted, though, it is entirely possible that loan officers' behavior is constrained optimal; we focus instead on the *distributional* consequences of their behavior.

Second, while we are assured by loan officers that the application-assignment algorithm is idiosyncratic (see Section IV.D for details), the bank does not disclose the exact algorithm. Nevertheless, we verify that the assignment-instrumented busyness measures are orthogonal to a comprehensive list of creditworthiness measures. More importantly, when loan officers are busier, their approval rates for high- and low-SES applicants often move in non-parallel or even *opposite* directions, which is a unique prediction of this attention-based mechanism.

Related Literature

Previous researchers have investigated how attention constraints impact financial decisions. For example, Hirshleifer, Levi, Lourie, and Teoh (2019) show that

[3]See Liberti and Petersen (2019) for a review of the literature.

[4]This rationale is aligned with the arguments of Philippon (2019), Dobbie, Liberman, Paravisini, and Pathania (2021), Bartlett, Morse, Stanton, and Wallace (2022), Fuster, Plosser, Schnabl, and Vickery (2019), and so forth Of course, there are also concerns that the application of big-data fintech may generate new distributional issues (e.g., Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2022)). Therefore, the net distributional effect of fintech might not be clear-cut (Morse and Pence (2021)).

financial analysts who suffer from fatigue resort to heuristic decisions when making forecasts. Huang, Huang, and Lin (2019) show that attention-constrained investors pay less attention to firm-specific news. Müller (2022) shows that bankruptcy-court congestion leads to lower recovery values in defaults and also impacts pre-default credit spreads.[5] Of greater relevance to lending decisions, Liao, Wang, Xiang, Yan, and Yang (2021) document that peer-to-peer investors tend to use "system one thinking" à la Kahneman (2011) under time pressure. In the broader literature on endogenous attention allocation,[6] our paper is related most closely to the seminal work on selective attention allocation by Bartoš, Bauer, Chytilová, and Matějka (2016). Our key innovation lies in testing the impact of decision-maker attention constraints and providing more direct field-based evidence pertaining to the distributional consequences of attention constraints.

This paper also contributes to the literature that investigates distributional issues in financial-resource allocation. Many studies have documented discriminatory practices in mortgage credit (Bayer, Ferreira, and Ross (2018); Bartlett et al. (2022); Giacoletti, Heimer, and Yu (2021); Ambrose, Conklin, and Lopez (2021); Huo, Sun, Tai, and Xuan (2024)), consumer credit (Montoya, Parrado, Solís, and Undurraga (2020); Dobbie et al. (2021)), bank lending (Fisman, Paravisini, and Vig (2017); Fisman, Sarkar, Skrastins, and Vig (2020)), auto loans (Charles, Hurst, and Stephens (2008); Butler, Mayer, and Weston (2023); Lanning (2021)), small business lending (Ongena and Popov (2016); Brock and De Haas (2023)), microlending (Beck, Behr, and Madestam (2018)), entrepreneurial finance (Hebert (2020); Ewens and Townsend (2020); Hu and Ma (2025); Zhang (2020)), and even housing returns (Goldsmith-Pinkham and Shue (2023); Kermani and Wong (2021)). Our findings provide empirical evidence of attention-based credit allocation that could conceivably function as one of the mechanisms underlying some of the findings reported in the aforementioned studies.[7]

The remainder of the paper is organized as follows: In Section II we present a simple model to illustrate our attention-based mechanism and derive testable predictions. In Section III we describe the data and relevant institutional details. In Section IV we present our main empirical results. In Section V we conclude the study.

II. Conceptual Framework

Building on Bartoš et al. (2016), we use a simple model to illustrate how decision-maker attention constraints can exacerbate financial inclusion concerns.

[5]Other papers that have applied the framework of endogenous attention allocation to financial settings include Peng (2005), Peng and Xiong (2006), Van Nieuwerburgh and Veldkamp (2010), Mondria (2010), Mondria and Quintana-Domeque (2012), Andrei and Hasler (2015), Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016), Hasler and Ornthanalai (2018), Huang et al. (2019), Liu, Peng, and Tang (2023), Shu, Tian, and Zhan (2022), and Hirshleifer and Sheng (2022).

[6]Gabaix (2019) and Mackowiak, Matejka, and Wiederholt (2023) provide extensive reviews.

[7]Interest in studying the distributional impact of machine learning and artificial intelligence has recently surged (Bartlett et al. (2022); Fuster et al. (2022); Jansen, Nguyen, and Shams (2025); D'Acunto, Ghosh, Jain, and Rossi (2021)).

Less interested readers can skip the model at little cost as we have explained the intuition in the introduction.

Model setup. Consider a risk-neutral loan officer faced with the task of deciding whether to approve an application to borrow 1 unit of capital for 1 period of time. Applicants come from a continuum of groups denoted by G, and the associated group identities are observable at zero cost. The officer makes 2 decisions: i) whether to incur an attention cost of c to learn more about the applicant, and ii) whether to approve or reject the application. Empirically, we think of the attention cost as the time and energy consumed in reading credit reports, scrutinizing the applicant's application forms, and so forth.

The interest rate $r > 0$ is fixed exogenously.[8] If the loan officer approves the application, the expected profit (before considering attention cost) is

$$-\text{distaste}_G + \underbrace{(1-p)\cdot r}_{\text{interest payments if paid back}} - \underbrace{p}_{\text{loss from default}} \tag{1}$$

where $\text{distaste}_G \geq 0$ captures group-specific *preference-based* attributes, and p is the default rate. For the sake of simplicity, equation (1) assumes a zero recovery rate upon default. We also assume risk neutrality and no time discounting. To ensure that lending can happen in equilibrium, we assume that $r > \text{distaste}_G$, so an applicant for whom the default probability is zero is worth lending to.

Apart from possible differences in distaste_G, groups can also differ in average credit quality. For every applicant, his default probability p decomposes into 2 components:

$$p = \overline{p}_G + p_I \tag{2}$$

where $\overline{p}_G$ is a group-specific average observable at no cost, and $p_I \sim N(0, \sigma^2)$ is an applicant-specific component that can be learned by paying the attention cost c.[9] As such, differences in distaste_G capture taste-based discrimination, and differences in $\overline{p}_G$ capture statistical discrimination.

Optimal loan officer behavior. We explain the solution intuitively and refer the reader to Appendix A for formal proofs. As illustrated in Graph A of Figure 2, loan officers can adopt any of 3 strategies. The optimal strategy falls into three regions and depends on the ex ante SES status of the group:

1. Applicants from *extremely high-SES groups* (low $\overline{p}_G$ or distaste_G) are *immediately approved* without information acquisition. While paying attention to them can enable the loan officer to screen out the occasional bad borrower, the probability that this occurs is sufficiently low that the cost of attention outweighs the benefit.
   - In our empirical setting, this should apply to only a very small subset of borrowers. The average approval rate over the full sample is only 34%,

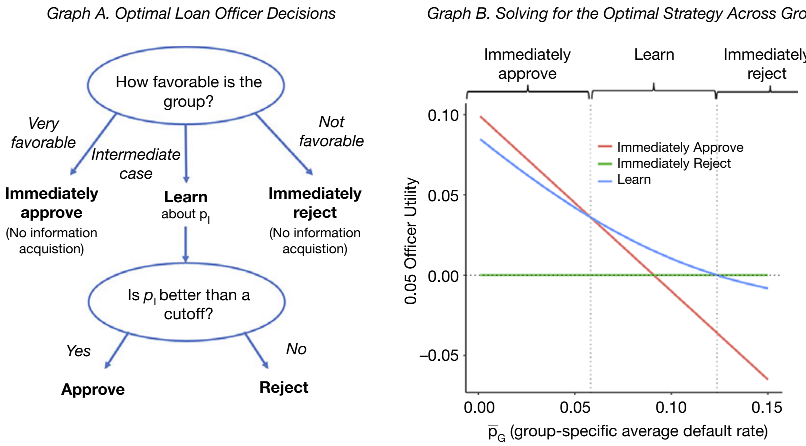
[8] This is true in our empirical setting. The loan officer needs to decide only whether to approve or reject the application.

[9] Technically, using normal distributions for p_I can lead to default rates above 1 or below 0. The results are qualitatively unchanged if we use modified zero-mean distributions with bounded support such as truncated normal distributions.

FIGURE 2

Illustration of The Model

Figure 2 Graph A shows the optimal loan officer decision process. At stage 1, the officer decides whether to incur attention cost c to learn applicant-specific quality information p_l, given knowledge of the applicant's group. Conditional on doing so, at stage 2 the officer decides whether to approve or reject the application. In Graph B, we plot the expected loan officer utility associated with the 3 strategies – immediately approve (red line), immediately reject (green line), and learn before making a decision (blue line) – as a function of the ex ante group-specific average default rate $\overline{p}_G$. The optimal decisions are divided into three regions annotated at the top. Model parameters: $\sigma = 0.08, r = 0.1, \text{distaste}_G = 0$, and $c = 0.02$.



indicating that most applicants are not considered worthy borrowers. A simple back-of-the-envelope analysis suggests that the occurrence of such extremely high-SES groups should be rare.[10]

2. For an applicant from an *intermediate group*, the loan officer will first conduct information acquisition and then approve the application if and only if the revealed information indicates sufficiently high creditworthiness.
3. An applicant from a *low-SES group* (high $\overline{p}_G$ or distaste_G) is *immediately rejected* without information acquisition. Even though paying attention can identify some good borrowers, the benefit is not large enough to justify the attention cost.

The two cutoffs between these three regions are graphically illustrated in Graph B of Figure 2. We plot the loan officer's expected utility when employing each of the three strategies as a function of the group-specific average default rate $\overline{p}_G$.[11]

[10]How good does an applicant group have to be for it to be worth approving without information acquisition? The average interest rate in our data is approximately 8.6%. If the bank's cost of capital is equal to China's central bank rate of 3.25% in our sample period, this would mean the bank can earn a cost-adjusted annual return of only 5.35% if the applicant does not default. In contrast, if the application defaults and we assume a 40% recovery rate as the loan is uncollateralized, the bank stands to lose 60%. Therefore, as long as the average application default rate is higher than $\frac{5.35\%}{5.35\% + 60\%} \approx 8\%$, the default loan officer's action without information acquisition is rejection.

[11]The socioeconomic status of a group is also impacted by distaste_G, but we show in Appendix A that variation in distaste_G is equivalent to variation in $\overline{p}_G$ with a different scaling parameter applied. Therefore, Graph B of Figure 2 illustrates the behavior of the various model regions without loss of generality.

Finding the optimal strategy in each region amounts to simply choosing the strategy with the highest expected utility.

The distributional consequences of higher attention costs. When attention cost c rises, the "immediately approve" and "immediately reject" regions both expand. This is easy to see in Graph B of Figure 2. A higher cost c causes a downward parallel shift in the blue curve. As a consequence, the blue curve's crossing point with the red and green curves will shift rightward and leftward, respectively. This result should be intuitive: When attention becomes more costly, loan officers are less likely to acquire information and are more likely to make a decision based on ex ante group attributes.

Therefore, increases in attention costs lead to *asymmetric* consequences for applicants from different groups. When loan officers are busier, applicants from low-SES groups will be rejected more often; applicants from extremely high-SES groups may even be accepted more often.[12] This discussion is formalized into the testable predictions presented below.

Testable predictions. Consider two groups, G_1 and G_2, where the former has lower SES (has higher $\overline{p}_G$ or higher distaste_G). Then:

1. (Average effect) G_1 will receive weakly less attention and be approved less often.
2. (Comparative statics) If loan officer attention cost c increases, then:

   (a) In the *base case* where the more favorable group (G_2) is in the intermediate region, both groups will receive weakly less attention, and the *gaps* between their attention and approval rates will weakly widen.

   (b) In the *special case* where the favorable group (G_2) is in the extremely high-SES region, then its approval rate will in fact increase.

It is worth emphasizing that we think the base case should dominate our empirical setting. As explained earlier with respect to the 3 possible loan officer strategies, there should be only a small subset of applicants whose SES is so high that they can be approved without loan officer attention.

What is the main innovation? As discussed at the start of this section, our model builds on Bartoš et al. (2016) who were the first to propose this "attention discrimination" mechanism. They also present experimental evidence for prediction 1 above: the differential attention received by higher or lower SES groups. Our main contribution is testing prediction 2: the comparative statics on attention cost c. In our empirical tests, we use orthogonal variations in loan officer workloads to perturb attention cost c.

[12]Approval/rejection decision quality also declines. To see this, consider applicants positioned exactly to the right of the boundary between the "immediately approve" and "learn" regions. Before an increase in attention cost, the loan officer acquires information and makes an informed decision, so the bad applicants—those whose default rates $\overline{p}_G + p_I$ are high—are screened out. After an increase in attention cost, those applicants are also approved without scrutiny. By the same reasoning, the decision quality for those at the boundary between "learn" and "immediately reject" regions also declines: Before an increase in attention cost, the good marginal applicants are approved, but after the increase, they are automatically rejected without review.

III. Data and Institutional Background

In this section, we describe the data and provide background information on the retail loan-screening process. We also compare the basic statistics between high- versus low-SES applicants. The main empirical results based on loan officer attention variation are reported in Section IV.

A. Data Source

We obtain internal retail-lending screening records from one of the largest national banks in China. The sample data cover approximately 146,000 loan applications screened by 92 loan officers working at the bank's headquarters office from Apr. 2013 through Apr. 2014. Borrowers include both wage/salary workers and self-employed individuals running small/microscale businesses. The loan terms and targeted borrowers are comparable to those associated with retail financing products in the United States. Loan maturity is 1–3 years; the median (mean) loan amount is 60,000 (66,461) Chinese RMB, which is equivalent to $9787 ($10,841) U.S. dollars and comparable to the average personal installment loan of around $16,000 in the U.S.[13] The average annual interest rate in our sample is 8.56%, which is also similar to the 2-year U.S. personal loan interest rate of about 10% over the same sample period.[14] Summary statistics are presented in Table 1 and variable definitions are listed in Appendix B.

Our data include all information that loan officers can see in each application during the screening process, which allows us to control for a rich set of applicant- and loan-level characteristics that are potentially related to the borrower's credit quality. The data include 111 variables extracted from application materials and 295 variables extracted from borrower personal credit reports issued by the Chinese Central Bank. These variables include almost all commonly used metrics for creditworthiness, such as leverage ratio, existing debt, credit history, income, and so forth.

More importantly, the data contain detailed timestamps for each step in the loan officer's screening and decision-making process, which allows us to infer the amount of attention paid by loan officers to each applicant.

B. The Loan-Screening Process

The 3-stage loan origination and screening process is illustrated in Figure 3. Stage 1, which occurs at the local bank-branch level, is not captured by our data. Our study focuses on stage 2, which generates workload variations via an external algorithm that is not affected by loan officer discretion, as well as stage 3, during which headquarters loan officers screen the applications and make lending decisions. We now describe the 3 stages of loan screening.

Stage 1. Application submission. Loan applications are sourced from local bank branches all over the country. Each applicant submits an application for a specific maturity and loan amount. The local bank branch manager ensures that the application materials are complete and determines the appropriate interest rate and terms

[13]Source: https://www.experian.com/blogs/ask-experian/research/personal-loan-study/.

[14]Source: https://www.federalreserve.gov/releases/g19/hist/cc_hist_tc_levels.html.

TABLE 1

Summary Statistics

Table 1 presents summary statistics. In Panel A we report the summary statistics for the full sample. In Panel B we compare the means of applicants in groups with high versus low social/economic status. See Appendix B for variable definitions.

Panel A. Summary Statistics for the Full Sample

	N	Mean	SD	10%	25%	50%	75%	90%
Officer Screening Activities								
Approval	145,982	0.342	0.474	0	0	0	1	1
ReviewTime (*min*)	145,977	30.674	40.615	2.433	6.712	18.354	36.536	72.392
StandardizedReviewTime	145,977	0.933	1.082	−0.552	0.488	1.068	1.476	2.113
Busyness	145,982	19.150	6.979	10	15	19	24	27
Predicted Busyness	145,982	17.323	5.241	10.408	13.866	17.531	20.756	23.873
LOO-Predicted Busyness	145,982	16.406	4.951	9.843	13.041	16.534	19.786	22.636
Assignment	145,982	17.621	9.410	5	11	18	24	30
Borrower Characteristics								
PublicEmployee	145,982	0.081	0.273	0	0	0	0	0
LocalResident	145,982	0.455	0.498	0	0	0	1	1
EmploymentCert	145,982	0.620	0.486	0	0	1	1	1
IncomeCert	145,982	0.342	0.474	0	0	0	1	1
RegularPay	145,982	0.117	0.321	0	0	0	1	1
HomeOwner	145,982	0.223	0.417	0	0	0	0	1
NoCreditHistory	145,982	0.173	0.379	0	0	0	0	1
LeverageRatio	145,982	0.268	0.850	0	0.017	0.103	0.276	0.543
OverdueMonth	145,982	1.073	1.829	0	0	0	1	3
CreditInqury	145,982	3.274	5.907	0	0	1	4	9
HasInvestmentAcc	145,982	0.007	0.081	0	0	0	0	0
SocialSecurity	145,982	0.406	0.491	0	0	0	1	1
Litigation	145,982	0.002	0.043	0	0	0	0	0
Peasant	145,982	0.114	0.317	0	0	0	0	1
NonCollege	145,982	0.296	0.457	0	0	0	1	1
Female	145,982	0.240	0.427	0	0	0	0	1
Age	145,982	35.767	8.258	25.458	28.951	34.723	42.145	47.866
Income (*RMB*)	145,982	57,131	112,254	8000	12,000	22,000	50,000	150,000
Loan characteristics								
LoanSize (*RMB*)	145,982	66,461	28,057	40,000	50,000	60,000	80,000	100,000
LoanToIncome	145,982	3.285	2.733	0.600	1.286	2.609	4.444	6.667
ShortTerm	145,982	0.279	0.449	0	0	0	1	1
InterestRate (%)	145,982	8.558	0.208	8.400	8.400	8.610	8.610	8.610

Panel B. Comparison Between the High- Versus Low-SES Groups

	SES Measure			
	SocialStatus		EconomicStatus	
	High-SES	Low-SES	High-SES	Low-SES
Approval	0.519	0.181	0.655	0.254
StandardizedReviewTime	1.169	0.719	1.131	0.877

for each application, but approval decisions need to be made by loan officers at bank headquarters in stage 3.

Stage 2. Assignment of applications to headquarters loan officers. After an application is completed in stage 1, it is stored electronically in the bank's systems and then distributed to the headquarters loan officers by a central workload-dispatcher algorithm over which loan officers have no control. As for how exactly the algorithm assigns applications, see Section IV.D for an in-depth discussion.

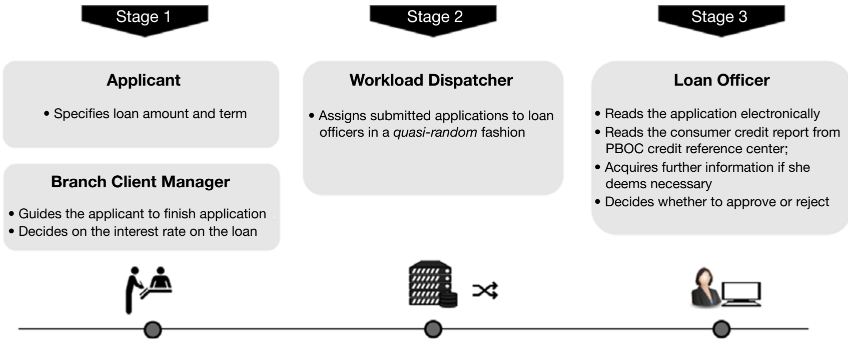
Stage 3. Headquarters loan officers make approval/rejection decisions. The assigned loan officer accesses applicant information electronically, evaluates the information, and decides whether to approve the application.[15] Our sample comprises 92 officers. Of a total of 145,982 applications, only 34.2% are approved, so the process is relatively selective. Our data include precise timestamps when

[15]This is the only decision she needs to make. Loan terms are already determined at stage 1.

FIGURE 3

Flow Chart of Loan Origination and Screening

Figure 3 illustrates that in stage 1, loan applications are submitted at regional bank branches across the country. Loan amounts, maturities, and interest rates are already determined at this stage. In stage 2, a central dispatcher algorithm assigns applications to headquarters loan officers. In stage 3, loan officers read each application and decide whether to approve or reject it.



applications are assigned to officers and when officers make decisions, enabling us to measure officers' attention allocation to each application by calculating the length of time they spend reviewing the application.

Key premise: loan officers are attention-constrained. A key premise of this attention-based mechanism is that decision-makers face attention constraints. In our sample, loan officers have to read lengthy documents within short periods of time. First, each application package includes an application form that spans 10–20 pages long, which includes the borrower's demographic information, personal wealth and income, the purpose of borrowing, and so forth. It also includes lengthy supplementary materials that are used to support the applicant's self-reported information. These materials could include third-party-issued official documentation such as photocopies of personal ID cards, employment certificates, property deeds, and bank statements. These additional documents can run into hundreds of pages. In addition, a credit report issued by the central credit bureau, which is about 10 pages long, is also attached. Meanwhile, we find that loan officers only spend a meager median (mean) of 18 (31) minutes per application.[16] Needless to say, this is not enough time to go over all application materials carefully.[17] Overall, our evidence suggests that loan officers are indeed facing attention constraints.

[16]We discuss in detail how we measure the review time in the next section.

[17]In private conversations, loan officers also told us that they are under time pressure and feel rather rushed in the loan screening process. Besides, the per-application review time in our sample is shorter than in similar loan review processes in the United States. To mention a crude comparison, when examining a U.S. commercial bank, Agarwal and Ben-David (2018) find that 133 loan officers screened 30,268 loan applications over 2 years (see Table 1). In our data, 92 loan officers screened 145,982 applications over 2 years. This implies that the average review time in the United States is $\frac{133 \times 2/30,268}{92 \times 1/145,982} \approx 13.9$ times longer than that in our data. In addition, Wei and Zhao (2022) show that the median processing time is 8–29 days in the U.S. mortgage market. However, this number includes processing time across all steps, from the submission of an application to the final origination of a loan, not just the review time spent by loan officers, and thus is not directly comparable to our review time measure.

C. Applicant Socioeconomic Status (SES)

Loan officers use socioeconomic labels to guide attention allocation. Given the abovementioned loan officer attention constraints, it is natural to hypothesize that they may use simple signals to decide how much time to spend reviewing each application. In private conversations, loan officers explain that they rely on a few easily observable socioeconomic labels as signals for guiding time allocation. Some of these labels are related to applicants' social statuses while others are related to applicants' economic statuses, and we call them "socioeconomic labels" collectively in subsequent discussions. These are zero-or-one indicator labels that can be easily observed on application forms that are verified by the branch officers in stage 1.

Two labels are usually considered signals of an applicant's *social status*:

1. *PublicEmployee*: whether the applicant works for the public sector. Chinese society treats public employees, including those working in the government, public schools or hospitals, state-owned firms, or any other government-sponsored institutions, as meriting higher social status.[18] Insofar as an applicant needs to fill in her position on the first page of the application form, this is a salient and easily observable signal for loan officers.
2. *LocalResident*: whether the applicant is a local city resident rather than a migrant worker. Local residents are typically thought to be of higher status than migrant workers (i.e., people who grow up in rural areas and migrate to work in a city) as the former generally have access to better public services such as education and healthcare because of local policy restrictions.[19] The Chinese "Hukou" (household registration) system makes it easy to distinguish local residents from migrant workers, making this another salient signal that loan officers use.

There are also 4 labels that reflect an applicant's *economic status*:

1. *EmploymentCert*: whether the applicant has an official certificate that verifies her position of employment. Such a certificate is considered acceptable to the bank only when i) the employer's official stamp and a top manager's signature are on the certificate, and ii) the employer's identity can be recognized and verified by the bank. This depends on whether the applicant works for a large employer. In contrast, employees of microscale businesses and self-employed entrepreneurs cannot. Thus, loan officers generally consider the availability of an employment certificate as a signal of superior economic status.
2. *IncomeCert*: whether the applicant's employer can provide an employer-issued income certificate. In practice, this depends on whether the applicant is in a long-term position with high job security. Short-term contractors or paid interns cannot provide this certificate.

[18]Public employees are colloquially described as "the insiders of the system" in China, and public positions are considered as high social status in Chinese society. For example, in 2019, 1.4 million applicants competed for around 24,000 government positions, suggesting that on average about 1 out of 60 candidates can land a job "inside the system."

[19]In fact, discrimination against migrant workers has been a long concern in China.

3. *RegularPay*: whether the income of the applicant is stable in terms of both timing and amount. This depends on the type of employment. For instance, salespeople usually have commission-based income which can be volatile and would not have this label.[20]
4. *HomeOwner*: whether the applicant owns real estate, which can be assessed via photocopies of property deeds.

While the social labels might (but do not have to) potentially relate to some taste-based prejudice by loan officers and the general society, it is not surprising that the economic labels are correlated with applicants' fundamental credit risk and thus can lead to cross-sectional differences in borrower financial access. However, when these socioeconomic signals are used by loan officers to determine attention allocation, our attention-driven mechanism suggests that the associated inclusion gap could be further widened, no matter which kind of motivation drives the differential perception among applicants. This difference-in-difference implication cannot be simply driven by fundamental credit risk differentials and is unique under this attention-based mechanism.

Applicants cannot change these labels. It is important to note that, under our institutional setting, the availability of those salient SES labels is determined by borrowers' *ability* to provide the corresponding documentation, rather than their *willingness*. At the time of loan application, the SES labels are exogenous to the applicant's discretion, as they are determined by the applicant's ex ante occupation type, migration status, and so forth. For instance, if a borrower is self-employed or works in a micro/small business, she typically cannot provide an employment certificate that can be considered "valid" by the bank. Instead, a certificate officially issued by a large public company or government entity is considered valid and indicates higher socioeconomic status. Also, applicants who work in the public sector (PublicEmployee = 1) are treated better by loan offices. While it is possible that people are well aware of this, it is too difficult to switch jobs to the public sector just for the sake of obtaining a loan.

Defining applicants from high and low SES backgrounds. Our attention-based mechanism, which is formally analyzed in Section II, suggests that applicants with fewer socioeconomic status labels may be considered to have lower SES and receive less attention from officers. That is, even if the credit quality of such an applicant is high enough to warrant approval, her application may still be hastily "passed up" by loan officers who are busy and intend to reserve their attention for applicants from higher socioeconomic backgrounds. We test this prediction formally in Section IV.

Given that loan officers are considering multiple social- and economic-status labels, we use a data-driven approach to summarize the combined effects of the aforementioned status indicators into 2 variables, "Social Status" and "Economic Status," to classify how applicants' socioeconomic status is perceived by loan officers. Specifically, we fit a linear probability model by regressing the approval

[20]In stage 1 of the loan screening process, the bank's local branch employees analyze the information from applicants' bank statements, create easily observable labels indicating income stability (or not), and add it to the application form.

TABLE 2

Higher Approval Probability for Applicants with More Social/Economic Labels

In Table 2, we estimate the relationship between loan approval probability and applicants' social and economic labels. The outcome variable equals 1 if the loan application is approved, and 0 otherwise. As discussed in Section III.C, *PublicEmployee* and *LocalResident* are indicators of applicant social status, while the other four indicators are applicant economic-status labels. Application-level controls include log(*Income*), log(*Loan/Income*), log(1 + *LeverageRatio*), log(1 + *OverdueMonth*), log(1 + *CreditInqury*), *HasInvestmentAcc*, *Female*, log(*Age*), *Peasant*, *NonCollege*, *SocialSecurity*, *Litigation*, *ShortTerm*, and log(*InterestRate*). See Appendix B for variable definitions. Standard errors are double-clustered at the week and officer levels. *** $p<1\%$, ** $p<5\%$, * $p<10\%$.

	Approval					
	1	2	3	4	5	6
PublicEmployee	0.246***					0.027***
	(17.062)					(3.985)
ResidentCert		0.467***				0.029
		(28.719)				(1.041)
StandardPay			0.419***			0.188***
			(22.675)			(10.223)
EmploymentCert				0.527***		0.459***
				(30.703)		(17.641)
IncomeCert					0.395***	0.020
					(23.722)	(1.123)
Application Controls	Yes	Yes	Yes	Yes	Yes	Yes
Officer-Month-Yr FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
Loan type FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	145,982	145,982	145,982	145,982	145,982	145,982
Adjusted R^2	0.140	0.265	0.170	0.354	0.222	0.361

probability on those SES labels in Table 2, and we compute the regression-predicted value of application approval for the 2 social-status labels and 4 economic-status labels, separately:[21]

$$\text{(3)} \qquad \begin{aligned} \text{SocialStatus}_i \equiv \widehat{\text{Approval}}_i|\{\text{PublicEmployee}_i, \text{LocalResident}_i\} \\ = \widehat{b}_{\text{PublicEmployee}} \cdot \text{PublicEmployee}_i \\ + \widehat{b}_{\text{LocalResident}} \cdot \text{LocalResident}_i \end{aligned}$$

$$\text{(4)} \qquad \begin{aligned} &\text{EconomicStatus}_i \equiv \\ &\widehat{\text{Approval}}_i|\{\text{EmploymentCert}_i, \text{RegularPay}_i, \text{IncomeCert}_i, \text{HomeOwner}_i\} \\ &= \widehat{b}_{\text{EmploymentCert}} \cdot \text{EmploymentCert}_i \\ &+ \widehat{b}_{\text{RegularPay}} \cdot \text{RegularPay}_i \\ &+ \widehat{b}_{\text{IncomeCert}} \cdot \text{IncomeCert}_i + \widehat{b}_{\text{HomeOwner}} \cdot \text{HomeOwner}_i \end{aligned}$$

In other words, these two variables are single-dimensional summaries of the multiple social and economic labels in a given application. For simplicity, in

[21]Here we use a version of the regression without additional controls. Our results are not sensitive to the exact methodology through which the socioeconomic status indicators are combined.

https://doi.org/10.1017/S0022109024000565 Published online by Cambridge University Press

subsequent analyses, we create 2 indicator variables, $\mathrm{High}-\mathrm{SES}(\mathrm{Social})_i$ and $\mathrm{High}-\mathrm{SES}(\mathrm{Economic})_i$, which equals 1 for applicants whose $\mathrm{SocialStatus}_i$ and $\mathrm{EconomicStatus}_i$ values, respectively, are above the sample median.[22] The correlation between the 2 High SES indicators is -0.133, suggesting they capture different information about the borrowers' profiles.

Credit quality distribution largely overlaps between the two borrower groups. As we can see from the definitions of the SES labels, borrowers from the low-SES group likely have lower credit quality and thus lower credit access on average. However, a pure cross-sectional difference in credit quality cannot explain the main implication of our paper, which is a difference-in-difference effect suggesting that the gap in attention allocation and financial inclusion between high- and low-SES borrower groups widen as loan officers' attention constraints get tighter. Also, even the average difference in approval rates across the high- versus low-SES borrower groups (18% vs. 51% by social status and 25% vs. 66% by economic status, as shown in Panel B of Table 1) seems too large to be fully justified by differences in credit risk. Comparing the credit quality distributions of the 2 borrower groups, we find substantial overlap between the high- and low-SES borrowers as shown in Figure IA2. In fact, we see a large share of rejected borrowers in the low-SES group actually have higher credit quality than the median high-SES borrowers who are approved. We discuss more details about the borrowers' credit quality distributions in the Appendix Section A.2.

IV. The Impacts of Loan Officer Attention Constraints

In this section, we formally test our main empirical prediction: When loan officers face tighter attention constraints, they disproportionately reduce attention on low-SES applicants and reject them much more frequently. We start by using a simple measure of loan officer busyness to proxy for their attention constraints and then construct instruments for the constraints that enable us to infer causal effects.

A. Measuring Loan Officer Attention-Constraints and Attention Allocation

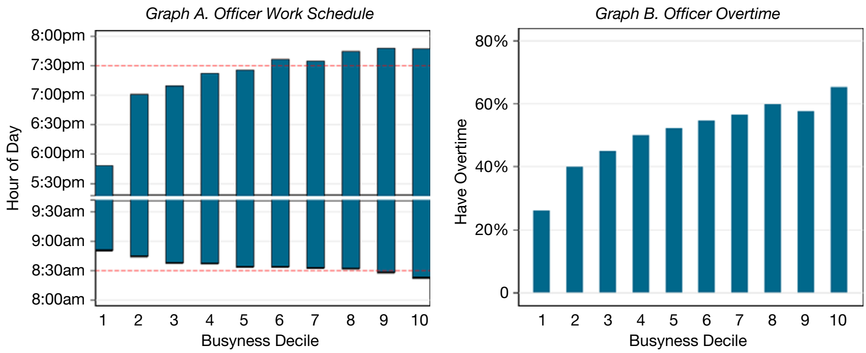
Measuring variations in attention constraints. To proxy for loan officer attention constraints, we compute an officer-day level variable $\mathrm{Busyness}_{j,d}$, which is defined as the number of applications officer j processes on day d. The reasoning is straightforward: The higher the number of applications the officer has to process, the less time she can afford to spend on each one. As shown in Table 1, the median officer processes 19 applications on a given day, and the 10th and 90th percentiles are 10 and 27 applications, respectively. Therefore, there are substantial variations in officer busyness and the concomitant time constraint on each application. In other

[22]In our baseline analyses, we fit the abovementioned linear probability models using the full sample. In the Supplementary Material Appendix Section A, we show that our empirical results remain robust when using out-of-sample fitting.

FIGURE 4

Officer Busyness and Work Schedules

In Figure 4, we sort the sample into deciles differentiated by officer busyness, which is defined as the number of applications an officer processes on a given day. Graph A plots the average time of a workday at which officers start and end their work. The start and end times are measured by the timestamps indicating when officers submit the first and last loan decisions on each day. Graph B plots the fraction of days on which officers work overtime, defined as working before 8:30AM or after 7:30PM (the red dashed lines in Graph A).



words, while loan officers are always busy (time spent on each application is always low), they sometimes become even busier.[23]

We argue that loan officer busyness is relevant to attention constraints. First, as shown in Figure 4, when officers are busier, they work longer hours, and the probability that officers work overtime rises from approximately 20% on the least busy days to over 60% on the busiest days.[24] Second, when officers are busier, they spend less time reviewing each application. This is reflected in both Graphs A and B of Figure 1, as well as in the subsequent analyses presented in this and the next section. Overall, these findings are consistent with our view that officers are more severely attention-constrained when they are busy.

Instrumented Variations in Attention Constraints Using *realized* loan officer busyness raises an endogeneity concern: Loan officers can set their own pace at work, which may lead to omitted variable problems. For instance, a loan officer who wants to relax on a particular day may choose to quickly reject most applications perfunctorily, leading to a spurious negative correlation between busyness and the officer's loan-approval rate. To be clear, such situations per se cannot lead to our difference-in-differences result, which shows the *differential* impact of attention constraints on loan officer decisions for applicants from high- and low-SES groups. To obtain these differential results, one needs to explain why a careless officer would rashly reject the low-SES borrowers but not the high-SES ones. That said, we take this endogeneity concern seriously.

[23]Even though mortgage applications are more complex and involve larger loan sizes, which may require more careful screening, it is worth noting that the average number of applications processed by our sample loan officers is significantly larger than the number processed by U.S. mortgage loan officers (e.g., https://www.bancorp.com/employment-opportunities/loan-processor/.)

[24]Supplementary Material Figure IB1 and Table IB2 also indicate that the longer an officer works on a given day, the less time she spends reviewing each individual application.

To resolve this endogeneity concern, we need to find a source of busyness variation that is external, that is, *not controlled* by loan officers. As described in Section III.B, loan applications are assigned to officers by a central dispatcher algorithm over which officers have no control. Apart from assurances from the bank we study, we also confirm that the assignment algorithm is external to loan officers by verifying that the number of assignments has no relationship with current or previous loan officer backlogs (Supplementary Material Table IA8).[25]

Using the number of assignments as an instrument, we capture the external variation in officer busyness through a first-stage regression at the officer–day level:

$$\text{Busyness}_{j,d} = a + \sum_{\tau=0}^{3} b_\tau \cdot \text{Assignment}_{j,d-\tau} + \varepsilon_{j,d} \tag{5}$$

where $\text{Assignment}_{j,d}$ is the number of applications assigned to loan officer j on working day d. We include 3 lagged working days because some applications are processed a few days after assignments are allotted. In Panel A of Table 3, we present details associated with this first-stage regression using a number of different specifications, such as with/without lagged assignment terms, and with/without fixed effects. The main specification we use in our later instrumental variable analyses is in column 4.[26] The instrument is strong and can explain more than 40% of the variation in realized officer busyness. As will be shown later, the F statistic for this first-stage regression is well above the Staiger-Stock rule-of-thumb threshold of 10 or the Stock-Yogo threshold of 16.38, confirming the strength of this instrument. Hereafter, we call the value predicted in regression (5) "predicted busyness."

We then consider a second instrument which can be thought of as a further refinement of the previous one. Even though we find no correlation between assignments and loan characteristics (explained in Section IV.D), one might still worry about correlations between assignments and unobservable loan quality. We argue that this concern is unlikely to explain our findings, for 3 reasons. First, we obtain the entire set of administrative records for these loan applications and control for a comprehensive list of group fixed effects and loan characteristics, and control for all the characteristics that are commonly considered in the screening practice. Second, to explain our difference-in-differences results, the confounding driver of assignments has to be negatively correlated with the credit quality of low-SES borrowers but uncorrelated, or even positively correlated, with the credit quality of high-SES borrowers. Third, in all our specifications, we control for week, branch, and loan-officer-year-month fixed effects, which enable us to control for all aggregate time-series variations in loan application volume or loan officer habits or

[25]Specifically, we are worried that loan officers may be able to *indirectly* influence the number of assignments they receive by working faster or slower. If an officer can face fewer assignments by having a larger backlog (through working more slowly), this would be a concern to our identification strategy.

[26]Our results are robust if we use those alternative specifications for the first-stage estimation.

TABLE 3

Predicting Loan Officer Busyness Using Assignments

In Table 3, we estimate the relationship between realized officer busyness on the number of applications assigned by the bank's workload dispatcher algorithm. The dependent variable $Busyness_{j,d}$ is the total number of applications processed by loan officer j on day d, $Assignment_{j,d}$ is the total number of assignments the loan officer receives, and $LOO-Assignment_{j,d}$ is the total number of assignments from other provinces she received. In Panel A we report the results obtained using total assignments at officer-day level, and in Panel B we report the results obtained using LOO-assignments. For columns 1 through 4 we do not include fixed effects, while for columns 5 and 6, we include officer- and officer-month-year fixed effects, respectively. Standard errors are double-clustered at the officer and month-year levels. We use specification (4) to compute the "predicted busyness" instrument presented in Section IV.A. T-statistics are reported in parentheses. $^{***}p<1\%$, $^{**}p<5\%$, $^{*}p<10\%$.

	Dependent Variable : $Busyness_{j,d}$								
	1	2	3	4	5	6	7	8	9
Panel A. Using Total Assignments to Predict Busyness									
$Assignment_{j,d}$	0.519***	0.418***	0.392***	0.367***	0.332***	0.239***	0.358***	0.329***	0.243***
	(11.070)	(11.813)	(12.118)	(11.163)	(10.267)	(8.564)	(10.659)	(10.334)	(8.151)
$Assignment_{j,d-1}$		0.216***	0.174***	0.163***	0.143***	0.088***	0.162***	0.144***	0.095***
		(5.929)	(5.284)	(5.441)	(4.622)	(3.489)	(6.102)	(5.132)	(3.730)
$Assignment_{j,d-2}$			0.125***	0.080***	0.064**	0.015	0.082***	0.067***	0.022
			(5.796)	(3.827)	(2.792)	(0.769)	(4.499)	(3.319)	(1.157)
$Assignment_{j,d-3}$				0.140***	0.119***	0.054***	0.142***	0.122***	0.061***
				(10.255)	(8.287)	(5.197)	(12.005)	(9.126)	(5.766)
Officer FE	No	No	No	No	Yes	No	No	Yes	No
Officer-Month-Yr FE	No	No	No	No	No	Yes	No	No	Yes
Week FE	No	No	No	No	No	No	Yes	Yes	Yes
Observations	9,498	9,498	9,498	9,498	9,498	9,498	9,498	9,498	9,498
Within R^2					0.422	0.197	0.508	0.408	0.200
Adjusted R^2	0.420	0.482	0.502	0.527	0.555	0.638	0.596	0.615	0.669
Panel B. Using LOO-Assignments to Predict Busyness									
$LOO-Assignment_{j,d}$	0.452***	0.395***	0.379***	0.363***	0.323***	0.201***	0.336***	0.293***	0.200***
	(8.275)	(8.551)	(9.391)	(9.124)	(8.281)	(9.347)	(9.354)	(8.744)	(9.214)
$LOO-Assignment_{j,d-1}$		0.157***	0.128***	0.126***	0.109***	0.060***	0.119***	0.099***	0.061***
		(4.716)	(4.302)	(4.536)	(3.985)	(3.065)	(4.705)	(4.038)	(3.218)
$LOO-Assignment_{j,d-2}$			0.106***	0.077***	0.063**	0.018	0.074***	0.057**	0.025
			(4.114)	(3.126)	(2.484)	(0.805)	(3.354)	(2.708)	(1.256)
$LOO-Assignment_{j,d-3}$				0.111***	0.093***	0.034***	0.108***	0.086***	0.042***
				(8.700)	(7.544)	(3.542)	(9.872)	(8.281)	(4.394)
Officer FE	No	No	No	No	Yes	No	No	Yes	No
Officer-Month-Yr FE	No	No	No	No	No	Yes	No	No	Yes
Week FE	No	No	No	No	No	No	Yes	Yes	Yes
Observations	145,982	145,982	145,982	145,982	145,982	145,982	145,982	145,982	145,982
Within R^2					0.320	0.122	0.344	0.259	0.119
Adjusted R^2	0.317	0.363	0.384	0.406	0.450	0.595	0.514	0.555	0.640

preferences. Therefore, we are effectively capturing the idiosyncratic variations which are very unlikely to have any systematic correlation with the credit quality gap between borrower groups.

Although this concern is not very realistic, we address it by constructing a loan-level, LOO instrument that is based on variations on loan officer workload from non-local markets that are not affected by local market conditions. The intuition is, since the loan officers we study work at the headquarters office, while applications are sourced from bank branches all over the country, if many assignments from province A make a given loan officer busy, this could affect her decision-making regarding applications from province B even when neither the quantity nor the quality of applications from province B changes. Specifically, when examining a loan officer's decision-making process on a loan application from a province, we construct a "Bartik-type" measure as the weighted average assignments to the same loan officer from all other provinces. By using this measure as our instrument *and directly control for the number of assignments from the source province*, we tease out the effect of loan officer attention constraints that is driven by *external* workload variations that are orthogonal to the local market conditions. As shown in Panel B of Table 3, this LOO instrument is also very strong, and it explains a similar share of variations in loan officer busyness as the first instrument.

Measuring attention allocation. With internal timestamps for each action, we can directly measure loan officer attention allocation by estimating how much time is spent reviewing each application.[27] To remove variations in application review times that are unlikely to reflect active loan officer choices,[28] we define "standardized review time" as the log deviation of review time from the median level within each Officer × Month-Year × Loan-Type × Bank-Branch group. Specifically, we compute

$$\text{StandardizedReviewTime} = \log\left(\frac{\text{ReviewTime}}{\text{MedianReviewTimebygroup}}\right) + \underbrace{\text{Medianlog}(\text{ReviewTime})}_{\text{fullsample}} \quad (6)$$

where the groups in the denominator of the first term are Officer × Month-Year × Loan-Type × Bank-Branch buckets. In other words, we remove review-time variations that are explained by interactions of all of the fixed effects we use in our regressions; these fixed effects, combined, explain 36% of log review time variations, as shown in column 4 of Supplementary Material Table IB1. The second term in equation (6) simply adds back the overall sample median of log review time. As reported in Table 1, the interquartile range of this attention measure (standardized review time) runs from 0.488 to 1.476.

B. The Impact on Attention Allocation

Having defined measures of loan officer attention constraints and attention allocation in the previous section, we now examine the impact of tighter attention constraints on attention allocation and approval rates.

In an exploratory analysis, we first simply plot (without controls) average attention and approval rates as a function of busyness. For Figure 1 Graphs A and B, we sort the sample into deciles differentiated by officer busyness and plot the average standardized review time for the high- and low-SES groups as measured by their social or economic status, respectively. Attention declines for both groups, but the decline is more noticeable for the low-SES group: When officers become

[27]We measure the time spent reviewing each as the number of minutes that elapses between 2 consecutive decisions rendered by the same loan officer. We subtract lunch breaks (12:00 to 13:00) and all non-working periods (including weekends, national holidays, and other days off). Our results are not sensitive to this specific method for measuring review time.

[28]For instance, less-experienced officers may take longer to process each application. Also, officers may become more proficient at processing applications over time, so we also include year-month fixed effects.

busier, they appear to shift attention *away* from low-SES applicants. Graphs C and D plot the approval rates. When officers become busier, the approval rate for the low-SES group declines steadily relative to the rate for the high-SES group. These patterns are consistent with the base case model predictions. Further, the approval rate for high-SES applicants actually appears to increase slightly with loan officer busyness, which may be consistent with the special case predictions where some extremely high-SES applicants are quickly approved without careful screening.

To formally investigate the effects of officer attention constraints, we now conduct regression analyses at the application level. For column 1 of Table 4, we regress standardized review time – our measure of loan officer attention allocation – on loan officer busyness decile, the *High-SES(Social)* dummy that indicates whether the applicant's *SocialStatus* is above the sample median, as well as the interaction between busyness decile and High-SES(Social). We then run similar tests for column 2, measuring SES based on the applicant's economic status. Consistent with the visual patterns displayed in Figure 1, when officers are busier, the attention they pay to applicants decreases, but the decrease is disproportionately larger for low-SES applicants. For instance, the results reported in column 1 indicate that, when officer busyness varies from the lowest to the highest decile, the attention paid to low-SES applicants declines by $(10-1)\times -0.059 \approx 53\%$. While it is unavoidable that officers will spend less time on each application when they are busier, the attention gap between high- and low-SES applicants increases by $(10-1)\times 0.017 \approx 15.3\%$, and these effects are statistically significant at the 1% level. In column 3, we show that the effects remain statistically and economically similar when both social and economic status are considered simultaneously.

We then re-examine these results using the 2 abovementioned instruments to capture variations in loan officer busyness. In columns 4 through 9 of Table 4, we use *Predicted Busyness* and *LOO-Predicted Busyness* to estimate how idiosyncratic variations in loan officer attention constraints affect their allocation of review time. The instrumented busyness measures are estimated in an earlier stage, and the F statistics are reported at the end of the corresponding columns. As shown in the table, the F statistics are very high, confirming our earlier argument that the 2 instruments are strong. For these 2SLS regressions, we estimate standard errors using the bootstrap method. Under these instrumented analyses, the effects are qualitatively and quantitatively similar to those estimated in ordinary least squares regression (OLS) analyses.

Using an auxiliary attention measure to provide complementary evidence. In Supplementary Material Appendix C, we use another loan officer action—conducting further due diligence when screening applicants—to proxy officer attention allocation.[29] This indirect measure yields the same conclusion,

[29] At the bank we study, a loan officer must select from a list of reasons when she renders a rejection. While some reasons for rejection indicate that the officer makes rejection decisions based on information already in hand (e.g., high leverage or a bad credit history), some others indicate that the loan officer, before rejecting the application, engaged in additional due diligence to gain information beyond what was readily available in the application package. For example, the loan officer can indicate that she called one of the applicant's references but was given inconsistent information during the call.

TABLE 4

Effects of Officer Attention Constraints on Review Time

For Table 4, we estimate how loan office attention constraints affect the time they spend on reviewing each loan application by applicants from high and low socioeconomic backgrounds. The dependent variable is the standardized review time for a loan application, defined as the logarithm of the excess time spent by officers in reviewing each application (equation (6)). *High-SES(Social)* and *High-SES(Economic)* are dummy variables indicating whether the applicant has above-median *SocialStatus* and *EconomicStatus*, respectively, and the definition is explained in Section III.C. For columns 1 through 3, *BusynessDecile* is the officer's daily busyness, defined as the number of applications processed on a given day, sorted into deciles. For columns 4 through 9, *BusynessDecile* is the officer's instrumented daily busyness, defined as the number of applications processed on a given day sorted into deciles and instrumented by the total or leave-one-out (LOO) number of applications assigned to the loan officer over the preceding 3 working days. For columns 4 through 6, we use assignment-predicted busyness; for columns 7 through 9, we use LOO assignment-predicted busyness. For the overall effect of loan office attention constraints on groups with high social or economic status, we calculate the sum of two groups of coefficients $(\beta_1+\beta_3)$ and $(\beta_1+\beta_5)$, and report the P-values of their T-tests. The regressions include officer × month-year fixed effects, week fixed effects, origination-bank-branch fixed effects, and loan-type fixed effects. Application controls include log(*Income*), log(*Loan/Income*), log(1 + *LeverageRatio*), log(1 + *OverdueMonth)*, log(1 + *CreditInqury)*, *HasInvestmentAcc*, *Female*, log (*Age*), *Peasant*, *NonCollege*, *SocialSecurity*, *Litigation*, *ShortTerm*, and log(*InterestRate*). See Appendix B for variable definitions. T-statistics are reported in parentheses. Standard errors are double-clustered at the week and officer levels. $^{***}p<1\%$, $^{**}p<5\%$, $^{*}p<10\%$.

	Dependent Variable: StandardizedReviewTime								
	Busyness Measure								
	Actual Busyness			Predicted Busyness			LOO-Predicted Busyness		
	1	2	3	4	5	6	7	8	9
β_1BusynessDecile	-0.059***	-0.058***	-0.061***	-0.025***	-0.024***	-0.029***	-0.020***	-0.016***	-0.022***
	(−17.248)	(−19.538)	(−16.167)	(−8.297)	(−9.248)	(−8.999)	(−6.347)	(−5.970)	(−6.395)
β_2High-SES(Social)	0.439***		0.413***	0.461***		0.434***	0.469***		0.444***
	(13.854)		(12.801)	(23.833)		(21.608)	(23.566)		(21.385)
β_3High-SES(Social) × BusynessDecile	0.017***		0.017***	0.013***		0.015***	0.013***		0.013***
	(3.949)		(3.922)	(4.722)		(5.152)	(4.783)		(4.484)
β_4High-SES(Economic)		0.249***	0.188***		0.285***	0.215***		0.291***	0.222***
		(13.496)	(9.599)		(13.311)	(10.031)		(14.021)	(10.912)
β_5High-SES(Economic) × BusynessDecile		0.017***	0.014***		0.013***	0.012***		0.012***	0.011***
		(6.639)	(5.117)		(4.139)	(3.645)		(3.817)	(3.740)
Application Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Local Busyness Controls	No	No	No	No	No	No	Yes	Yes	Yes
Officer-Month-Yr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	145,977	145,977	145,977	145,977	145,977	145,977	145,977	145,977	145,977
Adjusted R^2	0.076	0.046	0.082	0.074	0.044	0.082	0.075	0.045	0.082
$\beta_1+\beta_3$	-0.042***		-0.044***	-0.011***		-0.014***	-0.006***		-0.009***
P-value of $(\beta_1+\beta_3)$	(0.000)		(0.000)	(0.000)		(0.000)	(0.001)		(0.000)
$\beta_1+\beta_5$		-0.041***	-0.047***		-0.011***	-0.017***		-0.005**	-0.011***
P-value of $(\beta_1+\beta_5)$		(0.000)	(0.000)		(0.000)	(0.000)		(0.048)	(0.000)
First-stage F-Statistics				99.5	99.5	99.5	41.7	41.7	41.7

as officers are less likely to conduct due diligence for low-SES applicants when they are attention-constrained. In particular, by comparing the reasons loan officers cite when rejecting applications, we find that, when a loan officer rejects a high-SES applicant, she is much more likely to have engaged in further due diligence (e.g., searching up the applicant online) beyond simply browsing the documents already provided. In contrast, a loan officer is more likely to reject a low-SES applicant based on boilerplate reasons such as "leverage is too high." More importantly, such differences in rejection reasons become more prominent when loan officers are busier. In sum, this suggestive result complements our main finding, suggesting that low-SES applicants receive less attention allocation when loan officers face tighter attention constraints.

C. The Impact on Lending Decisions

In Table 5, we use similar specifications to estimate the impact of officer busyness on approval decisions. The results reported in column 1 indicate that, for applicants with low social status, increasing from the lowest to the highest busyness decile reduces approval probability by $(10-1)\times -0.009 \approx 8.1$ percentage points. This reduction is about 45% of the average approval rate for this group of applicants. Similarly, the results reported in column 2 show a decline in the approval rate by $(10-1)\times -0.011 \approx 9.9$ percentage points, which is about 39% of the group average. In contrast, the approval probability is roughly unchanged or even slightly increased for the high-SES groups. The results remain broadly similar when both SES measures are considered jointly for column 3.

We then use the instrumented busyness measures to estimate the effects of loan officer attention constraints on approval decisions. The results reported in columns 4 through 9 of Table 5 are robust when using both Predicted Busyness and LOO-Predicted Busyness. Overall, these results are consistent with our main prediction that, when loan officers face tighter attention constraints, low-SES applicants receive disproportionately less attention and are rejected more frequently. In some specifications, we even find evidence that high-SES groups experience higher approval rates, suggesting that the special-case model predictions may have some bite.

In Supplementary Material Table IB4, we further show that our main results hold up when we examine each of the 6 social or economic status labels separately.[30]

To further confirm that we are indeed utilizing the idiosyncratic variations in loan officer workload as instruments for their attention constraints, we also conduct an additional robustness check for all our instrumental variable analyses. Specifically, we use an alternative "residual" measure of the instrument by first regressing the assignment or LOO assignment on week and branch fixed effects, which fully control for time-series or locational fundamentals, and then we take the residual component for each of those 2 variables and use it as the instrument in our analyses. As shown in Supplementary Material

[30]Using these two instrumented measures of busyness, we reproduce Figure 1 in Figures IB3 and IB4 and reproduce Figure IB2 in Figures IB5 and IB6. We find qualitatively similar results in all cases.

TABLE 5

Effects of Attention Constraints on Approval Decisions

For Table 5, we estimate how loan office attention constraints affect their approval decisions on loan applications by high- and low-SES applicants. The dependent variable is a dummy variable indicating whether the officer approves the application. *High-SES(Social)* and *High-SES(Economic)* are dummy variables indicating whether the applicant's *SocialStatus* and *EconomicStatus* are above the median, respectively, and the definition is explained in Section III.C. For columns 1 through 3, *BusynessDecile* is the officer's daily busyness, defined as the number of applications processed on a given day, sorted into deciles. For columns 4 through 9, *BusynessDecile* is the officer's instrumented daily busyness, defined as the number of applications processed on a given day sorted into deciles and instrumented by the total or leave-one-out (LOO) number of applications assigned to the loan officer over the preceding 3 working days. For columns 4 through 6, we use assignment-predicted busyness; for columns 7 through 9, we use LOO assignment-predicted busyness. For the overall effect of loan office attention constraints on groups with high social or economic status, we calculate the sum of two groups of coefficients $(\beta_1+\beta_3)$ and $(\beta_1+\beta_5)$, and report the P-values of their T-tests. The regressions include officer × month-year fixed effects, week fixed effects, origination-bank-branch fixed effects, and loan-type fixed effects. Application controls include log(*Income*), log(*Loan/Income*), log(1 + *LeverageRatio*), log (1 + *OverdueMonth)*, log(1 + *CreditInqury)*, *HasInvestmentAcc*, *Female*, log(*Age*), *Peasant*, *NonCollege*, *SocialSecurity*, *Litigation*, *ShortTerm*, and log(*InterestRate*). See Appendix B for the variable definitions. *t*-statistics are reported in parentheses. Standard errors are double-clustered at the week and officer levels. $^{***}p<1\%$, $^{**}p<5\%$, $^{*}p<10\%$.

	Dependent Variable: Approval								
	Busyness Measure								
	Actual Busyness			Predicted Busyness			LOO-Predicted Busyness		
	1	2	3	4	5	6	7	8	9
β_1BusynessDecile	-0.009***	-0.011***	-0.010***	-0.004***	-0.003***	-0.006***	-0.004***	-0.003***	-0.006***
	(−6.415)	(−9.395)	(−6.964)	(−4.380)	(−4.046)	(−8.048)	(−5.144)	(−4.700)	(−9.396)
β_2High-SES(Social)	0.408***		0.375***	0.399***		0.367***	0.403***		0.371***
	(26.896)		(23.928)	(56.706)		(50.568)	(57.625)		(53.533)
β_3High-SES(Social) × BusynessDecile	0.007***		0.006***	0.009***		0.008***	0.008***		0.008***
	(3.326)		(3.289)	(7.241)		(7.018)	(6.688)		(7.141)
β_4High-SES(Economic)		0.373***	0.331***		0.383***	0.331***		0.385***	0.333***
		(36.447)	(31.948)		(37.182)	(35.370)			
β_5High-SES(Economic) × BusynessDecile		0.015***	0.011***		0.013***	0.012***		0.012***	0.011***
		(7.014)	(5.286)		(8.564)	(7.725)		(8.402)	(7.876)
Application Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Local Busyness Controls	No	No	No	No	No	No	Yes	Yes	Yes
Officer-Month-Yr FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan type FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	145,982	145,982	145,982	145,982	145,982	145,982	145,982	145,982	145,982
Adjusted R^2	0.269	0.217	0.338	0.272	0.219	0.342	0.272	0.219	0.342
$\beta_1+\beta_3$	−0.002		-0.004**	0.005***		0.002***	0.004***		0.001*
P-value of $(\beta_1+\beta_3)$	(0.251)		(0.012)	(0.000)		(0.000)	(0.000)		(0.055)
$\beta_1+\beta_5$		0.004*	0.002		0.010***	0.006***		0.009***	0.005***
P-value of $(\beta_1+\beta_5)$		(0.067)	(0.300)		(0.000)	(0.000)		(0.000)	(0.000)
First-stage F-Statistics				99.5	99.5	99.5	41.7	41.7	41.7

https://doi.org/10.1017/S0022109024000565 Published online by Cambridge University Press

Tables IA10 and IA11, our results are robust to using "residual" variation in instruments.[31]

It is worth noting that, in our regressions, we control for officer × month-year, week, bank branch, and loan-type fixed effects. Therefore, our findings do not stem from any differences regarding officer-specific preferences, branch-specific risk-management styles, or aggregate time trends. We also control for a comprehensive list of features that could be related to borrower creditworthiness; in unreported robustness checks, we also find that our results are not sensitive to the choice of controls. Supplementary Material Figure IB2 further shows that, conditional on all loan characteristics and fixed effects, the gaps between the high- and low-SES applicant groups in terms of both attention allocation and approval rate widen almost monotonically when loan officers get busier.

In addition, we also show that our results are similar under additional robustness checks. Specifically, we use propensity score matching to focus on comparable borrowers in the low- versus high-SES groups and find the results to be largely unchanged (Supplementary Material Tables IA5 and IA6).

D. Loan Officer Work Patterns and the Assignment Algorithm

By using assignments as the instrumental variable and controlling for week, branch, loan-type, and officer × month-year fixed effects, we end up identifying *idiosyncratic* variations in loan officer busyness. What drives the remaining variation? Here, we provide more background information about the assignment algorithm.

Although the bank we study does not disclose the exact implementation details of the algorithm, we can catch a glimpse from informal discussions with bank employees. According to the employees, the algorithm groups loan applications based on a variety of factors such as the branch where the application is submitted, the loan type, size categories, and so forth. Every day, a loan officer is randomly matched to one or several groups of applications.[32] The officer-group matching changes randomly over time, and as a consequence, the law of large numbers ensures that different loan officers' workloads are on average similar over a longer horizon, but they experience idiosyncratic variations in their workloads on a day-by-day basis as a result of the randomness of the assignments. While the bank does not disclose the basis of the officer-to-group matching, bank employees affirm that loan officers have no influence over the matching criteria and that the matching does not take into account the quality of applications. Therefore, this assignment algorithm generates idiosyncratic variations in officer attention constraints that are orthogonal to loan officer preferences and loan quality.

[31]We also show in Supplementary Material Table 3 that, after controlling for week and loan officer fixed effects, the within-group R^2 is still as high as about 20%, suggesting that our instrument is strong even when focusing on idiosyncratic variations.

[32]Banks use this "assignment by group" algorithm so that each loan officer can process a set of relatively homogeneous applications on a given day. This makes loan processing more efficient and lending standards more consistent.

https://doi.org/10.1017/S0022109024000565 Published online by Cambridge University Press

To test the claims of bank employees, we conducted tests to confirm that assignments are indeed uncorrelated with credit quality. We examined the relationship between assignments and a comprehensive list of observable loan characteristics and reported the results in Supplementary Material Table IA9. Consistent with our discussion with bank employees, the results provide no indication of a relationship between assignment volume and loan characteristics.

Assignment volume is a valid instrument for loan officer busyness because loan officers have an incentive to finish their backlogs. Although loan officers face no "hard constraints" in completing all assigned workloads within the same day, they do face "soft constraints" such that if the backlogs are not processed within 2–3 days, their supervisors may "have a conversation with them."[33] Consistent with this, we find that when loan officers are assigned heavier workloads, they work longer hours (Supplementary Material Figure IA3 and Table IB2) and spend less time reviewing each application (Supplementary Material Figure IB1).

Lastly, we note that the headquarters loan officers we study have incentives that are generally aligned with the bank. They are paid fixed salaries plus bonuses, and their bonuses are determined by metrics that take into account both the volume of applications they process and the realized default rate relative to the bank's target rate. We are not aware of any significant agency problems that might distort loan officer incentives.[34]

V. Conclusion

Insufficient financial inclusion of individuals from low SES backgrounds is a concern on both equity and efficiency grounds. Motivated by Bartoš et al. (2016), we propose that financial inclusion can be hindered by attention constraints on financial decision-makers. In the selection process, attention-constrained decision-makers may ration their attention allocation using ex ante socioeconomic labels. As a result, low-SES applicants may be given insufficient attention and be rejected more often. For applicants from very high-SES backgrounds, the reverse often applies: They may be quickly approved without careful review. As a result, this attention-based mechanism can lead to an inclusion gap between high- and low-SES borrowers.

We provide evidence for this mechanism using proprietary retail loan-screening records from a large national bank in China. Loan officers at the bank are time-constrained and spend a median of only 18 minutes on each loan application they review. Against this backdrop, applicants without certain socioeconomic labels are considered to have low socioeconomic status by loan officers who screen applicants and make lending decisions. The low-SES applicants receive less review time and their loan applications are more often rejected compared with those of otherwise similar applicants with more socioeconomic labels. Furthermore, when

[33]Also, severe delay in the loan processing may lead the applicants to complain, given that the expected processing time (including all other steps that we are not looking at) for a loan application is about 2 weeks.

[34]Also, unlike loan officers from the branch offices who usually also play "sales" roles (e.g., Giacoletti et al. (2021)), our sample loan officers do not face any quota pressure. Meanwhile, they also do not face any "upper limits" regarding how many loans to process in a day.

loan officers experience tighter attention constraints caused by orthogonal variation in their workloads, both review times and approval gaps between high- and low-SES applicants widen.

Our findings imply that, in human-based decision processes, organizational arrangements or technologies that relax attention constraints may help improve inclusion and promote diversity. Our findings also suggest that the rise of fintech may—if properly used—promote financial inclusion through pre-processing of applicant information and relieving decision-makers of attention bottlenecks. Moving beyond our immediate setting, many high-stake decisions are made by humans, and key decision-makers—such as court judges, college admissions officers, and so on—are often very busy. Therefore, while our study focuses on the impact of attention constraints on the allocation of financial resources, we suspect that similar mechanisms are at play in other settings that are potentially more consequential.

Appendix A. Analytical Results

A.1. Solving the Model

We first prove results with fixed distaste_G and varying $\overline{p}_G$, and then show that the results based on varying distaste_G are mathematically similar.

Optimal loan officer behavior. We first prove that the optimal loan officer decision is characterized by three regions with two cutoffs. For notational simplicity, we define the profit function (which includes the distaste term) in equation (1) as $\Pi(p)$:

$$\Pi(p) \equiv (r - \text{distaste}_G) - (1+r)\cdot p \tag{7}$$

where $p = \overline{p}_G + p_I$ is the applicant's default probability, r is the interest rate, and distaste_G reflects loan officer prejudice. Note that, if the loan officer does not acquire information about p_I, the expected profit is given by $\Pi(\overline{p}_G)$:

$$E_{p_I}[\Pi(\overline{p}_G + p_I)] = E_{p_I}[(r - \text{distaste}_G) - (1+r)\cdot(\overline{p}_G + p_I)] \tag{8}$$

$$\overset{E(p_I)=0}{=} (r - \text{distaste}_G) - (1+r)\cdot \overline{p}_G \tag{9}$$

The loan officer can choose 1 of 3 strategies: immediately accept (A), learn and then decide (L), or immediately reject (R). The expected utilities associated with these choices are

$$U_A(\overline{p}_G) = \Pi(\overline{p}_G) \tag{10}$$

$$U_L(\overline{p}_G) = E_{p_I}[\max(\Pi(\overline{p}_G + p_I), 0)] - c \tag{11}$$

$$U_R(\overline{p}_G) = 0. \tag{12}$$

These utilities are plotted in Graph B of Figure 2 as a function of $\overline{p}_G$. For each $\overline{p}_G$, the loan officer chooses the strategy $\in \{A, L, R\}$ that maximizes expected utility. Let us now discuss the three possible regions.

- Immediately accept region. For sufficiently small $\overline{p}_G$, $U_A > U_R$ because $\lim_{\overline{p}_G \downarrow 0} \Pi(\overline{p}_G) = r - \text{distaste}_G > 0 = U_R$. Further, as long as $c > 0$, as $\overline{p}_G$ decreases there exists sufficiently small $\overline{p}_G$ such that $U_A > U_L$. Intuitively, because the probability of rejection after information acquisition is sufficiently low, it is not worth paying the cost to acquire information.[35]

- Learning region. It is easy to show that U_L must cross with U_A once from below. We have just argued that, when $\overline{p}_G$ is sufficiently low, $U_L < U_A$. Also, as $\overline{p}_G$ increases, U_L converges to $-c$ in the limit while U_A diverges to negative infinity, so for sufficiently large $\overline{p}_G$, we must have $U_L > U_A$.

  To see that U_L crosses U_A only once, we just need to see that $\frac{dU_L(\overline{p}_G)}{d\overline{p}_G} > \frac{dU_A(\overline{p}_G)}{d\overline{p}_G}$:

$$\frac{dU_L(\overline{p}_G)}{d\overline{p}_G} = \frac{d}{d\overline{p}_G}\int_{-\infty}^{\frac{r-\text{distaste}_G}{1+r}-\overline{p}_G}[(r-\text{distaste}_G)-(1+r)(\overline{p}_G+p_I)]f(p_I)dp_I \tag{13}$$

$$=(-1)\times 0+\int_{-\infty}^{\frac{r-\text{distaste}_G}{1+r}-\overline{p}_G}\frac{d}{d\overline{p}_G}[(r-\text{distaste}_G)-(1+r)(\overline{p}_G+p_I)]f(p_I)dp_I \tag{14}$$

$$=\int_{-\infty}^{\frac{r-\text{distaste}_G}{1+r}-\overline{p}_G}-(1+r)f(p_I)dp_I \tag{15}$$

$$=-(1+r)\cdot P\left(p_I<\frac{r-\text{distaste}_G}{1+r}-\overline{p}_G\right) \tag{16}$$

$$>-(1+r)=\frac{dU_A(\overline{p}_G)}{d\overline{p}_G} \tag{17}$$

- Immediately reject region. If $\overline{p}_G$ is sufficiently high, U_A clearly becomes unboundedly negative while U_L converges to $-c$, both of which are lower than $U_R=0$.

Note that it is possible for the first two regions to have zero length under certain parameters.[36]

Comparative statics on c. We need to show simply that the 2 crossing points between the three regions move in desired directions when c varies.

- The first crossing $\overline{p}_G^{(1)}$ is defined by $U_A\left(\overline{p}_G^{(1)}\right)=U_L\left(\overline{p}_G^{(1)}\right)$. If c increases, this reduces $U_L\left(\overline{p}_G^{(1)}\right)$ but does not change $U_A\left(\overline{p}_G^{(1)}\right)$. Because U_A is a decreasing function, which means that $\overline{p}_G^{(1)}$ must rise.
- The second crossing $\overline{p}_G^{(2)}$ is defined by $U_L\left(\overline{p}_G^{(2)}\right)=U_R=0$. Recall that $U_L(\overline{p}_G)=E_{p_I}[\max(\Pi(\overline{p}_G+p_I),0)]-c$ and that the first component is a decreasing function with $\overline{p}_G$. Thus, increases in c must be offset by decreases in $\overline{p}_G^{(2)}$.

Parallel results when groups differ by distaste_G. We have derived results when varying $\overline{p}_G$. What if groups differ by distaste_G? Well, if we rearrange the profit function (7), we obtain

[35]The crossing point between the U_L and U_A may be negative, which is an infeasible value for $\overline{p}_G\in[0,1]$. When this happens, the "immediately approve" region has zero length.

[36]When c is very high, the learning region can disappear. If r is low, c is low, and if σ (the standard deviation of p_I) is high, the immediately accept region can disappear.

$$\Pi(p) = r - [\text{distaste}_G - (1+r)\overline{p}_G] - (1+r)\cdot p_I. \tag{18}$$

Note that distaste_G and $(1+r)\overline{p}_G$ enter into the formula in identical ways. Therefore, all results based on varying $\overline{p}_G$ also apply to varying distaste_G after adjusting for the $1+r$ scaling.

A.2. Uneven Distribution of Workloads is Suboptimal for Lender Profits

When attention constraints might be binding, it would be optimal for the bank to distribute workloads evenly to equalize the marginal benefits of attention across loan officers. This indicates that the empirically observed workload distribution method the bank employs, which leads to uneven distribution, is likely to be suboptimal from a profit-maximization perspective.

Let us modify the model slightly to analyze the impact of workload distribution. Suppose that there are a total of X applications to be assigned to N ex ante identical loan officers, and let the number of assignments be denoted as $\{X_1, X_2, \ldots, X_N\}$, such that $\sum_{i=1}^{N} X_i = X$. Suppose that each loan officer can read only an expected number of K applications per day and $KN << X$: in other words, the attention constraint is binding in aggregate. Assume that each application has group identity G (which determines $\overline{p}_G$ and distaste_G), drawn IID from some distribution, and that workload assignments cannot depend on actual group identities.

We now argue that the profit-maximizing approach is to allocate workloads evenly: $X_1 = X_2 = \ldots = X_N = \frac{X}{N}$. The proof involves a simple application of a "water-filling" argument. Note that, for each loan officer, the marginal benefit of paying attention to an application is a function of the application's group identity G. In the notation of Appendix Section A.1, the marginal benefit is given by

$$h(G) = (U_L(G) + c) - \max(U_A(G), U_R(A)),$$

where U_L is the expected profit if the loan officer pays attention to learn about the applicant before making a decision, and U_A and U_R are expected profits if the loan officer immediately accepts or rejects the application without paying it any attention. All groups can be re-ordered such that $h(G)$ becomes a decreasing function with the re-ordered group identities.

Clearly, for each loan officer $i = 1\ldots, N$, the expected marginal benefit of being able to pay attention to 1 additional applicant is decreasing with the volume of assignments X_i, as each loan officer always first pays attention to the group with the highest $h(G)$, followed by the second, and so forth. Therefore, the optimal decision for the bank is to assign workloads evenly.

Appendix B. Variable Definitions

Appendix B provides the definitions of the key variables used in our analyses.

Officer Screening Activities

Approval: Equals 1 if the officer has approved the application, and 0 otherwise.

ReviewTime: The number of minutes that the officer spends reviewing an application, measured as the time that has elapsed between the officer's previous decision and the current decision.

StandardizedReviewTime: The log of reviewing time divided by the median values for each officer×month-year×branch×loan type. See equation (6).

HasInfoAcquision: Equals 1 if the cited rejection rationale indicates that the loan officer has engaged in further due diligence (e.g., phone calls), and 0 otherwise.

Busyness: The total number of applications reviewed by an officer on a given day.

Predicted Busyness: The predicted number of applications an officer reviews on a given day using the total number of applications on the current day and on 3 lagged business days that are assigned to an officer. See equation (5).

LOO-Predicted Busyness: The predicted number of applications an officer reviews on a given day using the number of applications from other provinces on the current day and on 3 lagged business days that are assigned to an officer.

Assignment: The total number of applications assigned to an officer on a given day.

Backlog: The number of applications that have been assigned to an officer but have not yet been reviewed, at the beginning of a given day.

Borrower Socioeconomic Statuses

PublicEmployee: Equals 1 if the applicant works in the public sector, and 0 otherwise.

LocalResident: Equals 1 if the applicant provides certificates indicating recent places of residency, and 0 otherwise.

EmploymentCert: Equals 1 if the applicant provides certificates related to current employment, and 0 otherwise.

IncomeCert: Equals 1 if the applicant provides certificates related to income, and 0 otherwise.

RegularPay: Equals 1 if the applicant receives fixed salary payments, and 0 otherwise.

HomeOwner: Equals 1 if the applicant provides certificates related to housing property owned, and 0 otherwise.

Borrower Characteristics

LeverageRatio: The applicant's preexisting debt-to-income ratio.

NonCreditHistory: Equals 1 if the applicant has no credit history, and 0 otherwise.

OverdueMonth: The highest number of months over which the applicant has been overdue making payments in the most recent 2 years.

CreditInquiry: The number of inquiries into the applicant's credit history in the most recent 2 years.

HasInvestmentAcc: Equals 1 if the applicant has an investment account, and 0 otherwise.

SocialSecurity: Equals 1 if the applicant receives a social security allowance, and 0 otherwise.

Litigation: Equals 1 if the applicant has been involved in any legal proceedings, and 0 otherwise.

Peasant: Equals 1 if the applicant reports holding a permanent agricultural residence registration in an application, and 0 otherwise.

NonCollege: Equals 1 if the applicant has a non-college degree, and 0 otherwise.

Female: Equals 1 if the applicant is female, and 0 otherwise.

Age: The applicant's age.

Income: The applicant's total income.

Loan Characteristics

Loan/Income: The ratio of the amount of the loan for which the applicant has applied to the applicant's total income.

ShortTerm: Equals 1 if the term of the loan for which the applicant has applied is less than 3 years.

InterestRate: The interest rate of the loan for which the applicant has applied at origination.

Supplementary Material

To view supplementary material for this article, please visit http://doi.org/10.1017/S0022109024000565.

Funding statement

Huang acknowledges financial support from the Key Program of National Natural Science Foundation of China (NSFC Grant Number 72233003) and the General Program of National Natural Science Foundation of China (NSFC Grant Number 72372157). Tai acknowledges financial support from the Research Grants Council of the Hong Kong Special Administration Region, China (Project No.17504320).

References

Agarwal, S., and I. Ben-David. "Loan Prospecting and the Loss of Soft Information." *Journal of Financial Economics*, 129 (2018), 608–628.

Ambrose, B. W.; J. N. Conklin; and L. A. Lopez. "Does Borrower and Broker Race Affect the Cost of Mortgage Credit?" *Review of Financial Studies*, 34 (2021), 790–826.

Andrei, D., and M. Hasler. "Investor Attention and Stock Market Volatility." *Review of Financial Studies*, 28 (2015), 33–72.

Bartlett, R.; A. Morse; R. Stanton; and N. Wallace. "Consumer-Lending Discrimination in the FinTech Era." *Journal of Financial Economics*, 143 (2022), 30–56.

Bartoš, V.; M. Bauer; J. Chytilová; and F. Matějka. "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition." *American Economic Review*, 106 (2016), 1437–1475.

Bayer, P.; F. Ferreira; and S. L. Ross. "What Drives Racial and Ethnic Differences in High-Cost Mortgages? The Role of High-Risk Lenders." *Review of Financial Studies*, 31 (2018), 175–205.

Beck, T.; P. Behr; and A. Madestam. "Sex and Credit: Do Gender Interactions Matter for Credit Market Outcomes?" *Journal of Banking and Finance*, 87 (2018), 280–396.

Brock, J. M., and R. De Haas. "Discriminatory Lending: Evidence from Bankers in the Lab." *American Economic Journal: Applied Economics*, 15 (2023), 31–68.
Brown, J. R.; J. A. Cookson; and R. Z. Heimer. "Growing Up Without Finance." *Journal of Financial Economics*, 134 (2019), 591–616.
Butler, A. W.; E. J. Mayer; and J. P. Weston. "Racial Disparities in the Auto Loan Market." *Review of Financial Studies*, 36 (2023), 1–41.
Célerier, C., and A. Matray. "Bank-Branch Supply, Financial Inclusion, and Wealth Accumulation." *Review of Financial Studies*, 32 (2019), 4767–4809.
Charles, K. K.; E. Hurst; and M. Stephens. "Rates for Vehicle Loans: Race and Loan Source." *American Economic Review*, 98 (2008), 315–320.
D'Acunto, F.; P. Ghosh; R. Jain; and A. G. Rossi. "How Costly Are Cultural Biases?" Working Paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3736117 (2021).
Dobbie, W.; A. Liberman; D. Paravisini; and V. Pathania. "Measuring Bias in Consumer Lending." *Review of Economic Studies*, 88 (2021), 2799–2832.
Ewens, M., and R. R. Townsend. "Are Early Stage Investors Biased Against Women?" *Journal of Financial Economics*, 135 (2020), 653–677.
Fisman, R.; D. Paravisini; and V. Vig. "Cultural Proximity and Loan Outcomes." *American Economic Review*, 107 (2017), 457–492.
Fisman, R.; A. Sarkar; J. Skrastins; and V. Vig. "Experience of Communal Conflicts and Intergroup Lending." *Journal of Political Economy*, 128 (2020), 3346–3375.
Fuster, A.; P. Goldsmith-Pinkham; T. Ramadorai; and A. Walther. "Predictably Unequal? The Effects of Machine Learning on Credit Markets." *Journal of Finance*, 77 (2022), 5–47.
Fuster, A.; M. Plosser; P. Schnabl; and J. Vickery. "The Role of Technology in Mortgage Lending." *Review of Financial Studies*, 32 (2019), 1854–1899.
Gabaix, X. "Behavioral Inattention." In *Handbook of Behavioral Economics I: Applications and Foundations,* Vol 2. Amsterdam, Netherlands: North-Holland (2019), 261–343.
Giacoletti, M.; R. Heimer; and E. G. Yu. "Using High-Frequency Evaluations to Estimate Discrimination: Evidence from Mortgage Loan Officers." Working Paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3795547 (2021).
Goldsmith-Pinkham, P., and K. Shue. "The Gender Gap in Housing Returns." *Journal of Finance*, 78 (2023), 1097–1145.
Hasler, M., and C. Ornthanalai. "Fluctuating Attention and Financial Contagion." *Journal of Monetary Economics*, 99 (2018), 106–123.
Hebert, C. "Gender Stereotypes and Entrepreneur Financing." Working Paper, University of Toronto (2020).
Hirshleifer, D.; Y. Levi; B. Lourie; and S. H. Teoh. "Decision Fatigue and Heuristic Analyst Forecasts." *Journal of Financial Economics*, 133 (2019), 83–98.
Hirshleifer, D. A., and J. Sheng. "Macro News and Micro News: Complements or Substitutes?" *Journal of Financial Economics*, 145 (2022), 1006–1024.
Hu, A., and S. Ma. "Persuading Investors: A Video-Based Study." *Journal of Finance*, forthcoming (2025).
Huang, S.; Y. Huang; and T.-C. Lin. "Attention Allocation and Return Co-Movement: Evidence from Repeated Natural Experiments." *Journal of Financial Economics*, 132 (2019), 369–383.
Huo, D.; B. Sun; M. Tai; and Y. Xuan. "Lending Next to the Courthouse: Exposure to Adverse Events and Mortgage Lending Decisions." *Journal of Financial and Quantitative Analysis*, 59 (2024), 2375–2398.
Jansen, M.; H. Nguyen; and A. Shams, "Rise of the Machines: The Impact of Automated Underwriting." *Management Science*, forthcoming (2025).
Kacperczyk, M.; S. Van Nieuwerburgh; and L. Veldkamp. "A Rational Theory of Mutual Funds' Attention Allocation." *Econometrica*, 84 (2016), 571–626.
Kahneman, D. *Thinking, Fast and Slow.* New York, NY: Farrar, Straus and Giroux (2011).
Kermani, A., and F. Wong. "Racial Disparities in Housing Returns." NBER Working Paper No. 29306 (2021).
Lanning, J. A. "Testing Models of Economic Discrimination Using the Discretionary Markup of Indirect Auto Loans." Working Paper, Consumer Financial Protection Bureau (2021).
Liao, L.; Z. Wang; J. Xiang; H. Yan; and J. Yang. "User Interface and Firsthand Experience in Retail Investing." *Review of Financial Studies*, 34 (2021), 4486–4523.
Liberti, J. M., and M. A. Petersen. "Information: Hard and Soft." *Review of Corporate Finance Studies*, 8 (2019), 1–41.
Liu, H.; L. Peng; and Y. Tang. "Retail Attention, Institutional Attention." *Journal of Financial and Quantitative Analysis*, 58 (2023), 1005–1038.

Mackowiak, B.; F. Matejka; and M. Wiederholt. “Rational Inattention: A Review.” *Journal of Economic Literature*, 61 (2023), 226–273.
Mondria, J. “Portfolio Choice, Attention Allocation, and Price Comovement.” *Journal of Economic Theory*, 145 (2010), 1837–1864.
Mondria, J., and C. Quintana-Domeque. “Financial Contagion and Attention Allocation.” *Economic Journal*, 123 (2012), 429–454.
Montoya, A. M.; E. Parrado; A. Solís; and R. Undurraga. “Bad Taste: Gender Discrimination in the Consumer Credit Market.” Working Paper (2020).
Morse, A., and K. Pence. “Technological Innovation and Discrimination in Household Finance.” In *The Palgrave Handbook of Technological Finance*. London, UK: Springer International Publishing (2021), 783–808.
Müller, K. “Busy Bankruptcy Courts and the Cost of Credit.” *Journal of Financial Economics*, 143 (2022), 824–845.
Ongena, S., and A. Popov. “Gender Bias and Credit Access.” *Journal of Money, Credit and Banking*, 48 (2016), 1691–1724.
Peng, L. “Learning with Information Capacity Constraints.” *Journal of Financial and Quantitative Analysis*, 40 (2005), 307–329.
Peng, L., and W. Xiong. “Investor Attention, Overconfidence and Category Learning.” *Journal of Financial Economics*, 80 (2006), 563–602.
Philippon, T. “On Fintech and Financial Inclusion.” NBER Working Paper No. 26330 (2019).
Shu, T.; X. Tian; and X. Zhan. “Patent Quality, Firm Value, and Investor Underreaction: Evidence from Patent Examiner Busyness.” *Journal of Financial Economics*, 143 (2022), 1043–1069.
Stein, L. C., and C. Yannelis. “Financial Inclusion, Human Capital, and Wealth Accumulation: Evidence from the Freedman’s Savings Bank.” *Review of Financial Studies*, 33 (2020), 5333–5377.
Van Nieuwerburgh, S., and L. Veldkamp. “Information Acquisition and Under-Diversification.” *Review of Economic Studies*, 77 (2010), 779–805.
Wei, B., and F. Zhao. “Racial Disparities in Mortgage Lending: New Evidence based on Processing Time.” *Review of Corporate Finance Studies*, 11 (2022), 775–813.
Zhang, Y. “Discrimination in the Venture Capital Industry: Evidence from Two Randomized Controlled Trials.” Working Paper, Columbia University (2020).