Discrimination and Access to Capital: A Field Experiment in Ethiopia*

Shibiru Ayalew Shanthi Manian Ketki Sheth

July 26, 2024

In a large-scale field experiment in Ethiopia, we show that gender discrimination by loan officers is unlikely to contribute to gender gaps in capital access. In 3,696 evaluations for a business plan competition, randomized business-owner gender did not affect capital decisions, including loan consideration. Using an incentivized belief elicitation, we conclude that loan officers do not engage in statistical discrimination or taste-based discrimination. We confirm that loan officers' beliefs are accurate using machine learning to predict business outcomes after 18 months: gender does not meaningfully improve targeting of high-performing businesses, suggesting no trade-off between gender equity and effective capital allocation.

*Ayalew: Adama Science and Technology University, Ethiopia, shibekoo84@gmail.com; Manian: Washington State University, shanthi.manian@wsu.edu; Sheth: University of Tennessee, Knoxville, ksheth2@utk.edu. This study was funded by the Center for Effective Global Action and International Growth Center, and implemented in collaboration with the Entrepreneurship Development Institute in Ethiopia. The study was approved by IRBs at the University of California, Merced, and Washington State University and pre-registered on the AEA registry (AEARCTR-0006518). We are grateful to Prashant Bharadwaj, Aislinn Bohren, Laura Gee, Leah Lakdawala, Ted Miguel, and participants at numerous seminars and conferences for helpful feedback.

1 Introduction

Capital is a key accelerator for business growth and productivity, yet female business owners in many low-income countries are less likely to obtain formal financing and earn lower profits.[1] Observable differences between male and female entrepreneurs explain only a small portion of the profit gap, suggesting that gender discrimination may be an important, yet understudied, factor inhibiting the success of female entrepreneurship (Buvinic, 2018). Financial providers, such as loan officers, may discriminate against female business owners either due to prejudice or because they use gender as a proxy for business performance. Both factors are commonly believed to be more pronounced in contexts with large gender disparities, such as Ethiopia. On the other hand, financial providers are experts in their field, face real stakes in how their portfolio performs, and have access to a significant amount of information about businesses they are evaluating, all of which reduce the likelihood of discrimination.

In addition to equity considerations, the response to gender in capital allocation decisions may have implications for profit maximization. If loan officers aim to identify and target capital to high-performing businesses, and gender does in fact predict business performance, then gender discrimination may be an effective profit-maximizing strategy.[2] Indeed, standard theories of discrimination suggest that discrimination should only persist when it is profit-maximizing.[3] However, empirical studies have documented discrimination from prejudice and inaccurate beliefs, even when it reduces profits (Guryan and Charles, 2013).

These equity and profit-maximization implications depend not only on whether loan officers discriminate but also on the source of discrimination. Researchers generally distinguish between two possible sources of discrimination: preferences (i.e., taste-based discrimination (Becker, 1957)) or differing beliefs about female-owned versus male-owned businesses (i.e., belief-based or statistical discrimination (Phelps, 1972; Arrow, 1974; Aigner and Cain, 1977)). Each source of discrimination, as well as the accuracy of the beliefs held by loan officers, has different welfare impacts and policy implications (Bohren et al., 2023). Our study addresses each of these questions in turn: whether loan officers discriminate in making capital allocation decisions, the extent of each source of discrimination, and whether the beliefs held by

[1]See, for example, Blattman, Fiala and Martinez (2014); World Bank (2019); Hardy and Kagy (2020). In Ethiopia, male managers are more likely to take out loans and tend to borrow significantly more than female managers (World Bank, 2019).

[2]High-performing (i.e., profitable) businesses increase the likelihood of repayment and the returns to capital. We conducted numerous interviews with loan officers confirming that the profitability of a business is a key factor in loan decisions.

[3]The Becker model (i.e., taste-based discrimination) implies that discriminatory financial providers will only remain in the market if discrimination is profit-maximizing (Becker, 1957). The theory of statistical discrimination demonstrates that it can indeed be profit-maximizing if discrimination is based on accurate beliefs on gender and business performance (Phelps, 1972; Arrow, 1974; Aigner and Cain, 1977; Bohren, Imas and Rosenberg, 2019).

loan officers are accurate. By exploring these key questions, we go beyond measuring discrimination to evaluate the implications of our results for profit maximization (i.e., targeting high-performing businesses).

We study these questions by embedding a field experiment in the natural context of a large business plan competition in Ethiopia. We identify discrimination by randomly assigning business-owner gender in the evaluation of the competition. Business owners throughout Ethiopia completed an application form designed to mimic information commonly captured in an initial application for a loan. We recruited 84 loan officers from thirteen different financial institutions to evaluate 916 real businesses that applied to the competition. For each business, the loan officers provided a score used to determine the competition grants and prizes, decided whether to forward the application to their own lending institution to be considered for a loan, and predicted how the business would perform in the future (both with and without additional capital). The latter belief elicitation was incentivized for accuracy and allows us to determine the extent of each source of discrimination.[4] Each business application was evaluated multiple times, and each loan officer evaluated multiple businesses, resulting in a sample of 3,696 evaluations and an identification strategy that incorporates both business and loan officer fixed effects. This context allows us to study the key players of interest: real businesses interested in applying for capital, and loan officers that are regularly involved in making capital allocation decisions in the financial industry.

To assess the accuracy of the loan officers' beliefs and implications for profit maximization, we conducted a follow-up survey of the businesses 18 months after the competition. We use these data to test whether gender is a meaningful predictor of future business performance. First, we compare the loan officers' subjective beliefs with actual future business performance to determine the accuracy of those beliefs.[5] We then employ machine learning algorithms as a benchmark for the optimal response to gender if loan officers had access to accurate beliefs. Using this as a benchmark allows us to assess whether loan officers' beliefs are accurate within a framework of belief-based discrimination and determines the implications for targeting high-performing businesses. The three components of our research design—loan officers' capital allocation decisions, their subjective beliefs, and business outcomes measured in the follow-up survey—allow us to identify the overall extent of discrimination (i.e., gender equity), separately identify taste-based and belief-based discrimination, and assess implica-

[4] Without eliciting loan officers' beliefs, the response to gender we identify in the capital allocation decisions could be consistent with a continuum of belief-based discrimination and taste-based discrimination (Bohren et al., 2023). An alternative approach to identifying statistical discrimination is to vary the amount of information observed by the evaluator. However, Bohren et al. (2023) show that this method can only partially identify the source of discrimination.

[5] It is not possible to distinguish taste-based discrimination from statistical discrimination from this comparison alone, unless one explicitly *assumes* that loan officers have accurate beliefs (Bohren et al., 2023).

tions of capital allocation decisions for profit-maximization (i.e., targeting capital towards high-performing businesses).

We find no evidence that the loan officers discriminated against female-owned businesses. Randomly assigned business-owner gender did not affect loan officers' scores for awarding the competition prizes or their decision to forward the business to their own financial institution for consideration for a loan. The point estimates of gender differences in both decisions are small: less than .03 standard deviations in the competition score and less than .01 percentage points in sending applications to their own lending institution. We find no evidence for discrimination in any subset of the sample (e.g., based on characteristics of the business owner, the business, or the loan officers), and our standard errors rule out any meaningful differences in these capital allocation decisions by gender.

Consistent with this lack of gender discrimination, we find that the loan officers expected similar future business performance and returns to capital (i.e., future survival, profits, and assets) for both genders. Loan officers also reported no gender difference in loan repayment likelihood. By eliciting loan officers' subjective beliefs, we identify and rule out belief-based discrimination, demonstrating that loan officers exhibit no preference-based partiality and no belief-based partiality. The alignment between capital allocation decisions and expectations of business performance is also consistent with a model of decision-making based on beliefs: loan officers do not believe that business-owner gender is predictive of business performance, and so do not award capital differently.

We then turn to the accuracy of beliefs. We find that loan officers' beliefs about gender differences in future business performance were not statistically different from the actual mean gender differences in future business performance, as measured in our follow-up survey. Furthermore, just like the loan officers, standard machine learning algorithms do not prioritize gender in a set of optimal predictors of actual business profits measured 18 months after the competition. Rather, both the loan officers and our machine learning algorithms rely on other information in the application to target high-performing businesses. This set of results implies that the loan officers have accurate beliefs and that there is no meaningful trade-off between gender equity and allocating capital to the highest-performing businesses. These results are consistent with the theoretical underpinnings of discrimination: theory suggests that if loan officers' beliefs are accurate and gender is *not* predictive of business performance, then discrimination would *not* be profit-maximizing.

We provide a number of checks to confirm the validity and robustness of our results. We show that the loan officers were attentive and thorough in their evaluations: capital allocation decisions responded to business characteristics in the application and were correlated with expectations of future performance, loan officers' beliefs and decisions were predictive of the actual future performance of the business, and business performance expectations were higher conditional on the business receiving additional capital. Moreover, our finding of a lack of discrimination is robust to a battery of tests that modify the identifying variation, variable definitions, and sample selection. We further assess the generalizability of our study, particularly to decisions on an initial loan application, using the SANS framework (List, 2020).

Our primary contribution is to identify the role of gender discrimination in gender gaps in business finance. The bulk of this literature relies on observational studies that estimate residual gender differences in survey data after controlling for observable characteristics. These studies have found mixed results, without a clear consensus on the existence of gender discrimination in business capital decisions.[6,7] A recent exception is a clever lab-in-the-field experiment, framed as a training session, in which loan officers from Turkiye predicted actual lending decisions for previous real loan applications in which business-owner gender was randomized (Brock and De Haas, 2023). These decisions were incentivized for accuracy based on the real outcome of the loan. Consistent with our findings, Brock and De Haas (2023) do not find evidence for gender discrimination on the extensive margin (i.e., whether to approve a loan) and find that loan officers do not differ in their beliefs about loan repayment by gender. However, they do find that women receive stricter conditions on credit offers. This is similar to Alibhai et al. (2019), who find support for gender discrimination on the intensive margin by loan officers in Turkiye. One other recent exception is a concurrent study from the high-income country Chile that uses well-identified experimental variation on gender in a large field experiment and does find gender discrimination in the related context of consumer credit (Montoya et al., 2020).

We build on this literature by identifying gender discrimination in business finance using experimental variation in a large-scale real-stakes natural context. We incentivized loan officers to make decisions based on their own capital allocation preferences, and their decisions had meaningful consequences for real businesses. Our experimental approach improves upon observational studies where financial providers' incentives are strong but gender discrimination is not well identified. Our real-stakes setting improves upon experimental studies that yield convincing identification but often lack strong incentives directly aligned with loan officer preferences. To the best of our knowledge, our paper represents one of the first experimental

[6]For example, Muravyev, Talavera and Schäfer (2007) find that female-managed firms are less likely to obtain a bank loan across 34 countries, primarily representing Central and Eastern Europe. In contrast, Aterido, Beck and Iacovone (2013) find that across Sub-Saharan Africa, the gender gap in using formal bank credit, and being rejected conditional on applying for a loan, disappears after controlling for the firm characteristics. Beck and Cull (2014) find some evidence that female-owned firms are more likely to have bank loans in Africa, likely reflecting survival bias. See Klapper and Parker (2011) for a more thorough review.

[7]A related literature explores credit decisions when clients and loan officers share traits, which suggests that discrimination may be an underlying phenomenon (Fisman, Paravisini and Vig, 2017; Beck, Behr and Madestam, 2017).

estimates of gender discrimination in a real-world business capital allocation decision.

An additional key contribution of our paper is to use financial providers' elicited beliefs to link their decisions to theoretical frameworks for discrimination. Previous research typically has not observed underlying beliefs, making our study one of the first to measure and estimate the accuracy of these beliefs across our entire sample. Bohren et al. (2023) show that comparing subjective beliefs with the true underlying distribution of outcomes is the best way to identify between sources of discrimination. Our incentive-compatible belief elicitation allows us to directly study whether loan officers' beliefs about gender and business performance align with their actual decision-making. Furthermore, by collecting data on true business performance, we evaluate the *accuracy* of these beliefs, the critical component to identifying whether gender discrimination is a profit-maximizing strategy that can persist in equilibrium. We do so by using a novel approach in which we benchmark beliefs against machine learning algorithms (i.e., a "synthetic" judge for the competition). This comprehensive approach enables a deeper understanding of the interplay between beliefs, decision-making, and the implications of gender discrimination in the realm of business finance.

The rest of the paper proceeds as follows. Section 2 describes the context in which we implement our study and our experimental design, including our empirical strategy for identifying discrimination and differential beliefs. We present our findings in Section 3, and discuss the generalizability of these results in Section 4. Section 5 concludes.

2 Context and Experimental Design

2.1 Ethiopian context

Ethiopia generally performs poorly on global indicators of gender equality. For example, in the World Economic Forum's 2016 Global Gender Gap Report, Ethiopia ranked 109 of 144. This low rank was driven by sub-indices related to labor and education outcomes: they ranked 106 on economic participation and opportunity and 132 on educational attainment. These stark gender differences suggest that gender discrimination (both belief-based and alternative mechanisms such as social norms or prejudice) may be present in various contexts in Ethiopia.

After the agricultural sector, the most common way women participate in the labor force in Ethiopia (and in Sub-saharan Africa) is as entrepreneurs. This highlights the importance of gender gaps in capital and business performance. Based on data from the Ethiopia Socioeconomic Survey, the World Bank (2019) documents that male business managers are 3.7 percent more likely to borrow and borrow approximately 50 percent more than their female counterparts.[8] There is increasing acknowledgment of these gender gaps, which has driven pol-

[8] A business manager is defined as an individual within a household in charge of the decisions regarding the

icy responses. For example, Ethiopia has a financial inclusion policy that specifically targets gender gaps, and many lending institutions are encouraged to lend to female clients.

2.2 The Business Plan Competition

In 2019, the Entrepreneurship Development Institute (EDI)[9] launched a business plan competition, EthioSpur, to provide capital and other awards to promising businesses. Business plan competitions are an increasingly common method to stimulate entrepreneurial growth in developing countries. For example, during the time of our own competition, we were aware of two other business plan competitions in Ethiopia itself.

The competition's prizes were 300,000 ETB, 220,000 ETB, and 140,000 ETB for the top three businesses.[10] In addition, the top 20 businesses were awarded with media and marketing coverage, and the top 100 were awarded with a "fast track to credit," as described in Section 2.5. The competition was promoted on a national level via social media, SMS, and targeted outreach by EDI staff.

To participate in the competition, business owners simply had to be the majority business owner in the business, be operational for at least four months, have a business license, and complete the application form.[11] Any business that would seek capital (including loans) from a formal financial institution would be expected to meet these minimal requirements (discussed in Section 4 in more detail).

We partnered with EDI to study whether loan officers, recruited to judge the competition, discriminated against female-owned businesses during the judging process. To identify gender discrimination, we used an approach similar to an audit study: recruited loan officers were given a packet of applications to evaluate in which the gender of the applicant had been randomly assigned.

We intentionally designed the application form to collect the same information used by loan officers when making initial decisions on loan requests. The application form collected information on current business characteristics (e.g., industry, profits, years of operation, etc.) and a business expansion plan (e.g., description of plan, how awarded funds would be used, expected revenue). The form also collected additional information on the business owner (e.g.,

---

earnings from an enterprise.

[9]The Entrepreneurship Development Institute, formerly Entrepreneurship Development Center, is a key agency tasked by the government of Ethiopia to increase entrepreneurship and economic growth, with specific attention to the needs of women entrepreneurs. A key element of EDI's mission is to improve access to finance.

[10]This corresponds to approximately 9,375, 6,875, and 4,375 USD at the time of the competition. These amounts are within the range of expected loans for smaller businesses. For example, in Alibhai (2021), a dataset of 357 female entrepreneurs interested in borrowing from Wasasa MFI in Ethiopia, the median loan request was 200,000 ETB.

[11]Businesses were not required to have a license at the time of the application, but were informed that they would be required to get a business license to receive any prizes.

marital status, age, gender).[12] To ensure that applicants were truthful, they were informed that all information would be audited and verified for winning businesses. If a business was found to have provided false or misleading information, it would not only be disqualified from the competition, but also from all future EDI initiatives. The application was designed to be simple and available in multiple languages, and the application could be submitted online, in hard copy, or via email.

2.3 Applicants

The competition attracted 916 businesses. Table 1 provides summary statistics on the median business performance of the applicants, overall and by business owner gender. The majority of businesses that applied to the competition are relatively small, but they are more successful and larger than the median Ethiopian business. The median years in the industry is 5 years for both male and female businesses. The median profit for the previous month is 15,000 ETB (500 USD), the median number of employees is 3, the median value of assets is 240,000 ETB, and the median value of liabilities is 4,000 ETB. This profile reflects the type of business that we expect to apply for capital. For example, our sample is similar to that in a previous study of female loan applicants at Wasasa Microfinance Institution: their median monthly business profit was 15,000 ETB, median number of employees was 1, the median age of the business was 4 years, and the median value of reported assets was 150,000 ETB (Alibhai, 2021).[13] Below the median, we report the mean and standard deviation. The mean is generally much higher than the median, highlighting that there is a significant right tail of larger businesses that applied to the competition. Relative to male businesses, we find that female-owned businesses report lower profits, have fewer employees, and have fewer assets and liabilities. In Section 3.3, we discuss the relevance of this gender gap in more detail.

Table 2 provides summary statistics for the business owners themselves. 44 percent of applicants were female-owned businesses. The sample is highly educated: we observe that nearly 50 percent report having a bachelor's degree or higher, though this varies by business owner gender. We similarly see that female-owned businesses have more children (1.95 vs 1.62), though we see smaller differences by gender on marital status (54 percent are married or cohabitating). Both genders report high levels of being a household head (86%) and having a high self-reported risk preference (8.64).[14]

[12] The complete application form can be found in Appendix A.

[13] We are not aware of a more comprehensive representative sample of loan applicants in Ethiopia to which we can compare our sample. We take the similarity in our sample and that of the study at Wasasa MFI as support that our sample overlaps well with our population of interest: businesses that are requesting capital.

[14] This includes sharing the status of the household head with a spouse.

Table 1: Applicants: Business Median, Mean, and Standard Deviation

	(1) Total	(2) Male	(3) Female
Years in Industry	5	5	5
	6.04	5.93	6.16
	(4.59)	(4.84)	(4.28)
Profits (thousands of ETB)	15	17	12
	182.16	292.82	43.69
	(1,604.38)	(2,141.17)	(170.80)
Employees	3	3	2
	14.79	11.09	19.44
	(201.79)	(108.49)	(277.74)
Assets (thousands of ETB)	240	248	221
	1,760.02	2,245.59	1,142.47
	(15,964.60)	(20,855.22)	(5,067.66)
Liabilities (thousands of ETB)	4	4.5	3
	644.89	1,024.13	168.65
	(7,948.47)	(10,621.61)	(750.25)
Observations	911	510	401

For each variable, the first row reports the median, followed by mean and standard deviation in parentheses. Profits refer to reported profits from the previous month. Profits, Assets and Liabilities are shown in thousands of ETB.

Table 2: Applicants: Mean Owner Characteristics

	(1) Total	(2) Male	(3) Female
Female	0.44	-	-
	(0.50)		
Bachelors Degree or Higher	0.49	0.56	0.39
	(0.50)	(0.50)	(0.49)
Married/Cohabitating	0.53	0.54	0.53
	(0.50)	(0.50)	(0.50)
Number of Children	1.76	1.62	1.95
	(1.70)	(1.80)	(1.56)
Household Head	0.86	0.85	0.87
	(0.35)	(0.36)	(0.33)
Self-Reported Risk Preference	8.64	8.60	8.69
	(2.15)	(2.16)	(2.14)
Observations	911	510	401

Table reports mean and standard deviation. Self-reported risk preference ranges from 0 to 10, increasing in risk tolerance.

2.4 Loan Officers as Judges

The competition was judged by loan officers recruited from lending institutions (i.e., banks and microfinance institutions) across Addis Ababa. Institutions were asked to provide experts who met the following criteria: (i) involved in reviewing applications seeking capital from the institution, with specific attention to urban clients, capital for business purposes, and individual applicants or enterprises;[15] (ii) employed as a loan officer or a member of the loan approval committee; and (iii) employed for at least one year at the institution. The purpose of these criteria was to ensure that those recruited were from the relevant population for reviewing requests for capital. Thus, just as applicants were real businesses interested in growth and capital, judges were real experts who reviewed and evaluated loans for businesses as their primary profession.

The recruited loan officers spanned 13 different lending institutions, representing a significant portion of the institutions in the financial sector serving Addis Ababa. 14 percent were female and 65 percent were recruited from microfinance institutions. On average, the loan officers had been at their respective institution for five years, and in finance for 11 years.

The loan officers reviewed the applications remotely and completed an evaluation form for

[15]This is in contrast to "social collateral" loans in which a group receives a loan with joint liability, commonly found in microfinance.

each application they reviewed. They underwent an orientation that was generally done over the phone or Internet due to the COVID-19 pandemic. The orientation included reviewing all questions in the evaluation form, signing a contract and a "comprehension check" in which they were explicitly asked about how the information in each section of the evaluation form would be used.[16] This ensured that the loan officers were aware of the definitions and objectives of each question. To protect against concerns of social desirability bias, the loan officers were not told ahead of time that their evaluations would be used in a study on gender and finance or that they were participating in an experiment. All communication with the loan officers, including the orientation, was conducted by the local project manager who was blinded to the key question of interest and to the randomized gender assignment. The loan officers were compensated 2,500 ETB for their time upon satisfactory completion of their evaluations. They were asked to complete their review of applications in two weeks, but extensions were granted.

The recruitment process highlights that these loan officers had several incentives to conduct a thorough and thoughtful review, in addition to payment contingent on quality. They had been handpicked by their respective institutions to serve as judges, in some cases based on formal Memorandums of Understanding (MoUs) between EDI and their institutions. EDI is a highly respected agency in Ethiopia with a key focus on developing and maintaining relationships with many of the financial institutions that provided the judges. Thus, there would be a reputational consequence both within their institution and in the broader financial sector for poor performance on the task. In Section 3.2, we provide evidence of the quality of their evaluations.

2.5 Evaluation Form: Treatment Salience and Outcomes

Treatment Salience: The loan officers reviewed digitized application forms that were translated into English and shown in a standard format.[17] An example of an application form is shown in Appendix Figure A1, with identifying information redacted. The top of the form provided demographic information about the applicant, including gender. The remainder of the form showed information about the business and the business plan.

The evaluation form that the loan officer completed for each reviewed business was divided into four sections (see Appendix Figure A2). Section A was designed to ensure salience of the randomly assigned gender without revealing the research question. This section asks

[16] The contract included agreeing to undertake "to perform the services with the highest standards of professional and ethical competence and integrity," reviewing the expertise requirements of a judge, remuneration (including the bonus based on the accuracy of beliefs described in Section 2.5), and non-monetary benefits (recognition on both EDI's website and at the awarding event for the competition).

[17] An exception to the translation requirement was made for detailed business plan narratives submitted in Amharic, the most prevalent local language in Ethiopia.

the loan officer to confirm basic demographics of the applicant: ID, age, gender, total years of experience, and whether the applicant was also employed outside of the business. The loan officers were informed that this section was used to verify that the correct application was being reviewed. In addition to ensuring that the evaluator was aware of the randomly assigned gender of the business owner, we used this section as a check that the loan officer was paying attention to the information in the application. 98.5% of evaluations noted the gender correctly.

To confirm that gender was not revealed in other parts of the application, our local survey firm explicitly reviewed the digitized application materials and confirmed that there was no information in any digitized form that would reveal the gender of the applicant. In addition, in the sample of applicants, both genders are represented across all twenty industry categories included in the application form, suggesting that the industry alone would not reveal the gender of the applicant. Similarly, EDI was not concerned that businesses were gendered to the extent that the type of business would reveal the true gender of the business owner or cast doubt on the credibility of the randomly assigned gender of the business owner. This suggests that any observed gender discrimination would likely not be due to the business owner's gender being revealed or surprising.

Beliefs on Future Performance: Section B asked the loan officers to provide a prediction of the business' performance in January 2021, exactly one year after the submission of applications. Importantly, the majority of the evaluations were completed only a few months prior to January 2021. As a result, the loan officers were well aware of the shocks in the economy, including those related to the COVID-19 pandemic, at the time of their predictions. They were asked to provide these predictions for two scenarios: if the business did or did not win 300,000 ETB (i.e., the amount of the top prize) from the competition. The loan officers predicted the likelihood of survival, monthly profit, capital stock, and number of paid employees in these two scenarios.

Our interviews with loan officers indicated that a business's future profitability is a key metric used when deciding whether to allocate capital to a business at their institution. Loan officers said they aim to allocate loans toward more promising businesses, in part because the probability of repayment and returns to capital is increasing in business performance. In an exit survey of 43 loan officers who served as judges, 86 percent reported that growth potential (i.e., future profits) was either an important or very important factor when determining whether to approve a loan. We also asked the loan officers explicitly about the likelihood of repayment for a 3-year loan for 100,000 ETB.

This section on beliefs was incentivized for accuracy. The loan officers were informed that the person with the most accurate evaluations for Section B would receive 15,000 ETB

(500 USD). We intentionally did not provide details on how accuracy is determined for the bonus, consistent with Danz, Vesterlund and Wilson (2022) who find that false reports of beliefs are *lower* when subjects are simply told that the payment rule is designed to maximize their payment, rather than being given additional information of the details of the incentive-compatible payment calculation.[18] We similarly chose to have one large bonus, following the guidance from Charness, Gneezy and Halladay (2016) that randomly paying a large amount to one subject is as effective as paying a smaller amount to every subject. A conservative interpretation of our approach is that the loan officers' perception of our belief elicitation is equivalent to simply asking them for their best prediction. Charness, Gneezy and Rasocha (2021)'s review of the literature finds that such introspective belief elicitation performs equally well as "state-of-the-art" complex belief elicitation methods (e.g., quadratic scoring rule). This is particularly true with respect to accuracy, the extent to which elicited beliefs match the objective probabilities of an event, which is the purpose of our belief elicitation (Trautmann and van de Kuilen, 2014). Charness, Gneezy and Rasocha (2021) conjecture that increasing simplicity is an important dimension for improving belief elicitation.[19]

The loan officers were also informed that their responses in this section would have no bearing on the awarding of the capital from the competition.[20] Because we analyze beliefs as a mechanism underlying capital allocation decisions, we designed the evaluation form to allow for independence in the decision-making process between providing beliefs on business success and capital decisions. In this way, we ensured that loan officers had no incentive to manipulate their stated beliefs about business performance in order to influence capital allocation outcomes.[21]

[18] Danz, Vesterlund and Wilson (2022) compare false reports of subjects' beliefs when they are told "the payment rule is designed so that you can secure the largest chance of winning the prize by reporting your most accurate guess," to a treatment arm in which in addition to the statement, they were also given details on how the binarized scoring rule, a state-of-the-art elicitation, is calculated. Similarly, Charness, Gneezy and Rasocha (2021)'s review of the literature highlights that in most cases where complex scoring rules that are incentive-compatible to truth-telling are used, subjects are also explicitly told by the researchers that telling the truth will maximize their payment. This further suggests that the effectiveness of incentive-compatible structures may be driven by respondents responding to the researchers' claim that truth-telling maximizes payment.

[19] In general, there are two primary concerns with non-incentivized introspection: resorting to defaults or random responses, or explicit bias in reporting due to factors such as experimenter demand effects (Charness, Gneezy and Rasocha, 2021). In terms of the former, we show in Section 3.2 that elicited beliefs are both predictive of future performance and responsive to baseline information, confirming that the loan officers did not respond randomly. Section 3.2 provides evidence for the loan officers being thoughtful and attentive in their evaluations. In addressing the latter concern, Section 4 provides details on the limited role of experimenter demand and social desirability biases in our context.

[20] In the orientation, the loan officers were told that "this information is collected to understand what characteristics determine business success. We are collecting information from experts, such as yourself, since you are best equipped and knowledgeable to predict a business's success."

[21] Our incentive structure ensured that the loan officers did not face any external incentive to align their stated beliefs with their capital allocation preferences. However, we would not expect the outcomes themselves

Capital Allocation Decisions: Section D was the loan officer's overall score for the business that was used to determine the competition's prizes. The loan officers were asked to score the business on overall impression, value proposition, and entrepreneurial credibility with a range of 1 to 10 each. This was then aggregated into a final score using the following formula: $FinalScore = OverallImpression + .5 * (ValueProposition + EntrepreneurialCredibility)$. Importantly, the loan officers were informed that this was the *only* measure that would determine the competition's winners.

Following Section D, the loan officers were asked whether they wanted the applicant's information to be sent to their institution for consideration for a loan. This question served as a proxy for capital allocation decisions from the provider's own lending institution, particularly for decisions on the initial application for a loan request.

Identifying promising businesses as potential borrowers was a key part of discussions with lending institutions when forming partnerships for the competition, beginning with the initial recruitment of loan officers from lending institutions to serve as judges. Several institutions signed MoUs that explicitly committed the lending institution to provide identified applicants with a fast-track loan process.[22] Thus, the leadership at lending institutions was aware from the start of their involvement that identifying potential borrowers and facilitating access to loans was of key interest. Similarly, in the orientation for recruited loan officers serving as judges, they were informed that EDI would "let your institution know you recommend this person to be reviewed for a loan and forward this information to your institution if the applicant is interested in a loan." Hence, the decision to forward the applicant to their own lending institution was a meaningful proxy for that institution reviewing the applicant as a potential borrower.

Our main outcomes of interest for capital decisions are 1) the loan officer's final score from Section D, and 2) whether they requested the applicant's information be sent to their own institution for consideration for a loan.

Additional components in the evaluation form: Section C collected additional information about the loan officer's beliefs about the business owner. They were asked to evaluate the business owner's managerial skills, sources and amount of capital for the business, market demand for the business, and whether the business was the primary source of income for

to be independent. For example, if a loan officer perceives a business as being more likely to succeed, we would naturally expect they may prefer to allocate capital to it, particularly for loans.

[22] Though MoUs may not be easily legally enforceable in practice, they are meaningful agreements. This is supported by the fact that the institutions took time and negotiation to reach an agreement, the agreements differ by institution, and not all institutions were ready to make such a commitment. Ultimately, because of the disruptions caused by COVID-19 in the interim and the delays in the judging process, EDI did not track what happened to promising businesses after they were forwarded to the lending institutions. However, loan officers and lending institutions would have considered their decisions consequential at the time they were making them as they would not have known that EDI would not follow up with applicants.

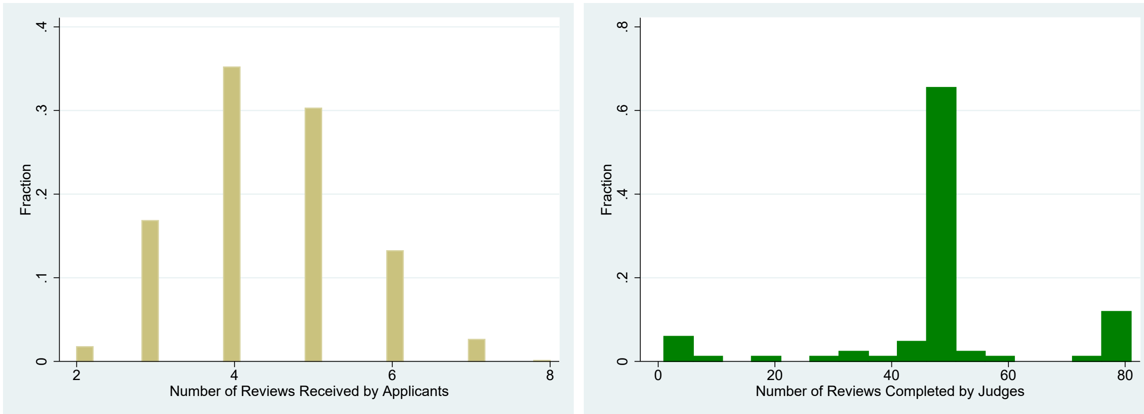

Figure 1: Reviews per Application and Evaluation per Loan Officer

the household. This section was not incentivized. It was designed to shed light on potential beliefs that did not affect business performance but could be influenced by gender and affect an evaluation of a business.

2.6 Random Assignment

We created duplicates of the original applications so that each application would be shown at least once as a male-owned business and as a female-owned business.[23] We then randomly assigned those applications to be shown as either male or female to each loan officer, in a random order, with each loan officer receiving 48 or 49 application forms (referred to as their "application packet").[24] Each application was reviewed multiple times with a randomly assigned business owner gender, and each loan officer evaluated multiple applications. Figure 1 illustrates that the median number of reviews per business was 4, and that the median number of evaluations completed by a loan officer is 48.

Due to COVID-19, there was a delay between the submission deadline to the competition and the evaluation process, and the evaluation process itself took longer than planned. The competition closed on January 20, 2020, and evaluations were conducted from September 2020 to December 2020 [25]. Thus, the loan officers were aware of COVID when they scored the applications and predicted business success. Loan officers benefited from the delay because they had more contextual information to predict business performance in January 2021, given information on past business performance in January 2020.

[23] Each application is duplicated two to eight times. The number of times they were duplicated was randomly determined, and the median number of reviews was 4.

[24] If it was the case that an applicant was assigned to the same loan officer twice, we simply dropped one of the application forms before providing the packet to the loan officer.

[25] A few loan officers also returned packets after the December deadline.

2.7 Estimating Equations: Identification of Gender Discrimination

We estimate whether capital allocation decisions differed when the business owner was randomly assigned to be shown as male using the following estimating equation:

$$Y_{ij} = \beta_1 * RandomlyAssignedMale_{ij} + \alpha_i + \alpha_j + \epsilon_{ij} \tag{1}$$

where $RandomlyAssignedMale$ indicates that applicant i assigned to loan officer j was shown as a male. The specification includes applicant and loan officer fixed effects and uses robust standard errors.[26] We study two pre-specified outcomes that reflect capital allocation decisions. The first outcome is the overall final score given to the application, which determined the winners of the business plan competition. The second outcome is an indicator of whether the loan officer selected the business application to be forwarded to their institution for consideration for a loan.

We next estimate Eq 1 on a pre-specified set of loan officer predictions of future business performance in the upcoming months: survival, profits, and assets. We estimate these for the loan officer's beliefs on expected business performance with and without having received additional capital. We use the differences in these predictions as a measure of the loan officer's expectations on the return to capital as a function of gender.

We limit our primary analysis sample to evaluations in which the loan officers completed all our pre-specified outcomes. Our primary analysis sample consists of 3,696 completed evaluations of 916 businesses by 84 loan officers. In this sample, 910 businesses were evaluated by multiple (2 to 8) financial providers, and 83 loan officers reviewed multiple applicants (2 to 79). 82 of these loan officers had variation in the gender of the applications they reviewed.

2.8 Ethical Considerations

As in all audit study designs, our methodology uses deception by randomizing the gender depicted in the application that a loan officer is reviewing. The justification for using deception in audit studies is that no alternative method exists to rigorously identify discrimination, as was the case in our setting. Given the scarcity of studies identifying gender discrimination in business finance and in low-income country settings, we argue that the benefits of the research justified the design. The study was approved by the IRB at UC Merced. It was also approved by the Entrepreneurship Development Institute, the local organization with whom we collaborated. EDI is a highly respected institution in Ethiopia and had a reputational stake in the study. All the loan officers who served as judges were debriefed and informed

[26] Since applications are randomly assigned to loan officers, there is no need to cluster at the loan officer level (Abadie et al., 2023).

after the completion of the study that demographic information was manipulated for research purposes in the applications they were reviewing.

Another ethical concern with audit studies is the time spent by experts in reviewing fake materials. In our case, experts were evaluating real businesses for a real business plan competition, and they were compensated for their time.[27]

3 Results

3.1 Identifying Discrimination

We find that the randomly assigned gender of the business owner did not affect capital allocation decisions by loan officers, neither for the capital prize in the competition nor for consideration of a loan at their own institution. Table 3, Column 1 finds that the final score, which was used to determine who would be awarded the competition prizes, is not statistically different whether the applicant was shown as male or female. In fact, when applicants were shown as male, they received slightly lower scores. The point estimate for the difference in scores is 0.105 points (on a scale from 2 to 20), which amounts to a difference of less than .03 standard deviations. The 95 percent confidence interval for the differences in scores is similarly very small (-.337 to .127), a range of merely -.07 to .03 standard deviations. These results suggest that loan officers did not discriminate by applicant gender in the allocation of capital in the business plan competition. Columns 2 through 4 document differences in each component of the final score (each ranging from 1 to 10), and we continue to find no meaningful differences by randomly assigned business-owner gender.

We then turn to the decision of whether the loan officer wanted to forward the application to their own institution. Randomly assigned business-owner gender did not affect the loan officers' decision to send the applicant's information to their own institution for consideration of a loan (see Table 3, Column 5). We also explicitly asked the loan officers for their beliefs about the applicant's ability to repay a loan and find no significant difference in their expectations of either strategic default or default due to lack of resources based on randomly assigned business-owner gender.[28] The point estimate on the difference in the decision to forward the applicant is less than .01 percentage points. This highlights that the loan officers did not

[27] An additional ethical concern is the scores given to the applicants for the business plan competition. If we had observed discrimination, there were two possible ways we would have proceeded: using only real gender or using only one gender when determining scores to award the competition prizes. However, since we did not observe gender discrimination, EDI chose to use all evaluations in determining the prizes.

[28] In 19 and 13 percent of evaluations, the loan officers believed the applicant would be unable to repay a loan or strategically default, respectively. Differences by randomized gender were .7 and .1 percentage points. The loan was described as being for 3 years for 100,000 ETB. This outcome was not a primary outcome in our pre-analysis plan, and was only pre-specified conditional on loan consideration being uninformative.

Table 3: Causal Effect of Gender on Capital Allocation Decisions

	(1) Score	(2) Overall Impress	(3) Value Prop	(4) Entrepreneurial	(5) Loan
Male	-0.105	-0.0478	-0.0550	-0.0596	0.00159
	(0.116)	(0.0611)	(0.0626)	(0.0650)	(0.0140)
Observations	3696	3696	3696	3696	3696
Female Mean	12.06	5.990	6.079	6.069	0.495

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, determined by Overall Impression (Overall Impress) + .5* Value Proposition (Value Prop) + .5*Entreprenuerial Credibility (Entreprenuerial). Each of these subscores is on an increasing scale of 1 to 10. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

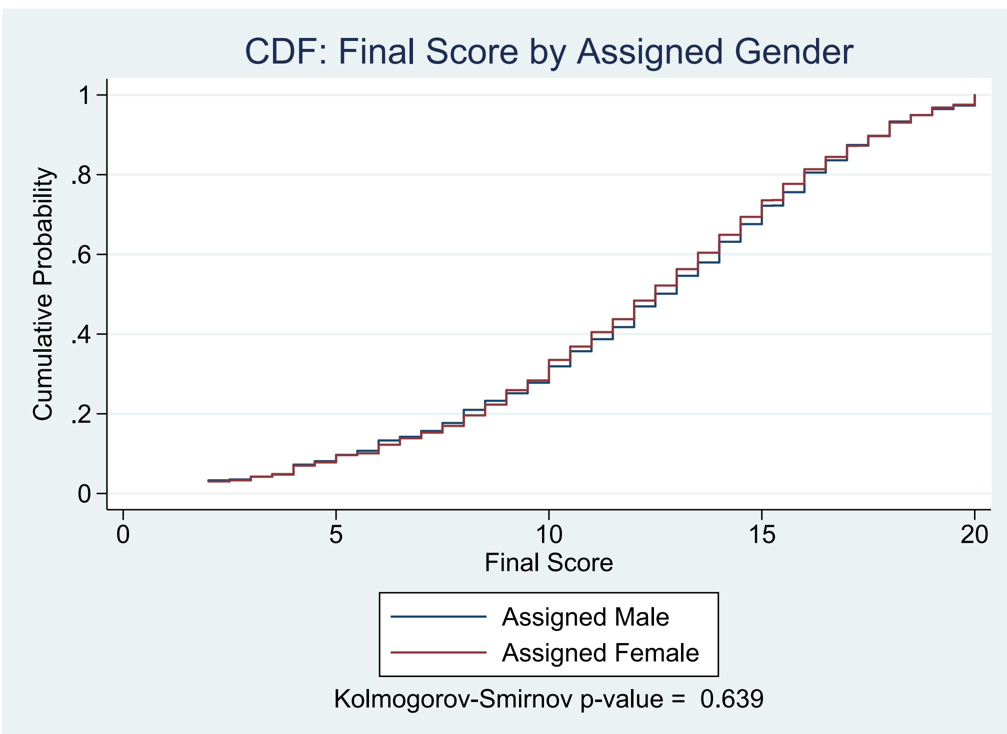

Figure 2: CDF of Final Score by Randomly Assigned Gender

discriminate even when making decisions relevant to their own institution.

The standard errors for both estimates are very small, allowing us to rule out any meaningful differences in how the application was treated as a function of the randomly assigned gender of the business owner. The similarities we observe across both capital allocation outcomes suggest that there is significant external validity across the two settings.

This lack of discrimination is consistent across the distribution of scores in the competition (Figure 2). The figure highlights that throughout the distribution of business quality, randomly assigned gender had no meaningful effect on the evaluation of the business. A Kolmogorov-Smirnov test finds no statistical difference between these two distributions, with a p-value of 0.639. Similarly, we find no differences in the variance of final scores by gender.[29] Note that the KS test also fails to reject the null of first-order stochastic dominance, indicating that the data is consistent with statistical discrimination alone (i.e., no taste-based discrimination), including no discrimination overall (Bharadwaj, Deb and Renou, 2024).[30]

We generally find no evidence of gender discrimination along several pre-specified dimensions of heterogeneity. We do not find discrimination conditional on the business owner's marital status, education, or number of children (see Table 4). We do observe gender discrimination against female widows for consideration of a loan. This is consistent with female widowhood signaling unique vulnerability and access to fewer resources.

We also find no evidence for discrimination conditional on business characteristics. First, we explore discrimination based on whether the business industry is male-dominated, a pre-specified characteristic. If female business owners face discrimination in male-dominated industries, which tend to be more profitable, this could be an important driver of the gender profit gap. We asked our local survey firm to have two employees review the products and services described in the application and categorize the business as belonging to an industry dominated by women, dominated by men, or neither.[31] Second, we present exploratory analyses based on business performance, as measured by profits and size. Even if there is no gender discrimination on average, if high-performing female business owners face discrimina-

[29] We test for differences in variance using the STATA command sdtest and robvar, reflecting the proposed tests by Levene (1960) and the alternative specifications proposed by Brown and Forsythe (1974).

[30] Bharadwaj, Deb and Renou (2024) propose that first-order stochastic dominance implies that taste-based discrimination must be present; i.e., that such a pattern in data cannot be supported by statistical discrimination alone.

[31] For each application, employees were requested to answer the following two questions with Yes, No, or Unsure: In your opinion, are over 90 percent of businesses that supply the main product described [in the application] run by women [men] (i.e., are over 90 percent of the business owners of such businesses female [male])? In practice, the employees appeared to categorize businesses as being dominated by a particular gender using a threshold lower than 90 percent. We use this question to define indicators for male or female industries for the businesses that were marked affirmative for each of these respective questions. 28 percent of applications were coded as female-dominated, 38 percent as male-dominated, 30 percent as unsure, and 3 percent were missing.

Table 4: Heterogeneity by applicant characteristics

	(1) Score	(2) Score	(3) Score	(4) Loan	(5) Loan	(6) Loan
Male	-0.168	-0.313	-0.143	-0.0144	-0.0343	-0.0115
	(0.191)	(0.423)	(0.177)	(0.0227)	(0.0530)	(0.0226)
Male × Married=1	0.152			0.0261		
	(0.249)			(0.0300)		
Male × Separated=1	-0.455			-0.0662		
	(0.511)			(0.0569)		
Male × Widowed=1	0.388			0.214**		
	(0.749)			(0.0930)		
Male × Highest Education		0.0254			0.00459	
		(0.0614)			(0.00772)	
Male × Number children			0.0482			0.00981
			(0.0736)			(0.00962)
Observations	3602	3605	3093	3602	3605	3093
Female Mean	12.06	12.06	12.06	12.06	12.06	12.06

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 2 to 20. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table 5: Heterogeneity by business performance

	(1) Score	(2) Score	(3) Score	(4) Loan	(5) Loan	(6) Loan
Male	-0.239	-0.0798	-0.120	0.0261	0.00108	-0.000677
	(0.215)	(0.117)	(0.132)	(0.0243)	(0.0149)	(0.0161)
Male × Male-dominated industry	0.247			-0.0336		
	(0.285)			(0.0335)		
Male × Female-dominated industry	0.137			-0.0407		
	(0.293)			(0.0350)		
Male × Baseline profits		0.0431			0.000548	
		(0.0594)			(0.00791)	
Male × Number employees			0.00185			0.000298
			(0.00980)			(0.00139)
Observations	3696	3367	3593	3696	3367	3593
Female Mean	12.06	12.06	12.06	0.495	0.495	0.495

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 2 to 20. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. Baseline profits are in units of 1,000,000 ETB. The number of employees is winsorized at the 99 percent level. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

tion, this could explain why they are not able to grow further. We do not find support for discrimination within gendered industries using our survey firm's categorization of industries as female-dominated or male-dominated, nor as a function of business baseline profits or the number of employees (see Table 5).

In addition to testing for discrimination within subsets of business type, we also look for differences by whether the judging loan officer was employed at a microfinance institution (MFI) or a woman. We find no heterogeneity in discrimination based on these characteristics (Table 6). Though MFIs often prioritize female clients, none of the MFIs that participated in the judging serve women exclusively. In our exit survey of loan officers (N = 43), no loan officer reported having a portfolio of borrowers of only one gender. The highest share of female borrowers in a loan officer's portfolio was 82.5 percent.[32]

By itself, these capital allocation decisions by the loan officers are consistent with a continuum of taste-based and belief-based discrimination. Consistent with a lack of belief-based discrimination, loan officers predict similar future business performance for applicants shown as male or female. As described in Section 2.4, the loan officers were asked to predict business performance one year after the application submission. Table 7 finds no difference in expectations of the business' profit (Column 1), survival likelihood (Column 2), or assets (Column

[32]We pre-specified additional loan officer characteristics for heterogeneity tests based on information collected in an exit survey. However, our response rate on the exit survey was only 63 percent (43 loan officers), and thus we do not report these additional tests.

Table 6: Heterogeneity by loan officer characteristics

	(1) Score	(2) Score	(3) Loan	(4) Loan
Male	-0.0540	-0.0694	0.0135	0.00291
	(0.223)	(0.126)	(0.0267)	(0.0155)
Assigned Male=1 × MFI=1	-0.0718		-0.0167	
	(0.281)		(0.0324)	
Assigned Male=1 × Fem. Loan Officer=1		-0.257		-0.00954
		(0.374)		(0.0429)
Observations	3695	3695	3695	3695
Female Mean	12.07	12.07	0.495	0.495

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. MFI is an indicator for whether the loan officer was employed at a microfinance institution. Fem. Loan Officer is an indicator for whether the loan officer was female. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

3) as a function of the business owner's gender. This lack of difference in expected business performance remains true for both predictions without additional capital (Panel A) and with additional capital (Panel B).[33]

These results are robust to comparing the CDF of expected profits and assets by gender (see Figure 3). In each scenario and outcome, Kolmogorov-Smirnov tests fail to reject equality across the two distributions. We also find no differences in the variance of these distributions by gender, except for profit predictions in the condition with capital, where we observe a slightly higher variance in expected profits with additional capital among female-owned businesses.[34] Taken as a whole, we find that the loan officers did not expect gender differences in a business's growth potential on average, even after receiving a capital infusion.

The finding that loan officers do not believe that gender is predictive of business performance is further supported by the lack of heterogeneity in capital allocation decisions as a function of missing information in the application. If loan officers believed that gender had predictive value in business performance, then the standard theory of statistical discrimination suggests that the loan officers would be more likely to consider gender in their capital allocation decisions when the application is missing key information. However, if the loan officers did not believe that gender is correlated with business performance, then we would expect to continue to see a lack of gender discrimination. We test for heterogeneity in the capital

[33] Appendix Table A14 finds no support for differences in the beliefs about return to capital by business owner gender. Appendix Table A1 includes beliefs on employment, an additional prespecified variable.

[34] This difference in variance is not robust to using winsorized levels of profit expectations.

Table 7: Effect of Gender on Business Performance Beliefs

	(1) Survival	(2) Win. Profits	(3) Win. Assets
Panel A: Without capital			
Male	-0.0944	1.665	60.09
	(0.636)	(4.208)	(46.85)
Observations	3696	3696	3696
Female Mean	50.47	42.41	778.4
Panel B: With capital			
Male	-0.0339	-8.534	52.75
	(0.666)	(7.895)	(65.42)
Observations	3696	3696	3696
Female Mean	60.08	84.57	1089.4

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Survival, Profits, and Assets are loan officer expectations with and without additional capital. Profit and Assets are in thousands of ETB. Survival is the probability of survival, from 0 to 100; Win. specifications winsorize the variables at 1 percent. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

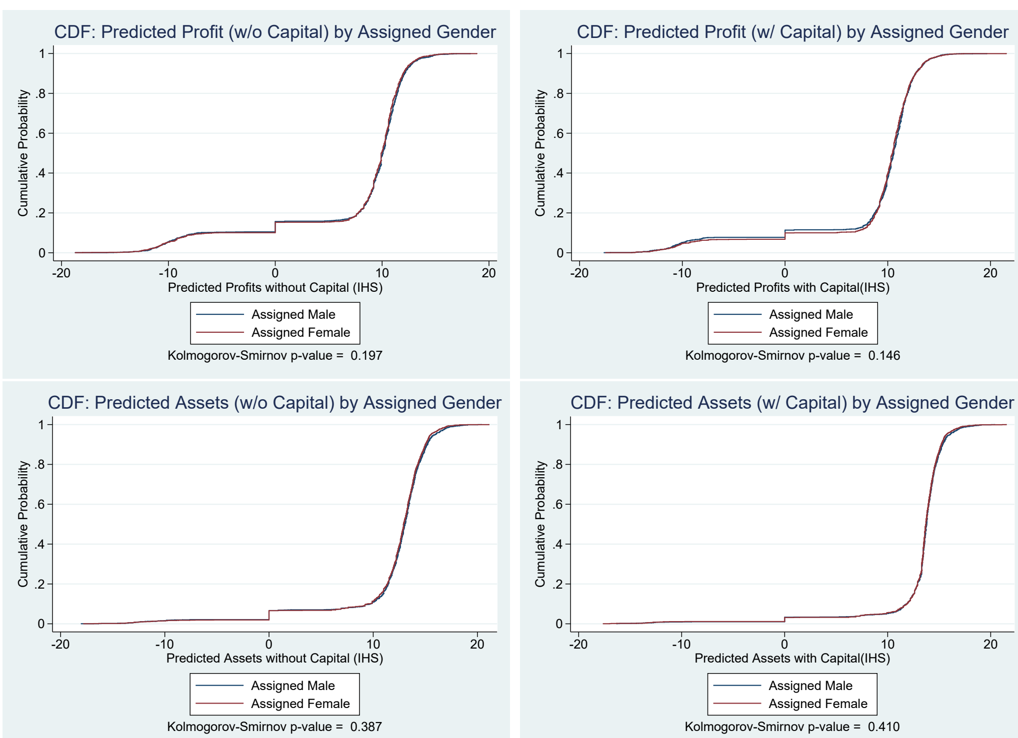

Figure 3: Distribution of Profits and Assets by Assigned Gender

Table 8: Heterogeneity by Missing Information in the Application

	(1) Score	(2) Loan
Male	-0.0230	-0.00498
	(0.139)	(0.0161)
Male=1 × Mssng Indx	-0.112	0.00895
	(0.140)	(0.00984)
Observations	3696	3696
Female Mean	12.06	0.495

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 2 to 20. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. Mssng Indx ranges from 0 to 8 and is a count of missing the following information in the application: profits, number of employees, total assets, total liabilities, years of operation, years of experience in the industry, projected employees, and projected revenue. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

allocation decision as a function of missing information, defined as a count of how many of the following data points are missing in the application: profits, number of employees, total assets, total liabilities, year of operation, years of experience in the industry, projected employees, and projected revenue. Table 8 finds no evidence for increasing discrimination against female-owned businesses in applications with more missing information.[35]

These results indicate that loan officers do not have belief-based partiality with respect to business owner gender. Combined with the lack of discrimination in the capital allocation decisions, we can conclude that loan officers exhibit neither belief-based partiality nor preference-based partiality (i.e., they do not engage in statistical discrimination or taste-based discrimination).

We conducted an additional battery of pre-specified robustness tests that confirm the lack of discrimination: weighting evaluations so that each loan officer has equal weight (Table A2 and A3); controlling for the order in which evaluations were assigned (Table A4 and A5); using the gender as reported by the loan officer (Table A6); excluding 5 percent of loan officers with

[35]As expected, missing information does reduce evaluations overall. Table 8 is unable to identify the effect of missing information as this is perfectly collinear with the application fixed effects. However, a specification that removes application fixed effects finds that missing information significantly reduced the score provided (coef = -.837 with a p-value of 0.000) and consideration for a loan (coef = -.07 with a p-value of 0.000), and does not yield coefficients that are statistically different from those reported in Table 8.

the least amount of variation in their final score (Table A11); limiting the sample to the first five applications given to the loan officer (Table A12); and removing loan officer fixed effects (Table A13). We also confirm robustness to limiting the sample to loan officers who passed various attention and internal consistency checks: correctly answering 75 and 100 percent of the verification questions (Table A7 and A8), baseline information in the application predicted the final score with a p-value of less than .15 (Table A9), and prediction of profits and firm survival with capital were higher than predictions without capital (Table A10). The main finding that there is no discrimination in the evaluation of businesses is remarkably robust.

3.2 Robustness of Evaluations

We provide several pieces of evidence that the loan officers were attentive and thorough when evaluating businesses. First, though randomly assigned gender did not affect evaluations, we find that the loan officers did consider other aspects of the business when evaluating the applicant. Table 9[36] shows that businesses with higher profits, greater assets, and business plans that projected greater employees and revenue were more likely to receive higher scores for the competition and to be forwarded to the lending institution. Evaluation outcomes are predicted by baseline business information, which indicates that the loan officers reviewed the businesses with effort and attention.

Second, the loan officers completed the initial verification section of the evaluation form with high accuracy. As described in Section 2.5, the loan officers were asked to verify the applicant's gender and other demographic characteristics before filling out the evaluation. They correctly indicated the applicant's gender in 98.5 percent of evaluations, the applicant's age in 97 percent of evaluations, the applicant's experience in 96 percent of evaluations, and the applicant's employment status in 95 percent of evaluations.

Third, evaluations were internally consistent in several ways. The loan officers predicted businesses would have better performance with capital than without in the vast majority of evaluations. In 92 percent of evaluations, the loan officers predicted that the business would be as or more likely to be operational in a year if they received additional capital than if they did not. We observe similarly high percentages of internally consistent evaluations with and without capital for the projected number of employees (93 percent), capital stock (93 percent), and profits (84 percent).

Fourth, businesses with stronger predicted performance were more likely to be awarded capital. Table 10 finds that loan officers provided higher scores and were more likely to consider for a loan those businesses that they believed were more likely to survive, have higher profits, and have greater assets. Using our endline survey, we also confirm that both the final score

[36]These estimations were pre-specified in our analysis plan.

Table 9: Baseline Business Characteristics Predictive of Capital Allocation Decisions

	(1) Score	(2) Loan
Profits (IHS)	0.136***	0.0104***
	(0.0201)	(0.00173)
Employees	0.000141	0.0000563**
	(0.000359)	(0.0000269)
Assets (IHS)	0.254***	0.0190***
	(0.0258)	(0.00255)
Liabilities (IHS)	-0.0108	-0.000353
	(0.0116)	(0.00135)
Initial Yr	0.0197*	0.00125
	(0.0114)	(0.00101)
Projected Employees	0.00313**	-0.0000309
	(0.00143)	(0.000104)
Projected Revenue (IHS)	0.224***	0.0137***
	(0.0290)	(0.00300)
Industry Exp.	0.0136	0.00239
	(0.0155)	(0.00187)
Observations	3696	3696
F	40.59	28.18
pvalue	4.41e-57	7.79e-35

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each independent variable is interacted with an indicator for the variable being missing, which is not shown. All independent variables are information reported by the applicant and viewed by the loan officer. Specifications include loan officer fixed effects, and standard errors are clustered by application.

Table 10: Final Score Correlates with Business Performance Beliefs

	(1) Survival	(2) Survival	(3) Win. Profit	(4) Win. Profit	(5) Win. Assets	(6) Win. Assets
Score	2.579***		7.733***		99.40***	
	(0.100)		(0.946)		(15.13)	
Loan		15.53***		53.45***		499.9***
		(0.830)		(8.251)		(88.49)
Observations	3696	3696	3696	3696	3696	3696

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Survival, Profits, and Assets are expectations of the loan officers with and without additional capital. Profit and Assets are in thousands of ETB. Survival is the probability of survival, from 0 to 100; Win. specifications winsorize the variables at the 1 percent. Specifications include loan officer fixed effects. Standard errors are clustered at the application.

and loan consideration decision were predictive of firm survival and profits 18 months after the competition (Appendix Table A16).

Finally, loan officers had significant variation within their own evaluations, suggesting that they were thoughtful in evaluating the information in the application. The average range of scores used by a loan officer in their evaluations is 13.8 out of a possible 18, and the average standard deviation for final scores within a given loan officer is 3.6. Loan officers recommended 50 percent of their businesses for consideration for a loan, on average, and all loan officers except five forwarded at least one application to their lending institution. None recommended all of the businesses they reviewed to be considered for a loan.[37]

3.3 Implications for capital targeting and accuracy of beliefs

A key contribution of our paper is to compare loan officers' decisions and beliefs with the actual gender gaps in business performance. Since loan officers did not discriminate by gender in their capital allocation decisions, if gender were meaningfully predictive of profits, this would suggest a trade-off between gender equity (i.e., the lack of gender discrimination) and profit maximization (i.e., targeting high-performing businesses). Comparing the loan officers' subjective beliefs with real future business performance also determines the accuracy of those loan officers' beliefs. Understanding the accuracy of loan officers' beliefs is critical, as inaccurate beliefs that result in different capital allocation decisions have significant policy and welfare implications.[38]

[37]These statistics are based on 83 loan officers that are used in our main sample and had more than 1 evaluation.

[38]Some of the analysis in this section deviates from our original pre-analysis plan. This is because the original pre-analysis plan considered only statistical significance, as opposed to meaningful differences in the theoretical

In models of accurate statistical discrimination, the justification for discrimination is that considering gender can improve predictions about the future performance of businesses. That is, if financial providers have accurate beliefs that female-owned businesses perform differently, conditional on all other observable information, belief-based discrimination can be profit-maximizing. In our case, loan officers' beliefs do not differ by business-owner gender. If these beliefs were accurate, this would imply that loan officers' decisions are consistent with targeting capital toward the highest-performing businesses. However, it may be the case that loan officers' beliefs are inaccurate, and business-owner gender truly has meaningful predictive value in identifying business performance. This would imply that the lack of gender discrimination in their capital allocation decisions (i.e., gender equity) comes at a cost of selecting lower-performing businesses (i.e., profit maximization). We assess this by connecting the loan officers' results to real performance outcomes using endline measures of business performance, based on a follow-up survey of applicants 18 months after the competition.[39]

A common way in which beliefs are tested for accuracy is to compare the beliefs to outcome data (Bohren et al., 2023; Manski, 2004). In our case, this is comparing loan officers' beliefs on future business performance with the actual gender difference in future business performance as measured in our follow-up survey. Relative to female-owned businesses, an OLS regression estimates that male-owned businesses were 2.85% (*standard error* = 2.35%) more likely to still be operational, earned 4,593 ETB (*standard error* = 2,364 ETB) more profit, and had 180,000 ETB (*standard error* = 130,900 ETB) more assets in our follow-up survey. For all three outcomes, we find that the real mean gender difference is within the 95% confidence interval of the loan officers' beliefs (Table 7). Thus, the direct comparison of subjective mean beliefs with mean outcomes fails to reject that the loan officers' beliefs are accurate.[40]

While this outcome-based test provides one measure of the accuracy of beliefs, it stops short of assessing whether loan officers' beliefs are different enough to be considered inaccurate in the framework of statistical discrimination. The motivation for assessing the accuracy of beliefs is that incorrect beliefs about gender differences in business performance may lead to misidentifying high-performing businesses and allocating capital away from these businesses. In our case, we found that loan officers did not believe business-owner gender predicted business performance. Therefore, the critical question of interest is whether loan officers *should*

context of discrimination or differences across estimations. Nonetheless, we do report main outcomes from the original pre-analysis plan in Appendix Table A16 and A17, and provide explanations for the limitations of the original pre-specified analysis plans where relevant. In general, the analysis put forth in this section reflects a more accurate and precise analysis plan that maintains the purpose of the original. In addition, we do not conduct simulations on policies to reduce discrimination, as we did not find support for discrimination.

[39] Appendix Table A15 confirms that our main results on discrimination in capital allocation are robust to the sample for which we successfully survey at endline.

[40] While loan officers had access to all of the business information in the competition's application, the regression for estimating the gender gap using the follow-up survey controls for only a subset of that information.

have responded to gender when identifying the most promising businesses, given true future business outcomes. That is, could the targeting of capital have been meaningfully improved by belief-based discrimination? If not, then we can conclude that the loan officers' beliefs were "accurate enough", as their beliefs did not result in a capital allocation decision that differed from what a model of accurate statistical discrimination would have predicted.

We determine whether loan officers "should" have responded to gender by assessing whether gender is an important variable when constructing an optimal prediction of business outcomes.[41] One possible prediction model is simply an OLS regression of business outcomes on all information shown to loan officers. A crude measure of the predictivity of this regression is the R^2.[42] We can quantify the predictive value of gender by comparing the R^2 in a regression with and without gender. Using a regression of winsorized endline profits on all quantifiable information shown to loan officers, we find that including gender increases the R^2 by only 1.9% (0.2447 versus 0.2401). This is consistent with the loan officers' beliefs being accurate: accounting for gender did not meaningfully improve model predictivity. However, a major concern with this naive linear regression is overfitting, where the model captures noise and random fluctuations in the data rather than the underlying relationship between the predictors and the outcome.

Machine learning methods aim to construct prediction models that reduce overfitting and thus minimize out-of-sample prediction error. A more robust way of quantifying the importance of gender is to use a machine learning algorithm to select optimal predictors of endline business profits. We consider the set of all quantifiable variables shown to the loan officers during the judging process and assess whether an algorithm will select gender as an important predictor from this set.

There are numerous potential algorithms for selecting optimal predictors of profits in this setting.[43] We provide results from three common methods (Athey and Imbens, 2019). We

[41] An alternative method would be to simply compare the survival and profit of female- versus male-owned businesses conditional on financial providers' evaluations. Such a strategy provides evidence on whether gender is a *statistically significant* factor for business performance, but does not tell us if it is a valuable predictor. Appendix Table A16 shows regression results estimating how business performance differs as a function of the true gender of the business owner, conditional on the business' capital allocation decisions by the financial providers. Appendix Table A17 estimates differences in gender on additional measures of business performance prespecified in the study's pre-analysis plan. Though we do not find statistically significant differences in the average likelihood of survival by business-owner gender, we do find that female businesses have lower profits even after controlling for financial providers' decisions. However, the magnitude is small and the predictive value is limited. In such cases, people might find it challenging to accurately predict business performance based solely on the gender of the business owner because the predictive power of the variable is limited. The small effect size might not translate into a meaningful or noticeable difference when making individual predictions, explaining why we do not observe differences in loan officers' beliefs despite these regression results.

[42] Note that statistical significance is different from predictivity (Lo et al., 2015). Because the R^2 is a measure of the distance between the values predicted by OLS and the true values in the sample, it is one possible measure of the predictivity of the model.

[43] We use profits winsorized at 99% as the target variable for all methods.

begin by focusing on the least absolute shrinkage and selection operator (LASSO), which has been widely used for variable selection in linear regression models in the economics literature (Baylis, Heckelei and Storm, 2021). The LASSO introduces a penalty that shrinks coefficients in the regression model toward zero, producing a small set of predictors with non-zero coefficients. The purpose of this penalty is to prevent overfitting, reduce the risk of multicollinearity, and enhance predictive accuracy with a simplified model. The magnitude of this penalty influences the number of variables selected (i.e., models with larger penalties select fewer variables). We use 10-fold cross-validation to optimally select this penalty, a process that relies on randomly splitting the data into 10 subsamples and selecting the penalty parameter that minimizes out-of-sample prediction error.[44] This randomness implies that the variables selected may be slightly different each time the algorithm is implemented. Therefore, to assess the importance of gender as a predictor, we conduct 1,000 trials of the LASSO to see which of the 94 variables in the application the algorithm will select most often to target the most profitable businesses.

A drawback of the LASSO and other penalized linear regression methods is that variable selection can be unstable when many of the predictors are highly correlated, which is true in our setting (Zhao and Yu, 2006). Tree-based methods provide a more robust alternative (Athey and Imbens, 2019). The idea of these approaches is to recursively split the data into subsets based on values of the predictors (e.g., male versus female, or above versus below a certain threshold), then estimate predictions in each subset. Each split aims to minimize the sum of in-sample squared errors across all subsets, and the average squared error reduces with each subsequent split. The subsets are called leaves, and the entire sequence of splits and subsets is called a tree. We provide results from two tree-based algorithms. First, the random forest averages over a large number of trees, where each tree is based on a bootstrap sample and the splits at each stage are based on a random subset of the covariates that change every split (Athey and Imbens, 2019). Second, extreme gradient boosting (XGBoost) similarly averages over a large number of trees, but each sequential tree is designed to reduce the prediction error of the previous tree (Chen and Guestrin, 2016). Tree-based methods do not perform variable selection explicitly. However, it is possible to measure the contribution of each variable to prediction accuracy by calculating the reduction in the variance of the target variable (a proxy for prediction accuracy) when that variable is used in a tree split. This measure, known as feature importance, can be used to rank the variables.[45]

The results from the machine learning exercise, presented in Table 11, suggest that gender

[44]For each fold, the algorithm fits a linear regression on the remaining nine folds, then calculates the prediction error on the tenth fold. This is done for each fold, creating a mean prediction error across the ten folds for each penalty parameter λ_k. The algorithm then selects the λ_k that minimizes this prediction error.

[45]Tree-based methods tend to perform well "out of the box" (Athey and Imbens, 2019). Therefore, we use the default tuning parameters for random forest in Stata and XGBoost in R.

Table 11: Optimal Predictors of Business Profits

Panel A: LASSO		
Variable	Share Included	Rank
Projected Revenue (second item)	0.746	1
Tax Liabilities	0.746	2
Projected Administrative Expenses	0.170	3
No Business License	0.126	4
Profit Missing	0.126	5
Female	0.050	12

Panel B: Tree-based methods		
Variable	Importance	Rank
Random Forest		
Total Assets	1.000	1
Profit	0.926	2
Total Expenditure	0.920	3
Fixed Assets	0.888	4
Second Revenue Item	0.861	5
Female	0.209	52
XGBoost		
Profit	0.176	1
Work Experience	0.060	2
Fixed Assets	0.054	3
Age	0.054	4
Salary Expenditures	0.032	5
Female	0.001	67

is not among the most important predictors of business outcomes. Panel A shows the variables most frequently selected by LASSO. LASSO tends to drop highly correlated predictors; consistent with this, the mean number of variables selected is 2.49, and only two variables are selected in over 50 percent of the 1,000 simulations. Gender is selected in just 5% of simulations. Panel B shows the top 5 variables in terms of feature importance in the random forest and XGBoost, as well as the importance and rank for gender. Gender ranks 52 in importance (out of 94 variables) in the random forest and 67 in importance in XGBoost.[46]

These results are consistent with the conclusion that loan officers' beliefs about gender are accurate: considering gender would not meaningfully improve their ability to target the most

[46]In an exit survey completed by a subset of the loan officers, they highlight current profits and growth potential as key criteria for approving a loan. These factors are also selected in Table 9, and randomly assigned gender does not affect expectations about growth potential (i.e., future business performance).

profitable businesses. As a result, there is no meaningful trade-off between gender equity and targeting of successful businesses.

The machine learning exercise serves as a clear benchmark for the optimal capital allocation decision by loan officers. By design, the objective function of the machine learning algorithm is the same as the belief elicitation of the loan officers: to predict the future performance of the business. Furthermore, the data utilized by the machine learning algorithm constitutes only a subset of the information available to financial providers (i.e., information that could be quantified). Considering that the additional information observed by the loan officers is independent of gender by design, it must be the case that gender is more (or equally) informative in the machine learning context than in the decision made by the loan officers. Lastly, there is a large literature highlighting that humans have limited resources when making decisions (i.e., limited attention).[47] Given these cognitive constraints, loan officers must rely on a limited set of criteria when making capital allocation decisions, while these constraints are less strict for machines. The fact that our machine learning exercise shared the same objective function as the loan officers, that loan officers saw even more information than the machine learning algorithm, and that machine learning is not restricted by human cognitive limitations collectively establish the machine learning exercise as a benchmark with which to compare the loan officers' decisions.

Comparing the machine learning predictions to loan officers' predictions demonstrates why loan officers did not discriminate despite the unconditional gender gap in profits. Just as our machine learning algorithm results selected key variables to identify successful businesses, loan officers also responded to key information about the business. The exercise highlights that the importance of gender in predicting business performance is dwarfed by the other detailed information provided in the application (and in a loan application more generally). If we consider the machine learning prediction model to be optimal, this suggests that loan officers' beliefs were not different enough from reality to affect how capital should be awarded.

Our results suggest that the decisions made by loan officers are consistent with a model of decision-making based on accurate beliefs. Loan officers do not believe that gender predicts business performance, and in accordance with these beliefs, they do not discriminate. Moreover, our machine learning results suggest that the predictive value of gender is not large enough to justify statistical discrimination. This suggests that loan officers' beliefs are accurate enough to respond to gender in a manner consistent with profit maximization and

[47]Limited attention has been the subject of a large theoretical and empirical literature in economics; see for example, Sims (2003) and reviews by Wiederholt et al. (2010), Gabaix (2019) and Maćkowiak, Matějka and Wiederholt (2023). Lieder and Griffiths (2020) discuss how these models fit into the broader psychological literature on human cognitive limitations. Bartoš et al. (2016) discuss endogenous allocation of attention as an aspect of discrimination; we do not find evidence for this type of effect given that we document an absence of discrimination.

that there is no meaningful trade-off between gender equity and targeting capital towards successful businesses.

4 External validity: Assessing generalizability to the broader financial sector

In this next section, we discuss the generalizability of this study using the SANS framework (List, 2020). We posit that our results are informative about gender discrimination in capital markets in Ethiopia and directly generalizable to decisions about initial applications for loans and capital grants. We also discuss the extent to which our results are relevant to questions about gender discrimination in low-income countries more broadly, and in business finance. The SANS framework offers a structured method for defining the contexts to which a study's findings can be generalized; as List (2020) points out, "all results are externally valid to some setting, and no result will be externally valid to all settings." The framework assesses generalizability through four factors: Selection, Attrition, Naturalness, and Scaling. The framework contends that if preferences, constraints, and beliefs in the research study align with those in the context of interest, then the findings of the research study can be generalized to that context. We focus on Selection and Naturalness. We have no Attrition by design: since applications were randomly assigned to loan officers, the randomized gender is necessarily orthogonal to loan officer attrition. Scaling is not applicable in our context.

Selection refers to the selection of the study group and how they compare to the underlying population of interest, and Naturalness refers to the naturalness of the choice task, setting, and timeframe—that is, the extent to which "the choice and outcome architecture [are] exchangeable between research and target settings" (List, 2020). Experimenter demand and social desirability bias are two key aspects that potentially affect naturalness, i.e., the generalizability of the decisions made by subjects. We begin by addressing these two concerns directly. We end by categorizing our study based on the stage of research in the general literature on the topic, which determines the weight of importance that should be put on the external validity assessment.

4.1 Experimenter Demand and Social Desirability

Our results are unlikely to be explained by either experimenter demand or social desirability bias, which are defined as motivations to please the researchers or the host of the competition (EDI), respectively.

Experimenter demand is unlikely because the loan officers did not know they were participating in an experiment.[48] They communicated exclusively with a project manager who was blinded to the key question of interest and to the randomized gender assignment.[49] Consequently, there existed no channel that would induce experimenter demand.

Several factors suggest that social desirability bias is unlikely to significantly reduce the generalizability of our results to the broader financial sector. First, EDI's primary objective is to promote entrepreneurship in all sectors. While their mission includes providing support for traditionally under-resourced populations, such as women, youth, and the rural sector, a large segment of EDI's clients are male and urban, reflecting their primary focus on general entrepreneurship.[50] Both broader messaging on EthioSpur and direct communication with lenders focused on identifying promising businesses and did not include any mention of inclusive growth or women.[51] Moreover, individual loan officers did not appear to be familiar with EDI's programs, and none of the loan officers was aware of EthioSpur before their recruitment. This highlights that the loan officers were unlikely to be familiar with the nuances of EDI's mission statement.

Second, the orientation provided the loan officers with a detailed description of EDI's objectives for each question asked in the evaluation form, none of which mentioned gender.[52] If the loan officers were prioritizing women or men in an effort to please EDI, this would imply that they were pursuing an unstated secondary objective at the cost of performing well on the

[48] The loan officers were simply told that "One of EDC's missions is to improve services and entrepreneurship growth by participating and conducting research. By participating as a judge in the EthioSpur Competition, you are consenting to share your contact information and de-identified responses with research partners who may use it to contact you for additional information and to conduct research." The loan officers were debriefed on the research after the competition was completed.

[49] The project manager was not informed about the gendered randomization until necessary for the debriefing of the loan officers.

[50] Note that since our study, EDI's focus on women has expanded. This is because EDI, a new entity after our project, is the combination of WEDP (a female-focused loan program) and EDC (the partner institution for this project at the time of the study). Other programs by EDC at the time of our study highlight their central focus on general entrepreneurship, with a sensitivity to the needs of underserved populations. For example, the Incubation Project, a contemporaneous initiative, was described on their website as follows: "the primary aim of this program is to develop innovative businesses that are being commercialized and be able to create sustainable profit to the owners, and create more jobs to the community." The only mention of women and youth was a statement saying they were eligible to apply.

[51] This includes promotion materials for the competition, recruitment letters and materials for the judging loan officers, orientation guide for the judging loan officers, contracts with the judging loan officers, and memorandums of understanding with lenders. For example, in the partnership request for loan officers to participate as judges, Ethiospur is described as a business plan competition "to promote the entrepreneurial spirit and provide support to promising entrepreneurs with a strong passion to grow their business" and that loan officers are requested because "the expertise in your organization would be extremely valuable in the evaluation of the business plan competition applications...consistent with our existing partnership to promote entrepreneurs with significant potential."

[52] For example, in the section about their beliefs on future performance, the loan officers were informed that "this information is collected to understand what characteristics determine business success" and for the final score a definition was provided for each sub-category (e.g., "Value proposition: your overall assessment of the profitability potential of the applicant's business plan.").

objectives explicitly requested by EDI.

Third, the loan officers' responses in the evaluation forms, except for loan recommendations, were confidential and de-identified to all EDI staff and researchers, except for the project manager with whom they interacted directly. To make this salient, evaluation forms never captured the loan officers' names and were tracked through an ID number.[53]

Fourth, if social desirability bias had affected loan officers' responses, we might expect them to have favored youth entrepreneurs. Although age was not randomized, we do not find that the loan officers favored younger entrepreneurs by giving them higher evaluations on any of our outcomes of interest.

And finally, each of our outcomes was incentivized differently, and yet our results are consistent across all three measures: monetary prize (beliefs), awarding of capital by the separate entity EDI (final score), and awarding of capital through a loan at one's own institution (loan consideration). If social desirability bias explained our results, this would suggest that an unstated, assumed priority of EDI dominated across multiple domains, even when the decision did not affect EDI. The complete absence of messaging on women in the competition, explicit evaluation criteria, anonymity in evaluations, and the consistency of our results across different objectives collectively suggest that social desirability biases are unlikely to explain our results or inhibit generalizability.

In addition, it is worth noting that EDI's mention of women in its mission statement is not unique in the broader financial sector of Ethiopia. Lending institutions such as microfinance institutions often have a focus on inclusive growth, and at the time of our study, Ethiopia had a national policy focused on improving financial access for women. Thus, to the extent that there is some focus on women, this reflects the goals of the broader financial sector.

4.2 Selection

Applicants: EthioSpur recruited businesses interested in capital, the key group that formal financial institutions support through business loans. The competition's information collection and eligibility requirements mirrored those of an initial loan application, ensuring that the types of businesses likely to seek loans would not be deterred from participating in the competition. Among our sample, 32 percent of businesses report having applied for a loan in the previous 12 months and 88 percent say they are interested in obtaining a loan through the competition, confirming that these businesses are interested in capital, including in the form of a loan. These businesses' interest in loans highlights that they reflect the relevant

[53]The loan officers were informed that their ID would be used to identify them if they were awarded the bonus that we are providing based on the accuracy of evaluations. Even in this case, individual evaluations would not be reviewed and remain anonymous.

population for understanding the beliefs and behaviors of loan officers.[54] Further supporting this, the applicants' business characteristics are similar to those in the only other dataset we could find of potential borrowers in Ethiopia, as discussed in Section 2.3 (Alibhai, 2021).

We intentionally designed the competition's application form to reflect the criteria used by loan officers when making initial capital lending decisions. This ensured that the capital allocation decisions in the business plan competition were based on the same information that loan officers typically have when making initial decisions for loans.[55] Since statistical discrimination models suggest that discrimination emerges because of the informational value of gender, the fact that the information environment was similar for the research and for an initial loan application enhances generalizability.

The minimal eligibility requirements for the competition were unlikely to exclude any business that would be interested in seeking capital in the broader financial sector (e.g., a loan), as they mirror the expectations and requirements in an initial loan application.[56]

In summary, the participation requirements imposed on our sample were designed to be analogous to initial loan requests, and thus the sampling process resembles that of typical loan applicants. By focusing on active entrepreneurs, our sample aligns with the core demographic of lending institutions, ensuring relevance to the broader financial sector.

Loan Officers: The loan officers who served as judges for the competition were real experts who reviewed and made credit decisions. The recruitment was diverse and reflected multiple banking institutions across Addis Ababa, including both microfinance institutions and commercial banks. As described above, there was no information in the competition, recruiting materials for the loan officers, or in agreements with the banks that suggests selection into the loan officer sample would be correlated with gender attitudes. In addition, the eligibility requirements for the loan officers ensured that the sample recruited was relevant to the broader population of interest (i.e., those who make capital decisions in the financial sector).

Thus, the businesses and loan officers both are sampled from the key population of interest.

Experimental Variation of Business Owner Gender: An important consideration revolves

[54] We observe no statistically significant difference by gender in whether the business had applied for a loan in the last 12 months, was interested in a loan through the competition, and the amount of capital they would request in such a loan. Similarly, Kolmogorov-Smirov tests find no statistical difference in the distributions of the amount requested by gender.

[55] We interviewed financial providers from nine different financial institutions on the criteria they used when evaluating businesses and reviewed their standard loan application forms to design the application for the competition.

[56] This includes the four-month operational history requirement. All loan applications we reviewed and loan officers we spoke to highlighted the importance of business history. This is true even for start-up firms. For example, in the sample of potential borrowers from Wasasa MFI, they ask loan applicants the source of their starting capital, and over 80 percent report their own or partner's savings. Less than 1 percent reported a loan from an MFI, and a loan from a formal financial institution was not even an option. This highlights that businesses that are seeking capital generally have already some operational history (Alibhai, 2021).

around whether altering the gender of business owners changed the distribution of gender-incongruent businesses. For example, female-owned businesses in male-dominated sectors may appear more frequently in our sample than in the true population of businesses. The concern is that this could diminish the applicability of our findings to real-world scenarios where there is a stronger correlation between the gender of business owners and the nature of their businesses. Our results suggest this is unlikely. Reduced generalizability would require that the loan officers treat these gender-incongruent businesses differently—i.e., evaluating female or male-owned businesses differently when in a sector dominated by the opposite gender. Our results in Table 5 suggest this is not the case. We observe no heterogeneity in the response to gender in male-dominated industries for any of our outcomes. This suggests that the experimental manipulation did not inhibit generalizability to the true distribution of businesses. Our context further suggests that the manipulation was unlikely to have created large gender incongruities: both genders are represented among all 20 industries listed in the application form, and EDI was not concerned that businesses were gendered to the extent that interpretation of the results would be impacted.

4.3 Naturalness

We next assess the naturalness of our three key outcomes of interest: beliefs, competition score, and loan consideration. Since we have already discussed the concerns of social desirability and experimenter demand above, we focus on the naturalness of the setting and the similarity of the decision environment to that of an initial loan application.

One of our key outcomes is beliefs about gender and future business performance. The elicited beliefs were independent of all other outcomes and incentivized for accuracy. Thus, there was no aspect of the research design that would have altered loan officers' beliefs on what would make a business successful. Another key outcome is the competition score. Similarly, this was the natural decision for the business plan competition—it determined the awarding of a grant and thus is directly generalizable to such decisions. Thus, the research design has no aspects that would change the preferences, constraints, and beliefs about how a business will perform in the future or who is deserving of capital grants.

A key consideration regarding the generalizability of our results to decisions made during an initial loan application is the possibility that loan officers may render different decisions for a loan as opposed to a grant, or when acting on behalf of their own employer. The loan forwarding decision was designed to address this concern. We consider this a proxy for the decision to accept an initial application for a loan request. This proxy is generalizable as long as the following holds: 1) the two measures are correlated, and 2) the relationship does not differ by business owners' gender. On the former, the two outcomes are mechanically correlated—to accept an initial loan application from the applicant, the loan officer must have access to their information, which is only possible by asking for the information to be forwarded. On the latter, our results suggest that there is no gender discrimination in deciding to forward to the lending institution.

It is illustrative to consider an example that would raise concerns about the generalizability of the loan forwarding decision. Suppose loan officers have some threshold of perceived quality that must be met to forward an applicant. We may reasonably assume that the decision to forward the applicant has a lower threshold than the initial approval of a loan application. If female-owned businesses are closer to this threshold, then the conversion from forwarding the applicant to approving the applicant will be lower for women, thus limiting the generalizability to initial loan decisions. But we do not see support for this. Among businesses that were forwarded, neither the loan officers' beliefs on the future business performance nor their scores in the competition differ by business owner gender, suggesting that female-owned businesses are not closer to the threshold. Given that loan officers' perceptions and evaluations of businesses were similar by gender, gender similarly should not affect the conversion of forwarded businesses to initial loan approval. Note that by itself, the lower threshold for forwarding an application does not reduce generalizability.

With respect to the language of our initial framework, because the decision to forward the applicant to their own institution is an intermediary step towards initial loan application approval, the preferences and beliefs for the decision mirror those for the decision on the initial application for a loan. While it may be the case that the constraint differs in that the decision to forward the applicant is less binding than providing a loan (i.e., further down the steps to final approval), our results suggest that this change in constraint is not affecting female-owned businesses differently. This suggests that our results of the causal effect of business-owner gender on the decision to forward the application to one's own institution is a reasonable proxy for approval of an initial loan application to continue to the next step in a loan review process.

4.4 Wave Consideration: The Weight of External Validity and Boundaries of Generalizability

The framework outlined in List (2020) highlights that expectations for external validity should be based on the stage of the body of research on a given topic.[57] We view our results as a Wave 1 insight for gender discrimination in low-income countries, establishing initial causality.

[57] List (2020) categorizes the stage of research into three waves: the first wave is efficacy and proof of concept, the second wage is underlying mechanisms, boundaries, and replications, and the third wave is measurement, mechanisms, and scaling.

Very few studies have been able to use randomized gender to isolate causality convincingly in low- or middle-income contexts. For the broader question of gender discrimination in business finance, we view the study as contributing to Wave 2, understanding underlying mechanisms of discrimination and broadening the exploration of boundary conditions. In the former wave, List (2020) argues that external validity should be viewed as “extra credit”, and in the latter wave, studies should vary subject populations, stakes, and other factors that are theoretically important and to expand the setting to realistic factors that mirror natural settings.

A useful benchmark is to consider the ideal experiment and its feasibility. To make a statement about gender discrimination in credit markets, the ideal experiment would randomize the gender of the business owner in a loan application and then follow the evaluation of that applicant through the entire loan review process. This would be nearly impossible in most low-income countries. In a context like Ethiopia, loan processes are conducted in person. In addition, the file for a potential borrower has a significant amount of information and is not digitized—loan files are generally many pages of information by the time the application is at its final stages. To our knowledge, only one concurrent study has successfully followed subjects through an entire loan process, using an innovative correspondence methodology with gender-balanced prospective borrowers in a consumer credit context. This study was conducted in a high-income country where the loan application and approval were entirely digital (Montoya et al., 2020). Given these limitations, it would be quite difficult to design a study in a low-income country closer to a natural context than our approach in this study.

Finally, theory provides some guidance for what our findings suggest for the possibility of gender discrimination in the later stages of a loan review process. Our context is most similar to the earlier stages of a loan or grant process, but capital requests often involve further steps and interactions, during which we may be concerned that gender discrimination could become a factor. As noted, loan processes generally include multiple visits and communications between a potential business and the lending institution.

However, standard economic models of discrimination, including statistical discrimination, taste-based discrimination, and discrimination based on violation of gender norms, suggest that discrimination should be even less likely at later stages of the process. A key prediction of models of statistical discrimination is that more information reduces reliance on gender as a signal (Aigner and Cain, 1977; Guryan and Charles, 2013). Since every interaction between a financial provider and a loan or grant applicant increases information, this should reduce statistical discrimination. For taste-based discrimination or discrimination due to violation of gender norms, backward induction suggests that a loan officer would not start a process that would be less likely to be successful due to their preferences. Whether this prediction holds true is an open question for future research, both in terms of its existence and also how it can be reconciled with the lack of discrimination observed in the early stages of the process.

5 Conclusion

This paper provides evidence on the role of gender discrimination in capital allocation decisions using a large-scale field experiment in the context of a high-stakes business plan competition in Ethiopia. We obtain clean identification of gender discrimination by randomizing the gender of applicants in the evaluation of the competition. We then evaluate the potential trade-off between gender equity and targeting capital to the highest-performing businesses. Taking our results as a whole, we do not find support for gender discrimination by capital providers as an explanation for gender gaps in access to capital. In a sample of 84 experienced loan officers representing 13 different financial institutions in Ethiopia, the randomly assigned gender of the business owner did not affect real capital allocation decisions, either in a high-stakes business plan competition or for consideration for a loan. These results are remarkably consistent regardless of business-owner, business, and loan officer characteristics.

We find that this lack of discrimination did not have a meaningful cost in terms of targeting capital to the highest-performing businesses. Using data we collected from the competition applicants 18 months after the competition, we use several machine learning algorithms to generate optimal predictions of future business performance, and we find that the gender of the business owner is not a key variable used in these predictions. This suggests that there is no meaningful trade-off between gender equity and targeting of successful businesses. Moreover, it implies that loan officers' beliefs were accurate enough to respond to gender in a manner consistent with profit maximization (i.e., targeting high-performing businesses). This set of results implies that loan officers have neither belief-based partiality (i.e., statistical discrimination) nor preference-based partiality (i.e., taste-based discrimination).

Our findings provide support to the theoretical argument that discrimination will not persist in contexts where it is not profit-maximizing. This is generally consistent with the experimental literature to date. For example, Brock and De Haas (2023) did not find discrimination on the margin for which loan officers were incentivized (*loan approval*), though they do find discrimination on a margin that was not incentivized (*conditions of the loan offer*).[58] This latter discrimination is mitigated by loan officer age and experience, two key characteristics associated with increased learning. Similarly, in the context of consumer credit in Chile, Montoya et al. (2020) find gender discrimination against women is mitigated in areas with greater market competition.

[58]The patterns provided in Brock and De Haas (2023) provide convincing evidence of gender discrimination on this margin of contractual details, but it remains an open question of whether the discrimination could be mitigated if loan officers had to pay a cost for making them.

Our results also suggest caution against assuming that patterns of gender discrimination align with patterns of gender disparities. In Ethiopia, like many low-income countries, gender gaps in access to finance persist and contribute to high levels of gender inequality (Klapper and Parker, 2011). In such contexts, a common assumption is that unequal gender norms will lead to gender discrimination and cause capital misallocation across equally productive men and women; however, we show that this is not necessarily the case. Indeed, empirical evidence has highlighted that in developing countries with high gender inequality, gender discrimination can favor women conditional on indicators of high ability (Delavande and Zafar, 2019; Ayalew, Manian and Sheth, 2021).

Our results suggest that gender discrimination in access to capital is not a key contributor to gender gaps in business performance and growth, highlighting the importance of future research on which factors do explain these gaps. One key consideration is gender differences in the demand for capital. It may be that the decision to apply for capital is an important driver of the gender gap, and also contributes to the lack of gender discrimination conditional on applying for capital. In our sample of applicants, the gender gap in business performance appears to be smaller than what we may expect in the broader population. This suggests that applying for capital may already be a signal for business success. Furthermore, to the extent that the decision to apply for a loan is a household decision, the household's choice to apply for capital for a female-owned business may be informative about productive assets in the household and household support for the business (Bernhardt et al., 2019). That is, applying for capital may signal both success and household support and responsibility for the loan. In addition, whether discrimination would occur for the marginal borrower, and whether perceived discrimination has a selection effect into the demand for capital, are also potential directions for future research.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2023. "When should you adjust standard errors for clustering?" *The Quarterly Journal of Economics*, 138(1): 1–35.

Aigner, Dennis J., and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review*, 30(2): 175.

Alibhai, Salman, Aletheia Donald, Markus Goldstein, Alper Ahmet Oguz, Alexander Pankov, and Francesco Strobbe. 2019. "Gender Bias in SME Lending Experimental Evidence from Turkey." *Policy Research Working Paper*, , (9100).

Alibhai, Salman, Rachel Cassidy Markus Goldstein Sreelakshmi Papineni. 2021. "Ethiopia-Psychometric Credit Scoring as an Alternative to Collateral for Women Entreprenuers in Ethiopia - Baseline and Endline Surveys 2021., Randomized Controlled Trial Impact Evaluation (WEDP ETH 2021)." World Bank Group. Ref: ETH_2018-2021_WEDP-IE_v01_M. DOI: 10.48529te4j-av44 (accessed June 14, 2024).

Arrow, Kenneth J. 1974. "THE THEORY OF DISCRIMINATION." *Discrimination in Labor Markets*, 1–33. Princeton:Princeton University Press.

Aterido, Reyes, Thorsten Beck, and Leonardo Iacovone. 2013. "Access to Finance in Sub-Saharan Africa: Is There a Gender Gap?" *World Development*, 47: 102–120.

Athey, Susan, and Guido W Imbens. 2019. "Machine learning methods that economists should know about." *Annual Review of Economics*, 11: 685–725.

Ayalew, Shibiru, Shanthi Manian, and Ketki Sheth. 2021. "Discrimination from below: Experimental evidence from Ethiopia." *Journal of Development Economics*, 151: 102653.

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka. 2016. "Attention discrimination: Theory and field experiments with monitoring information acquisition." *American Economic Review*, 106(6): 1437–1475.

Baylis, Kathy, Thomas Heckelei, and Hugo Storm. 2021. "Machine learning in agricultural economics." In *Handbook of Agricultural Economics*. Vol. 5, 4551–4612. Elsevier.

Becker, Gary S. 1957. *The economics of discrimination.* University of Chicago press.

Beck, Thorsten, and Robert Cull. 2014. "SME finance in Africa." *Journal of African Economies*, 23(5): 583–613.

Beck, Thorsten, Patrick Behr, and Andreas Madestam. 2017. "Sex and Credit: Is There a Gender Bias in Lending?" *Journal of Banking and Finance*, 87.

Bernhardt, Arielle, Erica Field, Rohini Pande, and Natalia Rigol. 2019. "Household matters: Revisiting the returns to capital among female microentrepreneurs." *American Economic Review: Insights*, 1(2): 141–160.

Bharadwaj, Prashant, Rahul Deb, and Ludovic Renou. 2024. "Statistical Discrimination and the Distribution of Wages." *Working Paper.*

Blattman, Chris, Nathan Fiala, and Sebastian Martinez. 2014. "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda." *The Quarterly Journal of Economics*, 129(2): 697–752.

Bohren, J Aislinn, Alex Imas, and Michael Rosenberg. 2019. "The dynamics of discrimination: Theory and evidence." *American economic review*, 109(10): 3395–3436.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope. 2023. "Inaccurate Statistical Discrimination: An Identification Problem." *The Review of Economics and Statistics.*

Brock, J. Michelle, and Ralph De Haas. 2023. "Discriminatory Lending: Evidence from Bankers in the Lab." *American Economic Journal: Applied Economics*, 15(2): 31–68.

Brown, Morton B., and Alan B. Forsythe. 1974. "Robust tests for the equality of variances." *Journal of American Statistical Association*, 69.

Charness, Gary, Uri Gneezy, and Brianna Halladay. 2016. "Experimental methods: Pay one or pay all." *Journal of Economic Behavior & Organization*, 131: 141–150.

Charness, Gary, Uri Gneezy, and Vlastimil Rasocha. 2021. "Experimental methods: Eliciting Beliefs." *Journal of Economic Behavior and Organization*, 189.

Chen, Tianqi, and Carlos Guestrin. 2016. "Xgboost: A scalable tree boosting system." 785–794.

Danz, David, Lise Vesterlund, and Alistair J. Wilson. 2022. "Belief Elicitation and Behavioral Incentive Compatibility." *American Economic Review*, 112.

Delavande, Adeline, and Basit Zafar. 2019. "Gender Discrimination and Social Identity: Evidence from Urban Pakistan." *Economic Development and Cultural Change*, 68.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. 2017. "Cultural Proximity and Loan Outcomes." *The American Economic Review*, 107(2): 457–92.

Gabaix, Xavier. 2019. "Behavioral inattention." In *Handbook of behavioral economics: Applications and foundations 1*. Vol. 2, 261–343. Elsevier.

Guryan, Jonathan, and Kerwin Kofi Charles. 2013. "Taste-Based or Statistical Discrimination: The Economics of Discrimination Returns to Its Roots." *Economic Journal*, 123(572): 417–32.

Hardy, Morgan, and Gisella Kagy. 2020. "It's Getting Crowded in Here: Experimental Evidence of Demand Constraints in the Gender Profit Gap." *The Economic Journal*, 130(631): 2272–2290.

Klapper, Leora F, and Simon C Parker. 2011. "Gender and the business environment for new firm creation." *The World Bank Research Observer*, 26(2): 237–257.

Levene, H. 1960. "Robust tests for equality of variances." In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling.* , ed. W. Hoeffding W.G. Madow I. Olkin, S.G. Ghurye and H.B. Mann, 278–292. Stanford University Press.

Lieder, Falk, and Thomas L Griffiths. 2020. "Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources." *Behavioral and brain sciences*, 43: e1.

List, John. 2020. "Non est Disputandum de Generalizability? A Glimpse into the External Validity Trial." *NBER Working Paper*, , (27535).

Lo, Adeline, Herman Chernoff, Tian Zheng, and Shaw-Hwa Lo. 2015. "Why significant variables aren't automatically good predictors." *Proceedings of the National Academy of Sciences*, 112(45): 13892–13897.

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt. 2023. "Rational inattention: A review." *Journal of Economic Literature*, 61(1): 226–273.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica.*

Montoya, Ana María, Eric Parrado, Alex Solís, and Raimundo Undurraga. 2020. "Bad taste: gender discrimination in the consumer credit market." IDB Working Paper Series.

Muravyev, Alexander, Oleksandr Talavera, and Dorothea Schäfer. 2007. "Entrepreneurs' gender and financial constraints: Evidence from international data." *Journal of Comparative Economics*, 37(2): 270–286.

Phelps, Edmund S. 1972. "The statistical theory of racism and sexism." *American Economic Review*, 62(4): 659–661.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of monetary Economics*, 50(3): 665–690.

Trautmann, Stefan T., and Gijs van de Kuilen. 2014. "Belief Elicitation: A Horse Race among Truth Serums." *Economic Journal*, 125(589).

Wiederholt, Mirko, et al. 2010. "Rational inattention." *The New Palgrave Dictionary of Economics (Online Edition ed.)*, 1–8.

World Bank. 2019. "Ethiopia Gender Diagnostic Report: Priorities for Promoting Equity."

Zhao, Peng, and Bin Yu. 2006. "On model selection consistency of Lasso." *The Journal of Machine Learning Research*, 7: 2541–2563.

A Appendix

Application ID: [redacted]

Business Owner's Characteristics

28 years old
Female
9 years of total work experience
Not currently employed outside of proposed business
Vocational training is the highest level of education completed
Never Married
4 total household members, including applicant
0 children

Business Characteristics

Owns 100% of the business
Description of product and/or services: [redacted]
5 years of experience in the industry
Business cards have been used for advertising
Has a current business license
13 years of operation
Rents primary business location
Yes, had a written financial record keeping system for the previous 4 months

Current Business Performance

25,950 birr PROFIT reported in previous month
- 84,000 birr total revenue reported in previous month
- 58,050 birr total expenses reported in previous month
  - 45,000 birr supply purchase expenses
  - 150 birr rental expenses
  - 7,000 birr salary expenses
  - 5,900 birr other expenses

9 paid full-time employee(s)
- 48 hours per week typically worked by employees

310,000 birr total assets
- 55,000 birr cash assets
- 210,000 birr fixed assets
- 45,000 birr other assets

0 birr total liabilities
- 0 birr loans payable within one year
- 0 birr loans with longer than one year duration
- 0 birr to trade creditors
- 0 birr in other liabilities

Top Three Customers

1. Name and phone number provided. 320,000 birr in revenue during the past year. Customer is an Organization.
2. Name and phone number provided. 21,000 birr in revenue during the past year. Customer is an Organization.
3. Name and phone number provided. 300,000 birr in revenue during the past year. Customer is an Organization.

Business Plan

[redacted]

Figure A1: Application form shown to loan officers

Confidential Evaluation of Applicant
Date of evaluation (DD/MM): ______________________ Judge ID: ____________________________

Section A: Application Verification (*For verification purposes only*)

Application ID:
Applicant's age: ☐ 18-25 ☐ 26-35 ☐ 36-45 ☐ 46-55 ☐ above 55 ☐ Information is missing
Applicant's gender: ☐ Male ☐ Female ☐ Information is missing
Applicant's total years of experience: ☐ 0-4 ☐ 5-9 ☐ 10-19 ☐ 20 or more ☐ Information is missing
Applicant employed outside of the proposed business: ☐ Yes ☐ No ☐ Information is missing

<u>Section B: Understanding Business Growth (*For determining judge bonus only*)</u>

*Suppose that the applicant receives **no capital** from the competition:*

What is the probability that this business will be operational in January 2021: ☐ 0-10% ☐ 11–20% ☐ 21–30% ☐ 31–40% ☐ 41-50% ☐ 51-60% ☐ 61-70% ☐ 71–80% ☐ 81-90% ☐ 91-100%
Assuming that the business is operational in January 2021, provide your best estimate of:
The number of operational hours in January 2021 will be: ☐ Less than in January 2020 ☐ Similar to January 2020 ☐ Greater than January 2020
The value of the business' capital stock in January 2021: Birr
The monthly profits or losses of the business in January 2021 (*Only one should be filled*). *Monthly Profit:* ______________________________ Birr *Monthly Loss:* ____________________________ Birr
The number of paid employees (excluding the owner) in January 2021:

*Suppose the applicant receives **300,000 ETB** from the competition:*

What is the probability that this business will be operational in January 2021: ☐ 0-10% ☐ 11–20% ☐ 21–30% ☐ 31–40% ☐ 41-50% ☐ 51-60% ☐ 61-70% ☐ 71–80% ☐ 81-90% ☐ 91-100%
Assuming that the business is operational in January 2021, provide your best estimate of:
The number of operational hours in January 2021 will be: ☐ Less than in January 2020 ☐ Similar to January 2020 ☐ Greater than January 2020
The value of the business' capital stock in January 2021: Birr
The monthly profits or losses of the business in January 2021 (*Only one should be filled*). *Monthly Profit:* ______________________________ Birr *Monthly Loss:* ____________________________ Birr
The number of paid employees (excluding the owner) in January 2021:

*If the applicant was instead given a **3-year 100,000 ETB loan**, which of the following do you believe is most likely?*

☐ Applicant will repay the loan: Applicant will have enough financial resources and will repay. ☐ Applicant will strategically default: Applicant will have enough financial resources, but will still not repay. ☐ Applicant must default: Applicant will not have enough financial resources to repay the loan.

Section C: Reviewing the Applicant

Rate applicant's managerial skills: ☐ very poor ☐ poor ☐ acceptable ☐ good ☐ excellent
Which do you expect that the applicant can access to cover shortfalls in demand? *Check all that apply.* ☐ Personal savings/assets Gifts/Loans from family or friends ☐ Business loans from microfinance B Busness loans from bank Government assistance
Estimate the total amount of additional capital the applicant can secure (from all sources): Birr
Applicant's business is most likely the primary source of income for the applicant's household? ☐ Yes ☐ No
Rate market demand of applicant's business: ☐ very low ☐ low ☐ medium ☐ high ☐ very high

<u>Section D: Determination of winner</u> Overall impression will be half the final score, and value proposition and entrepreneurial credibility will be the other half of the final score. ***This final score is the only measure that determines the competition winners.***
Final Score = Overall Impression + ½ *Value Proposition + ½ *Entrepreneurial Credibility.

OVERALL IMPRESSION:	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10
VALUE PROPOSITION:	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10
ENTREPRENEURIAL CREDIBILITY:	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7	☐ 8	☐ 9	☐ 10

***<u>Internal</u>*:** Should applicant's information be sent to your institution for loan consideration? ☐ Yes ☐ No

Figure A2: Evaluation Form

Table A1: Pre-specified secondary outcomes

	(1) Surv., w/o Cap	(2) Surv., w/ Cap	(3) Win. Assets, w/o Cap	(4) Win. Assets, w/ Cap	(5) Win. Jobs, w/o Cap	(6) Win. Jobs, w/ Cap	(7) Loan
Male	-0.0944	-0.0339	60.09	52.75	87.85**	162.0	0.00159
	(0.636)	(0.666)	(46.85)	(65.42)	(43.00)	(205.3)	(0.0140)
Observations	3696	3696	3696	3696	3696	3696	3696
Female Mean	50.47	60.08	778.4	1089.4	219.2	878.0	0.495

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. Survival, Assets, and Employees are expectations of the loan officers with and without additional capital. Assets are in thousands of birr. Survival is the probability of survival, from 0 to 100. Employees is the number of people employed by the business. Win. specifications winsorize the variables at the 1 percent. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A2: Primary outcomes with loan officer weights

	(1) Score	(2) Profit, w/o Cap	(3) Win. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Win. Profit, w/ Cap
Male	-0.0727	-30.13	0.701	-1141.5	-14.20*
	(0.116)	(77.97)	(4.197)	(990.2)	(8.323)
Observations	3696	3696	3696	3696	3696
Female Mean	11.96	52.98	42.74	713.8	86.19

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Specifications include loan officer and application fixed effects, and weight each loan officer equally. Robust standard errors in parentheses.

Table A3: Secondary outcomes with loan officer weights

	(1) Surv., w/o Cap	(2) Surv., w/ Cap	(3) Win. Assets, w/o Cap	(4) Win. Assets, w/ Cap	(5) Win. Jobs, w/o Cap	(6) Win. Jobs, w/ Cap	(7) Loan
Male	0.276	0.531	57.58	46.95	81.51**	93.86	0.00127
	(0.665)	(0.703)	(45.28)	(66.18)	(38.63)	(191.3)	(0.0144)
Observations	3696	3696	3696	3696	3696	3696	3696
Female Mean	49.74	58.10	771.5	1105.8	155.3	692.7	0.479

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. Survival, Assets, and Employees are expectations of the loan officers with and without additional capital. Assets are in thousands of birr. Survival is the probability of survival, from 0 to 100. Employees is the number of people employed by the business. Win. specifications winsorize the variables at the 1 percent. Specifications include loan officer and application fixed effects, and weight each loan officer equally. Robust standard errors in parentheses.

Table A4: Primary outcomes with ordering FE

	(1) Score	(2) Profit, w/o Cap	(3) Win. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Win. Profit, w/ Cap
Male	-0.114	-25.15	1.721	-948.8	-8.159
	(0.116)	(73.46)	(4.220)	(907.6)	(7.910)
Observations	3685	3685	3685	3685	3685
Female Mean	12.06	43.26	42.41	709.2	84.57

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Specifications include the order of the application presented to the loan officer, loan officer, and application fixed effects. Robust standard errors in parentheses.

Table A5: Secondary outcomes with ordering FE

	(1) Surv., w/o Cap	(2) Surv., w/ Cap	(3) Win. Assets, w/o Cap	(4) Win. Assets, w/ Cap	(5) Win. Jobs, w/o Cap	(6) Win. Jobs, w/ Cap	(7) Loan
Male	-0.108	-0.0501	57.63	50.45	90.41**	160.8	0.00108
	(0.639)	(0.670)	(47.14)	(65.39)	(42.93)	(206.1)	(0.0141)
Observations	3685	3685	3685	3685	3685	3685	3685
Female Mean	50.47	60.08	778.4	1089.4	219.2	878.0	0.495

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. Survival, Assets, and Employees are expectations of the loan officers with and without additional capital. Assets are in thousands of birr. Survival is the probability of survival, from 0 to 100. Employees is the number of people employed by the business. Win. specifications winsorize the variables at the 1 percent. Specifications include the order of the application presented to the loan officer, loan officer, and application fixed effects. Robust standard errors in parentheses.

Table A6: Reported gender

	(1) Score	(2) Profit, w/o Cap	(3) Wind. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Wind. Profit, w/ Cap
Reported Male	-0.114	-26.87	2.276	-990.6	-7.967
	(0.117)	(76.04)	(4.273)	(937.4)	(7.995)
Reported No Gender	-0.0292	-57.08	-2.715	-754.8	-32.09
	(0.451)	(93.75)	(16.26)	(701.2)	(31.22)
Observations	3696	3696	3696	3696	3696
Reported Female Mean	12.06	43.18	41.99	717.2	84.32

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Independent variables are those reported by the loan officer. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A7: 75 percent correct on verification questions

	(1) Score	(2) Profit, w/o Cap	(3) Wind. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Wind. Profit, w/ Cap
Male	-0.105	-24.26	1.665	-962.1	-8.534
	(0.116)	(74.49)	(4.208)	(911.4)	(7.895)
Observations	3696	3696	3696	3696	3696
Female Mean	12.06	43.26	42.41	709.2	84.57

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Observations limited to loan officers that corrected answered verification questions on at least 75 percent of their evaluations. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A8: 100 percent correct on verification questions

	(1) Score	(2) Profit, w/o Cap	(3) Wind. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Wind. Profit, w/ Cap
Male	1.625	19.06	19.06	36.21	36.21
	(2.518)	(22.94)	(22.94)	(60.90)	(60.90)
Observations	329	329	329	329	329
Female Mean	12.00	54.82	41.40	105.4	85.13

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Observations limited to loan officers that corrected answered verification questions on at 100 percent of their evaluations. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A9: Loan officers responsive to application information

	(1) Score	(2) Profit, w/o Cap	(3) Wind. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Wind. Profit, w/ Cap
Male	-0.111	-14.73	2.499	-842.5	-7.358
	(0.118)	(69.24)	(4.257)	(800.8)	(8.057)
Observations	3542	3542	3542	3542	3542
Female Mean	12.00	42.41	42.42	736.6	85.32

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A10: Loan officers responsive to capital

	(1) Score	(2) Profit, w/o Cap	(3) Wind. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Wind. Profit, w/ Cap
Male	-0.0409	-58.81	1.958	-1324.8	-3.953
	(0.132)	(107.1)	(4.197)	(1394.7)	(9.984)
Observations	2973	2973	2973	2973	2973
Female Mean	12.30	34.36	37.14	898.5	102.5

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Observations limited to evaluations in which predictions with capital were greater than predictions without capital. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A11: Excludes 5 percent of loan officers with lowest variance of final scores

	(1) Score	(2) Profit, w/o Cap	(3) Wind. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Wind. Profit, w/ Cap
Male	-0.105	-16.10	1.735	-864.0	-9.143
	(0.117)	(70.96)	(4.376)	(816.5)	(8.211)
Observations	3663	3544	3544	3592	3592
Female Mean	12.05	44.77	43.88	731.7	86.86

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Observations limited to loan officers with variation in outcomes at least at the 5th percentile. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A12: First five evaluations

	(1) Score	(2) Profit, w/o Cap	(3) Wind. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Wind. Profit, w/ Cap
Male	2.195	-29.66	-29.66	45.13	44.97
	(4.520)	(63.82)	(63.82)	(66.87)	(66.69)
Observations	410	410	410	410	410
Female Mean	11.76	62.90	45.10	8.228	93.27

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Observations are limited to the first round of applications given to the loan officer. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A13: No loan officers fixed effects

	(1) Score	(2) Profit, w/o Cap	(3) Wind. Profit, w/o Cap	(4) Profit, w/ Cap	(5) Wind. Profit, w/ Cap
Male	-0.0328	-14.25	4.011	-750.1	-1.344
	(0.138)	(64.67)	(4.331)	(781.4)	(8.266)
Observations	3693	3693	3693	3693	3693
Female Mean	12.07	43.19	42.33	710.2	84.55

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, ranging from 1 to 20. Profits are expectations of the loan officers with and without additional capital, and measured in thousands of birr. Win. specifications winsorize the variables at the 1 percent. Specifications include application fixed effects. Robust standard errors in parentheses.

Table A14: Effect of gender on return to capital

	(1) Survival	(2) Profit	(3) Win. Profit	(4) Profit (IHS)	(5) Assets	(6) Win. Assets	(7) Assets (IHS)
Male	0.0605	-937.9	-9.446*	-0.0932	-582.6	-1.745	-0.0584
	(0.513)	(851.3)	(5.687)	(0.0906)	(710.4)	(11.17)	(0.0982)
Observations	3696	3696	3696	3696	3696	3696	3696
Female Mean	9.610	665.9	41.70	1.856	926.3	241.0	4.776

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Survival, Profits, and Assets are the difference in expectations of the loan officers with and without additional capital (i.e., the return to capital). Profit and Assets are in thousands of birr. Survival is the probability of survival, from 0 to 100; Win. specifications winsorize the variables at the 1 percent; and IHS specifications transform the variables using the inverse hyperbolic sine transformation. Specifications include loan officer and application fixed effects. Robust standard errors in parentheses.

Table A15: Main specification robustness to endline sample

	(1)	(2)	(3)	(4)	(5)
	Score	Overall Impress	Value Prop	Entrepreneurial	Loan
Male	-0.0792	-0.0270	-0.0561	-0.0485	-0.00311
	(0.119)	(0.0632)	(0.0644)	(0.0671)	(0.0145)
Observations	3430	3430	3430	3430	3430
Female Mean	12.09	5.999	6.092	6.085	0.496

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Score is the final score in the business plan competition, determined by Overall Impression (Overall Impress) + .5* Value Proposition (Value Prop) + .5*Entreprenuerial Credibility (Entreprenuerial). Each of these subscores is on an increasing scale of 1 to 10. Loan indicates whether the application was forwarded by the loan officer to their own institution for loan consideration. Specifications include loan officer and application fixed effects. Sample is limited to those observed in endline survey. Robust standard errors in parentheses.

Table A16: Predicted value of gender for business performance

	(1) Firm Survival	(2) Firm Survival	(3) Profit	(4) Profit	(5) Win. Profit	(6) Win. Profit	(7) Firm Profits (IHS)	(8) Firm Profits (IHS)
Male	0.0269	0.0284	13317.3**	13645.5**	4539.4*	4922.0**	0.437	0.484
	(0.0245)	(0.0245)	(6094.1)	(6018.3)	(2345.1)	(2352.2)	(0.451)	(0.452)
Mean Final Score	0.00953**		2132.2		2366.1***		0.292***	
	(0.00421)		(1422.6)		(496.8)		(0.0863)	
Mean Loan Consideration		0.0825*		13033.9		22915.9***		2.994***
		(0.0439)		(21947.4)		(5487.5)		(0.944)
Constant	0.840***	0.839***	11827.8***	11638.3***	14144.8***	13991.0***	6.125***	6.107***
	(0.0186)	(0.0186)	(3053.6)	(3053.7)	(1524.4)	(1530.4)	(0.325)	(0.328)
r2	0.00702	0.00527	0.00950	0.00687	0.0379	0.0332	0.0157	0.0152
N	847	847	846	846	846	846	846	846

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mean final score (loan consideration) is the relative mean for the applicant, in which the score or loan consideration has been demeaned by the loan officer average score or consideration of loan. Robust standard errors in parentheses.

Table A17: Pre-specified endline variables

	(1) Ttl Bus Prof	(2) Win. Ttl Bus Prof	(3) Ttl Bus Prof (IHS)	(4) Hh Bus Prof	(5) Win. Hh Bus Prof	(6) Hh Bus Prof (IHS)	(7) Pers Income	(8) Win. Pers Income	(9) Pers Income (IHS
Male	15684.4**	7802.6***	0.431	11343.0*	3783.3	-0.0877	15684.4**	7802.7***	0.442
	(6306.1)	(2699.3)	(0.437)	(6470.7)	(3130.4)	(0.416)	(6306.1)	(2699.3)	(0.432)
Mean Final Score	2549.4*	2730.4***	0.248***	2829.2*	3030.9***	0.184**	2549.4*	2730.4***	0.243***
	(1435.9)	(575.6)	(0.0841)	(1453.9)	(642.5)	(0.0800)	(1435.9)	(575.6)	(0.0831)
Constant	14098.3***	15829.0***	6.767***	20606.9***	22448.2***	7.701***	14098.7***	15829.3***	6.839***
	(3293.5)	(1639.1)	(0.312)	(3524.6)	(2067.4)	(0.291)	(3293.5)	(1639.1)	(0.309)
r2	0.0125	0.0423	0.0124	0.0100	0.0329	0.00671	0.0125	0.0423	0.0123
N	846	846	846	844	844	844	846	846	846

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mean final score is the relative mean for the applicant, in which the score has been demeaned by the loan officer average score. Columns 1 to 3 reflect total business profits earned by the applicant from all their businesses, Columns 4 to 6 reflect all business profits earned by all members in the applicant's household, and Column 7 to 9 reflect all income earned by the applicant. Robust standard errors in parentheses.