Screening in Loan Guarantee Programs: Combining Contract Menus with Information Collection*

Yusik Andrew Kim †

November 26, 2025

Click here for the latest version.

Abstract

To support credit-constrained small businesses, governments use loan guarantee programs that insure lenders against default risk. However, these programs face challenges in allocating appropriate loan sizes due to limited information about borrowers. I study a screening mechanism that combines two common tools in lending—a menu of contracts and soft information collection (e.g., interviews and site visits)—to mitigate borrower information asymmetry. Firms sort themselves by risk: low-risk borrowers accept more intensive soft information collection to obtain larger loans, while higher-risk borrowers choose smaller, fully guaranteed loans to avoid such scrutiny. In effect, the level of information collection becomes part of the contract menu, encouraging risk-based self-selection and improving the agency's ability to allocate loan sizes. Using detailed data from South Korea's loan guarantee program, I evaluate the welfare implications of this joint mechanism. I find that offering a loan guarantee menu alone increases program welfare by 3.9%, while adding soft information raises it by 8.7%. These findings highlight the complementarity between contract menus and soft information in enabling more efficient loan allocation across credit markets.

Keywords: loan guarantees, screening, menu, soft information, small business lending

JEL Codes: D82, G21, G28, H81, L38

*I am grateful to Dan Ackerberg, Victoria Marone, Andrey Ordin, and Jackson Dorsey for their guidance, and to participants at the UT Austin IO workshop for valuable comments. I thank the Korea Federation of Credit Guarantee Foundations for providing data access. The views expressed are my own and do not necessarily reflect those of the Federal Reserve Board or the Federal Reserve System.

†Federal Reserve Board of Governors. andrew.kim@frb.gov

1 Introduction

Small businesses are the backbone of the economy, accounting for more than 50% of global employment (World Bank [2024]). However, they face significant challenges in obtaining loans due to information asymmetry between themselves and lenders. Lenders lack access to formal company credit scores or concrete accounting information (Greenbaum and Thakor [1995]), hindering their ability to accurately assess creditworthiness. This information gap, combined with higher average default rates and the relatively small size of requested loans, tends to deter lenders from engaging with small businesses. As a result, small businesses are often unable to get financing from private lending markets. To address this market failure, all developed countries have some form of government loan guarantee program (Beck et al. [2010], OECD [2017], World Bank [2020]). Most commonly, these programs insure private lenders against default risk by committing to cover a pre-specified portion of the outstanding balance in the event of default.[1]

In programs of this type, a key policy variable is the *guarantee rate*, meaning the portion of the loan that is insured. The guarantee rate significantly affects the size and number of loans that lenders are willing to make (Bachas et al. [2021]). Some businesses appear so risky that nearly a 100% guarantee rate is necessary to attract any private lending, and high guarantee rates (80 percent or above) are common. However, high guarantee rates can incentivize lenders to extend more credit than is appropriate. Governments therefore typically also specify a *maximum loan size* to which the guarantee rate may be applied. Because the appropriate level of credit may vary across borrowers of different types, it is common to allow the maximum loan size to vary with borrower characteristics (Bryan et al. [2024], Deelen and Molenaar [2004]).

The challenge for public guarantee agencies in setting appropriate maximum loan sizes for different types of businesses stems from the same information asymmetry that affects lenders (Saito and Tsuruta [2018]). Beyond the traditional approach of exerting effort to collect "soft information" about borrowers, an increasingly common additional approach used across countries is to offer borrowers a menu of loan guarantee contracts.[2] The basic trade-off presented to borrowers is typically between loan size and other lending conditions (such as interest rates or rejection rates).[3] The idea

[1] Guarantee schemes are more common than direct public lending as they require smaller fiscal outlays.

[2] The agency collects "soft information" due to the lack of standardized data such as official accounting records. This information, gathered through methods like on-site visits and in-depth interviews, assesses business potential and viability. Liberti and Petersen [2019] differentiate between "hard" information, which is quantifiable and objective, and "soft" information, which is qualitative and subject to interpretation variability.

[3] For example, the US SBA program offers 85% guarantees on smaller loans and 75% on larger ones, with higher guarantee rates leading to lower interest rates and higher funding probabilities. Similarly, Korea's KOREG guarantees

is that a borrower's choice of guarantee contract can reveal private information about their type, so that guarantee agencies can make better-informed decisions on loan sizes. Although such menus have become more prevalent over the past decade, their effectiveness remains largely unexplored in the economic literature, particularly regarding how this menu-based approach interacts with the process of soft information gathering.

This paper examines the welfare implications of using a loan guarantee menu as a screening mechanism for allocating loan sizes to small businesses, highlighting how its effectiveness is enhanced when combined with soft information collection. I first develop a two-stage model for screening small business borrowers within loan guarantee programs. In the first stage, a borrower faces a menu of loan guarantee contracts and selects a single contract. In the second stage, the guarantee agency gathers additional "soft information" about the borrower — the extent of which depends on the borrower's contract choice — and then uses this information, along with the contract choice, to set a maximum loan size. I estimate my model using data from the South Korean loan guarantee program, one of the largest such programs globally. This program provides an ideal environment to study these screening mechanisms in practice. A key feature of this program is that it allows a borrower to apply for either the "small" loan program, which has a higher guarantee rate (100%) but typically smaller loan sizes, or the "large" loan program, which has a lower guarantee rate (85%) and thus allows larger loan sizes. After a borrower selects the small or large loan program, a government guarantee agency collects additional "soft information" about the borrower via interviews and site visits, and finally assigns a maximum loan size.

The role of soft information collection in enhancing the effectiveness of loan guarantee menus is intuitive in the market I study. Agencies collect different amounts of soft information depending on whether the borrower opts for the "small" or "large" loan program, using a simplified process for the small loan program and a detailed assessment for the large program. This distinction incentivizes low-risk borrowers to apply for the large loan program, because they can disclose their true, low-risk profiles through detailed evaluations and secure larger loans. Conversely, higher-risk borrowers tend to opt for the small loan program, precisely to avoid intensive scrutiny; the less-intensive soft information collection in that track allows them to conceal their riskier profiles. In effect, the depth of scrutiny becomes part of the contract choice, amplifying the menu's screening power. As a result, the combination of soft information collection and the loan guarantee menu

100% for small loans and 85% for larger loans. Japan's CGC shifted from a universal 100% guarantee to options of 80% or 100% in 2017. Meanwhile, some countries like Finland and Switzerland do not offer such menus and instead use a uniform guarantee rate.

induces borrowers to self-select into contracts that reflect their true risk profiles, thereby enabling agencies to allocate loans more effectively. This intuition parallels patterns observed in the fintech sector, as discussed by Babina et al. [2024], where borrowers' decisions to share or withhold data signal their risk types.[4]

The paper begins by presenting evidence that guarantee agencies use soft information to determine loan sizes, showing that agencies tend to grant larger loans to borrowers with higher ex-post repayment rates, even after controlling for the borrower's choice between the "small" or "large" loan program and other observables. Furthermore, such correlation between ex-post repayment rates and loan sizes is stronger in the large loan program than in the small loan program. The higher correlation reflects more intensive soft information collection in the large loan program.

To provide evidence of borrowers sorting into different guarantee contracts, I examine whether repayment rates differ systematically among borrowers based on their contract choice. I find that borrowers opting for the large loan program have higher repayment rates than those choosing the small loan program, even after controlling for observables including the interest rate and the final loan size. The result suggests that borrowers are self-sorting into different loan guarantees, which means that their contract choice is informative of their risk profiles. Furthermore, the analysis suggests that the sorting effect is driven not only by how much soft information is collected but also by differing guarantee rates (85% or 100%), which influence the risk of borrower rejection by lenders.

I estimate the two-stage screening model, leveraging the observed correlations between loan size, borrower contract choice, and ex-post repayment rate. To determine the appropriate loan size, the government agency balances two objectives: maximizing the value added by the business and minimizing its own financial loss.[5] I model this trade-off using a flexible agency objective function, which is a weighted sum of these two outcomes, with the estimable weights representing the agency's priorities. Estimates reveal that, on average, the agency assigns weights of 39% to maximizing the business's value-added and 61% to minimizing its own losses.

I evaluate the impact of employing a loan guarantee menu by comparing the baseline scenario, which includes both a menu and soft information collection, with a counterfactual scenario that maintains soft information collection but offers only a uniform 100% guarantee program. The

[4] Babina et al. [2024] highlights that low-risk borrowers typically opt to share data to demonstrate their lower risk and improve their loan terms, while high-risk borrowers often choose not to share data to obscure their higher risk levels.

[5] Value-added refers to the net economic contribution of the business, accounting for the loan received. It represents the additional value generated by the business beyond the loan amount.

analysis shows that the value of the agency's objective function—which reflects both the value-added of small businesses and the agency's financial considerations—increases by 8.7% under the loan guarantee menu compared to the uniform program. This increase stems from the agency's ability to more effectively differentiate loan sizes based on borrower risk using the menu. In contrast, under the uniform program, low-risk borrowers are restricted to smaller loans, while high-risk borrowers receive larger loans due to being pooled together, reducing the agency's ability to differentiate effectively between borrower risk types.

To examine whether soft information collection enhances the effectiveness of the loan guarantee menu, I compare outcomes under the loan guarantee menu and a uniform program, both in the absence of soft information collection. The value of the agency's objective function is 3.9% higher with the loan guarantee menu than with the uniform program. The effectiveness of the menu is substantially reduced from 8.7% to 3.9% without soft information collection. This reduction is due to weakened sorting effects within the menu; without soft information collection, borrowers' choices become less informative of their risk types. This demonstrates the critical role that soft information collection plays in enabling borrowers to self-select into appropriate loan guarantee contracts, which in turn significantly improves the agency's objective.

While the empirical findings of this paper are directly applicable to loan guarantee programs, the use of a menu of contracts alongside soft information collection can be applied in broader financial markets, such as consumer finance and commercial banking. In these sectors, financial institutions employ similar screening mechanisms by offering a variety of contract terms and conducting personal interviews or assessments to gather additional information about borrower risk. The empirical setting of this study is well-suited to investigate these two prevalent screening mechanisms. Specifically, I observe detailed data on ex-post loan performance, which allows me to directly correlate loan performance with borrowers' contract choices and also with loan sizes conditional on those choices. By analyzing these relationships, I can isolate the informational value derived from the contract choice and from soft information collection, enabling me to explore the two screening mechanisms both individually and jointly, and highlighting their potential complementarity.

This paper contributes to the literature on screening in lending markets by examining the combined effects of a menu of contracts and soft information collection on borrower self-selection. While both screening mechanisms are well-documented as individual methods to mitigate information asymmetry, their interactive effects remain unexplored. The menu of contracts is examined in previous research including Adams et al. [2009] and Einav et al. [2012], who explore how auto dealerships use down payment options to screen borrowers, and studies such as those by Ioannidou et al. [2022], Taburet et al. [2024], Kawai et al. [2022], and Hertzberg et al. [2018] that investigate various lender tactics like secured versus unsecured loans, loan-to-value ratios, and loan maturity. The role of soft information collection is emphasized in works by Stiglitz and Weiss [1981], Panetta et al. [2009], Agarwal et al. [2011], and Wang [2020], who describe how lenders gather detailed and non-standard information to better assess borrower risks. This study brings the literature together by evaluating how soft information collection enhances the effectiveness of contract menus, providing valuable insights for designing more efficient screening mechanisms in lending markets.

My paper also relates to work examining the efficacy of public loan guarantee programs. Many studies focus on the U.S. Small Business Administration (SBA) loan guarantee program, where lenders independently set loan sizes—a contrast to the setting in this study where the government sets appropriate loan sizes (Brown and Earle [2017], Cox et al. [2021], Bachas et al. [2021], Stillerman [2022], Choi and Lee [2019]). Research on loan guarantee programs in countries such as the UK (Cowling [2010]), Chile (Mullins and Toro [2018]), France (Barrot et al. [2024]), and South Korea (Oh et al. [2009]) typically explores the broad impacts on banks and businesses without delving into the government's role in loan size determination. Although some studies (Panetta [2012], Deelen and Molenaar [2004], Columba et al. [2010], Kuo et al. [2011]) discuss the benefits of government involvement in the loan decision process, surprisingly little attention has been paid to the specific mechanisms for determining appropriate loan sizes, even though over 70% of loan guarantee programs include a government role in the loan size decision (Beck et al. [2010]). Research such as Bryan et al. [2024] underscores the importance of appropriate loan sizes, demonstrating that larger loans boost profits for more suitable businesses while causing declines for less suitable ones, which emphasizes the critical role of tailored loan allocations. My work fills this gap by investigating how guarantee agencies can use a loan guarantee menu to effectively allocate appropriate loan sizes, thereby enhancing overall program outcomes.

The paper proceeds as follows. Section 2 provides the institutional background of the South Korean loan guarantee program. Section 3 describes the dataset used for analysis and presents descriptive statistics. Sections 4 and 5 discuss the empirical model and the estimation process. Section 6 discusses the results obtained from these estimates. Section 7 presents the counterfactual policy simulations, and Section 8 concludes.

2 Institutional Background

The South Korean small business loan guarantee program, managed by the Korea Federation of Credit Guarantee Foundations (KOREG), aims to support small businesses by facilitating access to finance. KOREG oversees 17 regional foundations and 176 local agencies to implement this program. In 2014, KOREG guaranteed loans totaling 8.5 trillion Korean Won (approximately 8.5 billion USD, using a simplified exchange rate of 1,000 KRW to 1 USD), representing about 0.57% of South Korea's GDP.[6] Notably, even creditworthy borrowers often prefer KOREG's guaranteed loans because they offer lower interest rates than conventional bank loans. From 2012 to 2021, approximately 3.1 million small businesses—typically firms with fewer than 10 employees—utilized the loan guarantee program, highlighting its significant role in a sector that numbered around 4.1 million small businesses in South Korea by 2021.[7] While guaranteed loans offer lower interest rates, they often do not fully cover small businesses' financial needs, leading some borrowers to seek additional financing from the private lending market.

A typical loan guarantee process encompasses application, screening, funding, and repayment phases, which are discussed below.

2.1 Borrower Application

Small business owners seeking loan guarantees from KOREG submit their requested loan size. However, the agency determines the guaranteed loan size primarily based on its assessment rather than the borrower's request. On average, guaranteed loans amount to approximately $27,000, typically falling short of requested amounts by a factor of 2.5, with 95% of applicants receiving less than they asked for.

More importantly, borrowers can choose between the "small loan program" with a 100% guarantee rate and the "large loan program" with an 85% guarantee rate.[8] The guarantee rate is a critical factor in this choice, directly influencing the trade-offs borrowers must consider. The full guarantee (100%) in the small loan program ensures that lenders will fund these loans, typically at lower interest rates, but restricts borrowers to smaller loan sizes determined by the agency during

[6] The actual average exchange rate in 2014 was approximately 1,053 KRW per USD. However, all monetary values in this paper are converted using a simplified rate of 1,000 KRW to 1 USD to present data more clearly, as loan sizes typically cluster at intervals such as 10 million KRW (approximately 10k USD), 20 million KRW, etc. This rounded rate is used consistently throughout the analysis for clarity and ease of comparison.

[7] These figures do not reflect the share of all small businesses using the program, as frequent firm entry and exit complicate comparisons. Utilization also rose sharply during the COVID-19 pandemic.

[8] An 85% guarantee on a $100,000 loan and a 100% guarantee on an $85,000 loan both insure $85,000, but only the former leaves the lender exposed on the remaining $15,000.

subsequent evaluations. Conversely, the partial guarantee (85%) in the large loan program allows for larger loan sizes but introduces higher interest rates and a potential for loan rejection due to the 15% of the loan that remains at risk for the lender. Borrowers trade off between larger loan sizes and less favorable loan terms (higher interest rates and lower funding probabilities).

2.2 Agency Screening

Guarantee agencies screen small business borrowers and set appropriate maximum loan sizes based on each borrower's risk profile, business potential, and their choice of the guarantee contract—either the "small loan program" with a 100% guarantee rate or the "large loan program" with an 85% guarantee rate. While agencies have the authority to reject guarantee applications outright, such rejections are uncommon.[9] Instead, the maximum loan sizes are adjusted to match the risk and potential of the businesses: higher-risk or lower-potential borrowers may be offered as little as $5,000, whereas lower-risk, higher-potential applicants could receive up to $100,000. Notably, these maximum amounts almost always become the actual loan sizes disbursed, as most credit-constrained small business borrowers opt to take the maximum available amount. Following the screening process, agencies issue a guarantee contract to borrowers, specifying the maximum loan size and guarantee rate.[10]

Importantly, the borrower's choice of guarantee contract—either the small or large loan program—affects loan size decisions in two ways. First, all else being equal, the lower 85% guarantee rate in the large loan program allows agencies to offer larger loans compared to the 100% guarantee in the small loan program. Second, this choice serves as an informative signal about borrower risk types to the agencies, who then use this information to determine the most appropriate loan size.

The screening process then incorporates additional information, both "hard" and "soft". "Hard" information such as the owner's credit score, business age, and number of employees often falls short of fully capturing a small business's potential. Consequently, agencies also rely on collecting "soft" information to fill this gap. This information, including assessments of the business's potential and viability, is gathered through methods like on-site business visits and in-depth interviews, making the screening process thorough but time-consuming.

To manage resources effectively, agencies vary their screening efforts based on the program selected by borrowers. The "large loan program", with larger loan sizes averaging around $37,000,

[9]Borrowers classified as credit delinquents or those who have previously defaulted on government-guaranteed loans are typically rejected. Discussions with guarantee officers indicate that rejections occur in less than 10% of cases.

[10]The contract also details the maturity, repayment method, and guarantee fee.

prompts agencies to conduct detailed evaluations to mitigate financial risks adequately. Conversely, the "small loan program", involving smaller loan sizes averaging around \$19,000, utilizes a simplified evaluation process, due to the lower financial stakes.[11]

Agencies' loan decision-making processes balance two main goals: supporting small businesses and maintaining the guarantee program's financial sustainability. While they collect small fees from banks for the guarantees—costs typically passed on to borrowers—they also incur an average loss of over \$1,000 per loan, reflecting their commitment to economic growth through small business support. However, given their operational budget constraints, agencies must manage their resources carefully to maintain the program's financial sustainability.

2.3 Loan Funding and Repayment

After acquiring a guarantee contract, which includes the guarantee rate and the maximum loan size, a borrower visits a private lender to obtain a loan. At this stage, the lender conducts their own risk assessment of the guarantee contract to decide whether to approve or deny the loan, and to set the interest rate. A borrower holding an 85% guarantee rate faces the risk of lender rejection due to the unguaranteed 15% portion of the loan. In case of rejection, the borrower can reapply for the small loan program with a 100% guarantee rate, incurring costs such as time delays and adverse effects on her credit score resulting from the rejection. The risk of rejection influences the borrower's decision-making process, as she must weigh the higher funding probability and lower interest rates associated with a 100% guarantee rate against the larger loan size but increased rejection risk with an 85% guarantee rate.

In the event of borrower default, the government reimburses the lender for the guaranteed portion of the remaining loan balance. Defaults are reported to credit bureaus, and there may be legal action to recover the loan from the borrower. From the borrower's perspective, defaulting on a guaranteed loan has the same consequences as defaulting on any other loan—it damages the borrower's credit history and triggers debt collection proceedings.

3 Motivating Evidence

In this section I present evidence showing how public guarantee agencies in South Korea effectively utilize a loan guarantee menu alongside additional information collection to screen borrow-

[11] Although the guarantee rate is 15% lower in the large loan program than in the small loan program (85% vs. 100%), the total amount at risk for the agency is still higher in the large loan program because 85% of \$37,000 (\$31,450) exceeds 100% of \$19,000 (\$19,000).

ers. Specifically, I show how agencies allocate larger loans to more creditworthy borrowers based on screening outcomes. Additionally, I present evidence of borrower self-selection into different guarantee contracts based on their risk profiles, exploring the trade-offs that lead to this sorting. The analysis focuses on how differing guarantee rates—85% for the large loan program and 100% for the small loan program—significantly influence loan contract terms, including loan size, interest rate, and funding probability.

3.1 Data

This study utilizes administrative data from the Korea Federation of Credit Guarantee Foundations (KOREG) and its 15 regional foundations, merged using unique borrower and loan identifiers.[12] The dataset encompasses detailed loan characteristics, including interest rates, maturities, loan sizes, guarantee rates, as well as comprehensive borrower and lender information, and repayment outcomes. Additionally, it includes data on applications approved for guarantees by KOREG but not subsequently funded by lenders, facilitating an analysis of lender decision-making processes regarding the 85% and 100% guarantee rates. In this study, lender rejection is defined as instances where a borrower's application with an 85% guarantee rate is not funded, followed by a subsequent application for the "small" loan program with a 100% guarantee rate within six months.

For empirical analysis, the focus is restricted to loans with a 5-year maturity issued in 2014. This selection accounts for approximately 50% of the sample and mitigates potential confounding effects related to varying loan maturities and the economic disruptions caused by the COVID-19 pandemic.[13] The 5-year loans involve borrowers typically repaying the principal in equal installments every three months.

Furthermore, the analysis concentrates on first-time borrowers, representing approximately 71% of the dataset, to minimize biases arising from agencies' prior knowledge of repeat borrowers. The sample is further narrowed to "general guarantee products", excluding specialized and partnership guarantee products to maintain consistency in borrower choices and agency evaluations.

Table 1 presents summary statistics for the final loan-level dataset. On average, approximately 19% of small business borrowers default on their loans, underscoring the high-risk nature of these loans and the necessity for guarantee programs. Notably, borrowers choosing the "small loan program" are more likely to default (23.6%) compared to those choosing the "large loan program"

[12] The data excludes Sejong, which did not exist in 2014, and Jeju, which did not offer program menus at the time.

[13] 1-year maturities are also common, accounting for roughly 30% of the sample, but 1-year loans often include an extension option, complicating their analysis.

Table 1: Summary Statistics

	Small loan	Large loan	All
Guarantee structure			
Money "at risk" for agency (guarantee rate)	100%	85%	
Money "at risk" for lender	0%	15%	
Application / Lender funding			
Number of borrowers	17,860	18,742	36,602
Number of guarantees not funded	-	1,773	-
Loan contract			
Guaranteed loan size, mean ($)	18,589	36,530	27,384
Interest rate, mean (%)	3.50	3.88	3.68
Loan performance			
Repayment rate, mean (%)	87.5	93.7	90.2
Defaulted (%)	23.6	13.3	19.1
Agency's loss per borrower, mean ($)	1,553	635	1,165
Borrower attributes			
Business age (years), mean	3.25	4.74	4.01
Credit score, mean	793.9	839.7	817.4
Number of employees, mean	1.47	1.87	1.67
Home ownership (%)	32.1	47.1	39.8
Service industry sector (%)	88.5	86.2	87.3
Agency attributes			
Number of regional agencies		15	
Agency's capital fund, mean ($)		9,209	

Notes: The columns labeled "Small loan" and "Large loan" present data based on borrowers' choices between the "small loan program", which offers a 100% guarantee rate, and the "large loan program", with an 85% guarantee rate. The repayment rates, default rates, interest rates, and agency loss reported under the "Large loan" column only include loans that were actually funded by the lender, as outcomes from unfunded loans are not available. The "All" column includes outcomes for all guarantees that were eventually funded, incorporating those initially rejected by the lender under the large loan program but subsequently approved under the small loan program. The guaranteed loan sizes in the "All" column are based on the guarantee rates ultimately funded. Repayment rate indicates the proportion of each loan that has been repaid, while the defaulted percentage reflects the proportion of borrowers who defaulted. These metrics are calculated as averages per borrower. Agency's capital fund data is normalized per guarantee, based on the number of guarantees issued by each regional agency within the year.

(13.3%), suggesting that riskier borrowers tend to opt for the "small loan program". Additionally, some borrowers opting for the "large loan program", due to its associated partial guarantee rate of 85%, face rejection from lenders. The average guaranteed loan size under the "large loan program" is significantly larger at $36,530, compared to $18,589 for the "small loan program". However, this benefit comes with a trade-off: borrowers choosing the "large loan program" encounter higher interest rates (3.88%) compared to those choosing the "small loan program" (3.50%), reflecting the increased risk from the lender's perspective due to the lower guarantee rate. Despite the agency guaranteeing larger loan sizes for the "large loan program", the agency's loss per borrower is substantially lower for the "large loan program" ($635) than for the "small loan program" ($1,553).

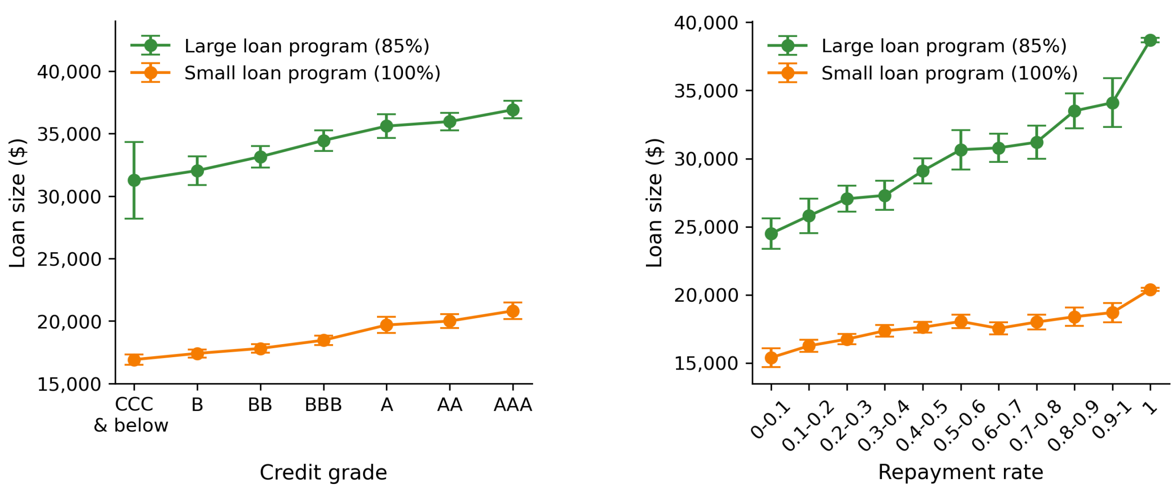
3.2 Evidence of Screening

In this section I show how guarantee agencies screen borrowers and allocate loan sizes based on screening outcomes, typically guaranteeing larger loans to more creditworthy borrowers.

Figure 1: Average loan size ($)

(a) Loan size vs. credit grade

(b) Loan size vs. repayment rate



Notes: In the left figure, the bins represent credit grades, ranging from CCC and below to AAA, where AAA represents the best grade. In the right figure, the repayment rate is discretized into eleven intervals: [0, 0.1), [0.1, 0.2), ..., [0.9, 1), and 1. The final interval exclusively includes borrowers who fully repaid their loans. For each interval, the average loan size, along with a 95% confidence interval, is plotted, distinguishing between loans for the "small loan program" and the "large loan program".

Figure 1a displays average loan sizes across different credit grades, showing that borrowers with higher credit ratings typically receive larger loans. Similarly, Figure 1b shows average loan size across repayment rates. The repayment rate, defined as the fraction of the loan that borrowers repay (ranging from 0 to 1), serves as a measure of loan performance. The figure reveals that larger loan sizes are positively correlated with higher repayment rates. In the next subsection, I show that the positive correlation between loan sizes and repayment rates persists even after controlling for observed borrower characteristics, including credit grades. This suggests that agencies are collecting additional soft information that predicts repayment behavior to determine loan sizes, highlighting the effectiveness of the agencies' screening methods.

To address potential concerns that larger loan sizes might themselves influence repayment rates, I utilize a regression discontinuity design detailed in Online Appendix C. The analysis confirms that increases in loan size do not significantly affect repayment rates, indicating that the observed correlation is primarily a result of the agencies' screening rather than the effect of loan size itself.

3.3 Guarantee Choice and Loan Contracts

In this section I examine how the choice between the "small" and "large" loan programs affects loan contracts, highlighting how different types of borrowers evaluate the trade-off differently.

I first investigate the sorting behavior of borrowers into different guarantee contracts: low-risk borrowers are more likely to choose the large loan program with an 85% guarantee, while high-risk borrowers often opt for the small loan program with a 100% guarantee. Then, I present descriptive evidence on the trade-offs associated with the guarantee choice, focusing on (1) lender funding probability, (2) interest rates, and (3) loan sizes.

Borrower sorting I examine how borrowers with a range of risk profiles self-select into different loan guarantee programs, using repayment rate as a proxy for risk. To quantify this relationship, column (1) of Table 2 reports results from the following regression model:

$$\lambda_i = \psi_l Large_i + X_i \Psi + v_i$$

Here, λ_i represents the ex-post repayment rate for borrower i, $Large_i$ is a binary indicator variable denoting whether borrower i chose the large loan program, and X_i is a vector of observable borrower characteristics, including credit score, business age, home ownership, number of employees, and the contract interest rate. Industry and regional fixed effects are also included. A positive ψ_l coefficient suggests that borrowers opting for the large loan program are associated with higher repayment rates, indicative of selection effects.[14] This evidence supports the notion that guarantee choice is informative of a borrower's lower risk profile.

[14]Loan size is not included as a control variable in this model because, as demonstrated in the Online Appendix C, it does not affect repayment outcomes.

Funding probability The probability of a loan being funded by the lender depends on the guarantee rate as shown in Table 1. Selecting the small loan program, which comes with a 100% guarantee rate, ensures full funding for the borrower. Conversely, choosing the large loan program, associated with an 85% guarantee rate, introduces a potential risk of lender rejection due to the 15% of the loan that remains unguaranteed.

Table 2: Reduced Form Analysis

Variable	**Repayment Rate (1)**	**Funded (2)**	**Interest Rate (3)**	**Loan Size (4)**
Large	0.023		0.533	13.362
	(0.003)		(0.019)	(0.444)
Repayment rate		1.784	-0.047	0.935
		(0.074)	(0.012)	(0.327)
Large ×			-0.208	2.466
Repayment rate			(0.021)	(0.474)
Credit score	3.05e-04	0.006	-5.88e-04	0.011
	(9.94e-06)	(<0.001)	(2.21e-05)	(<0.001)
Business age	0.002	0.046	-0.001	0.093
	(<0.001)	(0.004)	(0.001)	(0.010)
Home-owner	0.051	0.043	-0.020	2.466
	(0.003)	(0.047)	(0.008)	(0.123)
Num of employees	-0.001	0.001	3.02e-05	1.062
	(0.001)	(0.001)	(4.91e-05)	(0.027)
Interest rate	-0.01			
	(0.001)			
Industry FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	36,602	18,742	34,829	36,602
R^2	0.082		0.235	0.413
Log likelihood		-6521.5		

Notes: Standard errors (clustered by region) are shown in parentheses. The variable "Large" is an indicator variable representing the large loan program as opposed to the small loan program. Column (1) reports OLS estimates where the dependent variable is the repayment rate. Column (2) reports the estimated coefficients of the Probit model, where the dependent variable indicates whether the 85% guarantee is funded by the lender. Column (3) reports OLS estimates where the dependent variable is the interest rate conditional on the loan being funded. Column (4) uses OLS to analyze loan size, including all guarantees, regardless of funding status. For guarantees initially rejected but later funded under 100% guarantee rate, the repayment rate associated with the subsequent 100% guarantee rate is used.

Within the dataset, there are applications approved for the large loan program with an 85% guarantee by the agency but later canceled before the disbursement of funds. A subset of these applicants subsequently reapplies and successfully secures a loan with the small loan program with a 100% guarantee. These specific instances of application cancellations, followed by successful reapplications under a 100% guarantee rate, serve as a measure of lender rejections for the 85% guarantee.

To analyze the funding probability associated with 85% guarantees, I employ the following probit model:

$$Funded_i = \mathbf{1}[\psi_\lambda \lambda_i + X_i \Psi + v_i \geq 0]$$

$Funded_i$ is a dummy variable indicating whether the loan application with an 85% guarantee (under the large loan program) for borrower i is funded by the lender. The variable λ_i represents the borrower's repayment rate, and X_i includes the same set of observable characteristics as previously mentioned, excluding the contract interest rate.[15]

The results, presented in Column (2) of Table 2, suggest that lower-risk borrowers are more likely to secure funding under a 85% guarantee. The coefficient ψ_λ measures the correlation between borrowers' repayment rates and their probability of securing funding, serving as a proxy for lenders' ability to evaluate unobserved risks when deciding to accept or reject loan applications.

Interest rate Interest rates are also influenced by the guarantee choice due to the associated guarantee rates. To analyze the impact of choosing between the small loan program (100% guarantee) and the large loan program (85% guarantee) on interest rates, the following regression model is employed

$$r_i = \psi_l Large_i + \psi_\lambda \lambda_i + \psi_{\lambda l}(\lambda_i \times Large_i) + X_i \Psi + \xi_i$$

r_i denotes the interest rate for borrower i, $Large_i$ as an indicator variable for whether borrower i chooses the large loan program with 85% guarantee rate, λ_i is the repayment rate, and X_i is the same vector of controls as above.

The results, detailed in Column (3) of Table 2, reveal that 85% guarantees typically lead to higher interest rates than 100% guarantees, as indicated by a positive ψ_l. Additionally, the negative $\psi_{\lambda l}$ coefficient indicates that the increase in interest rates associated with choosing a 85% guarantee is smaller for low risk borrowers. This suggests that high-risk borrowers benefit more from 100% guarantees compared to low-risk borrowers, because low risk borrowers are already eligible for favorable interest rates and thus have a smaller benefit from 100% guarantees.

Loan size The final loan size set by the agency is influenced by the borrower's choice between the small and large loan programs. To analyze the impact of the guarantee choices on loan sizes, I

[15] For applications that were initially not funded under the 85% guarantee and subsequently secured under the 100% guarantee, I use the repayment rate (λ_i) observed from these successfully funded loans to analyze the funding probability associated with the original 85% guarantee.

use the following regression model:

$$L_i = \psi_l Large_i + \psi_\lambda \lambda_i + \psi_{\lambda l}(\lambda_i \times Large_i) + X_i\Psi + \xi_i$$

L_i represents the loan size for borrower i, $Large_i$ indicates whether borrower i chooses the large loan program, λ_i is the repayment rate, and X_i contains the same control variables as before.

In the column (4) of Table 2, a positive ψ_l indicates that loans under the large loan program tend to be larger than those under small loan program. Furthermore, a positive $\psi_{\lambda l}$ suggests that the correlation between loan size and repayment rate is significantly higher under the large loan program. This finding is consistent with the more rigorous screening efforts for the large loan program, which enables agencies to accurately align loan sizes with the borrowers' risk levels.

Interpretation of the result This analysis suggests the trade-offs borrowers face when choosing between the large and small loan programs. The large loan program, despite offering potentially larger loans, comes with an 85% partial guarantee, which can result in a lower funding probability and higher interest rates compared to the small loan program's full (100%) guarantee.

For low-risk borrowers, the benefits of choosing the large loan program are pronounced. Typically, lenders provide favorable loan terms to low-risk borrowers, such as low interest rates and high funding probabilities, regardless of the guarantee rate. Thus, the incremental benefits of the full guarantee in terms of interest rates are less significant for the low-risk borrowers. Instead, the larger loan sizes available through the large loan program are more appealing. Moreover, these low-risk borrowers can leverage the detailed screening process of the large loan program to reveal their true type, potentially securing even larger loans.

Consequently, guarantee agencies anticipate this sorting behavior in risk types and allocate even larger loan sizes to borrowers opting for the large loan program. This heterogeneity in borrower preferences helps sustain the separation in equilibrium.

4 Empirical Model

This section develops a screening model of loan guarantee menus and agency loan size decisions. A simplified conceptual framework that provides the basic intuition is presented in the Online Appendix A. The empirical model here formalizes that intuition by introducing richer borrower heterogeneity and incorporating the objectives of regional government agencies. Specifically, it details the interactions between borrowers, indexed by i, and regional government agencies, indexed by j. Conditional on observables, borrowers are characterized by two-dimensional types: their inherent repayment type η_i and their preference shock on guarantee rate denoted as ϵ_i. Agency j does not observe the borrower types, but the agency observes the guarantee choice G_i ("small" or "large" loan program) and also receives a noisy signal s_i^G on borrower's repayment type η_i.

Agencies differ in their priorities (τ_j), balancing the economic value added from supporting small businesses against their financial returns. Conditional on the borrower's guarantee choice and the additional signal the agency receives, the agency determines the guaranteed loan size. Then loan funding and borrower repayment are realized.

Figure 2 depicts the overall structure of the model. A more detailed description of the timing of the model is as follows:

1. Borrower i, with private information about their repayment type and guarantee preference, is matched with a regional agency j.

2. The borrower selects a guarantee program G_i, which can be the small loan program with a 100% guarantee rate or the large loan program with an 85% guarantee rate.

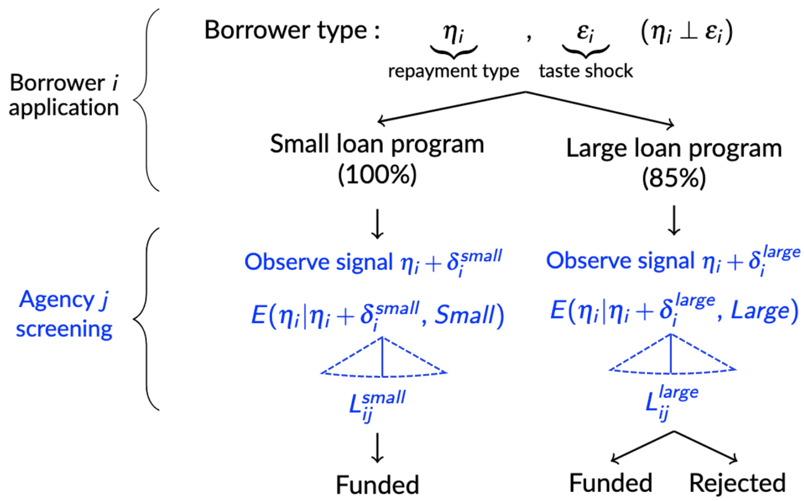
3. Following the borrower's guarantee choice, agency j collects additional information and receives a noisy signal s_i^G on borrower i's repayment type. The precision of the signal is tied to the guarantee choice G_i. Conditional on both the guarantee choice G_i and the signal s_i^G, the agency decides on the loan size (L_{ij}) for the borrower.

4. The lender evaluates the guaranteed contract and makes a decision on whether to provide funding.

5. Conditional on loan funding, the borrower repays according to their repayment type η_i and random repayment shock υ_i.

To keep the focus on the screening game between the borrower and the government agency, and not on the borrower's choice of lenders, I make a key assumption: lenders are considered homogeneous and perfectly competitive. This assumption implies that the borrower's interest rate and the probability of funding are the same across all lenders, abstracting away from the complexities of the borrower's choice in lenders.

Figure 2: Overview of the model



4.1 Borrower Utility from Loan Contract

In the model, a borrower i obtains a loan of size L_i with an interest rate of r_i, which is then invested into their small business. Each business i has a stochastic investment technology that produces output as a function of investment size L_i:

$$f_i(L_i) = \lambda_i A_i L_i^{\alpha}$$

where λ_i represents the repayment rate, or the fraction of the loan that is actually repaid ex-post. This factor also serves as a proxy for the borrower's overall productivity—those who repay more are assumed to be more productive. A_i is a productivity shifter that enhances the output, accounting for factors that affect the business's output not related to the repayment. α denotes the concavity of the production function, illustrating diminishing returns to additional investment.

This formulation simplifies the complex dynamics of business productivity by directly linking it to the borrower's repayment behavior. It assumes that the productive output is directly proportional to the loan repayment rate, a simplifying assumption that abstracts from the more complex and less predictable aspects of business operations. For instance, if a borrower fully repays the loan, the output is maximized at $A_i L_i^{\alpha}$; conversely, if no repayment is made, the output is zero.

The borrower's ex-post utility from the loan contract is given by:

$$U_i(\lambda_i, L_i, r_i) = \underbrace{\lambda_i \cdot A_i L_i^{\alpha}}_{total\ output} - \underbrace{\lambda_i \cdot L_i}_{total\ principal\ repayment} - \underbrace{a(\lambda_i, r_i) \cdot L_i}_{total\ interest\ payment} - \underbrace{\mathbf{1}[\lambda_i < 1] \cdot D_i}_{default\ cost}$$

Here, the interest payment function $a(\lambda_i, r_i) \cdot L_i$ converts the total interest payments into a form aligned with the repayment rate λ_i and interest rate r_i. D_i represents the default cost incurred when repayment falls below full, encapsulating financial penalties and other negative consequences.

A more detailed derivation and justification of this utility function, including its microfoundations, is provided in Online Appendix D. In the microfoundations, the repayment rate λ_i approximates the "productive life-cycle" of a firm with respect to its investment returns. For example, a 50% repayment rate implies that the business is effectively producing outputs for 50% of the loan's maturity period before defaulting. Therefore, λ_i scales both repayment and output of the small business borrowers. This formulation of borrower utility positions λ_i as the sole source of information asymmetry between the agency and the borrower. By serving both as the repayment rate and a scaling factor for the production function, λ_i simplifies the screening problem to a single dimension. This focus on a single variable makes the model tractable, avoiding the complexities typically associated with multi-dimensional screening models.

4.2 Borrower Repayment Rate

The repayment rate λ_i for a borrower i is determined by two key factors: the inherent repayment type of the borrower, denoted as η_i, and an exogenous random shock, υ_i. In our model, η_i indicates the borrower's quality. A high η_i value (high type) suggests high repayment ability and low risk of default, whereas a low η_i value (low type) indicates lower repayment ability and higher risk. Each borrower knows their own repayment type η_i, but this information is not directly observable by the guarantee agency. The random shock υ_i, representing unforeseeable fluctuations in repayment capacity post-loan origination is unknown to both the agency and the borrowers. By definition, these shocks are uncorrelated with the borrower's repayment type, ensuring that they represent truly exogenous influences on repayment rates.

I assume that these components are additively separable, allowing us to construct λ_i as the sum of the borrower's repayment type and the random shock. However, since λ_i represents a repayment rate, it must logically be constrained within the range $[0, 1]$, signifying the fraction of the principal that is repaid. Thus, I censor λ_i to ensure it remains within this valid range:

$$\lambda_i = min\{max\{\underbrace{\eta_i}_{borrower's\,type} + \underbrace{\upsilon_i}_{random\,shock}, 0\}, 1\}$$

This censoring of λ_i rationalizes a mass point at 1, where many borrowers fully repay, while also capturing the variations in repayment rates between 0 and 1. This approach preserves the valuable data on these variations, which a simple binary model—distinguishing only between full repayment and any level of default—would ignore.

This formulation implies that the borrower repays a portion $\lambda_i L_i$ of the loan principal back to the lender. For the analysis, I consider λ_i to be exogenous, unaffected by variations in loan size and interest rates. This assumption is common in small business loan literature, as demonstrated by studies such as Cox et al. [2021] and Stillerman [2022]. The empirical findings, detailed in the Online Appendix C, provide support for this assumption by showing that increases in loan size do not significantly affect the repayment rate.[16]

4.3 Borrower Utility from Guarantee

The guarantee option the borrower opts for—the "large loan program" with an 85% guarantee rate or the "small loan program" with a 100% guarantee rate—significantly affects the terms of the loan contract (L_i, r_i) and, consequently, the utility derived from it. L_{ij}^{small} and L_{ij}^{large} represent the loan sizes for borrower i set by agency j for the small and large loan program, respectively. Likewise, r_i^{small} and r_i^{large} are the interest rates assigned to borrower i by the lender under the small and large loan programs, respectively.

The small loan program, offering the full 100% guarantee rate, eliminates lender risk, resulting in lower interest rates r_i^{small} and ensures funding for all borrowers. However, under this guarantee, the agency sets a smaller loan size L_{ij}^{small} because it absorbs all the risk. In contrast, the large loan program, with an 85% guarantee rate, involves shared risk between the lender and the guarantee agency, leading to higher interest rates r_i^{large} and potential uncertainty in loan funding, with $P_{i,85}^{funding}$ indicating the probability of a loan being funded by the lender under the large loan program with an 85% guarantee. The advantage of the large loan program is that it allows the agency to set a larger maximum loan size L_{ij}^{large}.

L_{ij}^{small} and L_{ij}^{large} are equilibrium objects, determined through interactions between borrowers and the guarantee agency. The borrowers form beliefs about the size of the loan they might receive based on the guarantee choice, conditional on their borrower type η_i. To simplify the model and focus on the interaction between the borrower and the guarantee agency, I assume that the interest rates (r_i^{small}, r_i^{large}) are known to the borrower. This assumption eliminates the need to model the lender's decision-making process regarding interest rates, allowing me to concentrate on how

[16]It is noteworthy that allowing λ_i to vary with loan size and interest rates might introduce non-monotonicities in the mapping from signals to loan sizes, which would complicate the analysis.

borrowers and the agency strategically choose the guarantee type (small or large loan program) and loan sizes.[17]

4.3.1 Lender's Funding Probability

Lenders decide on funding loans under the large loan program, which offers an 85% guarantee, by evaluating a noisy signal related to the borrower's repayment type η_i, along with a random noise component ζ_i. The decision rule can be expressed as:

$$Funded_i = \mathbf{1}[\kappa_\eta \cdot \eta_i + \zeta_i > 0]$$

The guaranteed loan is approved and funded by the lender if it meets this condition; otherwise, it's rejected. Note that $\kappa_\eta \cdot \eta_i + \zeta_i$ can be interpreted as the lender's imperfect signal about the borrower's repayment type η_i. Here κ_η is a constant parameter that captures the screening precision of the lender, with higher values of κ_η indicating more precise screening capabilities. ($P_{i,85}^{fund}$) is taken as exogenous and known by the borrower. In contrast, loans under the small loan program with 100% guarantees are always funded (i.e. $P_{i,100}^{fund} = 1$), consistent with the data, reflecting the lender's perception of negligible risk due to the complete backing by the government agency.

Note that loan size does not influence the decision rule, indicating a form of risk neutrality with respect to loan size. Lenders set interest rates based on factors that ensure a positive profit margin, focusing on the borrower's repayment rate rather than the potential magnitude of losses in case of a default. The model further implies that the size of the loan guaranteed by the agency does not serve as a signal to lenders regarding borrower risk, aligning with insights from discussions with lending officers involved in funding loan guarantees. This assumption simplifies the analysis by allowing us to treat the funding probability as exogenous.[18]

4.3.2 Expected Utility with the Small and Large Loan Program

Given the 100% guarantee by the government agency in the small loan program, lenders are inclined to fund these loans without hesitation, eliminating the possibility of rejection. Therefore the borrower's expected utility under the small loan program, EU_{ij}^{small} is straightforward and does

[17] In Korea, the agencies operate a a system known as "Interest Rate Informer" that provides data on the average interest rates for each guarantee rate (85% and 100%), offered by various lenders, conditioned on the borrower's credit grade. This system helps borrowers to understand the interest rates they should anticipate, supporting the rationalization of this assumption.

[18] Introducing loan size into the funding decision would complicate the model significantly, as the funding probability would then depend on the agency's screening and loan size decisions, which would, in turn, affect the borrower's choice of guarantee rate and further influence the screening outcomes.

not consider rejection risk:

$$EU_{ij}^{small} = \int\int U_i(\lambda_i, L_{ij}^{small}, r_i^{small}) dF_{\lambda_i, L_{ij}^{small}|\eta_i}$$

For the large loan program, which offers a 85% guarantee, the borrower's expected utility, EU_{ij}^{large}, incorporates the probability of lender funding, represented by $P_{i,85}^{fund}$. If funded, the borrower benefits from the larger loan. If rejected, she incurs significant costs, captured by the term c, such as decreased credit scores and delays in securing a loan, and then reapplies for the small loan program with a 100% guarantee:

$$\begin{aligned} EU_{ij}^{large} = & P_{i,85}^{fund}(\eta_i) \int\int \Big(U_i(\lambda_i, L_{ij}^{large}, r_i^{large}) dF_{\lambda_i, L_{ij}^{large}|\eta_i} \\ & + \big(1 - P_{i,85}^{fund}(\eta_i)\big) \int\int \big(U_i(\lambda_i, L_{ij}^{small}, r_i^{small}) - c \big) \Big) dF_{\lambda_i, L_{ij}^{small}|\eta_i} \end{aligned}$$

4.3.3 Borrower Guarantee Choice

Borrowers choose the large loan program over the small loan program if the expected utility from the large loan program EU_{ij}^{large} exceeds that from the small loan program EU_{ij}^{small}. The condition for opting for the large loan program is given by:

$$\begin{aligned} EU_{ij}^{large} - EU_{ij}^{small} = {} & P_{i,85}^{fund}(\eta_i) \int\int U_i(\lambda_i, L_{ij}^{large}, r_i^{large}) dF_{\lambda_i, L_{ij}^{large}|\eta_i} \\ & + \big(1 - P_{i,85}^{fund}(\eta_i)\big) \cdot \big(EU_{ij}^{small} - c\big) - EU_{ij}^{small} > 0 \\ \iff E\Big(\underbrace{\lambda_i \cdot \triangle(A_i \cdot L_{ij}^{\alpha} - L_{ij})}_{\text{diff in net output}} |\eta_i\Big) - E\Big(\underbrace{\triangle interest_{ij}}_{\text{diff in interest}} |\eta_i\Big) & - \underbrace{\Big(\frac{1 - P_{i,85}^{fund}(\eta_i)}{P_{i,85}^{fund}(\eta_i)}\Big) \cdot c}_{\text{disutility from rejection}} + \epsilon_i > 0 \end{aligned}$$

This inequality describes the borrower's decision process, weighing the expected difference in net output and interest payments along with the disutility associated with rejection. The decision is influenced by the borrower's repayment type η_i, indicating that borrowers with different risk profiles might assess this trade-off differently. Here, we introduce ϵ_i as a preference shock, capturing the idiosyncratic preferences of each borrower toward the large loan program versus the small loan program, uncorrelated with the borrower's repayment type η_i. This assumption allows us to capture borrower behaviors and preferences that are not tied to their risk characteristics.

4.4 Government Agency Objective

The government loan guarantee agency, denoted by j, balances two objectives: to support small businesses in achieving their appropriate loan size while also minimizing the financial burden by reducing losses from the loan guarantee.

4.4.1 Support for Small Businesses

The agency seeks to maximize the value generated by small businesses, such as output, employment, and local economic growth. The value added by supporting a business i is captured as:

$$VA_i = \underbrace{\lambda_i A_i L_i^{\alpha}}_{\text{firm output}} - L_i$$

The value-added effect from supporting a small business is quantified by the difference between the business output $\lambda_i A_i L_i^{\alpha}$ and the loan size L_i. The optimal loan size that would maximize value-added, $L_i^{*VA} = (\alpha A_i \lambda_i)^{\frac{1}{1-\alpha}}$, is not feasible due to practical constraints. The agency does not know the borrower's repayment rate, λ_i, and it operates under budget limitations that necessitate careful management of potential losses. These constraints require the agency to balance the value-added objective with financial sustainability.

4.4.2 Financial Sustainability

To ensure the program's sustainability, the agency also aims to minimize the losses from the guarantees. The profit, or potentially the loss, from issuing a guarantee is determined by incorporating the repayment rate λ_i, the fees, and the guarantee rate g_i, which is directly associated with the borrower's guarantee choice G_i. Specifically, g_i=85% for the large loan program ($G_i = large$) and g_i=100% for the small loan program ($G_i = small$). The agency's profit or loss from borrower i can be expressed as:

$$\pi_i = (-1 + \lambda_i + fee_i) \times g_i \times L_i$$

4.4.3 Information Acquisition and Loan Size Decision

The agency determines the loan size for each borrower using an objective function to optimize the value-added by small businesses VA_i and the agency's operational profit or loss π_i. The parameter τ_j represents the weight that agency j assigns to the small business value-added component, acknowledging that agencies might weight this aspect differently:

$$\Pi_{ij} = \tau_j \cdot \underbrace{VA_i}_{\text{SB value-added}} + (1 - \tau_j) \cdot \underbrace{\pi_i}_{\text{agency profit}}$$

This use of a hybrid objective function is similar to the approach in Timmins [2002], where a regulator's preferences are modeled as a combination of social benefits and operational profits. While π_i is defined as the agency's operational profit, if borrowers default on their loans it turns negative and becomes a cost to the agency. Therefore, the agency seeks to maximize the value added by supporting small businesses while minimizing potential losses arising from defaults.

The sole source of information asymmetry in this decision-making process is the borrowers' ex-post repayment rates λ_i. During the agency's screening process, the focus is on gaining information on the borrower's repayment type, η_i, as it is the screenable component of ex-post repayment rate λ_i. The agency receives two types of noisy signals regarding η_i. The first signal comes from the borrower's guarantee choice (G_i), of either the small or the large loan program. The second signal comes from the agency's soft information collection, yielding a signal $s_i^G = \eta_i + \delta_i^G$, where δ_i^G represents the noise of the signal. Importantly, I allow the variance of δ_i^G to differ between the large loan program ($G_i = large$) and the small loan program ($G_i = small$) to reflect the agency using different levels of screening effort. The large loan program undergoes a more detailed evaluation, while the small loan program has a simplified evaluation.

Using these two signals, the agency updates its beliefs about the borrower's repayment type using Bayes' rule, and consequently the repayment rate. The optimal loan size is then calculated by maximizing the expected objective:

$$L_{ij}^{*agency} = argmax\, E(\Pi_{ij}|G_i, s_i^G) = \Big(\frac{\alpha\tau_j E(\lambda_i|G_i, s_i^G)A_i}{\tau_j + (1-\tau_j)\cdot(1 - E(\lambda_i|G_i, s_i^G) - fee_i)\cdot g_i}\Big)^{\frac{1}{1-\alpha}}$$

In the model I assume that the government loan guarantee agency optimizes each loan individually rather than adopting a global objective with a budget constraint. This is consistent with the operational reality that guarantee officers assess borrowers one at a time, without immediate consideration of previous decisions. This approach is not only more aligned with the practical workings of such agencies, which often operate under soft budget constraints that can be adjusted in response to economic conditions like the COVID-19 pandemic, but also simplifies the mathematical complexity of the model. Avoiding a global optimization that requires anticipating all potential loans within a period, which is unrealistic and impractical, this method focuses on the agency's ability to balance the support for businesses and the agency's financial sustainability through the parameters τ_j and $(1-\tau_j)$, effectively capturing the hybrid objectives of the agency's operations.

4.5 Equilibrium

The equilibrium in this model is defined as a Perfect Bayesian Equilibrium (PBE), where all players-borrowers and government agencies-act optimally based on consistent beliefs updated through Bayesian inference in response to the actions and information signals observed throughout the game. Borrowers choose between full and partial guarantee contracts based on their repayment type and expectations about agency responses. Government agencies adjust loan sizes based on these choices and the signals they received. The formal conditions for equilibrium existence, including the verification of the single-crossing property and separation, are provided in Online Appendix F.

5 Identification and Estimation

5.1 Parameterization

I now provide details on the parameterization of the model. As I discussed above, the repayment rate λ_i is influenced by two key factors: the repayment type η_i and a random shock υ_i. The repayment type η_i follows a normal distribution, $N(\mu_i, \sigma_\eta^2)$, where μ_i represents the mean of this distribution and is known to the agency. The mean μ_i is derived as a linear function of observed borrower characteristics, formulated as $\mu_i = X_i\Gamma$. This vector X_i includes a constant, the borrower's credit score, the age of the business, and an indicator for whether the borrower owns a home. These observables are chosen because they are key predictors of the borrower's repayment risk. The random shock υ_i, representing unforeseeable fluctuations in repayment capacity post-loan origination, follows a normal distribution, $N(0, \sigma_\upsilon^2)$, unknown to both the agency and the borrower.

Regarding the lender's funding rule on partial guarantees, the model incorporates a linear function of the same borrower observables X_i along with the repayment type η_i. Additionally, ζ_i, which follows a Type 1 extreme value distribution ($T1EV$), is included to account for the noisy signal in the lender's evaluation process. This assumption simplifies the borrower's decision for a partial guarantee over full guarantee, expressed through the following utility comparison:

$$EU_{ij}(p) - EU_{ij}(f) > 0 \iff E\Big(\lambda_i \cdot \underbrace{\triangle(A_i \cdot L_{ij}^{\alpha} - L_{ij})}_{\text{diff in output}} |\eta_i\Big) - E\Big(\underbrace{\triangle interest_{ij}}_{\text{diff in interest}} |\eta_i\Big)$$

$$- \underbrace{exp(-K_X \cdot X_i - \kappa_\eta \cdot \eta_i) \cdot cost}_{\text{disutility from rejection}} + \epsilon_i > 0$$

The potential productivity of a small business, A_i, is modeled as a linear function of the number of employees and an indicator variable for whether the business operates within the service industry. This parameterization captures the influence of labor capacity and sector specifics on business productivity. The preference shock on guarantee rate ϵ_i follows $N(0, \sigma_\epsilon^2)$.

The parameter τ_j, representing the agency's preference for balancing small business value added against own profits, is formulated as a linear function of the agency's capital fund, normalized by the number of guarantees issued in a given year. This relationship captures the notion of soft budget constraints: agencies with more substantial capital funds may prioritize the small business's value added over their own profitability.

Finally, the agency updates its beliefs about the borrower's repayment type using Bayes' rule and receives two types of noisy signals regarding η_i. The first signal is the borrower's choice of guarantee rate G_i—either the small loan program or the large loan program. The second signal from agency's soft information collection is denoted as $s_i^G = \eta_i + \delta_i^G$, where $\delta_i^G \sim N(0, \sigma_G^2)$ represents the noise in the agency's screening process, with σ_G^2 indicating the screening precision. Importantly, this precision varies based on the guarantee choice G_i: σ_{small}^2 for the small loan program and σ_{large}^2 for the large loan program. This variation reflects the agency's strategy of conducting more detailed evaluations for the large loan program due to the larger loan sizes involved, compared to more simplified evaluations for the small loan program.

While $\eta_i|s_i^G \sim N(\mu_i + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_G^2} s_i^G, \frac{\sigma_\eta^2\sigma_G^2}{\sigma_\eta^2+\sigma_G^2})$, incorporating the guarantee choice makes the posterior distribution $\eta_i|G_i, s_i^G$ analytically intractable. Therefore, to estimate the posterior beliefs about the borrower's repayment type—and consequently the repayment rate $E(\lambda_i|G_i, s_i^G)$—I employ a numerical simulation approach. The specifics of this simulation method are elaborated in the Online Appendix G.3.

5.2 Identification

I now discuss the identification of the model's parameters.

The variance of the ex-post repayment rate recovers the sum $\sigma_\eta^2 + \sigma_\upsilon^2$ but it does not allow us to distinguish between these two components individually. To address this challenge, I initially make an assumption that $\sigma_\upsilon^2 = 0$. This assumption of perfect foresight implies that borrowers are completely aware of their repayment rate. Although strong and somewhat unrealistic, this assumption is a practical starting point because it simplifies the identification arguments. Under this assumption, all observed variation in the ex-post repayment rates can be attributed directly to differences in the borrower's repayment type η_i, which simplifies the analysis. Moreover, in the descriptive evidence section, I demonstrated that borrower choices are significantly correlated with their repayment rates, suggesting that borrowers possess considerable private information. By assuming full awareness, the complexities of separately identifying the effects of this private information from random shocks are circumvented. Future stages of the research will explore different values for σ_υ^2 to assess the robustness of the model's predictions and enhance understanding of how uncertainties in the borrower repayment rate influence the effectiveness of the screening process.

Further, the lender's funding probability for 85% guarantee rate, $P_{i,85}^{fund}$, is informed by the observed funding rate. Given the borrower's repayment type and rejection probability, the cost of rejection is informed by a negative correlation between rejection probability and the choice of the large loan program. The business productivity shifter A_i is identified by borrower guarantee choices. Specifically, when borrowers opt for the large loan program, they receive larger loan sizes at higher interest rates. Their willingness to accept increased interest costs for these larger loans identifies A_i, highlighting their expectation of sufficient returns to offset the higher cost.

The parameter α, which reflects the concavity of the production function, is informed by how guarantee agencies allocate loans to borrowers with observably different risk profiles. As α approaches 1, the production function becomes less concave, indicating that the marginal output generated from increasing loan sizes remains relatively constant. Consequently, agencies are incentivized to allocate substantially larger loans to observably low risk borrowers (high μ_i) because the expected return on these larger loans remains high. Conversely, as α decreases, the production function becomes more concave, leading to a more rapidly diminishing returns to loan size. Therefore, agencies allocate more similar loan sizes to both observably low-risk and high-risk borrowers. The difference in loan sizes between observably low-risk and high-risk borrowers identifies the concavity parameter α.

However, the current model assumes that agencies primarily focus on maximizing their profit and the economic value added by the businesses they support. In practice, as public guarantee programs, these agencies may also take into consideration of equity concerns, aiming to provide same opportunities to all small businesses. Such equity considerations could impact loan allocation decisions, leading agencies to limit the size of loans offered to observably safer borrowers, thereby potentially driving the estimation of α downward. Consequently, separately identifying the impacts of fairness from the concavity of the borrower's production function becomes challenging.[19]

The parameter τ_j reflects the agency's preference for maximizing the economic output (value added) of small businesses over minimizing its own financial losses. I identify τ_j by observing how the agency adjusts loan sizes based on business productivity characteristics (A_i)—such as employee count and industry type—that influence the potential value added but are uncorrelated with repayment risk (η_i). Conditional on the borrower's repayment type and guarantee rate, if the agency consistently allocates larger loans to businesses with higher A_i, this indicates a higher τ_j, revealing that the agency places greater weight on supporting productive businesses even at the expense of higher financial risk.

Finally, similar to that in Wang [2020], the screening precision from the agency's soft information collection, σ_{small} and σ_{large}, is identified by the correlation in loan sizes and ex-post repayment rate. Consider two conditional loan distributions: $(L_i|X_i, G_i, \lambda_i = 1)$ for loans that were fully repaid, and $(L_i|X_i, G_i, \lambda_i < 1)$ for loans that defaulted. A decrease in the standard deviation of signal noise increases the separation between these conditional distributions, thereby informing the precision of information from the business evaluation. It is important to note that while I interpret this screening precision as soft information collected by guarantee officers, it might also capture hard information not explicitly included in the dataset or not controlled for.

5.3 Estimation

I estimate the model by maximum likelihood by matching the probabilities of various guarantee outcomes as observed within the dataset. The guarantee outcomes of interest are fourfold: borrower repayment λ_i, borrower's guarantee choice G_i, the loan size decision made by the agency L_i, and the eventual funding outcome by the lender $Funded_i$. The likelihood function that encapsulates the joint probability of these guarantee outcomes, conditional on the model parameters Θ is given

[19] If equity considerations are indeed a fundamental part of the agency's broader objectives, this does not detract from the core aim of this study, which is to analyze how screening mechanisms—through menus and soft information collection—affect the agency's comprehensive objectives. By examining the effects of counterfactual scenarios on these objectives, whether they emphasize equity or focus solely on the output of firms, my analysis continues to reflect the diverse goals that public guarantee programs aim to fulfill. Nonetheless, incorporating these broader considerations does make it more challenging to assign structural significance to the parameters.

by:

$$\begin{aligned}\mathcal{L}(\lambda_i, G_i, L_i, Funded_i|\Theta) = &Pr(\lambda_i|\Theta) \times Pr(G_i|\lambda_i, \Theta)\\ &\times Pr(L_{ij}|G_i, \lambda_i, \Theta)\\ &\times Pr(Funded_i|L_i, G_i, \lambda_i, \Theta)\end{aligned}$$

Here, Θ represents the set of all model parameters, including those related to the borrowers' repayment types, the agency's screening accuracy, and the lender's funding criteria. Further details on the full likelihood specification, and the estimation algorithms are provided in the Online Appendix.

6 Results

6.1 Model Estimates

Repayment type Table 3-A presents the estimates for the repayment type. As expected, borrowers with a higher credit score, longer-established business, and homeownership repay more on their loan contracts. However, our estimate of σ_η indicates substantial unobserved borrower risk, as these observable factors explain less than 10 percent of the total variance in repayment behavior. The considerable unexplained variance underscores the importance of employing advanced screening mechanisms to better understand and manage the diverse repayment capacities of borrowers.

Investment technology Table 3-B presents the estimates for the borrower's investment technology. The concavity of the borrower's production function (α) is estimated to be 0.896. This indicates a nearly linear relationship between loan size and output, which is reasonable given that loans typically range from \$10,000 to \$70,000, suggesting limited diminishing returns within this range. This estimate is slightly lower but still comparable to the estimates (0.91-0.93) from Cox et al. [2021]. Additionally, the technology shifter A_i increases with the number of employees, representing the size of the business, and is larger for businesses in non-service sectors, reflecting sector-specific productivity differences.[20]

Funding probability The estimates from Table 3-C indicate that the rejection rate for loans with an 85% guarantee rate varies significantly based on borrower repayment types, with lower

[20] Non-service sectors mainly include manufacturing, construction, agriculture, mining, and forestry, while service sectors mainly encompass hospitality and food services, wholesale and retail trade.

Table 3: Model Parameter Estimates

	Panel A Repayment		**Panel B Investment**		**Panel C Funding for 85%**	
	Estimate	**S.E.**	**Estimate**	**S.E.**	**Estimate**	**S.E.**
Borrower Risk Covariates						
Credit Score	0.001	<0.001			0.012	<0.001
Business Age	0.007	<0.001			0.059	0.005
Homeownership	0.146	0.006			0.015	0.004
Repayment Type (η)					0.398	0.037
Technology Shifter and Concavity						
Employee Count ($A_{employee}$)			0.079	0.006		
Service Industry ($A_{service}$)			-0.159	0.058		
Concavity (α)			0.896	0.003		
Cost of Rejection						
Cost (c)					6.513	0.163
Repayment/Preference Shock						
Standard Deviation (σ_η, σ_ϵ)	0.854	0.009	11.336	0.721		
Constant ($\bar{\eta}$, $\bar{A}$, $\bar{\kappa}$)	1.138	0.011	2.059	0.073	-9.178	0.252
	Panel D - Agency Objective & Screening Precision					
	Estimate	**S.E**		**Quantiles**		
Weight on VA Relative to π				**25%**	**50%**	**75%**
Constant ($\bar{\tau}$)	0.268	0.003				
Capital Fund (τ_{fund})	0.013	0.002				
Values of τ_j				0.353	0.396	0.423
Soft Info Screening Precision						
Large Loan Program $\frac{1}{(\sigma_{large})^2}$	0.099	0.003				
Small Loan Program $\frac{1}{(\sigma_{small})^2}$	0.025	0.001				

Notes: Panel A displays estimates for the repayment type, and Panel B displays estimates of the borrowers' investment technology from the guarantee rate choice equation. Panel C shows estimates of the funding probability for 85% guarantee rate. Panel D displays estimates from the agency side of the model. The quantiles of τ_j represent weighted distributions based on the number of borrowers per region. Standard errors based on the inverse of the numerical hessian of the log-likelihood function.

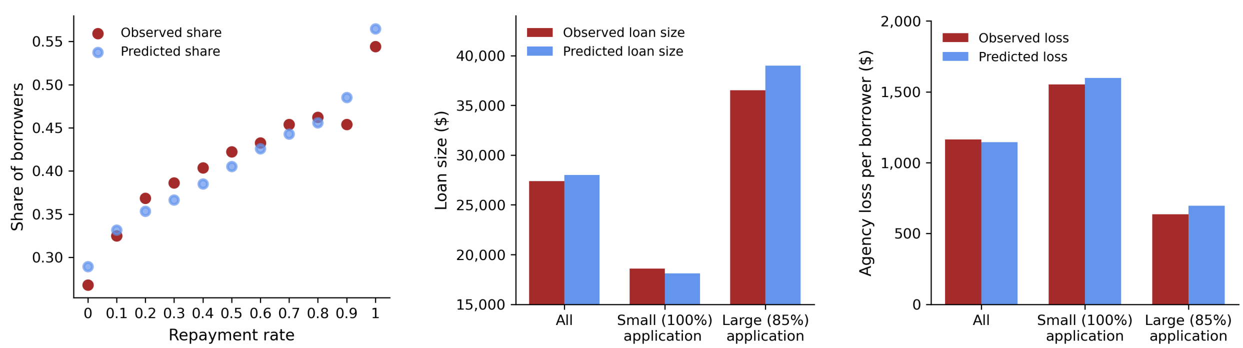
repayment type borrowers more likely to be rejected. The cost of being rejected, captured by the parameter c, is valued at \$6,513. This cost may be overestimated, as the dataset primarily captures explicit rejections and does not consider preemptive non-applications by borrowers who anticipate rejection at the 85% guarantee rate after informal discussions with lenders. This omission could lead to an underestimation of the true rejection rate. The underestimation could inflate the perceived cost of rejection, as the cost, multiplied by the rejection rate, must justify borrowers' reluctance to opt for the large loan program. Therefore, the estimated rejection rates require careful interpretation in the counterfactual analysis, which will be further discussed in the next section.

Agency objective and soft information screening precision The estimates from Table 3-D imply that agencies assign an average relative weight (τ_j) of 39% to the value-added generated by small businesses and the remaining 61% to their own profit. The weighting varies among regional agencies and tends to increase with the level of the agencies' capital funds relative to the number of guarantees they are backing. This suggests that agencies with larger capital funds (relative to the number of guarantees) are better positioned to prioritize the value-added of small businesses, while those with smaller capital funds must focus more on reducing financial losses, likely due to tighter budget constraints.

The final parameter of interest is the screening precision of guarantee agencies via soft information collection. The precision for the large loan program is estimated to be four times higher than for the small loan program, consistent with institutional practices where agencies perform more detailed evaluations for the large loan program. As the numerical values for screening precision are difficult to interpret directly, they will be further explored in counterfactual analyses. These analyses will show how the screening precision influences the loan size guaranteed to different borrowers, in turn affecting market outcomes such as the agency's losses from guarantees and the value generated by small businesses.

Figure 3: Model fit: average outcome

(a) Share of "large" program (b) Loan size ($) (c) Agency loss ($)



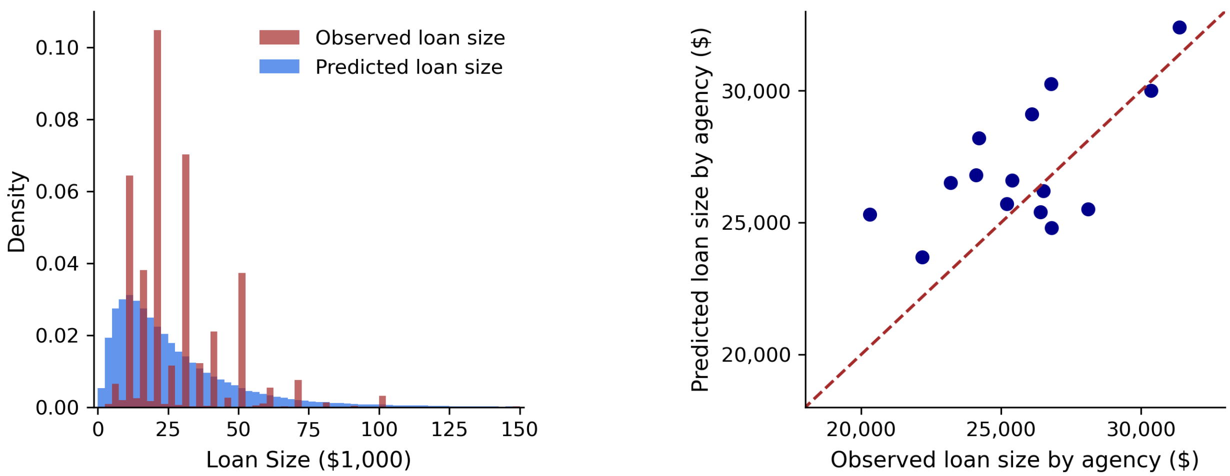
Model fit I examine model fit by using the estimated parameters to simulate equilibrium outcomes and compare simulated to observed outcomes. I describe the simulation procedure in the Online Appendix H. Figure 3a shows a close match between simulated and observed shares of borrowers opting for the large loan program, conditional on the repayment rate. In terms of loan size, Figure 3b shows that, while the model accurately predicts average loan sizes for the small loan program, the prediction for the large loan program is approximately \$2,300 higher than observed. Figure 3c shows that the simulated average agency loss per borrower (\$1,146) closely matches the observed average in the data (\$1,165).

Figure 4a illustrates the distribution of loan sizes, contrasting the continuous predictions of our model against the observed discrete bunching at \$10,000, \$20,000, and so forth. This highlights our model's abstraction from certain real-world financial practices. Figure 4b compares predicted and observed average loan sizes across regional agencies. It demonstrates that, despite lacking agency-specific fixed effects and not aligning perfectly on an agency-by-agency basis, our model effectively captures the general trend by modeling the preference parameter τ_j based on agencies' budgets.

Figure 4: Model fit: loan size distribution

(a) Loan size distribution

(b) Average loan size across regional agencies



7 Counterfactuals

To evaluate the benefits of employing a loan guarantee menu alongside information collection, this section outlines three counterfactual experiments: (i) the agency offers a uniform guarantee program while continuing to collect soft information; (ii) the agency maintains the loan guarantee menu but stops soft information collection; and (iii) the agency offers a uniform program and also stops soft information collection. These scenarios are compared against the status quo, in which the agency utilizes both the loan guarantee menu and soft information collection. In the status quo, the average value-added per borrower is $9,398, while the average loss per agency per borrower is $1,146, resulting in an average value of the agency objective function of $2,935 per borrower. Notably, the primary focus for scenario (iii) is its comparison with scenario (ii) to specifically isolate the effect of offering a loan guarantee menu versus uniform program under no soft information collection.

Table 4 presents the findings from three counterfactual experiments, detailing average loan sizes by guarantee rate and comparing crucial outcomes per borrower against the status quo. The analysis focuses on the agency's average financial loss per borrower, the economic contribution (VA) generated by the small businesses, and the agency objective.

Table 4: Outcomes Under Counterfactual Policies

Policy	**Avg loan size**		**Outcome**		
	Small (100%)	**Large (85%)**	**Loss**	**VA**	**Hybrid obj**
Status quo (menu + soft info)	$18,120	$36,750	$1,146	$9,398	$2,935
(i) Uniform program + soft info	$26,456	-	+$139 (+12.1%)	-$385 (-4.1%)	-$235 (-8.0%)
(ii) Menu + no soft info	$22,568	$29,801	+$307 (+26.8%)	-$493 (-5.2%)	-$376 (-12.8%)
(ii-a) Fix borrower's choice	$17,582	$35,423	+$197 (+17.2%)	-$182 (-1.9%)	-$190 (-6.5%)
(iii) Uniform program + no soft info	$26,371	-	+$431 (+37.7%)	-$542 (-5.8%)	-$471 (-16.1%)

Notes: The average loan size columns display the average funded loan size. The Small (100%)" and Large (85%)" columns reflect data based on borrowers' initial guarantee choices. For borrowers initially choosing the large loan program with an 85% guarantee rate, if they are rejected, they subsequently receive a 100% guarantee rate with the small loan program. The average loan size for these cases is calculated by accounting for the subsequent 100% guarantee rate loans. The outcome columns on the right are calculated as an average per borrower. The Status quo" row refers to the baseline, and the outcomes in the three counterfactual scenarios below show differences compared to the status quo. Parentheses indicate the percentage change compared to the baseline.

7.1 Uniform Guarantee Program with Soft Information Collection

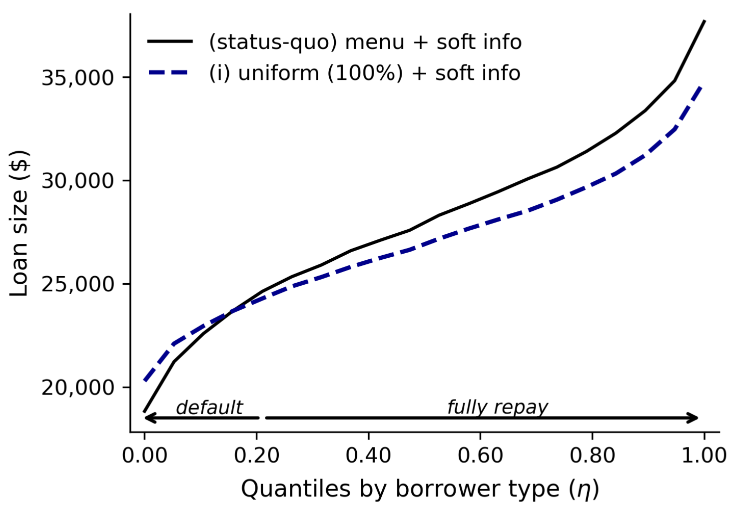
Counterfactual scenario (i) explores the impact of replacing the loan guarantee menu with a uniform program, while continuing soft information collection. Given that agencies can no longer adjust screening precision for soft information collection based on borrowers' guarantee choices, I retain σ_{large}, the higher precision level, to establish an upper boundary for the agency welfare. This ensures a conservative assessment of the impact of eliminating the loan guarantee menu.[21]

[21] Using σ_{small} as the precision level, characterized by lower precision, would worsen outcomes, making it challenging to discern whether changes are due to the uniform program or diminished screening precision in soft information collection.

To search for the most plausible guarantee rate associated with the uniform program, I conduct a simulation exercise evaluating the agency objective function across uniform rates from 85% to 100%. Higher guarantee rates increase the value added by small businesses by enhancing credit access, but lead to greater agency losses due to increased default risks. The rejection rates between the known 85% and 100% guarantees are linearly interpolated to estimate intermediate values. The results detailed in Online Appendix I suggest minimal variations in the value of the agency objective function across the rates, with differences amounting to less than $35 per borrower. The results suggest the optimal rate is 96%, and the 100% guarantee rate closely approximates this maximum value of the agency objective function. For simplicity and clarity, the 100% rate is used as the primary counterfactual in the main text.[22]

The welfare implications of switching from the loan guarantee menu to a uniform program with a 100% guarantee rate are reported in Table 4, focusing on the agency's perspective regarding agency's financial losses and the value generated by the small businesses. Under this policy, the agency's loss per borrower increases by $139 (12.1%), and the value added by businesses declines by $385 (4.1%), resulting in an 8% reduction in the value of the agency objective function.

Figure 5: Average loan size in uniform program (100% guarantee)



The reduction in agency objective function is attributed to changes in loan size distribution under the uniform guarantee policy. As shown in Figure 5, the uniform program pools all borrowers

[22]It should be noted that if the rejection rate for the 85% guarantee rate has been underestimated, as previously discussed, the outcomes for employing any uniform guarantee rate below 100% could be worse, since more borrowers might end up unfunded. Furthermore, under a uniform 85% guarantee program, riskier borrowers are likely to be pooled together, which could prompt agencies to adjust their rejection practices to manage increased risk. This scenario is not accounted for in my estimates, leading to potential bias. The 100% guarantee rate, where no rejections occur, thereby becomes a more suitable and reliable choice for the main counterfactual analysis.

together, which limits the agency's ability to differentiate loan sizes based on a borrower's type. This pooling effect results in a reduction in loan size for high-type borrowers, who benefit from larger loans under the large loan program with an 85% guarantee rate. In the uniform program, these high-type borrowers generate less value-added due to receiving smaller loans. Conversely, low-type borrowers, who default, now receive larger loans by being pooled with higher types, leading to greater losses for the agency.

This analysis shows the benefit of employing a loan guarantee menu, a strategy that is becoming increasingly common in many countries. Programs with a menu enable high type borrowers to secure larger loans that maximize their economic contribution. At the same time, these programs ensure that lower type borrowers, though contributing on a smaller scale, still gain access to necessary credit but with reduced loan sizes. Employing a menu of rates enables guarantee agencies to maximize the value generated by small businesses, while simultaneously managing their own risk exposure to ensure the program's sustainability.

7.2 Loan Guarantee Menu without Soft Information Collection

In counterfactual policy (ii), I explore the effects of discontinuing soft information collection while maintaining a loan guarantee menu. The agencies still observe basic borrower characteristics (X) such as credit score, business age, and industry: data that are readily available at the application stage. However, it stops collecting costly soft information like revenue projections and assessments of borrower's trustworthiness, typically obtained through site visits and interviews (i.e. $\sigma_G \to \infty$).[23]

The elimination of such information collection while maintaining a loan guarantee menu leads to an increase of \$307 (26.8%) in agency losses per borrower and a decrease of \$493 (5.2%) in value added by businesses. These changes result in a 12.8% reduction in the value of the agency objective function, as detailed in Table 4.

This reduction in the agency objective function can be attributed to two effects. First, without soft information collection, the agency cannot effectively differentiate between borrowers based on repayment types, which influences loan sizes conditional on the guarantee choice; this is what I call the "direct information effect". Second, borrowers adjust their expectations regarding loan sizes associated with each guarantee choice, prompting them to reoptimize their guarantee choices. This adjustment alters the self-selection based on repayment type and impacts their loan sizes.; I refer

[23]This scenario represents a realistic change in information collection practices. A complete cessation would result in more severe consequences.

to this as the “sorting effect”. Figure 6a illustrates how the share of borrowers opting for the large guarantee program changes in the absence of soft information collection. The “menu + info” line represents the share under the status quo, whereas the “menu + no info” line displays the shares with no soft information collection. The noticeably flatter slope indicates reduced sorting among borrowers due to the absence of soft information collection. The share of those who choose the large loan program among high repayment type borrowers has decreased, while it has increased among low repayment type borrowers, indicating a diminished self-selection based on repayment type.

The diminished sorting effect is primarily due to eliminating the difference in screening precision from soft information collection across the two guarantee choices. Under the status quo, the difference influences borrower choices: high type borrowers opt for the large loan program to capitalize on higher precision, revealing their true type and thereby securing larger loans, while low type borrowers choose the small loan program to obscure their true type. In the absence of soft information collection these incentives no longer exist, resulting in less pronounced sorting among borrowers. This logic also applies when the screening precision for the small loan program and the large loan program is set equally, either by enhancing the precision of the small loan program or reducing that of the large loan program. As demonstrated in Online Appendix J, making the screening precision uniform across guarantee choices similarly reduces sorting, confirming that the difference in screening precision between the small and large loan program is the primary factor influencing the borrower sorting. Figure 6b shows the distribution of loan sizes under both the small and large loan program, comparing scenarios without soft information collection to the status quo. With soft information collection, the differences in loan sizes are notably larger between high and low type borrowers. This is especially true in the large loan program, due to higher screening precision from soft information collection. Without soft information collection, loan sizes become more uniform across each guarantee choice, illustrating a “direct information effect” where the agency cannot differentiate between borrower types. It is noteworthy that the gap between loan sizes for the small and large loan program narrows due to the reduced sorting effect. In the status quo, high type borrowers typically select the large loan program, helping them secure larger loans due to the significant information the agency derives from their guarantee choice. Without soft information collection the choice becomes less informative, leading to less disparity in loan size between the guarantee choices.

Figure 6: Loan guarantee menu without soft information collection

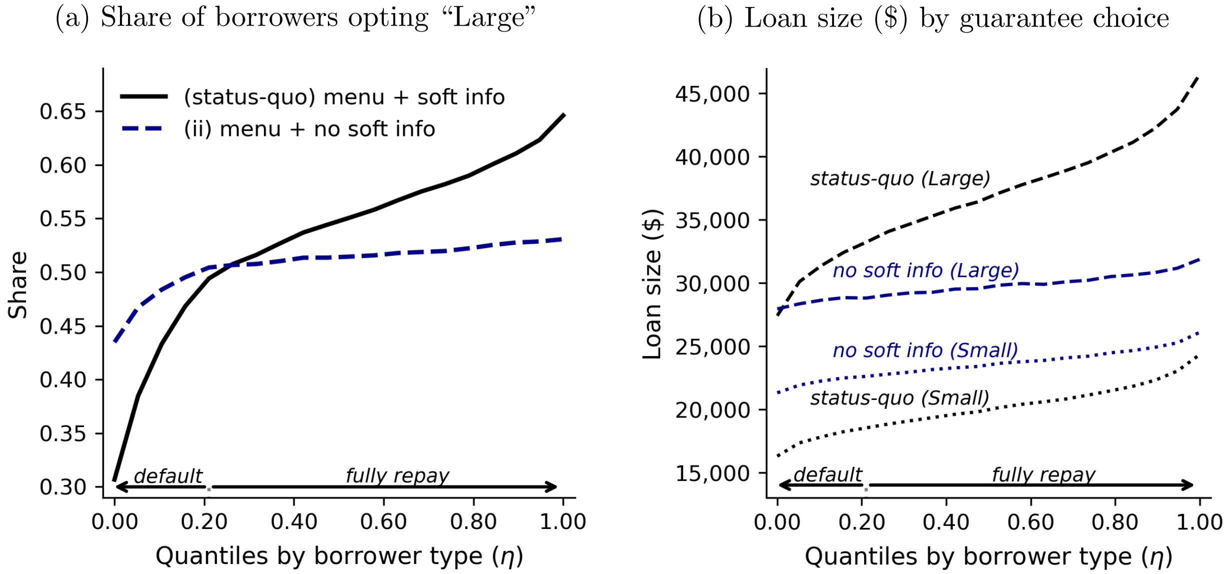
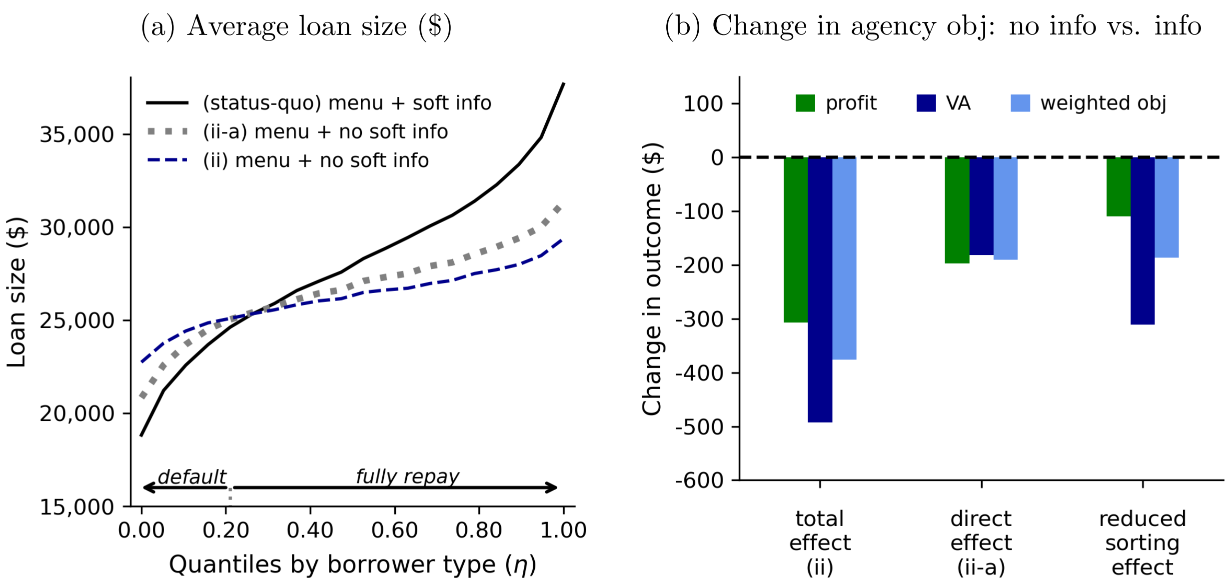


Figure 7: Decomposition of the effect from absence of information collection



Decomposition of reduced sorting effect To isolate the "reduced sorting effect" from the "direct information effect", I conduct a decomposition exercise. In counterfactual scenario (ii-a), I simulate the loan allocations while keeping borrowers' guarantee choices the same as in the status quo. Figure 7a depicts the loan size distribution for these scenarios. "Menu + info" corresponds to the status quo loan allocation, "(ii) menu + no info" to the allocation where borrowers re-optimize their guarantee choice in the absence of soft information collection, and "(ii-a) menu + no info" to the allocation where borrowers maintain the guarantee choices. The figure demonstrates that the differentiation in average loan sizes between high and low repayment type borrowers is reduced in scenario (ii-a) compared to the status quo, and diminishes further in scenario (ii) due to the effect of reduced self-selection. Regarding the outcomes for scenario (ii-a), the agency's losses increase by \$197 and value-added decreases by \$182, resulting in an 6.5% reduction in the value of the agency objective function. This outcome represents the direct effect of losing information. The additional 6.3% decline in the agency objective function, when moving from scenario (ii-a) to (ii), is attributed to the reduced sorting effect within the loan guarantee menu. This decomposition is displayed in Figure 7b. It shows that approximately 50% of the total effect can be attributed to the sorting effect.

Reduced effectiveness of loan guarantee menu under no soft information collection The comparison of counterfactual scenarios (iii) and (ii) quantifies the diminished effect of the loan guarantee menu in the absence of soft information collection. Transitioning from (iii) a uniform program to (ii) the loan guarantee menu results in only a \$95 (3.9%) increase in the value of the agency objective function. This increase is significantly smaller than the \$235 (8.7%) improvement when the loan guarantee menu is employed alongside soft information collection (from counterfactual (i) to the status quo). These findings underscore that the effectiveness of loan guarantee menu is greatly enhanced when used in conjunction with soft information collection.

This analysis presents a new perspective for countries that adopt different approaches to loan guarantee programs, such as the SBA Advantage Loan Program, the primary loan guarantee scheme in the US. While this program does offer a loan guarantee menu, it employs a uniform maximum loan size for each option within the menu, without adjusting for individual borrower characteristics based on additional information collection. Consequently, the responsibility for assessing borrower risk is largely delegated to lenders. The program could potentially overlook the benefits of enabling borrowers to self-select into different menu options through adjusted loan sizes based on soft information collection, which could enhance the effectiveness of loan guarantee schemes.

8 Conclusion

Government-backed loan guarantee programs are crucial for facilitating access to credit for small businesses. The effectiveness of these programs is significantly influenced by how the guaranteed loans are allocated, with the aim of maximizing economic benefits for small businesses while maintaining the financial soundness of the program.

I analyze data from the South Korean loan guarantee program, demonstrating how using a loan guarantee menu, along with soft information collection, significantly enhances the allocation of loans. I find that a loan guarantee menu, when accompanied by soft information collection, leads to significant sorting among borrowers.

However, the effectiveness of a loan guarantee menu diminishes in the absence of soft information collection. Without soft information collection, the sorting effect is notably reduced, highlighting the critical role of information collection in enabling borrowers to self-select into different guarantee options and thereby ensuring a more appropriate loan allocation.

These findings underscore the critical roles of both soft information collection and the strategic use of loan guarantee menus as complementary methods to enhance the efficiency of loan guarantee programs. While directly applicable to government loan guarantee schemes for small businesses, the implications of this research can extend to the broader financial sector, such as in mortgage lending. In mortgage markets, lenders not only offer a variety of contract options to elicit self-revelation of borrowers' risk levels, but also actively collect detailed information about a borrower's income, employment history, assets, and more. My findings highlight the potential complementarity of these two screening mechanisms, and the framework I present here can be applied in similar markets to evaluate the effects.

References

W. Adams, L. Einav, and J. Levin. Liquidity constraints and imperfect information in subprime lending. *American Economic Review*, 99(1):49–84, 2009.

S. Agarwal, B. Ambrose, S. Chomsisengphet, and C. Liu. The role of soft information in a dynamic contract setting: Evidence from the home equity credit market. *Journal of Money, Credit and Banking*, 43(4):633–655, 2011.

T. Babina, S. Bahaj, G. Buchak, F. De Marco, A. Foulis, W. Gornall, F. Mazzola, and T. Yu. Customer data access and fintech entry: Early evidence from open banking. *NBER Working Paper*, 2024.

N. Bachas, O. Kim, and C. Yannelis. Loan guarantees and credit supply. *Journal of Financial Economics*, 139(3):872–894, 2021.

J. Barrot, T. Martin, J. Sauvagnat, and B. Vallee. The labor market effects of loan guarantee programs. *The Review of Financial Studies*, 2024.

T. Beck, L. Klapper, and J. Mendoza. The typology of partial credit guarantee funds around the world. *Journal of Financial Stability*, 6:10–25, 2010.

J. Brown and J. Earle. Finance and growth at the firm level: Evidence from sba loans. *Journal of Finance*, 72(3):1039–1080, 2017.

G. Bryan, D. Karlan, and A. Osman. Big loans to small businesses: Predicting winners and losers in an entrepreneurial lending experiment. *American Economic Review*, 114(9):2825–2860, 2024.

B. Choi and H. Lee. Heterogeneous elasticities of bank loan supply: Evidence from sba loan guarantee expansion. *Working Paper*, 2019.

F. Columba, L. Gambacorta, and P. Mistrulli. Mutual guarantee institutions and small business finance. *Journal of Financial Stability*, 6(1):45–54, 2010.

M. Cowling. The role of loan guarantee schemes in alleviating credit rationing in the uk. *Journal of Financial Stability*, 6(1):36–44, 2010.

N. Cox, E. Liu, and D. Morrison. Market power in small business lending: A two-dimensional bunching approach. *Working Paper*, 2021.

G. Crawford, N. Pavanini, and F. Schivardi. Asymmetric information and imperfect competition in lending markets. *American Economic Review*, 108(7):1659–1701, 2018.

L. Deelen and K. Molenaar. *Guarantee Funds for Small Enterprises: A Manual for Guarantee Fund Managers*. International Labour Organization (ILO), 2004.

L. Einav, M. Jenkins, and J. Levin. Contract pricing in consumer credit markets. *Econometrica*, 80(4):1387–1432, 2012.

S. Greenbaum and A. Thakor. *Contemporary Financial Intermediation*. The Dryden Press, New York, 1995.

A. Hertzberg, A. Liberman, and D. Paravisini. Screening on loan terms: Evidence from maturity choice in consumer credit. *The Review of Financial Studies*, 31(9):3532–3567, 2018.

V. Ioannidou, N. Pavanini, and Y. Peng. Collateral and asymmetric information in lending markets. *Journal of Financial Economics*, 144:93–121, 2022.

K. Kawai, K. Onishi, and K. Uetake. Signaling in online credit markets. *Journal of Political Economy*, 130(6):1585–1629, 2022.

C. Kuo, C. Chen, and C. Sung. Evaluating guarantee fees for loans to small and medium-sized enterprises. *Small Business Economics*, 37:205–218, 2011.

J. Liberti and M. Petersen. Information: Hard and soft. *The Review of Corporate Finance Studies*, 8(1):1–41, 2019.

G. Mailath. Incentive compatibility in signaling games with a continuum of types. *Econometrica*, 55(6):1349–1365, 1987.

W. Mullins and P. Toro. Credit guarantees and new bank relationships. *Working Paper*, 2018.

OECD. Evaluating publicly supported credit guarantee programmes for smes. 2017.

I. Oh, J. Lee, A. Heshmati, and G. Choi. Evaluation of credit guarantee policy using propensity score matching. *Small Business Economics*, 33:335–351, 2009.

F. Panetta, F. Schivardi, and M. Shum. Do mergers improve information? evidence from the loan market. *Journal of Money, Credit, and Banking*, 41(4):673–709, 2009.

I. Panetta. An analysis of credit guarantee schemes: Suggestions provided by literature. In P. Leone and G. Vento, editors, *Credit Guarantee Institutions and SME Finance*, chapter 2, pages 11–37. Palgrave Macmillan, London, 2012.

A. Saito and D. Tsuruta. Information asymmetry in small and medium enterprise credit guarantee schemes: Evidence from japan. *Applied Economics*, 50(22):2469–2485, 2018.

J. Stiglitz and A. Weiss. Credit rationing in markets with imperfect information. *American Economic Review*, 71(3):393–410, 1981.

D. Stillerman. Loan guarantees and incentives for information acquisition. *Working Paper*, 2022.

A. Taburet, A. Polo, and Q.-A. Vo. Screening using a menu of contracts: A structural model for lending markets. *Working Paper*, 2024.

C. Timmins. Measuring the dynamic efficiency costs of regulators' preferences: Municipal water utilities in the arid west. *Econometrica*, 70(2):603–629, 2002.

J. Wang. Screening soft information: Evidence from loan officers. *RAND Journal of Economics*, 51(4):1287–1322, 2020.

World Bank. Map of sme-support measures, 2020.

World Bank. Ifc and small and medium enterprises. 2024.

Online Appendix

A Conceptual Framework

In this section, I present a simplified framework to illustrate how borrowers choose from a menu of loan guarantee contracts within loan guarantee programs, with a particular focus on how their self-selection is influenced by the practice of additional information collection. Each contract offers different combinations of loan sizes and guarantee rates. I demonstrate how varying levels of information collection induce borrowers of different types to choose distinct contracts—some opting for the large loan program with a lower (85%) guarantee rate and others choosing the small loan program with a higher (100%) guarantee rate. Since there is a continuum of borrowers, no individual borrower can affect the menu of contracts set by the guarantee agency. For simplicity, I assume that the menu of contracts is exogenously given, reflecting the agency's objectives: by making this assumption I can focus exclusively on borrower behavior and clearly present the underlying concepts graphically. In the empirical model section, I endogenize the loan sizes associated with each guarantee rate, with the agency setting the loan sizes to maximize its objective.

A.1 Borrower's Demand for Guarantee Contract

A borrower seeks a loan to invest in her small business. There is a continuum of borrower types, denoted by $\eta \in [\eta_{\min}, \eta_{\max}]$, where η represents the borrower's private information about her productivity and default risk. Higher values of η indicate higher productivity and a lower likelihood of default. Consider two types from this continuum: a "high" type η^h and a "low" type η^l, with $0 < \eta^l < \eta^h$. The types are privately known to the borrowers but not directly observable by the guarantee agency or lenders.

A borrower chooses a loan guarantee contract (L, g) from the agency's menu, where L is the loan size and g is the guarantee rate. The agency may collect additional information about the borrower and receive a signal that could influence the adjustment of the final loan size L. Then, the lender evaluates the guarantee contract and decides whether to fund the loan, and at what interest rate $r(g, \eta)$. If the loan is funded, the borrower invests the money from the loan, produces output and repays a fraction $\lambda(\eta)$ of the principal and interest payments.

A "high" type borrower(η^h), who poses a lower default risk, receives a higher lender funding probability $P^F(g, \eta)$ and a more favorable interest rate $r(g, \eta)$ compared to a "low" type borrower(η^l), for any given guarantee rate g:

$$P^F(g,\eta^h) > P^F(g,\eta^l) \quad \text{and} \quad r(g,\eta^h) < r(g,\eta^l) \quad ; \quad \forall g \in [0,1]$$

Higher guarantee rates reduce the lender's exposure to borrower default, leading to an increase in funding probability and a decrease in interest rates. Moreover, higher guarantee rates disproportionately benefit "low" type borrowers, who have higher default risks. That is, the marginal impact of an increase in the guarantee rate is greater for "low" type borrowers:

$$\frac{\partial P^F(g,\eta^l)}{\partial g} > \frac{\partial P^F(g,\eta^h)}{\partial g} > 0 \quad \text{and} \quad \frac{\partial r(g,\eta^l)}{\partial g} < \frac{\partial r(g,\eta^h)}{\partial g} < 0 \quad ; \quad \forall g \in [0,1]$$

This illustrates that increases in the guarantee rate have a more significant effect on improving funding probabilities and lowering interest rates for borrowers with higher default risks (Stillerman [2022]).[24]

The borrower uses a production technology to generate output based on the loan size L:

$$F(L) = A(\eta) \cdot L^\alpha$$

where $A(\eta)$ is a technology shifter that increases with the borrower's type η, such that $A(\eta^h) > A(\eta^l)$. The parameter α represents the concavity of the production function. The borrower repays a fraction $\lambda(\eta)$ of the principal and interest payments on the loan, known as the repayment rate. Higher-type borrowers have higher repayment rates, i.e., $\lambda(\eta^h) > \lambda(\eta^l)$. It is assumed that the increase in productivity outweighs or is proportionate to the increase in repayment rates, such that $\frac{A(\eta^h)}{\lambda(\eta^h)} \geq \frac{A(\eta^l)}{\lambda(\eta^l)}$.

The expected utility of a borrower of type η who chooses a guarantee contract (L,g) is:

$$U(L,g) = P^F(g,\eta) \cdot \Big[A(\eta) \cdot L^\alpha - \lambda(\eta) \cdot [1 + r(g,\eta)] \cdot L\Big]$$

If the loan is funded, the borrower obtains the net benefit of production minus repayment. If the loan is not funded, the utility is zero. The expected utility increases with both the loan size L and the guarantee rate g, as a larger loan size increases output and a higher guarantee rate enhances the probability of funding while reducing the interest rate.

[24] For illustration, consider a "high" type borrower who is likely to repay the loan in full. For this borrower, the guarantee rate has little impact on the lender's risk assessment because the likelihood of default is low. Conversely, a "low" type borrower presents a higher risk of default. Increasing the guarantee rate substantially reduces the lender's potential loss from this borrower, significantly improving the funding probability and reducing the interest rate.

Given these considerations, we can establish the following lemma regarding borrowers' preferences:

Lemma 1 *The marginal rate of substitution of loan size for guarantee rate, defined as* $|MRS_{L,g}| = \frac{MU_L}{MU_g}$*, is steeper for the "high" type* (η^h) *than for the "low" type* (η^l)*, i.e.,* $|MRS^h_{L,g}| > |MRS^l_{L,g}|$

Proof. See Online Appendix B

Intuition: The "high" type borrower values additional loan size more relative to an increase in the guarantee rate because she has higher productivity and lower default risk. Therefore, she is willing to trade off a higher guarantee rate in exchange for a larger loan size. In contrast, the "low" type borrower places a higher value on an increased guarantee rate to improve her funding probability and reduce the interest rate, given her higher default risk. This difference in their marginal rates of substitution influences borrowers to self-select into different contracts based on their risk profiles, potentially facilitating a separation between high and low-risk types in their contract choices.

A.2 Graphical Analysis

Figure A.1 illustrates the indifference curves of the "high" type ($U_h(L, g)$) and the "low" type ($U_l(L, g)$) borrowers in the space of loan size L (horizontal axis) and guarantee rate g (vertical axis). In these figures, the indifference curves for the "high" type borrower are steeper, indicating a higher willingness to substitute guarantee rate for loan size. The "low" type borrower's indifference curves are flatter, reflecting a stronger preference for higher guarantee rates due to higher default risk.[25]

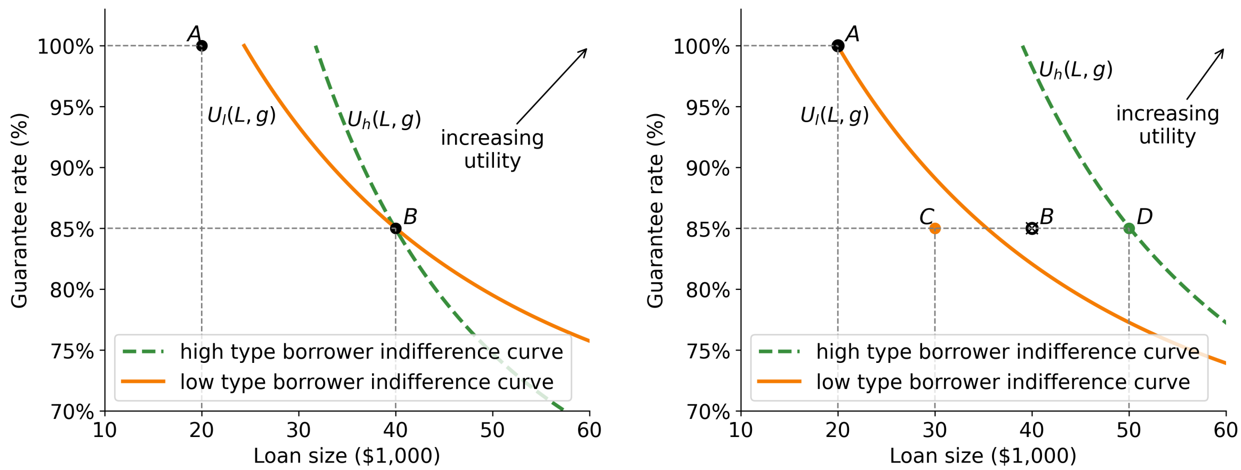
The guarantee agency sets the menu of loan guarantee contracts to maximize its objectives across the entire continuum of borrower types, meaning it does not adjust the menu to accommodate or separate any specific types. Therefore, from the perspective of individual borrowers, the menu is treated as given, with no single borrower able to influence it. Here I provide graphical examples to illustrate how borrowers can be sorted into different contract options. These examples show how the two borrower types choose between the large loan program with a lower (85%) guarantee rate and the small loan program with a higher (100%) guarantee rate under specific menu conditions, demonstrating the potential sorting outcomes.

In Figure A.1a, I consider a scenario where the guarantee agency does not collect additional

[25] The utility functions reflect non-homothetic preferences. As these curves shift rightward—indicating increased utility levels—the marginal rate of substitution decreases. This change signifies a diminishing willingness to substitute loan size for guarantee rate, influenced by the borrowers' concave production functions. As loan size increases, the marginal productivity and thus the incremental utility derived from additional loan amounts diminish, making higher guarantee rates relatively more valuable.

Figure A.1: Conceptual framework of contract menu

(a) No separation without information collection (b) Separation with information collection



information about borrowers. In this setting, the agency must offer the same menu of contracts to all types, as it cannot distinguish between them. Suppose the agency offers an equilibrium menu consisting of two contracts: point A, which represents a smaller loan size with a higher guarantee rate ($\bar{L}_{small}$, 100%), and point B, which represents a larger loan size with a lower guarantee rate ($\bar{L}_{large}$, 85%).[26] These points reflect the maximum loan sizes for each guarantee rate that the agency is willing to offer when it considers borrowers as an average risk type, without distinguishing between "high" and "low" type borrowers. In this specific example, focusing on our two borrower types—the "high" type and the "low" type—both prefer the large loan at point B over the small loan at point A. Therefore, in this scenario, both types choose the same contract, and there is no separation based on their types.[27]

In contrast, Figure A.1b illustrates a scenario where the same two borrower types from the previous example show separation in their choices when the agency collects additional information to determine the final loan size. The agency sets different screening precisions for each contract option: the large loan program with a lower (85%) guarantee rate involves detailed evaluations, leading to informative signals about the borrower's type, while the small loan program with a

[26] The contract menu options represented by points A and B in this figure, and points C and D in Figure A.1b, are chosen to provide a simplified framework to illustrate potential borrower sorting effects. These points ideally reflect the agency's objective function, which is to balance support for small businesses with the agency's financial sustainability. However, in this conceptual framework, they are selected based on plausible assumptions rather than derived from a detailed model of the agency's strategy. The modeling and estimation of the agency's objective function are thoroughly incorporated into the empirical model discussed in the main text of this paper.

[27] Note that the scenario described here involves specific examples to illustrate the potential effects of additional information collection. However, it's important to recognize that with different borrower types or contract options, separation might still occur even without additional information collection.

higher (100%) guarantee rate involves simplified evaluations, resulting in uninformative signals. For illustrative purposes, consider that for the large loan program, the agency obtains an informative signal (i.e., a high type signal s^h_{large} such that $Pr(\eta^h|s^h_{large}) > Pr(\eta^h)$). Conversely, for the small loan program, the agency receives an uninformative signal s^h_{small} (i.e., a high type signal s^h_{small} such that $Pr(\eta^h|s^h_{small}) = Pr(\eta^h)$).[28] The agency adjusts the final loan sizes based on these signals.

- **High type borrower**: For the large loan program, the expected loan size increases from point B to point D due to the informative signal from detailed screening, which likely reveals her favorable characteristics. The loan size for the small loan program remains at point A, as the uninformative signal does not alter the agency's initial assessment.

- **Low type borrower**: For the large loan program, the expected loan size decreases from point B to point C as the detailed screening exposes her higher risk profiles. The loan size for the small loan program remains at point A, unchanged due to the uninformative signal.

The difference in screening precision for each loan program leads to distinct guarantee choices between the two borrowers: the high type borrower opts for the large loan program at point D, where detailed screening is more likely to accurately reveal their high type status and reward them with an even larger loan. Conversely, the low type borrower favors the small loan program at point A, where the utility is higher compared to point C, as less intensive screening helps them to hide their low type status.

It is important to note that while these examples focus on specific borrower types, separation between borrower types can indeed occur even without additional information collection, depending on the shapes of their indifference curves and the agency's menu of contracts. However, combining a menu with information collection can further facilitate and enhance this separation. The key empirical question is how much this combination increases the effectiveness of screening and improves loan allocation outcomes according to borrower types.

This conceptual framework highlights the main forces behind screening using a menu of loan guarantee contracts and differential information collection. All borrowers prefer larger loan sizes and higher guarantee rates, but the menu presents a trade-off: larger loan sizes come with lower guarantee rates and more intensive screening, and vice versa. Different borrowers evaluate this trade-off differently based on their risk profiles, leading to potential separation through self-selection. The

[28]The observed disparity in signal precision between the small and large loan programs is consistent with empirical findings, which are further discussed and analyzed in the empirical model and results sections of this paper.

agency's strategy of varying screening precision across contracts facilitates this separation, enhancing its ability to allocate loan sizes more effectively according to borrower types.

My model in section 4 incorporates the guarantee agency's objectives and endogenizes the loan sizes associated with each guarantee rate. It also accounts for the varying screening precisions across contracts, allowing for a comprehensive analysis of how these factors influence borrower sorting and loan allocation outcomes.

B Marginal Rate of Substitution of Loan Size for Guarantee Rate

This subsection demonstrates that the marginal rate of substitution of loan size for guarantee rate ($MRS_{L,g}$) is steeper for the "high" type borrower (η^h) compared to the "low" type borrower (η^l).The expected utility of a borrower who chooses a guarantee contract with a particular guarantee rate, g, and loan size, L as defined in the Online Appendix A, is given by:

$$U(L,g) = P^F(g,\eta) \cdot \Big[A(\eta) \cdot L^\alpha - \lambda(\eta) \cdot [1 + r(g,\eta)] \cdot L\Big]$$

From this utility function, the $MRS_{L,g}$ for a borrower type η can be expressed as:

$$|MRS^\eta_{L,g}| = \frac{MU_L}{MU_g} = \frac{\Big[\alpha \cdot \frac{A(\eta)}{\lambda(\eta)} \cdot L^{\alpha-1} - (1 + r(g,\eta))\Big]}{\frac{\frac{\partial P^F(g,\eta)}{\partial g}}{P^F(g,\eta)} \cdot \Big[\frac{A(\eta)}{\lambda(\eta)} \cdot L^\alpha - (1 + r(g,\eta)) \cdot L\Big] - \frac{\partial r(g,\eta)}{\partial g} \cdot L}$$

The goal is to establish the relationship:

$$|MRS^h_{L,g}| > |MRS^l_{L,g}|$$

I proceed by first taking the reciprocal of both sides of the inequality, which reverses the direction of the inequality. Following this, both sides are multiplied by L^{-1}, which is a positive quantity and thus preserves the direction of the inequality. This transformation yields the revised inequality:

$$\frac{1}{|MRS^h_{L,g}| \cdot L} < \frac{1}{|MRS^l_{L,g}| \cdot L}$$

where $\frac{1}{|MRS_{L,g}| \times L}$ for each borrower type η is expressed as:

$$\frac{1}{|MRS^\eta_{L,g}| \times L} = \frac{\frac{\frac{\partial P^F(g,\eta)}{\partial g}}{P^F(g,\eta)} \cdot \Big[\frac{A(\eta)}{\lambda(\eta)} \cdot L^\alpha - (1 + r(g,\eta)) \cdot L\Big] - \frac{\partial r(g,\eta)}{\partial g} \cdot L}{\Big[\alpha \cdot \frac{A(\eta)}{\lambda(\eta)} \cdot L^{\alpha-1} - (1 + r(g,\eta))\Big]}$$

This expression is evaluated in three parts. Each part will confirm the relationship established in the inequality.

1. $\frac{\frac{\partial P^F(g,\eta^H)}{\partial g}}{P^F(g,\eta^H)} < \frac{\frac{\partial P^F(g,\eta^L)}{\partial g}}{P^F(g,\eta^L)}$, as the elasticity of funding probability with respect to guarantee rate is larger for the "low" type borrowers.

2. $\frac{\left[\frac{A(\eta^H)}{\lambda(\eta^H)} \cdot L^{\alpha-1} - (1+r(g,\eta^H))\right]}{\left[\alpha \cdot \frac{A(\eta^H)}{\lambda(\eta^H)} \cdot L^{\alpha-1} - (1+r(g,\eta^H)\right]} < \frac{\left[\frac{A(\eta^L)}{\lambda(\eta^L)} \cdot L^{\alpha-1} - (1+r(g,\eta^L))\right]}{\left[\alpha \cdot \frac{A(\eta^L)}{\lambda(\eta^L)} \cdot L^{\alpha-1} - (1+r(g,\eta^L)\right]}$. The observation here is that both the numerator and the denominator are greater for the "low" type borrower (right-hand side of the inequality) due to lower repayment rates and higher interest rates. However, this inequality holds because $0 < \alpha < 1$. For any positive X, Y, the fraction $\frac{X-Y}{\alpha X - Y}$ becomes smaller as X increases. Conversely, the fraction becomes larger as Y is larger.

3. $-\frac{\frac{\partial r(g,\eta^H)}{\partial g}}{\left[\alpha \cdot \frac{A(\eta^H)}{\lambda(\eta^H)} \cdot L^{\alpha-1} - (1+r(g,\eta^H)\right]} < -\frac{\frac{\partial r(g,\eta^L)}{\partial g}}{\left[\alpha \cdot \frac{A(\eta^L)}{\lambda(\eta^L)} \cdot L^{\alpha-1} - (1+r(g,\eta^L)\right]}$ is satisfied as the decrease in interest rate with respect to the guarantee rate is more pronounced for the "low" type borrowers.

Therefore, since each component of the inequality is greater for the "low" type borrower (right-hand side), the overall expression holds true. This implies that marginal rate of substitution of loan size for guarantee rate is steeper for the "high" type borrower (η^h) compared to the "low" type borrower (η^l). (i.e. $|MRS^h_{L,g}| > |MRS^l_{L,g}|$)

C Exogeneity of Repayment Rate: A Regression Discontinuity Approach

This section employs a regression discontinuity design to explore the causal relationship between loan size and borrowers' repayment rates, aiming to validate the exogeneity of repayment rates. This validation is crucial as it supports the assumption that repayment rates, used as proxies for borrower type in the model developed later, are not influenced by loan size or other variables. It also demonstrates that the observed correlation between loan size and repayment rate, as noted in the main text, arises from agency screening.

The analysis leverages a unique feature of South Korea's credit rating system, which, prior to 2020, categorized borrowers into grades from AAA to D based on their credit scores, with grade AAA representing the highest creditworthiness. The details of this grade score mapping are provided in Table C.1. These credit grades significantly influenced loan sizes for "special guarantee products," which were designed with targeted, narrow policy objectives for specific small business sectors and were only available for limited periods following specific events or conditions. This contrasts with "general guarantee products," which are available to all borrowers at any time. From

Table C.1: Credit Grade Mapping

Credit Grade	**Credit Score**
AAA	900 ∼ 1000
AA	870 ∼ 899
A	840 ∼ 869
BBB	805 ∼ 839
BB	750 ∼ 804
B	665 ∼ 749
CCC	600 ∼ 664
CC	515 ∼ 599
C	445 ∼ 514
D	0 ∼ 444

the perspective of the borrower, the consequences of defaulting on a loan are the same regardless of whether the loan is a special or general guarantee, suggesting that the impact of loan size on repayment behaviors is likely consistent across both types of guarantees.

The screening process for assigning loan sizes to these "special guarantee products" was notably simpler and less detailed, focusing largely on the borrower's credit grade. This simpler screening process creates a natural setting for employing a regression discontinuity design, particularly due to the observable jumps in loan sizes at these credit grade thresholds.

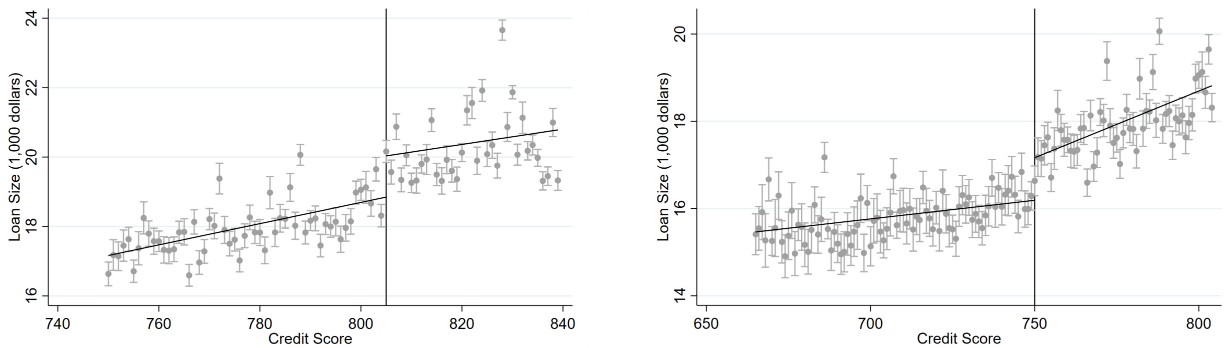
Illustrated in Figure C.2a, there's a clear jump in loan sizes at credit grade cutoffs for these "special guarantee products." The discontinuity in loan sizes increases by \$1.9k at the BB to BBB cutoff and \$1.2k at the B to BB cutoff, both with highly significant p-values of <0.001. This observation provides a foundation for a regression discontinuity analysis, suggesting that borrowers near these thresholds are essentially similar, except for the loan size they are allocated based on their grade. The aim of this analysis is to investigate whether these notable increases in loan sizes at the grade boundaries correspond to changes in repayment rates.

However, Figure C.2b reveals no significant discontinuities in repayment rates at these credit grade boundaries, with p-values of 0.275 for the cutoff between BB and BBB, and 0.243 for the cutoff between B and BB, indicating that the loan size increases associated with grade classification do not substantially influence borrowers' repayment rate. Given this finding, it is reasonable to infer that marginal increases in loan size likely do not affect repayment rates under general guarantees either. This supports the conclusion that the observed correlation between larger loan sizes and higher repayment rates, as discussed in the preceding subsection 3.2, is primarily a result of the agencies' strategic screening, rather than the effect of loan size itself.

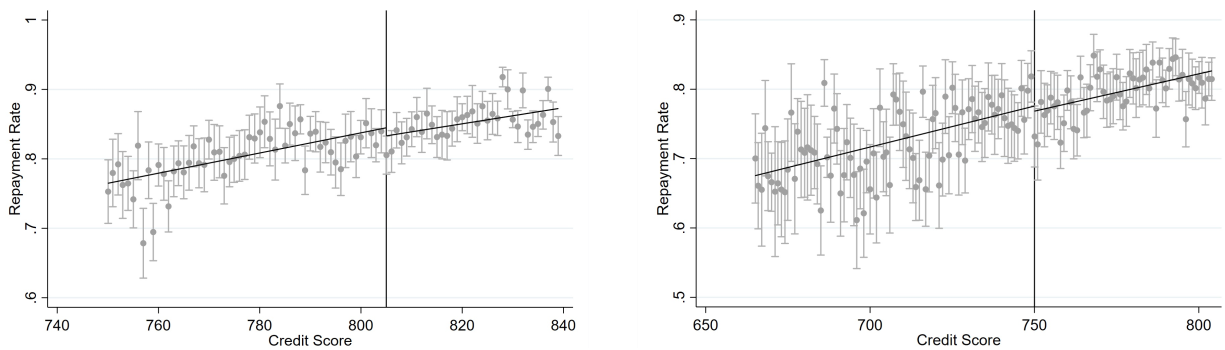
To formally estimate these patterns, I implement the fuzzy regression discontinuity design

Figure C.2: Exogeneity of repayment rate

(a) Loan size response at credit grade thresholds



(b) Repayment rate response at credit grade thresholds



Notes: The top two figures depict the discontinuity in loan sizes at the credit grade thresholds between BB and BBB (left), and B and BB (right). The bottom two figures illustrate the corresponding responses in repayment rates. Both sets of plots employ bin scatter plots to group data into bins along the credit score axis, effectively reducing variability and enhancing the visualization of trends and discontinuities.

described below.

C.1 Empirical Strategy and Results

Given the institutional setting described above, this study exploits discrete credit-grade thresholds to identify the causal effect of loan size on repayment behavior, employing a fuzzy regression discontinuity (RD) design with a continuous endogenous treatment variable. In the South Korean Special Guarantee program, a borrower's credit score, a continuous measure ranging from 0 to 1000, serves as the running variable determining credit grade (AAA, AA, A, BBB, BB, etc.) via pre-specified cutoffs.

While crossing a given credit-score threshold does not deterministically assign a specific loan size, borrowers just above the threshold typically receive larger loans than those just below it. Loan officers, however, retain some discretion, adjusting loan amounts based on additional borrower-specific characteristics such as liquidity needs, available collateral, and subjective assessments of creditworthiness. Thus, the threshold does not yield a sharp discontinuity in loan size, but rather a probabilistic shift in the distribution, resulting in a fuzzy RD scenario.

Formally, we model the relationship between repayment rate (the outcome variable, denoted λ_i) and loan size (the endogenous treatment, L_i) as follows:

$$\lambda_i = \alpha + L_i \tau \epsilon_i$$

In this specification, τ, the coefficient of interest, captures the causal impact of loan size on repayment rate.

To credibly estimate τ, we leverage the discontinuities at the credit-grade thresholds. Specifically, we apply a fuzzy RD estimator, constructing our identification strategy on two fundamental discontinuities at the threshold credit score . First, we estimate the discontinuity in the loan size, L_i, around the threshold:

$$\lim_{x \downarrow c} \mathbb{E}[L_i \mid \text{Score} = x] - \lim_{x \uparrow c} \mathbb{E}[L_i \mid \text{Score} = x] \tag{1}$$

Second, we estimate the corresponding discontinuity in the repayment outcome, λ_i:

$$\lim_{x \downarrow c} \mathbb{E}[\lambda_i \mid \text{Score} = x] - \lim_{x \uparrow c} \mathbb{E}[\lambda_i \mid \text{Score} = x] \tag{2}$$

Thus, our fuzzy RD estimate for the causal effect of loan size on repayment rate is computed as the ratio of these two discontinuities:

$$\tau = \frac{\lim_{x \downarrow c} \mathbb{E}[Y_i \mid \text{Score} = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i \mid \text{Score} = x]}{\lim_{x \downarrow c} \mathbb{E}[L_i \mid \text{Score} = x] - \lim_{x \uparrow c} \mathbb{E}[L_i \mid \text{Score} = x]}$$

This equation provides the local average treatment effect (LATE), capturing the impact of an incremental increase in loan size on repayment performance for borrowers at the credit grade threshold.

Table C.2 summarizes the formal fuzzy RD estimates using local polynomial regressions, both pooled across thresholds (with normalized cutoffs) and separately for each individual cutoff. The first-stage estimates confirm strong and statistically significant discontinuities in loan sizes at nearly all considered thresholds. For example, in the pooled specification, crossing the credit threshold results in an average loan-size increase of approximately 1,588 dollars (p-value < 0.001). However, the second-stage estimates consistently show insignificant effects on repayment rates. The pooled fuzzy RD estimate of loan size on repayment is 0.007, which is statistically insignificant (p-value =

0.422).

Additionally, we examine whether there are other important discontinuities at the credit-grade cutoffs that might affect our results, particularly interest rates set by lenders. Online Appendix Table C.3 confirms that there are no significant discontinuities in interest rates at these thresholds, suggesting that lenders use the full credit score information rather than discrete credit-grade categories when determining interest rates, further giving credibility to the fuzzy RD design to study the causl effect of loan size on repayment rates.

Overall, these findings robustly indicate that, although borrowers above credit score thresholds receive substantially larger loans, these variations in loan amounts do not translate into any significant changes in repayment performance. Thus, the positive correlation between loan size and repayment rates observed in the data likely arises entirely from screening by lenders rather than any causal effect of larger loans. We excluded the thresholds AAA–AA, CCC–CC, CC–C, and C–D due to either minimal first-stage discontinuities (AAA–AA) or insufficient data in the lower credit grades (CCC–CC, CC–C, and C–D), ensuring our analysis focuses on institutionally relevant and empirically credible credit-grade cutoffs.

Table C.2: Fuzzy RD Estimates of the Effect of Loan Size on Repayment Rates.

Threshold	**Obs.**	$\boldsymbol{\Delta L}$	$\hat{\boldsymbol{\tau}}$
Pooled	28,509	1,588.4*** (379.41)	0.007 (0.008)
AA–A	11,056	2,327.6*** (854.51)	0.015 (0.010)
A–BBB	11,425	1,630.5* (869.34)	0.036 (0.025)
BBB–BB	13,600	1,708.7* (945.24)	-0.016 (0.019)
BB–B	11,007	1,371.2** (590.38)	-0.017 (0.021)
B–CCC	6,416	1,402.9** (679.40)	0.035 (0.029)

Notes: ** p<0.01, * p<0.1. ΔL refers to the first-stage jump in loan size at each credit score threshold. $\hat{\tau}$ denotes the estimated effect of loan size on repayment rate from the second-stage fuzzy RD regression. The pooled estimate normalizes credit score cutoffs to zero and adjusts loan size by subtracting the average around each threshold. Each observation is assigned to its closest cutoff based on midpoints between adjacent credit grade boundaries. All estimates are from conventional fuzzy RD local linear regressions using a triangular kernel with local polynomial order 1. Effective sample sizes reflect observations within the optimal bandwidth around each cutoff. Standard errors are reported in parentheses.

C.2 Testing for Other Discontinuity at Credit Grade Cutoffs

One potential concern for the validity of the fuzzy regression discontinuity (RD) design employed above is the the potential existence of other types of discontinuities at the credit grade thresholds. Most notably, if interest rates charged to borrowers also exhibited discontinuities at these thresholds, it could confound our interpretation of the loan-size discontinuity. Specifically, lenders might alter interest rates discretely at the grade cutoffs, thereby indirectly affecting borrower repayment behavior and complicating the isolation of loan-size effects.

To address this possibility, Table C.3 presents sharp RD estimates examining the presence of discontinuities in interest rates at each credit grade cutoff, as well as a pooled estimate combining all thresholds. These estimates are generated using local linear regressions (order $p = 1$) with a triangular kernel, closely following the main empirical strategy.

The results indicate no statistically significant discontinuities in interest rates at any of the individual credit grade thresholds, nor in the pooled specification. For instance, the pooled discontinuity estimate is 0.038 percentage points, with a standard error of 0.040, which is not statistically significant. Likewise, individual threshold estimates are also statistically indistinguishable from zero. Thus, interest rates evolve smoothly across thresholds, reinforcing the plausibility and validity of our fuzzy RD approach exploiting loan-size discontinuities for identification.

Table C.3: Discontinuity in Interest Rates at the Credit Grade Cutoff

Threshold	**Obs.**	**Δ Interest Rate**
Pooled	28,492	0.03755 (0.04001)
AA–A	11,046	0.05538 (0.0721)
A–BBB	11,418	0.17975 (0.08904)
BBB–BB	13,595	0.01763 (0.0800)
BB–B	11,003	-0.23731 (0.08368)
B–CCC	6,410	-0.02308 (0.07861)

Notes: These estimates are obtained using sharp RD designs via local linear regressions (p=1) with a triangular kernel. Observations are assigned based on the running variable (credit score) and the credit grade cutoff. Standard errors are shown in parentheses.

D Microfoundation for Borrower Utility from a Loan Contract

Borrower utility from a loan contract in our model is derived from a sequential decision-making process over T periods. A borrower has a stochastic investment technology that produces output each period as a function of loan size L_i.

$$f_{it}(L_i) = S_t \times z_i L_i^{\alpha}$$

Each period, the borrower's output is influenced by an exogenous shock S_t, determining business success ($S_t = 1$) or failure ($S_t = 0$). For periods where the business succeeds ($S_t = 1$), it generates an output of $z_i L_i^{\alpha}$. z_i is a productivity shifter, and the parameter α captures the concavity of the production function. Conversely, in the event of business failure ($S_t = 0$), the output plummets to zero, mirroring the cessation of operational activities. Failure in any period leads to a persistent state of non-success in all subsequent periods ($S_{t+k} = 0$ for $k > 0$).

At each period t, the borrower is confronted with a choice to repay the loan or not, based on which option maximizes their linear utility from consumption. (i.e $u(C) = C$) This decision-making process is formalized as the utility maximization problem:

$$V_t(S_t) = \max \left\{ \underbrace{C_t + \beta E V_{t+1}}_{\text{utility when repaying}} \; , \; \underbrace{-D_i}_{\text{utility when defaulting}} \right\}$$

$$s.t. \quad C_t + \left(\frac{L_i}{T} + r_i \left(L_i - \sum_{\tilde{t}=1}^{t-1} \frac{L_i}{T} \right) \right) \leq S_t \times z_i L_i^{\alpha}$$

In this model, the repayment amount $\frac{L_i}{T}$ represents an equal amortization of the principal L_i over T periods. The interest payment for each period is calculated based on the remaining principal balance, where r_i is the interest rate applied to the unpaid portion of the principal. Such a structure is consistent with conventional loan repayment schemes in loan guarantee programs in South Korea.

Borrowers are presented with a decision each period: to repay the loan, which entails consuming the residual output C_t after repayment, or to default. Defaulting would result in the forfeiture of assets, a drop to zero consumption $C_t = 0$, and the incurrence of a default cost, represented by D_i. The term D_i encompasses costs associated with defaulting, such as the detrimental impact on the borrower's credit score. The model presupposes that borrowers will default when facing negative consumption, as defaulting is considered less detrimental than accumulating debt with no corresponding consumption. This is expressed as $-\left(\frac{L_i}{T}+r_i\left(L_i-\sum_{\tilde{t}=1}^{t-1}\frac{L_i}{T}\right)\right)<-D_i$, indicating that the utility loss from defaulting is less than the disutility of negative consumption.

Working backward from the final period, the decision rule—or policy function—dictates that borrowers repay if $S_t=1$ and default if $S_t=0$, aligning repayment decisions directly with the success state of the business. We convert flow utility V_t into stock utility over T periods, U_i, by aggregating utilities from each period. This stock utility accounts for total output, total repayment, total interest payment, and default costs (applied if any period results in failure).

$$U_i=\underbrace{\sum_{t=1}^{T}\mathbf{1}[S_t=1]\beta^t z_i L_i^{\alpha}}_{total\,output}-\underbrace{\sum_{t=1}^{T}\mathbf{1}[S_t=1]\beta^t\frac{L_i}{T}}_{total\,principal\,repayment}-\underbrace{\sum_{t=1}^{T}\mathbf{1}[S_t=1]\beta^t r_i\left(L_i-\sum_{\tilde{t}=1}^{t-1}\frac{L_i}{T}\right)}_{total\,interest\,payment}-\underbrace{\mathbf{1}[\exists t:S_t=0]\beta^{min\{t:S_t=0\}}D_i}_{default\,cost}$$

I simplify the representation by defining $T\times z_i=A_i$ to represent the total potential productivity. I also introduce $\lambda_i=\frac{\beta(1-\beta^k)}{T(1-\beta)}$ as the effective repayment rate, which accounts for discounting over the repayment term. The expression $a(\lambda_i,r_i)\cdot L_i$ transforms the total interest payments into a function of λ_i, aligning it with the effective repayment rate. Furthermore,$d(\lambda_i,D_i)$ represents the discounted default cost, applicable if a default occurs at any point during the loan's term.

$$U_i(\lambda_i,L_i,r_i)=\underbrace{\lambda_i\cdot A_i L_i^{\alpha}}_{total\,output}-\underbrace{\lambda_i\cdot L_i}_{total\,principal\,repayment}-\underbrace{a(\lambda_i,r_i)\cdot L_i}_{total\,interest\,payment}-\underbrace{d(\lambda_i,D_i)}_{default\,cost}$$

Note that the borrower's utility is a function of three pivotal elements: the repayment rate λ_i, the loan size L_i, and the interest rate r_i. In the context of a 5-year loan term and a low-interest-rate environment, I argue that discounting effects are minimal, justifying the approximation $\lambda_i\approx\frac{k}{T}$ for simplicity ($\beta\approx1$). This simplification enhances the tractability of the model without significantly detracting from the accuracy of the main findings.

E Interest Payment Structure

This section outlines the typical structure for interest payments associated with 5-year loans under the study. Borrowers generally repay the principal in equal installments every three months. Interest is calculated on the remaining principal amount, thus as borrowers continue to repay, the interest burden decreases. Given this structure, a borrower who fully repays the principal over the term of the loan will have paid an amount equivalent to 2.625 times the nominal interest rate, denoted by r_i.

It is important to clarify that the repayment rate $lambda_i$, which represents the fraction of the principal repaid, does not directly translate to a payment of $\lambda_i\times2.625\times r_i$. The reason for this is that the loan repayment is amortized. To accurately calculate the interest payments when a borrower repays λ_i of the loan, we employ the function $a(\lambda_i,r_i)$. This amortization calculator determines the actual interest payments, where $a(\lambda_i,r_i)\times L_i$ gives the total amount paid by the borrower.

While we use a quarterly repayment interval for consistency in our calculations, it is acknowledged that loan terms can occasionally differ, including variations in the timing of repayments and changes in interest rates. However, these variations are relatively infrequent and do not substantially alter the typical borrowing and repayment behaviors captured in our model. This approach allows for a standardized analysis across diverse loan agreements, simplifying the calculation process while providing a reliable approximation of borrower obligations under common loan conditions.

F Equilibrium

The equilibrium in this model is defined as a Perfect Bayesian Equilibrium (PBE), which involves two players: the borrowers i and the government agency j. The setup is conditional on the borrower's observable characteristics X_i, which are known to both players. Each borrower is characterized by a two-dimensional type space $\Theta_i = \{(\eta_i, \epsilon_i) \in \mathbb{R}^2\}$, where η_i represents the borrower's repayment type and ϵ_i denotes the preference shock for guarantee rate. The agency's type space, $\Theta_j = \{(s_j^{small}, s_j^{large}) \in \mathbb{R}^2\}$, is determined by signals from borrower screening, where s_j^{small} and s_j^{large} correspond to signals for the small and large loan program, respectively.

F.1 Strategies

- Borrower's strategy, given by $\sigma_i : (\eta_i, \epsilon_i; X_i) \to G_i \in \{small, large\}$, dictates their choice between the small or large loan program, factoring in their type and observed characteristics.
- Agency's strategy, $\rho_j : (G_i, s_j^G; X_i) \to L_j \in \mathbb{R}^+$, determines the loan size based on the borrower's guarantee choice G_i and the corresponding signal s_j^G, which varies (s_j^{small} for the small loan program and s_j^{large} for the large loan program).

F.2 Beliefs

- Borrowers' beliefs (b_i): Borrowers form beliefs about the agency's signals, $b_i(s_j^G|\eta_i)$ based on their repayment type η_i. These beliefs dictate their expectations about potential loan sizes under each guarantee option, influencing the choice between the small and large loan program. The decision between the small and large loan program hinges on balancing two expectations: the expected change in business output due to the loan size, $E(\underbrace{\triangle(A_i \cdot L_i^\alpha - L_i)}_{\text{diff in output}}|\eta_i, X_i)$, against the expected difference in interest payments between the two options $E(\underbrace{\triangle interest}_{\text{diff in interest}}|\eta_i, X_i)$.
- Agency's beliefs (b_j): The agency forms beliefs $b_j(\eta_i|G_i, s_j^G)$ about the borrower's repayment type η_i based on the borrower's guarantee choice G_i and the screening signals received. These beliefs guide the agency in updating its expectations regarding the borrower's repayment rate, represented by $E(\lambda_i|G_i, s_{ij}^G)$, and subsequently the loan size decision.

F.3 Sequential Rationality

Sequential rationality ensures that each player's strategy is optimal given their beliefs and the strategies of other players, taking into account the information available at each decision point. This requires that:

- For Borrowers: Each borrower's strategy of choosing between a full or partial guarantee must be optimal, based on their expectations about the agency's response and the potential outcomes. Specifically, for any borrower type $(\eta_i, \epsilon_i) \in \Theta_i$, their strategy σ_i^* must maximize their expected utility, considering the agency's subsequent actions and the borrower's beliefs about the agency's signal. Mathematically, this is expressed as:

$$\forall(\eta_i, \epsilon_i) \in \Theta_i, \quad \sigma_i^*(\cdot|\eta_i, \epsilon_i) \in argmax_\sigma \int_{s_j^G} U_i(\eta_i, \epsilon_i, s_j^G, \sigma_i, \rho_i^*) b_i^*(s_j^G|\eta_i) ds_i^G$$

Here, U_i is the utility function for the borrower, and μ_i^* is the borrower's beliefs about the agency's signal, conditioned on their own type and observed characteristics.

- For the Agency: The agency's strategy in determining the loan size must be optimal, considering the borrowers' guarantee choices and the agency's beliefs about the borrowers' repayment types. For each borrower guarantee choice $G_i \in \{small, large\}$ and signal $s_j^G \in \Theta_j$, the agency's strategy ρ_j^* should maximize its expected utility based on its beliefs about the borrower's type. Formally, this is represented as:

$$\forall G_i \in \{small, large\}, \forall s_j^G \in \Theta_j, \quad \rho_j^*(\cdot|G_i, s_j^G) \in argmax_\rho \int_{\eta_i} U_j(\eta_i, G_i, \rho_j) b_j^*(\eta_i|G_i, s_j^G) d\eta_i$$

In this equation, U_j denotes the utility function for the agency, and b_j^* signifies the agency's beliefs about the borrower's repayment type, influenced by the received signal and the borrower's chosen guarantee.

F.4 Existence of Separating Equilibrium

I address the existence of a separating equilibrium. Mailath [1987] outlines sufficient conditions for such an equilibrium. A pivotal condition is the single-crossing property applied to the borrower's utility function:

$$\frac{V(\eta, \epsilon, \tilde{\eta}, small; X) - V(\eta, \epsilon, \tilde{\eta}, large; X)}{\frac{\partial}{\partial \tilde{\eta}} V(\eta, \epsilon, \tilde{\eta}, G; X)} \text{ is monotone in } \eta$$

Here, $V(\eta, \epsilon, \tilde{\eta}, G; X)$ represents the expected utility for a borrower of type (η, ϵ) choosing a guarantee G, with the agency perceiving the borrower's repayment type as $\tilde{\eta}$. The vector X denotes a set of borrower characteristics known to the agency. This condition intuitively holds with the borrower's utility structure. The numerator represents the increase in the borrower's utility from selecting the small loan program over the large loan program, while the agency's beliefs ($\tilde{\eta}$) remain constant. The utility increase from higher funding probabilities and reduced interest rates under the small loan program is more pronounced for borrowers with lower type. Hence, the numerator should be decreasing in η. The denominator corresponds to the utility gain when agency perceives the borrower to be a better type, holding fixed the borrower's guarantee choice

$\bar{G}$. An improvement in the agency's beliefs increases the loan size, which is particularly beneficial for higher types due to their increased productivity with equivalent loan sizes. Thus, the denominator should increase with η. These two forces suggest that the single-crossing condition described should indeed be monotonically decreasing in η.

Similar to the approach in Kawai et al. [2022], I estimate the model parameters initially without verifying the existence of a separating equilibrium. Following the estimation, I then assess whether the single-crossing condition is upheld at the estimated parameter values. This procedure confirms that, at the estimated values, the condition sufficient for separation is indeed satisfied.

G Estimation

This section provides further detail on the maximum likelihood approach discussed in Section 6. The log-likelihood function is conditional on observed outcomes in the data.

G.1 Interest Rate Prediction

An empirical challenge in the model is that I only observe the interest rates associated with the guarantee rate each borrower has actually received. However, the borrower choice model needs the counterfactual interest rates—what borrowers would have been offered had they been funded under an alternative guarantee rate (either 85% or 100%). This issue is common in banking literature, where they often need to predict potential interest rate outcomes under different hypothetical scenarios, such as if borrowers had chosen other lenders or if they had or had not provided collateral.

Following common practices in the field, as illustrated in studies by Adams et al. [2009], Crawford et al. [2018], and Ioannidou et al. [2022], I employ a predictive approach to estimate these counterfactual interest rates. I utilize the following Ordinary Least Squares (OLS) regression model with a comprehensive set of controls:

$$r_i = \psi_l Large_i + \psi_\lambda \lambda_i + \psi_{\lambda l}(\lambda_i \times Large_i) + X_i \Psi + \xi_i$$

This model predicts the counterfactual interest rate r_i a borrower would likely have received under a counterfactual guarantee type. The model achieves an R-squared of 0.46, with detailed regression results reported in Online Appendix K. After predicting these rates, I use both the predicted and the actual observed interest rates to compute $E\Big(\triangle interest_{ij}|\eta_i, X_i\Big)$, the expected difference in interest payments under different guarantee scenarios, as demonstrated in the subsequent section.

G.2 Estimation Procedure

In estimating the parameters on the borrower side, I exploit a key simplification: the borrower's decision-making can be analyzed as a single-agent optimization problem with respect to the expected loan size. This simplification is possible because the equilibrium loan sizes are directly observable in the dataset, which circumvents the need to solve for equilibria during the estimation phase.

More specifically, as described in subsection 4.3.3, a borrower's guarantee choice is influenced by the expected difference in business output due to the loan, $E\Big(\triangle(A_i \cdot L_{ij}^{\alpha} - L_{ij})|\eta_i, X_i\Big)$, and the expected difference in interest payments, $E\Big(\triangle interest_{ij}|\eta_i, X_i\Big)$. These expectations are directly estimated from the data, leveraging the equilibrium loan sizes L_{ij}, which are observable. To estimate these differences, I conduct separate regressions for both the small loan program and the large loan program to predict the expected output and interest payments for each guarantee condition. Using these predicted values, I then compute the expected differences in both output and interest costs between the the small loan program and the large loan program. The details of the estimation procedure are provided in Online Appendix G.

For counterfactual analyses, however, the equilibrium is resimulated. I approach this by iteratively simulating the borrower's guarantee choice and the agency's loan size decision, while keeping the estimated parameters fixed. This iterative process involves solving for a fixed point, where the borrower's expected difference in loan size aligns closely with the simulated loan size, within a predefined tolerance level. Details of this procedure are provided in the Online Appendix H.

G.3 Likelihood of the observed loan size

While other components of the joint-likelihood are relatively straightforward, forming the likelihood of the observed loan size $\mathcal{L}(L_i|G_i, \lambda_i; \Theta)$ set by the agency for borrower i presents a non-trivial challenge. The likelihood is derived from the distribution of the information signal (s_i^G) the agency received from the borrower, which I aim to estimate. This process involves inverting the observed loan size L_i back to the underlying information signal s_i^G. Such inversion involves inverting the agency's beliefs about the repayment rate, $E(\lambda_i|G_i, s_i^G)$, to the information signal, s_i^G. Due to the analytical intractability of directly inverting $E(\lambda_i|G_i, s_i^G)$ to s_i^G, a simulation-based approach is employed.

1. Take N pairs of iid standard normal draws for the borrower repayment type conditional on the information signal, $\eta|s_i^G$, and the preference shock, ϵ_i. For any given set of model parameters, I then scale these draws up or down.

2. Generate a grid of parameters for the model. For each parameter set, simulate $\eta|s_i^G$ from the distribution $N(\mu_i + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_G^2} s_i^G, \frac{\sigma_\eta^2\sigma_G^2}{\sigma_\eta^2+\sigma_G^2})$ by scaling the simulated draws from step 1. Simultaneously, simulate ϵ_i from $N(0, \sigma_\epsilon^2)$ by scaling the simulated draws generated in step 1.

3. Retain those draws that opt for the large loan program, i.e., pairs of draws that satisfy the inequality specified in Section 5.3.3. Using these filtered draws for η_i, compute the mean to approximate $E(\lambda_i|large, s_i^{large})$.

4. Using the simulation grid, create an interpolated inversion function f_{large}^{-1}: $E(\lambda_i|large, s_i^{large}) \rightarrow \mu_i + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_{large}^2} s_i^{large}$.

5. Repeat the process to generate an interpolated inversion function for the small loan program scenario.(i.e. f_{small}^{-1}: $E(\lambda_i|small, s_i^{small}) \rightarrow \mu_i + \frac{\sigma_\eta^2}{\sigma_\eta^2+\sigma_{small}^2} s_i^{small}$) The only difference is that this function retains those draws that opt for the full guarantee rate, i.e., pairs of draws that do not satisfy the inequality specified in Section 5.3.3.

Using this interpolated inversion function, the likelihood is then formed as follows, where $\eta_i = \mu_i + u_i$ represents the borrower's type with μ_i being the common knowledge and u_i representing the private information, the source of information asymmetry between the agency and the borrower:

$$
\begin{aligned}
\mathcal{L}(L_i|G_i, \lambda_i; \Theta) &= \phi\left(\frac{f_G^{-1}(E(\lambda_i|G_i, s_i^G)) - \frac{\sigma_\eta^2 u_i}{\sigma_\eta^2+\sigma_G^2} - \mu_i}{\frac{\sigma_\eta^2 \sigma_G}{\sigma_\eta^2+\sigma_G^2}}\right) \times \left|\frac{dE(\lambda_i|G_i, s_i^G)}{dL}\right| \times \left|\frac{df_G^{-1}(E(\lambda_i|G_i, s_i^G))}{dE(\lambda_i|G_i, s_i^G)}\right| \times \frac{1}{\frac{\sigma_\eta^2 \sigma_G}{\sigma_\eta^2+\sigma_G^2}} \\
&= \phi\left(\frac{f_G^{-1}\left(\frac{L^{1-\alpha}((1-\tau_j)(1-fee_i)g_i+\tau)}{L^{1-\alpha}g_i(1-\tau_j)+\alpha\tau A}\right) - \frac{\sigma_\eta^2 u_i}{\sigma_\eta^2+\sigma_G^2} - \mu_i}{\frac{\sigma_\eta^2 \sigma_G}{\sigma_\eta^2+\sigma_G^2}}\right) \\
&\times \left|\frac{-(g_i(1-fee_i)(1-\tau_j)+\tau)(1-\alpha)\left(-L^{2-\alpha}(A\alpha\tau + L^{1-\alpha}g_i) + L^{3-2\alpha}g_i(1-\tau_j)\right)}{L^2(A\alpha\tau_j + L^{1-\alpha}g_i(1-\tau_j))^2}\right| \\
&\times \left|\frac{df_G^{-1}(E(\lambda_i|G_i, s_i^G))}{dE(\lambda_i|G_i, s_i^G)}\right| \times \frac{1}{\frac{\sigma_\eta^2 \sigma_G}{\sigma_\eta^2+\sigma_G^2}}
\end{aligned}
$$

Note that $\frac{df_G^{-1}(E(\lambda_i|G_i, s_i^G))}{dE(\lambda_i|G_i, s_i^G)}$ is also calculated numerically using the interpolated inversion function.

G.4 Maximum Likelihood Estimation Process

The MLE estimation was done using Python. The complete process can be summarized in the following steps:

1. Using OLS regression, predict $E(L_{ij}^\alpha|\eta_i, X_i)$ and $E(L_{ij}|\eta_i, X_i)$ as quadratic functions of borrowers' repayment rate λ_i—a censored version of η_i—and observed characteristics X_i for both partial and full guarantee loans. These regressions inform the equilibrium expected differences in loan sizes and interests as in $E\Big(\triangle(A_i \cdot L_{ij}^\alpha - L_{ij})|\eta_i, X_i\Big)$ and $E\Big(\triangle interest_{ij}|\eta_i, X_i\Big)$ of the borrower guarantee choice model.

2. Compute the joint log-likelihood, incorporating the likelihood of L_{ij} using the interpolated inversion functions generated earlier.

3. The log-likelihood function is not globally concave and includes flat sections, which pose challenges for computational maximization routines. To enhance the probability of identifying the global maximum, I conduct a global search algorithm that emphasizes shifting away from potential local extrema. I use the dual annealing function from Python's scipy.optimize library for global optimization. This method generalizes the traditional simulated annealing algorithm, which is designed to avoid getting trapped in local minima by performing random steps and controlled reheating, effectively exploring a broad parameter space. In my application, dual annealing performs 1000 random "steps" or iterations, to robustly explore the global search space. By opting to set no local search to False, the method automatically includes a subsequent local search phase using the L-BFGS-B algorithm.

H Simulation Details

For all the simulations using the estimated model, I follow the procedure we describe below:

1. Draw 50 sets of shocks for each borrower in the sample. This includes the borrower's repayment type (η_i), preference (ϵ_i), the noise in the agency's signal for both the small (δ_j^{small}) and large (δ_j^{large}) loan programs, and the lender funding shock (ζ_i) using the estimated distribution. Drawing multiple shocks per borrower essentially increases the number of simulations, similar to increasing the sample size, which helps reduce simulation errors. Since the analysis focuses on average outcomes, expanding the number of simulations does not alter the results but ensures more reliable and smoother outcomes.

2. For each simulated borrower, calculate the interest rates applicable for loans with the small and large loan program.

3. Simulate borrower's guarantee choices and agency's loan size decisions. This step requires solving a fixed point problem because the the borrowers take the expectation of the loan size for each guarantee choice conditional on their repayment type, and the agency forms beliefs on the borrower's repayment type conditional on the borrower's guarantee choice. I proceed by: (i) computing the conditional expectation of the repayment type (η_i) based solely on the information signal ($\eta_i + \delta_i^G$), (ii) computing the agency's loan size for each guarantee choice, (iii) computing the borrower's expected loan size for each guarantee choice, (iv) computing the simulated conditional repayment type, and (v) repeating (ii)-(iv) until convergence of loan size.

4. Simulate lender rejection for the large loan program with 85% guarantee rate. Borrowers whose applications are rejected under 85% guarantee rate are then offered loans the small loan program with a 100% guarantee rate.

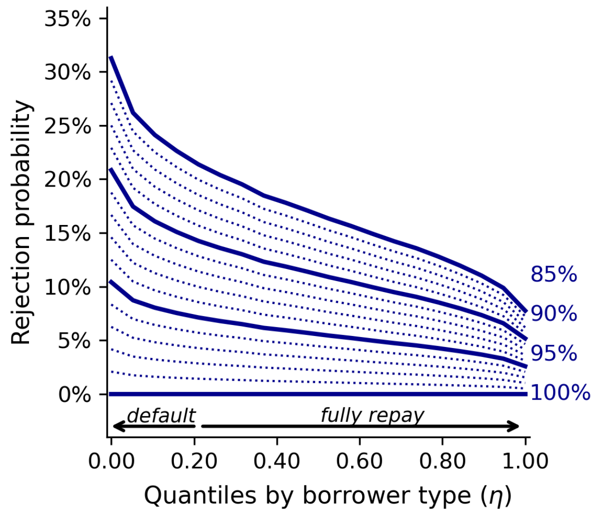
I Further Detail on Counterfactual (i) Uniform Guarantee Rates

This section presents detailed analyses related to the impact of employing uniform program of guarantee rates between 85% and 100%, supplementing the main findings discussed in the paper. Figure I.3a illustrates the linearly interpolated rejection rates between the 85% and 100% guarantee rates, and Figure I.3b shows how these rates affects the average approved loan size. This interpolation provides insights into how varying guarantee rates could potentially influence lender behaviors and affect the average loan size.

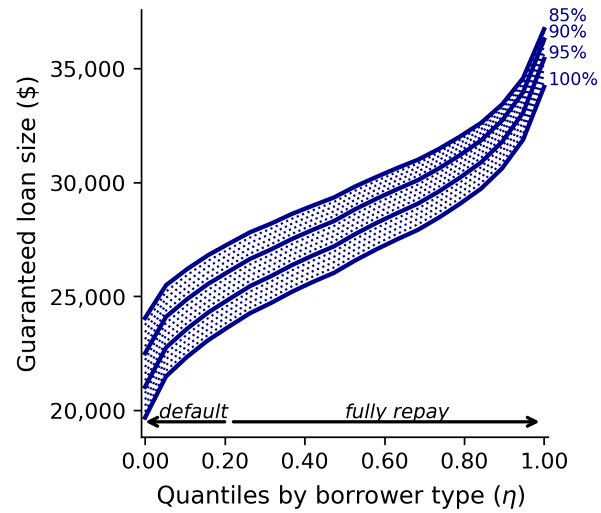
Further decomposition of the agency's objectives is detailed in Figures I.3c and I.3d, which analyze the effects of different uniform guarantee rates on agency loss per borrower and the value added by small businesses, respectively. The figures show that as the guarantee rate increases, both the agency's losses per borrower and the value added by small businesses rise. This demonstrates a trade-off from the agency's perspective: higher guarantee rates lead to greater losses due to increased default risk coverage, but they also enhance the value added by facilitating greater access to credit for businesses. Despite the trade-offs, the differences in agency hybrid objectives across various guarantee rates are relatively minor, as shown in Figure I.4. While a 96% uniform guarantee rate maximizes the agency's objective, the 100% guarantee rate achieves an objective close to this maximum. For simplicity and clarity in presentation, this analysis employs

Figure I.3: Outcome across Different Uniform Guarantee Rates

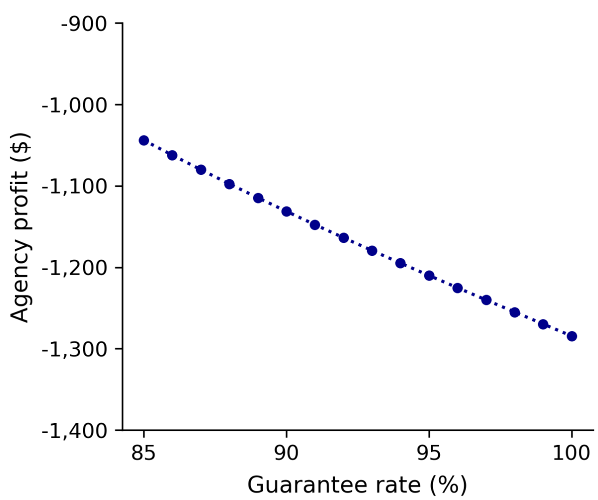
(a) Rejection probability across guarantee rates



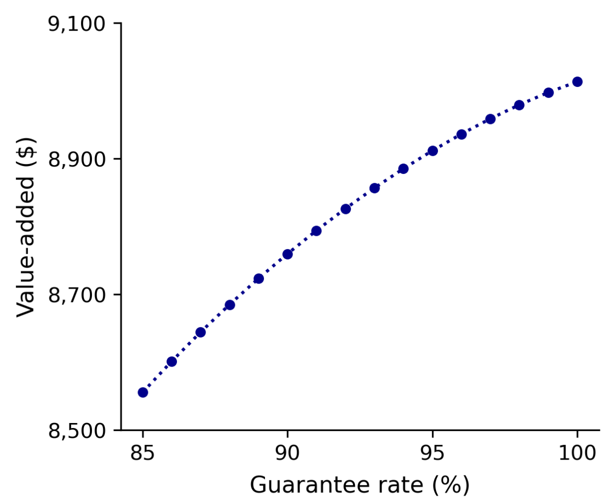
(b) Approved loan size ($) across guarantee rates



(c) Agency loss ($) across guarantee rates

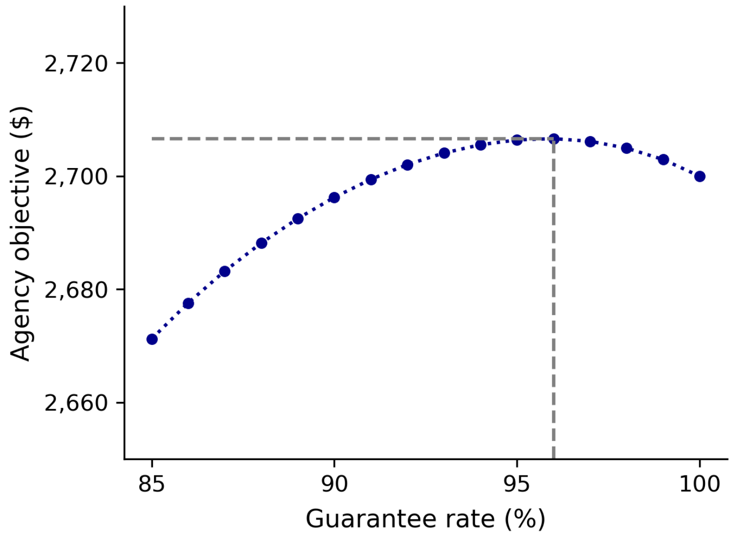


(d) Value-added ($) across guarantee rates



a 100% uniform guarantee rate as the main counterfactual, which streamlines the discussion by eliminating the possibility of lender rejection.

Figure I.4: Agency objective ($) across uniform guarantee rates (85% - 100%)

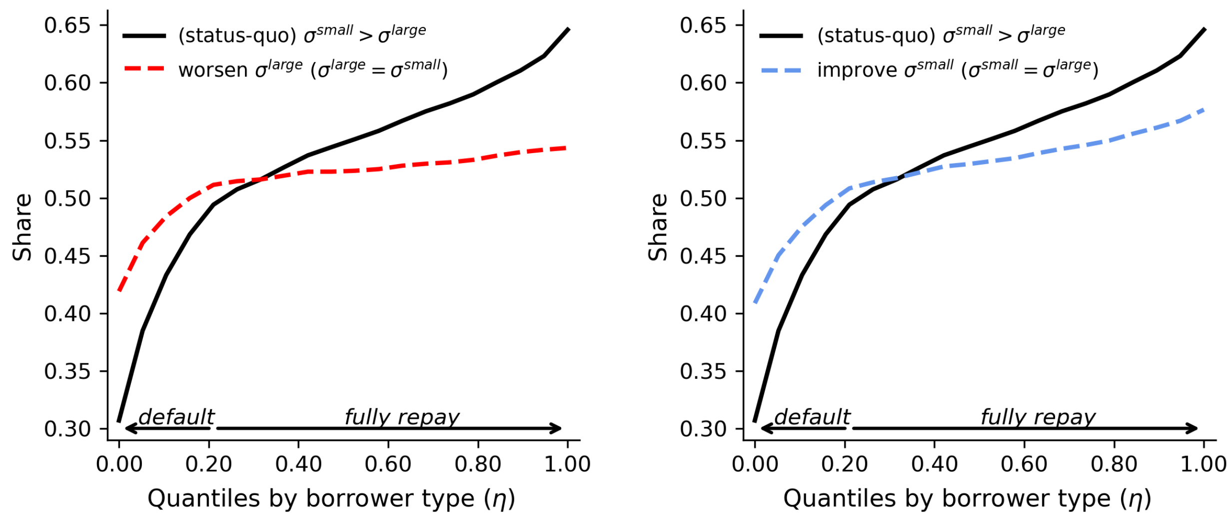


J Counterfactual: Uniform Screening Precision Across Guarantee Rates

This section investigates the effects of equalizing screening precision of soft information collection ($\sigma_{small} = \sigma_{large}$) for the small and large loan program on borrower sorting across different guarantee options. I perform two counterfactual exercises: one by reducing the precision of the large loan program to match that of the small loan program, and another by enhancing the precision of the small loan program to align with the precision of the large loan program. These adjustments aim to test the impact of uniform screening precision on the selection behavior of borrowers. Figures J.5a and J.5b display the share of borrowers opting for the the large loan program under each scenario. Both figures indicate that setting the screening precision equally across guarantee options results in reduced sorting among borrowers, suggesting that differences in screening precision in soft information collection significantly influence borrower decisions and are a key driver of self-selection in the loan guarantee menu.

Figure J.5: Share of borrowers opting for “large”

(a) Uniformly low precision (b) Uniformly high precision



K Interest Rate Prediction

Figure K.6: Interest rate distribution

(a) Interest rate distribution (b) Interest rate by grade

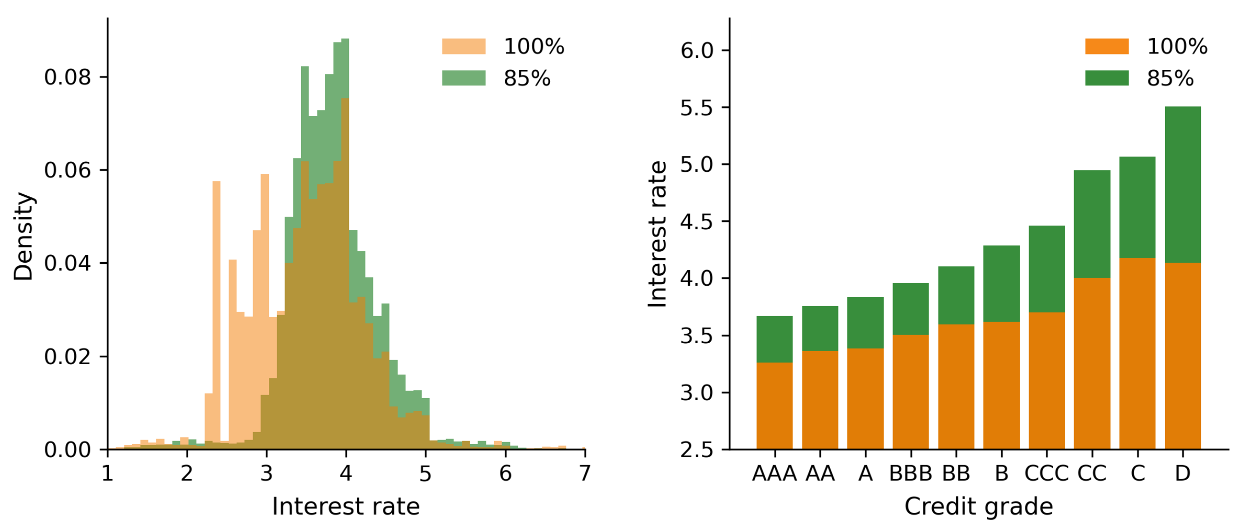


Table K.4: Interest rate prediction (OLS)

Variable	**Coeff**	**S.E.**
Constant	3.4412	0.155
Repayment rate	-0.0650	0.026
Credit score	-0.0005	$<$0.001
Business age	0.0009	0.001
Home-ownership	-0.0168	0.010
Number of employees	-0.0050	0.003
Debt ($10k)	-0.0007	$<$0.001
Owner age	0.0062	0.001
$\mathbf{1}[g_i = 85\%]$	0.8934	0.050
$\mathbf{1}[g_i = 85\%]$ $\times$ Repayment rate	-0.1927	0.044
$\mathbf{1}[g_i = 85\%]$ $\times$ Credit score	-0.0002	$<$0.001
$\mathbf{1}[g_i = 85\%]$ $\times$ Business age	-0.0057	0.002
$\mathbf{1}[g_i = 85\%]$ $\times$ Home ownership	-0.0396	0.017
$\mathbf{1}[g_i = 85\%]$ $\times$ Number of employees	0.0050	0.003
$\mathbf{1}[g_i = 85\%]$ $\times$ Debt ($10k)	0.0008	$<$0.001
$\mathbf{1}[g_i = 85\%]$ $\times$ Owner age	-0.0040	0.001
Region FE	Yes	
Industry FE	Yes	
Bank FE	Yes	
Observations	34,829	
R^2	0.455	

Notes: This table presents OLS regression results predicting nominal interest rates. The regression model includes controls, dummy variables indicating a guarantee rate of 85%, and fixed effects for region, industry, and bank. Standard errors are clustered by bank and region.