Why Hire Loan Officers? Examining Delegated Expertise

James Wang‡

University of Michigan

October 2014

Job Market Paper

Abstract

I develop and estimate a model of expertise that weighs the tradeoffs of delegating loan decisions to loan officers. While loan officers can screen soft information, their preferences may distort decisions away from the lender’s objective. The model features rich heterogeneity (e.g. risk aversion, willingness to exert effort, and overconfidence) and addresses possible endogeneity with a natural experiment. I can recover the loan officer’s primitives using loan decisions and resulting profits with data from a Chinese lender. I find that despite the costs, the average loan officer contributes three times his pay in additional annual profits compared to a counterfactual that maximizes profits conditional only on the borrower’s hard information. I further explore additional counterfactuals that eliminate some of the distortions due to preferences.

Keywords: Credit Markets, Delegated Expertise, Principal Agent, Overconfidence

* I am indebted to my thesis committee Francine Lafontaine, Daniel Ackerberg, Ryan Kellogg, Ying Fan, and Sugato Bhattacharyya for guidance. I want to thank the lender and especially Zane Wang, Sanyu Zhou, Hao Hu, and Yiqi Mao for providing data access and support. I am also grateful to seminar participants at the University of Michigan and China Credit Analytics for helpful comments including Aaron Flaaen, Evan Hernstadt, Cristina Iftimie, Max Kapustin, and Johannes Norling. Generous financial support was provided by the Haber Fellowship, CIBE Research Grant, MITRE Research Grant, University of Michigan Rackham Research Grant, and the University of Michigan Regent’s Fellowship. All errors are my own.

‡University of Michigan, Department of Economics, jtabw@umich.edu
1 Introduction

Many organizations, including universities, investment funds, and even boxing commissions, employ experts to screen uncodified information. For example, admissions counselors can read essays, analysts can interpret conference calls, and boxing judges can look for fatigue and aggression. However, these experts may impose costs for their organizations due to distortions caused by their preferences. In this paper, I examine delegating application screening and loan size choice to loan officers in the context of a Chinese lender that specializes in personal and small business loans.

To weigh the tradeoffs, I develop a structural model of delegated expertise that incorporates rich heterogeneity. While loan officers can screen uncodified information, they may be risk averse or unwilling to exert effort. Additionally, behavior inconsistent with a simpler model motivates a third type of heterogeneity where some loan officers may even exhibit overconfidence. I estimate the model using data on loan amounts and resulting profits while accounting for possible endogeneity with the random assignment of borrower applications to loan officers. Using the model primitives, I explore automated lending without loan officers where profits are maximized conditional only on the borrower’s codified information. Additional counterfactuals examine lending without some of the distortions to compare profitability both across and within loan officers. To the best of my knowledge, this is the first paper examining delegated decision-making in the context of a structural model.

Lending is a particularly interesting setting in which to study delegated expertise for a number of reasons. First, the costs and benefits of delegation to the lender can be objectively quantified using loan profit and salary data. Second, agency costs can have important negative effects on borrowers as well as lenders. For example, Banerjee and Newman (1993) find that a lack of credit could lead to poverty traps, while Fan et al. (2013) and Nanda (2008) find that financial constraints for commercial lending also negatively impact firm entry, profit, and survival. And third, lenders have a clear alternative. Since the 1980’s, lending markets worldwide have undergone drastic transformations from interview-based to risk-based pricing (Johnson 2004). Today, while loan officers are still used in mortgage markets (Tzioumis and Gee 2013), business financing (Agarwal and Ben-David 2014), and consumer loans (Karlan and Zinman 2009), some products such as revolving credit exclusively use risk-based pricing. As a result, credit scoring provides a natural benchmark against which to assess the value of delegation.

I preview three results. First, there is substantial heterogeneity in risk aversion, overconfidence, and cost of effort across loan officers. This leads to large differences in average loan amounts and resulting profits. Second, loan officers increase annual profits by 147,000 RMB. Karlan and Zinman (2010) and Morduch (1998) also find negative impacts on job retention, income, and mental outlook from insufficient credit. However, the welfare effects are more ambiguous when examining certain kinds of high APR loans in developed countries. For example, Melzer (2011) finds that use of US payday lending leads to increased difficulty in repaying household bills.
per loan officer compared to a counterfactual that maximizes profits conditional only on the borrower’s codified information. Given an average annual compensation of 45,000 RMB, loan officers thus contribute roughly three times their pay in additional annual profits. Lastly, I find that counterfactuals eliminating the distortions caused by the loan officer’s preferences lead to 224,000 to 448,000 RMB higher annual profits compared to the codified information only case.

These results provide an empirical basis for models of delegated expertise. While Demski and Sappington (1987), Bhattacharya and Pfleiderer (1985), and others have examined the theory (see Stracca 2006 for a survey), empirical work has focused almost exclusively on testing comparative static predictions (see Gruber 1996 or Brown et al. 1996) and not modeling primitives. Misra and Nair (2011) and Paarsch and Shearer (2009) estimated structural models of behavior in the broader principal agent literature, but were largely concerned with the effort policy function and not with additional forms of heterogeneity. Einav et al. (2013) and Edelberg (2006) have found compelling evidence of the effectiveness of credit scoring compared to loan officers alone. However, the comparison considered here is not to eliminate the machines, but to examine the additional value of experts working with these machines. I also contribute to the literature that uses econometric tools to optimize firm decisions. Cho and Rust (2008) and Mantrala et al. (2006) used structural models to find profitable policy counterfactuals for auto firms, while Agarwal and Ben-David (2014) and Cole et al. (2013) used experimental designs to increase loan officer effort and lender profits. Authors have also tried to find evidence of behavioral overconfidence. Malmendier and Tate (2004) and Barber and Odean (2001) pre-classify agents as overconfident, while Grubb (2009) and Conlin et al. (2007) take model predictions to aggregate data. By examining the initial loan rating and ex-post performance jointly, I can develop a measure of overconfidence for each loan officer without presupposing its existence. In addition, the possible offsetting interaction between overconfidence and risk aversion is explicitly modeled.2

In the next section, I describe the data and context in more detail. In Section 3 I develop the model incorporating a risk neutral lender, borrowers with a large demand for borrowing, and utility-maximizing loan officers. Loan officers choose an effort level for screening and subsequently observe a noisy signal about the borrower. The loan officer then chooses a loan size. Section 4 presents the estimation strategy and intuition for identification. In Section 5, I use the estimated parameters to evaluate counterfactual scenarios. I first examine automated lending and then move on to counterfactuals that eliminate the distortions due to the loan officer’s preferences. Section 6 concludes.

---

2Goel and Thakor (2008) find that models that do not separately account for risk aversion and overconfidence may confound identification of both.
2 Lending Environment

I study a large Chinese lender with 46 sales offices spread across the country. The lender offers unsecured cash loans to households and small businesses.\(^3\) Average loan sizes are significantly higher than the micro loans studied in the development literature and are slightly less than half of the average borrower’s annual salary income. Self-reported loan purposes range from weddings to home appliances to restaurant furnishings. The borrowing population is not financially at-risk, and has access to other financing options such as credit cards, home and vehicle loans, and other cash-based lenders.\(^4\) The credit card market in China is characterized by low rates of merchant take-up and high transaction fees. As a result, there has been a large growth in popularity of cash-based lending in China. See Ayyagari et al. (2010) for a more detailed survey of the Chinese financing industry.

2.1 Data Description

Table (1) presents some summary statistics for the data. The data sample covers loans made from December 2011 to January 2014 and includes 31,954 borrowers with application and repayment data. Because some loans have not yet completed, the repayment data is censored for about 22,000 borrowers. I discuss methods to accommodate data censoring in Section 4. Average loan sizes are about 33,450 RMB, which is roughly $5,400 at current exchange rates. These loan amounts are substantial both in absolute terms and as a proportion of average salary income of 71,000 RMB.\(^5\)

Each loan product is advertised with an APR and payment length, and products are offered across locations based on local laws and the competitive landscape. New products are introduced and retired frequently at each of the various locations. Loan lengths range from 12 to 36 months with the median at 24 months. The average APR is 48% with a large portion of the interest coming from application fees, maintenance fees, priority rush fees, early repayment, and late

\(^3\)In many parts of the developing world including China, the formal differences between small businesses and households may be small. Morduch (1998) finds that small business loans are often used for consumption smoothing as well as investment purchases. The lender’s underwriting differences between the two segments is minimal. Lending to state owned enterprises and large firms is primarily handled by traditional banks.

\(^4\)Broecker (1990) finds that competition among different lenders could decrease the average credit-worthiness of a lender’s portfolio through adverse selection. In this environment, this concern is somewhat alleviated by credit agency reporting. Some of the reported information include credit reports, credit applications, and external loan terms.

\(^5\)Beyond direct deposited salary income, the lender counts many sources of additional income such as non-deposited salary, social security payments, business income, housing assistance, tax payments, and others. Detailed asset information such as housing, vehicles, and insurance policies are also collected to estimate net worth. There is also separate income accounting for certain worker types such as specialized employees or government workers whose primary compensation may be through reimbursement. The result is that individuals with low amounts of stated payroll may still have large sources of income and are approved for large loans. For comparison, per capita GDP in Shanghai is 82,000 RMB, and is 38,000 RMB in China as a whole.
### Table 1: Summary Statistics

<table>
<thead>
<tr>
<th>Loan Terms</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan Amount (000’s RMB)</td>
<td>33</td>
<td>5</td>
<td>60</td>
<td>12</td>
</tr>
<tr>
<td>Requested (000’s RMB)</td>
<td>124</td>
<td>3</td>
<td>300</td>
<td>103</td>
</tr>
<tr>
<td>Monthly Payment (000’s RMB)</td>
<td>2.2</td>
<td>0.3</td>
<td>5.7</td>
<td>1.1</td>
</tr>
<tr>
<td>Payment Length (Months)</td>
<td>25</td>
<td>12</td>
<td>36</td>
<td>7</td>
</tr>
<tr>
<td>APR (%)</td>
<td>48%</td>
<td>33%</td>
<td>62%</td>
<td>7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Borrower Characteristics</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Assets (000’s RMB)</td>
<td>587</td>
<td>1</td>
<td>7,358</td>
<td>1,337</td>
</tr>
<tr>
<td>Salary Income (000’s RMB)</td>
<td>71</td>
<td>4</td>
<td>642</td>
<td>267</td>
</tr>
<tr>
<td>External Debt (000’s RMB)</td>
<td>160</td>
<td>0</td>
<td>1,923</td>
<td>856</td>
</tr>
<tr>
<td>Age</td>
<td>38</td>
<td>18</td>
<td>58</td>
<td>9</td>
</tr>
<tr>
<td>Credit Card Utilization</td>
<td>41%</td>
<td>0%</td>
<td>100%</td>
<td>38%</td>
</tr>
<tr>
<td>Credit Card Limit (000’s RMB)</td>
<td>20</td>
<td>0</td>
<td>718</td>
<td>156</td>
</tr>
<tr>
<td>Proportion w/ Credit Card</td>
<td>76%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion Female</td>
<td>28%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. APR is inclusive of fees. Financial variables are from verified credit reports. Estimated assets include non-payroll sources, social security payments, vehicle and other durable goods as well as business income. Debt is the sum of the external debt load including credit, housing, and auto loans as identified by a credit report. As of August 2014, $1 is RMB 6.18.

---

**2.2 Loan Process**

The borrower’s first step is to fill out a loan application which includes identification, demographics, financial documents, and verified references from coworkers, friends, and family. The borrower records the loan amount requested as well as the purpose of the loan. Following the application, local branch employees verify the borrower’s income, debt, and asset information using bank statements and credit reporting. The employees will also make home and workplace inspections to assess the borrower’s environment. Data is collected on the number of TV sets, air conditioning units, square footage, and other details. One purpose of these somewhat unorthodox inspections is to find signs of flight risk or fraud.

The process collects a large amount of hard and soft information. I follow Petersen (2004) in defining hard information as quantitative, codified, and with no ambiguity in interpretation. Examples of hard information include age, income, occupation, number of TV sets, and the payment fees.\footnote{Karlan and Zinman (2009) examine a similarly positioned unsecured cash-based lender in South Africa and find APR rates of over 200%. The overall default rate for the South African lender is 30% for first time borrowers. The studied lender’s overall default rate is less than 10%. The average APR for US credit cards is 15%.

---

- **Petersen (2004)**
- **Karlan and Zinman (2009)**
borrower’s measure of credit quality from credit scoring. Soft information on the other hand is either difficult or costly to codify. Two loan officers may also have honest disagreements when interpreting soft information such as written notes, photographs, or the feel of the home. Since I observe the exhaustive set of codified borrower data, I further define soft information as the uncodified data available to the loan officer but not to the lender or myself.

After the sales office excludes high risk borrowers such as those with incorrect identification or unverified references, the application is sent to the central office for loan underwriting. Due to incomplete reporting standards from the different sales offices, I do not have information on rejections that happen before reaching the central office. Once the central office receives the borrower’s file, it is randomly assigned to a loan officer. In practice, a batch of applications will enter the office and be distributed to loan officers without presorting. This assignment procedure will be crucial in identifying the borrower repayment function. See Section 4 for evidence supporting random assignment.

During screening, the loan officer may wish to verify additional details with calls to the sales office or the borrower. Once screening is complete, the loan officer chooses a loan amount given the interest rate and payment length. If this loan size is larger than the borrower’s requested amount, then the full amount requested is approved. If the loan size is smaller, then the borrower is underfunded. Once the determination is complete, the borrower may sign the terms of the loan with no further recourse for adjustment. The sales offices themselves play no part in adjusting loan terms. The entire process from initial application to loan disbursement can take between 3 to 7 days.

The lender utilizes a variety of tools to incentivize repayment. Subsequent loans may come with more attractive terms including lower fees and reduced APR. Additional carrots also come from higher future loan amounts and a simpler approval process. Sticks may be collection calls from external collectors or litigation. Accounts delinquent for more than 90 days or 3 payment periods are generally packaged and sold to collection agencies. The loan officer has no contact with the borrower after loan origination.

2.3 Borrower Liquidity Demands

One unique feature of the application is that borrowers are asked to request a loan amount. Table (1) shows that this value is generally three times the approved amount with a large amount of variation. Figure (1) shows the average requested amount and approved amount for loans by credit quality, which is an internal measure of borrower quality from credit scoring. Higher values

---

7 Some loan products such as high net worth lending, college credit, or rapid turnaround loans do have specialized loan officers for screening. For the loan products in my sample, no additional specialization occurs during distribution.

8 While there is no explicit upper bound on loan size, the highest value I observe is 60,000 RMB. Loan officers may need to acquire supervisor approval for these large amounts. Ghosh et al. (2013) examine a model where the delegated pricing authority of the agent depends on the agent’s local knowledge.

9 This lender was established in the last 4 years and over 98% of loans are made to first-time borrowers. Because secondary loans may come with additional information, I restrict all of my analysis to a borrower’s first loan.
Notes: Sample includes 282 first-time borrowers from June 2012 examined by 8 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Borrower applications are randomly assigned to loan officers. Credit quality is an internal measure of borrower quality from credit scoring - higher values indicate safer borrowers. For each level of credit quality, the average loan amount and requested amount is averaged over a bandwidth interval of 5.5. As of August 2014, $1 is RMB 6.18.

of credit quality indicate safer borrowers. Over 90% of borrowers were approved for loan amounts smaller than requested. The underwriting department states that their goal is to maximize loan profits based on the borrower’s repayment ability rather than trying to satisfy liquidity demands.

While such gaps could be the result of a strategic game where borrowers request larger loans and anticipate underfunding, it is unlikely to explain all of the variation. Early repayment fees and immediate first month payment make the costs of excessive borrowing non-trivial. Also, the average borrower may not be able to predict the loan officer’s own lending preference. While a small proportion of borrowers reject loans because the amounts were too small, there were no cases where borrowers refused a loan for being too large. This gap is then suggestive of a large demand for borrowing. I interpret this as evidence that borrowers are willing to take on additional loan amounts, and that the constraint on loan sizes occurs on the lender side.\footnote{Adams et al. (2009) examine a model where credit rationing occurs due to information asymmetries between the borrower and the lender. They find that moral hazard can cause the lender to restrict loan sizes. In addition to this effect, the model in this paper will attribute some of the constraints to loan officer characteristics.}
3 Model of the Loan Officer

In this section, I develop a delegated expertise model featuring a risk neutral lender, borrowers with a large demand for borrowing, and utility-maximizing loan officers. Borrowers desire as much credit as the company can provide, and this willingness does not differ by borrowers. The lender’s objective is to maximize expected profit, and in the absence of loan officers would approve identically-sized loans to each borrower conditional on the observed borrower information. However, the lender may do better by delegating to loan officers.

The loan officer wishes to maximize his expected utility from lending by first choosing a screening effort level, observing a signal about the borrower, and then deciding on a utility-maximizing loan amount. Since borrowers apply for an advertised interest rate and payment length, I take the loan terms beyond loan size as given. Modeling the borrower’s choice of loan products and endogenizing the interest rate is beyond the scope of this paper. Loan officers will be heterogeneous across three dimensions: risk aversion, overconfidence, and cost of effort. I first introduce the model without overconfidence, and then I address why this behavioral factor is necessary before completing the full exposition.

3.1 Borrower’s Repayment Function and Credit Scoring

When borrower $i$ receives a loan of size $L_i$, he will ultimately repay the proportion $\eta^*_i$ of the total monthly payments. Total repayment to the lender is then $RL_i \eta^*_i$ where $R$ is the interest rate of the loan.\(^{11}\) $\eta^*_i$ is inclusive of penalty fees so that ultimately, some borrowers may return to the lender more than $RL_i$.

Repayment $\eta^*_i$ is determined by four parts: loan amount $L_i$, hard information $x_i$, soft information $u_i$, and an exogenous error $\epsilon_i$. The hard information $x_i$ contains all of the codified borrower data and loan terms such as payment length, APR, date, and city. The function $m(L_i, x_i)$ represents the mean repayment rate conditional on loan amount and hard information. I model the effect of credit scoring as the lender and loan officers can both observe $m(L_i, x_i)$ without ambiguity.\(^{12}\) The partial effects of $m(L_i, x_i)$ will be estimated from the data with no constraints imposed on the signs.

Soft information $u_i$ symbolizes uncodified information about the borrower such as ability, personality, and job prospects. The exogenous error $\epsilon_i$ represents unpredictable post-origination

---

\(^{11}\)For a fixed stream of monthly payments, the present discounted value of a loan if repaid in full is given by $PV_i = \frac{apr}{1-(1+apr)^{-N}} \times \frac{1-(1+i)^{-N}}{1-i}$ $L_i = RL_i$ where $apr$ is the monthly APR of the loan, $N$ is the payment length, and $i$ is the discount rate. The interest rate of the loan is given by $R = \frac{apr}{1-(1+apr)^{-N}} \times \frac{(1+i)^{-N}}{1-i}$. $i$ is set at the People’s Bank of China (PBC) base interest rate of 6% annually. The monthly payment $MP_i$ is given by $MP_i = \frac{apr}{1-(1+apr)^{-N}} \times \frac{1}{1-i}$. The Chinese inflation rate is between 3 and 6%.

\(^{12}\)This understates the value of loan officers in two ways. One is that the $m(L_i, x_i)$ contains variables not used in credit scoring such as time controls, higher order terms, interacted terms, and fixed effects. The other is that loan officers are only needed to interpret soft information. To the extent that loan officers have additional value to the lender, the bias will only increase the value of delegation.
shocks such as job loss, health incidents, or particularly effective loan collectors. Soft information is distributed $N(0, \sigma_u^2)$ and the exogenous error is distributed $N(0, \sigma_\epsilon^2)$. The two terms are further assumed by construction to be mean independent. The loan officer’s loan size choice $L_i$ will be endogenously determined and correlated with soft information $u_i$. Combining the mentioned factors, the borrower will repay the proportion

$$\eta^*_i = m(L_i, x_i) + u_i + \epsilon_i$$

of the total monthly payments. Loan profits are given by $\pi_i = R L_i \eta^*_i - L_i$.

### 3.2 Loan Officer’s Objective

Figure (2) illustrates the steps of the loan officer’s problem. Loan officers first choose a screening effort, observe a noisy signal, and then decide on a utility-maximizing loan amount. Individual loan officers indexed by $j$ have utility over compensation given by a constant absolute risk aversion (CARA) exponential utility function $u(y_j) = -e^{-r_j y_j}$. $r_j$ is a heterogeneous risk aversion parameter, and $y_j$ is the loan officer’s compensation net of effort costs. The loan officer’s annual salary averages 42,000 RMB a year and also includes a 3 to 4,000 RMB bonus that is contingent on loan performance. I model the compensation with a base salary and a bonus proportional to total loan profits so that compensation is given by $\alpha + \beta \sum \pi_{ij}$.\(^{13}\)

In practice, the actual compensation scheme is most likely non-linear with limited liability constraining downward risk, total bonuses upward constrained by overall lender performance, and deferred compensation over multiple calendar periods (Cole et al. 2013). While the proportional bonus is admittedly an approximation, the structure is consistent with a reduced form model of career concerns or corporate obedience where the loan officers are incentivized to maximize loan profits. Conversations with the lender, supervisors, and the loan officers themselves all reveal that maximizing loan profits is the objective. The loan officers themselves believe that .5% to 1% of loan profits accrue as the year-end bonus.

Net compensation is given by $y_j = \alpha + \beta \sum \pi_{ij} - \text{cost}_j (\sigma_j^2)$ where $\text{cost}_j (\sigma_j^2)$ is the disutility in money of the screening effort level. Since $\sigma_j^2$ determines the variance of a noisy screening signal, lower values of $\sigma_j^2$ indicate higher effort and greater precision. $\text{cost}_j (\sigma_j^2)$ is assumed to be convex.

\(^{13}\)The lender is vague about the bonus structure even to the loan officers themselves. One reason may be to avoid loan officers exploiting the scheme. See Berg et al. (2013) for evidence that loan officers may manipulate information to pass an approval threshold and increase their pay.
and strictly decreasing with \(\text{cost}_j \left(\sigma_j^2\right) > 0\) for any \(\sigma_j^2 > 0\), \(\text{cost}'_j \left(\sigma_j^2\right) < 0\), and \(\text{cost}''_j \left(\sigma_j^2\right) > 0\). Since the total portfolio profit \(\sum \pi_{ij}\) is linear in individual profits and each borrower’s stochastic repayment \(u_i + \epsilon_i\) is uncorrelated, maximizing utility from total loans is equivalent to maximizing utility from each individual loan given the exponential utility function. Although the average loan officer is screening 700 applications a year, he may still have reason to be risk averse over individual loans. For example, a defaulted loan may warrant extra scrutiny regardless of the performance of the remaining portfolio. Combining everything, a loan officer tries to maximize expected utility given by

\[
EU \left[ L_{ij} \right] = \int -e^{-r_j \left(\alpha + \beta \pi_{ij} - \text{cost}_j \left(\sigma_j^2\right)\right)} f \left(\pi_{ij}\right) d\pi
\]

(2)

### 3.3 Screening Soft Information

Screening yields a noisy signal \(\omega_{ij}\) about the borrower’s soft information \(u_i\) such that \(\omega_{ij} = u_i + \delta_{ij}\) where \(\delta_{ij}\) is an additive noise term that is distributed \(N \left(0, \sigma_j^2\right)\) and mean independent of \(u_i\). The precision of this signal depends on the loan officer’s effort level where lower values of \(\sigma_j^2\) indicate a more precise signal. Heterogeneity in screening effort will be driven by a loan officer specific cost of effort parameter \(d_j\) embedded in the cost of effort function \(\text{cost}_j \left(\sigma_j^2\right)\). The motivation is that some loan officers may be more effective at screening than others because of experience or natural ability.

Observing \(\omega_{ij}\), the loan officer then develops a posterior belief about the distribution of soft information and exogenous error \(u_i + \epsilon_i\). This posterior is distributed

\[
u_i + \epsilon_i|\omega_{ij} \sim N \left(\frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2}\omega_{ij}, \frac{\sigma_u^2\sigma_j^2}{\sigma_u^2 + \sigma_j^2} + \sigma_\epsilon^2\right)
\]

(3)

with a non-zero mean and a smaller variance than what the lender observes. Since the conditional variance is decreasing with greater screening effort and a lower \(\sigma_j^2\), loan officers exerting greater effort can sharpen their signal.

### 3.4 Optimal Loan Size and Effort

When choosing the optimal loan size, the loan officer has to balance two forces. One is that higher amounts of \(L_{ij}\) directly increase the repayment value \(RL_{ij}\) when the loan is repaid in full. However, increasing \(L_{ij}\) may change the fraction repaid \(\eta_i^*\). The marginal effect of loan size could impact repayment due to factors under the borrower’s control as well as circumstance. For example, borrowers may have a greater incentive to default on larger loans through strategic default or simply because larger loan payments have a greater chance of pushing a borrower into delinquency. It is important to note that I do not directly model the borrower’s default decision.

---

14 This follows from the conditional distribution of the normal distribution. The covariance between \(u_i + \epsilon_i\) and \(\omega_{ij}\) is \(\sigma_u^2\) since \(u_i\) and \(\delta_{ij}\) are assumed to be mean independent.
The estimated marginal effect can be considered a reduced form treatment of the borrower’s optimal response to greater loan amounts.\textsuperscript{15}

Figure 3: Expected Loan Profit for Average Borrower

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{PDV of Net Profit (000's RMB) vs Loan Amount (000's RMB)}
\end{figure}

\textbf{Notes:} Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Borrower repayment parameters are from the main MLE specification. The average borrower is constructed by setting all covariates to their average values for a 24 month loan with an APR of 48%. The discount rate is set to the People’s Bank of China base interest rate of 6%. As of August 2014, $1 is RMB 6.18.

Figure (3) shows the expected profit estimated from the parametrized model as a function of the loan amount for a 24 month loan at an APR of 48%. At small loan amounts, increases in loan size increase profits. Past the maximum, further increases will decrease profits as the fraction repaid decreases. While the model does not incorporate unobserved heterogeneity in $\partial n^* / \partial L$, the parametrized function will allow $\partial n^* / \partial L$ to vary according to payment lengths and APR. This allows the marginal effect for two otherwise identical borrowers to differ because of different loan terms.\textsuperscript{16} After parameterizing the mean repayment $m(L_{ij}, x_i)$, the optimal loan $L^*_{ij}$ can be solved for as a function of the loan officer’s structural parameters and the observed signal.\textsuperscript{17}

\textsuperscript{15}This effect is also known as borrower moral hazard. See Gine et al. (2012) for a model that explicitly outlines the borrower’s repayment decision through a private action.

\textsuperscript{16}Adams et al. (2009) find that some households are even more responsive to down payments than the borrowed amount indicating that accounting for short term liquidity effects such as monthly payments is extremely important.

\textsuperscript{17}This loan size can also be zero, but this rarely happens since by the time the loan officer examines the file, the borrower’s application has already passed inspections at the sales office level.
The observed signal is itself the output of a prior screening effort decision. Balancing the cost of effort with increased precision, the loan officer must choose a utility-maximizing level of screening effort for all of his loans. Solving backwards, the loan officer’s optimal loan size given the signal is \( L_{ij}^* (\omega_{ij}) = \arg \max E \left[ L_{ij} \right] \) and therefore the value function is \( E \left[ L_{ij}^* (\omega_{ij}) \right] \). The loan officer chooses the screening effort level \( \sigma_j^2 \) to maximize this expectation over the distribution of \( \omega_{ij} \) and \( x_i \).

Specifically, the loan officer is solving the maximization given by

\[
\sigma_j^{2*} = \arg \max \int \int E \left[ L_{ij}^* (\omega_{ij}) \right] f \left( \omega_{ij} | \sigma_j^2 \right) f \left( x_i \right) d\omega d x
\]

While a more robust model may exhibit variation in screening effort \( \sigma_j^2 \) across different levels of hard information \( x_i \), the constant effort level here allows for a more tractable estimation strategy. This interpretation is also consistent with a story where loan officers have many borrower applications to screen and limited time. Before even opening the application and observing the borrower’s hard information, the loan officer must decide on some amount of time to spend on each file.

### 3.5 Parametrized Model

I first parametrize the borrower’s repayment function to be \( m(L_{ij}, x_i) = \gamma_L L_{ij} + \gamma_{MP} MP_{ij} + \gamma_{PV} PV_{ij} + x_i' \Gamma \) where \( MP_{ij} \) is the monthly payment and \( PV_{ij} \) is the present value of the loan if repaid in full. Conditional on the payment length \( N \), \( apr \), and the discount rate \( i \), the monthly payment and present value terms are linear functions of loan size. This implies that the marginal effect of loan size on repayment can be re-written as

\[
\frac{\partial \eta_i^*}{\partial L} = \gamma = \gamma_L + \gamma_{MP} \frac{\text{apr}}{1 - (1 + \text{apr})^{-N}} + \gamma_{PV} \frac{\text{apr}}{1 - (1 + \text{apr})^{-N}} \frac{1 - (1 + i)^{-N}}{i}
\]

This gives repayment to be

\[
\eta_i^* = \gamma L_{ij} + x_i' \Gamma + u_i + \epsilon_i
\]

where \( \gamma \) varies across different values of APR and payment lengths. Combining equation (6) with the first order condition gives the optimal loan size to be

\[
L_{ij}^* = k_j \left( x_i' \Gamma - \frac{1}{R} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} \right)
\]

where \( k_j = \left[ r_j \beta R \left( \frac{\sigma_u^2 \sigma_j^2}{\sigma_u^2 + \sigma_j^2} + \sigma_u^2 \right) - 2 \gamma \right]^{-1}.^{18} \)

I term \( k_j \) to be the officer effect, which multiplies the effect of hard information \( x_i' \Gamma \) and the loan officer’s posterior belief about soft information \( \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} \). Loan officers with high risk

---

\(^{18}\)Integrating the expected utility relies on the properties of the moment generating function of the normal distribution. See Appendix A for the derivation.
aversion have a larger officer effect, which leads to greater average loan amounts. Similarly, the officer effect increases as borrower repay a larger proportion of their loans or if the exogenous portion of repayment decreases. Without additional information about the compensation scheme, \( \alpha \) is not identified and \( \beta \) cannot be separately identified from \( r_j \beta \).

To solve for the optimal effort level, the expectation of the value function in equation (4) can be approximated using a first order Taylor expansion. Cost of effort is parametrized to be \( \text{cost}_j(\sigma^2_j) = \frac{d^2_j}{\sigma_j^2} \). Taking the first order condition gives the utility-maximizing effort level to be

\[
\sigma^2_j^* = \left( r_j \beta R \sigma^2 - 2\gamma \right) \sigma^2_u \left( \frac{r_j^2}{2} \beta R \left( x'_i \Gamma - \frac{1}{R} \right) \sigma^2_u - r_j \beta R \left( \sigma^2_u + \sigma^2_e \right) + 2\gamma \right)^{-1}
\]

with the derivation in Appendix B.

While difficult to interpret, \( \sigma^2_j^* \) is increasing in the cost of effort parameter \( d_j \) indicating that higher cost of effort loan officers choose higher values of \( \sigma^2_j \). The effect of the bonus rate \( \beta \) on \( \sigma^2_j^* \) is non-linear and non-monotonic.\(^{19}\) Conditional on the other model parameters, the screening effort level \( \sigma^2_j^* \) is isomorphic to the cost of effort parameter \( d_j \) for \( \sigma^2_j > 0 \) and when \( \gamma < 0 \). To simplify the model, I proceed to estimate the screening effort level \( \sigma^2_j^* \) instead of the cost of effort \( d_j \). Estimation will not rely upon the relationship in equation (8). The derived expression can be considered as illustrative for how cost of effort \( d_j \) affects the screening effort level \( \sigma^2_j \). This approach also avoids functional form assumptions for cost of effort as well as the distribution of the borrower’s hard information.\(^{20}\)

### 3.6 Overconfidence

Equation (7) predicts that as loan officers become more risk averse, the officer effect \( k_j \) should decrease leading to smaller average loan amounts. Since expected profit should be higher for risk neutral loan officers, loan officers with smaller average loan amounts should be associated with lower loan profits. Similarly, the variance of \( L^*_{ij} \) is decreasing in \( \sigma^2_j \). This means that the loan officers exerting lower effort are those with the smallest variance. Loan officers that have large deviations must have received a precise signal about the borrower. Both of these predictions can be non-parametrically tested, and will serve as the motivation for introducing overconfidence.

---

\(^{19}\)Under certain conditions, Holmstrom and Milgrom (1987) established that the linear compensation scheme is the optimum incentive contract for principal agent models. For delegated expertise models however, Stoughton (1993) and Bhattacharya and Pfleiderer (1985) find that linear contracts suffer from the irrelevance result where the agent’s optimum effort level is not a function of a linear bonus. To see this, note that as \( \gamma \) approaches 0 and the compensation plan becomes linear in loan amounts, the bonus rate \( \beta \) disappears from equation (8). An agent can always undo the incentive effects of a linear scheme because the agent has costless control of a linear action \( L_{ij} \) after realization of the signal (Stracca 2006). Stoughton (1993) instead propose a quadratic contract that asymptotically approximates the optimal incentive scheme. The compensation plan I use is linear in \textit{profits} but quadratic in \textit{loan amounts}. This distinction enables the lender to motivate screening effort and avoid the irrelevance result.

\(^{20}\)Equation (8) is derived for the case of a single borrower type. With additional borrowers, an additional expectation needs to be taken over the distribution of the borrower’s hard information.
Table 2: Effect of Average Loan Amount and Variance on Loan Profits

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: Loan Profits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan officer average loan size</td>
<td>0.315**</td>
</tr>
<tr>
<td>conditional on loan terms and hard</td>
<td>(0.15)</td>
</tr>
<tr>
<td>information.</td>
<td></td>
</tr>
<tr>
<td>Variance of loan officer’s loan</td>
<td>-0.084***</td>
</tr>
<tr>
<td>sizes conditional on loan terms</td>
<td>(0.02)</td>
</tr>
<tr>
<td>and hard information.</td>
<td></td>
</tr>
<tr>
<td>Observations (Uncensored)</td>
<td>8,531</td>
</tr>
</tbody>
</table>

Notes: Standard errors in parenthesis. Data includes 8,531 uncensored first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Average loan size is the coefficient of the loan officer indicator dummy from a regression of loan amount on the full set of controls. Variance is the variance of the residuals from the above regression. As of August 2014, $1 is RMB 6.18.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Table (2) shows the effect of the loan officer’s average loan amount and variance on loan profits. Conditional on the borrower’s hard information, loan officers that have a higher average loan amount have higher profits. This supports the risk aversion prediction that links lower loan amounts and lower profits. However, a higher variance of loan amounts is associated with lower average profits. This motivates an extension of the model that can account for the empirical finding that loan officers with large loan deviations are less profitable. One candidate is that the loan officers are overconfident.²¹

Moore and Healy (2008) describe one form of overconfidence as an excessive precision in beliefs. In lab studies, McKenzie et al. (2008) find that participant’s 90% confidence intervals include the true value only half of the time. I interpret these results to mean that loan officers may be overconfident in their signal, and think that their screening effort is actually \( \sigma^2 \) instead of \( \sigma_j^2 \). While a rational Bayesian believes that soft information and the exogenous error are distributed following equation (3), an overconfident loan officer falsely perceives the posterior to instead be

²¹Beyond overconfidence, some alternative explanations may explain why the variance of loan amounts is inversely related to loan profits. For example, loan officers may be rent-seeking so that the borrowers approved for larger loans are those that offered kickbacks to the loan officers. Alternatively, individual loan officers may have random preference shocks for some borrowers. Fisman et al. (2012) find some evidence that loan officers originate loans at higher rates for borrowers from similar backgrounds. Given the large amount of borrowers, the difficulty in contracting, and the extensive controls on observables, overconfidence may be a more plausible explanation. Note that the random preference shocks must also be correlated with the borrower’s repayment ability so that the the borrowers given larger loans are less likely to repay. If not, then a risk averse loan officer could actually increase his profits with a preference shock to a safe borrower.
\[ u_i + \epsilon_i \omega_{ij} \sim N \left( \frac{\sigma_{\omega_{ij}}^2}{\sigma_v^2 + c_j^2}, \frac{\sigma_v^2 c_j^2}{\sigma_v^2 + c_j^2} + \sigma^2 \right) \]  

(9)

If \( c_j^2 < \sigma_j^2 \), this means that the loan officer is overconfident and believes the signal is more precise than reality. If \( c_j^2 > \sigma_j^2 \), this means that the loan officer must be underconfident since the true variance \( \sigma_j^2 \) is smaller than the belief. The optimal loan amount with overconfidence is given by

\[ L_{ij}^* = k_j \left( x_i \Gamma - \frac{1}{R} + \frac{\sigma_v^2}{\sigma_v^2 + c_j^2} \omega_{ij} \right) \]  

(10)

where \( k_j = \left[ r_j \beta R \left( \frac{\sigma_v^2 c_j^2}{\sigma_v^2 + c_j^2} + \sigma^2 \right) - 2\gamma \right]^{-1} \). Conditional on overconfidence and the other parameters, as \( \sigma_j^2 \) increases in size and loan officers exert less screening effort, the variance of loan sizes will now become larger and loan profits will decrease.

Overconfidence also interacts with risk aversion. Overconfidence increases the officer effect \( k_j \) and hence average loan sizes in a similar mechanism to decreased risk aversion. For a risk averse loan officer, overconfidence counteracts the tendency to choose smaller loan amounts. In some cases with an especially risk averse loan officer, some amount of overconfidence may actually increase the loan officer’s profits. Put another way, the distortion caused by overconfidence may offset some of the effects of risk aversion. This is also why the separate identification of overconfidence apart from risk aversion is difficult (Goel and Thakor 2008), but is not an obstacle here because overconfidence and risk aversion also have different effects on the variance.

4 Estimation Strategy

Equations (6) and (10) jointly determine the borrower’s repayment and the loan officer’s loan decision respectively, and define a model that can be estimated by maximum likelihood. There are three main challenges to estimation: endogenous loan sizes, censored repayment histories, and identification of the loan officer characteristics. I address these in turn.

4.1 Identification of \( \gamma \)

Without an instrument, \( L_{ij} \) in equation (6) is endogenous. This is because loan amounts are correlated with the soft information the loan officer observes. One solution is to use the random assignment process as exogenous variation in loan amounts. The intuition is that identical borrowers are assigned to different loan officers with different preferences. The loan officer’s average loan amounts are functions of their own characteristics but orthogonal to each borrower’s soft information. This means that the differences in repayment across loan officers can be causally attributed to differences in their average loan amount. The necessary assumption is
that $E \left[ u_i + \epsilon_i | \sigma^2_i, r_j, c^2_j, x_j \right] = 0$ indicating that the unobserved components cannot be correlated with the loan officer’s characteristics conditional on hard information.\textsuperscript{22}

### Table 3: First Stage and Test of Correlated Observables

<table>
<thead>
<tr>
<th>21 Loan Officer Indicators</th>
<th>Dependent Variable: Loan Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joint F-Test</td>
<td>17.18 18.91 19.48 19.59 22.57</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.000 0.000 0.000 0.000 0.000</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.373 0.420 0.440 0.482 0.525</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loan Officer Mean Loan Amount</th>
<th>Dependent Variable: Loan Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.155 0.150 0.153 0.154 0.156</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.000 0.000 0.000 0.000 0.000</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.367 0.413 0.433 0.476 0.518</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Controls</th>
<th>78 84 85 129 265</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year by Month, City, Product</td>
<td>✓ ✓ ✓ ✓ ✓</td>
</tr>
<tr>
<td>Application Variables</td>
<td>✓ ✓ ✓ ✓</td>
</tr>
<tr>
<td>Internal Credit Quality</td>
<td>✓ ✓ ✓</td>
</tr>
<tr>
<td>Financial Variables</td>
<td>✓ ✓</td>
</tr>
<tr>
<td>Inspection Variables</td>
<td>✓</td>
</tr>
</tbody>
</table>

Notes: Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Top panel shows the first stage test of correlation between loan officers and loan amounts. Second panel shows the test of correlated observables. Mean loan amount is the loan officer’s average loan amount excluding borrower $i$ or $\frac{1}{N-1} \sum_{i = 1}^N L_{ij}$. Application variables include loan amount requested, loan purpose, and transformations. Internal credit quality is an internal measure of borrower quality from credit scoring. Financial variables include income, wealth, taxes, social security, credit reports, external debt, and transformations. Inspection variables include home financing, home living arrangement, extensive occupation details, home furnishing, tenure in workplace, payment length dummies, and others. As of August 2014, $1$ is RMB 6.18.

This assumption is supported by the random assignment of borrower applications to loan officers. The exclusion restriction is also satisfied since loan officers have no impact on repayment except through the loan amount. Loan collection is handled entirely by the local branch offices with no contact from the original loan officer. A test for the strength of these loan officer instruments is the first stage F-test shown in the top panel of Table (3). The test statistic across all specifications shows that the instruments have sufficient correlation with loan sizes and that there are large differences in average loan sizes across loan officers.

Furthermore, it is not necessary that the assignment be unconditionally random. The identification assumption is still valid as long as the borrowers are randomly assigned conditional on hard information. Loan officers may specialize in borrowers from specific cities or time

\textsuperscript{22}One possible violation is if loan officers collaborate with each other. For example, a loan officer on a particularly hard to read application may consult with others. The data does not allow me to identify collaboration in this way, but the large workload precludes this type of joint inspection from frequently happening.
windows, but cannot specialize in cases with particularly high or low values of $u_i + \epsilon_i$ within those categories. While the lender indicates that applications are only assigned based on time and city, this assumption can be tested.

I follow Maestas et al. (2013) and test for correlated observables to see if loan officers specialize in applications with specific attributes. I construct a loan officer’s average loan amount excluding borrower $i$ as $\frac{1}{N_i-1} \sum_{j} L_{ij}$. By examining the regression of $L_{ij}$ on this variable along with additional covariates, I can test for correlated observables because only borrower characteristics that are correlated with a loan officer’s case mix should change the variable’s coefficient when added to the regression. Table (3) shows the results of this test in the second panel for increasing amounts of hard information. The stable values across specifications provides evidence that loan officers do not specialize in observed borrower types beyond time and city. Extending the logic, this also supports the assumption that borrowers do not specialize within borrower types either.

4.2 Censored Repayment Data

Another empirical challenge is censoring in the repayment data. I do not observe the final repayment proportion $\eta^*_i$ but a censored variable $\eta_i$ because many of the outstanding loans have not yet completed by the end of the study period. Only around 9,500 loans are complete. I account for this by estimating with a Tobit specification. Specifically, if $d_i$ is the current proportion of payments made, then the repayment $\eta_{ij}$ I observe is given by

$$\eta_{ij} = \begin{cases} 
\eta^*_{ij} & \text{if Uncensored} \\
 d_i < \eta^*_{ij} & \text{if Censored} 
\end{cases}$$

(11)

For example, if the study period ends with a borrower making 6 of 12 required payments, then $\eta_{ij} = d_i = .5$ ignoring adjustments due to discounting. The likelihood function accounts for the fact that $\eta^*_i$ is greater than $d_i$ since the lender would not refund payments back to the borrower. Uncensored contracts are those where no further payments are expected and include completed contracts and delinquent borrowers. With the lender engaged in business throughout the year, the censoring point $d_i$ only depends on the date of loan origination and is assumed to not be correlated with $\eta^*_i$ or $L^*_{ij}$.

4.3 Identification of the Loan Officer Characteristics

I first give some intuition for the identification of the loan officer parameters. Since borrower applications are randomly assigned to loan officers, the distribution of soft information and

\footnotesize
\begin{itemize}
  \item Given the volume of cases, it would be difficult and impractical for a pre-screener to give the borrowers with lower than expected $u_i + \epsilon_i$ to certain loan officers. The lender asserts that borrowers are not pre-sorted beyond time and city since applications from the same sales office are frequently batched together.
  \item I rely on the lender’s definition of a defaulting loan being 90 days past due or three payment periods. After the 90 day mark, the loans are usually packaged and sold to collections agencies. In some cases, the delinquent loans go to litigation.
\end{itemize}
Figure 4: Average Loan Amount for Selected Officers by Credit Quality
24 Month Loans at 48% APR from June 2012

Notes: Sample includes 282 first-time borrowers from June 2012 examined by 8 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Borrower applications are randomly assigned to loan officers. Credit quality is an internal measure of borrower quality from credit scoring - higher values indicate safer borrowers. For each level of credit quality, the loan officer’s average loan amount is averaged over a bandwidth interval of 5.5. As of August 2014, $1 is RMB 6.18.

exogenous error $u_i + \epsilon_i$ across loan officers is the same. If all loan officers were identical, then conditional on the hard information $x_i$, the mean and variance of loan sizes across all loan officers should be constant. However, heterogeneity in risk aversion leads to differences in average loan amounts. More risk averse loan officers have a smaller officer effect $k_j$, which leads to lower average loan sizes. This is graphically depicted in Figure (4) where Officer A has approved a larger average loan amount for every level of the lender’s internal measure of credit quality from credit scoring. Conditional on the other parameters, the model would indicate Officer A to be less risk averse than Officer B.

As loan officers exert greater screening effort, the variance of loan sizes decreases conditional on the other parameters. Identification then comes from the variance of loan sizes. This is highlighted in Figure (5) where Officer C has a smaller variance than Officer D. Conditional on the other parameters including overconfidence, then Officer C must be exerting greater screening effort.

Identifying overconfidence requires both the initial assessment and the outcome. Two loan of-
Notes: Sample includes first-time borrowers from June 2012 examined by 2 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Borrower applications are randomly assigned to loan officers. Credit quality is an internal measure of borrower quality from credit scoring - higher values indicate safer borrowers. For each level of credit quality, the loan officer’s loans are plotted with a 95% confidence interval. As of August 2014, $1 is RMB 6.18.

Loan officers equal in every dimension except overconfidence should have different repayment outcomes. For the same loan amount, the loan officer that is overconfident must have lower average profits. Figure (6) illustrates that Officer E’s loan portfolio has a much higher level of delinquency than Officer F. If this higher delinquency level corresponds to lower loan profits, then Officer F must be less overconfident. The strength of this approach is that overconfidence and screening effort is identified jointly from the loan officer’s initial loan amount and the ex-post loan outcome.

4.4 Identification of Soft Information and Exogenous Error

The variance of loan repayment identify the sum $\sigma_u^2 + \sigma_\epsilon^2$, but not the individual components. This is because there is no observable source of variation affecting one and not the other. To address this, I use a change of variables that allows me to estimate the model parameters as

---

25 If there was a loan officer that could perfectly observe the soft information $u_i$, then the variance of that loan officer’s loan amounts would be able to separate $\sigma_u^2$ and $\sigma_\epsilon^2$. For this reason, the counterfactual exercises cannot create a loan officer that has a perfectly accurate signal.
Figure 6: Loan Defaults for Selected Officers by Credit Quality
24 Month Loans at 48% APR from June 2012

Defaulted Borrower

Officer E

Riskier <= Credit Quality => Safer

Officer F

Riskier <= Credit Quality => Safer

Notes: Sample includes first-time borrowers from June 2012 examined by 2 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Borrower applications are randomly assigned to loan officers. Credit quality is an internal measure of borrower quality from credit scoring - higher values indicate safer borrowers. Defaulted loans are delinquent for more than 2 payment periods. As of August 2014, $1 is RMB 6.18.

functions of the soft information $\sigma_\theta^2$. This change rewrites the model using a reduced set of variables and an identical likelihood function.

Instead of estimating the full set of parameters $\left(\Gamma, \gamma, r_j, \sigma_j^2, c_j, \sigma_u^2, \sigma_\epsilon^2\right)$, I instead estimate the reduced set of $\Gamma, \gamma, \hat{r}_j = r_j\beta$, $\tilde{c}_j^2 = \frac{c_j^2}{\sigma_\theta^2} + \frac{1}{\sigma_\theta^2} - 1$, $\tilde{\sigma}_j^2 = \frac{\sigma_j^2}{\sigma_\theta^2} + \frac{1}{\sigma_\theta^2} - 1$, and $\tilde{\sigma}_\epsilon^2 = \sigma_u^2 + \sigma_\epsilon^2 - 1$. I refer to these new parameters as adjusted versions. The adjusted risk aversion is equal to the product $r_j\beta$ and is not a function of $\sigma_\theta^2$ indicating that risk aversion cannot be separated from the bonus rate. The change of variables still preserves comparability between overconfidence and screening effort. For example, $\tilde{c}_j^2 < \tilde{\sigma}_j^2$ still implies that the loan officer is overconfident, and $\tilde{c}_j^2 > \tilde{\sigma}_j^2$ still indicates underconfidence. While I do not get an estimate of the magnitude of $\sigma_u^2$ versus $\sigma_\epsilon^2$, this change of variables still allows me to make comparisons between loan officers and perform counterfactual exercises.
4.5 Maximum Likelihood

Deriving the actual likelihood function is straightforward. For ease of exposition, I write the likelihood in terms of the full set of parameters and not the adjusted set. Appendix C contains the proof of their equality. The likelihood of observing $L^*_{ij}$ and $\eta_i$ conditional on the censoring point $d_i$ and hard information $x_i$ is $\mathcal{L}(L^*_{ij}, \eta_i | x_i, d_i)$. Bayes’ rule expands this to the product $\mathcal{L}(L^*_{ij} | x_i, d_i) \times \mathcal{L}(\eta_i | L^*_{ij}, x_i, d_i)$. Since the stochastic element in both pieces is additively normal, the individual likelihoods can be written in the familiar form

$$
\mathcal{L}(L^*_{ij} | x_i, d_i) = \frac{1}{\sigma_v} \phi \left( \frac{L_{ij} - w_{ij}}{\sigma_v} \right)
$$

$$
\mathcal{L}(\eta_i | L^*_{ij}, x_i, d_i) = \begin{cases} 
\frac{1}{\sigma_{u+e|v}} \phi \left( \frac{\eta_i - h_{ij}}{\sigma_{u+e|v}} \right) & \text{if Uncensored} \\
\Phi \left( \frac{h_{ij} - d_i}{\sigma_{u+e|v}} \right) & \text{if Censored}
\end{cases}
$$

where $w_{ij} = x'_i \Gamma k_j - \frac{k_j}{R}$, $\sigma^2_v = k^2_j \frac{\gamma^4 \sigma^2_u + \sigma^2_e}{\sigma^2_u + \sigma^2_e}$, $h_{ij} = \gamma L_{ij} + x'_i \Gamma + \frac{\sigma^2_u + \sigma^2_e}{\sigma^2_u + \sigma^2_e} (L_{ij} - w_{ij})$, and $\sigma^2_{u+e|v} = \frac{\gamma^2 \sigma^2_u + \sigma^2_e}{\sigma^2_u + \sigma^2_e}$. This likelihood function can be thought of as a Tobit model with an endogenous regressor and an observation-specific censoring point.

5 Results and Counterfactuals

I present the estimates for the adjusted model parameters $(\Gamma, \gamma, \gamma_e, c^2_j, \sigma^2_j, \sigma^2_e)$. The parameters allow me to construct a lending model without loan officers and based entirely on hard information. I can also examine counterfactuals designed to mitigate the costs of risk aversion, overconfidence, and low screening effort. Each counterfactual along with the status quo is compared to the hard information only case to determine the value of delegation.

5.1 Borrower Repayment Function

Table (4) lists some of the estimates for the empirical repayment function $\eta^*_i$. Following equation (5), loan size, monthly payment, and the present value of full repayment all affect repayment through $\gamma$. Higher borrowed amounts decrease repayment, while a higher monthly payment increases the proportion returned. With repayment rates around 85%, it is not surprising that higher monthly payments lead to greater loan recovery. While the individual $\gamma$ estimates are difficult to compare because of different interest rates and payment lengths, $\gamma$ for the average borrower with a 24 month loan and an APR of 48% is -.5%. In other words, an increase of 1,000 RMB in loan amount leads to the borrower repaying -.5% less of the total monthly payments.26

26The magnitude can be compared to the literature, although the comparison is not straightforward since most studies investigate the impact on binary default and not repayment. In addition, differences in institutional setting make cross-country comparisons difficult. Despite the difficulties, the estimates are roughly consistent. In the US subprime auto financing market, Adams et al. (2009) find that a 1% increase in loan amounts leads to a 1.6% increase in the chance of default. By running the same specification with default instead of repayment, I find in Wang (2014) that a similar 1% increase in loan size leads to a .6% increase in the chance of default. Other estimates of $\gamma$ find
Table 4: Estimates of the Borrower Repayment Function

<table>
<thead>
<tr>
<th>Loan Terms</th>
<th>Coefficient</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan Amount</td>
<td>-0.016</td>
<td>0.000</td>
</tr>
<tr>
<td>Monthly Payment of Loan</td>
<td>0.038</td>
<td>0.000</td>
</tr>
<tr>
<td>Present Discounted Value of Loan</td>
<td>0.006</td>
<td>0.000</td>
</tr>
<tr>
<td>Variance of soft information and error $\sigma_u^2 + \sigma_{\epsilon}^2$</td>
<td>0.115</td>
<td>0.000</td>
</tr>
</tbody>
</table>

γ for a 24 month loan with an APR of 48%.

<table>
<thead>
<tr>
<th>Selected Hard Information</th>
<th>Coefficient</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salaried Income</td>
<td>1.6e-5</td>
<td>0.000</td>
</tr>
<tr>
<td>External Debt</td>
<td>1.1e-6</td>
<td>0.059</td>
</tr>
<tr>
<td>Credit Card Utilization</td>
<td>-0.008</td>
<td>0.101</td>
</tr>
<tr>
<td>Management Position</td>
<td>0.024</td>
<td>0.000</td>
</tr>
<tr>
<td>Female Indicator</td>
<td>0.002</td>
<td>0.071</td>
</tr>
<tr>
<td>Amount Requested</td>
<td>1.7e-4</td>
<td>0.000</td>
</tr>
<tr>
<td>Dorm or Rental Indicator</td>
<td>-0.019</td>
<td>0.000</td>
</tr>
<tr>
<td>Credit Quality</td>
<td>1.0e-5</td>
<td>0.838</td>
</tr>
<tr>
<td>Air Conditioner Units</td>
<td>0.004</td>
<td>0.001</td>
</tr>
<tr>
<td>Met Family/Coworkers</td>
<td>0.008</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Notes: Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Borrower repayment parameters are from the main MLE specification. The monthly payment varies with payment length and APR. PDV is the total repayment if the loan is repaid in full and includes fees. Additional controls include year by month, city, product fixed effects, and additional application, financial, and inspection variables. As of August 2014, $1 is RMB 6.18.

To put this into context, the average loan in the sample is 34,450 RMB for 24 months at an APR of 48%. This gives a monthly payment of 2,190 RMB with a total present value of 49,500 RMB if repaid in full. By extending an additional 1,000 RMB to this borrower, repayments decrease by 247 RMB for the inframarginal loan amounts. Profits also increase through the additional interest associated with the higher loan amount. These two effects can be seen in Figure (3). Note that the lender’s average loan amount is just short of the profit-maximizing level indicating that increases in loan amount may be profitable. Karlan and Zinman (2010) attribute marginal loan profitability to overly conservative risk assessment. I take this as further evidence that loan officers as a whole are risk averse.

The variance estimate of $u_i + \epsilon_i$ of .115 in Table (4) shows that there is still substantial variation in repayment. As $u_i + \epsilon_i$ increases, more of the variation in repayment cannot be explained by hard information or credit scoring. As $\sigma_u^2$ increases relative to $\sigma_{\epsilon}^2$, the potential either very small or sometimes positive magnitudes. Dobbie and Skiba (2013) find that increases in loan size lead to a slight decrease in default when examining US payday lenders. However, their estimates cannot reject no change.
value of loan officers increases as more of the unexplained variation can be screened. However, the relative contribution of soft information or exogenous error is not identified.

Some of the additional covariates are also of interest. Managers and women all repay at above average rates. Dorm or rental indicators in China generally imply young, migrant, or transient workers, and it is not surprising that their repayment rates are correspondingly lower than average. Some of the home inspection items such as air conditioner units or conversations with family and coworkers also indicate that household wealth and openness are associated with higher rates of repayment.  

5.2 Loan Officer Estimates

Table (5) presents the estimates of the adjusted variables $\hat{r}_j$, $\hat{c}_j^2$, and $\hat{\sigma}_j^2$. Interpreting these estimates is difficult as they are functions of $\beta$ and $\sigma_u^2$. However, some conclusions can be drawn. The estimated values for $\hat{r}_j = r_j \beta$ are all positive indicating that all of the loan officers are risk averse. This means that the loan amounts are smaller than the profit-maximizing level for all of the loan officers. Another conclusion is that there is substantial heterogeneity across loan officers in risk preferences. Table (6) displays a likelihood ratio test for a restricted model of constant risk aversion. The data is sufficient to reject this null.

The second and third columns in Table (5) shows the estimate of overconfidence and screening effort for all of the loan officers. The change of variables still allows comparisons to be made between these loan officers. Adjusted overconfidence $\hat{c}_j^2 = \frac{c_j^2}{\sigma_u^4} + \frac{1}{\sigma_u^2} - 1$ is uniformly less than adjusted screening effort $\hat{\sigma}_j^2 = \frac{\sigma_j^2}{\sigma_u^4} + \frac{1}{\sigma_u^2} - 1$. This indicates that all of the loan officers are overconfident although this result is imprecise. It is informative to characterize how an underconfident loan officer would behave. Holding screening effort $\sigma_j^2$ constant, as $c_j^2$ increases, the variance and mean of loan amounts will decrease, but the variance of repayment will not change. Underconfident loan officers therefore have smaller but worse performing loans.

Table (6) shows specification tests for constant levels of overconfidence and for constant screening effort. Both of these nulls are rejected at the 10% level. Another specification test is the null that loan officers have accurate perceptions about their screening effort or that $\hat{c}_j^2 = \sigma_j^2$ for all of the loan officers. The likelihood ratio test rejects this hypothesis convincingly. Together, these results provide evidence that there is heterogeneity in risk aversion, overconfidence, screening effort, and that loan officers do not accurately perceive their screening effort. While the point

---

27 I urge caution when interpreting the coefficients on hard information since many of the included variables are higher order terms, interactions, or functions of existing variables. For instance, the lender's credit score variable credit quality is itself a function of salaried income and characteristics. The purpose of this exercise is to mimic the lender's credit scoring model and not to assign causal interpretations to borrower characteristics apart from loan amount.

28 There are no constraints placed on $\hat{r}_j$ during estimation. Having a negative $\hat{r}_j$ may indicate risk loving behavior by some loan officers. Agarwal and Ben-David (2014) find that some loan officers take excessive risks given that their screening effort conveys little information. These excessive risks may be attributed to risk aversion if overconfidence is not explicitly modeled.
Table 5: Estimates of the Loan Officer Parameters

<table>
<thead>
<tr>
<th>Officer #</th>
<th>Estimate of Risk Aversion</th>
<th>Estimate of Overconfidence</th>
<th>Estimate of Screening Effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>0.004</td>
<td>124.9 ***</td>
<td>129.7 *</td>
</tr>
<tr>
<td>#2</td>
<td>0.011 ***</td>
<td>476.5</td>
<td>1917.0</td>
</tr>
<tr>
<td>#3</td>
<td>0.018 ***</td>
<td>96.4 ***</td>
<td>107.8 **</td>
</tr>
<tr>
<td>#4</td>
<td>0.010 ***</td>
<td>262.4 ***</td>
<td>753.2 *</td>
</tr>
<tr>
<td>#5</td>
<td>0.001</td>
<td>151.6 ***</td>
<td>181.1 **</td>
</tr>
<tr>
<td>#6</td>
<td>0.009 ***</td>
<td>294.2 **</td>
<td>968.2</td>
</tr>
<tr>
<td>#7</td>
<td>0.005 **</td>
<td>440.0 ***</td>
<td>1767.0</td>
</tr>
<tr>
<td>#8</td>
<td>0.008 ***</td>
<td>186.5 ***</td>
<td>324.1</td>
</tr>
<tr>
<td>#9</td>
<td>0.006 **</td>
<td>199.7 ***</td>
<td>320.5 *</td>
</tr>
<tr>
<td>#10</td>
<td>0.010 ***</td>
<td>243.3 ***</td>
<td>550.2 **</td>
</tr>
<tr>
<td>#11</td>
<td>0.010 ***</td>
<td>267.8 ***</td>
<td>620.4 **</td>
</tr>
<tr>
<td>#12</td>
<td>0.005 *</td>
<td>289.5 **</td>
<td>744.0</td>
</tr>
<tr>
<td>#13</td>
<td>0.010 ***</td>
<td>78.9 ***</td>
<td>87.7 *</td>
</tr>
<tr>
<td>#14</td>
<td>0.008 ***</td>
<td>210.8 ***</td>
<td>434.1 *</td>
</tr>
<tr>
<td>#15</td>
<td>0.001</td>
<td>223.2 ***</td>
<td>444.8</td>
</tr>
<tr>
<td>#16</td>
<td>0.015 ***</td>
<td>196.9 **</td>
<td>272.2</td>
</tr>
<tr>
<td>#17</td>
<td>0.006 **</td>
<td>268.0 *</td>
<td>555.1</td>
</tr>
<tr>
<td>#18</td>
<td>0.010 ***</td>
<td>297.1 ***</td>
<td>1026.0 **</td>
</tr>
<tr>
<td>#19</td>
<td>0.007 ***</td>
<td>462.7</td>
<td>1708.0</td>
</tr>
<tr>
<td>#20</td>
<td>0.007 ***</td>
<td>288.7 ***</td>
<td>583.6 **</td>
</tr>
<tr>
<td>#21</td>
<td>0.011 ***</td>
<td>200.7 ***</td>
<td>389.6</td>
</tr>
</tbody>
</table>

Notes: Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Estimates are from the main MLE specification and are the adjusted variables in terms of $\beta$ and $\sigma_u^2$.

*** Significant at the 1 percent level.
**  Significant at the 5 percent level.
*   Significant at the 10 percent level.

estimates convey little meaning in isolation, the value of these parameters is in constructing counterfactual lending models.

5.3 Counterfactual Lending - No Delegation

The structural model allows for counterfactual exercises examining lending under different scenarios. One useful exercise is to examine lending without delegating to loan officers. How do outcomes change for a risk neutral lender if it could only use credit scoring and no soft information?\(^{29}\) Given equation (6), the profit-maximizing loan amount for the lender conditional only on hard information is given by

\(^{29}\)Heider and Inderst (2012) examine a model where competition between lenders causes lenders to disregard soft information and rely solely on observable characteristics. The role of the loan officer changes to a salesperson prospecting for borrowers rather than as evaluators. Berger et al. (2005) find that the costs of collecting soft information rise with the size of the lender. Past a certain volume, lenders may not find it worthwhile to continue to collect soft information.
Table 6: Likelihood Ratio Specification Tests

<table>
<thead>
<tr>
<th>Null:</th>
<th>Likelihood Ratio Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan officers have constant risk aversion $r_j = r$</td>
<td>$\chi^2 = 471$, $P$-Value $= 0.000$</td>
</tr>
<tr>
<td>Loan officers have constant overconfidence $c_j^2 = c^2$</td>
<td>$\chi^2 = 31$, $P$-Value $= 0.051$</td>
</tr>
<tr>
<td>Loan officers have constant screening effort $\sigma_j^2 = \sigma^2$</td>
<td>$\chi^2 = 32$, $P$-Value $= 0.042$</td>
</tr>
<tr>
<td>Loan officers are not overconfident $c_j^2 = \sigma_j^2$</td>
<td>$\chi^2 = 446$, $P$-Value $= 0.000$</td>
</tr>
</tbody>
</table>

Notes: Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Likelihood ratio test statistics are on the parameter estimates from the main MLE specification.

\[
L_i^{\text{Lender}} = (-2\gamma)^{-1} \left( x_i' \Gamma - \frac{1}{R} \right) \tag{13}
\]

Contrast this with the loan officer’s loan amount in equation (10). The lender effect $(-2\gamma)^{-1}$ is much larger than the officer effect $k_j$ indicating that a risk neutral lender would approve higher average loan sizes. Furthermore, without access to soft information, the lender’s posterior belief of $u_i + \epsilon_i$ is equal to 0 instead of $\frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} \omega_{ij}$.

For each loan, expected profit can be calculated as

\[
E[\pi_i^{\text{Lender}}] = RL_i^{\text{Lender}} \left( \gamma L_i^{\text{Lender}} + x_i' \Gamma \right) - L_i^{\text{Lender}} \tag{14}
\]

Without access to soft information, the average loan amount is 38,000 RMB, which is 14% higher than the average loan amount of 33,450 RMB in the status quo. Also, the variance of loan sizes is 50% smaller than the status quo because there is no longer any variation within observable groups. The hard information only case also has average loan profits of 11,000 RMB. These baseline results can be compared to the status quo and different counterfactuals that eliminate the distortions caused by the loan officer’s preferences. For example, how would loan amounts and profits be affected if loan officers behaved with less risk aversion, exerted greater effort, or properly updated their beliefs?

5.4 Counterfactual Lending - Status Quo

The status quo model uses the estimated parameters and matches the observed loan amounts. Using the estimated loan officer characteristics, I first solve for the signal that the loan officer must have observed to choose $L_{ij}$ for that particular borrower. By knowing loan officer $j$’s parameters, borrower $i$’s hard information, and the approved loan amount, I can invert for the soft information signal. This signal relates both the borrower’s belief about expected profit as well as the loan’s actual expected profit.
More specifically, for each loan officer, the officer effect \( k_j \) is calculated using the adjusted parameters where 
\[
\tilde{r}_j = \frac{\tilde{r}_j}{1 + \tilde{c}_j} + \sigma_\epsilon^2 - 2\gamma
\]
With the officer effect and equation (10), the original signal is given by
\[
\frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} = L_i^s/k_j^{S,Q} - x'_i \Gamma + \frac{1}{R}
\]  
(15)

Because loan officers are overconfident, the correct posterior mean of \( u_i + \epsilon_i \) is not given by equation (15) but is instead \( \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} \). Using the adjusted parameters, the correct posterior mean can be expressed as
\[
\frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} = \frac{\sigma_u^2 + \sigma_j^2}{\sigma_u^2 + \sigma_j^2} \times \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij}
\]
(16)

which gives the expected profit conditional on the signal to be
\[
E \left[ \pi_i^{S,Q} \right] = RL_i^{S,Q} \left( \gamma L_i^{S,Q} + x'_i \Gamma + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} \right) - L_i^{S,Q}
\]  
(17)

Table (7) shows that the status quo with risk averse loan officers has smaller loan amounts than the hard information only case with a risk neutral lender. However, average loan profits are about 200 RMB higher and significant at the 10% level due to a more profitable allocation to borrowers. With an average compensation of 45,000 RMB and roughly 700 screened borrowers per year, the average loan officer contributes 140,000 RMB in additional profits over relying solely on hard information. Roughly speaking, the average loan officer is contributing three times his pay. However, this value varies across loan officers. Figure (7) shows this additional profit for all of the loan officers individually. While the individual estimates are imprecise, the majority of the point estimates are profitable.

5.5 Counterfactual Lending - Minimum Estimated Risk Aversion

One counterfactual is to examine lending if loan officers behaved with less risk aversion. Specifically, I can construct a lending model where all of the loan officers behaved with risk aversion equal to the least risk averse loan officer. The procedure first requires calculating an updated officer effect \( \tilde{r}_j \) and then re-calculating the loan amount \( L_{ij}^{MinRA} \). Since the adjusted risk aversion estimate \( \tilde{r}_j = r_j \beta \) is just a scalar multiple of the bonus rate \( \beta \), the loan officer with the lowest value of \( r_j \) also has the lowest value of \( \tilde{r}_j \). Each officer effect is given by
\[
k_j^{MinRA} = \left[ r_{Min} \beta \left( \frac{\epsilon_j^2}{1 + \epsilon_j^2} + \sigma_\epsilon^2 \right) - 2\gamma \right]^{-1}
\]
. After solving for the original signal in equation (15), the updated loan amount and expected loan profits are given similarly to equations (10) and (17).
Table 7: Counterfactuals and Comparison to Hard Information Only

<table>
<thead>
<tr>
<th>Counterfactual</th>
<th>Additional Loan Amount (000's RMB)</th>
<th>Additional Loan Profit (000's RMB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status quo with estimated parameters.</td>
<td>-4.56 *** (1.63)</td>
<td>0.21 * (0.11)</td>
</tr>
<tr>
<td>Loan officers with $r_j = r_{\text{min}}$.</td>
<td>-0.57 (1.00)</td>
<td>0.32 *** (0.04)</td>
</tr>
<tr>
<td>Loan officers with $c_j^2 = \sigma_j^2$.</td>
<td>-4.62 *** (1.64)</td>
<td>0.35 *** (0.12)</td>
</tr>
<tr>
<td>Loan officers with $\sigma_j^2 = \sigma_{\text{min}}^2$.</td>
<td>-4.54 *** (1.63)</td>
<td>0.64 *** (0.13)</td>
</tr>
</tbody>
</table>

Notes: Bootstrapped standard errors in parenthesis. Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Hard information only is risk neutral lending with no loan officers and no soft information. Average loan amounts equal 38,000 RMB and average profits are 11,000 RMB. For each counterfactual, each borrower’s loan amount is re-optimized and expected profit is calculated. As of August 2014, $1 is RMB 6.18.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

This would increase average loan sizes and profits since the original loan was approved with a higher risk aversion parameter. However, this effect is not ambiguously beneficial to the lender and it could be possible that some loans actually see a decrease in expected profit if loan officers behaved with lower risk aversion. To see this, suppose an extremely overconfident loan officer received a signal about the borrower’s soft information. Overconfidence leads him to believe the borrower’s quality is higher than reality, but this tendency is counteracted by risk aversion which decreases average loan amounts. By behaving with less risk aversion, the overconfidence effect may dominate the risk aversion effect. The result may be that some loan amounts increase past the point of profitability. The effectiveness of the counterfactual is then largely an empirical question.

Table (7) reports that the average loan size of 37,440 RMB is significantly higher than the status quo, and statistically indistinguishable from the hard information only case. Profits are 320 RMB higher for every loan. With 700 applications a year, moving to risk neutrality brings an additional profit of 224,000 RMB over the hard information only case. This is about five times the average loan officer’s compensation. Figure (8) shows the profit comparison for individual loan officers. The majority now statistically out-perform the hard information only case with some achieving additional profits of over 1,000 RMB per loan. These superstars now contribute
Figure 7: Additional Profit over Hard Information Only
Lending profit under the status quo.

Notes: 95% confidence interval constructed from bootstrapped standard errors. Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Status quo lending is with the estimated parameters. Hard information only is risk neutral lending with no loan officers and no soft information. Each borrower’s loan amount is re-optimized and expected profit is calculated. As of August 2014, $1 is RMB 6.18.

an additional 15 to 16 times their compensation to the lender.

5.6 Counterfactual Lending - No Overconfidence

Another counterfactual is to look at changes in loan sizes and loan profits if loan officers behaved without overconfidence. Instead of the estimated \( \sigma^2_j \), a lending model can be constructed with the loan officer’s true \( \sigma^2_j \). The procedure for determining counterfactual loan profits follows the counterfactual on risk aversion, which requires calculating an updated officer effect \( k_j \). Because the adjusted overconfidence \( \tilde{c}_j = \frac{\tilde{c}_j}{\sigma^2_u} + \frac{1}{\sigma^2_u} - 1 \) has the same functional form as the adjusted screening effort \( \tilde{\sigma}_j = \frac{\sigma^2_j}{\sigma^2_u} + \frac{1}{\sigma^2_u} - 1 \), a loan officer with accurate perceptions has \( \tilde{c}_j = \tilde{\sigma}_j \). The officer effect is given by

\[
k_j^{\text{NoOC}} = \left[ \tilde{r}_j R \left( \frac{\tilde{c}_j}{1+\sigma^2_j} + \tilde{\sigma}^2_j \right) - 2\gamma \right]^{-1},
\]

and the updated loan amount is given by
Figure 8: Additional Profit over Hard Information Only
Lending profit if loan officers behaved with \( r_j = r_{\text{min}} \).

Notes: 95% confidence interval constructed from bootstrapped standard errors. Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Counterfactual lending is if all loan officers behaved with the minimum estimated risk aversion. Hard information only is risk neutral lending with no loan officers and no soft information. Each borrower’s loan amount is re-optimized and expected profit is calculated. As of August 2014, $1 is RMB 6.18.

\[
L_{ij}^{\text{NoOC}} = k_j^{\text{NoOC}} \left( x_i' \Gamma - \frac{1}{R} + \frac{\sigma^2_u}{\sigma^2_u + \sigma^2_j} \omega_{ij} \right) \tag{18}
\]

where the adjusted signal can be calculated from equation (16) and expected profit is similar to equation (14).

Without overconfidence, average loan sizes must decrease but so might profits due to two effects. The first is because an overconfident loan officer has a low value of \( c^2_j \) compared to \( \sigma^2_j \) indicating a larger officer effect \( k_j \). Overconfidence alleviates the effect of risk aversion so that by reducing overconfidence, loan amounts and profits may decrease further.\(^{30}\) The second effect is because overconfidence also changes the posterior belief \( \frac{\sigma^2_u}{\sigma^2_u + \sigma^2_j} \omega_{ij} \). Risky borrowers with

\(^{30}\)Even with a precise signal, it may be possible that some loan officers are not willing to increase loans to the point where marginal profit equals zero. This fear may be driven by herd behavior or career concerns where a failing loan given to an observably risky borrower may be blamed on the loan officer. See Borenstein et al. (2012) for a career concerns model where the agent willfully deviates from the profit-maximizing outcome to avoid possible blame.
\( \omega_{ij} < 0 \) are given loans too small, and safer borrowers with \( \omega_{ij} > 0 \) are given loans excessively large. This effect also interacts with risk aversion so that with accurate beliefs about screening effort, risky borrowers are given larger loans and safer borrowers are given smaller loans. The lender may not even be willing to reduce overconfidence in such scenarios if it exacerbated risk aversion and lowered profits. Leaving the loan officers overconfident may be effective at offsetting the distortion caused by risk aversion. Undoubtedly, a more effective solution is to attenuate risk aversion and overconfidence jointly. While the theoretical prediction of counteracting overconfidence is ambiguous, the effects can still be measured empirically.

![Figure 9: Additional Profit over Hard Information Only](image)

Lending profit if loan officers behaved with \( c_j^2 = \sigma_j^2 \).

**Notes:** 95% confidence interval constructed from bootstrapped standard errors. Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Counterfactual lending is if all loan officers behaved with no overconfidence. Hard information only is risk neutral lending with no loan officers and no soft information. Each borrower’s loan amount is re-optimized and expected profit is calculated. As of August 2014, $1 is RMB 6.18.

Table (7) shows that the impact on average loan amounts is virtually unchanged, but profits are 350 RMB higher on average. The average loan officer is able to contribute 5.5 times his compensation in additional profits over the hard information only case. Figure (9) shows this for each loan officer separately indicating that with correct beliefs, most loan officers are statistically profitable for the lender. However, the point estimates for some of the loan officers decrease due
to less overconfidence. For example, loan officer 3 sees his profits decrease by more than 30 RMB per loan as his loan amounts decrease further away from the profit-maximizing level. Note that loan officer 3 also sees a large increase in profits in the counterfactual addressing risk aversion indicating that for particularly high levels of risk aversion, some amount of overconfidence may be profitable.

5.7 Counterfactual Lending - Highest Estimated Screening Effort

One last counterfactual compares loan amounts and profits when loan officers exhibit greater screening effort. If all loan officers behaved with the highest estimated screening effort, then the signal would be more precise and lead to a more efficient allocation of loan amounts to borrowers. The procedure begins by recognizing that the officer effect $k_j$ does not change from the status quo since $k_j$ is not directly a function of screening effort. Also, the loan officer with the lowest value of $\sigma^2_j$ must also have the lowest adjusted $\tilde{\sigma}^2_j$. Since the realization of $\omega_{ij}$ would change with greater screening effort, a different approach must be taken in calculating the adjusted signal. The adjusted signal in the status quo is distributed $\tilde{\sigma}^2_u \tilde{\sigma}^2_j \omega_{ij} \sim N(0, \frac{\tilde{\sigma}^4_u}{\tilde{\sigma}^2_u + \tilde{\sigma}^2_j})$, while a signal coming from a loan officer with variance $\sigma^2_{Min}$ would be distributed

$$\frac{\sigma^2_u}{\sigma^2_u + \sigma^2_{Min}} \omega_{ij}^{High\ Effort} \sim N\left(0, \frac{\sigma^4_u}{\sigma^2_u + \sigma^2_{Min}}\right)$$ (19)

To maintain consistent comparisons with the status quo lending, I rescale the adjusted signal so that the two expressions are distributed similarly.

$$\frac{\sigma^2_u}{\sigma^2_u + \sigma^2_{Min}} \omega_{ij}^{High\ Effort} \sim \left[\frac{\sigma^2_u + \sigma^2_j}{\sigma^2_u + \sigma^2_{Min}}\right] \times \frac{\sigma^2_u}{\sigma^2_u + \sigma^2_{Min}} \omega_{ij}^{S.Q}$$ (20)

Eqn 16

This normalizes the value of the counterfactual signal $\omega_{ij}^{High\ Effort}$ without having to draw a stochastic value for $\omega_{ij}^{High\ Effort}$. The procedure is not yet complete since the loan officer actually observes a different signal because of overconfidence. This value can be calculated as

$$\frac{\sigma^2_u}{\sigma^2_u + \sigma^2_j} \omega_{ij}^{High\ Effort} = \left[\frac{\sigma^2_u + \sigma^2_{Min}}{\sigma^2_u + \sigma^2_j}\right] \times \frac{\sigma^2_u}{\sigma^2_u + \sigma^2_{Min}} \omega_{ij}^{High\ Effort}$$ (21)

Eqn 20

The loan officer’s new loan amount and expected profit is given by

31 The benefit to this approach instead of simply drawing from the distribution in equation (19) is that this allows for a consistent comparison with the status quo. As $\sigma^2_{Min}$ approaches $\sigma^2_j$, the counterfactual values equal the status quo values.
\[
E \left[ \pi_i^{High\,Effort} \right] = L_i^{High\,Effort} = k_j^{S,Q} \left( x_i^{\Gamma} - \frac{1}{R} \frac{\sigma^2}{\sigma^2 + c_j^{\omega_i^{High\,Effort}}} \right) - L_i^{High\,Effort}
\]

Figure 10: Additional Profit over Hard Information Only
Lending profit if loan officers behaved with \( \sigma^2_j = \sigma^2_{min} \).

Notes: 95% confidence interval constructed from bootstrapped standard errors. Data includes 31,954 first-time borrowers from December 2011 to January 2014 across 46 branches examined by 21 loan officers. Borrowers apply for a loan amount with a set APR and payment length. Approved loan amounts are decided by the loan officers. Counterfactual lending is if all loan officers behaved with the highest estimated screening effort. Hard information only is risk neutral lending with no loan officers and no soft information. Each borrower’s loan amount is re-optimized and expected profit is calculated. As of August 2014, $1 is RMB 6.18.

This change should not affect average loan sizes since a more accurate signal leads to larger loans for safer borrowers and smaller loans for riskier borrowers. Average profits increase because a more efficient allocation leads to greater profits. Table (7) finds that average profits are 640 RMB higher in the counterfactual case indicating that the average loan officer would contribute 10 times his annual compensation in additional profits. Figure (10) shows this effect for each loan officer and finds that almost all of the loan officers are above the Mendoza line.\(^{32}\)

\(^{32}\)An expression in baseball that gives the batting average where players below this threshold cannot contribute positively to the team regardless of other strengths. The loan officer would have to contribute about 65 RMB per loan above the hard information only case to equal his average 45,000 RMB compensation.
5.8 Implementation

The practical implementation of these counterfactuals may be difficult. Because these depend on knowing the parameters of the individual loan officers, there may be a ratchet effect as loan officers vary their behavior anticipating that the lender is learning about them (Gibbons 1987). In addition, the composition of the loan officers may change in response to these attempts to alter their behavior (Lazear 2000). However, with that said, some of the counterfactuals may be possible to implement with post-origination scaling. For example, one way to mitigate risk aversion is to increase each loan by a set percentage based on the loan officer’s officer effect $k_j$. By scaling up each loan, the lender can possibly eliminate the cost associated with risk aversion. However, this requires that the loan officer to truthfully report an unaltered loan amount that the lender subsequently adjusts upwards. One possible solution could be for the lender to compensate the loan officer according to the initial loan amount rather than the adjusted loan amount which may mean calculating the borrower’s counterfactual repayment.

For overconfidence, if simply informing the loan officers of their behavioral bias is insufficient, then it may also be possible to use ex-post loan adjustments. Overconfidence leads loan officers to exaggerate the precision of their signal, which means that safer borrowers are given loans too large, and risky borrowers are given loans that are too small. The lender could intervene by compressing loan amounts conditional on the borrower’s hard information so that larger loans are reduced and smaller loans are increased. As long as the loan officer still has the proper incentives to truthfully report the initial loan amount, then it could be possible to eliminate the cost of overconfidence. However, care must be taken so that the design can account for offsetting interactions. For example, mitigating overconfidence may actually decrease loan profits through its interaction with risk aversion.

However, increasing screening effort may be much more difficult and require structural changes in hiring, training, and compensation. Equation (8) shows that the effect of the bonus rate $\beta$ on the optimal screening effort to be non-linear and non-monotonic. Increasing the bonus rate may also exacerbate the costs of risk aversion since greater profit sharing may encourage fewer risks. In addition, the costs of inducing additional effort may not be profitable for the lender. Another possibility is to potentially increase the amount of hard information and reduce soft information by codifying more of the collected data as in Berg et al. (2013). This could reduce the amount of information that loan officers need to examine when screening, which may lead to more precise signals.

---

33 Ackerberg and Botticini (2002) find that matching between principals and agents may be important in determining the optimal structure of contracts.
6 Conclusion

This paper has examined the value of delegating to loan officers exhibiting three sources of heterogeneity: risk aversion, willingness to exert effort, and overconfidence. To weigh the costs and benefits of delegation, I developed a structural model where loans officers were delegated two choices. The first was a costly screening effort decision used to analyze soft information unobserved by credit scoring, and the second was a loan size choice that incorporated both soft and hard information. The model featured multiple dimensions of heterogeneity, modeled the borrower’s repayment, and accounted for potential endogeneity with the random assignment of borrower applications to loan officers. I found three main results.

First, there was substantial heterogeneity in risk aversion, overconfidence, and cost of effort across loan officers. This led to large differences in average loan amounts and resulting profits. Second, delegating to loan officers was more profitable than relying solely on hard information. Estimates showed that the average loan officer with an annual compensation of 45,000 RMB contributed 147,000 RMB more in annual profits compared to the automated alternative. This effect was heterogeneous with some loan officers unable to contribute additional value, while others were extremely valuable. Lastly, if the lender could have somehow eliminated the costs due to preferences, then average annual profits would have increased by 224,000 to 448,000 RMB compared to the automated alternative.

I view these results as highlighting the value of experts even when decisions can be based on a large amount of hard information. Subjective evaluation of soft information may be even more important in settings where hard information is difficult or costly to obtain. This is not to say that delegating to experts is always valuable. For example, Gruber (1996) finds that passive investment management generally outperforms active investment in mutual funds. In addition, other lenders may find no use for loan officers since collecting soft information such as appearance may be prohibited. With that said, in spite of the costs of delegation, these loan officers are on average very profitable for the lender.

More broadly, an insight of this paper is that it is difficult to evaluate expertise without jointly modeling a number of factors both environmental and innate. Consideration must be given to the information collection process, the expert’s preferences, the compensation scheme, as well as the efficacy of the alternative model.\footnote{Another reason to hire experts could be to avoid directly specifying a rule that may skirt regulation. For example, some universities may use admissions counselors to implement quotas that would be controversial if programmed explicitly in an automated model.} Beyond lending, this work provides a valuable framework that is useful in evaluating experts working elsewhere in program eligibility or diagnostic assessment. I view the study of this type of decision-making as an important avenue for further study because by studying the possible distortions, interventions can be taken to improve outcomes.
References


Appendix

A Loan Officer’s Optimal Loan Size

Conditional on the hard information \( \mathbf{x}_i \), the observed signal \( \omega_{ij} \), and the loan officer’s structural characteristics, loan officers choose the loan size \( L_{ij}^* \) to maximize expected utility. Utility is given by \( u(y_j) = -e^{-r_jy_j} \) where \( y_j = \alpha + \beta \sum_i \pi_{ij} - cost_j \left( \sigma_j^2 \right) \). \( r_j \) is risk aversion, \( c_j^2 \) is the overconfidence, and \( \sigma_j^2 \) is the screening effort level. Each borrower’s repayment is independent and given by equation (6).

\[
E[L_{ij}] = \int -e^{-r_j \left( \alpha + \beta \sum_i \pi_{ij} + \beta \pi_{ij} - cost_j(\sigma_j^2) \right)} f(\pi_{ij}) d\pi \\
= -e^{-r_j \left( \alpha + \beta \sum_i \pi_{ij} + \beta RL_{ij} \left( \gamma L_{ij} + \mathbf{x}_j' \Gamma \right) - \beta L_{ij} - cost_j(\sigma_j^2) \right)} \int e^{-r_j \beta \left( u_i + e_i \right)} f(\pi_{ij}) d\pi
\]

If \( x \) is normally distributed, then \( \int e^{tx} f(x) dx = e^{\mu x + \frac{1}{2}t^2\sigma^2} \) by the moment generating function where the posterior mean and variance are given by equation (16).

\[
= -e^{-r_j \left( \alpha + \beta \sum_i \pi_{ij} + \beta RL_{ij} \left( \gamma L_{ij} + \mathbf{x}_j' \Gamma \right) - \beta L_{ij} - cost_j(\sigma_j^2) \right)} e^{-r_j \beta RL_{ij} \frac{u_j}{\sigma_u + c_j} \omega_{ij} + \frac{1}{2} \left( r_j \beta RL_{ij} \right)^2 \left( \frac{\sigma_u^2 c_j^2}{\sigma_u^2 + c_j^2} + \sigma^2 \right)} \\
= -e^{-r_j \left( \alpha + \beta \sum_i \pi_{ij} + \beta RL_{ij} \left( \gamma L_{ij} + \mathbf{x}_j' \Gamma \right) - \beta L_{ij} - cost_j(\sigma_j^2) + \beta RL_{ij} \frac{u_j}{\sigma_u + c_j} \omega_{ij} - \frac{1}{2} \left( \beta RL_{ij} \right)^2 \left( \frac{\sigma_u^2 c_j^2}{\sigma_u^2 + c_j^2} + \sigma^2 \right) \right)}
\]

Maximizing the expression above is equivalent to maximizing the expression inside the inner parenthesis.

\[
0 = \gamma L_{ij} + \mathbf{x}_j' \Gamma + \gamma L_{ij} - \frac{1}{R} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} - r_j \beta RL_{ij} \left( \frac{\sigma_u^2 c_j^2}{\sigma_u^2 + c_j^2} + \sigma^2 \right)
\]

\[
L_{ij}^* = \left[ r_j \beta R \left( \frac{\sigma_u^2 c_j^2}{\sigma_u^2 + \sigma_j^2} + \sigma_j^2 \right) - 2\gamma \right]^{-1} \left( \mathbf{x}_j' \Gamma - \frac{1}{R} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} \right)
\]

B Loan Officer’s Optimal Screening Effort Level

Cost of effort is parametrized as \( cost_j \left( \sigma_j^2 \right) = \frac{d_j^2}{\sigma_j^2} \) and is a decreasing function of a loan officer specific cost of effort term \( d_j \). Once the effort level is chosen, a signal \( \omega_{ij} \) is observed, and the optimal loan amount is given by equation (7). Loan officers do not recognize that they are overconfident when choosing screening effort. For ease of exposition, I assume that there is a single borrower with hard information \( \mathbf{x}_i \). Since the errors are not correlated, the results
Solving backwards then, the loan officer’s indirect utility function from observing \( \omega_{ij} \) is given by

\[
v(L_{ij}(\omega_{ij})) = -e^{r_j \left( \alpha + \beta RL_{ij}^* (\gamma L_{ij}^* + x_i') - \beta L_{ij}^* \text{cost} \left( \sigma_j^2 \right) + \beta RL_{ij}^* \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \omega_{ij} - \frac{r_j}{4} \left( \beta RL_{ij}^* \right)^2 \left( \frac{\sigma_u^2 \sigma_j^2}{\sigma_u^2 + \sigma_j^2} + \sigma^2 \right) \right)}
\]

\[
E[U[\sigma_j^2]] = \int v(L_{ij}(\omega_{ij})) f(\omega_{ij} | \sigma_j^2) \, d\omega
\]

\[
\approx E \left[ v(L_{ij}(\bar{\omega})) + \frac{dv}{d\omega} (L_{ij}(\bar{\omega})) (\omega_{ij} - \bar{\omega}) \right]
\]

\[
= v(L_{ij}(\bar{\omega})) - e^{r_j \left( \alpha + \beta RL_{ij}^* (\gamma L_{ij}^* + x_i') - \beta L_{ij}^* \text{cost} \left( \sigma_j^2 \right) - \frac{r_j}{4} \left( \beta RL_{ij}^* \right)^2 \left( \frac{\sigma_u^2 \sigma_j^2}{\sigma_u^2 + \sigma_j^2} + \sigma^2 \right) \right)}
\]

Maximizing this expression is equivalent to maximizing the term inside the parenthesis.

\[
cost_j^2(\sigma_j^2) = \frac{\partial L_{ij}^*}{\partial \sigma_j^2} \left( 2 \beta R \gamma L_{ij}^* + \beta Rx_i' \Gamma - \beta - r_j (\beta R)^2 L_{ij}^* \left( \frac{\sigma_u^2 \sigma_j^2}{\sigma_u^2 + \sigma_j^2} + \sigma^2 \right) \right) - \frac{r_j}{2} (\beta RL_{ij}^* \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2})^2
\]

where the last step is because \( L_{ij}^* (\omega_{ij}) = k_j (x_i' \Gamma - \frac{1}{R}) \).

\[
\frac{d^2}{\sigma_j^2} = \frac{r_j}{2} (\beta RL_{ij}^*)^2 \left( \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2} \right)^2
\]

\[
\frac{d_j}{\sigma_j^2} = \left( \frac{r_j}{2} \right)^{\frac{1}{2}} \beta R \left[ r_j \beta R \left( \frac{\sigma_u^2 \sigma_j^2}{\sigma_u^2 + \sigma_j^2} + \sigma^2 \right) - 2\gamma \right]^{-1} \left( x_i' \Gamma - \frac{1}{R} \right) \frac{\sigma_u^2}{\sigma_u^2 + \sigma_j^2}
\]

\[
= \left( \frac{r_j}{2} \right)^{\frac{1}{2}} \beta R \left( x_i' \Gamma - \frac{1}{R} \right) \sigma_u^2 \left[ r_j \beta R \sigma_u^2 \sigma_j^2 + (r_j \beta R \sigma_u^2 - 2\gamma) \left( \sigma_u^2 + \sigma_j^2 \right) \right]^{-1}
\]

Cross multiplying gives
\[
\frac{d_j}{(\frac{r_j}{2})^{\frac{1}{2}} \beta R \left(x_i' \Gamma - \frac{1}{R} \right) \sigma_u^2} = \frac{\sigma_j^2}{r_j \beta R \sigma_u^2 \sigma_j^2 + (r_j \beta R \sigma_u^2 - 2 \gamma) \left( \sigma_u^2 + \sigma_j^2 \right)}
\]

\[
\frac{(r_j \beta R \sigma_u^2 - 2 \gamma) \sigma_u^2}{d_j} = \frac{1}{r_j \beta R (\sigma_u^2 + \sigma_j^2) - 2 \gamma + (r_j \beta R \sigma_u^2 - 2 \gamma) \frac{\sigma_j^2}{\sigma_u^2}}
\]

Simplifying gives the final optimal screening effort to be

\[
\sigma_j^{2*} = (r_j \beta R \sigma_u^2 - 2 \gamma) \sigma_u^2 \left( \frac{(r_j \beta R \sigma_u^2 - 2 \gamma) \sigma_u^2}{d_j} - r_j \beta R \left( \sigma_u^2 + \sigma_j^2 \right) + 2 \gamma \right)^{-1}
\]

### C Change of Variables

Since I cannot separately identify \(\sigma_u^2\) apart from \(\sigma_j^2\), I perform a change of variables and reduce the parameter set by one. I show that the likelihood can be written entirely using this reduced set showing that \(\sigma_u^2\) is not identified and that the reduced parameter set describes the same model.

Change of variables from \((\Gamma, \gamma, r_j, \sigma_j^2, c_j^2, \sigma_u^2, \sigma_\epsilon^2)\) to \((\Gamma, \gamma, \tilde{r}_j, \tilde{c}_j^2, \tilde{\sigma}_j^2, \tilde{\sigma}_\epsilon^2)\) where

\[
\begin{align*}
\tilde{r}_j &= r_j \beta \\
\tilde{c}_j^2 &= \frac{c_j^2}{\sigma_u^2} + \frac{1}{\sigma_u^2} - 1 \\
\tilde{\sigma}_j^2 &= \frac{\sigma_j^2}{\sigma_u^2} + \frac{1}{\sigma_u^2} - 1 \\
\tilde{\sigma}_\epsilon^2 &= \sigma_u^2 + \sigma_\epsilon^2 - 1
\end{align*}
\]

The likelihood in equation (12) can be written entirely as functions of \(k_j\), \(\sigma_v^2\), \(w_{ij}\), \(\sigma_{u+\epsilon|v}\), and \(h_{ij}\). I show that these functions can be written in terms of the adjusted parameters.

\[
k_j = \left[ r_j \beta R \left( \frac{\sigma_u^2 c_j^2 + \sigma_j^2}{\sigma_u^2 + c_j^2} \right) - 2 \gamma \right]^{-1} = \tilde{r}_j \beta R \left( \frac{\tilde{c}_j^2}{1+c_j^2} + \tilde{\sigma}_\epsilon^2 \right) - 2 \gamma = \tilde{k}_j
\]

\[
\sigma_v^2 = k_j^2 \frac{\sigma_v^2 \sigma_j^2}{\sigma_u^2 + c_j^2} = \tilde{k}_j^2 \frac{\tilde{c}_j^2}{1+c_j^2} = \tilde{\sigma}_v^2
\]

\[
w_{ij} = x_i' \Gamma k_j - \frac{k_j}{R} = x_i' \Gamma \tilde{k}_j - \frac{\tilde{k}_j}{R} = \tilde{w}_{ij}
\]

\[
\sigma_{u+\epsilon|v}^2 = \frac{\sigma_u^2 \sigma_j^2}{\sigma_u^2 + c_j^2} + \sigma_\epsilon^2 = \frac{\tilde{\sigma}_j^2}{1+c_j^2} + \tilde{\sigma}_\epsilon^2 = \tilde{\sigma}_{u+\epsilon|v}^2
\]
\[ h_{ij} = \gamma L_{ij} + \mathbf{a}_i \mathbf{r} + \frac{\sigma_u^2 + \sigma_j^2}{(\sigma_u^2 + \sigma_j^2) k_j} (L_{ij} - w_{ij}) = \gamma L_{ij} + \mathbf{a}_i \mathbf{r} + \frac{1 + \sigma_j^2}{(1 + \sigma_j^2) k_j} (L_{ij} - w_{ij}) = \tilde{h}_{ij}. \]

These functions can be inverted to solve for the true officer parameters so that

\[
\begin{align*}
  r_j &= \tilde{r}_j / \beta \\
  c_j^2 &= \sigma_u^4 \tilde{c}_j^2 + \sigma_u^4 - \sigma_u^2 \\
  \tilde{\sigma}_j^2 &= \sigma_j^2 \sigma_u^4 + \sigma_u^4 - \sigma_u^2 \\
  \sigma_{\epsilon}^2 &= 1 + \sigma_{\epsilon}^2 - \sigma_u^2
\end{align*}
\]