

1 Model Setup

We take model 4 and assume that the repayment probability is distributed according to a parametric family of functions. Specifically, we assume the Kumaraswamy distribution.

1.1 Kumaraswamy Distribution

The probability density function is:

$$f(h) = \alpha\beta h^{\alpha-1}(1-h^\alpha)^{\beta-1} \quad (1)$$

The cumulative distribution function is:

$$F(h) = 1 - (1-h^\alpha)^\beta \quad (2)$$

Note that for $\alpha = \beta = 1$, this reduces to the uniform distribution. The expected value is given by:

$$\mathbb{E}(\alpha, \beta) = \beta B\left(1 + \frac{1}{\alpha}, \beta\right) = \frac{\beta\Gamma(1 + \frac{1}{\alpha})\Gamma(\beta)}{\Gamma(1 + \frac{1}{\alpha} + \beta)} \quad (3)$$

where $B(\cdot)$ is the beta function and Γ is the gamma function.

1.2 Parameters to Identify

The elements we need to identify from our model are:

- λ : Switching cost
- α, β : Distribution of types

1.3 Observable Data

For each borrower $i = 1, \dots, N$, we observe:

- r_i^{win} : The interest rate charged by the winning bank (i.e., the bids conditional on winning)
- $W_i \in \{1, 2\}$: The bank from which the customer took the loan, where $W_i = 1$ indicates the incumbent won and $W_i = 2$ indicates the entrant won
- $D_i \in \{0, 1\}$: Whether the borrower defaulted ($D_i = 1$) or repaid ($D_i = 0$)

2 Estimation Strategy with Limited Observables

In practice, the researcher observes only a subset of the model's variables.

2.1 Unobservable Variables

The following are *not* observed by the researcher:

1. **Latent type** h_i : The borrower's true default probability
2. **Loser's bid**: The interest rate offered by the bank that lost (either r_1 if entrant won, or r_2 if incumbent won)
3. **Bank profits**: The realized profits are not directly observed

The parameter vector to be estimated is $\theta = (\alpha, \beta, \lambda, h_{min}, h_{max})$.

2.2 Likelihood Function

Given the model structure, we can construct the likelihood of observing the data. For borrower i :

Incumbent wins ($W_i = 1$): The incumbent wins when $r_1(h_i) \leq r_2 + \lambda$. Given the equilibrium strategy $\sigma(h; \theta)$ and mixed strategy $G(x; \theta)$, the probability that the incumbent wins with type h is:

$$\Pr(W_i = 1 \mid h_i; \theta) = 1 - G(\sigma(h_i; \theta) - \lambda; \theta) \quad (4)$$

The conditional density of the winning rate is:

$$f(r_i^{win} \mid W_i = 1, h_i; \theta) = \begin{cases} 1 & \text{if } r_i^{win} = \sigma(h_i; \theta) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

since the incumbent plays a pure strategy.

Entrant wins ($W_i = 2$): The entrant wins when $r_2 < r_1(h_i) - \lambda$. The probability is:

$$\Pr(W_i = 2 \mid h_i; \theta) = G(\sigma(h_i; \theta) - \lambda; \theta) \quad (6)$$

The conditional density of the winning rate comes from the mixed strategy:

$$f(r_i^{win} \mid W_i = 2, h_i; \theta) = g(r_i^{win}; \theta) \quad (7)$$

where $g(\cdot)$ is the density of $G(\cdot)$.

Default outcome: Conditional on type h_i :

$$\Pr(D_i = 1 \mid h_i) = h_i, \quad \Pr(D_i = 0 \mid h_i) = 1 - h_i \quad (8)$$

Complete likelihood: Since h_i is unobserved, we integrate it out:

$$\begin{aligned} \mathcal{L}_i(\theta) = & \int_{h_{min}}^{h_{max}} f(r_i^{win} \mid W_i, h; \theta) \cdot \Pr(W_i \mid h; \theta) \\ & \cdot \Pr(D_i \mid h) \cdot f(h; \alpha, \beta) dh \end{aligned} \quad (9)$$

The log-likelihood for the sample is:

$$\ell(\theta) = \sum_{i=1}^N \log \mathcal{L}_i(\theta) \quad (10)$$

2.3 Estimation Procedure

Maximum Likelihood Estimation:

1. For a given parameter vector θ :
 - Compute the equilibrium strategies $\sigma(h; \theta)$ and $G(x; \theta)$ numerically (following the algorithm in Section “Equilibrium computation” in `model4_adapting Engelbrecht et al.tex`)
 - For each observation (r_i^{win}, W_i, D_i) , evaluate the likelihood $\mathcal{L}_i(\theta)$ by numerical integration over h
2. Maximize the log-likelihood:

$$\hat{\theta} = \arg \max_{\theta} \ell(\theta) \tag{11}$$

using numerical optimization (e.g., Nelder-Mead, BFGS)

3. Compute standard errors from the inverse Hessian or via bootstrap

2.4 Identification

Key identification arguments:

- **Switching cost λ :** Identified from the gap between incumbent and entrant rates. In equilibrium, the incumbent charges $r_1(h) = \lambda + 1/\mu(h)$ while the entrant’s support starts at $x_{min} = 1/E[1 - H]$. The minimum observed incumbent rate minus the minimum observed entrant rate reveals λ .
- **Distribution parameters (α, β) :** Identified from:
 1. The distribution of winning rates conditional on winner identity
 2. The correlation between winning rates and default outcomes
 3. The switching probability (fraction of borrowers choosing the entrant)
- **Support bounds (h_{min}, h_{max}) :** The observed range of rates (conditional on winner) restricts the support of h via the equilibrium pricing function $\sigma(h)$.

3 Random thoughts

- Switching costs does not generate heterogeneity in the rates paid by old customers, whereas information asymmetries do. Maybe the variance of rates paid by old customers can tell us about the degree of learning by the bank.
- If we can determine the switching probability implied by the model, then we can compare it to the observed in the data and this can reveal something about the relative importance of switching costs vs information asymmetries. For example if there is no information asymmetries, then the equilibrium is in pure strategies and there is no switching. If there are information asymmetries there will be switching.
- Given that we observe the winning bids conditional on winning, we can try to obtain some moments from the model and compare them to the data.
- Maybe if one calculates the probability of switching conditional on ex-post repayment there are some moments that we can try to match. Note that the lemons are the ones that switch, hence switchers should have a lower repayment rate than non-switchers.
- Note that the home bank plays a pure strategy. For any h they play $\sigma(h)$ and whenever the home bank wins we observe $\sigma(h)$, and the home bank has a higher likelihood of winning when h is low. if we adjust for this winning probability we can obtain the distribution of $\sigma(h)$ and then given a switching cost obtain the distribution of h .
- To identify the switching costs we can use the difference between the lowest bid by the entrant and the lowest rate by the incumbent. The difference should be equal to the switching cost.