Gender Discrimination in the Consumer Credit Market in Chile: Experimental Evidence from a Correspondence Study with Real Borrowers. A Pre-analysis Plan

Ana María Montoya
University of Chile

Eric Parrado
IADB

Raimundo Undurraga
University of Chile

Alex Solís
Uppsala University

January 15, 2019

This is a pre-analysis plan for the research project "Gender Discrimination in the Consumer Credit Market in Chile: Experimental Evidence from a Correspondence Study with Real Borrowers". The objective is to pre-specify in a precise way the analysis to be run before examining the data. In doing so, we will follow the guidelines offered by Olken (2015). This will be available for public at https://www.socialscienceregistry.org under the corresponding code associated to this project.

Basic Design: We recruited a sample of 404 potential borrowers in Chile, and matched male and female profiles on demographics, incomes, employment status, and credit history information so that we get a balanced sample of testers across gender. We then asked each tester to email 4 randomly assigned loan requests (i.e., with a random loan amount between $1 and $9 millions CLP and a random loan term between 12 and 60 months) to 4 randomly assigned loan officers. Our unit of analysis are loan requests. We test whether response rates, approval rates, and credit conditions of approved loans varied across borrower's gender.

Randomization: We randomly design loan requests combining loan amount, loan term, and standardized-text. We then create 4 slots per tester and stratify the slots by region and tester's gender. Then, within each region, we randomly assign each loan request to each slot. We then stratify the sample of loan officers by region, bank, and loan officer's gender, and within each strata we randomly assign one tester-loan-request slot to one loan officer. The number of loan officers is lower than the number of loan requests, so each loan officer may receive more than one loan request. Each tester cannot match with more than one loan officer per bank.

1. **General considerations**: our prior is that there exists gender discrimination against female applicants, thus hypothesis testing procedure will only consider one-sided test in favor of male applicants. This is to the cost that we acknowledge the reader in advance that unexpected results pointing to discrimination effects against

males (the opposite effect) will just be interpreted as evidence of no discrimination against female applicants.

2. **Primary Outcome Variables**: external margin

   - If loan request was responded: dummy variable that equal to 1 if the submitted loan request was responded by the assigned loan officer, and 0 otherwise.
   - If loan request was approved: dummy variable that equal to 1 if the submitted loan request was approved by the assigned loan officer, and 0 otherwise.

3. **Co-primary Outcome Variables**: internal margin, i.e., conditional on that the loan is approved

   - Approved Amount: approved amount, measured in Thousands of CLP
   - If the Requested Loan Amount is lower than the Approved Loan Amount: dummy variable that equal to 1 if the requested loan amount is lower than the approved amount, and 0 otherwise.
   - Approved Term: approved term, measured in months.
   - If the Requested Loan Term is lower than the Approved Loan Term: dummy variable that equal to 1 if the requested loan term is lower than the approved loan term, and 0 otherwise.
   - Loan Payment: amount of loan payment, measured in thousands of CLP.
   - Interest Rate: interest rate offered for the approved loan.
   - CAE: Annual Equivalent Charge, a summarized measure of costs involved by law in every loan offer when this is approved, which includes loan amount, interest rate, loan term, operational costs, and credit insurances.

4. **Inclusion/Exclusion Rules**:

   - Our analysis will be divided in two parts. We first identify attriter and non-attriter loan requests. Attriter loan requests are loan requests that were assigned to the tester but not submitted to the assigned loan officer. Non-attriter loan requests are loan requests that were assigned to the tester and effectively submitted to the assigned loan officer. Our baseline analysis will consider the full sample (attriters and non-attriters loan requests). Our impact analysis will only consider non-attriter loan requests.

5. **Main Analysis**:

   We will test whether the primary and co-primary outcomes of interest vary across loan requests submitted by male testers and loan requests submitted by female testers. We will estimate the following linear regression model (when the DV is binary, we will implement robustness checks using a Logit specification):

   $$Y_{lijkt} = \alpha + \beta Gender_i + \theta Gender_j + \mu_k + \delta_l + \zeta_t + \eta X_i + \pi Z_j + \varepsilon_{ijklt} \quad (1)$$

where li will stand for loan request l submitted by tester i, j will index loan officers, k will index region-banks where loan officer works, and t will index the week in which the loan request is submitted.

- Y_{lijkt} is any of the outcomes under study, and β is the coefficient associated to a dummy variable that equals 1 if the loan request is submitted by a male tester and 0 otherwise. A negative β would indicate that loan officers discriminate against female borrowers as the probability of responding or approving a loan request submitted by a female is lower compared to the probability of responding or approving a loan request submitted by a male. Our hypothesis is that there is discrimination against females (not against males), thus we will report one-sided test of the null hypothesis $H_0 : \beta > 0$. We expect loan requests will be correlated at the region-bank level and thus standard errors are clustered by region-bank category.

- We will control for stratification variables in all the specifications, i.e., by $Gender_j$, which is a dummy equal to 1 if the loan officer is male and 0 otherwise, which controls for unobserved differences between male and female loan officers that may affect the decision about whether respond/accept a loan request; by μ_k, which is a vector of region-bank fixed effects that will capture the average unobservable differences across banks in different regions like loan policies, criterias for eligible borrowers, and standard procedures of loan processing imposed by the bank. Controlling for these two sets of variables is important since randomization was conducted within blocks conformed by them, and thus recovering causal effects on gender discrimination is conditional on adjusting by them. In our main specification, we will also control by δ_l, which is a vector of loan type fixed effects that will include loan amount and loan term fixed effects. Finally, we will also include ζ_t, which is a week fixed effect. These time fixed effects capture unobserved differences across weeks in which the loan requests were submitted, like contemporaneous economic shocks and variations in the credit policies adopted by banks over time. Finally, ε_{lijkt} is the error term.

- We expect tester and loan officer characteristics will be well balanced, thus we will include them in a second specification only to test the robustness of β coefficient as well as to increase the statistical power. X_i, the vector of pre–treatment characteristics of the tester will include age, if married, monthly wage, if self-employed, and if the individual is client of the bank where the assigned loan officer works; Z_j, the vector of pre–treatment characteristics of the loan officer will include age, if has higher education degree, and years of experience in the banking sector.

6. **Subgroup analysis**:

- **Heterogeneous effects across Loan amount**: We will test whether the differences across male and female loan requests in terms of response and approval rates vary across loan amount requested. Loan amounts go from 1 to 9 millions CLP. The median loan amount should be close to $5M, thus in order to gain

3

power we will group the loan amounts in two categories: 1-4 M (small loans) and 5-9 M (large loans). In order to estimate heterogeneous effects, we will add to the main specification a dummy that equals 1 if the requested loan amount is \$4 M or less and zero otherwise, as well as an interaction term of this dummy variable and the dummy by loan request's gender. Our hypothesis is that the larger is the requested loan amount, the larger will be the gender discrimination against female borrowers and so expect that the coefficient associated with the interaction term be positive. Hence, we will report one-sided test of the null hypothesis $H_0 : \gamma \leq 0$, where γ is the coefficient associated to the interactive variable.

- **Heterogeneous effects across Individual Monthly Wage**: We will test whether the differences across male and female loan requests in terms of response and approval rates vary across the monthly wage of the individual. We will divide the sample in two groups: above and below the median wage. In order to estimate heterogeneous effects across this dimension, we will add to the main specification a dummy that is equal to 1 if the individual monthly wage is below the median and zero otherwise, as well as an interactive variable that combine this dummy with the dummy by loan request's gender. Our hypothesis is that the poorer is the individual, the larger will be the gender discrimination against female borrowers and so expect that the coefficient associated with the interaction term be negative. Hence, we will report one-sided test of the null hypothesis $H_0 : \gamma > 0$, where γ is the coefficient associated to the interactive variable.

- **Heterogeneous effects across Loan Officer's gender**: We will test whether the differences across male and female loan requests in terms of response and approval rates vary across the Loan Officer's gender. In order to estimate heterogeneous effects across this dimension, we will add to the main specification a dummy that is equal to 1 if the loan officer is male and zero otherwise, as well as an interactive variable that combine this dummy with the dummy by loan request's gender. Our hypothesis is that male officers are more likely to discriminate against female borrowers than female officers, thus we expect that the coefficient associated with the interaction term be negative. Hence, we will report one-sided test of the null hypothesis $H_0 : \gamma > 0$, where γ is the coefficient associated to the interactive variable.

- **Heterogeneous effects across Loan Officer's gender preference**: We will test whether the differences across male and female loan requests in terms of response and approval rates vary across the Loan Officer's gender preference. In order to elicit whether the loan officer manifest any particular gender preference for client's portfolio, we offer them 5 possible options: 20%-80% males-females; 40%-60% males-females; 50%-50% males-females; 60%-40% males-females; and 80%-20% males-females. We will estimate heterogeneous effects across this dimension by adding to the main specification a dummy that is equal to 1 if the loan officer prefer a portfolio of clients composed by a majority of males (60% male-40% females or 80% male-20% females) and zero otherwise, as well as an

interactive variable that combine this dummy with the dummy by loan request's gender. Our hypothesis is that pro-male officers are more likely to discriminate against female borrowers than not-pro-male officers, thus we expect that the coefficient associated with the interaction term be negative. Hence, we will report one-sided test of the null hypothesis $H_0 : \gamma > 0$, where γ is the coefficient associated to the interactive variable. Additionally, we will examine whether male pro-male loan officers discriminate more against female borrowers than female pro-male loan officers. Interactive variables will be added to the main specification accordingly. Our hypothesis is that male pro-male officers are more likely to discriminate against female borrowers than female pro-male officers, thus we expect that the coefficient associated with the triple interaction term be negative. Hence, we will report one-sided test of the null hypothesis $H_0 : \alpha > 0$, where α is the coefficient associated to the triple interactive variable.

- **Salience Experiment**: our experimental design includes a salience experiment attempting to test whether attitudes toward gender discrimination among loan officers can be modified through information treatments. In particular, a half of our sample of loan officers were randomly sent an anti-discrimination message. We will test whether the differences across male and female loan requests in terms of response and approval rates vary across whether the Loan Officer received the treatment message or not. In order to estimate heterogeneous effects across this dimension, we will add to the main specification a dummy that is equal to 1 if the loan officer received the treatment message and zero otherwise, as well as an interactive variable that combine this dummy with the dummy by loan request's gender. Our hypothesis is that treatment officers are less likely to discriminate against female borrowers than control officers, thus we expect that the coefficient associated with the interaction term be positive. Hence, we will report one-sided test of the null hypothesis $H_0 : \gamma \leq 0$, where γ is the coefficient associated to the interactive variable.

References

Olken, B., 2015. Promises and perils of pre-analysis plans. Journal of Economic Perspectives 29, 61–80.