

Technical Analysis Appendix

Technical Analysis

a. The Approach

The method we employ to measure racial discrimination in mortgage lending is called *the residual difference* approach. In a nutshell, this approach decomposes racial gaps in loan-rejection rates between the component that can be explained and that which cannot be explained by racial differences in characteristics. The unexplained gap is interpreted as being "discrimination." Alternatively, the residual difference method permits the estimation of minority loan-rejection rates when minorities are "treated" like equally qualified whites.

Because one of the qualifications that enters the lending decision is the probability that the loan will satisfy secondary market saleability requirements, we focus on producing a *proxy* for this probability measure using data on loans actually acquired by GSEs. The method used to assess discrimination in HMDA data also can be used to measure what first-time home-buying percentages among blacks would be had blacks been endowed with white characteristics. This intermediate step permits us to evaluate how much of the anticipation of GSE acquisition of minority loans -- as measured by our proxy -- derives from racial differences in borrower characteristics and how much derives from possibly different treatment by GSEs of minority and nonminority first-time homeowner loans.

b. The Variables

Let R^k be the loan-rejection rate of the k th group. For an individual loan applicant this value is equal to 1 if the borrower is denied a loan; it is equal to 0 otherwise. Across all members of the k th group this value can be summed and divided by the number of loan applicants in the k th group to yield the percentage of k -group applicants who are denied loans.

The loan rejection rate depends on the following variables:

- W = a vector of characteristics of the loan, including:
FHA, VA, vs. conventional loans; amount requested.
- X = a vector of characteristics of the applicant(s), including:
a proxy for bad credit, gender of applicant and coapplicant, race of applicant and coapplicant, income.
- Y = a vector of characteristics of the lender, including:
OCC, FRB, OTS, NCUA, etc; large or small lender.

Z = a vector of characteristics of the Census tract, including: median housing prices, family income, poverty rates, mobility rates, percentage minorities, age distribution, percentage owner-occupied units.

$P(NS)$ = the probability that the loan will not be sold on the secondary market.

$P(ft)$ = the probability that a loan purchased by a GSE is a first-time home-buyer loan.

Table 1-2 defines these variables in more detail.

c. The Model

Consider the following logistic regression equation for the probability that a black loan applicant is rejected:

$$R^b = \frac{1}{1 + \exp(\tilde{\alpha}^b + \alpha^b W + \beta^b X + \gamma^b Y + \delta^b Z + \eta^b P(NS)^b)} \quad (1)$$

where $\tilde{\alpha}$, α , β , γ , and δ are coefficients to be estimated. Similarly, we have the following for white applicants:

$$R^w = \frac{1}{1 + \exp(\tilde{\alpha}^w + \alpha^w W + \beta^w X + \gamma^w Y + \delta^w Z + \eta^w P(NS)^w)} \quad (2)$$

A measure of *discrimination* is the percentage of the rejection gap that cannot be explained by the W s, X s, Y s, Z s and $P(NS)$ s. Or,

$$D = \frac{\tilde{R}^b - R^b}{R^w - R^b}, \quad (3)$$

where R^b is the equal-treatment value of loan-denial rates for blacks. This rate is equal to the value of the rejection rates for blacks when blacks have the same coefficients (or treatment) as equally qualified whites, or

$$\tilde{R}^b = \frac{1}{1 + \exp(\tilde{\alpha}^w + \alpha^w W + \beta^w X + \gamma^w Y + \delta^w Z + \eta^w P(NS)^b)} \quad (4)$$

The problem, however, with this analysis is that we do not directly observe $P(NS)$, the probability that the loan is not sold to Fannie Mae or Freddie Mac or elsewhere on the secondary market. Focusing specifically on GSE purchases, we only know whether an *approved* loan is sold. A proxy for $P(NS)$ can be obtained from data on GSE purchases by using a derivation from Bayes' rule.

Note that what is observed is the conditional probability that a loan is sold to GSEs given that it was approved. We do not know the unconditional probability that a loan will be sold to a GSE from either the HMDA data--which does not include details of which loans were sent for review and which were declined by the secondary purchaser--or from the HUD-GSE data, which only includes approved loans. However, we know from Bayes' rule (Freund 1972, 62) that

$$P(S^*O) = \frac{P(S) @ P(O^*S)}{P(O)} , \quad (5)$$

where S means that the loan was sold and O means that the loan was originated. Note that all loans sold by the lender must have been originated so we can obtain a measure of the unconditional probability that a loan will not be sold from

$$P(NS) = 1 - P(S) = 1 - P(S^*O) @ P(O) . \quad (6)$$

A proxy for the lender's probability that the loan will not be sold can be estimated by the product of two separately obtained probabilities: a) the conditional probability of selling a loan on the secondary market; and b) the probability that the loan was originated. The second of these two probabilities is, unfortunately, tautologically related to the dependent variable, R . However, previous year's data from a given lender can be used to estimate for each racial group the likelihood of rejection controlling for location. Specifically, for the j th lender and the i th Census tract for racial group k , the estimated probability of origination in time t is the origination rate for the j th lender in the i th Census tract for racial group k in period $t-1$. In other words,

$$\hat{P}_{it}^{jk}(O) = P_{i\&1}^{jk}(O) . \quad (7)$$

This can be computed simply for all lenders that made loans in given Census tracts. To simplify matters further and to account for the continuing controversy over whether smaller lenders who make very few minority loans should be subject to CRA rules, we restrict the analysis to those lenders with at least 250 loans in years t and $t-1$ and at least one minority applicant in Census tract i in year t if there is at least one minority applicant in Census tract

t in year $t-1$. Previous analysis suggests that by restricting the sample to lenders with 250 or more loans, one eliminates or reduces the number of cases where the above estimate cannot be obtained.¹

The first of the two probabilities, the conditional probability that a loan is sold on the secondary market, most likely is a function of the decision apparatus of the buyers of these loans. Fortunately, the HUD-GSE data are well-suited for incorporating GSE decision making into the computation of lender outcomes. This is accomplished by a two-step method.

First we estimate the probability that the GSE purchases a first-time buyer loan. Table 4 lists our results. Denote this probability by $P(ft)$. Then we can estimate for each group k :

$$P(ft)^k = f(W', X', Y', Z') \quad (8)$$

where primes denote the fact that these independent variables are a proper subset of those listed in the rejection equation. This equation can be estimated using the HUD-GSE Census tract data to obtain coefficients that, once applied, permit the calculation of predicted values of $P(ft)$ for each loan that is sold. This predicted value along with other variables relevant to the lenders' and secondary-market buyers' decisions are then included to predict the probability that the loan is sold:

$$P^k(S^*O) = f(W'', X'', Y'', Z''); \hat{P}(ft) \quad (9)$$

where the double prime denotes that these independent variables are proper subsets of those listed in the rejection equation, but are different from those in the first-time homeowner sale equation (Table 6 lists these results).

Armed with a measure of the conditional probability that a loan is sold on the secondary market and an estimate of the probability that a loan is sold, we can then derive a proxy for

¹Of course, this estimation forces a year-to-year equality in origination rates in particular Census tracts for particular racial groups. An alternative is to estimate a growth parameter \tilde{n} from data on all Census tracts for all years from 1992-96 to obtain

$$\hat{P}_{it}^{jk}(O) = \tilde{n}^k P_{it\&1}^{jk}(O) .$$

Knowing the values for \tilde{n} for each racial group would be important and valuable in its own right for policymaking. But, it also provides a useful adjustment to the strong assumption that there is no growth in origination rates for specific racial groups.

the probability that a loan will not be sold on the secondary market and thus is rejected by the lender. This probability of rejection is estimated separately for minorities and nonminorities, and the residual difference analysis proceeds as detailed previously.

The novel aspect of this investigation is the preliminary estimation of the probability that a GSE-purchased loan is a first-time home-buyer loan. The logic of including the predicted value of this variable in the loan-sale equation is that loans that can meet the stringent requirements of GSEs often are less costly and move more quickly.² Some analysts claim that GSEs, and particularly Freddie Mac, fail to pursue minority first-time home-buyer loans in the same proportions as the industry at large (Lind, February 1996, April 1996).

To test the hypothesis that the problem is not discrimination by lenders but discriminatory behavior by GSEs as secondary market underwriters, we can predict from equations (8) and (9) *equal-treatment* probabilities of nonsale of loans in equation (6), holding constant the predicted probability of originating the loan. In other words, we can isolate the impact that possible discrimination by GSEs plays in contributing to racial gaps in rejection rates. Obviously, if the marginal impacts of $P(NS)$ on the rejection rates are large, then changes in the probabilities of sale arising from possible discrimination by GSEs could have nontrivial impacts on rejection rates.

²Freddie Mac (1996) asserts that its new "automated underwriting" assures that loan costs are lower and that processing is faster. Still, both Freddie Mac and Fannie Mae are known to exert substantial influence on overall market underwriting standards.