Department of Economics

ON THE ROBUSTNESS OF RACIAL DISCRIMINATION FINDINGS IN MORTGAGE LENDING STUDIES

Judith A. Clarke
Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

Marsha J. Courchane
ERS Group, 2100 M Street NW, Suite 810
Washington, D.C., U.S.A. 20037

Nilanjana Roy
Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

October, 2005

Abstract

The binary logistic regression or logit link model is commonly used to test for racial disparate treatment in fair lending studies undertaken by government agencies, including the Office of the Comptroller of the Currency (OCC) and the Federal Reserve Board (FRB). Ensuring race neutrality in lending remains a concern of regulators and consumer advocates. Improving the understanding of any shortcomings of either bank internal models or regulatory agency models will enable those participants in the mortgage industry to better serve the needs of consumers. We explore this issue using five bank studies undertaken by the OCC. We consider the impact of the logit link assumption, as this determines how race affects the likelihood of loan approval, by moving to three other links: probit, gompit and complementary log log; the latter two are examples of asymmetric links. As our data sets have been obtained using stratified sampling procedures, which has been typical at the OCC, rather than being drawn via simple random sampling, moving away from the logit link complicates estimation; it is no longer possible to use a standard estimation command with an adjustment for stratum effects. Our results reveal that the choice of link function, despite exhibiting similar sample fit, can influence findings of disparate treatment at the nominal level of significance commonly accepted as the legal standard. We also find that the use of a resampling method, which aims to better approximate the finite sample null distribution, for obtaining p-values typically leads to support for discrimination more often than arises from use of the standard normal approximation.

Keywords: Logit; Fair lending; Stratified sampling; Binary response; Semi-parametric maximum likelihood; Pseudo log-likelihood; Profile log-likelihood; Without replacement resampling; Bootstrapping.

1. <u>Introduction</u>

A number of programs that exist today in the U.S. housing finance market have their origin in the Depression era. During that period, the focus was on providing affordable housing to **every** American, as is clear from the National Housing Act of 1934, which states that "the Congress affirms the national goal that every American family be able to afford a decent home in a suitable environment". This focus has changed little over the decades, as is evident from the following statement obtained from the U.S. Housing Policy statement 42 U.S.C. §12702:

The objective of national housing policy shall be to reaffirm the long-established national commitment to decent, safe, and sanitary housing for every American by strengthening a nationwide partnership of public and private institutions able ….(6) to provide every American community with a reliable, readily available supply of mortgage finance at the lowest possible interest rates…

One potential obstacle to this goal of making **every** American family able to afford a decent home is disparate treatment in the mortgage lending market against minorities. Such discrimination can take on various forms including the turning down of a loan application based on certain personal characteristics of the applicant such as race, age, and gender[1], which is prohibited under U.S. laws. Data collected by the Federal Financial Institutions Examination Council (FFIEC) under the Home Mortgage Disclosure Act (HMDA), enacted by the Congress in 1975, are designed to help regulators enforce such fair lending laws. Results from this data indicate that loan approval rates for minority applicants have been and continue to be lower than those of white applicants, but this evidence alone does not allow one to infer that lending discrimination exists, as account must be taken of differences in variables representing creditworthiness.

A large empirical literature exists that has augmented the HMDA data with additional information, including variables on creditworthiness, and has applied statistical models to

---

[1] Discrimination in mortgage lending can take other forms, e.g., prescreening, unfavorable terms for an approved loan and redlining. Our concern is with discrimination in the loan approval process.

estimate the effect of race on the probability of loan approval[2]. One such work is the well-known study by Munnell et al. (1992), commonly known as the Boston Federal Reserve Study (BFS), which found evidence of racial discrimination in mortgage lending[3]. The BFS, and its published version, have received several criticisms (Zandi, 1993; Day and Liebowitz, 1998; Harrison, 1998; Horne, 1994, 1997; Stengel and Glennon, 1999; among others). Comprehensive reviews are given in Ross and Yinger (1999, 2002); they suggest that the findings hold up reasonably well despite the criticisms. One crucial concern with the BFS study is its use of aggregated data across a number of institutions, which might be inappropriate as it assumes that institutions have similar underwriting criteria (e.g., Stengel and Glennon, 1999; Courchane et al., 2000b).

Several government regulatory agencies monitor lending practices of institutions, including the Office of the Comptroller of the Currency (OCC), which uses statistical models and bank-specific approaches as part of most fair lending reviews; see Courchane et al. (2000b) and Dietrich (2005) for a detailed description of OCC review practice. In the instances from which the data in our paper are drawn, regulators from the OCC met with bank underwriters to understand the factors that the bank considers in assessing a loan application. Data were then recorded on relevant variables for a subset of applicants from which statistical models are estimated to arrive at a conclusion regarding racial discrimination in mortgage lending at that bank. The underwriting process was estimated using a binary outcome logistic regression or logit model for a bank, with the dependent variable being loan outcome (approved or denied) for applicants against a set of regressors, such as loan-to-value ratio, debt-to-income ratio, one or more credit score variables, and various dummy variables representing insufficient funds to close, race (e.g., White, African American, Hispanic American) etc. Tests were undertaken on the race coefficients to examine whether the probability of loan approval differs across races; if applicants from one race with similar profiles as applicants from another race face a statistically significant lower probability of approval, then disparate treatment exists.

---

[2] See, for example, LaCour-Little (1999) for an excellent review of the literature on discrimination in mortgage lending. See, also, Courchane et al. (2000b), who examine the use of statistical analysis as a tool for detecting discrimination.
[3] A revised version was published as Munnell et al. (1996). The final conclusion remained unchanged.

The logistic regressions we examine were estimated using a sample from all loan applications processed by the bank that are either approved or denied. It would be preferable to use the information from all such applications, but this data, though collected and kept by the banks for a legally mandated time period, are typically unavailable electronically, which necessitates the drawing of a sample, given the time and cost that would be involved for the assessor to manually code the population data. Population statistics on whether the applicant's loan is approved or denied and on some other covariates, including race, are readily available, being required information by HMDA. This information enables stratification of the applications, with a sample then being formed by randomly selecting from each of the strata rather than a simple random sample from all applications, as is usually assumed to be the case. Stratification can result in substantial gains in efficiency when estimating parameters and in power when undertaking inference; Clarke and Courchane (2005) illustrate such possibilities for fair lending studies. Heterogeneity of characteristics across strata and homogeneity within strata will magnify the gains of stratified sampling over simple random sampling.

How the applications are stratified prior to drawing the sample has implications for estimation and inference. If the applications are divided into strata by race, with a sample then formed from randomly drawing from each of the racial strata, then there are no implications on estimation and inference; such a scheme is sometimes called exogenous sampling (ES), as the stratification is exogenous to the dependent variable of interest. Estimation and inference may or may not be affected when the applications are first divided into strata by loan outcome (those with loans approved and those with loans denied), with the sample then formed from randomly selecting applicants from these two groups; for example, there are implications for the estimation of all parameters for some binary response models (e.g., probit) but only for the stratum effects for others (e.g., logit)[4]. This sampling scheme is often called choice-based sampling (CBS) or case-

---

[4] Given the popularity of logit and probit, we used them as examples of models impacted by stratification. This may suggest that it is the error term distribution that determines whether stratification affects all parameters, which is only partially the story; the key is whether our regression model is a so-called multiplicative intercept model. In a multiplicative intercept model the log-odds-ratio of the probability of a loan being approved is a known function of the regression parameters and the predictors plus a variation-independent intercept parameter; logit satisfies this assumption whereas probit does not. See, for example, Hsieh et al. (1985), Weinberg and Wacholder (1993) and Rabinowitz (1997).

control sampling or endogenous sampling, as the stratification is on the choice or dependent variable, often "cases" of disease and "controls" in medical applications.

The researcher can stratify applications endogenously by loan outcome and exogenously by race (sometimes called stratified choice-based sampling - SCBS), which has implications for estimation of, and inference on, racial dummy variable coefficients when using logit but on all parameters when using probit; e.g., Cosslett (1981b), Scott and Wild (2001). Clarke and Courchane (2005) show, using Monte Carlo experiments, that stratified choice-based sampling is usually preferable to exogenous sampling or choice-based sampling in logistic regressions for fair lending.

The impact of the choice of a logistic regression (that is, a choice of a logit link function) and the use of asymptotic distributions to approximate unknown finite sample distributions, when undertaking hypothesis tests, on the discrimination outcome remain unexplored. We study these issues in this paper. We examine whether, ceteris paribus, the outcome of the non-discrimination test in regression models remains unchanged when we move to other common link functions: probit, gompit and complementary log log (the latter two are examples of asymmetric links), allowing for the impact of stratified choice-based sampling on obtaining estimators and their variance-covariance matrices. We also consider the sensitivity of the inference results to the use of asymptotic approximation theory by undertaking bootstrap experiments, which aim to better approximate the finite sample null distribution, to estimate p-values; we tailor the simulations to take account of the finite sub-population and the adopted stratified sampling procedure. The question of robustness or non-robustness, as the case may be, of findings of discrimination from statistical models commonly used is of importance to the regulatory agencies and the bodies to whom cases are referred (the Department of Justice and the Department of Housing and Urban Development), bank officials and the consultants hired by these banks, academic researchers, and consumers affected by possible discrimination.

This paper is organized into the following sections. Section 2 presents our model setup, including a discussion of the link functions; section 3 considers estimation methods and hypothesis testing procedures when the data are stratified both, endogenously, by the dependent variable and, exogenously, by our categorical race

covariate; section 4 details our data, including particulars on covariates; section 5 provides the empirical results and section 6 concludes.

2. Binary response model, cdfs and link functions

Our adopted statistical models arise from bank-specific examinations that aim to model underwriting practices. A regression models whether a loan is approved or denied as a function of covariates such as loan-to-value ratio (LTV), debt-to-income ratio (DTI), income, one or more credit score variables, dummy variables used to capture bad credit, insufficient funds to close, and race (e.g., White, African American, Hispanic American). More generally, for each bank, we assume a binary outcome dependent variable, y_j, which takes values $y_j = 0$, when a mortgage loan application is denied, and $y_j = 1$, when it is approved; j=1,...,N, the number of applicants whose loan applications have been denied or approved[5]. We suppose K race categories with a vector x_j, of dimension K, which contains categorical dummy variables that describe the race of an applicant: $x_{jk} = 1$ if the j'th applicant belongs to racial group k (k=1,…,K), 0 otherwise; then, $x_j = [x_{j1}, x_{j2}, …, x_{jK}]'$. There is an additional q-dimensional vector z_j that contains the other discrete and continuous variables that describe characteristics of the loan application. Our aim is to estimate a binary response model of the form:

$$h(P_1(w_j;\beta)) = w_j'\beta \qquad , \qquad\qquad j=1,2, ..., N \qquad\qquad (2.1)$$

where, for i=0,1, $P_i(w_j;\beta) = pr(y_j = i \mid w_j;\beta)$ and β is a p-dimensional coefficient vector (p=K+q); $\beta=[\beta_1, \beta_2, …, \beta_K, \beta_{K+1}, …, \beta_p]'$. Having appropriately estimated (2.1), whether a bank is deemed to be racially discriminating is determined by testing whether the impacts of the racial categorical variables are equal; i.e., we test the K!/(2((K-2)!)) distinct null hypotheses, $H_0^j : \beta_j - \beta_k = 0$, j≠k, j, k=1,…,K; against, usually, a one-sided alternative hypothesis (e.g., that discriminatory treatment is towards African Americans over Whites and so on).

---

[5] Typically, applications that were incomplete or withdrawn are excluded from data sets, and approved, but not yet accepted, loans are treated as accepted.

In model (2.1), the vector w_j is partitioned as $w_j' = [x_j'\ z_j']$ and h(.) is the link function. We can equivalently write (2.1) as:

$$P_1(w_j;\beta) = h^{-1}(w_j'\beta) = F(w_j'\beta) \qquad (2.2)$$

where F(.) denotes a cumulative distribution function (cdf) so the link function is the inverse of the cdf. Statistical analyses undertaken by fair lending regulators have, to our knowledge, exclusively considered a logistic cdf, which corresponds to the logit link function[6]:

$$\text{logit}(P_1(w_j;\beta)) = \log[P_1(w_j;\beta)/(1- P_1(w_j;\beta))] = w_j'\beta \qquad (2.3)$$

or, equivalently,

$$P_1(w_j;\beta) = \exp(w_j'\beta)/(1+\exp(w_j'\beta)).$$

Another link function, though not commonly adopted in fair lending studies, is the normit, which results in the probit regression:

$$\Phi^{-1}(P_1(w_j;\ \beta)) = w_j'\beta \qquad (2.4)$$

or, equivalently,

$$P_1(w_j;\beta) = \Phi(w_j'\beta)$$

where Φ is the cumulative distribution function of a standard normal variate.

The logistic cdf has fatter tails than the probit cdf, appoaching zero and one more slowly; we illustrate cdfs or probability plots in Figure One, with corresponding density functions in Figure Two. Often this difference is unimportant, particularly when the predicted probabilities of a loan approval are not extreme. However, when there are large numbers of observations or when many of the predicted probabilities are close to zero or one, choice of a logit or a normit link can lead to different conclusions. The bank data sets we examine range from 145 to 420 observations, not particularly large compared to the thousands of observations often used in the estimation of binary response models, which may lead to little difference between logit and probit models. We provide the percentage distribution of the predicted probabilities for our five banks, denoted as Bank 1 to Bank 5 for confidentiality reasons, using logit in Table 1; a significant percentage of

---

[6] An exception in the academic literature is Courchane et al. (2000a), who adopt generalized entropy models. As our goal is to examine robustness to the choice of link, we do not incorporate this approach.

the predictions are close to one for Banks 2, 3 and 4, suggesting our exploration of probit may be warranted.

The probability $P_1(w_j;\beta)$ approaches zero and one at the same rate for the logit and probit models, as their links are symmetric. This may be a questionable assumption for the sub-populations of bank applications, which feature few denials compared to approvals; an asymmetric link that leads $P_1(w_j;\beta)$ to approach zero and one at different rates may be preferred. Incorrectly assuming a symmetric link will likely lead to substantial bias in our coefficient estimates and detrimentally affect our disparate treatment test; e.g., Czado and Santner (1992). There are many possible classes of parametric links that lead to such skewness; e.g., Stukel (1988) and Chen et al. (1999). We consider two common asymmetric links: gompit and cloglog. The gompit model is:

$$-\log(-\log(P_1(w_j;\beta))) = w_j'\beta \tag{2.5}$$

or, equivalently, $P_1(w_j;\beta) = \exp(-\exp(-w_j'\beta))$, with $P_1(w_j;\beta)$ approaching zero faster than one. The cloglog, or complementary log-log, model is:

$$\log(-\log(1-P_1(w_j;\beta))) = w_j'\beta \tag{2.6}$$

or, equivalently, $P_1(w_j;\beta) = 1 - \exp(-\exp(w_j'\beta))$, with $P_1(w_j;\beta)$ approaching one faster than zero. We illustrate in Figures One and Two.

Table 1: Distribution of Illustrative Predicted Probabilites of Loan Approval[*]

Range for Predicted Probability									
0-<0.10	0.10-<0.20	0.20-<0.30	0.30-<0.40	0.40-<0.50	0.50-<0.60	0.60-<0.70	0.70-<0.80	0.80-<0.90	0.90-<1
Bank 1: N=7013, n=332									
5.1%	7.5%	6.3%	7.2%	11.7%	7.5%	8.1%	12.7%	19.9%	13.9%
Bank 2: N=2959, n=245									
5.7%	3.7%	2.4%	3.3%	2.0%	7.8%	6.1%	6.9%	12.2%	49.8%
Bank 3: N=939, n=340									
8.2%	4.1%	3.5%	2.1%	2.1%	3.2%	5.9%	5.0%	14.1%	51.8%
Bank 4: N=3550, n=420									
10.7%	3.8%	4.8%	4.0%	4.0%	5.5%	6.0%	6.7%	15.0%	39.8%
Bank 5: N=1976, n=228									
1.8%	1.8%	2.6%	2.2%	2.2%	11.0%	16.2%	29.8%	26.3%	3.1%

* The probabilities are from the OCC logistic regressions.

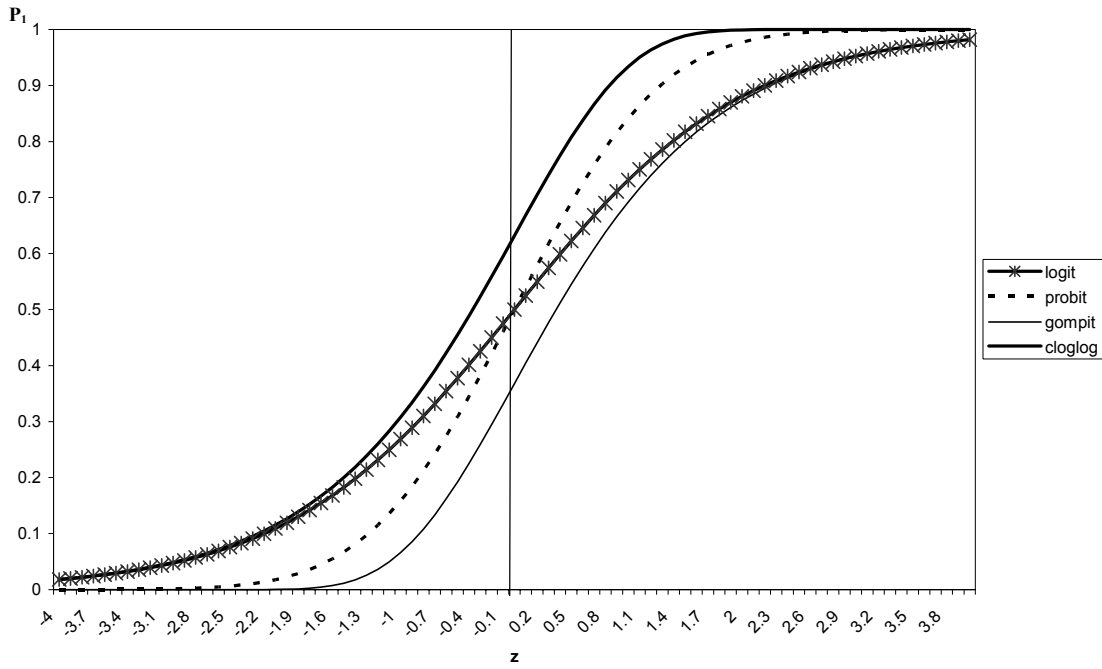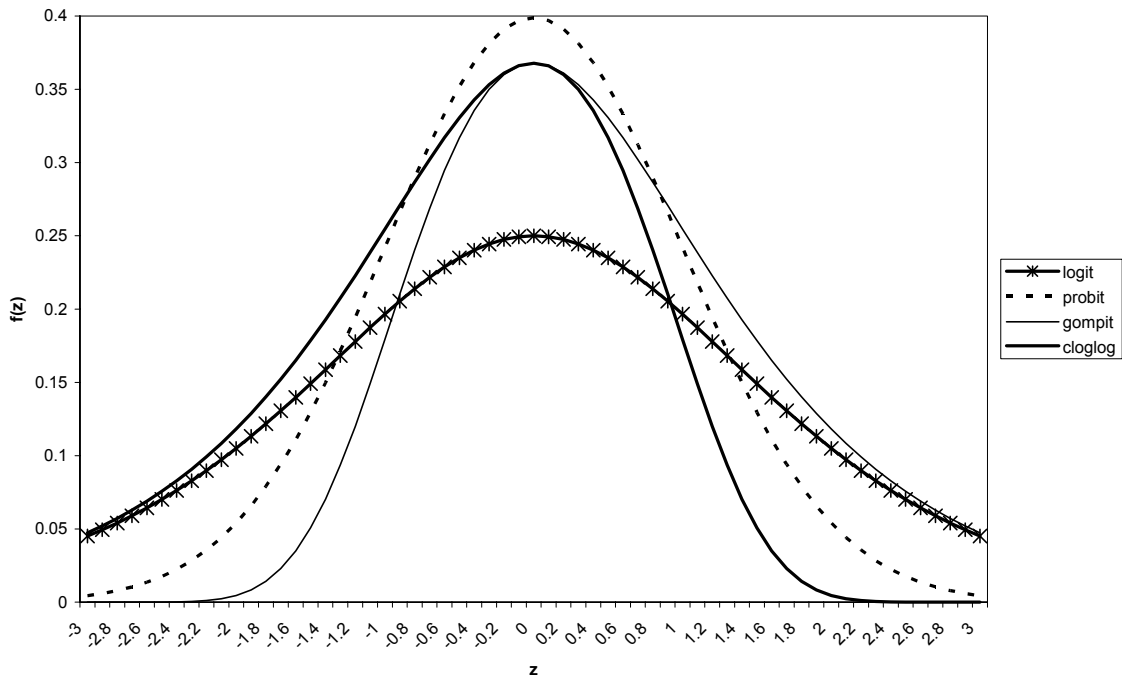**Figure One: Probability Plots (Cumulative Distribution Functions)**

Figure Two: Probability Density Functions

3. Estimation under stratified sampling

The likelihood, for an arbitrary cdf or link function, assuming that the applicant is from a simple random sample from all those who present at the bank, is:

$$L^{SRS} = \prod_{j=1}^{N} pr(y_j, w_j \mid \beta) = \prod_{j=1}^{N} pr(y_j \mid w_j; \beta) g(w_j).$$

(3.1)

As we assume that $g(w_j)$ does not involve β, we can base estimation on:

$$L = \prod_{j=1}^{N} pr(y_j \mid w_j; \beta) = \prod_{j=1}^{N} P_1(w_j; \beta)^{y_j} P_0(w_j; \beta)^{(1-y_j)}.$$

(3.2)

Unfortunately, in practice, all of the variables in w_j are not electronically available to the examiner. For cost and efficiency reasons, a sample of size n is drawn from the N available, files are pulled for the chosen applicants, data are manually coded and regression models estimated; the final step is to undertake the disparate treatment hypothesis tests based on the sample of n applicants, rather than the subpopulation of size N. Simple random sampling (SRS) would be one way to draw the n applications; then the likelihood contribution of interest is of the form of (3.2) with N replaced by n. This approach, however, ignores additional information that the bank is required to collect: loan disposition and racial class are electronically available for each subpopulation member because of HMDA requirements. Let $N_{i,k}$ be the number of applicants in racial class k with $y_j=i$, i=0,1, k=1,2,...,K; $\sum_{i=0}^{1} \sum_{k=1}^{K} N_{i,k} = N$. This additional information can be partially or fully incorporated into a sampling scheme:

(i) Exogenous sampling (ES), for which the information on the race of each applicant only is used prior to drawing the sample. Let N_{+k} be the number of applicants in race k; $N_{+,k}=N_{0,k}+N_{1,k}$, $\sum_{k=1}^{K} N_{+,k} = N$. ES randomly draws $n_{+,k}$ applicants from the $N_{+,k}$ available; $\sum_{k=1}^{K} n_{+,k} = n$. This sampling scheme ignores that the outcome of the loan application is known for each applicant.

(ii) Choice-based sampling (CBS) or case-control sampling, for which the information on y_j only is used prior to drawing the sample. Let $N_{i,+}$ be the number of applicants with

$y_j = i$; $N_{i,+} = \sum_{k=1}^{K} N_{i,k}$, $N_{0,+} + N_{1,+} = N$. We randomly select $n_{i,+}$ applicants from the $N_{i,+}$

available with CBS, ignoring the race of each applicant when forming the sample; $n_{0,+} + n_{1,+} = n$.

(iii) Stratified choice-based sampling (SCBS), for which the prior knowledge on the race *and* outcome of the loan application for each case is used prior to drawing the sample. Under SCBS, we draw $n_{i,k}$ applicants from the $N_{i,k}$ available, $i=0,1$, $k=1,2,...,K$;

$$\sum_{i=0}^{1} \sum_{k=1}^{K} n_{i,k} = n \,.$$

Using the known prior information about race and loan outcome in these ways can change the likelihood for β, as applicants are potentially sampled at different rates from strata. Under ES, the assessor draws an applicant j from race k (that is, fixes x_{jk}) and observes whether the application has been approved or denied; corresponding values for z_{jk} are then recorded. The likelihood under such a sampling scheme is:

$$L^{ES} = \prod_{k=1}^{K} \prod_{j=1}^{n_{+,k}} pr(y_{jk}, z_{jk} \mid x_{jk}; \beta)\, g(x_{jk})$$

$$= \prod_{k=1}^{K} \prod_{j=1}^{n_{+,k}} pr(y_{jk} \mid w_{jk}; \beta) g(z_{jk}) g(x_{jk}) \qquad (3.3)$$

using Bayes' Rule. As $g(z_{jk})$ and $g(x_{jk})$ are assumed not to depend on β, we can write (3.3) as:

$$L^{ES} = \left[\prod_{k=1}^{K} \prod_{j=1}^{n_{+,k}} pr(y_{jk} \mid w_{jk}; \beta) \right] \left[\prod_{k=1}^{K} \prod_{j=1}^{n_{+,k}} g(z_{jk}) g(x_{jk}) \right] = L_1^{ES} L_2^{ES} \qquad (3.4)$$

so that estimation for, and inference on, β can be based solely on L_1^{ES}; that is, on:

$$L_1^{ES} = \prod_{k=1}^{K} \prod_{j=1}^{n_{+,k}} pr(y_{jk} \mid w_{jk}; \beta) = \prod_{j=1}^{n} pr(y_j \mid w_j; \beta) \qquad (3.5)$$

which corresponds to the likelihood under SRS; i.e., we can assume that a sample obtained via ES is a simple random sample (see, e.g., Manski and Lerman, 1977, and Cosslett, 1981a,b).

However, such is not the case with CBS or SCBS; e.g., Anderson (1972), Prentice and Pyke (1979), Cosslett (1981a,b), Scott and Wild (2001) and the many references

therein. We define two strata with CBS: those cases with $y_j=1$ and those with $y_j=0$, from which we draw two independent random samples: $n_{i,+}$ applicants from those units with $y_j=i$. We then record the corresponding w_{ij} values, where we now subscript w by i, in addition to j, given the sampling scheme used: $i=0,1$, $j=1,2,...,n_{i,+}$. The likelihood for such a sampling scheme is:

$$L^{CBS} = \prod_{i=0}^{1} \prod_{j=1}^{n_{i,+}} pr(w_{ij} \mid y_j = i)$$

$$= \prod_{i=0}^{1} \prod_{j=1}^{n_{i,+}} pr(y_j = i \mid w_{ij})g(w_{ij})/pr(y_j = i)$$

$$= \prod_{i=0}^{1} \prod_{j=1}^{n_{i,+}} P_i(w_{ij};\beta)g(w_{ij})/pr(y_j = i) \qquad (3.6)$$

using Bayes' Rule. As $pr(y_j=i)=\int P_i(w_{ij};\beta)dG(w_{ij})$, where G(.) denotes the appropriate marginal distribution function, we cannot separate out $g(w_{ij})$ when estimating β.

Similarly, under SCBS, where there are S=2K strata, from which we sample $n_{i,k}$ units with $y_j=i$ and x_j is such that the case belongs to race k, which we denote by $x_j \in k$, and we then record the corresponding w_{ijk} values; the k subscript noting that the case belongs to the k'th race class, $k=1,2,...,K$, $i=0,1$, $j=1,2,...,n_{i,k}$. The likelihood is then:

$$L^{SCBS} = \prod_{i=0}^{1} \prod_{k=1}^{K} \prod_{j=1}^{n_{i,k}} pr(w_{ijk} \mid y_j = i, x_j \in k)$$

$$= \prod_{i=0}^{1} \prod_{k=1}^{K} \prod_{j=1}^{n_{i,k}} pr(y_j = i \mid w_{ijk}, x_j \in k)g(w_{ijk} \mid x_j \in k)/pr(y_j = i \mid x_j \in k)$$

$$= \prod_{i=0}^{1} \prod_{k=1}^{K} \prod_{j=1}^{n_{i,k}} P_i(w_{ijk};\beta \mid x_j \in k)g(w_{ijk} \mid x_j \in k)/pr(y_j = i \mid x_j \in k) \qquad (3.7)$$

using Bayes' Rule, from which we again cannot separate out the density for w, given the dependence of $pr(y_j=i \mid x_j \in k)$ on β. We write (3.7) in log-likelihood form:

$$\ell = \sum_{i=0}^{1} \sum_{k=1}^{K} \sum_{j=1}^{n_{i,k}} \log P_i(w_{ijk};\beta \mid x_j \in k) + \sum_{i=0}^{1} \sum_{k=1}^{K} \sum_{j=1}^{n_{i,k}} \log g(w_{ijk} \mid x_j \in k)$$

$$- \sum_{i=0}^{1} \sum_{k=1}^{K} \sum_{j=1}^{n_{i,k}} \log pr(y_j = i \mid x_j \in k) \qquad (3.8)$$

As our focus is on SCBS, we explore the maximization of (3.8), although the discussion applies equally to samples generated by CBS. We use semiparametric maximum likelihood estimation, where the term "semiparametric" is taken to mean that we parametrically model $P_i(w_{ijk};\beta|\ x_j \in k)$ (for example, using one of the links provided in the previous section) and we nonparametrically model $g(w_{ijk}|x_j \in k)$; e.g., Cosslett (1981b), Lawless et al. (1999), Neuhaus et al. (2002), Scott and Wild (1991, 1997, 2001), among others. The literature proposes two routes for solving for estimates for β: maximizing either a profile log-likelihood or a pseudo log-likelihood. The former, considered in the next sub-section, leads to maximum likelihood estimates irrespective of the form of the link function, but is less user-friendly in the sense of not being straightforward to code in standard packages (e.g., EViews, SPSS and Stata). The alternative path of maximizing a pseudo log-likelihood, which we detail in sub-section 3.2, is uncomplicated to code, but, for many common link functions, has severe computational issues. Accordingly, we also consider a computationally simpler estimator, which is consistent, but not usually asymptotically efficient, that is available via the pseudo log-likelihood route.

3.1 A profile log-likelihood route

Without proof (see Scott and Wild, 2001), the profile log-likelihood for β $\ell_P(\beta) = \ell(\beta, \hat{g}(\beta))$, after nonparametrically modeling the density of w by replacing its (unknown) cumulative probability distribution with its empirical distribution[7], is:

$$\ell_P(\beta) = \ell^*(\beta, \rho(\beta)) = \sum_{j=1}^{n} \Big\{ (1 - y_j)\log(1 - P_1(w_j;\beta)) + y_j \log P_1(w_j;\beta) -$$

$$\sum_{k=1}^{K} \Big(S_{jk} \log\big[\mu_{0k}P_0(w_j;\beta) + \mu_{1k}P_1(w_j;\beta)\big] - \big(m_{1k}\rho_{1,k} -$$

$$(m_{0k} + m_{1k})\log(1 + \exp(\rho_{1,k}))\big)/n\Big\} \qquad (3.9)$$

---

[7] The empirical distribution is the maximum likelihood estimate of an unknown distribution function; e.g., Kiefer and Wolfowitz (1956).

where: $m_{ik} = (N_{i,k}-n_{i,k})$; $\mu_{ik} = N_{+,k} - \dfrac{m_{ik}(1+\exp(\rho_{1,k}))}{(\exp(\rho_{1,k}))^i}$; $S_{jk}=1$ if the j'th applicant

belongs to stratum k, 0 otherwise; i=0,1, k=1, …, K, j=1,2, …, n. Excluding variance-covariance matrix parameters, this objective function has (p+K) unknown parameters, p from β and K from $\rho_{1,1}$ … $\rho_{1,K}$, which arise from the nonparametric modeling of the density of w; these additional parameters relate to unconditional probabilities. Specifically, let $Q_{i,k}$ be the unconditional probability that y=i in stratum k

with $\displaystyle\sum_{i=0}^{1} Q_{i,k} = 1$, then $\rho_{i,k}=\log(Q_{i,k}/Q_{0,k})$ with $\rho_{0,k}=0$.

The criterion (3.9) is highly non-linear in β and ρ ($=[\rho_{1,1}$ … $\rho_{1,K}]$), although, for fixed β, the ρ parameters are orthogonal, as each involves only observations from the relevant stratum. We apply the iterative routine suggested by Scott and Wild (2001, p.18) to solve for the maximum likelihood solutions, say $\hat{\beta}_{PR}$ and $\hat{\rho}_{PR}$; throughout this paper, a subscript "PR" will refer to a statistic or a p-value obtained by means of the profile log-likelihood.   Specifically, the additional sub-population information on $N_{i,k}$ provides initial, consistent, estimates of $\rho_{1,1}$ … $\rho_{1,K}$, say $\overline{\rho}_{1,1}…\overline{\rho}_{1,K}$, which are then used to maximize (3.9) for estimates of β, say β^*. With β fixed at β^*, we again maximize (3.9) to obtain new ρ estimates and so on until we converge to $\hat{\beta}_{PR}$ and $\hat{\rho}_{PR}$. When solving for β, our algorithm used the score vector and information matrix provided by Scott and Wild (2001, p.18). Convergence usually resulted in fewer than five such major loops, with ten major loops being the highest number required for our data sets.

3.2 A pseudo log-likelihood route

Without proof (e.g., Cosslett, 1981b; Scott and Wild, 2001), when we model g(.) nonparametrically, maximizing ℓ is equivalent to maximizing the pseudo log-likelihood function:

$$\ell^* = \sum_{i=0}^{1} \sum_{k=1}^{K} \sum_{j=1}^{n_{i,k}} \log P_i^*(w_{ijk};\beta,\kappa_k) \qquad (3.10)$$

with $P_i^*(w_{ijk};\beta,\kappa_k)$ defined by setting

$$\text{logit}\,P_i^*(w_{ijk};\beta,\kappa_k) = \text{logit}\,P_i(w_{ijk};\beta \mid x_j \in k) + \log \kappa_k . \qquad (3.11)$$

The parameter κ_k is the ratio of the sampling rates for race class k:

$$\kappa_k = \left(\frac{n_{1,k}}{\text{pr}(y_j = 1 \mid x_j \in k)} \right) \Big/ \left(\frac{n_{0,k}}{\text{pr}(y_j = 0 \mid x_j \in k)} \right) \qquad (3.12)$$

and

$$\text{logit}\,P_1(w_{1jk};\beta \mid x_j \in k) = \log \left(\frac{P_1(w_{1jk};\beta \mid x_j \in k)}{P_0(w_{0jk};\beta \mid x_j \in k)} \right). \qquad (3.13)$$

The objective function (3.10) is termed a "pseudo log-likelihood" because in general it is not equal to the log-likelihood ℓ; they are equal at their maximums.

A logit link model with stratum constants has $\text{logit}\,P_1(w_{1jk};\beta \mid x_j \in k) = \beta_k + z_j'\overline{\beta}$, where $\overline{\beta}$ is a q-dimensional parameter vector, with P_1^* also logistic in form:

$\text{logit}\,P_1^*(w_{1jk};\beta,\kappa_k) = \beta_k + \log \kappa_k + z_j'\overline{\beta}$, k=1,2,...,K. Clearly, we cannot identify both β_k and κ_k without the use of supplementary information; separate identification is crucial as our disparate treatment hypothesis tests are of the equality of race stratum effects. These results hold more generally, indeed for any multiplicative intercept model where $\text{logit}\,P_1(w_{1jk};\beta \mid x_j \in k) = \beta_k + H(z_j,\overline{\beta})$.

The probit regression model, on the other hand, is an example of a non-multiplicative model as $\text{logit}\,P_1(w_{1jk};\beta \mid x_j \in k) = H(\beta_k + z_j'\overline{\beta})$; we have

$$\text{logit}\,P_1(w_{1jk};\beta \mid x_j \in k) = H(\beta_k + z_j'\overline{\beta}) = \log \left[\frac{\Phi(\beta_k + z_j'\overline{\beta})}{1 - \Phi(\beta_k + z_j'\overline{\beta})} \right]$$

so that

$$\text{logit}\,P_1^*(w_{1jk};\beta,\kappa_k) = \log \left[\frac{\kappa_k \Phi(\beta_k + z_j'\overline{\beta})}{1 - \Phi(\beta_k + z_j'\overline{\beta})} \right] \qquad (3.14)$$

and

$$P_1^*(w_{1jk};\beta,\kappa_k) = \left[\frac{\kappa_k \Phi(\beta_k + z_j'\overline{\beta})}{1 - \Phi(\beta_k + z_j'\overline{\beta}) + \kappa_k \Phi(\beta_k + z_j'\overline{\beta})} \right]. \qquad (3.15)$$

The gompit link and cloglog link models are also non-multiplicative intercept models and can be similarly expressed.

Thus, $\kappa_1,...,\kappa_K$ are identifiable in a non-multiplicative intercept model, such as probit, gompit and cloglog, although there may be some multicollinarity issues that might cause convergence concerns. To further complicate computational matters, the stationary point of (3.10) occurs at a saddlepoint in the combined parameter space; Cosslett(1981b), Scott and Wild (2001). Therefore, to summarize where we are at, there is an identification issue in multiplicative intercept models and computational concerns associated with a saddlepoint solution that needs addressing to proceed to obtain the maximum likelihood estimators under SCBS via the pseudo log-likelihood.

This may suggest that it is preferable to avoid working with the pseudo log-likelihood (3.10), which only requires minor coding in packages (e.g., EViews, Gauss, SAS, SPSS, Stata). However, the supplementary information available on sub-population stratum totals enables us to consistently estimate κ_k; specifically:

$$\hat{\kappa}_k = \left(\frac{n_{1,k}}{N_{1,k}} \right) \Big/ \left(\frac{n_{0,k}}{N_{0,k}} \right) \tag{3.16}$$

is a consistent estimator of κ_k. Use of this rule for the logit link leads to the estimator of β examined by Clarke and Courchane (2005) in their fair lending study; this estimator is known to be in fact the maximum likelihood estimator of β (Anderson, 1972; Prentice and Pyke, 1979)[9]. That is, for the logit link, maximum likelihood estimates of all the parameters, except stratum constants, are obtained by estimating the logit model as if it were from a simple random sample; a simple minor adjustment provides the maximum likelihood estimates of stratum constants.

With non-multiplicative links, use of $\hat{\kappa}_k$ will lead to a consistent, but not necessarily asymptotically efficient, estimator of β - we denote this estimator as $\hat{\beta}_{PS}$[10] - a pseudo log-likelihood one-step estimator; hereafter, a subscript "PS" will refer to a statistic or p-

---

[9] This holds for multiplicative intercept models with a complete set of stratum constants; e.g., Hsieh et al. (1985).
[10] It is, in fact, one form of the Manski-McFadden (1981) estimator.

value obtained via the pseudo log-likelihood. Obtaining the maximum likelihood estimator requires iteration, taking account that we are locating a saddlepoint. There are computational difficulties with this, while it is computationally straightforward to obtain $\hat{\beta}_{PS}$. Comparing outcomes for our disparate treatment tests using the (consistent but asymptotically inefficient) one-step pseudo log-likelihood estimator, $\hat{\beta}_{PS}$, and the maximum likelihood estimator obtained by iteration via the profile log-likelihood, $\hat{\beta}_{PR}$, is instructive, as the former is easier to code. It may well be that the gains in efficiency do not lead to practically significant changes in test outcomes.

3.3 Variance-covariance matrix

Our discussion, so far, has centered on estimation of the parameters of the response model under SCBS. Testing the null hypotheses of interest also requires variance-covariance matrices for our estimators obtained from the profile and pseudo log-likelihood routes.

When using the pseudo log-likelihood procedure for either the logit link or another multiplicative intercept model, a consistent estimator of $\text{var}(\hat{\beta}_{PS})$, say $\text{var}_{est}(\hat{\beta}_{PS})$, is given by (e.g., Scott and Wild, 1986):

$$\text{var}_{est}(\hat{\beta}_{PS}) = \text{var}^*(\hat{\beta}_{PS}) - \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \tag{3.17}$$

where $\text{var}^*(\hat{\beta}_{PS})$ is the inverse of the pseudo-information matrix for $\hat{\beta}_{PS}$, assuming simple random sampling, and A is a (K×K) diagonal matrix with elements:

$$\left\{ a_k = \left[\left(\frac{1}{n_{0,k}} + \frac{1}{n_{1,k}} \right) - \left(\frac{1}{N_{0,k}} + \frac{1}{N_{1,k}} \right) \right] \right\} ; \text{ k=1,2, …, K.} \tag{3.18}$$

The first term is the reduction in variance from stratifying, while the second term is the increase in variance arising from using $\hat{\kappa}_k$ to estimate κ_k.

When using a non-multiplicative intercept model, such as probit, gompit and cloglog, the one-step estimator $\hat{\beta}_{PS}$ is obtained by maximizing the pseudo log-likelihood (3.10) with $\hat{\kappa}_k$ as the estimator of κ_k, and a consistent estimator of $\text{var}(\hat{\beta}_{PS})$ is given by

var*($\hat{\beta}_{PS}$), the inverse of the pseudo-information matrix; see, e.g., Scott and Wild (2001).

The disparate treatment null hypotheses – $H_0^j : \beta_j - \beta_k = 0$, j≠k, j, k=1,…,K, are tested

using $t_{PS}^j = (\hat{\beta}_{PS,j} - \hat{\beta}_{PS,k}) / s.e.(\hat{\beta}_{PS,j} - \hat{\beta}_{PS,k})$, where $\hat{\beta}_{PS} = [\hat{\beta}_{PS,1}, \hat{\beta}_{PS,2}, …, \hat{\beta}_{PS,K}, …,$

$\hat{\beta}_{PS,p}]'$ and $s.e.(\hat{\beta}_{PS,j} - \hat{\beta}_{PS,k}) = \sqrt{var*(\hat{\beta}_{PS,j}) + var*(\hat{\beta}_{PS,k}) - 2cov(\hat{\beta}_{PS,j}, \hat{\beta}_{PS,k})}$. It

follows (e.g., Cosslett, 1981b, Scott and Wild, 2001), that the limiting null distribution

for t_{PS}^j is standard normal, from which p-values are usually obtained. We provide these

p-values in section 5, in addition to bootstrapped p-values, the methodology for which is

in the next sub-section.

As we use the analytic score vector and Hessian matrix in our routine to solve for the

maximum likelihood estimator, $\hat{\beta}$, by way of the profile log-likelihood, we estimate this

estimator's asymptotic covariance matrix as the inverse of the information matrix,

evaluated at the maximum likelihood estimates; see. e.g., Scott and Wild (2001, pp. 14-

15). Obvious modifications to the statistics t_{PS}^j provides the rules we use to examine

$H_0^j : \beta_j - \beta_k = 0$ when using the profile log-likelihood estimators; let t_{PR}^j denote this test

statistic, which has a standard normal limiting null distribution, from which we find p-

values. We do not provide bootstrapped p-values for these cases, although this is

feasible, because, as we shall see, there is little practical qualitative difference between

the test outcomes using the one-step pseudo and profile log-likelihood methods.

3.4 <u>Bootstrapped p-values</u>

To allow for the finite sub-population of N applicants presenting at a bank and the use

of SCBS to form the sample of n applicants, when forming our bootstrapped p-values we

take the following steps, primarily suggested by Booth et al. (1994) and Presnell and

Booth (1994).

Step 1: The first step is to create an empirical subpopulation for a bank. Let $f_{i,k} = n_{i,k}/N_{i,k}$

so that $N_{i,k} = g_{i,k} n_{i,k} + s_{i,k}$, $0 \le s_{i,k} \le n_{i,k}$, $g_{i,k} = int(1/f_{i,k})$, i=0,1, k=1,2,…,K. If $g_{i,k}$ is an integer

for all i,k then we can create a unique empirical subpopulation by combining $g_{i,k}$ copies

of the kth stratum's sample; e.g., Gross (1980). More often than not, this is not possible,

as, typically, one or more $g_{i,k}$ are not integers. Then, we create an empirical subpopulation by combining $g_{i,k}$ copies of the appropriate stratum's sample with a without replacement sample of size $s_{i,k}$ from the original sample.

Step 2: We draw B without replacement resamples of size n, stratified as per the original sample, from the empirical subpopulation; i.e., each resample has stratum denial ratios that match the original sample. For a particular link choice, we estimate the regression models for each resample, forming B values of the K!/(2((K-2)!)) test statistics to examine $H_0^d : \beta_j - \beta_k = 0, j \neq k$, j,k=1,...,K, d=1,..., K!/(2((K-2)!)); denote the bootstrapped statistics as $t_1^d, \dots t_B^d$. As our data may not have been drawn from a subpopulation that satisfies H_0^d, we follow the advice of Hall and Wilson (1991) by centering when forming these bootstrapped statistics, which has the effect of increasing power. That is: we form

$$t_{PS,i}^d = \frac{(b_i^j - b_i^k) - (\hat{\beta}_{PS,j} - \hat{\beta}_{PS,k})}{se(b_i^j - b_i^k)} \; ; \; i=1,\dots,B$$

where b_i^j is the estimate of β_j from the ith bootstrap resample and so on[11].

Step 3: Let $t_{PS,samp}^d$ be the statistic value for the original sample for testing H_0^d. The bootstrapped p-value is then the simulated number of rejections obtained by comparing $t_{PS,1}^d \dots t_{PS,B}^d$ with $t_{PS,samp}^d$; e.g., the bootstrapped p-value is

$$p^d = (1/B)\sum_{i=1}^{B} I(t_{PS,i}^d < t_{PS,samp}^d) \text{ when the alternative hypothesis is } H_a^d : \beta_j - \beta_k < 0.$$

Step 4: Repeat *Steps 2* and *3* for each bank using the other links.

We follow the pretesting method advocated by Davidson and MacKinnon (2000) to choose B, the number of bootstraps; typically, this led to B=99 for our chosen 5% nominal level of significance.

4. Data

---

[11] As we are sampling from a finite subpopulation, we resample without replacement, rather than with replacement, as the latter would not be consistent with our original data collection.

The data used in this research were collected by the OCC in the course of several fair lending examinations in the late 1990s.[12] The data here come from five separate national banks geographically distributed from the East to the West and the Midwest. The logistic approach requires that a certain minimum number of files be distributed across the values of the explanatory variables for that variable to be viable. To allow the use of many underwriting variables, the examinations were restricted to bank loan products that had at least 50 minority denials. Each statistical model was structured to best reflect banks' underwriting procedures in the approval or denial of a mortgage application. The logit estimations use as independent variables a combination of explicit elements collected from bank loan files and variables that were created from the primary data to measure credit worthiness.

The data elements typically collected at the examination include a number of elements from a credit bureau report on the borrower applicant such as credit score, public record information (bankruptcies, foreclosures, and unpaid collections and judgments), numbers and types of credit derogatories in the past 24 months, levels of current debts, and numbers of trade lines. Additional collected variables might include employment status (self-employed or not), the presence of gift funds, levels of reserves, job stability and the particular loan program through which the application was submitted. Each bank might use these variables differently in their credit decisions and an attempt was made to replicate the underwriting guidelines as closely as possible. A number of variables used by the banks to assist with their decisions were also created by the examination team. These include variables such as debt-to-income ratio, loan-to-value ratio, threshold variables indicating whether or not the applicant exceeded the bank's guidelines for ratios, insufficient funds to close (a reserves measure) and often a variable that indicated whether the applicant did or did not receive private mortgage insurance if they had applied for it. Sometimes an explanation variable was included that simply noted if the bank had asked for, received, or accepted explanations for credit bureau or underwriting elements from the applicant. At a few institutions, the existence of a banking relationship for the applicant was included as an independent variable and

---

[12] Examination procedures at the OCC may no longer follow these procedures. See, e.g., Dietrich (2005), an economist at the OCC, for up-to-date information.

for those with very low values of LTV an indicator variable noting that that low value served as a compensating factor in the lending decision was included. For some loan programs, collateral needs were noted. Table 2 provides a list of regressors included in the model specifications for our five banks, and brief broad meanings of the covariates are given in Table 3.   The specific definition of each variable used in the modeling process depends on bank-specific factors; e.g., DTI is a one/zero binary regressor with a threshold DTI ratio determining the switch for one bank, while it is the actual DTI ratio for another bank. Despite these differences, the list indicates that many variables are common across institutions.

Table 2: Explanatory Variables

| Variable | Bank | | | | |
	Bank 1	Bank 2	Bank 3	Bank 4	Bank 5
Credit score	×	×	×	×	×
LTV	×	×	×	×	×
Public record	×		×		×
Insufficient funds			×	×	
DTI	×	×	×	×	×
HDTI			×		
PMI			×		
Bad credit	×		×	×	×
Gifts/grants	×		×		
Relationship			×		
Income/savings				×	
Explanation	×		×		
Gender		×			
White	×	×	×	×	×
African American	×	×		×	×
Hispanic American	×		×	×	×

As previously noted, the modeling is undertaken using a stratified sample from the subpopulation of applicants presenting at the bank. A predetermined number of denied and approved loan applications, further stratified by racial group, are drawn. This sampling strategy results in sample racial stratum denial rates that differ from those for the subpopulation. We provide denial rates in Figure Three. Racial groups are denoted as follows: Whites - k=1; African Americans - k=2; Hispanic Americans - k=3. We see

Table 3: Broad Variable Definitions

Variable	Definition
Credit Score	Derived from the bank's underwriting guidelines manual. Typically, a specified procedure is used to calculate a score variable, combining information across obtained credit bureau scores and the applicant and any co-applicant.
LTV	Loan-to-value ratio. May also be a dummy variable equal to 1 if the loan-to-value ratio exceeds specific guidelines; otherwise 0.
Public record	Public record information, created to be approximately uncorrelated with the bad credit variable.
Insufficient funds	Dummy variable equal to 1 if there were not sufficient funds to close.
DTI	Debt-to-income (gross) ratio. May also be a dummy variable equal to 1 if DTI value exceeds bank guidelines; otherwise 0
HDTI	House payment-to-income (gross) ratio
PMI	Dummy variable equal to 1 if the applicant applied for private mortgage insurance and was denied
Bad credit	Derived from bank specific information on credit records. Equal to 1 if a bad credit element is observed, or this variable may be number of derogatories or delinquencies depending upon the underwriting standards of the bank.
Gifts/grants	Sum of gifts and grants, which may provide down payment information.
Relationship	Dummy variable equal to 1 if the applicant has any type of relationship with the bank, such as deposits or previous loan at the bank.
Income/savings	Income and savings information
Explanation	Various dummy variables equal 1 if the bank asked for, received, or accepted explanations for credit bureau or other underwriting elements; 0 otherwise
Gender	Dummy variable equal to 1 if the applicant is Female; 0 otherwise
White	Dummy variable equal to 1 if the applicant is White; 0 otherwise
African American	Dummy variable equal to 1 if the applicant is African American; 0 otherwise
Hispanic American	Dummy variable equal to 1 if the applicant is Hispanic American; 0 otherwise

Figure Three: Bank Denial Ratios
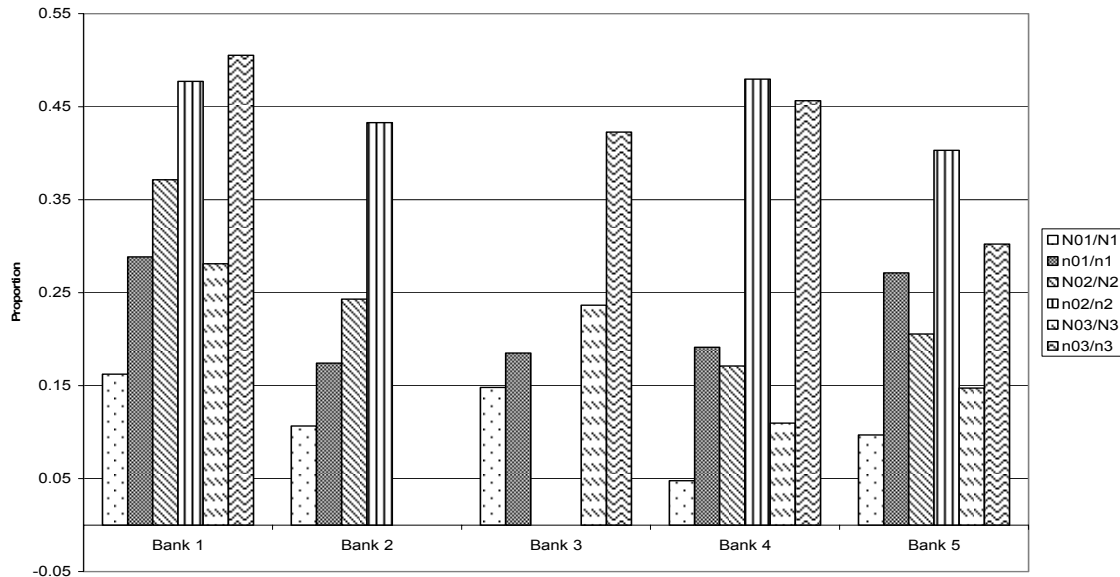
that for Banks 1,4 and 5 there are three racial strata (K=3), while for Banks 2 and 3 there are only two (K=2). The subpopulation measures are denoted by "N", the sample measures by "n", and denial of a loan application by "0"; e.g., "N01" is the number of denied whites loans, "n2" is the number of African Americans in the sample, and so on. We observe denial rates for African Americans that always exceed those for Whites and, when present, the denial rates for Hispanic Americans fall between those for African Americans and Whites.

5. Results

   To assess the impact of the link function on the disparate treatment test outcome, we estimated the five bank-specific models, with the covariates summarized in Table 2, using the estimators $\hat{\beta}_{PS}$ and $\hat{\beta}_{PR}$ for the four links detailed in section 2; recall that these two estimators are equivalent for the logit link but not for the other three studied links. We used Gauss, with the MAXLIK sub-routine, to obtain the maximum likelihood estimates from the profile log-likelihood, and EViews, Stata and Gauss – to satisfy ourselves that results were similar across standard packages – to obtain the one-step estimates from the pseudo log-likelihood.

Prior to comparing p-values, we detail two measures of fit for the models, as these may provide guidance on link preference for a particular bank. The value of the average log-likelihood function offers one way to assess model fit. Table 4 provides this information, with the measures given relative to the average log-likelihood value for the logit link; e.g., a number less than one indicates that the logit link has a smaller average log-likelihood value. Although the results suggest this measure is similar across link functions, with the average log-likelihood values being different by at most 7%, there is some preference for the logit link for Bank 5, the gompit link for both Banks 1 and 4 and the probit link for Bank 2.   Interestingly, for Bank 3, the logit link is outranked by the other link choices when using the profile approach, but is favoured with the pseudo route. Usually, irrespective of bank, the cloglog link is least preferred; this link, we recall, leads to $P_1(w_j; \beta)$ approaching one faster than zero.

Table 4: Relative average log-likelihood values

Bank/ Method	Regression Model			
	logit	probit	gompit	cloglog
Bank 1				
profile	1	1.000	0.989	1.002
pseudo	1	1.003	0.999	1.014
Bank 2				
profile	1	0.999	1.000	0.999
pseudo	1	0.999	1.003	1.020
Bank 3				
profile	1	0.983	0.983	0.983
pseudo	1	1.008	1.007	1.025
Bank 4				
profile	1	1.002	0.989	1.004
pseudo	1	1.026	0.999	1.056
Bank 5				
profile	1	1.000	1.001	1.001
pseudo	1	1.005	1.004	1.007

As the logit link's average profile log-likelihood and pseudo log-likelihood values are identical, the numbers in Table 4 also provide one measure of loss, for the non-multiplicative links, in using the one-step pseudo log-likelihood approach over the profile log-likelihood method. For the banks we examine, the loss in average log-likelihood

value is at most 5.2% with the average loss being 1.6%; this suggests, according to this measure, that it may be practically reasonable to work with the computationally easier pseudo log-likelihood.

Another commonly reported measure of model performance is the percentage correctly predicted, obtained by comparing the predicted and observed outcomes of the binary response. Classification of the predicted probabilities into 0/1 outcomes is achieved by relating them to a chosen cutoff value and counting the matches of observed and predicted outcomes; a classification is "correct" when the model predicts the applicant's loan disposition.   We provide this information in Tables 5a and 5b, using three cutoff values – the standard value of "0.5", a reasonable choice in samples with a balance of 1/0 outcomes, "sf", which is the frequency of y=1 observations in the sample, and "spf", which is the frequency of y=1 observations in the subpopulation; Table 5a presents the outcomes from the pseudo log-likelihood approach, while those from the profile log-likelihood route are given in Table 5b. As our subpopulations are unbalanced, as are also the samples despite the OCC's oversampling of denials, the "spf" and "sf" cutoffs are likely more realistic and sensible; e.g., Cramer (1999).

We observe only minor differences, usually, between the profile and pseudo log-likelihood percentages. Even for the few cases when there are practical differences, it is often less than two percentage points, although we do observe significant differences for the cloglog link for some banks.

The outcomes highlight the influence of the choice of cutoff value; when it is "0.5" or "sf", the models do far better at predicting approvals than denials, while their performance is more equitable when "spf" is adopted as the cutoff. Indeed, for some banks with "spf", the models do better at predicting denials rather than approvals. Overall, the models correctly classify, approximately, 65% to 90% of outcomes, irrespective of cutoff value.

Contrasting prediction abilities across links, we observe little difference, although, as expected, given its asymmetry, the gompit's ability to predict loan approvals is the same or higher than for the other links, with an associated minor loss, usually, in predicting

Table 5a: Percentage correctly predicted from pseudo log-likelihood route

Bank/	Loan Outcome											
Cutoff	Denied (y=0)				Approved (y=1)				Overall			
Value	logit	Probit	gompit	cloglog	logit	probit	gompit	cloglog	logit	probit	gompit	cloglog
Bank 1												
0.5	45.9%	45.9%	45.1%	42.1%	94.5%	95.0%	95.5%	95.5%	75.0%	75.3%	75.3%	74.1%
sf	57.9%	57.9%	57.1%	56.4%	89.9%	89.9%	90.5%	89.9%	77.1%	77.1%	77.1%	76.5%
spf	78.2%	78.9%	78.2%	80.5%	65.3%	64.3%	66.8%	61.8%	70.5%	70.2%	71.4%	69.3%
Bank 2												
0.5	41.7%	40.0%	40.0%	28.3%	96.2%	96.8%	96.8%	93.5%	82.9%	82.9%	82.9%	77.6%
sf	73.3%	73.3%	66.7%	61.7%	90.3%	87.6%	91.4%	81.1%	86.1%	84.1%	85.3%	76.3%
spf	86.7%	88.3%	86.7%	76.7%	77.8%	76.8%	78.4%	70.3%	80.0%	79.6%	80.4%	71.8%
Bank 3												
0.5	60.5%	58.1%	55.8%	58.1%	97.2%	97.2%	97.6%	97.2%	87.9%	87.4%	87.1%	87.4%
sf	76.7%	76.7%	72.1%	77.9%	90.6%	89.8%	92.5%	87.4%	87.1%	86.5%	87.4%	85.0%
spf	83.7%	83.7%	82.6%	83.7%	83.9%	82.3%	85.8%	79.1%	83.8%	82.6%	85.0%	80.3%
Bank 4												
0.5	42.1%	37.6%	44.4%	27.1%	100%	100%	100%	100%	81.7%	80.2%	82.4%	76.9%
sf	56.4%	54.1%	56.4%	45.9%	98.6%	99.3%	98.6%	100%	85.2%	85.0%	85.2%	82.9%
spf	82.0%	84.2%	80.5%	85.0%	80.1%	78.0%	80.1%	75.6%	80.7%	80.0%	80.2%	78.6%
Bank 5												
0.5	15.3%	9.7%	13.9%	6.9%	99.4%	99.4%	99.4%	99.4%	72.8%	71.1%	72.4%	70.2%
sf	29.2%	25.0%	23.6%	22.2%	96.8%	96.8%	96.8%	96.8%	75.4%	74.1%	73.2%	73.2%
spf	61.1%	61.1%	61.1%	65.3%	68.6%	67.9%	69.2%	64.1%	66.2%	65.8%	66.7%	64.5%

Table 5b: Percentage correctly predicted from profile log-likelihood route

Bank/	Loan Outcome								Overall			
Cutoff	Denied (y=0)				Approved (y=1)							
Value	logit	probit	gompit	cloglog	logit	probit	gompit	cloglog	logit	probit	gompit	cloglog
Bank 1												
0.5	45.9%	46.6%	45.1%	41.4%	94.5%	95.0%	96.0%	95.5%	75.0%	75.6%	75.6%	73.8%
sf	57.9%	59.4%	57.1%	50.4%	89.9%	89.4%	91.5%	90.5%	77.1%	77.4%	77.7%	76.8%
spf	78.2%	78.9%	76.7%	80.5%	65.3%	63.8%	66.3%	60.8%	70.5%	69.9%	70.5%	68.7%
Bank 2												
0.5	41.7%	41.7%	30.0%	26.7%	96.2%	96.2%	92.4%	94.6%	82.9%	82.9%	77.1%	78.0%
sf	73.3%	73.3%	51.7%	66.7%	90.3%	89.2%	85.4%	82.7%	86.1%	85.3%	77.1%	78.8%
spf	86.7%	90.0%	78.3%	76.7%	77.8%	75.7%	70.8%	69.7%	80.0%	79.2%	72.7%	71.4%
Bank 3												
0.5	60.5%	58.1%	55.8%	58.1%	97.2%	97.2%	97.6%	97.2%	87.9%	87.4%	87.1%	87.4%
sf	76.7%	76.7%	72.1%	77.9%	90.6%	89.8%	92.5%	87.4%	87.1%	86.5%	87.4%	85.0%
spf	83.7%	83.7%	81.4%	83.7%	83.9%	82.7%	85.8%	92.5%	83.8%	82.9%	84.7%	90.3%
Bank 4												
0.5	42.1%	37.6%	37.6%	56.4%	100%	99.7%	99.7%	100%	81.7%	80.0%	80.0%	86.2%
sf	56.4%	52.6%	56.4%	46.6%	98.6%	98.3%	98.6%	99.7%	85.2%	83.8%	85.2%	82.9%
spf	82.0%	82.0%	79.7%	86.5%	80.1%	76.7%	80.5%	75.3%	80.7%	78.3%	80.2%	78.8%
Bank 5												
0.5	15.3%	9.7%	13.9%	6.9%	99.4%	99.4%	99.4%	99.4%	72.8%	71.1%	72.4%	70.2%
sf	29.2%	26.4%	23.6%	25.0%	96.8%	96.8%	97.4%	97.4%	75.4%	74.6%	74.1%	74.6%
spf	61.1%	62.5%	61.1%	63.9%	68.6%	67.3%	68.6%	63.5%	66.2%	65.8%	66.2%	63.6%

denials. The logit link often correctly predicts more denied loans than the other links, although there is usually little difference between this link's ability and that of the cloglog link when using the profile log-likelihood route. Interestingly, for our data sets, the probit link never dominates another link in predictive ability. This is an important practical finding, as researchers often arbitrarily choose between probit and logit for their analysis. Our findings illustrate that there can be gains, albeit minor, in one link over another.

When comparing overall classification ability, irrespective of loan disposition, our findings suggest that the computationally easier logit link is likely a good starting choice. That the gompit link usually dominates at predicting approvals, while the logit is often preferred for predicting denied loans correctly, suggests the need for further exploration of link choice in future research.

We now turn our attention to the hypothesis tests for racial discrimination. We examined one or three hypotheses. For Banks 2 and 3, as there are only two racial classes (K=2), the null is $H_0^1 : \beta_1 - \beta_2 = 0$ against $H_a^1 : \beta_1 - \beta_2 > 0$ for Bank 2, which reflects favourable loan disposition towards Whites over African Americans, while the alternative hypothesis is specified as $H_a^1 : \beta_1 - \beta_2 < 0$ for Bank 3, reflecting possible discrimination against Whites in favour of Hispanic Americans. For the other banks – Banks 1, 4 and 5 – with three racial classes (K=3), we test $H_0^2 : \beta_1 - \beta_3 = 0$ versus $H_a^2 : \beta_1 - \beta_3 > 0$ and $H_0^3 : \beta_2 - \beta_3 = 0$ versus $H_a^3 : \beta_2 - \beta_3 \neq 0$, in addition to H_0^1 against $H_a^1 : \beta_1 - \beta_2 < 0$. The alternative hypothesis H_a^2 reflects that discrimination, should it exist, is typically in favor of Whites. As there is usually no a priori expectation when testing whether an African American is favored over a Hispanic American, we examined a two-sided alternative hypothesis.

We report p-values for t-ratios for testing the nulls using the standard normal (SN) distribution, the limiting null distribution, and the bootstrap procedure, outlined in sub-section 3.4. Table 6 provides results. The legal standard for a statistically significant race effect is two or three standard deviations, which suggests a nominal 5% or 1%

significance level[14]. Such a choice effectively gives the benefit of doubt to the bank, as it implies a belief in nondiscrimination unless the sample evidence is extreme in suggesting otherwise. We adopt a 5% level here; in Table 6, we highlight rejection cases for this level by use of the bold font.

Examination of the p-values reveals that although general similarities exist in the pattern of outcomes, differences are also evident. Consistent findings arise for racial disparate treatment for Bank 2 and Bank 1 (when examining Whites versus Hispanic Americans), nondiscrimination of African Americans against Whites for Bank 5, Hispanic Americans against Whites for Bank 4, and African Americans against Hispanic Americans for Banks 1 and 5, whereas variations in outcome occur for the other cases. For instance, the SN p-values suggest that Bank 5 does not favor Whites over Hispanic Americans, irrespective of whether we work from the profile or pseudo log-likelihoods and regardless of link choice, while we conclude discrimination is possible for this bank using the bootstrapped p-values.

Banks 3 and 4, when testing H_0^1, provide examples that illustrate the importance of link: comparing the PR SN p-values, the logit and probit models suggest discrimination for Bank 3, whereas we conclude nondiscrimination for the gompit and cloglog models. This results despite the similar fit of the probit, gompit and cloglog models. The SN p-values from the profile log-likelihood for Bank 4, when testing H_0^1, show the impact of a small, seemingly insignificant, difference in fit on test outcomes, using the legal standard level of 5%. The average profile log-likelihood value for the gompit model was marginally better than for the other links, and we now observe that the gompit profile SN p-value is marginally higher than 5% for the gompit model, while the SN p-values for the other links are less than 5%.

Our results show that there is usually no qualitative difference in test outcomes between the SN p-values, for the probit, gompit and cloglog models, from the pseudo and profile routes[15]. This is useful for practitioners, as obtaining estimates via the pseudo

---

[14] See, e.g., Kaye and Aicken (1986). LaCour-Little (1999) provides a useful commentary on this matter.
[15] When comparing the SN p-values via these two methods, we do not automatically expect the profile SN p-values to be smaller than those from the one-step pseudo route, because, although the profile estimator has higher precision than the pseudo estimator, at least asymptotically, coefficient estimates also change, which may result in a smaller (in magnitude) test statistic.

Table 6: P-values for testing for racial disparate treatment

Bank: p-value	Regression Model			
	Logit	probit	gompit	cloglog
$H_0^1 : \beta_1 - \beta_2 = 0$ vs. $H_a^1 : \beta_1 - \beta_2 > 0$ *				
Bank 1: PS SN p-value	**0.000**	**0.040**	0.101	**0.032**
Bank 1: PS boot p-value	**0.000**	**0.000**	**0.000**	**0.000**
Bank 1: PR SN p-value	**0.000**	**0.000**	**0.008**	**0.000**
Bank 2: PS SN p-value	**0.000**	**0.004**	**0.001**	**0.022**
Bank 2: PS boot p-value	**0.000**	**0.000**	**0.000**	**0.000**
Bank 2: PR SN p-value	**0.000**	**0.000**	**0.000**	**0.000**
Bank 3: PS SN p-value	**0.050**	0.248	0.156	0.287
Bank 3: PS boot p-value	**0.000**	0.122	**0.010**	0.145
Bank 3: PR SN p-value	**0.050**	**0.044**	0.136	0.689
Bank 4: PS SN p-value	**0.036**	0.085	0.079	0.111
Bank 4: PS boot p-value	**0.010**	**0.010**	**0.000**	**0.000**
Bank 4: PR SN p-value	**0.036**	**0.031**	0.052	**0.006**
Bank 5: PS SN p-value	0.591	0.622	0.716	0.565
Bank 5: PS boot p-value	0.505	0.535	0.798	0.509
Bank 5: PR SN p-value	0.591	0.529	0.726	0.492
$H_0^2 : \beta_1 - \beta_3 = 0$ vs. $H_a^2 : \beta_1 - \beta_3 > 0$				
Bank 1: PS SN p-value	**0.000**	**0.012**	**0.012**	**0.017**
Bank 1: PS boot p-value	**0.000**	**0.000**	**0.000**	**0.000**
Bank 1: PR SN p-value	**0.000**	**0.000**	**0.000**	**0.000**
Bank 4: PS SN p-value	0.411	0.397	0.424	0.361
Bank 4: PS boot p-value	0.616	0.283	0.419	0.343
Bank 4: PR SN p-value	0.411	0.349	0.455	0.223
Bank 5: PS SN p-value	0.285	0.214	0.229	0.364
Bank 5: PS boot p-value	**0.000**	**0.000**	**0.030**	**0.010**
Bank 5: PR SN p-value	0.285	0.238	0.285	0.246
$H_0^3 : \beta_2 - \beta_3 = 0$ vs. $H_a^3 : \beta_2 - \beta_3 \neq 0$				
Bank 1: PS SN p-value	0.054	0.754	0.353	0.972
Bank 1: PS boot p-value	0.495	0.687	0.121	0.691
Bank 1: PR SN p-value	0.054	0.888	0.069	0.958
Bank 4: PS SN p-value	0.149	0.265	0.219	0.366
Bank 4: PS boot p-value	**0.000**	**0.020**	**0.000**	**0.030**
Bank 4: PR SN p-value	0.149	0.182	0.057	0.145
Bank 5: PS SN p-value	0.569	0.445	0.282	0.492
Bank 5: PS boot p-value	0.394	0.414	0.283	0.485
Bank 5: PR SN p-value	0.569	0.590	0.360	0.735

Notes: PS = pseudo log-likelihood; PR = profile log-likelihood; SN = standard normal; Boot = bootstrap

* The alternative hypothesis for Bank 3 is $H_a^1 : \beta_1 - \beta_2 < 0$

log-likelihood is substantially easier than from the profile log-likelihood. However, there are some cases (e.g., Banks 3 and 4 when testing H_0^1) for which the gain in efficiency does lead to a smaller p-value, which results, at the 5% legal standard, in a switch in test outcome to discrimination. We note, however, that these cases result when using links that showed marginally less favourable fit.

In addition, we observe that the bootstrapped and SN p-values are markedly different; the bootstrapped p-values are usually smaller than the SN p-values, which suggests a finite-sample null distribution for the t-ratio that is thinner tailed than the standard normal; i.e., platykurtic. Such a feature leads us to support the nondiscrimination null when using the SN p-values, for a given nominal level of significance, more than should be the case, which has serious fair lending policy implications. It is evident even when using the logit link, as has been standard in the fair lending empirical literature. For example, the findings for Bank 4 when testing H_0^3 change from supporting the nondiscrimination hypothesis to rejecting it as we move from using SN p-values to bootstrapped p-values. As tests based on bootstrapped p-values are generally believed to perform better than those based on approximate asymptotic distributions, these findings support adoption of bootstrapping to generate p-values in statistical analysis for racial discrimination.

Moreover, despite our use of consistent estimators of the parameter vector, finite-sample bias, known to be present, likely differs across the links and between the profile and pseudo methods. This is also possibly contributing to different test outcomes. Benefits of adopting bias-reduction techniques, such as bootstrapping and jackknifing, would be worth exploring in future research.

6. <u>Concluding remarks</u>

Concerns regarding racial disparate treatment in mortgage lending have not abated over the years, despite legislation and efforts by regulators. Much of the evidence arises from formal statistical analyses of lender underwriting behavior that attempt to model the probability of application approval, with borrower's race as one of the factors. The question "Is race a statistically significant determinant of the likelihood of approval, after

controlling for lender underwriting criteria?" is, usually, then addressed as a test of statistical significance of the race effect. The most common framework assumes a logit link function, which implies linearity in the log-odds ratios.

Given the social, economic, political and legal ramifications of disparate treatment, it is important to understand any shortcomings of, and lack of robustness of outcomes from, the statistical models. Debate has focused on omitted variables and possible simultaneity bias, but the issue of link function has received little, if any, attention. Our study begins the exploration of this question by comparing the logit disparate treatment test outcomes with those from probit, gompit and cloglog links.

We show evidence that discrimination findings can be sensitive to this choice, despite most of the links providing similar sample fit. We show bank-specific examples for which a conclusion of discrimination (nondiscrimination) with the logit link became a nondiscrimination (discrimination) finding with one of the other links, using the legal standard (nominal) significance level of 5%. Because of the possibility of such a feature arising, we recommend that regulators, bank officials and others interested in testing for racial disparate treatment routinely examine links aside from the logit. Our inability to identify a clear and consistent link preference also suggests that this issue be the focus of future research.

For most of the cases presented, qualitative disparate treatment test outcomes are not sensitive to whether we estimate via the one-step pseudo log-likelihood, which accords with a consistent, but asymptotically inefficient, coefficient estimator, or the profile log-likelihood, from which we obtain maximum likelihood estimates. This has computational advantages for applied researchers. However, we do observe cases for which the profile route would lead to a conclusion of disparate treatment, whereas non-discrimination arises from the pseudo log-likelihood method. As these arose when using links that were marginally sub-optimal from a fit perspective, such situations are possibly less likely in practice.

Our investigation highlights the importance of the method adopted to determine p-values. For some banks, the resampling and standard normal approximations lead to consistent findings of discrimination or nondiscrimination, whereas for other banks, the two procedures result in divergent conclusions. Predominantly, we observe that the

resampling distribution has thinner tails than the standard normal distribution. Such a feature has important policy implications, given prevailing usage of this approximation from asymptotic theory in fair lending studies - if the resampling p-values are more accurate than the standard normal p-values, regulators, bank officials, consumers and court officials need to be aware that the latter may be significantly overstated. For such a case, an outcome of nondiscrimination might indeed be discrimination.   Given this potential event, we advocate that resampling procedures become standard practice when determining p-values in fair lending studies.

References

Anderson, J.A., 1972, Separate sample logistic discrimination, *Biometrika*, 59, 19-35.

Booth, J.G., Butler, R.W. and P. Hall, 1994, Bootstrap methods for finite populations, *Journal of the American Statistical Association*, 89, 1282-1289.

Browne, L.E. and G.M.B. Tootell, 1995, Mortgage lending in Boston – a response to the critics, *New England Economic Review*, September/October, 53.

Chen, M.-H., Dey, D.K. and Q.-M. Shao, 1999, A new skewed link model for dichotomous quantal response data, *Journal of the American Statistical Association*, 94, 1172-1186.

Clarke, J.A. and M.J. Courchane, 2005, Implications of stratified sampling for fair lending binary logit models, *Journal of Real Estate Finance and Economics*, 30, 5-31.

Cosslett, S., 1981a, Maximum likelihood estimators for choice-based samples, *Econometrica*, 49, 1289-1316.

Cosslett, S., 1981b, Efficient estimation of discrete-choice models, in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications* (pp. 51-111), Cambridge, Mass.: MIT Press.

Courchane, M., Golan, A. and D. Nickerson, 2000a, Estimation and evaluation of loan discrimination: An informational approach, *Journal of Housing Research*, 11, 67-90.

Courchane, M., Nehbut, D. and D. Nickerson, 2000b, Lessons learned: Statistical techniques and fair lending, *Journal of Housing Research*, 11, 277-295.

Cramer, J.S., 1999, Predictive performance of the binary logit model in unbalanced samples, *Journal of the Royal Statistical Society*, C48, 85-94.

Czado, C. and T.J. Santner, 1992, The effect of link misspecification on binary regression inference, *Journal of Statistical Planning and Inference,* 33, 213-231.

Davidson, R. and J.G. MacKinnon, 2000, Bootstrap tests: How many bootstraps?, *Econometric Reviews*, 19, 55-68.

Day, T.E. and S.J. Liebowitz, 1998, Mortgage lending to minorities: Where's the bias?, *Economic Inquiry*, 36, 3-28.

Dietrich, J., 2005, Under-specified models and detection of discrimination: A case study of mortgage lending, *Journal of Real Estate Finance and Economics*, 31, forthcoming.

Gross, S.T., 1980, Median estimation in sample surveys, *American Statistical Association Proceedings of the Survey Research Methods Section*, 181-184.

Hall, P. and S.R. Wilson, 1991, Two guidelines for bootstrap hypothesis testing, *Biometrics*, 47, 757-762.

Harrison, G.W., 1998, Mortgage lending in Boston: A reconsideration of the evidence, *Economic Inquiry*, 36, 29-38.

Horne, D.K., 1994, Evaluating the role of race in mortgage lending, *FDIC Banking Review*, 7, 1-15.

Horne, D.K., 1997, Mortgage lending, race and model specification, *Journal of Financial Services Research,* 11, 43-68.

Hsieh, D.A., Manski, C.F. and D. McFadden, 1985, Estimation of response probabilities from augmented retrospective observations, *Journal of the American Statistical Association* 80, 651-62.

Kaye, D.H. and M. Aicken (eds.), 1986, *Statistical Methods in Discrimination Litigation*, New York: Marcel Dekker.

Kiefer, J. and J. Wolfowitz, 1956, Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters, *Annals of Mathematical Statistics*, 27, 887-906.

LaCour-Little, M., 1999, Discrimination in mortgage lending: A critical review of the literature, *Journal of Real Estate Literature*, 7, 15-49.

Lawless, J.F., Kalbfleisch, J.D. and C.J. Wild, 1999, Semiparametric methods for response-selective and missing data problems in regression, *Journal of the Royal Statistical Society*, B61, 413-438.

Manski, C.F. and S.R. Lerman, 1977, The estimation of choice probabilities from choice based samples, *Econometrica*, 45, 1977-1988.

Manski, C.F. and D. McFadden, 1981, Alternative estimators and sample designs for discrete choice analysis, in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications* (pp. 2-50), Cambridge, Mass.: MIT Press.

Munnell, A.H., Browne, L.E., McEneaney, J. and G.M.B. Tootell, 1992, Mortgage lending in Boston: Interpreting HMDA data, Federal Reserve Bank of Boston, Working Paper 92-7.

Munnell, A.H., Browne, L.E., Tootell, G.M.B. and J. McEneaney, 1996, Mortgage lending in Boston: Interpreting HMDA data, *American Economic Review*, 86, 25-53.

Neuhaus, J., Scott, A.J. and C.J. Wild, 2002, The analysis of retrospective family studies, *Biometrika*, 89, 23-37.

Prentice, R.L. and R. Pyke, 1979, Logistic disease incidence models and case-control studies, *Biometrika*, 66, 403-411.

Presnell, B. and J.B. Booth, 1994, Resampling methods for sample surveys, Technical Report #470, Department of Statistics, University of Florida.

Rabinowitz, D., 1997, A note on efficient estimation from case-control data, *Biometrika*, 84, 486-8.

Ross, S. and J. Yinger, 1999, Does discrimination exist? The Boston Fed study and its critics, in M.A. Turner, F. Skidmore (eds.), *Mortgage Lending Discrimination: A Review of Existing Evidence* (pp. 43-75), Washington, D.C.: The Urban Institute.

Ross, S. and J. Yinger, 2002, *The Color of Credit: Mortgage Discrimination, Research Methodology, and Fair-Lending Enforcement*, Cambridge, Mass.: MIT Press.

Scott, A.J. and C.J. Wild, 1986, Fitting logistic models under case-control or choice based sampling, *Journal of Royal Statistical Society,* B48, 170-182.

Scott, A.J. and C.J. Wild, 1991, Fitting logistic regression models in stratified case-control studies, *Biometrics*, 47, 497-510.

Scott, A.J. and C.J. Wild, 1997, Fitting regression models to case-control data by maximum likelihood, *Biometrika*, 84, 57-71.

Scott, A.J. and C.J. Wild, 2001, Maximum likelihood for generalised case-control studies, *Journal of Statistical Planning and Inference*, 96, 3-27.

Stengel, M. and D. Glennon, 1999, Evaluating statistical models of mortgage lending discrimination: A bank-specific analysis, *Real Estate Economics*, 27, 299-334.

Stukel, T.A., 1988, Generalized logistic models, *Journal of the American Statistical Association*, 83, 426-431.

Wacholder, S. and C.R. Weinberg, 1993, Prospective analysis of case-control data under general multiplicative-intercept risk models, *Biometrika*, 80, 461-465.

Zandi, M., 1993, Boston Fed's bias study was deeply flawed, *American Banker*, 19 August.