The Effects of Choice-based Sampling and Small-sample Bias on Past Fair Lending Exams

Jason Dietrich

Economic and Policy Analysis Working Paper

June 2001

The Effects of Choice-based Sampling and Small-sample Bias on Past Fair Lending Exams

Jason Dietrich

Office of the Comptroller of the Currency
Economic and Policy Analysis Working Paper

June 2001

Abstract: The Office of the Comptroller of the Currency uses choice-based sampling and limited sample sizes for statistically modeled fair lending exams. Both choice-based sampling and small sample sizes introduce bias into the maximum likelihood logit estimator, the standard estimator used by the OCC. This study applies results from Amemiya (1980) and Scott and Wild (1991) to estimate these biases for 16 recent exams.

The results show that of 29 tests of the null hypothesis of no racial effect conducted during the 16 exams, the outcome of two would change if small sample bias were taken into account, and the outcome of six would change if choice-based sampling bias were taken into account. Overall, the bias from choice-based sampling is generally larger. Although this study does not attempt to establish whether better sampling strategies would have changed examination conclusions based on any of the 29 hypothesis tests, the findings show that such strategies would have prescribed more thorough manual follow-up reviews for at least five of the 29 tests.

The views expressed in this paper are those of the author alone, and do not necessarily reflect those of the Office of the Comptroller of the Currency or the Department of the Treasury. The author would like to thank Amber Jessup, Irene Fang, Jeff Brown, and Frank Dwyer for their insightful comments and editorial assistance.

Please address correspondence to Jason Dietrich, Economist, Risk Analysis Division, Office of the Comptroller of the Currency, 250 E. Street, S.W., Washington, DC 20219 (phone: 202-874-5119; e-mail: jason.dietrich@occ.treas.gov)

I. Introduction

The sample sizes and sampling strategies that regulators choose when conducting fair lending examinations influence the amount of resources they expend on each bank, which banks receive the most intense follow-up file review, and perhaps even an examination's conclusions. Whether for a manual file review or a statistically modeled exam, sampling loan applications requires one of the largest resource expenditures during the exam. Two ways of lowering these costs are simply to draw smaller samples and to use stratified random sampling techniques. Although these approaches save money, they can make conclusions less reliable. Small samples lower the precision of estimates and introduce bias into the maximum likelihood (ML) logit estimator, the standard tool used to model the underwriting decision-making process during statistically modeled exams. Stratified random sampling increases the precision of estimates, but introduces bias into the ML logit estimator if stratification is based on an endogenous variable. When considering lower-cost methods of monitoring fair lending compliance, regulators should know how reliable each method's results are.

This study will help determine the extent of the bias caused by small samples and stratified sampling. I measure the effects of small samples and stratified random sampling on 16 statistically modeled fair lending exams conducted by the Office of the Comptroller of the Currency (OCC) over the past five years.[1] Specifically, I measure the small-sample and stratified sampling bias in the ML logit results and determine how the

[1] All 16 exams focused on whether race was a factor in the underwriting decision.

resultant hypothesis tests and conclusions might change after adjusting for these biases. These findings are a starting point for reassessing and improving statistically modeled fair lending exams.

The remainder of the paper is laid out as follows. Section II summarizes the OCC's sampling guidelines. Section III discusses small-sample properties of the logit estimator and summarizes the bias approximation formula developed by Amemiya (1980) for simple random sampling. Section IV discusses stratified sampling and summarizes the correction developed by Prentice and Pyke (1979) and Scott and Wild (1986, 1991, 1997) introduced by stratification on an endogenous variable. Section V applies the two bias measures to past fair lending exams. Section VI discusses alternative sampling and modeling methodologies for achieving unbiased and precise estimates with minimal sample sizes, and section VII concludes the discussion.

II. OCC's Sampling Guidelines

The OCC's fair lending program has followed three general sampling guidelines for statistically modeled fair lending exams.[2] First, examiners consult the economics' staff about conducting a statistically modeled exam if at least two racial groups of interest each have a minimum of 50 denials and 50 approvals. Although clearly stated in the OCC's fair lending guidelines, these minimum thresholds have been only loosely followed. For 16 statistically modeled exams the OCC has conducted since 1995, a total of 90 action/race strata were analyzed. Of these 90, nine (10.0 percent) had fewer than 50 observations in the population. Further, after eliminating unusable observations

[2] The Federal Reserve's matching program was used during some exams, but only during targeting to help identify the scope of the exam. See Avery *et al.* (1997) for a detailed description of this matching program.

identified after sampling, 58 (64.4 percent) had fewer than 50 observations in the estimation sample. Although many of these were only slightly below 50, there were still 28 (26.7 percent) with fewer than 40 observations.

Choice-based sampling is the OCC's second general sampling guideline in its fair lending program. Choice-based sampling means that sampling depends on the bank's choices regarding loan applications. In other words, the composition of the sample depends on the numbers of approved and denied loans in the population. The OCC's approach was slightly more complicated than pure choice-based sampling, since they sampled on race in addition to action. This entails drawing a random sample of applications from each action/race strata relevant to the exam. Choice-based sampling is well-suited for fair lending analyses, since it allows over-sampling of groups such as minority approvals and denials, which commonly have limited numbers of observations. In addition, stratification increases the precision of estimates and allows smaller samples to be drawn. The one drawback to this sampling strategy is that it introduces bias into multivariate estimators. With the ML logit estimator, this bias is restricted to the constant and race coefficients, and bias corrections for both the coefficient and standard error estimates are available.[3]

The OCC's third general sampling guideline concerns determination of the sample size and strata allocations. The OCC used the following equation to calculate the strata sample sizes,

[3] See Prentice and Pyke (1979) and Scott and Wild (1986, 1991 and 1997).

$$n_{ri}^{*} = \frac{N_{ri}\left[\sum_{r=1}^{R}\sum_{i=0}^{1}\frac{N_{ri}}{N^{2}}\sigma_{ri}^{2}\right]}{(.05*m/1.645)^{2}+\left[\sum_{r=1}^{R}\sum_{i=0}^{1}\frac{N_{ri}}{N^{2}}\sigma_{ri}^{2}\right]} \qquad (1)$$

where *N* is the total population size; N_{ri} is the number of observations in the population with race=*r* and action=*i; m* is the population mean of a user-specified sampling variable, typically loan amount; and σ_{ri}^{2} is the population variance of the user-specified sampling variable for observations where race=*r* and action=*i*.[4] Using this specification, the estimated sample mean of the user-specified sampling variable using stratified sampling will differ from the true population mean by no more than 5 percent with 90 percent confidence. Further, the resultant strata proportions will equal the population strata proportions.

The calculated strata sizes using equation (1) were often less than 50, the minimum threshold needed to conduct a statistically modeled exam. When this occurred, these strata were inflated to more than 50 (a few extra were included as a contingency in case some observations were unusable), or to the population strata size if the population strata size was less than 50. Therefore, the actual stratified samples used in the estimations typically did not possess the same strata proportions as the population and the mean relation mentioned previously did not hold exactly.

[4] See Cochran (1977).

III. Small-sample Bias

Amemiya showed that the ML logit estimator is biased away from zero in small samples. All else equal, the ML logit estimator tends to provide a better fit to models with larger absolute population parameters. Therefore, there is a tendency for the estimator to overfit the data, causing increased bias the further the true population parameter is from zero. Amemiya constructs an approximate measure of this bias to the order of n^{-1} for the ML logit estimator with simple random sampling. I briefly outline his results here. Define the probability that applicant i is denied credit as

$$P(y_i = 1) = \frac{1}{1 + e^{-x_i^{'}\beta_0}} = P_i \tag{2}$$

where x_i is a known K-dimensional vector of explanatory variables for individual i, β_0 is the corresponding vector of unknown population parameters, and i = 1, 2, …, N. Differentiating the log likelihood function for the ML logit estimator with respect to β yields the following set of normal equations,

$$\sum_{i=1}^{N} (y_i - F_i)x_i = g(\beta, y) = 0 \tag{3}$$

which are non-linear in β. Here, F_i equals P(y_i = 1) evaluated at the estimated β instead of the true population value β_0. Solving this non-linear system of equations for β yields the ML logit estimator, $\hat{\beta} = h(y)$. Expanding $h(y)$ around $P=(P_1, P_2, \ldots, P_N)'$ for y using a third order Taylor series, and taking expectations, Amemiya shows the approximate order n^{-1} small-sample bias is,

$$BS = \frac{1}{2}(X^{'}D_1X)^{-1}X^{'}D_2D_4 1 \qquad (4)$$

where $D_1 = diag[P_i (1 - P_i)]$, $D_2 = diag[2P_i - 1]$, D_4 = the Hadamard product of I and $D_1^{1/2}X(X^{'}D_1X)^{-1}X^{'}D_1^{1/2}$, and 1 is an Nx1 vector of ones.[5]

The magnitude of Amemiya's approximation for small-sample bias in the ML logit estimator is a function of the sample size, the true population parameter values and the distributional characteristics of the explanatory variables. Sample size has the expected negative effect on the magnitude of bias, since the ML logit estimator is a consistent estimator. The true population parameter values have positive absolute effects on small-sample bias. As stated earlier, the absolute magnitude of bias increases as the true population parameter moves further from zero, since the ML logit estimator tends to overfit models with larger population values. Lastly, increases in the means, variances and covariances of the explanatory variables generally increase the magnitude of bias, all else held constant. In general, the means and variances have a much stronger effect on the bias than the covariances.

The calculation of small-sample bias using equation (4) is complicated somewhat by the fact that the approximate bias is a function of the unknown population parameters. Therefore, instead of providing a point estimate of bias, I provide a range of bias estimates based on the standard deviation of the estimated coefficient, holding all other coefficients at their estimated values. Since the absolute magnitude of small-sample bias

[5] The Hadamard product is cell-wise multiplication of two conforming matrices. Multiplying by the identity matrix, I, therefore simply yields a diagonal matrix consisting of the diagonal elements from the other matrix in the Hadamard product.

is positively related to the absolute magnitude of the true population parameter, a positive estimated coefficient will likely be a combination of the true population parameter and a positive bias term, while a negative estimated coefficient will likely be a combination of the true population parameter and a negative bias term. To incorporate this into the small-sample bias approximation, the range of true population parameters used to calculate the approximate bias for positive coefficient estimates is minus three standard deviations to plus one standard deviation from the coefficient estimate. For negative coefficient estimates, the range for the true population parameters is minus one standard deviation to plus three standard deviations from the coefficient estimate. The main underlying point here is that the true population parameter is most likely between the estimated coefficient value and zero.

A second factor complicates calculation of the approximations of small-sample bias: The formula is based on simple random sampling, and the OCC used choice-based sampling for all statistically modeled fair lending exams. Little is known about the small-sample properties of the logit estimator with choice-based sampling, and I apply equation (4) as if simple random sampling methods were used.

IV. Choice-based Sampling Bias

Choice-based sampling is a second source of bias affecting past statistically modeled fair lending exams. The OCC has used an ML logit estimator for every statistically modeled exam that it has conducted. With the ML logit estimator, the bias introduced by choice-based sampling is restricted to the constant and the coefficients for explanatory variables that were sampled on. Given that the OCC sampled on action and race, the bias in past fair lending results was therefore restricted to the constant and race estimates. Prentice and Pyke, and Scott and Wild provide a measure of this bias, and I briefly outline their results here. Using the same setup as for small-sample bias, let the probability that $y_i = 1$ equal the logistic cumulative distribution function. Assume that choice-based sampling is used to sample on action and race, and that a full set of race indicator variables, α_g ($g = 1, G$), are included in the model. Prentice and Pyke show that the estimated coefficients on the race variables from a ML logit estimator contain the following bias,

$$E(\hat{\alpha}_g) = \alpha_g + \log\left\{\frac{n_{1g}P(Y=0 \mid g)}{n_{0g}P(Y=1 \mid g)}\right\} = \alpha_g + o_g \qquad (5)$$

where, n_{0g} and n_{1g} are the numbers of sample observations from race g with $y = 0$ and $y = 1$, respectively. Scott and Wild go on to show that log $\{(n_{1g} / N_{1g}) / (n_{0g} / N_{0g})\}$ goes in probability to o_g and therefore provides a measure of choice-based sampling bias. Here, N_{0g} and N_{1g} are the numbers of population observations from race g with $y = 0$ and $y = 1$, respectively. For the typical fair lending model where one race is omitted and used as the base of comparison, simple algebra can be used to construct the appropriate bias-corrected race estimates,

$$\hat{\alpha}_g = \hat{\alpha}_{uc,g} - \log[(n_{1g}/N_{1g})/(n_{0g}/N_{0g})] + \log[(n_{1o}/N_{1o})/(n_{0o}/N_{0o})] \qquad (6)$$

where $\hat{\alpha}_{uc,g}$ = the uncorrected logit estimate, and n_{0o} and n_{1o} are the numbers of sample observations from the omitted race with $y = 0$ and $y = 1$, respectively.

The direction and magnitude of choice-based sampling bias depend on the ratios of the strata sample sizes to the strata population sizes. For the omitted race, over-sampling observations with $y = 1$ and under-sampling observations with $y = 0$ has a negative effect on choice-based sampling bias. Alternatively, for the included race, over-sampling observations with $y = 1$ and under-sampling observations with $y = 0$ has a positive effect on the bias. The magnitudes of these effects depend on the ratios of over-sampling and under-sampling for each race. If the choice-based sampling is proportional, so that the sampling proportions equal the population proportions by strata, the bias will equal zero. In this case, the issue becomes one solely of efficiency.

In addition to biasing the coefficient estimates, choice-based sampling also biases the standard error estimates for the constant and race. Scott and Wild derive the following consistent estimator for the variance/covariance matrix of the bias-corrected estimates,

$$Cov(\hat{\phi}) = \vartheta(\hat{\phi})^{-1} - \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \quad (7)$$

where, $\vartheta(\hat{\phi})^{-1}$ is the inverse of the observed information matrix and D=diag(d_1, …, d_g) with $d_g = n_{1g}^{-1} - N_{1g}^{-1} + n_{0g}^{-1} - N_{0g}^{-1}$. The three lower-right elements of the right-hand-side matrix are zero, since the coefficient and standard error estimates for only the g strata sampled upon are affected by choice-based sampling. Again, formulations for the case where one race is omitted from the estimation and a constant is included can be obtained with simple algebraic manipulations.

V. Effects On Previous Exam Results

This section uses the approximations of bias presented in sections III and IV to estimate the small-sample and choice-based sampling bias for 16 of the statistically modeled fair lending exams the OCC has conducted since 1995. Each of these exams utilized the minimum requirement of 50 applications per action/race strata, as well as choice-based sampling with no bias correction. Table 1 presents the population sizes and percentages by action and race, while table 2 presents the corresponding sampling results. The percentages presented in table 2 represent the difference between the percentage sampled and the true population percentage. They therefore convey whether a particular action/race strata was over-sampled (positive) or under-sampled (negative), and by what amount. As the tables show, the population sizes and distributions varied considerably across exams, and the OCC selected a wide range of sample and strata sizes. Total population sizes ranged from 939 for Bank 7 to 32,375 for Bank 2, with approved loans for race 1 consistently making up the majority of applications. The merit of using choice-based sampling is apparent: many strata have small population sizes, occasionally less than 50. This explains why earlier I described the OCC's policy of a minimum of 50 applications per action/race strata as "loosely followed."

Sample sizes ranged from 212 for Bank 3 to 716 for Bank 5. Banks 1, 3, 4, 6, and 16, which have the smallest sample sizes, will likely be the most affected by small-sample bias. This relationship is not exact, however, since the approximation for small-sample bias is also a function of the true population parameters and the distributional characteristics of the explanatory variables. For choice-based sampling bias, banks most susceptible to positive (negative) bias are those where the over-sampling of denials is high (low) relative to the over-sampling of approvals for a given race included in the model and where the over-sampling of approvals is high (low) relative to the over-sampling of denials for the omitted race. Banks 7, 10, and 12 all show sampling patterns

Table 1: Population sizes for past statistically modeled fair lending exams

		Race 1				Race 2				Race 3				Race 4			
		Approved		Denied		Approved		Denied		Approved		Denied		Approved		Denied	
	Total	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Bank 1*	8535	5151	60.4	809	9.5	1471	17.2	589	6.9	368	4.3	147	1.7				
Bank 2*	32375	25136	77.6	3016	9.3	1810	5.6	603	1.9	1408	4.4	402	1.2				
Bank 3*	2674	1250	46.7	479	17.9	378	14.1	567	21.2								
Bank 4	2960	2353	79.5	282	9.5	246	8.3	79	2.7								
Bank 5*	3632	2321	63.9	162	4.5	820	22.6	329	9.1								
Bank 6	7013	5123	73.1	992	14.1	220	3.1	130	1.9	394	5.6	154	2.2				
Bank 7	939	627	66.8	109	11.6	155	16.5	48	5.1								
Bank 8	3550	2612	73.6	131	3.7	329	9.3	68	1.9	365	10.3	45	1.3				
Bank 9	19903	14587	73.3	2798	14.1	451	2.3	278	1.4	1333	6.7	456	2.3				
Bank 10	1436	515	35.9	477	33.2	115	8.0	202	14.1	21	1.5	44	3.1	15	1.0	47	3.3
Bank 11	1261	776	61.5	211	16.7	193	15.3	81	6.4								
Bank 12	9040	4321	47.8	3455	38.2	504	5.6	540	6.0	67	0.7	153	1.7				
Bank 13	1637	942	57.5	75	4.6	35	2.1	35	2.1	148	9.0	63	3.8	249	15.2	90	5.5
Bank 14	26138	18728	71.7	3014	11.5	1111	4.3	526	2.0	2152	8.2	607	2.3				
Bank 15	3970	2787	70.2	308	7.8	453	11.4	130	3.3	227	5.7	65	1.6				
Bank 16	1977	1257	63.6	135	6.8	340	17.2	88	4.5	133	6.7	24	1.2				

* Only population strata proportions were available for these banks, so the population strata sizes had to be estimated using estimates of the total population size computed from publicly available HMDA data.

Table 2: Sample sizes for past statistically modeled fair lending exams

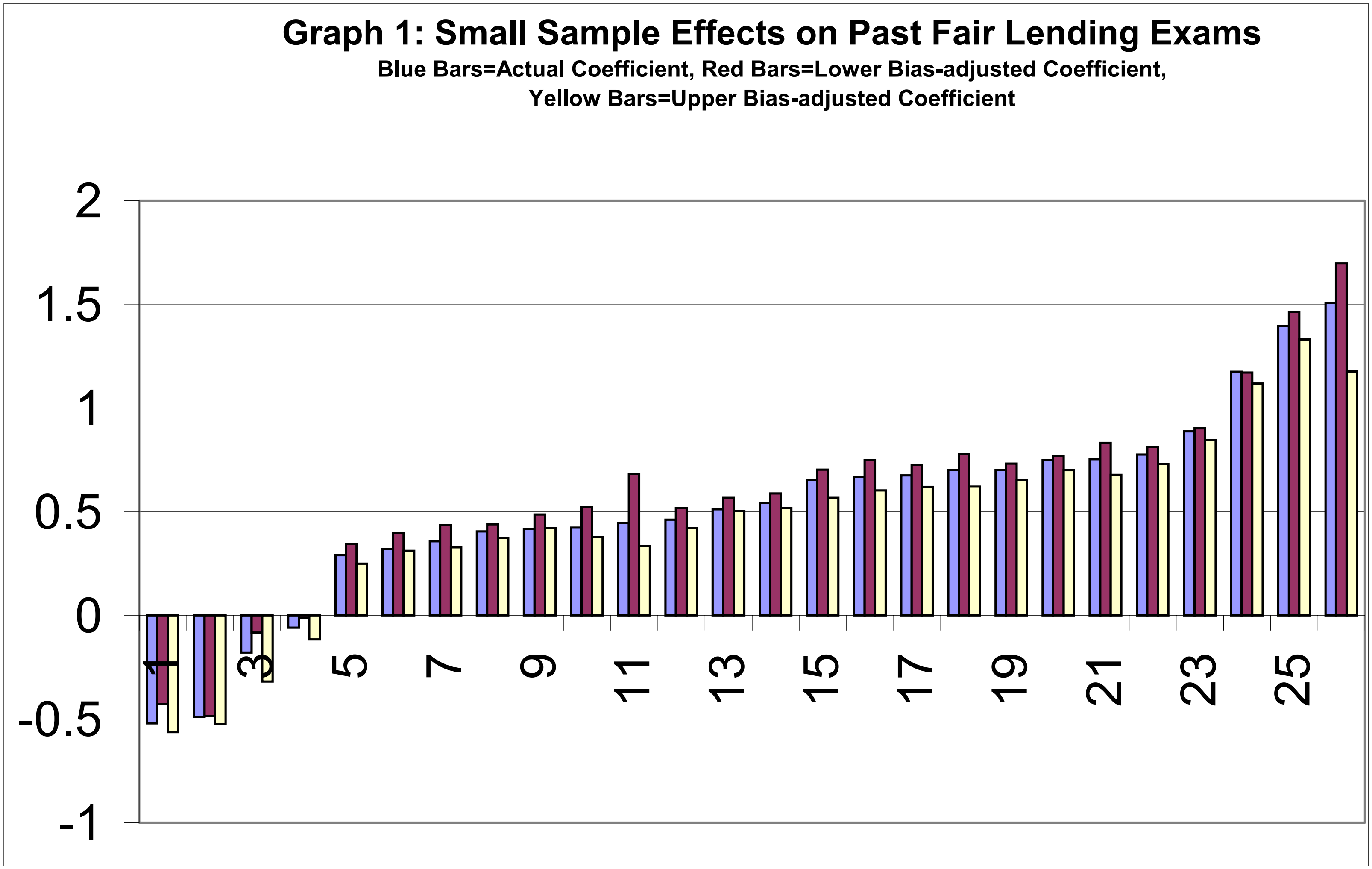
		Race 1				Race 1				Race 2				Race 3			
		Approved		Denied		Approved		Denied		Approved		Denied		Approved		Denied	
	Total	n	%Δ	n	%Δ	n	%Δ	N	%Δ	N	%Δ	N	%Δ	n	%Δ	N	%Δ
Bank 1	259	58	-38.0	36	4.4	43	-0.6	41	8.9	42	11.9	39	13.4				
Bank 2	362	125	-43.1	46	3.4	43	6.3	48	11.4	51	9.7	49	12.3				
Bank 3	212	89	-4.7	40	1.0	40	4.8	43	-0.9								
Bank 4	284	156	-24.6	43	5.6	48	8.6	37	10.3								
Bank 5	716	423	-4.8	44	1.6	149	-1.8	100	4.9								
Bank 6	232	90	-34.3	22	-4.6	33	11.1	27	9.7	41	12.1	19	6.0				
Bank 7	345	201	-8.5	45	1.4	57	0.0	42	7.1								
Bank 8	425	186	-29.8	48	7.6	51	2.7	47	9.2	50	1.5	43	8.8				
Bank 9	326	147	-28.2	42	-1.2	35	8.4	34	9.0	41	5.9	27	6.0				
Bank 10	496	204	5.2	76	-17.9	90	10.1	60	-2.0	14	1.3	22	1.3	9	0.8	21	0.9
Bank 11	392	197	-11.2	107	10.6	42	-4.6	46	5.3								
Bank 12	458	220	0.2	90	-18.5	25	-0.1	39	2.5	41	8.3	43	7.7				
Bank 13	461	154	-24.1	52	6.7	29	4.2	31	4.6	50	1.8	49	6.8	48	-4.8	48	4.9
Bank 14	307	132	-28.7	38	0.9	35	7.1	36	9.7	36	3.5	30	7.5				
Bank 15	472	213	-25.1	55	3.9	71	3.6	56	8.6	43	3.4	34	5.6				
Bank 16	233	90	-25.0	32	6.9	40	0.0	27	7.1	30	6.2	14	4.8				

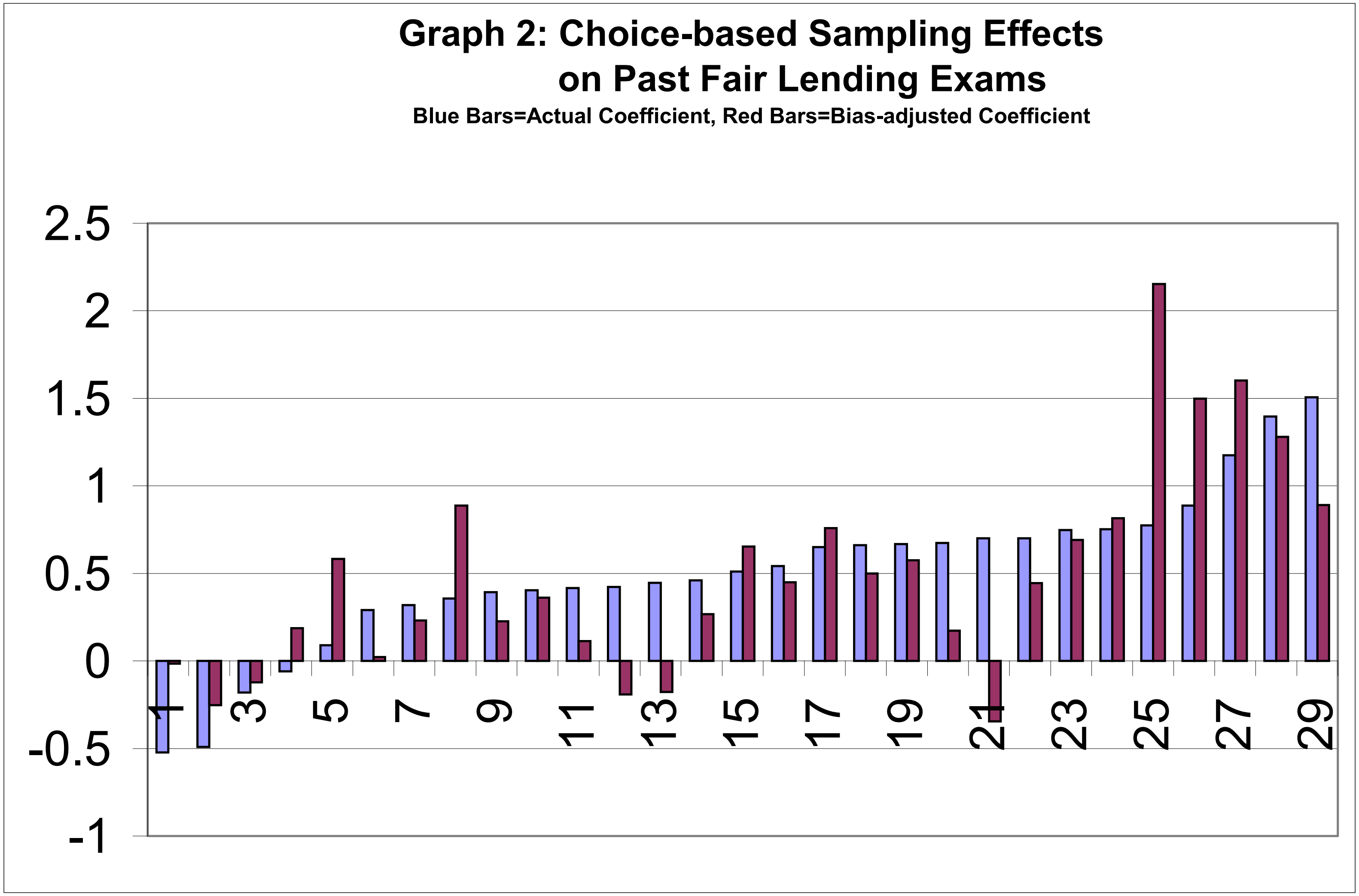
suggesting potential for positive choice-based sampling bias, while banks 1, 3, and 13 show sampling patterns suggesting potential for negative bias.

Graph 1 presents the approximations of small-sample bias and graph 2 of choice-based sampling bias. In graph 1, each racial estimate from the 16 past statistically modeled fair lending exams is represented by three bars.[6] The first bar shows the estimated coefficient from the actual exam. The second and third bars represent two estimates of coefficients adjusted for small-sample bias based on a range of assumptions about the true population parameter values. The main result of graph 1 is that small-sample bias appears to have had little effect on past statistically modeled fair lending exams. The results suggest that even smaller samples could have been examined without sacrificing the reliability of the results.

In addition to the magnitude of bias it is also important to examine whether tests of the null hypothesis of no racial effect change when the bias measures are incorporated. Computing revised t-statistics with bias-adjusted coefficients and the actual standard error estimates, the outcome of two of the 26 hypothesis tests would change. In each instance, a race coefficient significant at the 95 percent confidence level was no longer significant once the estimate of small-sample bias was taken into account. This represents a worst-case scenario, since the range of true population parameters used to calculate the bias approximations deploy "reasonable" to its extreme.

[6] Data needed to calculate small-sample bias were not available for three of the 29 racial estimates from the 16 statistically modeled exams analyzed in this study.


Graph 1: Small Sample Effects on Past Fair Lending Exams
Blue Bars=Actual Coefficient, Red Bars=Lower Bias-adjusted Coefficient,
Yellow Bars=Upper Bias-adjusted Coefficient
2
1.5
1
0.5
0
-0.5
-1
1
3
5
7
9
11
13
15
17
19
21
23
25


Graph 2: Choice-based Sampling Effects
on Past Fair Lending Exams
Blue Bars=Actual Coefficient, Red Bars=Bias-adjusted Coefficient
2.5
2
1.5
1
0.5
0
-0.5
-1
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29

Graph 2 presents the effects of choice-based sampling on the coefficient estimates from past statistically modeled exams. As with graph 1, the first bar in each set of two represents the actual coefficient estimate from the exam. The second bar shows the choice-based sampling bias-adjusted coefficient. The effects of choice-based sampling are generally much larger than those for small samples. They range from -1.4 to 1.0, and 22 out of 29 are greater than .10 in absolute value. Incorporating the choice-based sampling bias estimates shown in graph 2, in addition to bias-adjusted standard error estimates, the outcome changed for six of 29 tests of the null hypothesis of no racial effect. Five of these changed from not significant to significant, while one changed from significant to not significant. These results suggest that the bias of choice-based sampling significantly altered past fair lending exams, generally in the banks' favor.

When the OCC finds statistical evidence of disparate treatment, a thorough manual review of files is conducted to support or refute the statistical finding. When no statistical evidence is found, fewer follow-up resources are typically allocated. Although incorrect statistical evidence of discrimination at one bank is reason for some concern, that bank's files received a thorough and objective follow-up examination. The OCC's greatest concern may be that a thorough follow-up review of the files was not conducted for five banks whose statistics incorrectly showed no evidence of disparate treatment.

VI. Alternatives

Given the evidence of small-sample and choice-based sampling bias in past fair lending exams, what alternative sampling and estimation methodologies are available to ameliorate these problems? First, a quick and simple fix for the OCC's current fair lending program would be to increase sample sizes and to correct the bias of choice-based sampling. Although the OCC's limited resources may prevent the agency from increasing sample sizes, a good argument could be made for conducting fewer exams of higher quality.

Second, the OCC could explore alternative sampling approaches. Scheuren and Sangha (1998), Giles and Courchane (2000), and Dietrich (2001) have all examined how different sampling strategies affect the small-sample properties of the ML logit estimator. Using Monte Carlo simulation in the context of a fair lending analysis, each found that balanced stratified sampling by action and race performs better than alternative strategies. Giles and Courchane and Dietrich specifically recommend using this balanced sampling approach in addition to the choice-based sampling bias correction suggested by Scott and Wild.

Finally, the OCC could explore alternative-modeling approaches as well. Three general modeling approaches have been taken in the literature. First, some studies construct closed-formed expressions for bias up to a certain order of accuracy and then simply subtract this bias estimate from the standard maximum likelihood estimates (MLE). Amemiya develops a measure of bias to the order n^{-1} for the ML logit estimator, and Rilstone *et al.* (1996) develop a second-order formulation of bias for general non-linear models. Second, some studies use re-sampling techniques to estimate bias, which again is subtracted from the MLEs. Ferrari and Cribari-Neto (1998) and MacKinnon and Smith (1998) both take this approach. Third, alternative estimators to the ML logit can be used. Imbens (1992) suggests a method of moments estimator for discrete choice models with choice-based sampling. Golan (1998) uses a non-parametric approach, which chooses parameter estimates to maximize an entropy measure. The author argues that this method yields more precise estimates and works on smaller samples than traditional maximum likelihood methods. These sampling and modeling approaches are just some of the ways the OCC might improve its current fair lending program.

VII. Conclusion

The statistical-modeling segment of the OCC's fair lending program is subject to biases from two sampling-related sources, small sample sizes and choice-based sampling. The results of this study suggest that both sources have affected the results of the OCC's statistically modeled fair lending exams. Choice-based sampling had the greatest impact on the estimated coefficients, altering the original hypothesis tests in six of 29 cases. In five of these six altered tests, the actual exam results the OCC used in its conclusions favored the bank. Given the presence of these biases, regulators should use caution when interpreting statistical findings, should continue to use manual file reviews to support or refute statistical results, and should search for improved sampling and modeling approaches that increase the reliability of the results.

References

Amemiya, Takeshi. 1980. The n^{-2}-order Mean Squared Errors of the Maximum Likelihood and the Minimum Logit Chi-square Estimator. *The Annals of Statistics* 8(3):488-505.

Avery, Robert B., Patricia E. Beeson, and Paul S. Calem. 1997. Using HMDA Data as a Regulatory Screen for Fair Lending Compliance. *Journal of Financial Services Research* 11:9-42.

Cochran, W.G. 1977. *Sampling Techniques*. 3rd Ed. New York: Wiley.

Dietrich, Jason. 2001. How Low Can You Go?: An Optimal Sampling Strategy for Fair Lending Exams. Mimeo. Office of the Comptroller of the Currency.

Ferrari, Silvia L.P., and Francisco Cribari-Neto. 1998. On Bootstrap and Analytical Bias Corrections. *Economics Letters* 58:7-15.

Giles, Judith A., and Marsha J. Courchane. 2000. Stratified Sample Design for Fair Lending Binary Logit Models. Econometrics Working Paper EWP0007, Freddie Mac.

Golan, Amos. 1998. Estimation and Evaluation of Loan Discrimination -- An Informational Approach. Mimeo. American University Department of Economics.

Imbens, Guido W. 1992. An Efficient Method of Moments Estimator for Discrete Choice Models With Choice-based Sampling. *Econometrica* 60(5):1187-1214.

MacKinnon, James G., and Anthony A. Smith, Jr. 1998. Approximate Bias Correction in Econometrics. *Journal of Econometrics* 85:205-30.

Prentice, R.L., and R. Pyke. 1979. Logistic Disease Incidence Models and Case-control Studies. *Biometrika* 66(3):403-11.

Rilstone, Paul, V.K. Srivastava, and Aman Ullah. 1996. The Second-order Bias and Mean Squared Error of Nonlinear Estimators. *Journal of Econometrics* 75:369-95.

Scheuren, F.J., and B.S. Sangha. 1998. Interaction of Sample Design and Disparate Treatment Analysis in Bank Lending. Mimeo. Policy Economics and Quantitative Analysis, Ernst & Young LLP.

Scott, A.J., and C. J. Wild. 1986. Fitting Logistic Models under Case-control or Choice Based Sampling. *Journal of the Royal Statistical Society* Series B 48(2):170-82.

Scott, A.J., and C. J. Wild. 1991. Fitting Logistic 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.