Chapter 7

Under-specified Models and Detection of Discrimination: A Case Study of Mortgage Lending

Jason Dietrich
March 2026

Author's Note

This article is a reproduction of, "Under-specified Models and Detection of Discrimination: A Case Study of Mortgage Lending," *Journal of Real Estate Finance and Economics,* Springer Nature, **31**, 83–105 (2005). Note that some parts of this article, especially specific details of the OCC's exam processes, may be dated and no longer accurate.

I. Introduction

Omitted variable bias is a common problem afflicting empirical studies of racial discrimination. As a result, critics are quick to dismiss findings of discrimination as simply the effects of omitted variables that are correlated to race. Given the consequences of findings of discrimination, it is important to assess the existence and extent of omitted variable bias in racial estimates from statistical models. This is difficult to achieve directly, precisely because limited data and information are the cause of the bias. However, information about the extent of omitted variable bias can be obtained indirectly by identifying situations where omitted variables are known to be minimal, creating omitted variable bias by excluding relevant variables, and measuring the impact.

Fair lending exams conducted by the Office of the Comptroller of the Currency (OCC) present just such a situation. During fair lending exams, the OCC has access to lenders' policies, underwriters, and loan applications, which allows it to identify all of the economic factors a lender considers during the underwriting decision-making process. This allows the OCC to minimize problems of omitted variable bias when estimating statistical models to test for discrimination. Access to this information also creates an opportunity to assess how racial estimates from various fair lending studies may be affected by variables unavailable to these studies, but available to the OCC. Although, results from this type of analysis cannot be generalized to other fair lending or discrimination studies with certainty, compelling arguments can be made, especially as the number and variety of models analyzed, as well as the consistency of the patterns found increases.

This study uses data from 18 fair lending examinations recently conducted by the OCC to analyze the effects of omitted variables on racial estimates from statistical models of underwriting decisions.[1] Three specific questions are addressed. First, what does theory suggest about the expected direction of omitted variable bias in discrimination models? Specific attention is paid to differences in the effects of omitted variables between linear and non-linear estimators because the type of estimator affects hypotheses about the direction of omitted variable bias. Second, how would omitted variables have affected the statistical results from past fair lending exams? These results directly assess the magnitude and direction of omitted variable bias in fair lending models and address criticisms that omitted variable bias influenced previous fair lending studies showing evidence of discrimination. Third, how do omitted variables affect the identification of outliers for manual review? Reviewing a model's outliers is an important part of any modeling exercise because it identifies information potentially missed by the model and provides a second source of supporting evidence for the statistical analysis.

The remainder of the paper is organized as follows. Section II provides background information and a summary of the OCC's approach to statistically modeled fair lending exams. The third section examines the omitted variable bias problem in the logit estimator, summarizing the theoretical work on this topic presented in Lee (1980, 1982). The fourth section examines the impact omitted variables would have had on estimates of the effect of race, and the corresponding χ^2 test statistic from statistical models used during past exams. The fifth section focuses on the impact omitted variables would have had on the predicted probability of denial and the identification of outliers for manual review for these same exams. The paper concludes with a general discussion of the main findings.

[1] The data used in this study were collected as part of official OCC fair lending exams. They are therefore strictly confidential and cannot be released to the public or shared with other researchers.

II. Background

The Financial Institutions Reform, Recovery and Enforcement Act (FIRREA) passed by Congress in 1989 required banks to report data on the underwriting decision on applications for home mortgage products, as well as on the race of the applicants. These data allowed the public one of the first opportunities to calculate denial disparities by race at the individual level. Denial disparities, which almost always indicate that minorities are denied at a higher rate than whites, have been, and continue to be used as indicators of lending discrimination. Although these data are useful for developing an initial characterization of mortgage lending, using them as evidence of discrimination is dangerous as it does not consider differences in creditworthiness factors across racial groups.

Munnell et al. (1992, 1996) is commonly cited as the first study to control thoroughly for applicants' creditworthiness in an individual-level analysis of discrimination in the underwriting process.[2] Incorporating information on 38 factors that mortgage underwriters and lenders in Boston indicated were important in their decision-making process, the authors found that black and Hispanic applicants were still approximately 60 percent more likely to be denied credit than similarly-situated white applicants.

Munnell et al. was a clear improvement over previous studies of discrimination in mortgage lending. However, even with the extensive list of variables the authors used to capture applicants' creditworthiness, one set of responses still focused on biases introduced by omitted information.[3] There are three standard sources of bias related to omitted information -- omitted variable bias, endogeneity bias, and selection bias. Omitted variable bias occurs when an

[2] Additional studies that examined individual-level rejection probabilities, but that did not control for as many factors as Munnell et al. (1992, 1996) include Black et al. (1978), King (1980), Schafer and Ladd (1981), and Warner and Ingram (1987). LaCour-Little (1999) provides a comprehensive review of the fair lending literature.

[3] See Chapter 5 of Ross and Yinger (2002) for a comprehensive discussion of the effects of omitted variable bias on the results in Munnell et al. (1992, 1996).

exogenous variable that is related to both the dependent variable and at least one independent variable, is excluded from the model. Omitted variable bias is the focus of this paper.

Endogeneity bias occurs when a model includes an endogenous variable as an independent variable and the unobservable determinants of the endogenous variable are correlated with the unobservable determinants of the dependent variable. The unobservable determinants comprise the omitted information. Ross and Yinger (2002) analyze the effects of endogeneity bias in models of discrimination in mortgage lending and find evidence that racial estimates are biased toward zero. Endogeneity bias may affect the estimates of omitted variable bias presented in this study. However, it is difficult to disentangle bias due to endogeneity from bias due to omitted variables, so this paper makes the simplifying assumption that all of the independent variables are exogenous. In essence, this narrows the focus of the statistical analysis on treatment of applications once all information gathering and modifications are complete, instead of treatment of applicants from the time they apply.

Finally, selection bias occurs when the sample used in the estimation is not representative of the population. The reasons for, and likelihood, of applications being included in the sample comprise the omitted information. Selection bias is not an issue for this study, because the population for each exam is consumers who apply for a mortgage at the lender being examined, and the objective is to estimate the level of discrimination in the approve/deny decision for this population. Assuming population data are used, or appropriate sampling is conducted, selection bias should not be an issue. Selection bias is an issue if the population is defined as all persons, and the objective of the analysis is to estimate the level of discrimination in the approve/deny decision for this population, because only consumers who apply for a mortgage are included in the analysis.

Critics of Munnell et al. identified three general sources of omitted variable bias. The first source consists of specific factors in the Munnell et al. dataset that were relevant to the underwriting decision and correlated with race (Liebowitz, 1993; Zandi, 1993; Harrison, 1998; Day and Liebowitz, 1998). The three most commonly cited omitted variables are net worth, the bank's assessment of whether the applicant met its credit standards, and whether the bank was able to verify information provided by the applicant. The second source consists of idiosyncratic factors that determine the underwriting decision, regardless of the applicant's creditworthiness (Horne, 1994, 1997). One common idiosyncratic factor is that the collateral did not meet policy guidelines. These factors can have a large impact on the model estimates and are typically identified only by a thorough manual review of the application. The third source consists of bank-specific policy information (Stengel and Glennon, 1999). Stengel and Glennon provide evidence suggesting it may be inappropriate to aggregate across banks when analyzing underwriting decisions, because banks consider unique sets of factors, and apply unique weights to these factors when making their underwriting decisions. For all three sources of omitted variables suggested by critics of Munnell et al., when the additional variables were incorporated, the effect of race was mitigated or eliminated.[4]

The work by Munnell et al. and subsequent criticisms about omitted variables have shaped the OCC's current policies for using statistical models during fair lending examinations. The OCC employs a bank-specific approach that attempts to replicate the underwriting decision.[5] The intuition behind this approach is to control for all economic factors the bank uses in its underwriting decision-making process and then attribute any remaining systematic denial disparity to prohibited factor reasons. The first, and most important step in this process is to meet with the bank's underwriters and review the bank's underwriting policies. The goal is to identify all factors the bank considers in its decision-making process and understand how these

[4] Browne and Tootell (1995) respond to many of the criticisms of Munnell et al. (1992). Carr and Megbolugbe (1993) provide support for the robustness of the results in Munnell et al. (1992) as well.

[5] The objective of the OCC's modeling approach is to detect disparate treatment only, and not disparate impact. Ross and Yinger (2002), especially chapters 2 and 10 provide a more detailed discussion of disparate impact.

factors are used. The next step is data gathering and cleaning. Clearly, the model will only be as good as the data, so examiners conduct extensive data cleaning. The third step is to estimate an initial model and identify outliers for review.[6] The manual review of outliers is used to gather information potentially missed by the model, items such as fatal characteristics in the application, incorrect values in the data, and additional variables that should have been included in the original model. Other than issues regarding model mis-specification that can be corrected using variables already available electronically, information from the outlier review only affects applications included in the manual review. New information is typically not gathered on applications that were not reviewed manually. The final step in the process is to estimate a final model specification.[7] If no additional information is obtained during the outlier review, the preliminary and final model specifications will be identical. Conclusions about the role of race during the underwriting decision-making process are based on both the final statistical model results and the manual review of outliers.

The third step in the OCC's modeling process, -- identification and review of outliers -- warrants additional comment. The amount of information gathered during the manual review of outliers, and how that information is used, is related to the method for gathering the initial data and the extent of data cleaning. If data are gathered manually from the loan applications, less information is typically gained from the outlier review, because the examiners identify much of this information during data entry. In this case, information from the outlier review typically adds little to the model, and the outlier review is used more for supporting evidence

[6] The manual review of outliers is typically the first instance during an exam in which prohibited factors are considered as a possible explanation for an underwriting decision. Economists occasionally include prohibited factors in the preliminary model to help identify applications that have a large impact on the model estimates, but this is the exception rather than the rule.

[7] During some fair lending exams, information from an outlier review is incorporated back into the model, and additional outliers are identified and reviewed. This iterative process can continue until the model captures all additional information gained from reviewing outliers.

than to improve the model. Alternatively, if a bank has extensive electronic data, more information is typically gained from the outlier review, because these electronic datasets rarely contain information on idiosyncratic factors. In this case, information from the outlier review is incorporated back into the data to improve the preliminary model. As the number of banks maintaining electronic data has increased, information gathered during the outlier review has become increasingly important to the reliability of the models.

III. Omitted variable bias: Theory

The first question this study examines is what does theory suggest about the expected direction of omitted variable bias in fair lending models? For the OLS estimator, the direction of omitted variable bias depends on the product of the effects of the omitted variables on the outcome, and the correlations of the omitted variables and the included variables. Although we typically cannot explicitly calculate the bias because we rarely have data for the omitted variables, we can generally determine the direction of bias based on theorized signs of the two component effects.

For fair lending analyses, the omitted variables are typically measures of creditworthiness. With an assumption that minorities are generally less creditworthy, one could argue that the omitted variable bias in the racial estimate will typically be positive.[8] This is an important assumption, because if we cannot reject the null hypothesis of no racial effect using an under-specified model, then one could argue that we would not be able to reject the null hypothesis in a fully-specified model either, because the racial estimate will only get smaller.[9] This is one of the main arguments for estimating an initial under-specified model for fair

[8] Munnell et al. (1996) presents evidence suggesting minorities are typically less creditworthy than whites. The authors summarize their findings on page 31 by stating, “As reported in other surveys, black and Hispanic applicants have considerably less net wealth, liquid assets, and income than whites and they have weaker credit histories.”

[9] This argument ignores the effects of omitted variable bias on the standard error estimate.

lending exams using only data available in electronic form. If data need to be entered manually, the potential cost savings could be large, because additional data do not need to be gathered or cleaned.

On theoretical grounds, the argument for using under-specified models is flawed, because it is based on results from the OLS estimator, and not the non-linear estimators that are typically used for fair lending analyses. Relative to the standard omitted variable bias result for the OLS estimator, little work has examined how omitted variables affect non-linear estimators. Lee (1980, 1982) are two studies that have looked specifically at these issues. Lee shows that for a multinomial logistic probability model where the omitted variable (z) is dichotomous and the included variable (r) is discrete, the coefficient on r in the under-specified model will be biased upward, if the effect of z on the dependent variable (y), and the correlation of r and z conditional on y have the same sign; biased downward, if the effect of z on y and the correlation of r and z conditional on y have opposite signs; and unbiased, if z has no effect on y, or r and z are independent conditional on y.

Comparing Lee's findings with the standard OLS omitted variable bias results indicates the errors that can occur if the OLS results are applied to the logit estimator. Estimates of the effect of r in an under-specified model may be biased downward, even if the effect of z on y and the correlation between r and z unconditional of y are of the same sign; biased upward, even if the effect of z on y and the correlation between r and z unconditional of y are of opposite signs; and biased either upward or downward, if r and z are independent unconditional of y. For the latter case, the bias will be upward if the effect of r on y is negative and downward if the effect of r on y is positive.

The preceding theoretical results are specific to a model with two independent variables, a discrete variable that is included and a dichotomous variable that is omitted. Fair lending models typically have many discrete, dichotomous and continuous independent variables, so application of these results is somewhat limited.[10] [11] [12] Regardless, Lee's results do demonstrate that errors can occur when applying omitted variable bias results for the OLS estimator to omitted variable bias effects for a logit estimator, which is commonly used in discrimination studies.

IV. Omitted Variable Bias

The second question this study examines is how would omitted variables have affected the statistical results from past fair lending exams? It is difficult to address this question by directly applying the analytic results from Lee (1980, 1982) to the fair lending models the OCC estimates, because the number of independent variables in the model and the number of omitted variables is considerably larger than the models Lee analyzed. Further, it is also impossible to make reliable generalizations about the correlations between omitted variables and race needed to determine the direction of omitted variable bias. As an alternative, this study uses results from 18 statistically modeled fair lending exams the OCC recently conducted to demonstrate how omitted variables would have affected the results for those exams.

---

[10] In addition, Lee's results assume the independent variables are independently distributed. This is likely not the case for fair lending models, because various measures of creditworthiness from an applicant's credit bureau report are correlated.

[11] Lee does provide two extensions to his results that may be useful for typical fair lending models. First, the previous results are still valid if additional independent variables are included in the model, as long as the omitted variable z and these additional variables are independent conditional on r and y. Second, if the omitted variable z is continuous instead of dichotomous, the effect on the previous results depends on the distribution of z conditional on y and r. If z, conditional on y and r, is normally distributed, then omitting z from the multinomial logistic probability model will yield the same results as above.

[12] All of these findings for the logit estimator do not necessarily hold for the probit estimator. See Yatchew and Griliches (1984) for omitted variable bias results specific to the probit estimator.

Each exam included in this study focused on whether race played a role in a lender's underwriting decisions. The OCC estimated the following general model specification for each exam:

$$P(D = 1) = X\beta + R\delta \tag{1}$$

Matrix D indicates whether a mortgage application was denied. Matrix X consists of a set of control variables capturing the legitimate economic factors a lender considers during its underwriting process. Different lenders consider different factors, and the OCC estimates bank-specific models, so the composition of X differs across exams. In general, though, X contains measures of product characteristics, creditworthiness or credit history, employment/income stability, assets, and compensating factors. During data gathering and cleaning, efforts are made to account for counteroffers, so that X contains the information the lender considered when making the decision listed in D. Table 1 lists all of the independent variables used in this study. The matrix R contains race. Asian, black, Hispanic, and white were the races analyzed during these exams. The number of races analyzed during any one exam ranged from two to four. The error term for the underlying latent variable measuring the true creditworthiness of applicants was assumed to be drawn from a logistic cumulative distribution, which is the assumption for using a logit estimator.[13] [14]

---

[13] Although population data is occasionally used, the OCC typically uses sample data to estimate these models. Dietrich (2001) summarizes the sampling strategy the OCC uses, as well as the resultant sample compositions and sizes. In addition, Dietrich (2001) assesses the extent of choice-based sampling and small sample bias on the same set of exams used in this study.

[14] In reality, the mortgage process consists of determination of a number of jointly endogenous variables, such as the underwriting decision, loan-to-value ratio, monthly payments, term, and presence of co-signors. Therefore, one could legitimately argue that equation (1) is not appropriate. However, a single equation model has historically been the standard in discrimination studies. To the extent that this study attempts to address effects of omitted variable bias in other fair lending and discrimination studies, it is therefore appropriate to use as the benchmark of comparison here.

Table 1. List of independent variables.

Variables used in most exams	Variables used in selected exams
LTV	Co-applicant information
Debt ratios (front- and back-end)	Joint application
Score (bureau and custom)	Loan amount
Total trades	Term
Payment history of revolving accounts	Time at residence
Payment history of installment accounts	Lien status
Payment history of mortgage accounts	Satisfactory trades
Charge-offs	Inquiries
Collections	Alternative credit used
Public Records (bankruptcies, judgments, foreclosures, tax liens)	Explanation for derogatory accounts
Insufficient funds	Assets
Compensating factors	Reserves
Income	Source of down-payment (savings, gift funds, grants)
Owner occupancy	CRA loan
Time at job	Outcome from DU or LP
	PMI
	Cash-out amount
	Verification of Information
	Grade (Prime or Sub-prime)

Notes: This table lists the variables included as independent variables in at least one of the 18 exams included in this study. The actual form of each variable used for specific exams varied widely. For example, depending on the exam, LTV entered as a continuous, categorical, and 0/1 (met policy cutoff) variable, and was also interacted with other variables, such as compensating factors. In addition to differences in form, specific definitions of variables varied widely across exams. Creditworthiness variables were especially affected by this, because of the variety of timing and severity measures that were used across exams. For exams where a particular variable was not included in the final model specification, it may still have been incorporated into the overall statistical analysis.

The final model specification used for each exam is referred to in this study as the fully-specified model for that exam.[15] Therefore, there are 18 fully-specified models. Although the OCC goes to great lengths to build models that accurately reflect lenders' decision-making processes, it is difficult to argue that these models truly replicate the actual data generating process. However, to the extent that these models incorporate information beyond what is typically used in fair lending studies, they do provide a good opportunity to examine how racial estimates from these other studies are affected by omitting such information.

[15] Implicit in this assumption is that the model and error distribution specifications used during the fair lending exams are correct as well.

Starting from these fully-specified models, this study estimates various under-specified models and compares the estimates for each racial variable in these models to their corresponding values from the fully-specified models. Loan-to-value ratio (LTV), debt-to-income ratio (DTI), and credit score variables, as well as all variables the Home Mortgage Disclosure Act (HMDA) requires to be collected, are assumed to be available during every fair lending study. Therefore, these variables were never omitted as long as the lender considered them in its decision-making process. These variables will be referred to as the core variables. The remaining variables in the fully-specified models will be called the omittable variables. Using the omittable variables, under-specified models are created in three ways. First, the entire group of omittable variables is omitted at the same time. Second, only one omittable variable is omitted at a time from the fully-specified model. Third, information on idiosyncratic characteristics identified during file review is omitted.[16] [17]

Using the three under-specified model specifications, along with their corresponding fully-specified model specifications, three different sets of results measuring the effects of omitted variables are presented: 1) effects on model fit using Likelihood Ratio tests (LR), 2) magnitudes of estimated bias in the racial coefficients and corresponding χ^2 statistics used to test the null hypothesis that the population parameter for race equals zero, and 3) effects on tests of the null hypothesis that the omitted variable bias in the racial coefficients equals zero. All of these results are based on the final dataset and model specification used for the particular exam.

[16] Regressors with limited variation are another source of omitted variable bias that is a problem even though data are available for all relevant regressors. Typically, the small number of observations in which the decision is affected by such a factor are identified during the file review and eliminated from the sample.

[17] Often, categorical variables were used in place of continuous measures during a fair lending exam. For example, instead of using a continuous measure of reserves, the economist constructing the model may have created four indicator variables measuring quartiles of reserves to capture potential non-linearities in the reserves effect. For this study, the group of categorical variables measuring reserves would be treated as one variable, "Reserves," when omitting variables to create the under-specified models.

IV.1. *Model Fit*

The first item of interest is to determine the statistical importance of the omittable variables for recent fair lending exams. Specifically, do these variables matter? To answer this question, this study conducts LR tests to examine how omitting variables affects models' goodness of fit. Table 2 presents the results for 18 statistically modeled fair lending exams the OCC recently conducted. Each test uses a 95 percent confidence level and tests the null hypothesis that the model fit does not deteriorate when omitting variables.[18] The first column of the table lists the exam number. The second column indicates the number of omittable variables used in the final model specification for the particular exam. The third column shows the LR test statistic comparing a model with only the core variables and one with only a constant. These results indicate the statistical importance of the core variables, for comparison purposes. The fourth column shows the number of times the null hypothesis of no deterioration of goodness of fit can be rejected when excluding the omittable variables separately from the fully-specified model with replacement. For these results, multiple under-specified models were estimated for each exam, so aggregate results of the hypothesis tests are presented instead of the actual LR test statistics. For example, for Exam 12 there were eight omittable variables, and therefore eight under-specified models. For seven of these under-specified models, the LR test rejected the null hypothesis, suggesting that the model's fit worsened when each of these seven variables were omitted from the fully-specified model. The fifth column shows the LR test statistic when all omittable variables are excluded from the fully-specified model simultaneously. The final column shows the LR test statistic comparing the fully-specified

[18] Although model fit is used to help identify the optimal model specification, it is not the primary reason for including or excluding a particular variable. Because the OCC's modeling strategy is to replicate the underwriting decision-making process, the signs, and occasionally the magnitudes, of the coefficient estimates are just as, if not more, important. Therefore, the likelihood ratio value will not necessarily increase when omitting variables merely because of the modeling process used.

Table 2. Likelihood Ratio tests (LR) of under-specified models for past fair lending exams.

Exam	# of omittable variables	LR statistics from excluding core variables	# of rejections of LR tests from excluding one omittable variable at a time	LR statistics from excluding all omittable variables simultaneously	LR statistics from excluding information from outlier review
1	4	22.00*	3 of 4	116.48*	--
2	4	84.97*	3 of 4	88.86*	--
3	2	103.97*	1 of 2	7.86*	36.51*
4	8	45.60*	7 of 8	339.11*	--
5	4	81.11*	3 of 4	25.67*	--
6	7	127.67*	2 of 7	59.30*	96.93*
7	2	184.32*	2 of 2	57.21*	--
8	3	132.04*	3 of 3	43.86*	--
9	8	330.64*	5 of 8	41.51*	--
10	9	63.87*	4 of 9	178.08*	--
11	8	179.65*	4 of 8	154.11*	-45.53*
12	8	28.76*	7 of 8	137.75*	--
13	4	129.59*	3 of 4	44.42*	66.39*
14	2	51.32*	2 of 2	26.77*	--
15	7	887.89*	4 of 7	78.39*	--
16	14	152.10*	7 of 14	100.62*	78.65*
17	10	96.99*	3 of 10	40.72*	46.19*
18	8	393.33*	5 of 8	678.39*	689.70*

Notes: This table conveys the statistical significance of variables included in fair lending models. Four specific null hypotheses are tested using an LR statistic: H0: all core variables = 0 in a model with only core variables and a constant (column 3), H0: non-core variable = 0 in a fully-specified model (column 4), H0: all non-core variables = 0 in a fully-specified model (column 5), and H0: information from file review has no effect (column 6). Core variables include LTV, DTI, credit score, and all HMDA variables. Non-core variables include any additional variable included in the final model specification used during the exam.
*Indicates rejection of the null hypothesis at the 95 percent confidence level.

model incorporating information from the outliers to the fully-specified model excluding information from the outliers. As noted earlier, for some fair lending exams, information gathered during the outlier review was incorporated back into the dataset prior to estimating the final model specification. For these exams -- 3, 6, 11, 13, 16, 17, and 18 -- the necessary data is therefore available to examine the effects of this source of omitted information.

On the whole, excluding the group of core variables, the group of omittable variables, and information gathered during file review cause the model fit to deteriorate, suggesting these variables are statistically important. The null hypothesis of no deterioration can be rejected at the 95 percent confidence level for every exam and every type of under-specified model, as indicated by an asterisk in the table. One contrary result is found in Exam 11 where the fit improved by excluding file review information. This was caused primarily by a data entry problem identified during the review of outliers that increased the number of usable observations by 39. This is not typical of the type of adjustments made using information from the outlier review.

Although the core variables appear to have the greatest impact on model fit, the omittable variables have a larger impact for six of the 18 exams, and the file review information has the largest impact for one exam. This is somewhat surprising, because the core variables are often believed to be the main drivers of the underwriting decision. Also, the results for the omittable variables and the file review information are conditional on the core variables already being included in the model.

The LR test results from excluding one omittable variable at a time from the fully-specified model are not as strong as those from the other two types of under-specified models, or those from dropping the core variables. Of the 112 under-specified models created by excluding one omittable variable, the null hypothesis could be rejected in only 68 (60.7 percent) instances. This suggests that the addition of particular variables did not always improve the model's fit. This is not overly surprising, because the OCC places more emphasis on the bank's policies than on statistical importance when deciding whether to include variables in the model.

Overall, the evidence from Table 2 suggests that the variables included in the final model specifications for these 18 fair lending exams improved the fit of the model, and were therefore statistically important as well as theoretically important.

IV.2. *Magnitudes of Bias*

Having shown that the set of core and omittable variables, as well as information from file reviews are statistically important, the next step is to examine how racial estimates from past fair lending exams would have changed if information or data for the omittable variables were not available. As a starting point, it is useful to summarize the magnitudes of bias caused by each type of under-specified model.[19] Table 3a and Table 3b present summary statistics for the magnitudes of bias for the racial coefficient estimates and the corresponding χ^2 statistics used to test the null hypothesis that the population parameter for race equals zero.[20]

Table 3a shows how omitted variables would have affected the racial estimates for 18 statistically modeled fair lending exams the OCC recently conducted. Specifically, these results show how the estimated effects of race on the log odds of an application being denied a mortgage differ between under-specified and fully-specified models. Using Exam 12 as an example, the final model specification for that exam included three racial variables and eight variables this study defined as potentials for omission. Dropping each of these variables separately from the fully-specified model yields eight under-specified models and a total of 24 racial estimates. The first set of results in Table 3a shows that the bias for 13 of these estimates was negative, the average bias was 0.013 and the range was -0.121 to 0.220. Omitting all eight variables at the same time yields a 9th under-specified model and three more racial estimates. The second set of results in the table shows that the bias for only one of these estimates was negative, the average bias was 0.160 and the range was -0.144 to 0.364. Finally, excluding information gathered during file review yields a 10th under-specified model and three more

[19] Ross and Yinger (2002) raise concerns about the potential endogeneity of many of the control variables commonly included in fair-lending models, and how this may bias racial estimates toward zero. To the extent that some of the core and omittable variables used in this study are endogenous, some of the bias attributed to omitted variables, may be due to endogenous variables.

[20] The measures of omitted variable bias used in this study are for individual lenders and therefore are not directly comparable to the pooled analysis in Munnell et al. (1992, 1996).

Table 3a. Magnitudes of omitted variable bias in coefficient estimates.

			Only one omittable variable excluded at a time					*All omittable variables excluded simultaneously*					*Information from file review excluded*				
Exam	*# of races*[a]	*# of omittable variables*	*Mean*	*Min*	*Max*	*# neg*	*# pos*	*Mean*	*Min*	*Max*	*# neg*	*# pos*	*Mean*	*Min*	*Max*	*# neg*	*# pos*
1	2	4	0.087	−0.129	0.235	3	5	0.489	0.072	0.905	0	2	–	–	–	–	–
2	2	4	0.074	−0.074	0.360	3	5	0.390	0.230	0.549	0	2	–	–	–	–	–
3	1	2	0.062	−0.032	0.157	1	1	0.148	0.148	0.148	0	1	0.146	0.146	0.146	0	1
4	1	8	0.284	−0.105	1.02	2	6	0.320	0.320	0.320	0	1	–	–	–	–	–
5	2	4	−0.069	−0.205	0.102	5	3	−0.170	−0.257	−0.081	2	0	–	–	–	–	–
6	1	7	0.067	−0.014	0.292	2	5	0.450	0.450	0.450	0	1	0.325	0.325	0.325	0	1
7	2	2	0.005	−0.179	0.105	2	2	0.018	−0.141	0.177	1	1	–	–	–	–	–
8	2	3	−0.012	−0.138	0.078	3	3	−0.077	−0.145	−0.009	2	0	–	–	–	–	–
9	3	8	0.012	−0.313	0.188	9	15	0.201	0.045	0.467	0	3	–	–	–	–	–
10	1	9	0.015	−0.062	0.122	6	3	0.161	0.161	0.161	0	1	–	–	–	–	–
11	2	8	−0.002	−0.158	0.326	9	7	0.018	−0.147	0.183	1	1	−0.435	−0.919	0.049	1	1
12	3	8	0.013	−0.121	0.220	13	11	0.160	−0.144	0.364	1	2	–	–	–	–	–
13	2	4	0.025	−0.068	0.184	4	4	−0.052	−0.174	0.069	1	1	0.156	0.136	0.175	0	2
14	2	2	0.142	−0.112	0.453	2	2	0.295	0.247	0.343	0	2	–	–	–	–	–
15	1	7	0.030	−0.056	0.117	2	5	0.332	0.332	0.332	0	1	–	–	–	–	–
16	2	14	0.040	−0.042	0.345	11	17	0.537	0.269	0.804	0	2	0.023	−0.050	0.096	1	1
17	1	10	0.044	−0.096	0.182	3	7	0.398	0.398	0.398	0	1	0.587	0.587	0.587	0	1
18	1	8	0.009	−0.180	0.139	4	4	0.059	0.059	0.059	0	1	0.367	0.367	0.367	0	1
Total	31	112	0.043	−0.313	1.02	84	105	0.194	−0.257	0.905	8	23	0.091	−0.919	0.587	2	8

Notes: This table conveys summary statistics for the magnitude of omitted variable bias in racial coefficient estimates when using under-specified models. Under-specified models are constructed by excluding one non-core variable at a time from the final model specification used during the exam, by excluding all non-core variables simultaneously, and by excluding information gathered during file review. Core variables include LTV, DTI, credit score and all HMDA variables. Non-core variables include any additional variable included in the final model specification used during the exam.

[a] Column does not include the excluded race variable in the count.

Table 3b. Magnitudes of omitted variable bias in chi-squared test statistics.

			Only one omittable variable excluded at a time					*All omittable variables excluded simultaneously*					*Information from file review excluded*				
Exam	*# of races*[a]	*# of omittable variables*	*Mean*	*Min*	*Max*	*# neg*	*# pos*	*Mean*	*Min*	*Max*	*# neg*	*# pos*	*Mean*	*Min*	*Max*	*# neg*	*# pos*
1	2	4	2.12	−1.33	11.14	1	7	49.60	37.71	61.39	0	2	–	–	–	–	–
2	2	4	1.71	−0.95	6.29	3	5	15.10	13.51	16.77	0	2	–	–	–	–	–
3	1	2	0.89	−0.22	2.01	1	1	1.98	1.98	1.98	0	1	1.09	1.09	1.09	0	1
4	1	8	6.21	−0.44	21.57	1	7	193.8	193.8	193.8	0	1	–	–	–	–	–
5	2	4	−0.22	−1.61	2.03	5	3	−0.45	−1.27	0.38	1	1	–	–	–	–	–
6	1	7	−0.07	−0.24	0.04	5	2	0.53	0.53	0.53	0	1	−0.63	−0.63	−0.63	1	0
7	2	2	1.06	−0.05	2.57	1	3	2.19	−0.08	4.46	1	1	–	–	–	–	–
8	2	3	0.56	−1.02	2.14	2	4	1.41	−0.76	3.58	1	1	–	–	–	–	–
9	3	8	0.02	−1.05	0.93	15	9	0.58	−0.06	1.14	1	2	–	–	–	–	–
10	1	9	0.08	−0.00	0.32	4	5	0.84	0.84	0.84	0	1	–	–	–	–	–
11	2	8	0.22	−0.95	1.61	6	10	0.88	−0.42	2.17	1	1	3.26	−0.53	7.06	1	1
12	3	8	3.23	−2.37	13.82	9	15	37.05	22.32	44.62	0	3	–	–	–	–	–
13	2	4	0.38	−0.19	2.19	4	4	0.40	−0.32	1.12	1	1	0.51	0.32	0.69	0	2
14	2	2	0.66	−0.13	2.23	2	2	0.77	−0.45	1.99	1	1	–	–	–	–	–
15	1	7	−0.03	−0.12	0.14	5	2	0.39	0.39	0.39	0	1	–	–	–	–	–
16	2	14	−0.04	−0.71	0.52	16	12	2.87	2.31	3.42	0	2	−0.06	−0.09	−0.03	2	0
17	1	10	0.04	−0.00	0.13	1	9	1.04	1.04	1.04	0	1	2.41	2.41	2.41	0	1
18	1	8	0.38	−0.30	2.68	4	4	3.17	3.17	3.17	0	1	16.93	16.93	16.93	0	1
Total	31	112	0.91	−2.37	21.57	85	104	16.43	−1.27	193.8	7	24	2.72	−0.63	16.93	4	6

Notes: This table conveys summary statistics for the magnitude of omitted variable bias in Chi-squared test statistics for testing the null hypothesis that the population parameter for race equals zero. Under-specified models are constructed by excluding one non-core variable at a time from the final model specification used during the exam, by excluding all non-core variables simultaneously, and by excluding information gathered during file review. Core variables include LTV, DTI, credit score and all HMDA variables. Non-core variables include any additional variable included in the final model specification used during the exam.
[a]Column does not include the excluded race variable in the count.

racial estimates. Data were not available to estimate this under-specified model for exam 12, so all cells in the third set of results are missing.

There are two main results in Table 3a. First, although the racial estimates are more likely to be higher for the under-specified models (positive bias), there are a large number of instances where they are lower (negative bias). Looking at the results in which only one variable is omitted at a time, the bias is negative for 84 (44.4 percent) of the racial estimates, and four exams -- 5, 10, 11, and 12 -- had fewer racial estimates showing positive bias than negative bias. The occurrence of negative bias in the other two types of under-specified models is slightly lower as 8 (25.8 percent) racial estimates showed negative bias for under-specified models created by omitting all omittable variables simultaneously, and 2 (20.0 percent) showed negative bias for under-specified models created by omitting information from file review.

The first main result is very important, because some researchers argue that omitted variable bias in these types of models is typically positive. If this is true, then reliable conclusions of no discrimination could be made with an under-specified model that showed no discrimination. In addition, under-specified models that showed discrimination, could be strongly criticized as misleading due to omitted variable bias. The results in Table 3a cast doubt on these arguments.

The second main finding in Table 3a concerns the magnitudes of the bias. As expected, the bias is generally largest for under-specified models created by excluding all omittable variables. Overall, the bias ranges from -0.257 to 0.905 with an average of 0.194. Therefore, on average, the estimated effect of race on the log odds of being denied is 0.194 higher when all omittable variables are excluded. Information from file review has the next largest impact, which again, is not surprising considering that information gathered during file review is typically important to the analysis and has a large impact on the estimates. The bias due to omitting information from file review ranges from -0.919 to 0.587 with an overall average of

0.091. Excluding the anomalous result for Exam 11, the minimum bias is -0.050. Omission of one variable at a time shows the smallest effects with an overall average bias of 0.043.

In addition to affecting coefficient estimates, omitted variables also affect standard error estimates. As a result, it is not enough to look only at changes in coefficient estimates to determine how omitted variables would affect conclusions drawn for an exam; the effects on χ^2 test statistics need to be examined as well. Using data from the 18 statistically modeled exams, Table 3b presents the effects omitted variables would have had on the χ^2 test statistics used to test the null hypothesis that the population racial parameter equals zero. The format of Table 3b is the same as Table 3a.

Similar to the coefficient results from Table 3a, although the χ^2 test statistics are more likely to be higher for the under-specified models (positive bias), there are a large number of instances where they are lower (negative bias). Looking at the results in which only one variable is omitted at a time, 85 (45.0 percent) of the χ^2 test statistics are smaller for the under-specified model than for the fully-specified model. Five exams -- 5, 6, 9, 15, and 16 -- had fewer χ^2 test statistics that were larger for the under-specified model. The occurrences of negative bias are again lower for the other two types of under-specified models with 7 (22.6 percent) estimates showing lower test statistics for under-specified models created by omitting all omittable variables simultaneously, and 4 (40.0 percent) showing lower test statistics for under-specified models created by omitting information from file review.

The magnitudes of bias show similar patterns to the counts of positive and negative bias. Under-specified models created by omitting all omittable variables show the largest effects with an overall mean bias of 16.43. Although, this mean is pulled up somewhat by a mean of 193.8 for Exam 4. Under-specified models created by omitting information from file review showed the next largest impact with an overall mean of 2.72, and omitting one variable at a time had the smallest impact with an overall mean of 0.91. To put these results in perspective, 3.84 is the threshold for rejecting the null hypothesis with 95 percent confidence.

Overall, the results in Table 3b again cast doubt on the argument that a statistically insignificant race estimate in an under-specified model will necessarily be statistically insignificant in a fully-specified model. This uncertainty is too high to make this generalization for fair lending examination purposes.

IV.3. *Tests of Hypothesis that Omitted Variable Bias = 0*

The next step is to formally test the existence and direction of omitted variable bias in racial estimates. To do this, this study employs a series of t-tests of the null hypothesis that the difference between an estimated racial coefficient from an under-specified model and the estimated racial coefficient from the corresponding fully-specified model is equal to zero. Rejecting this null hypothesis in favor of the alternative hypothesis that the difference is positive suggests omitted variable bias is positive. Alternatively, rejecting this null hypothesis in favor of the alternative hypothesis that the difference is negative suggests omitted variable bias is negative.

One shortcoming of this approach is it assumes the sampling distributions of the two estimators are independent, which is clearly not the case. However, it is reasonable to assume that, if the racial estimate from the fully-specified model is large, the racial estimate from the under-specified model will be large as well. In other words, the two estimators are positively correlated. If this assumption is correct, the standard deviation of the difference between the two coefficient estimates, which is used to construct the t-statistic, will be over-stated if this correlation is not accounted for. Therefore, using the t-test, the percentage of instances when the null hypothesis of no difference is rejected in favor of the alternative hypotheses of either positive or negative bias will provide only a lower bound of the true rejection percentages.

Table 4 presents the t-test results for each exam using a 95 percent confidence level. The first column of the table lists the exam number. The second column indicates the number of racial groups, not including the omitted group, that were included in the final model specification for the exam. The third column indicates the number of omittable variables used in the final model specification for the particular exam. For the actual hypothesis test results, three columns are presented for each type of under-specified model. The first column in each set of three presents the number of one-sided tests where the null hypothesis of no bias was rejected in favor of an alternative hypothesis of positive bias. The second column in each set of three presents the number of one-sided tests where the null hypothesis of no bias was rejected in favor of an alternative hypothesis of negative bias. Finally, the third column in each set of three presents the number of two-sided tests where the null hypothesis could not be rejected.

Using Exam 12 as an example, the final model specification for that exam included three racial variables and eight variables this study defined as potentials for omission. Dropping each of these variables separately from the fully-specified model yields eight under-specified models and a total of 24 racial estimates. The first set of results show that the null hypothesis was rejected in favor of the alternative hypothesis of positive bias for 10 of these 24 estimates, was rejected in favor of the alternative hypothesis of negative bias in 10 estimates, and could not be rejected for four estimates. Dropping all omittable variables simultaneously from the fully-specified model yields one additional under-specified model with three racial estimates. Looking at the second set of results, the null hypothesis was rejected in favor of the alternative hypothesis of positive bias for two of these three estimates and in favor of the alternative hypothesis of negative bias for the remaining estimate. Finally, data were not available for

Table 4. T-tests of the null hypothesis that omitted variable bias equals zero using a 95 percent confidence level.

			Only one omittable variable excluded at a time[a]			*All omittable variables excluded at the same time*			*Information from file review excluded*		
Exam	*# of races*[b]	*# of omittable variables*	*H_a: Bias > 0*	*H_a: Bias < 0*	*H_0 could not be rejected*	*H_a: Bias > 0*	*H_a: Bias < 0*	*H_0 could not be rejected*	*H_a: Bias > 0*	*H_a: Bias < 0*	*H_0 could not be rejected*
1	2	4	4	1	3	2	0	0	–	–	–
2	2	4	4	1	3	2	0	0	–	–	–
3	1	2	1	0	1	1	0	0	1	0	0
4	1	8	5	1	2	1	0	0	–	–	–
5	2	4	1	5	2	0	2	0	–	–	–
6	1	7	2	0	5	1	0	0	1	0	0
7	2	2	2	1	1	1	1	0	–	–	–
8	2	3	2	3	1	0	1	1	–	–	–
9	3	8	8	5	11	2	0	1	–	–	–
10	1	9	3	3	3	1	0	0	–	–	–
11	2	8	4	6	6	1	1	0	1	1	0
12	3	8	10	10	4	2	1	0	–	–	–
13	2	4	3	3	2	1	1	0	2	0	0
14	2	2	2	2	0	2	0	0	–	–	–
15	1	7	4	1	2	1	0	0	–	–	–
16	2	14	10	1	17	2	0	0	1	1	0
17	1	10	4	1	5	1	0	0	1	0	0
18	1	8	3	2	3	1	0	0	1	0	0
Total	31	112	72 (38.1%)	46 (24.3%)	71 (37.6%)	22 (71.0%)	7 (22.6%)	2 (6.5%)	8 (80%)	2 (20%)	0 (0%)

Notes: This table conveys the number of times that the null hypothesis of no omitted variable bias is rejected in favor of the alternative hypotheses of either positive or negative bias. Under-specified models for these tests are constructed by excluding one non-core variable at a time from the final model specification used during the exam, by excluding all non-core variables simultaneously, and by excluding information from file review. Core variables include LTV, DTI, credit score and all HMDA variables. Non-core variables include any additional variable included in the final model specification used during the exam.

[a] One-sided t-tests were used for the "Ha: bias > 0" and "Ha: bias < 0" columns, while two-sided tests were used for the "H0 could not be rejected" columns.

[b] Column does not include the excluded race variable in the count.

Exam 12 to estimate an under-specified model by excluding information from file review, so these cells are missing.

The results in Table 4 suggest that the omitted variable bias is non-negative in general. However, the results are not consistent enough to support with certainty conclusions of whether discrimination exists when under-specified models are used. The null hypothesis of no bias was rejected at the 95 percent confidence level in favor of the alternative hypothesis of negative bias in 24.3 percent of the under-specified models created by dropping single omittable variables and 22.6 percent of the under-specified models created by dropping all omittable variables. Rejecting the null hypothesis in favor of the alternative hypothesis of negative bias was the most prevalent outcome for two exams (5 and 8). For nearly 38 percent of the tests using under-specified models created by dropping one omittable variable, the null hypothesis of no bias could not be rejected at the 95 percent confidence level. This compares to only 6.5 percent for the tests using under-specified models by dropping all omittable variables. As previously noted, however, both of these percentages are likely inflated given that the test used does not account for correlation between the estimators.

The results in Table 4 also suggest that information gathered during the outlier review is important with the null hypothesis of no bias being rejected in each of the 10 tests. In eight of the 10 cases, the bias is positive, so the estimated racial effects decline when this additional information is incorporated. This is considerably higher than the corresponding percentages for the other two types of under-specified models, highlighting the importance of information gained during the outlier review.

In addition to the t-test results presented in Table 4, the effects of each potential omitted variable were examined across exams to determine if generalizations could be made about the direction of bias in the race estimate if a particular variable was omitted. For example, number of mortgage delinquencies was used in six of the 18 models examined in this study. Among the under-specified models created by dropping this variable, 53 percent of the race estimates and 46 percent of the t-statistics were higher than the corresponding estimates from the fully-specified models. Clearly, if regulators knew that a mortgage delinquency variable was omitted from the model, they could not determine the expected direction of bias in the race estimates with much certainty. Overall, of the variables that were used during at least two exams, only one showed the same effect on the race estimates for all under-specified models -- a positive bias for compensating factors --, and only one showed the same effect on the t-statistic estimate for all under-specified models -- a positive bias for employment verification. These results suggest that generalizations about the direction of omitted variable bias cannot be made with much certainty about specific variables.

V. Outlier analysis

The third question this study addresses is how do omitted variables affect the identification of outliers for manual review? In addition to providing an estimate of pattern and practice of disparate treatment, the models the OCC estimates also identify outliers for manual review. A fully-specified model will accurately identify all of the questionable applications that require an in-depth manual review. The question examined here is whether an under-specified model will lead to the same rank ordering of risk and identify the same set of outliers. There are two potential errors that can occur when using an under-specified model to identify outliers, inappropriate outliers can be identified, and appropriate outliers can be missed. The first error leads to inefficient resource expenditures as applications are reviewed that do not need to be reviewed. This error will not affect the fair lending conclusions. The second error will affect the fair lending conclusions, however, because questionable applications that should be reviewed manually are not reviewed.

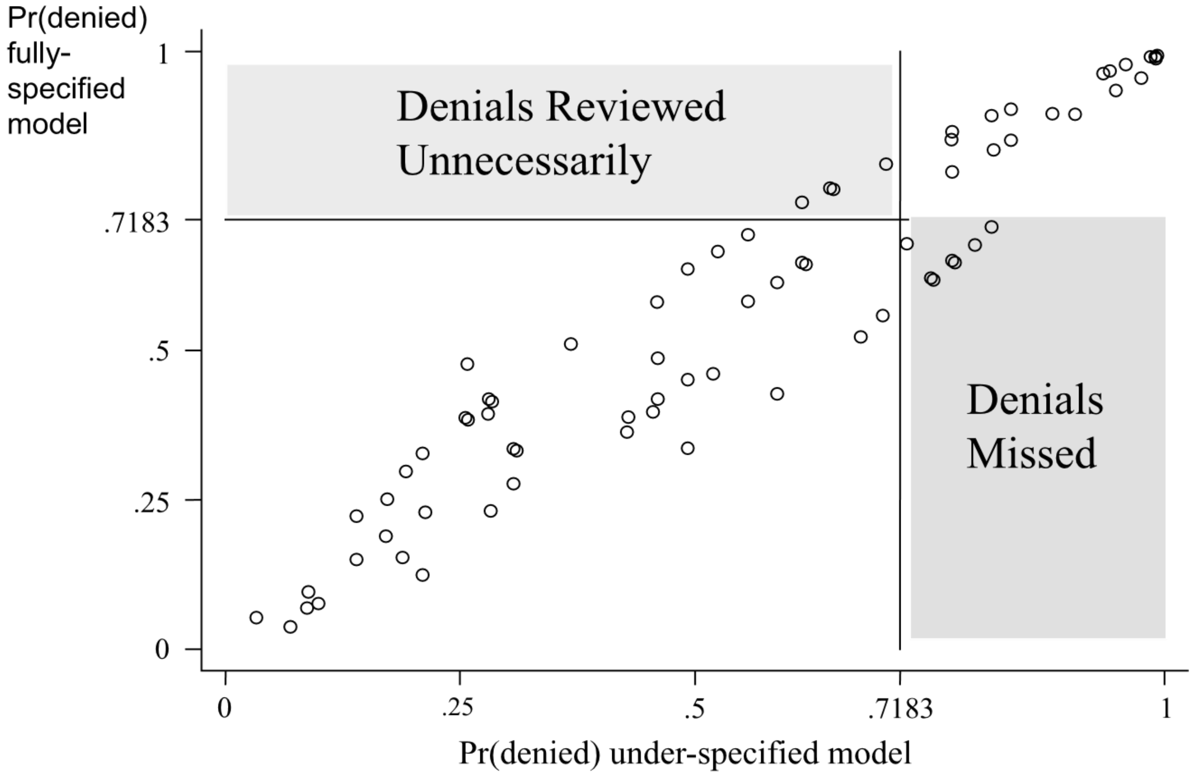
Similar to a credit score, statistical models provide a predicted probability of denial for each applicant, which rank orders the applicants' risk of default based on the factors included in the model. Applicants with a high predicted probability of denial that were approved, and applicants with a low predicted probability of denial that were denied are questionable applications that need to be reviewed manually.[21] To identify outliers, some cutoff probability first needs to be specified. A common approach is to use the actual approval rate for the population, because from the lender's perspective, the applicants near this cutoff represent the marginal applicants. During fair lending exams, denied applications are typically over-sampled, so using the population approval rate as the cutoff value will overstate the number of denied outliers and understate the number of approved outliers. As a result, this study uses the sample approval rate as the cutoff for each exam.[22]

Figure 1 depicts the two errors that can occur when using under-specified models to identify outliers for manual review. Using data from Exam 3, the graph presents a scatter plot of the predicted denial probabilities for denied applicants from the fully-specified model versus the predicted denial probabilities for denied applicants from the under-specified model created by omitting all omittable variables. The fully-specified results are plotted on the vertical axis. The sample approval rate for this exam, 71.83 percent, is used as the cutoff to identify outliers. The denied outliers for the fully-specified model will be those applications with a low predicted denial probability, i.e., those below the 71.83 percent line. Looking now at the predicted denial probabilities for the under-specified model, the denied outliers will be those to the left of the 71.83 percent line. Therefore, applications in the upper left quadrant would have been reviewed manually using the under-specified model even though they were not questionable. This measures the degree of inefficiency from using an under-specified model. Applications in the

[21] In addition, standardized differences in regression coefficients due to dropping individual observations can also be used to identify those outliers where race appears to have had a large impact in the decision.

[22] A cutoff of 0.5 is commonly used for fair lending exams as well.

Figure 1. Denied outliers from fully- and under-specified models.



lower right quadrant would not have been manually reviewed using the under-specified model even though they were questionable. This measures the degree of error in the fair lending conclusion from using an under-specified model. An analogous graph could also be created for approved outliers. For that graph, applications with predicted denial probabilities above and to the right of the cutoff lines for the fully- and under-specified models, respectively, would constitute outliers. The upper left quadrant would now contain the questionable applications that would have been missed had an under-specified model been used and the lower right quadrant would contain the applications that would have been reviewed manually even though they were not questionable.

Table 5 presents the outlier analysis results for all 18 exams in this study. Similar to the omitted variable bias results stated previously, all of these results are based on the final dataset and model specification used for the particular exam. Race is not included in any of the

Table 5. Outlier comparison between fully- and under-specified models.

Exam	Total outliers for fully-specified model	Total outliers for under-specified model	Applications that were outliers in the fully-specified model but not under-specified model[a]	Applications that were outliers in the under-specified model but not the fully-specified model[b]	Applications that were outliers in both models	Applications that were not outliers in either model
1	49	82	20 (40.8%)	53 (25.2%)	29	157
2	69	100	21 (30.4%)	52 (17.7%)	48	241
3	62	54	12 (19.4%)	4 (1.8%)	50	218
4	91	144	5 (5.5%)	58 (9.3%)	86	567
5	41	51	4 (9.8%)	14 (7.3%)	37	177
6	43	66	3 (7.0%)	26 (8.6%)	40	276
7	64	81	11 (17.2%)	28 (7.8%)	53	333
8	57	76	8 (14.0%)	27 (10.0%)	49	242
9	59	68	10 (16.9%)	19 (4.3%)	49	418
10	61	125	10 (16.4%)	74 (22.4%)	51	257
11	60	112	21 (35.0%)	73 (18.3%)	39	325
12	103	159	25 (24.3%)	81 (22.6%)	78	277
13	57	59	13 (22.8%)	15 (6.0%)	44	235
14	55	63	5 (9.1%)	13 (7.3%)	50	165
15	201	214	22 (10.9%)	35 (2.5%)	179	1377
16	55	87	11 (20.0%)	43 (13.4%)	44	278
17	26	42	5 (19.2%)	21 (12.4%)	21	149
18	128	217	48 (37.5%)	137 (12.6%)	80	947
Total	1281	1800	254 (19.8%)	773 (10.4%)	1027	6639

Notes: This table conveys the number of outliers that would have been missed, and the number that would have been reviewed unnecessarily had an under-specified model been used. Under-specified models are constructed by excluding all non-core variables simultaneously from the exam's final model specification.
a The denominator for the percentages is the number of outliers in the fully-specified model.
b The denominator for the percentages is the number of applications that were not outliers in the fully-specified model.

estimations, and sample approval rates for each exam are used as cutoffs in determining the outliers. The first column of the table lists the exam number. The second and third columns show the number of outliers identified using the fully-specified model and the under-specified model created by dropping all omittable variables, respectively. The fourth column contains the number of outliers in the fully-specified model that were not outliers in the under-specified models. The results are also shown as a percentage of the outliers from the fully-specified model to convey more effectively the magnitude of questionable applications that would have been missed had the under-specified model been used. The fifth column contains the number of outliers in the under-specified model that were not outliers in the fully-specified model. The results are also shown as a percentage of the applications that were not outliers from the fully-specified model to convey more effectively the magnitude of applications that would have been reviewed unnecessarily had the under-specified model been used. Finally, columns 6 and 7 show the number of applications that were outliers in both models and the number that were not outliers in either model, respectively.

The number of outliers is larger for the under-specified model for every exam, except Exam 3. This is expected because fewer relevant factors are included in the model to explain variation in the dependent variable. The percentage of outliers from the fully-specified model that would not have been reviewed manually had the under-specified model been used ranges from 5.5 percent for Exam 4 to 40.8 percent for Exam 1. The average across all exams is 19.8 percent. The number of outliers that would have been manually reviewed had the under-specified model been used, but not reviewed, if the fully-specified model had been used ranges from 1.8 percent for Exam 3 to 25.2 percent for Exam 1. Across all exams, 773 applications (10.4 percent) would have been unnecessarily reviewed, an average of nearly 43 outliers per exam. In summary, using an under-specified model instead of a fully-specified model would lead to missing an average of 19.8 percent of the questionable applications that should be reviewed and reviewing an average of 43 applications per exam that were not questionable and therefore that should not have been reviewed. The uncertainty caused by the first and the resource expenditures caused by the second provide further support for the use of fully-specified models during fair lending exams.

VI. Conclusion

A long-standing criticism of studies testing for discrimination is that omitted variables introduce an upward bias into estimates of the effect of race on the outcome of interest, thereby creating the illusion of statistical evidence of discrimination. With access to bank policies, underwriters, and all data that underwriters report that they use in their decision-making process, the OCC is in a unique position to diminish greatly problems of omitted variable bias. Access to this information also affords the opportunity to assess the impact of omitted variables on fair lending models that, in general, are afflicted by omitted variable problems.

Using data from 18 statistically modeled fair lending exams the OCC recently conducted, this study examined the effects of omitting variables available to regulators, but often unavailable to researchers. There were two main findings. First, omitted variables introduced bias into the racial estimates, but the direction of bias was not consistently positive as commonly thought. The null hypothesis of no bias was rejected at the 95 percent confidence level in favor of the alternative hypothesis of negative bias in 24.3 percent of the under-specified models created by dropping single omittable variables and 22.6 percent of the under-specified models created by dropping all omittable variables. Although omitted variables were considerably more likely to introduce no bias or positive bias, the results are not strong enough to be able to make these generalizations with a high degree of certainty. Second, omitted variables have an important effect on the identification of outliers to review manually. Estimating an under-specified model instead of a fully-specified model resulted in missing an average of 19.8 percent of the applications that should have been reviewed, and reviewing an average of 43 applications per exam that should not have been reviewed. These results provide some indication of the inefficient use of resources and the uncertainty introduced into conclusions from estimating an under-specified model.

Overall, the findings in this study suggest that variables often omitted by researchers do have an important impact on both the estimate of the effect of race and on the identification of outliers to review. They also show that it is dangerous to make generalizations about the potential direction of bias based on assumptions about the correlations between omitted variables and race. Although care must be taken when generalizing the results presented in this study to other fair lending or discrimination studies, the findings do provide compelling evidence that racial estimates may not always be biased upward due to omitted variables.

References

Black, H. A., R. Schweitzer, and L. Mandell. (1978). "Discrimination in Mortgage Lending," American Economic Review 68(2), 186-191.

Browne, L. E., and G. Tootell. (1995). "Mortgage Lending in Boston: A Response to the Critics," New England Economic Review pp. 53-78, Sept./Oct.

Carr, J. H., and I. F. Megbolugbe. (1993). "The Federal Reserve Bank of Boston Study on Mortgage Lending Revisited," Journal of Housing Research 4(2), 277-313.

Day, T. E., and S. J. Liebowitz. (1998). "Mortgage Lending to Minorities: Where's the Bias?" Economic Inquiry 36(1), 3-28.

Dietrich, J. (2001). "The Effects of Choice-based Sampling and Small-sample Bias on Past Fair Lending Exams," Economic and Policy Analysis Working Paper 2001-2, Office of the Comptroller of the Currency.

Harrison, G. W. (1998). "Mortgage Lending in Boston: A Reconsideration of the Evidence," Economic Inquiry 36(1), 29-38.

Horne, D. K. (1994). "Evaluating the Role of Race in Mortgage Lending," FDIC Banking Review 7(1), 1-15.

Horne, D. K. (1997). "Mortgage Lending, Race, and Model Specification," Journal of Financial Services Research 11(1/2), 43-68.

King, M. A. (1980). "An Econometric Model of Tenure Choice and Demand for Housing as a Joint Decision," Journal of Public Economics 14, 137-159.

LaCour-Little, M. (1999). "Discrimination in Mortgage Lending: A Critical Review of the Literature," Journal of Real Estate Literature 7, 15-49.

Lee, L. F. (1980). "Specification Error in Multinomial Logit Models: Analysis of the Omitted Variable Bias," Discussion Paper No. 80-131, Center for Economic Research, Department of Economics, University of Minnesota.

Lee, L. F. (1982). "Specification Error in Multinomial Logit Models: Analysis of the Omitted Variable Bias," Journal of Econometrics 20, 197-209.

Liebowitz, S. (1993). "A Study That Deserves No Credit," Wall Street Journal A14, Volume CCXXII, Number 44, September 1.

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

Munnell, A. H., L. E. Browne, J. McEneaney, and G. M. B. Tootell. (1996). "Mortgage Lending in Boston: Interpreting HMDA Data," American Economic Review 86(1), 25-53.

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

Schafer, R., and H. F. Ladd. (1981). Discrimination in Mortgage Lending, Cambridge, MA: MIT Press.

Stengel, M., and D. Glennon. (1999). "Evaluating Statistical Models of Mortgage Lending Discrimination: A Bank-Specific Analysis," Real Estate Economics 27(2), 299-334.

Warner, A., and J. Ingram. (1987). "A Test of Discrimination in a Mortgage Market," Journal of Bank Research 13, 116-124.

Yatchew, A., and Z. Griliches. (1984). "Specification Error in Probit Models," The Review of Economics and Statistics 66, 134-139.

Zandi, M. (1993). "Boston Fed's Bias Study Was Deeply Flawed," American Banker CLVIII, Number 159, page 13, August 19.