

Response to Critiques of “Mortgage Discrimination and FHA Loan Performance”

James A. Berkovec
Glenn B. Canner
Stuart A. Gabriel
Timothy H. Hannan

Abstract

This response discusses the critiques of Ross, Galster, and Yinger, with emphasis on the major issues of omitted variables and the potential for statistical discrimination. In our opinion, these concerns do not invalidate the use of loan performance data to investigate discrimination. Loan performance studies may not provide the complete and final answers to all research issues in lending discrimination, but they should be an integral part of the overall research program designed to understand discrimination and its effects.

We thank Stephen Ross, George Galster, and John Yinger (RGY) for their careful readings of our work and appreciate the opportunity to respond to their critiques. It appears to be the opinion of RGY that loan performance analyses cannot be used to identify discrimination in mortgage lending. We disagree.

Our disagreement with RGY is that their critiques appear to be focused less on the model and empirical analysis contained in our article (BCGH) than on the overall role of loan performance studies for evaluating discrimination. The BCGH results do not *prove* that discrimination does not exist; any claim to the contrary is obviously exaggerated. However, *proof* is a very high standard for any empirical analysis. In our opinion, studies of default and loan performance do add value to the debate about mortgage discrimination. Although performance studies may not provide the complete and final answer to all relevant research issues in lending discrimination, they should be an integral part of the overall research program designed to understand discrimination and its effects.

A brief summary of some aspects of the BCGH article that were not fully discussed by Ross, Galster, or Yinger will be useful. First, and most important, are the empirical findings in the BCGH work. Our primary empirical result is that Federal Housing Administration (FHA) mortgage default rates are higher for black borrowers than for borrowers from

other racial or ethnic groups. This result holds even after controlling for all borrower-, loan-, and property-related characteristics available in the data. Furthermore, this result holds for all loan origination years considered in the analysis and for a large variety of subsamples of the data. Thus the finding of higher default rates for black borrowers appears to be very robust.

In addition to default rates, BCGH also examine racial differences in the observed dollar amount of losses resulting from default. Our primary intent is to determine whether losses would give a different pattern of racial differences than default rates would give. In particular, does it seem likely that black borrowers default at a higher rate but have equivalent or lower overall losses when defaults occur? Examination of loss rates given default indicates that this is not the case. On average, black borrowers appear both to default more frequently and to have more severe losses when default occurs. Thus overall default losses, the product of the default probability, and the average loss given default are higher for black borrowers in comparison with other groups of FHA borrowers, after controlling for other available loan characteristics. This finding confirms the results of the default-rate analysis.

The analysis of losses from default also provides limited evidence that the higher default rate for black borrowers is not due simply to differential treatment of delinquent borrowers. The basic argument is that increased forbearance for one group of borrowers would be likely to result in a lower relative default rate, all else being equal, because some delinquent borrowers in the favored group would recover over time. However, the cases in which borrowers eventually default, despite forbearance, should result in higher average losses because their financial circumstances deteriorate. The fact that this pattern of relatively lower losses resulting from default does not occur for black borrowers suggests that differential forbearance policies do not explain the racial differences in default rates. Obviously, our data are inadequate to prove that discriminatory handling of distressed borrowers does not occur. However, results are not consistent with systematic discrimination against blacks in loan resolutions. Improved data that allow for more direct exploration of the process of moving from loan delinquency into default and foreclosure would certainly be desirable. However, such data are not available at this time.

The empirical analysis described in BCGH indicates that black borrowers have both higher observed default rates and higher loan losses on mortgages, after controlling for available indicators of loan quality. RGY do not dispute this conclusion. Furthermore, it seems unlikely that the performance differentials can be explained by unequal treatment of borrowers after loans are originated. Then the question becomes: What does the evidence on default rates reveal about discrimination in underwriting decisions?

According to the model developed by BCGH, discrimination in underwriting, by holding minorities to higher standards, results in improvements in observed minority FHA loan performance when compared with a nondiscriminatory baseline. Performance effects for FHA loans are shown to occur whether the discriminatory standards are used to evaluate applicants for FHA loans or for conventional loans. The prediction of improved observed loan performance resulting from discrimination follows directly from the assumption that discrimination acts to exclude differentially the weakest applicants of the affected group. Thus discrimination alters the performance of the marginal borrower. This fact does not necessarily imply that discrimination is revealed by comparing average performance among groups, a point that is well established (Peterson, 1981) and is repeated by BCGH and RGY.

The extension in the BCGH study shows that this performance differential from discrimination occurs even after conditioning on observable credit quality characteristics, due to a shift in the distribution of unobserved characteristics of accepted borrowers. In the case at hand, unobserved characteristics refer to those factors that are used in underwriting decisions—hence observed by lenders—but do not appear in the data available to outside analysts. This theoretical framework for underwriting decisions is based on the discrete choice models commonly used in accept/reject studies such as the Boston Federal Reserve’s analysis (Munnell, Brown, McEneaney, and Tootell, 1992). Using this framework, discrimination is shown to affect loan performance through a more severe truncation of the unobserved factors at the margin, regardless of the level of observable credit characteristics. The test for discrimination entails looking for the shift in unobservable factors that is induced by discrimination, through a comparison of loan performance in targeted and nontargeted groups after controlling for observable differences in credit quality.

When discussing his case 1, Galster notes that the test for discrimination depends on the existence of unobserved factors. This argument is true but basically irrelevant, because without unobserved characteristics discrimination is readily observable by simple comparison of the marginal borrowers, who are easily identified. Furthermore, it is hard to dispute the assertion that unobserved factors are important. In the absence of such factors, a perfect fit of accept/reject models of loan originations should be obtained. Existing models of lending decisions do not come close to explaining all of the variation in loan decisions. Until better explanatory power in these models is achieved, arguments that depend on perfectly observing all factors influencing loan decisions should be discounted. It is our belief that explicit consideration of unobservable factors is an important aspect of the model.

Ross notes that the BCGH model requires that unobserved factors (unmeasured characteristics) in underwriting decisions must influence loan performance. This is correct. To the extent that these unobserved factors are simply “random noise” uncorrelated with default, the selection effects associated with underwriting will have no bearing on loan performance. As Ross notes, direct econometric evidence of the effects of the unobserved factors on default does not exist. However, there is ample evidence that many of the observable factors used by underwriters, such as loan-to-value ratio, are strongly associated with loan performance. The BCGH presumption is that underwriting decisions are based on observable and unobservable factors that generally are indicative of loan performance. The relationship need not be perfect; the model predictions are maintained as long as underwriting decisions tend to select the less-risky loans for approval.¹

According to the BCGH model, discrimination always causes this shift in the distribution of the unobservable component of credit quality and, subject to the discussion above, a corresponding shift in expected loan performance. The remaining issue is the observability of the shift. In other words, when can researchers say that they have found it? The analysis in the BCGH study makes two additional assumptions. The first assumption is that the unobservable factors known to lenders are not correlated with race; the second is that the “true” loan performance among racial groups is equal, after controlling for all creditworthiness characteristics observed by lenders. In other words, the influential factors in default that are unobserved by lenders are also uncorrelated with race. Together, these assumptions imply that the performance of loans to all groups, after conditioning for observed loan characteristics, would be equal in the absence of discrimination. In this case, the improved performance of minority loans that is induced by discrimination is revealed

by better performance of the affected borrower group relative to identical loans made to other groups. Thus the conclusions in our study reject the joint hypothesis of lending discrimination along with equal “true” loan performance.

Most critiques by RGY concern the appropriateness of—and the effects of violations of—the two assumptions underlying equal observed performance of minority and nonminority borrowers in the absence of discrimination. In the context of the BCGH theoretical model, the relevant assumptions can be termed *no omitted variables* and *no statistical discrimination*. These assumptions will be discussed separately because of their differing implications for discrimination. However, the overall effect of violations of these assumptions on empirical estimates in BCGH will be quite similar, because they introduce a correlation between borrower race and the error term in a default equation, even if race is not used in underwriting decisions.

The first issue is that of omitted variables. In the current context, the implication is that the unobserved factors known by lenders and used in the underwriting decisions (the ε term in the BCGH model) are correlated with race. This implication results in a potential omitted variable problem in the BCGH default and loss models, because borrower race is correlated with the model error term as a result of the omitted factor. Omitted variables are a potential problem in any empirical study, and this one is no exception. The issue is particularly problematic in discrimination studies, since the important measurement is the correlation of race with the model residual.

Because minority groups tend, on average, to have riskier profiles of observable characteristics than nonminorities, it would not be surprising to find the same effects among the unobserved factors. Thus the probable direction of bias from omitted variables in the BCGH analysis is toward not finding discrimination. This is acknowledged in the article, and attempts are made to assess the degree of bias associated with the known important omitted factor, borrower credit history. However, other important omitted factors may also exist. Thus RGY are correct in pointing out that the BCGH results are subject to potential bias from omitted variables.

It is important to recognize, however, that the same potential problem with omitted variables occurs in other discrimination studies. In particular, accept/reject studies of mortgage origination are equally subject to omitted variables bias when important factors are omitted from the empirical model. However, in the case of accept/reject studies, the bias is in the opposite direction: toward a finding of discrimination where none exists. It is incorrect to argue that there is a fundamental problem with omitted variables in default studies that does not exist in accept/reject studies. If one believes in the existence of important omitted variables correlated with borrower race, no matter what data exist, one should never accept a test for discrimination that relies on correlations between race and a model residual. This argument appears to eliminate all statistical tests of discrimination based on observed data.

The second issue is the potential for “statistical discrimination,” whereby minorities may actually be worse credit risks on average, even though borrower creditworthiness indicators observed by lenders are equivalent across racial groups. In other words, race remains predictive of loan performance even after controlling for everything—observable and unobservable—that lenders know at the time of loan origination. Lenders, therefore, have incentives to engage in statistical discrimination by using race in underwriting precisely because it is predictive of default. Use of borrower race in underwriting to equalize marginal default rates among racial groups is illegal, but some would question whether it qualifies as discrimination.

The potential for statistical discrimination is one explanation for the results contained in the BCGH analysis. Although this potential is the major focus of the three critiques by RGY, the BCGH article makes no claims in regard to the results if this “true” racial difference in loan performance exists.² In the BCGH model, this racial difference in loan performance implies that the default rate $D(C)$ will depend on race. Thus at given levels of creditworthiness (C), minority groups will have an expected loan performance that is worse than that of nonminorities. In terms of the model, any requirement of higher creditworthiness standards (based on C) for minorities improves the relative performance of affected groups. However, since the baseline—with a common underwriting standard—shows a poorer loan performance for minorities, more discrimination (a bigger differential standard) is needed before a positive performance differential is observed for the affected group.

In this case, one cannot say whether race is used in underwriting unless the baseline performance differential can be quantified. Even if one assumes that lenders make full use of statistical discrimination and equate marginal default rates among racial groups, default rates conditioned on observable characteristics could differ between groups. Full statistical discrimination is consistent with higher, equal, or lower conditional default rates for minorities, depending on assumptions about the distributions of unobservable factors and “true” default differences. Thus, without imposing considerably more structure on the problem, it is difficult to say much about the existence or extent of statistical discrimination. All we can really say is that some degree of bias beyond the statistical level would be revealed as better performance for the affected group. Thus the BCGH results indicate that the extent of bias is not large enough to overwhelm any “true” differential in expected loan performance.

This basic point is made in various ways in all three critiques. There does not appear to be any substantive disagreement about the analytic effects of potential statistical discrimination in the BCGH model or results. However, RGY believe that the possibility that “true” default rates may depend on race invalidates the approach of using loan performance data to investigate discrimination. Again, we strongly disagree.

One reason for disagreement is our belief that loan performance studies *do* have the potential to identify the practice of statistical discrimination by lenders. To the extent that statistical discrimination is a result of correlations between race and performance factors that can be observed after loans are made, performance studies can detect the potential and actual use of statistical discrimination in underwriting. For example, suppose that “true” default rates for minorities were elevated because of a greater unmeasured risk of future unemployment—perhaps due to labor market discrimination. Then lenders might have incentives to statistically discriminate and impose higher standards on minorities to compensate for this increased risk. Analysts may measure the increased unemployment risk after the fact—perhaps by using actual employment experience—and include this factor in a loan performance study. If race is strongly related to loan performance through unemployment risk, inclusion of an actual unemployment risk variable should strongly influence the measured effect of race on loan performance. If, in the absence of information about unemployment risk, lenders were using race to statistically discriminate, then after controlling for unemployment risk the affected group should exhibit improved performance relative to other racial groups. Such an outcome is consistent with the prediction of the BCGH model. If, however, the potential existed but lenders were not statistically discriminating, inclusion of unemployment risk should remove an adverse effect of race on loan performance. This type of augmented loan performance study should be able to distinguish situations of actual and potential statistical discrimination from discrimination based on bias or from no discrimination at all.

The BCGH analysis has not been interpreted as an augmented loan performance study, even though some variables such as the 1990 census measures and State-level dummy variables are partially ex post observations. Clearly, there is potential for using improved data in a loan performance study to get better identification of the factors used in loan underwriting. An integrated effort to collect data on loan applications, approvals, and performance, along with borrower characteristics appropriate to monitor default decisions (as suggested in the critique by Ross), would clearly aid this endeavor. Having said that, we do not believe that the lack of such a model and data indicates that the “default approach is a fatally flawed method for studying discrimination.”

Furthermore, loan performance studies have value even if they are only able to determine whether statistical discrimination rather than lender bias is the major issue in lending discrimination. RGY seem to indicate that statistical discrimination is the cause of much of the observed disparity in mortgage lending rejection rates. However, this has not been the focus of much of the public discussion about the issue nor the basis for enforcement actions. Greater consensus regarding the importance of statistical discrimination would allow research and policy efforts to be better focused on identifying, measuring, and eliminating discrimination in mortgage lending. Whether or not statistical discrimination is the reason for disparities in mortgage lending may be an open question, but it is hard to see how this issue can be resolved without studies of loan performance.

Our interpretation of the BCGH results is limited: Our conclusion is that FHA loan performance data do not support a finding of widespread systematic discrimination in mortgage lending due to lender prejudice. The results of our research do not *prove* that there is no discrimination; no single study could do that. Some forms of lending discrimination would not affect loan performance. The BCGH results may be driven by omitted variables and/or a higher “true” default rate for black borrowers, potentially masking a limited amount of lender bias. The indication from BCGH, however, is that omitted variables and/or varying default risks—and not systematic racial bias—are the major factors in lending differentials.

Notes

1. A comparison of the financial and credit-related attributes of denied and accepted applicants in the Boston Federal Reserve study supports the view that underwriting decisions tend to exclude the least-creditworthy applicants.
2. Here *true* simply refers to expected default rates of the applicant population conditioned on lender observations of creditworthiness prior to screening by lenders. It is not meant to imply that race is the causal factor but rather that a variable influential in determining defaults is correlated with race but unobserved by the lender.

References

- Munnell, Alicia H., Lynn E. Browne, James McEneaney, and Geoffrey M.B. Tootell. October 1992. “Mortgage Lending in Boston: Interpreting HMDA Data.” Federal Reserve Bank of Boston Working Paper No. 92–7.
- Peterson, Richard L. Autumn 1981. “An Investigation of Sex Discrimination in Commercial Banks’ Direct Consumer Lending,” *The Bell Journal of Economics* 12.