# The Influence of Bureau Scores, Customized Scores, and Judgmental Review On the Bank Underwriting Decision-Making Process

M. Cary Collins
*University of Tennessee*

Keith D. Harvey
*Boise State University*

Peter J. Nigro
*The Office of the Comptroller of the Currency*

*In recent years, commercial banks have moved toward automated forms of underwriting and away from the judgmental review of loan applications. This paper employs unique bank loan-level data from a scoring lender to assess the impact of different underwriting approaches on applicant outcomes. To determine whether automated underwriting exhibits a potential "disparate impact" across income strata [high-income versus low- to moderate-income (LMI)], we compare outcomes created under two scoring approaches relative to a "judgmental" underwriting approach. We find that strict application of this custom scoring model leads to higher denial rates for LMI borrowers when compared with both a naive judgmental system and a bureau-scoring approach. We also identify the custom scorecard variables that produce the disparities in applicant denial rates. These results suggest that financial regulators should focus more resources on the evaluation and study of customized scoring models. Future research should exam-*

---

The opinions expressed are those of the authors and do not necessarily represent those of the Office of the Comptroller of the Currency. The authors thank William Lang, David Nebhut,and Gary Whalen for useful comments. Correspondence: Peter J. Nigro, Policy Analysis Division, The Office of the Comptroller of the Currency, 250 E Street, SW, Washington, DC 20219; Telephone:(202) 874-4799; E-Mail: peter.nigro@occ.treas.gov.

*ine additional ways to reduce or minimize these denial differentials. For example, in testing alternative scoring approaches, the institution or agency should include an assessment of the impact of employing non-traditional creditworthiness variables (e.g., rent or utility payments) on the approval/denial decision for LMI applicants relative to upper-income applicants.*

## Introduction

Statistically based credit decision-making systems were pioneered during the late 1950s but only saw mainstream use during the 1990s as the depth and breadth of electronic credit information increased.[1] These statistically based techniques are commonly referred to as "credit scoring" models. Initially, scoring models were employed in the consumer-credit portfolios of most major banks and credit card issuers to increase the speed of the credit decision, enhance the uniformity of the decision process, and reduce the overall costs of decision making. The relative homogeneity of this type of credit and the wide availability of performance data on applicants made the initial implementation of scoring models by these lenders successful and profitable.

In recent years, scoring models have migrated to other areas of the lending portfolio, including mortgage and small business loans. In fact, the use of credit scoring models has become widespread in the mortgage-lending industry over the past 10 to 15 years. In addition to its use in the underwriting process, credit scoring is also employed by secondary market purchasers of mortgage loans as a means of pricing credit or default risk, including development and use of such models by the government-sponsored enterprises (GSEs) and by providers of private mortgage insurance.

The growing importance of credit scoring in the allocation of mortgage credit led to the current debate about scoring's impact on the flow of credit to certain segments of the applicant population. Proponents of credit scoring and of the "automated underwriting" process that benefits from its use, argue that it lowers the overall cost of making credit available to consumers, while simultaneously increasing the speed and objectivity of the underwriting decisions.[2] For example, Calem and Wachter (1999) argue that scoring benefits lenders and borrowers alike by increasing the efficiency of the credit review process and reducing the likelihood of delinquency. Detractors of credit scoring models argue, however, that the underwriting variables employed and the weights assigned to each variable are based on the payment performance of traditional consumers.[3] As such, scores generated by these models may not accurately portray the creditworthiness of underrepresented groups in the applicant pool, such as low-income and minority applicants — groups that constitute a larger fraction of first-time home buyers. In particular, scoring models typically omit certain nontraditional indicators of credit performance, such as rent and utility payment histories, which are important components of credit performance for many low-income applicants.[4] A primary conjecture of this paper is that custom credit scoring systems that employ application-level information in addition to credit report data yield a disparity in low-income denials relative to upper income denials, since these scoring systems neglect compensating factors, or additional creditworthiness-related attributes, that are more common for low-income applicants.[5] To test this hypothesis, we employ unique data on unsecured, home improvement loans from a large lender using an overlay system of both custom and credit bureau scores in the underwriting process.[6]

## Credit Scoring and Potential Disparate Impact

What are credit scores? Credit scores are statistically derived measures of creditworthiness that rank credit applicants according to their degree of credit or default risk.[7] A score is typically associated with an odds ratio, addressing the question: How many applicants are likely to exhibit payment streams that become delinquent (or default) at the corresponding score? Although the models do not predict the absolute level of risk or which borrowers within a score range are likely to perform poorly, the literature has shown them to be effective tools for rank-ordering the risk of applicants (See Avery, Bostic, Calem, and Canner, 1996; Freddie Mac, 1996; and Pierzchalski, 1996).

Previous literature assessing the influence of credit scoring in the underwriting process is sparse and focuses solely on the role of bureau scores in that process. For example, Avery, Bostic, Calem, and Canner (2000) examine several statistical issues related to credit scoring using aggregate data. They find significant variation in bureau scores across a number of economic, geographic, and demographic groups, suggesting that the omitted variables and under-representation issues warrant further attention.

This paper extends the Avery *et al* research by examining loan-level data and underwriting decisions of a bank employing a custom scoring model. To our knowledge, this is the first paper to evaluate the underwriting decisions derived from customized, credit bureau and judgmental approaches and to demonstrate how these outcomes vary by income group.[8] We find that the custom-scorecard decisions lead to even larger disparities in high income versus low- to moderate-income denial rates than those disparities created using either the credit bureau score or the "judgmental" model approach. These results suggest that the issues of both omitted variable bias and the under-representation of certain subpopulations (*e.g.*, LMI) in model development may be even greater for some customized models.

## Data Description and Empirical Methods

This paper analyzes 1996 data on 2,266 unsecured, home improvement loan applications drawn from a large regional lender's activities in a single Metropolitan Statistical Area (MSA). As such, the pool of credits is relatively homogeneous and the underwriting standards relatively stable across time. The application-level data also include information on the income from the application, the bureau and custom credit scores, and the score attributes, or individual score loadings, for all applicants. LMI individuals are defined as those with incomes below the U.S. Department of Housing and Urban Development's 1996 MSA median income for this geography, while upper-income applicants are defined as those with incomes at or above the MSA median income.[9]

### *Descriptive Statistics*

Table 1 contains the breakdown of the 2,266 applications by income group. There are 1,698 applications from LMI applicants, accounting for 74.9 percent of the sample and 568 applications from upper-income applicants, accounting for 25.1 percent of the sample. As such, the sample has a reasonable balance and sufficient representation for both groups to perform hypothesis testing. Table 2 describes the data overall and then bifurcated by income group. From the mean difference test on application income, we see that LMI applicants have significantly lower incomes than upper-income applicants ($23,176 vs. $81,267), as expected. The mean difference tests for credit bureau scores and for custom credit scores are more revealing, however, as LMI applicants overall have significantly lower credit scores relative to upper-income applicants for both the custom (189 vs. 218) and bureau (661 vs. 678) score measures.[10]

Table 2 also provides the difference of means tests and distributional comparisons for the individual attributes, or factor loadings, of the custom scorecard. These attributes include: Time at current address, Number of bank trade lines, Finance company credit inquiries, Overall credit inquiries, Number of times 30 to 60 days late, Applicant income, Trade lines opened in less than 1 year, Highest revolving credit limit, Number of satisfactory credits, and Age of the credit bureau trade file in months. The two groups, LMI and upper income, reveal significant statistical differences across each of these attributes except Finance company credit inquiries and Overall credit inquiries. The only scorecard attribute on which LMI applicants fare better is Time at current address. This longer stay in residence may indicate a lack of upward mobility by these applicants, providing some evidence that risk characteristics may not be the same across income strata.

Table 3 contains the frequencies of application outcomes shown three different ways: Actual outcomes, Credit bureau-scored outcomes, and Custom-scored outcomes. Panel A reveals the breakdown of actual outcomes into approvals and denials, by income group as rendered by the lender. Of 1,698 LMI applicants, 890 were approved representing an approval rate of 52.41 percent. Of 568 upper income applicants, 431 were approved, resulting in an approval rate of 75.88 percent. The difference between these two approval rates, representing a disparity of 23.47 percentage points in favor of upper-income applicants, is statistically significant at the 99 percent level (Chi square statistic of 96.40).

Panel B of Table 3 details the approval and denial breakdown scenario should the applicants have been judged solely on the merits of the credit bureau score, using the cutoff score of 651 to create a denial rate for the group, which is identical to the overall actual denial rate for the sample, 41.7 percent. In this instance, the disparity in denial rates narrows from the original case of 23.47 percentage points to a disparity of 11.4 percentage points (44.52 percent LMI denials vs. 33.10 percent upper-income denials). The final panel of Table 3, Panel C, contains the approval and denial breakdown scenario should the applicants have been judged solely on the merits of the custom credit score, using the cutoff score of 191 to recreate the overall actual denial rate of 41.7 percent in the sample. In this instance, the difference in denial rates between LMI and upper-income applicants widens relative to the other two instances studied, to more than 30 percentage points, and represents the largest disparity among these three possible underwriting scenarios.

To recap, the results from Table 3 show that the actual bank decisions strike a middle ground between those dictated by the strict application of a custom credit score and the strict application of a credit bureau score. The most noticeable result, however, is the lower disparity that emerges through the strict application of the credit bureau score. This result may reflect the fact that credit bureaus do not have similar information sets relative to the banks (e.g., income, time at address) or alternatively, credit bureaus may be more concerned with disparate impact issues.[11]

### *Empirical Methods*

To assess the benefit of adding non-traditional credit information either to the credit bureau score approach or to the custom credit score approach, we use a three-step process for hypothesis testing. First, using the entire dataset, we create two separate logistic regression models — a judgmental model and a custom score model. The "judgmental" underwriting model employs the factors that a typical judgmental underwriter considers in making a credit determination, including the bureau score, the number of major derogatories, the number of minor derogatories, and whether a prior relationship existed with this lender. As a result, this judgmental model compliments the generic credit bureau score with a typical list of bank-specific factors an underwriter may employ in gauging the likelihood of loan repayment for a particular applicant. The resulting applicant-level probabilities of approval are used to "score" the applications based on the *prima facie* probability of denial of 41.7 percent. We classify the outcomes from the judgmental system in the same manner as we do for the custom score model. The custom score approach solely uses the attributes of the custom score model. We compare the denial disparity rates across these two underwriting models for LMI versus upper-income individuals to determine which of these approaches offers the lower disparity.

Second, since scholars and practitioners alike have shown underwriting models to perform differently for the overall population when compared with "marginal" credit applicants, we examine a similar set of tests for a group of marginal applicants. These marginal applicants represent those that have either characteristics or scores that are closer to the cutoff than the scores of the general applicant population. We define marginal applicants as applicants receiving an aggregate custom score of between 195 and 210. This group includes 444 applicants, or 20 percent of the overall sample.

Finally, after comparing the disparities across the two underwriting approaches, we identify three factors (the number of finance company inquiries, length of credit history, and applicant income) in the custom score model that drive the denial rate disparity between LMI and upper-income applicants in a custom scoring approach. We remove these three factors to determine the effect on the probabilities of approval across the two groups.[12]

## Results

### *Overall Sample*

Table 4 contains the outcomes for the judgmental underwriting model for the full sample. Using the weights derived from the logistic regression, the probability of approval derived from the judgmental model for each applicant, and the *prima facie* cutoff of 41.7 percent, we classify each of the applicants as either an approval or a denial in Panel A of Table 4. The resulting disparity in denial rates for the two groups is 13.0 percentage points, substantially smaller than the 30.8 percentage points disparity presented in the final panel of Table 3, where the custom score was used as the sole underwriting criteria.

Panel B of Table 4 contains the outcomes from the logistic regression using the 10 factors from the custom score underwriting model. Using the same classification approach, we separate the applicants into two groups based on income, as either approvals or denials, in Panel C of Table 4. The resulting denial of 24.7 percentage points disparity between these two groups is roughly double the judgmental model disparity reported in Panel A.

In sum, both approaches – judgmental and scorecard – result in significant disparities between LMI and upper-income applicants. On further review, however, the results show that the smaller of the two disparities is for the judgmental credit underwriting approach, affirming the null hypothesis that judgmental systems reduce the denial disparity between LMI and upper-income applicants over a custom credit scoring system, *ceteris paribus*.

### *Marginal Applicant Focus*

As previously discussed, the benefit or disbenefit of credit scoring tends to be illustrated best when reviewing marginal applicants. Marginal applicants are those with credit scores that are at or near the cutoff for denial. For our tests on the marginal sample, we show the breakdown across income groups in Panel A of Table 5. For the LMI group, the actual denial percentage from the decision file is 23.7 percent and for the upper income group, the denial percentage is 22.05 percent. As expected for the marginal sample, these two proportions across income groups are not statistically different at the weakest permissible statistical level of 90 percent.

Panel B of Table 5 employs the same judgmental underwriting model constructed for the overall sample in Table 4 and applies it to the subpopulation of marginal applicants. Using the *prima facie* cutoff of 23.4 percent for the entire marginal sample, we classify each of the applicants as either an approval or a denial in Panel B of Table 5. Of the 317 LMI applicants, 63 are denied using this rule, resulting in a denial rate of 19.9 percent. Forty-one of the 127 upper-income applicants are denied using this rule, resulting in a denial rate of 32.3 percent. The stated disparity is 12.4 percent in favor of LMI applicants and is statistically significant.

Panel C of Table 5 shows the outcomes for the 10-factor custom score model for the marginal sample. Here the disparity remains in favor of LMI applicants, but the size of the disparity is reduced by more than 60 percent to 4.7 percent versus the disparity from the judgmental system. This result provides further support for the null hypothesis that judgmental systems result in lower denial disparity between low-income and upper-income applicants compared with custom systems.

In our final test, we re-estimate the custom factor model after omitting the three variables argued by credit scoring detractors as likely to result in disparate impact. These variables include Applicant income, the Number of finance company inquiries, and the Highest revolving credit limit. Detractors have argued that income and credit limits are not robust predictors of creditworthiness and, therefore, should be scaled in a manner that better reflects the borrower's ability to pay, such as the debt-to-income ratio and the current debt-to-credit limit. Finally, given that low-income and minority borrowers are more likely to use nontraditional financial providers (e.g., finance companies), the inclusion of the Number of finance company inquiries has also been attacked as having the potential to disparately impact these groups by creating abnormally high rates of incidence.

For both samples, the full and the marginal, the seven-factor model results in significantly lower disparities than those derived from the full 10-factor custom model, confirming the hypothesis that use of these three variables increases the likelihood of denial for LMI applicants. For example, the full sample denial disparity in the seven-factor model is 14.4 percentage points, compared with a disparity of 24.7 percentage points from application of the 10-factor model (results are not shown due to space constraints). Similarly, the marginal group disparity is

10.5 percentage points in favor of LMI applicants in the seven-factor model, which is more than twice the 4.7 percent favorable disparity of the 10-factor model. When comparing these custom model disparities with the outcomes of the judgmental model in both the full sample and the marginal sample, however, the judgmental results are still more favorable to LMI applicants.

## Conclusions

As a result of the underwriting evolution toward the use of credit scoring in mortgage lending, scoring is at the forefront of the policy debate surrounding fair lending and potential disparate impact. Using 1996 loan application data on home improvement loans from a large commercial bank, we develop a framework for examining whether this system of credit scoring leads to more significant denial disparities between LMI applicants versus upper-income applicants when compared with disparities observed from a judgmental underwriting approach. We hypothesize that custom credit scoring systems result in larger disparities for LMI applicants, since these models neglect compensating or nontraditional credit factors that are more common for LMI applicants. Our findings confirm this hypothesis.

These findings are important for the current policy debate over the effect of credit scoring on LMI applicants. Proponents of credit scoring technology point out that scoring improves the objectivity of the loan decision and lowers the overall cost and time required to underwrite loans. Scoring detractors, however, are concerned that these models lack sufficient flexibility and often omit information important to the credit profile of LMI and minority applicants. Our findings lend support to the latter argument. In sum, we show that use of a custom credit score as the sole criteria in underwriting home improvement loan applications results in larger denial disparities between LMI versus upper-income applicants, *ceteris paribus*.

Finally, the results have important implications for bank supervision. Currently, bank supervisors, financial regulators, and researchers focus their fair lending concern on dealing with disparate treatment. The next generation of fair lending research, however, should begin to tackle issues related to the potential disparate impact of credit scoring. This need is especially high for internally-developed scoring models that have not been subject to much external scrutiny.

Although this paper focuses on the implementation of a single credit scoring model and the resulting underwriting disparities, future research must extend these results with loan performance data. Such research would determine if the inclusion of potentially discriminatory variables resulted from business necessity, in that these variables significantly influence the likelihood of default or default loss. If the potentially discriminatory variable(s) shows a strong relation to delinquency or default, research should assess the adequacy of alternative credit scoring variables that have a smaller adverse impact on certain segments of the population while maintaining or improving the predictiveness of delinquency or default.

*M. Cary Collins is Associate Professor of Finance and Director of the Financial Institutions Center at the University of Tennessee. His research focuses primarily on financial institutions and financial markets. As an author of over 20 publications and research papers, his work has appeared in numerous academic and professional journals, including Financial Management, The Financial Review, Journal of Economics & Finance, Journal of Real Estate Finance and Economics, Journal of Financial Research, Journal of Financial Services Research, and Journal of Real Estate Research. His research and opinions have appeared in The Wall Street Journal, The Atlanta Journal Constitution, and numerous other periodicals. Collins received his undergraduate degree from Concord College, West Virginia, his M.B.A. from Virginia Tech, and his Ph.D. from the University of Georgia.*

*Keith D. Harvey is Assistant Professor of Finance at Boise State University in Idaho. He received his doctorate from the University of Tennessee. Prior to pursuing his doctorate, he worked for a regional bank in Texas. His research focuses on financial institutions and corporate finance.*

*Peter J. Nigro is Senior Financial Economist in the Policy Analysis Division at the Office of the Comptroller of the Currency. His research interests include fair lending issues, credit-scoring models, mutual funds, and the economics of small business. He has published articles in numerous academic and professional journals, including Business Economics, Journal of Financial Services Research, and Managerial and Decision Economics. He received a bachelor's degree from the College of the Holy Cross, an M.A. in economics from the University of Southern California, and a Ph.D. in economics from Boston College.*

**Table 1**

Description of the unsecured, home improvement application sample for 1996. The table contains both the overall sample description across income groups and the break out of the sample by low- to moderate-income and upper income applicants. The 1996 HUD median income for the applicants MSA was used to create the bifurcation.

| Classification | No. of Applications | Fraction of Sample (%) |
|---|---|---|
| All income groups | 2266 | 100% |
| Low- to moderate-income group | 1698 | 74.9 |
| Upper-income group | 568 | 25.1 |

**Table 2**

Panel statistics for differences between low- to moderate-income and upper-income groups in 1996. The table contains the averages for the LMI group and the upper-income group, plus the Wilcoxon 2-sample test for differences across income groups and the Kolmogorov-Smirnov (KSa) test for distributional differences across income groups. The 1996 HUD median income was used to create the bifurcation.

| **Variables** | **LMI group** | **Upper income group** | **Wilcoxon (Z statistic)** | **Kolmogorov Smirnov (Ksa)** |
|---|---|---|---|---|
| Applicant income | $23,176 | $81,267 | 34.1*** | 20.6*** |
| Credit bureau score | 661.3 | 678.1 | 4.4*** | 2.6*** |
| Custom credit score | 189.8 | 218.7 | 14.3*** | 6.7*** |
| Custom Score Factors | | | | |
| Time at current address (months) | 175.2 | 98.8 | -8.5*** | 4.7*** |
| No. of Bank trade lines | 1.5 | 1.8 | 4.1*** | 1.8*** |
| No. of Finance company inquiries | 1.3 | 1.1 | -0.0 | 0.4 |
| No. of Overall inquiries | 2.9 | 2.9 | -0.5 | 0.5 |
| No. times 30-60 days late | 1.5 | 1.0 | -5.6*** | 2.5*** |
| Applicant income (monthly) | $1,931 | $6,772 | 35.7*** | 20.6*** |
| No. of Trade lines opened in <1 year | 2.1 | 2.6 | 4.8*** | 1.9*** |
| Revolving credit limit | $14,208 | $31,855 | 17.5*** | 7.8*** |
| No. of satisfactory trade lines | 10.7 | 18.8 | 18.6*** | 8.0*** |
| Age of trade file in months | 184.3 | 186.1 | 2.7*** | 3.1*** |

*  Indicates significance at the 0.10 level.
**Indicates significance at the 0.05 level.
***Indicates significance at the 0.01 level.

**Table 3**

2x2 matrix of outcomes for all 2,266 applications. Panel A contains the actual outcomes by income class for the population of 2,266 applications. Panel B contains the matrix of outcomes by income class, using the credit bureau cutoff score of 651 and a prima facie denial rate of 41.70 percent as the sole criteria for determining approval or denial. Panel C contains the matrix of outcomes by income class, using the custom credit score of 191 and a prima facie denial rate of 41.70 percent as the sole criteria for determining approval or denial. The row percentages appear to the right of each cell. The cumulative counts for each row appear as the right-most figures, while the cumulative counts for each column appear in the bottom row.

Panel A: Actual Outcomes

| **Income class** | **Approved** | **% Approved** | **Denied** | **% Denied** | **Total** |
|---|---|---|---|---|---|
| Low- to moderate-income | 890 | 52.41 [a] | 808 | 47.59 [a] | 1698 |
| Upper income | 431 | 75.88 [a] | 137 | 24.12 [a] | 568 |
| Total | 1321 | 58.30 | 945 | 41.70 | 2266 |

Denial Rate Disparity (Low vs. high) 23.47%

Panel B: Credit Bureau Score Outcomes

| **Income class** | **Approved** | **% Approved** | **Denied** | **% Denied** | **Total** |
|---|---|---|---|---|---|
| Low- to moderate-income | 942 | 55.48 [b] | 756 | 44.52 [b] | 1698 |
| Upper income | 380 | 66.90 [b] | 188 | 33.10 [b] | 568 |
| Total | 1322 | 58.34 | 944 | 41.66 | 2266 |

Denial Rate Disparity (LMI vs. high) 11.42%

Panel C: Custom Score Outcomes

| **Income class** | **Approved** | **% Approved** | **Denied** | **% Denied** | **Total** |
|---|---|---|---|---|---|
| Low- to moderate-income | 856 | 50.41 [c] | 842 | 49.59 [c] | 1698 |
| Upper income | 461 | 81.16 [c] | 107 | 18.84 [c] | 568 |
| Total | 1317 | 58.12 | 949 | 41.88 | 2266 |

Denial Rate Disparity (LMI vs. high) 30.75%

a The Chi-square statistics for the difference in the two proportions across groups is 96.40 and significant at the 0.01 level.
b The Chi-square statistic for the difference in the two proportions across groups is 22.86 and significant at the 0.01 level.
c The Chi-square statistic for the difference in the two proportions across groups is 165.34 and significant at the 0.01 level.

**Table 4**

2x2 matrix of outcomes for all 2,266 applications using weights from logistic regression for the various types of underwriting. Panel A contains outcomes by income class for the population of 2,266 applications using the judgmental model. Panel B contains the matrix of outcomes by income class using the custom scoring regression model weights. The row percentages appear to the right of each cell. The cumulative counts for each row appear as the right-most figures, while the cumulative counts for each column appear in the bottom row.

Panel A. Judgmental Model Outcomes

| **Income class** | **Approved** | **% Approved** | **Denied** | **% Denied** | **Total** |
|---|---|---|---|---|---|
| Low- to moderate-income | 934 | 55.01 * | 764 | 44.99 * | 1698 |
| Upper income | 386 | 67.96 * | 182 | 32.04 * | 568 |
| Total | 1320 | 58.25 | 946 | 41.75 | 2266 |

Denial Rate Disparity (LMI vs. high) 12.95%

Panel B. Custom Score Model Outcomes

| **Income class** | **Approved** | **% Approved** | **Denied** | **% Denied** | **Total** |
|---|---|---|---|---|---|
| Low- to moderate-income | 884 | 52.06* | 814 | 47.94 * | 1698 |
| Upper income | 436 | 76.76 * | 132 | 23.24 * | 568 |
| Total | 1320 | 58.25 | 946 | 41.75 | 2266 |

Denial Rate disparity 24.70%

* The Chi-square statistic for the difference in the two proportions across groups is 123.129 and significant at the 0.01 level.

**Table 5**

Panel A contains actual outcomes by income class for 444 marginal applicants. Panel B and Panel C contain 2x2 matrix of outcomes for all 444 marginal applications weights from the judgmental and custom logistic regression models. Panel B contains the matrix of outcomes by income class using the custom scoring regression model weights. The row percentages appear to the right of each cell. The cumulative counts for each row appear as the right-most figures, while the cumulative counts for each column appear in the bottom row.

Panel A. Actual outcomes by income class for those 444 applicants in the marginal region.

| **Income class** | **Approved** | **% Approved** | **Denied** | **% Denied** | **Total** |
|---|---|---|---|---|---|
| Low- to moderate-income | 242 | 76.34 | 75 | 23.66 | 317 |
| Upper income | 99 | 77.95 | 28 | 22.05 | 127 |
| Total | 341 | 76.80 | 103 | 23.20 | 444 |

Denial Rate disparity 1.61%

Panel B. Judgmental Model Outcomes for Marginal Applicants

| **Income class** | **Approved** | **% Approved** | **Denied** | **% Denied** | **Total** |
|---|---|---|---|---|---|
| Low- to moderate-income | 254 | 80.13% | 63 | 19.87% | 317 |
| Upper income | 86 | 67.72 | 41 | 32.28 | 127 |
| Total | 340 | 76.58 | 104 | 23.42 | 444 |

Denial Rate Disparity (LMI vs. High) 12.41%

Panel C. Custom Score Model Outcomes for Marginal Applicants

| **Income class** | **Approved** | **% Approved** | **Denied** | **% Denied** | **Total** |
|---|---|---|---|---|---|
| Low- to moderate-income | 247 | 77.92 | 70 | 22.08 * | 317 |
| Upper income | 93 | 73.23 * | 34 | 26.77 * | 568 |
| Total | 340 | 76.58 | 104 | 23.42 | 444 |

Denial Rate disparity 4.69%

* The Chi-square statistic for the difference in the two proportions across groups is 123.129 and significant at the 0.01 level.

## Notes

[1] See Edward M. Lewis (1994), *An Introduction to Credit Scoring* for a history of credit scoring models.

[2] Fair Isaac, one of the primary developers of scoring models employed by banks, estimates that when a bank changes from a judgmental to a scoring system they have a 20 to 30 percent increase in the number of applicants accepted with no increase in the loss rate. This is in addition to the reduction in processing costs and faster turnaround time.

[3] Traditional consumers include upper-income individuals with fairly extensive and long lasting credit histories.

[4] Banks and the credit bureaus do not collect or employ nontraditional forms of creditworthiness, such as information on payment history of utility bills and rent payments in their scoring models. Thus, it is argued by detractors of credit scoring that these models may not gauge adequately the true risk of this segment of the population.

[5] In recent years, these issues have been compounded by the fact that some subprime lenders, those that typically serve nontraditional groups, have neglected to report positive information on payment histories to credit bureaus to keep these profitable, but "high risk," customers captive. This omitted information further reduces the strength of scores for lower-income applicants. See "Credit Bureaus Move Against Lenders that Withhold Information," *American Banker,* December 30, 1999.

[6] Lending discrimination can take one of three forms: overt discrimination, disparate treatment, and disparate impact. Given the nature of how scoring models are developed, the first two forms of discrimination generally are moot when fully implementing a scoring model. Thus, disparate impact is the primary concern. However, the best way to determine whether scoring models have a potential disparate impact is to employ performance data. Unfortunately, we have been unable to obtain this information.

[7] This paper is concerned solely with the influence of scoring on the approval/denial decision. Scoring type models are used by banks for various other functions, including increasing or decreasing credit lines or loan rates and in the loan monitoring process.

[8] Van Order and Zorn (2000) examine mortgage loan default and loss rates by income levels and find that lower-income neighborhoods experience somewhat higher loss and default rates. Mills and Lubuele (1994), using a limited data set, conclude that LMI mortgages perform better than their high-income counterparts. Neither of these studies, however, controls for applicant credit history.

[9] The median MSAincome is not reported to protect the identity of the lender.

[10] Bureau scores can range anywhere from 400 to more than 800, while the custom card under investigation is scaled in a different manner and ranges from 50 to 276.

[11] The focus of fair lending exams at commercial banks typically analyzes disparate treatment issues with very little focus on disparate impact.

[12] Detractors of scoring models have argued against using debt to income, rather than income as a measure of ability to pay. Separately, other detractors have argued that both the use of finance company inquiries and income should not be used in models because of their high correlation with applicant race.

## References

"Automated Underwriting: Making Mortgage Lending Simpler and Fairer for America's Families", *Freddie Mac Report*, September 1996.

Avery, Robert, Raphael W. Bostic, Paul Calem, and Glenn Canner, "Credit Risk, Credit Scoring, and the Performance of Home Mortgages," *Federal Reserve Bulletin*, 82 (7), 1996, pp. 621-648.

_______. "Credit Scoring: Statistical Issues and Evidence," *Real Estate Economics*, 28 (1), pp. 105-134.

Calem, Paul and Susan M. Wachter, "Community Reinvestment and Credit Risk: Evidence from an Affordable-Home-Loan Program," *Real Estate Economics*, 27 (1), 1999, pp. 105-134.

"Credit Bureaus Move Against Lenders That Withhold Information," *American Banker*, December 30,1999.

"Giving No Credit Where Credit is Due," *ACORN Report*, December 1998.

Lewis, Edward M., An Introduction to Credit Scoring, *The Athena Press* (1994).

"Measuring Credit Risk: Borrower Credit Scores and Lender Profiles," *Fannie Mae Letter*, 1995, LL09-95.

Mills, E.S. and L.S. Lubuele, "The Performance of Residential Mortgages in Low- and Moderate-Income Neighborhoods," *Journal of Real Estate Finance and Economics*, 9(3), 1994, pp. 245-260.

Pierzchalski, L, “Guarding Against Risk,” *Mortgage Banker* (June), 1996, pp. 38-45.

“The Predictive Power of Selected Credit Scores,” *Freddie Mac Letter*, 1995, (Freddie Mac: McLean Virginia).

Van Order, Robert and Peter Zorn, “Income, Location and Default: Some Implications for Community Lending,” *Real Estate Economics*, Vol. 28 (3), 2000, pp. 385-404.