Testing Credit Scoring Models for Statistical Bias: Ushering in a New Era of Transparency

Contents

INTRODUCTION

In the consumer credit market, one of the most important challenges is to ensure credit scoring models are free from what is known as "statistical bias." Statistical bias concerns about a particular credit score model arise if the probability of default (PD) varies between various groups of consumers even though members of the respective groups receive the same score with the same model.

If there are significant differences in the probability of default between any population segments, the credit score model then has a statistical bias (either positive or negative) towards one or more such consumer groups. This, in turn, suggests that there is preferential treatment or mistreatment of those particular population segments.

Of course, this issue is related to a broader conversation around the enforcement of The Equal Credit Opportunity Act (ECOA), implemented by Federal Reserve Board's Regulation B (12 CFR 202), which prohibits discrimination in any aspect of a credit transaction on the basis of specific population classifications.

Another area of particular concern is the level of minority homeownership. According to a recent Washington Post story, "Forty-three percent of blacks owned homes in 2017, according to an annual report from the Joint Center for Housing Studies of Harvard University. In contrast, 72 percent of whites did, a gap that has mostly widened during the past three decades."

VantageScore®
a higher level of confidence

Credit scoring models have been cited as a contributor to the problem. Indeed, recent comments from Washington D.C. Attorney General Karl Racine exemplify these concerns:

> "Classic FICO in particular, perpetuates historic discrimination in mortgage lending to the detriment of communities that are protected by the Fair Housing Act, 42 U.S.C. §§ 3601"
>
> – Washington D.C. Attorney General Karl Racine, FHFA Credit Score RFI

In the past, these conversations were all too often held in private. Today's marketplace demands greater transparency. VantageScore has always held itself to the highest standards of transparency and competitors are encouraged to follow our example.

This paper examines whether or not the latest VantageScore credit scoring model - VantageScore 4.0[1] - evenly distributes PD across population segments and provides a methodology for best practices for lenders to test their own portfolios. The examination includes testing ethnic populations as well as new scoring (universe expansion) vs. mainstream scoring consumers.

EXECUTIVE SUMMARY REGARDING VANTAGESCORE 4.0

1. For both credit card and first mortgage loans, default curves are statistically aligned among ethnic groups at each credit score value and among the overall population.
2. For all ethnic groups, there is near alignment of default curves, indicating an unbiased model.
3. For score bands with sufficient sample sizes[2], there is no discernible difference in the probability of default among consumers who are conventionally scored by credit score models (mainstream consumers) and consumers who are not conventionally scored in the credit industry (e.g., universe expansion consumers).
4. More broadly, credit scoring models can and should be publicly scrutinized for bias towards population segments, including minority groups and those who do not exhibit traditional credit behaviors.

METHODOLOGY

A formal statistical test that determines if there are differences in default probabilities between population segments is the "Chi-Square" test for multiple probabilities. To perform this test on VantageScore 4.0, score bands are created to ensure that there is a sufficient sample in each band for statistical robustness. In each score band (i.e. 500-520, 521-540, etc.), the Chi-Square Probability Comparison Test is applied to assess if there are statistically significant differences in the actual default probabilities among population segments.

The two asset classes selected for the test are credit card and first mortgage. These were chosen for their relative size and because they are representative of mainstream lending.

In order to do this test, a "critical value" must be established. In this test, the critical value is the demarcation point at which it is statistically decided whether the default rate associated with a particular population segment is significantly different than that of the whole population. A critical value of 95% was chosen[3].

If the differences between the population segments and whole population default probabilities are large (i.e., statistically significant, at or beyond the 95% demarcation), then there is a demonstrated measurable bias. If not, then there is no measurable impact. The test is performed across all score bands. If even one band fails the test, then there is an implication of bias for the model as a whole.

1 All previous VantageScore models have also been tested and were proven not to produce statistical bias. Results of those tests are available publicly at www.VantageScore.com.

2 As the score bands get higher, the number of consumers in the universe expansion population declines.

3 In Chi-Square testing, it is best practice to use a 95% critical value.

Lower and upper thresholds are also produced, based on the Chi-Square test values for each score band, to determine where each population segment is considered within normal population boundaries. These thresholds build confidence curves around the difference in probabilities of default between population segments for each score band.

For more details on this methodology, please refer to "Testing Methodologies for Credit Score Models to Identify Statistical Bias toward Protected Classes", May 2014 (www.vantagescore.com/statbiasWP).

VANTAGESCORE 4.0 STATISTICAL BIAS: ETHNICITY

The statistical bias test is applied on ethnic protected class population segments, i.e., African-American (AOMC) and Hispanic-American (AOHC) populations.[4] Census data based on zip code was appended to the consumer as a proxy for ethnicity.

CHI-SQUARE TEST STATISTICS

For each VantageScore 4.0 score band, no Chi-Square test statistics calculation is larger than the critical value (see Figure 1 for credit card and Figure 2 for first mortgage), indicating there are no measurable differences between the groups at each credit score band and the overall population default rates for credit card or first mortgage.

Figure 1: Statistical bias ethnicity test statistic for credit card

African-American Population

VantageScore 4.0 Range	Start Point	821	801	781	761	741	721	701	681	661	641	621	601	581	561	541	521	500
	End Point	850	820	800	780	760	740	720	700	680	660	640	620	600	580	560	540	520
Test Chi-Square		0.590	0.028	0.085	3.121	8.307	6.848	2.731	3.858	1.898	5.492	0.001	0.154	1.579	0.968	0.001	1.304	0.058
Critical Value		8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844
Is Test -> Critical Value (if "Yes" then Statistical Bias)		NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

Hispanic-American Population

VantageScore 4.0 Range	Start Point	821	801	781	761	741	721	701	681	661	641	621	601	581	561	541	521	500
	End Point	850	820	800	780	760	740	720	700	680	660	640	620	600	580	560	540	520
Test Chi-Square		0.007	1.608	3.396	0.369	1.806	6.804	1.968	0.070	0.156	0.023	0.202	1.856	0.979	0.180	0.058	0.978	0.050
Critical Value		8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844
Is Test -> Critical Value (if "Yes" then Statistical Bias)		NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

Figure 2: Statistical bias ethnicity test statistic for first mortgage

African-American Population

VantageScore 4.0 Range	Start Point	821	801	781	761	741	721	701	681	661	641	621	601	581	561	541	521	500
	End Point	850	820	800	780	760	740	720	700	680	660	640	620	600	580	560	540	520
Test Chi-Square		1.736	3.482	0.059	2.496	4.190	0.896	6.619	4.578	1.317	4.275	0.404	0.424	1.567	0.998	1.294	0.008	0.487
Critical Value		8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844
Is Test -> Critical Value (if "Yes" then Statistical Bias)		NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

Hispanic-American Population

VantageScore 4.0 Range	Start Point	821	801	781	761	741	721	701	681	661	641	621	601	581	561	541	521	500
	End Point	850	820	800	780	760	740	720	700	680	660	640	620	600	580	560	540	520
Test Chi-Square		3.056	1.255	1.831	0.044	3.311	0.001	0.104	0.282	1.126	0.237	1.113	0.723	0.403	0.149	0.235	0.533	0.430
Critical Value		8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844	8.844
Is Test -> Critical Value (if "Yes" then Statistical Bias)		NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

4 Due to inadequate sample sizes, other ethnic populations were included in the Non AOMC/AOHC populations.

STATISTICAL BIAS PROFILE

For both credit card (Figure 3) and first mortgage (Figure 4), graphical comparisons show all ethnic class default curves are statistically aligned. Both ethnic group curves, African-American (AOMC) and Hispanic-American (AOHC), fall well within their confidence intervals (Figure 3).

Figure 3: Statistical bias profile by ethnicity on credit card products

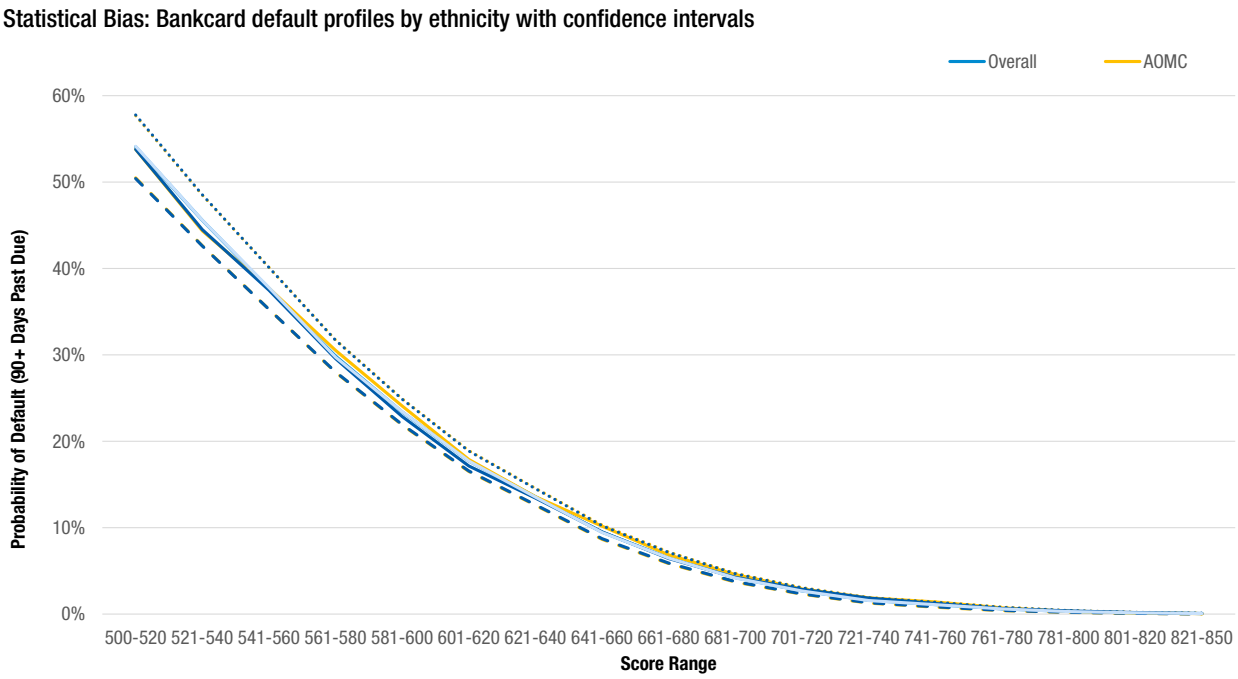
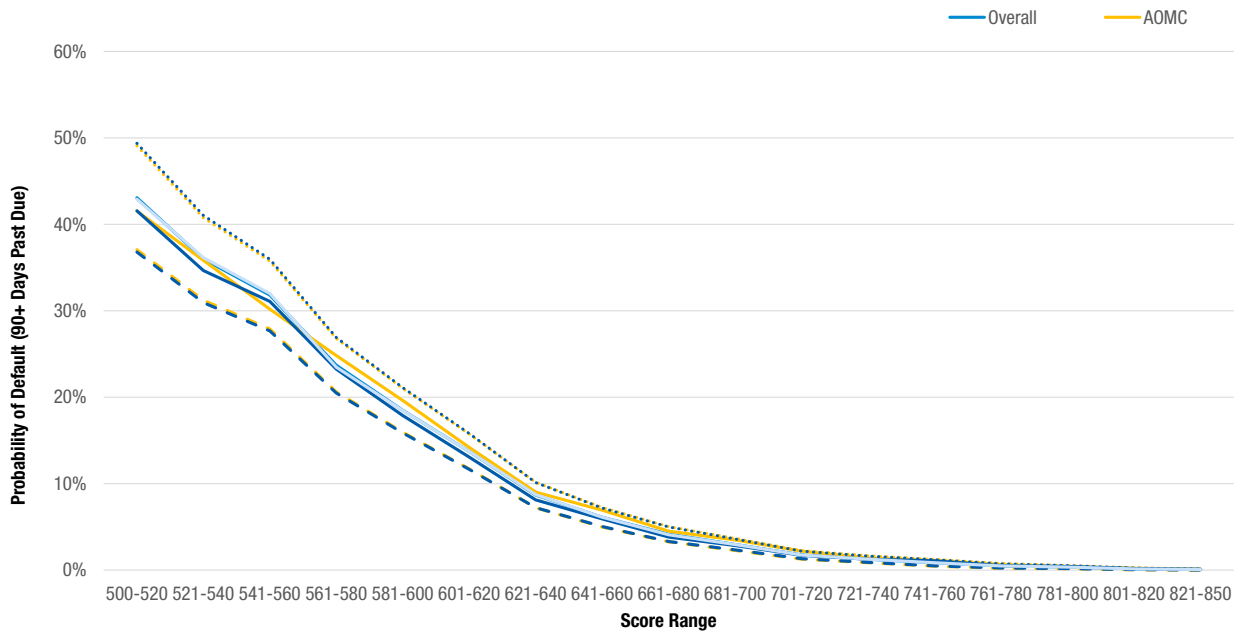


Figure 4: Statistical bias profile by ethnicity on first mortgage products



VANTAGESCORE 4.0 STATISTICAL BIAS: NEW SCORING (UNIVERSE EXPANSION) VS. MAINSTREAM SCORING CONSUMERS

The statistical bias test is applied on the universe expansion vs. mainstream consumer population segments for the score ranges that had a sufficient sample for a robust statistical Chi-Square test. Using the same methodology, there were no discernible differences in the probability of default within each score band between consumers who are scored by conventional credit score models (mainstream consumers) and consumers who are not conventionally scored in the credit industry (e.g., universe expansion consumers). Note that there is insufficient sample size for score bands above 660.

Figure 5:

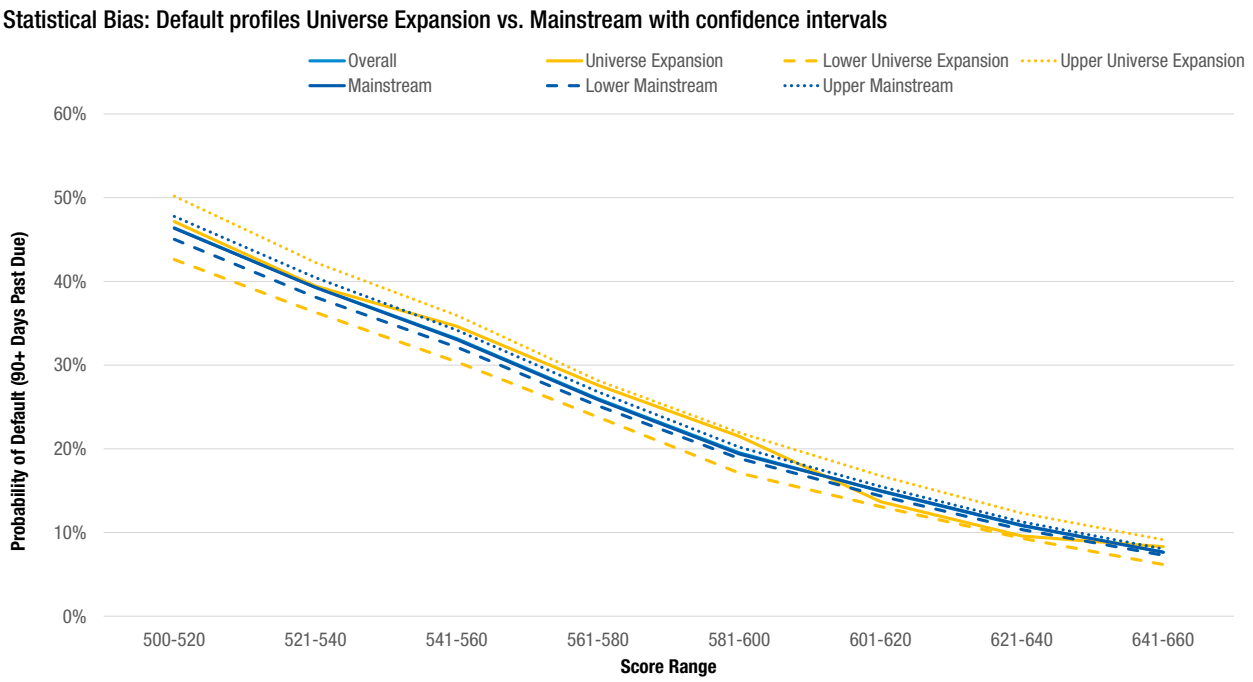


Figure 6:

Universe Expansion

VantageScore 4.0 Range	Start Point	641	621	601	581	561	541	521	500
	End Point	660	640	620	600	580	560	540	520
Test Chi-Square		1.276	3.866	2.689	7.101	3.597	1.804	0.016	0.245
Critical Value		7.477	7.477	7.477	7.477	7.477	7.477	7.477	7.477
Is Test -> Critical Value (if "Yes" then Statistical Bias)		NO	NO	NO	NO	NO	NO	NO	NO

Mainstream

VantageScore 4.0 Range	Start Point	641	621	601	581	561	541	521	500
	End Point	660	640	620	600	580	560	540	520
Test Chi-Square		0.024	0.093	0.067	0.246	0.102	0.071	0.000	0.010
Critical Value		7.477	7.477	7.477	7.477	7.477	7.477	7.477	7.477
Is Test -> Critical Value (if "Yes" then Statistical Bias)		NO	NO	NO	NO	NO	NO	NO	NO

For both credit card and first mortgage there is no evidence of bias toward protected classes when using VantageScore 4.0.

CONCLUSION

From regulators to consumer advocates and civil rights groups, and from lenders to the consumers they serve, all stakeholders in the consumer credit marketplace are deeply concerned about fair lending and ECOA violations.

The answer is transparency. By openly disclosing model test results, whether for statistical bias, predictive performance or other important measures, the roots of potential discrimination can be identified and eradicated.

Since its inception, VantageScore has promoted transparency, especially as it applies to the predictiveness of its models and treatment of consumers. The credit scoring model marketplace would be wise to follow this example.

The findings of the testing methodology subject to this white paper communicates two important messages: (1) credit scoring models can and should be publicly scrutinized for bias towards population segments, including minority groups and those who do not engage in traditional credit behaviors, (2) default rates associated with using the VantageScore 4.0 model are aligned when comparing ethnic population segments to the overall population and when comparing those who are conventionally scoreable versus those who are newly scoreable.

The VantageScore credit score models are sold and marketed only through individual licensing arrangements with the three major credit reporting companies (CRCs): Equifax, Experian and TransUnion. Lenders and other commercial entities interested in learning more about the VantageScore credit score models, including the VantageScore 4.0 credit score model, may contact one of the following CRCs listed for additional assistance:

VantageScore
August 2018

www.vantagescore.com

EQUIFAX®
Call 1-888-202-4025
http://VantageScore.com/Equifax

experian.
Call 1-888-414-4025
http://VantageScore.com/Experian

TransUnion
Call 1-866-922-2100
http://VantageScore.com/TransUnion