Chapter 8

Marginal Effects Estimates for Continuous Race Proxies in Limited Dependent Variable Models

Jason Dietrich
March 2026

I. Introduction

Most analyses of discrimination require individual-level demographic data. In many situations, self-reported demographic data are not available, so researchers rely on proxies. The Bayesian Improved Surname Geocoding (BISG) proxy developed by the RAND Corporation has gained significant prominence as the standard proxy for analyses of racial discrimination, particularly for credit markets. Since its development, a substantial amount of research has assessed the accuracy of the BISG proxy, compared the impact of using proxies in continuous form versus a variety of discrete forms, and explored alternative proxies based on different methodologies or different sets of predictors.[1] However, for analyses of discrimination using BISG proxies, little to no research has focused on what specific disparity measure to use, the underlying characteristics of disparity measures, or how to interpret disparity measures. This paper begins to fill that gap by analyzing the underlying characteristics of one specific disparity measure – marginal effects – when using a logistic estimator to estimate a limited dependent variable model with continuous BISG proxies. This is a common analytical approach economists use to test for discrimination in approval/denial decisions on credit applications, so the results of this paper will inform the statistical analyses regulators and industry conduct as part of supervisory fair lending exams.

Constructing marginal effects estimates of continuous variables in limited dependent variable models is not a new or novel topic. Marginal effects at the mean (MEM) and average marginal effects at actual values for all variables (AME) are two standard approaches researchers have used for these types of analyses for many years. However, the specific use case of using

[1] For research on the accuracy of BISG proxies see Greenwald et.al. (2023). For research comparing the impact of using proxies in continuous form versus a variety of discrete forms see Zhang (2016). For research exploring alternative proxies based on different methodologies or different sets of predictors see Voicu (2018).

continuous BISG proxies in a limited dependent variable model to test for discrimination introduces some unique issues that make further exploration of the underlying characteristics of marginal effects estimates in this modeling scenario an interesting and useful exercise. First, the set of six BISG proxies sum to one, so changing the values of one proxy variable necessitates an offsetting change in one or more of the other proxy variables. The potential impact this required offsetting change has on marginal effects estimates would not occur for other stand-alone continuous variables such as FICO score. Second, the distributions of BISG proxy variables tend to be bimodal with a significant mass point near 0. Since marginal effects estimates using a logistic estimator are application-specific, the unique distributions of BISG proxies will impact the distributions of the application-level marginal effects estimates used to generate the MEM and AME. Finally, given that marginal effects estimates in a limited dependent variable model are a function of all other variables in the model, as well as their coefficient estimates, the joint distributions of the proxy variable of interest and all other variables in the model will impact the marginal effects estimates. As one example, applications with one or two predominant proxies will impact marginal effects estimates differently than applications with larger values for three or more of the proxy variables. This potential impact is especially interesting given the recent increases in demographic diversity as comparisons of the 2010 and 2020 Census data show.[2] As a second example, the marginal effects estimates for the race proxy variable of interest will partially reflect impacts of each non-race proxy variable in the model. This partial impact is akin to a disparate impact effect from these variables, which is of particular interest to regulators.

In this paper we explore the underlying characteristics of marginal effects estimates for continuous BISG proxy variables when using a logistic estimator to estimate a limited dependent

[2] See for example, Updating BISG Proxies: How 2020 Census Data Impacts Fair Lending Risk

variable model of underwriting decisions on credit applications. Specifically, we explore how the unique aspects of this use case affect analyses and interpretations of marginal effects estimates. We focus on two specific questions: 1) how do proxy variables for other racial/ethnic groups impact the marginal effects estimates for a given proxy variable of interest and 2) how much of the marginal effects estimate for a given race proxy variable is due to indirect impacts from other variables in the model? We begin by re-creating the formal theoretical framework of marginal effects estimates for a logistic estimator. The theoretical framework shows the types of characteristics and patterns that occur for predicted probabilities of denial and marginal effects estimates when using a logistic estimator and continuous proxies. We then use simulated BISG proxy data that reflects the type of race proxies that occur during fair lending analyses, along with simulated data for FICO score and approve/deny decisions, to generate empirical examples illustrating these theoretical characteristics and patterns. Finally, we conclude with a set of recommendations for best practices on how to evaluate and interpret marginal effects estimates for these types of analyses.

The main takeaways in this report, which apply specifically to using a logistic estimator to estimate limited dependent variable models including continuous BISG proxies, are:

- The predicted probabilities of denial and application-specific marginal effects estimates for race will be a function of all other race proxy variables and their coefficient estimates, which in turn will impact aggregate marginal effects estimates for race. Therefore, it is important to analyze and understand the joint distributions of the race proxy variables, as well as their coefficient estimates.

- The predicted probabilities of denial and application-specific marginal effects estimates for race will be a function of all other variables in the model and their coefficient estimates, which in turn will impact aggregate marginal effects estimates for race. It is important to analyze and understand these indirect impacts.

II. Theoretical Framework

For all analyses we use a logistic estimator to estimate a limited dependent variable model of the probability that a lender denies applications for credit. Berkson (1944) originally developed the logistic estimator, so the underlying formulations of the estimator are nearly one hundred years old and can be found in numerous econometric textbooks and research papers. However, since these formulations provide the foundation for understanding the underlying characteristics of the marginal effects disparity measure, which is the focus of this report, it is instructive to re-generate all of the formulas in detail here.

Let y^* be an unobserved continuous latent measure of the probability that a lender denies applications for credit. Applications are indexed from j = 1 to n, but we exclude the application index throughout for ease of exposition.[3] The probability of denial is a function of several creditworthiness factors captured by matrix X; demographic group membership captured by five race proxy variables, R_i, where i = 1 to 5; and an error term, ϵ.

$$y^* = \beta_0 + \theta X + \sum_{i=1}^{5} \beta_i R_i + \epsilon \quad (1)$$

For the following discussion, let $Z = \beta_0 + \theta X + \sum_{i=1}^{5} \beta_i R_i$. Although we do not observe y^*, we do observe y, a 0/1 denial indicator variable such that,

$$y = \begin{cases} 1 \; if \; y^* > 0 \\ 0 \; otherwise \end{cases} \quad (2)$$

The probability of denial is then,

$$P(y = 1) = P\,(y^* > 0) = \; P(\epsilon > \; -Z) = 1 - \mathrm{F}(-\mathrm{Z}) \quad (3)$$

where F() is a cumulative distribution function. The functional form of F() depends on assumptions about the distribution of the error term ϵ. If we assume the distribution of the error

[3] The variables y^*, y, and R_i; the variables in matrix X; and the error term (ϵ) are application specific, so the j subscript is dropped for those terms. All of the parameters of the model (β_0, θ, β_i) are constant across applications, so the j subscript is not relevant.

term is logistic or extreme value, we have a logistic estimator.[4] Further, if F() is symmetric, which it is for the logistic distribution, then,

$$1 - F(-Z) = F(Z), \text{ so } P(y = 1) = F(Z) \tag{4}$$

The cumulative logistic distribution function is,

$$P(y = 1) = F(Z) = \frac{\exp(Z)}{1+\exp(Z)} \tag{5}$$

This paper focuses specifically on the marginal effects disparity measure. For a logistic estimator, the marginal effect of R_i on the probability of denial is the derivative of F(Z) w.r.t. R_i,

$$\frac{\partial F(Z)}{\partial R_i} = \frac{((1+e^{\beta_0+\theta X+\Sigma \beta_i R_i}) * \beta_i * e^{\beta_0+\theta X+\Sigma \beta_i R_i}) - (e^{\beta_0+\theta X+\Sigma \beta_i R_i} * \beta_i * e^{\beta_0+\theta X+\Sigma \beta_i R_i})}{(1+e^{\beta_0+\theta X+\Sigma \beta_i R_i})^2} \tag{6}$$

Upon simplification we get,

$$\frac{\partial F(Z)}{\partial R_i} = \frac{\beta_i * e^{\beta_0+\theta X+\Sigma \beta_i R_i}}{(1+e^{\beta_0+\theta X+\Sigma \beta_i R_i})^2} = \frac{\beta_i}{1+\exp(Z)} F(Z) = \beta_i * \left(1 - F(Z)\right) * F(Z) \tag{7}$$

Equation (7) has two key features. First, the marginal effect of R_i on the probability of denial is application-specific, although as noted above we have dropped the application j subscripts for ease of exposition. Second, via Z, the marginal effects are a function of all variables in the model, as well as the coefficient estimates on those variables. As a point of comparison, using an OLS estimator to estimate a linear model, the estimated marginal effect of R_i on the outcome is constant across all applications, and not a function of any other variables in the model or their coefficient estimates.

Given that marginal effects estimates are application-specific for continuous race proxies in a limited dependent variable model estimated using a logistic estimator, a fundamental question is how to combine application-specific marginal effects estimates into one aggregate marginal effect estimate. MEM and AME are two standard aggregation strategies economis use.

[4] If we assume the error term is normal, we have a probit estimator.

For the MEM measure of the marginal effect of R_i, equation (7) is first used to generate application-specific marginal effects, holding all variables except R_i, at their means for all applications. For these calculations and for most calculations that follow below, the β and Θ values are the estimated coefficients from the logistic regression. MEM is then generated as the average of these marginal effects estimates across all applications. The MEM is not a valid disparity measure for the modeling scenario we focus on, because, for any given application, the sum of the actual proxy value for R_i and the averages of the proxy variables for each of the other races will not sum to 1, which violates an identity of BISG proxies. For the AME measure of the marginal effect of R_i, equation (7) is first used to generate application-specific marginal effects, keeping all other variables at their actual values for all applications. AME is then generated as the average of these marginal effects estimates across all applications. The AME does not violate any underlying data identities of BISG proxies.

To help characterize the marginal effects estimates and identify the best aggregation approach, it is useful to bound the application-specific marginal effects estimates. For the standard s-shaped curve for predicted probabilities of denial from a logistic regression, the maximum and minimum application-specific marginal effects estimates will occur at inflection points of F(Z), which can be identified by setting the second derivative of F(Z) w.r.t. R_i equal to 0. The second derivative of F(Z) w.r.t. R_i is,

$$\frac{\partial^2 F(Z)}{{R_i}^2} = \frac{\beta_i}{1+\exp(Z)} \frac{\partial F(Z)}{\partial R_i} + F(Z)\,[\frac{(1+\exp(Z))*0-\beta_i*(0+\beta_i\exp(Z))}{(1+\exp(Z))^2}] \tag{8}$$

Upon simplification we get,

$$= \frac{\beta_i^2 F(Z)}{(1+\exp(Z))^2}(1-\exp(Z)) \tag{9}$$

The second derivative equals 0 when β_i, F(Z), or (1 – exp(Z)) equals zero. The case where β_i= 0 is not interesting, since the marginal effect of R_i would be 0 by definition. F(Z) equals 0 when exp(Z) = 0, which occurs in the limit as Z goes to negative infinity. Since this is unlikely to occur, this is not an interesting case either. That leaves (1 – exp(Z)). This term equals 0 if exp(Z) = 1, which occurs when Z = 0. Using equation (5) with Z = 0 we have,

$$P(y = 1) = F(0) = \frac{\exp(0)}{1+\exp(0)} = \frac{1}{2} \qquad (10)$$

Equation (10) suggests that the maximum and minimum application-specific marginal effects estimates occur for applications with a predicted probability of denial equal to 0.5. This has interesting implications for fair lending analyses using continuous race proxies. When discrimination has occurred and a given race has a large and strong impact on the probability of denial, it is common for the predicted probabilities of denial to tend toward 0 when the values of the given race proxy are low and toward 1 when the values of the given race proxy are high. Applicants with mid-level values for the given race proxy provide a weaker signal of an applicant's race that can create uncertainty for a biased decisionmaker, which would likely lead to mid-level predicted probabilities of denial. Based on equation (10), the first two subsets of applicants with low and high race proxy values would therefore tend to have lower estimated marginal effects, and applicants with mid-level proxy values would tend to have larger estimated marginal effects. Thus, per equation (10), applicants with mid-level proxy values would have the largest impact on the aggregate marginal effects estimates.

Plugging the result from equation (10) back into the equation for marginal effects (equation (7)) we get,

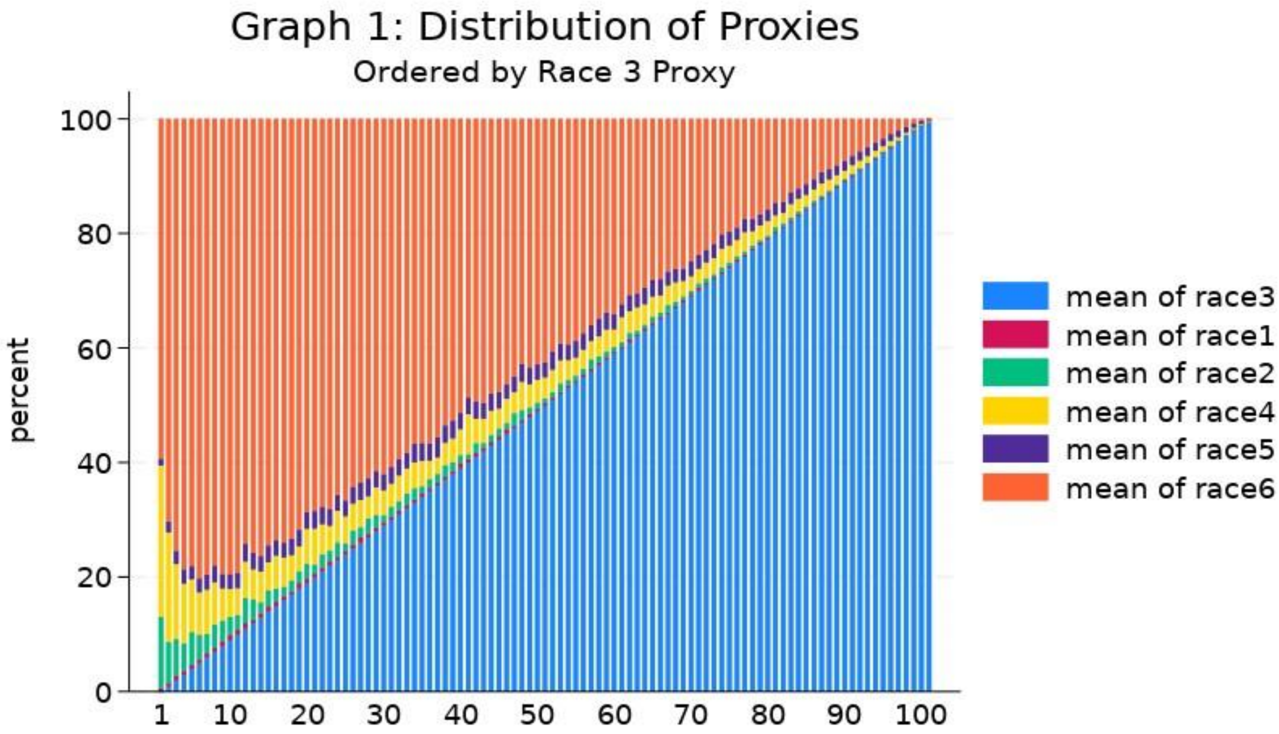
$$\frac{\partial F(0)}{\partial R_i} = \frac{\beta_i}{1+\exp(0)} F(0) = \frac{\beta_i}{2} * \frac{1}{2} = \frac{\beta_i}{4} \qquad (11)$$

Equation (11) suggests that, if $\beta_i > 0$, the maximum application-specific marginal effects estimate is $\frac{\beta_i}{4}$. Using equation (7), the minimum will be 0, since F(Z) and (1 + exp(Z)) are both non-negative for all Z, and F(Z) goes to 0 in the limit as Z goes to negative infinity. Similarly, if $\beta_i < 0$, the minimum application-specific marginal effects estimate is $\frac{-\beta_i}{4}$ and the maximum is 0. Therefore, the application-specific marginal effects estimates are bounded by $\frac{-\beta_i}{4}$ and $\frac{\beta_i}{4}$. We show this graphically in the empirical analysis below.

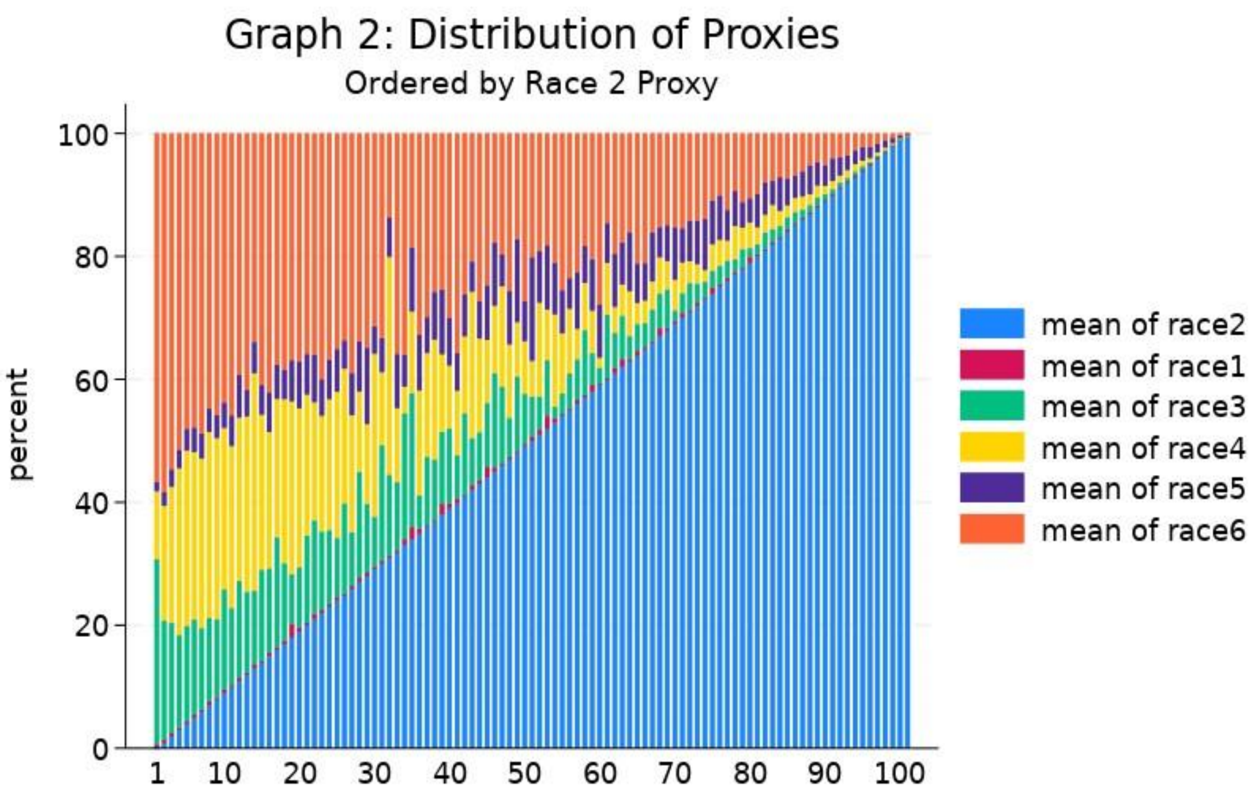
Before moving on to the empirical analysis, we present three additional theoretical results that will help inform our choices of parameter values for the data generating process (DGP) we use to simulate data for the analysis. First, the relationship between a continuous control variable and the predicted probabilities of denial from a logistic estimator will tend to be s-shaped. For a given continuous variable, this s-shaped relationship will be clear and pronounced when the predicted probabilities of denial are near 0 for the smallest values of the continuous variable and near 1 for the largest values of the continuous variable. Using equation (5), the predicted probabilities of denial will be near 0 when Z is approximately -4 or smaller, and near 1 when Z is approximately 4 or greater. In a very simplified model with no X variables, and only two races with non-zero proxy values -- R_1 and the omitted race --, there will be a clear and pronounced s-shaped relationship between R_1 and the predicted probabilities of denial when β_1 is at least 8 and the constant is -4. In this very simplified model, Z would range from -4 (= -4 + (0*8)) when R_1 is near 0 to 4 (= -4 + (1*8)) when R_1 is near 1, and the s-shaped relationship between R_1 and the predicted probabilities of denial would be clear and pronounced. This is the first result we use to inform our choices of parameter values in the DGP below.

The second result we utilize in the DGP is the impact of the constant on the location and shape of the s-shaped curve. Continuing the example of the very simplified model, a higher constant value will shift the s-shaped curve to the right and truncate the lower, left tail of the curve. A lower constant value will have the opposite effect, shifting the s-shaped curve to the left and truncating the upper, right tail of the curve. Again, we use this result to inform our choices of parameter values in the DGP below.

The final result we utilize in the DGP below focuses on how additional variables in a model affect the s-shaped relationship between R_i and the predicted probabilities of denial, which in turn will impact the estimated marginal effects of R_i. Adding variables to a model introduces vertical variation to the s-shaped curve for R_i. The extent of this vertical variation depends on the distributions of the additional variables, as well as their coefficient estimates. The next section discusses this vertical variation in detail and provides graphical examples. Here, we use the set of six race proxies to provide an example of the types of differences in the distributions of additional variables that can drive this vertical variation. Graph 1 presents the distribution of race proxies for each value of R_3 using the simulated data we generate in the next section. We multiple all proxies by 100 to transform the data to percentages, so all proxies range from 0 to 100 here. The horizontal axis shows the values R_3 rounded to the nearest integer. The vertical axis shows the means of each race proxy variable for all applicants with each specific value of R_3. As an example of how to interpret Graph 1 look at the distribution of proxy values when R_3 is 10%. After rounding R_3 for all applicants in the full dataset, there are 240 applicants with a value of 10%. For these 240 applicants, the average of R_3 is 10% by definition, as shown by the blue bar. The averages of the proxies for races 1, 2, 4, 5, and 6 for the 240 applicants are 0.6%, 1.4%, 6.0%, 2.7%, and 79.3%, respectively. The length of the vertical bars for each race at a value of 10% for R_3 reflects these percentages. Since the set of six proxies sums to 100 for each applicant, these six means sum to 100 as well. As a point of comparison, Graph 2 shows the same types of distributions for R_2. Fair lending analyses focus on comparing treatment across groups (minority to non-minority for example), so the key to understanding Graphs 1 and 2 is to


Graph 1: Distribution of Proxies
Ordered by Race 3 Proxy
percent
0
20
40
60
80
100
1
10
20
30
40
50
60
70
80
90
100
mean of race3
mean of race1
mean of race2
mean of race4
mean of race5
mean of race6

R_3


Graph 2: Distribution of Proxies
Ordered by Race 2 Proxy
percent
0
20
40
60
80
100
1
10
20
30
40
50
60
70
80
90
100
mean of race2
mean of race1
mean of race3
mean of race4
mean of race5
mean of race6

R_2

focus on the composition of the alternative races. As noted above, applicants with a value of 10% for R_3, on average, have a value of 79.3% for R_6, and small percentages for each of the other race proxies. Based on this proxy distribution, these applicants can be viewed to a large extent as either race 3 or race 6. This distribution of proxy values may impact how the lender treats these applicants, and will impact the predicted probabilities of denial and estimated marginal effects of R_3. Looking now at Graph 2, applicants with a value of 10% for R_2 have a very different distribution of proxies. For these applicants, on average, values for R_6, R_3, and R_4 are approximately 47%, 14%, and 26%, respectively, with small percentages for each of the other race proxies. Again, this distribution of proxy values may impact how the lender treats these applicants, and will impact the predicted probabilities of denial and estimated marginal effects. The main point here is that the distributions of other variables in the model, whether it be other race proxies as in the examples here or FICO score as in examples below, need to be explored and understood since they will impact the predicted probabilities of denial, as well as the estimation and interpretation of the estimated marginal effects of race.

III. DGP for Simulated Data

The analysis below presents empirical examples illustrating the concepts developed in the Theoretical Framework section above. Due to confidentiality restrictions, we cannot use actual exam data and results as real-world examples here, so instead we use simulated data. The primary advantage of using simulated data is that we control all aspects of data generation, so we know exactly what the results should be, which is useful when illustrating concepts. The primary disadvantage is that the results will be specific to the assumptions of the DGP and may not be generalizable. This is not as much of a concern here, since the primary objective is to illustrate theoretical concepts and highlight general analytical issues that Economists should consider when using race proxies for fair lending analyses.

As an initial step of the DGP we randomly generate data for the set of six race proxies (R_1 through R_6) for a sample of 55,755 applicants.[5] Consistent with the properties of BISG proxies, we ensure that all values for each proxy variable are between 0 and 1 and that the sum of the six proxy values sum to 1 for each applicant. In addition, we generate proxies with similar distributional characteristics (means and shapes) as found in race proxy data for real-world fair lending analyses.

As a second step of the DGP we then utilize a parsimonious model of the probability of denial as a function of just a constant and five race proxy variables. Specifically, we define Z from equation (1) above as,

$$Z = -4 + (0.1 * R_1) + (3 * R_2) + (8 * R_3) + (5 * R_4) + (0.1 * R_5) \qquad (12)$$

The choice of parameter values for each race proxy variable, as well as their relationships to the constant value, are particularly important to the DGP, since they have a significant impact on all subsequent results. Using the three results discussed at the end of the last section, we chose parameter values that generate output clearly demonstrating the types of characteristics and patterns detailed in the Theoretical Framework section. Given the theoretical basis for these characteristics and patterns, they will occur to some extent for all real-world applications involving a logistic estimator and continuous race proxies. However, given our approach of choosing parameter values that accentuate these characteristics and patterns, the results for real-world applications will likely be much more muted compared to the simulation results presented here.

---

[5] We use a sample size of 55,755 applicants because it is a large sample, which facilitates illustrating some of the concepts of interest, and it is the sample size of a recent exam involving race proxies that we supported.

To help understand our choices of parameter values in equation (12), we walk through a specific example focused on R_3. To start, we focus on the subset of applications where R_3 and R_6 (the omitted race) sum to approximately 1. For these applications, the proxy values for all other races are close to 0 and Z is determined primarily by the term -4 + (8 * R_3). As noted above, if Z ranges from -4 to 4 across the range of R_3, then the relationship between R_3 and the predicted probabilities of denial will be a clear s-shaped curve. Given that R_3 ranges from 0 to 1, by setting the parameter on R_3 to 8 and the constant to -4 we get Z = -4 when R_3 is near 0 (Z = -4 + (8 * 0) = -4) and Z = 4 when R_3 is near 1 (Z = -4 + (8 * 1) = 4). Therefore, for this subset of applications, with these parameter values we get a clear and pronounced version of the standard s-shaped relationship between R_3 and the predicted probabilities of denial, approximately centered at R_3 equal to 0.5. For all other applications where the proxy values for the other races are not all near 0, the (8 * R_3) term will still impact Z and provide the s-shaped curve, but now the term (0.1 * R_1) + (3 * R_2) + (5 * R_4) + (0.1 * R_5) will also impact Z, which provides vertical variation to the s-shaped curve. There are two primary drivers of the amount of vertical variation, the four parameter values and the values and distributions of the four race proxies. We chose the specific parameter values to be able to illustrate a variety of effects of the other race proxies on the marginal effects estimates for R_3. With the parameter values we chose, R_4 will have the biggest impact on the vertical variation, since R_4 has the largest parameter value (= 5). Similarly, R_1 and R_5 will have the smallest impact on the vertical variation, since they both have the smallest parameter values (= 0.1). Different choices of parameter values will lead to different effects. In addition to the parameter values, the values and distributions of the other four race proxy variables will also drive vertical variation. Given that all parameter values in equation (12) are positive, higher values of each race proxy variable will increase the vertical variation as well.

The total, net impacts of each of the other race proxy variables and their parameters will depend on the distributions of the proxies, similar to those shown in Graphs 1 and 2.

As the final step of the DGP, we generate a 0/1 indicator variable showing whether each application was denied credit. Using the values of Z from equation (12) we calculate F(Z) for each application and then generate the 0/1 denied flag using random draws from a binomial distribution with mean F(Z). At this point, we have a sample of data with six race proxy variables, as well as a 0/1 denied flag reflecting the relationships between the proxy variables and underwriting decisions from our DGP, so the simulated dataset for the analysis is complete. Although this is a simulated dataset, for the statistical analysis that follows, it can be viewed just like any other dataset we might have for any fair lending analysis, especially since our objective is to illustrate analytical concepts and not formally test for discrimination.

Table 1 presents summary statistics of the race proxy variables, Z, F(Z), and the 0/1 denied flag. As a reminder, for real-world applications, Z and F(Z) would be unknown. Overall, the summary statistics reflect the type of data found during typical fair lending analyses. Table 2 presents the results from a logistic regression of the probability of denial on the race proxies. In general, the coefficient estimates on each of the race proxy variables in the model are generally similar to the corresponding parameter values in Z (equation 12) for the DGP. As for the

Table 1: Summary Statistics of Simulated Data

Variable	Obs	Mean	Std. dev.	Min	Max
race1	55,755	.0050014	.0263882	0	.9992888
race2	55,755	.0605451	.2087537	0	1
race3	55,755	.2518156	.3408553	0	1
race4	55,755	.1329268	.2956868	0	1
race5	55,755	.0190916	.028378	0	.5991936
race6	55,755	.5306194	.3925479	0	1
z	55,755	-1.135415	2.665014	-4.005054	4.026016
FofZ	55,755	.3538696	.3702957	.0178972	.9824676
denied	55,755	.3538696	.4781736	0	1

Table 2: Logistic Regression of the P(Denial) Using Simulated Data

```
Logistic regression                                    Number of obs =   55,755
                                                       LR chi2(5)    = 38985.49
                                                       Prob > chi2   =   0.0000
Log likelihood = -16737.367                            Pseudo R2     =   0.5380

-----------------------------------------------------------------------------
denied | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------+---------------------------------------------------------------------
race1  |   .0045017   .8214796     0.01   0.996    -1.605569    1.614572
race2  |    2.93255   .0596325    49.18   0.000     2.815672    3.049427
race3  |    8.03107   .0695548   115.46   0.000     7.894745    8.167395
race4  |   5.005458   .0495973   100.92   0.000     4.908249    5.102666
race5  |    .228164   .5373053     0.42   0.671    -.8249352    1.281263
_cons  |  -4.005054   .0383634  -104.40   0.000    -4.080245   -3.929863
----------------------------------------------------------------------------
```

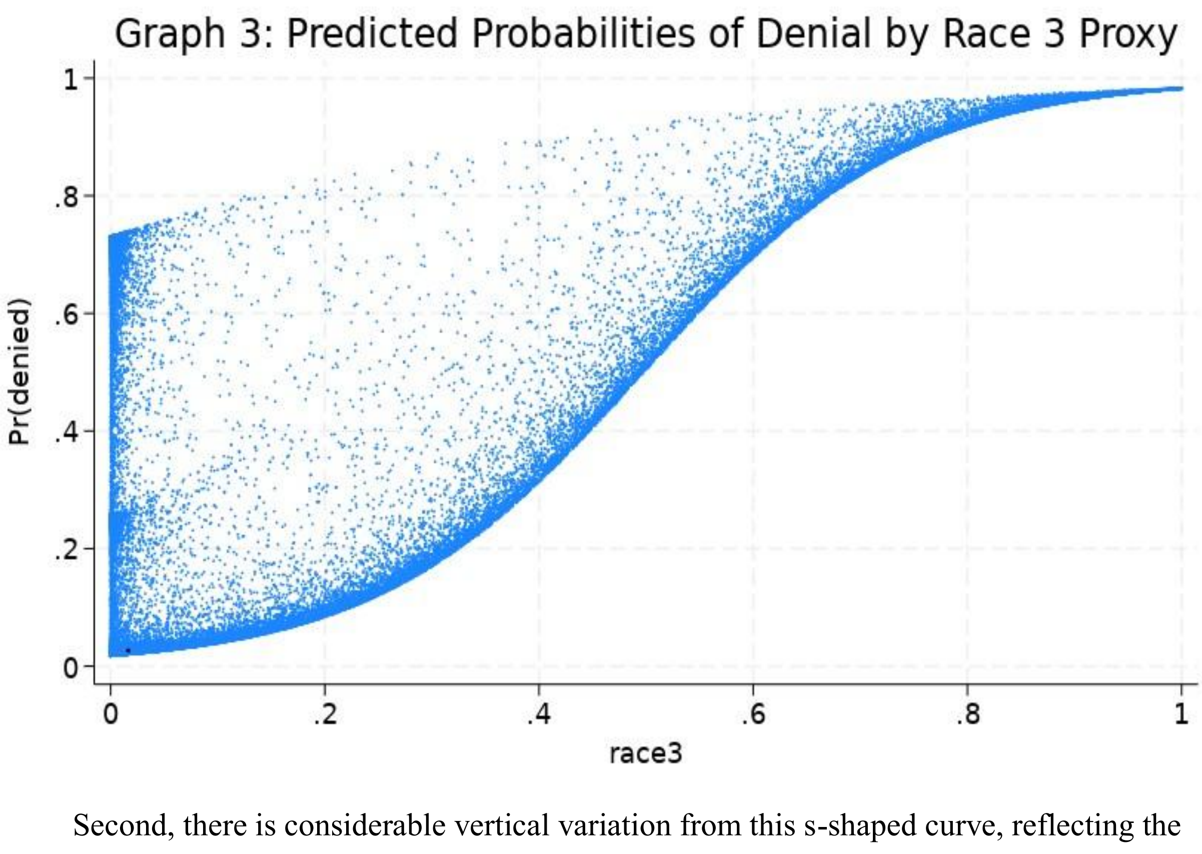
marginal effects estimates, the AME for races 1 through 5 are, 0.0004, 0.2684, 0.7351, 0.4581, and 0.0209, respectively. The AME conveys how a small change in a race proxy variable impacts the predicted probability of denial on average. For example, if R_3 changes by 0.001 (for example, from 10.0% to 10.1%), the predicted probability of denial would change by 0.0007351 (for example from 30.0% to 30.07351%) on average. Although these marginal effects estimates may appear small, they quickly become larger and more meaningful if you consider larger changes to the proxy values.

IV. Empirical Analysis and Results

This section presents empirical examples of the concepts from the Theoretical Framework section above. In addition to illustrating many of the general theoretical concepts, we also address two specific questions: 1) how do the proxy variables for other races impact marginal effects estimates for a given race and 2) how do non-proxy variables impact marginal effects estimates for race?

Question 1: How do the proxy variables for other races impact marginal effects estimates for a given race?

Using the logistic regression from Table 2, we generate, for each application, predicted probabilities of denial, as well as estimated marginal effects for each race. Using these results, we then generate a series of graphs illustrating the characteristics and patterns from the Theoretical Framework section above. Focusing first on race 3, Graph 3 presents a scatterplot of the predicted probabilities of denial and R_3 for each application in our sample. There are four items of note in Graph 3. First, the lower, right edge of the graph shows the s-shaped relationship between R_3 and the predicted probabilities of denial that is standard for logistic regressions. Again, we were able to generate this clear s-shaped curve in large part, because of our choices of parameter values in the DGP.

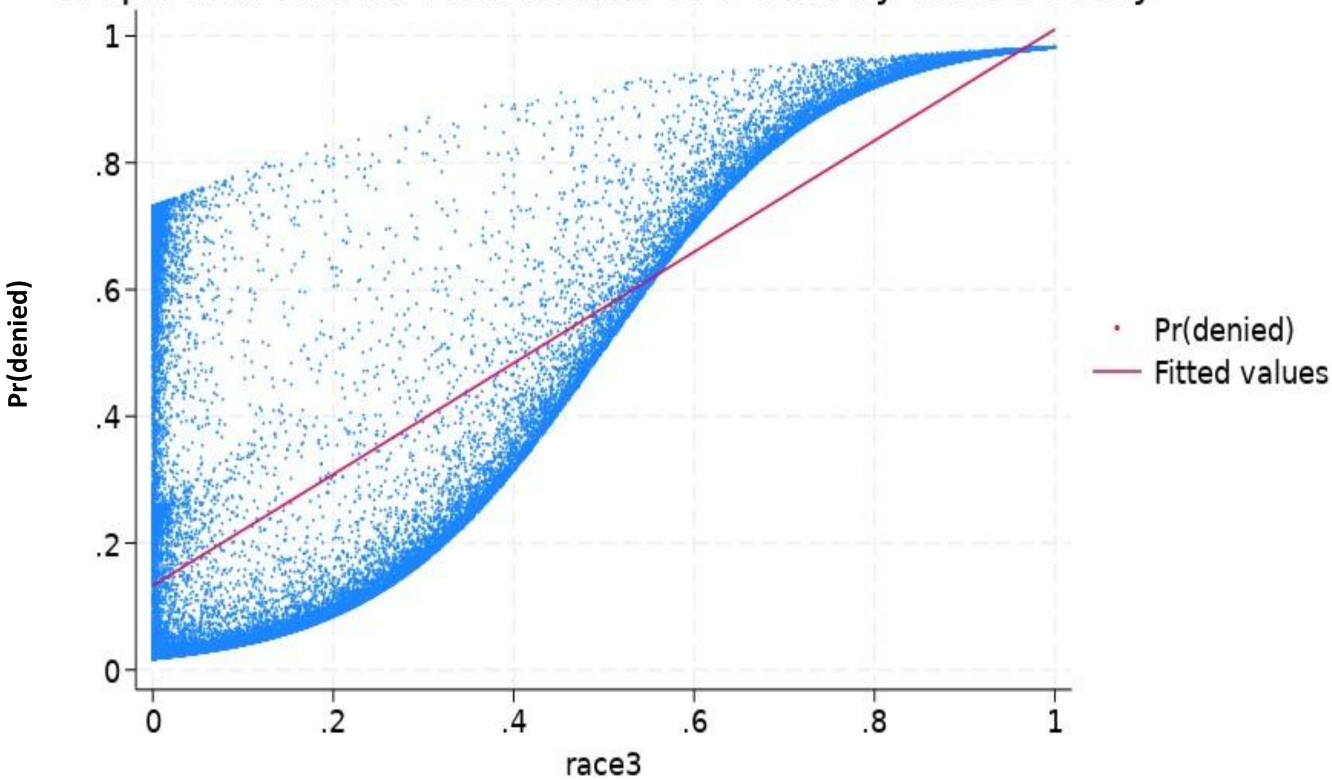


Second, there is considerable vertical variation from this s-shaped curve, reflecting the impacts of the other race proxy variables on the predicted probabilities of denial. To illustrate the drivers of this variation, we focus on one point on the x-axis, specifically where R_3 equals 20% (0.2). For all applications with a value of 0.2 for R_3 and a value of 0.8 for R_6, the predicted probability of denial will be approximately 0.15, which is on the s-shaped curve. For most applications with a value of 0.2 for R_3, the value for R_6 will be less than 0.8. Since the set of proxies sums to 1 for each application, this means that the values for at least one of the other race proxies must be non-zero for these applications. Higher values for each of these other race proxies will impact the predicted probability of denial by an amount that depends on the value of the race proxy and the estimated coefficient for that race proxy from the logistic regression in Table 2. Specifically, the vertical variation for each given value of R_3 will be a function of (0.0045017 * R_1) + (2.93255 * R_2) + (5.005458 * R_4) + (0.228164 * R_5). If this additional term is positive (which it is in our example since all terms are positive), the predicted probabilities of denial will increase (as shown in Graph 3), and if this term is negative, the predicted probabilities of denial will decrease. We should note that, since R_3 is fixed at 0.2 in this example, a higher value for one of the other race proxies implicitly means that R_6 is lower by the same amount. Since R_6 is the omitted race in the logistic regression, this offsetting change does not directly affect the calculation of the predicted probabilities of denial via Z or F(Z).

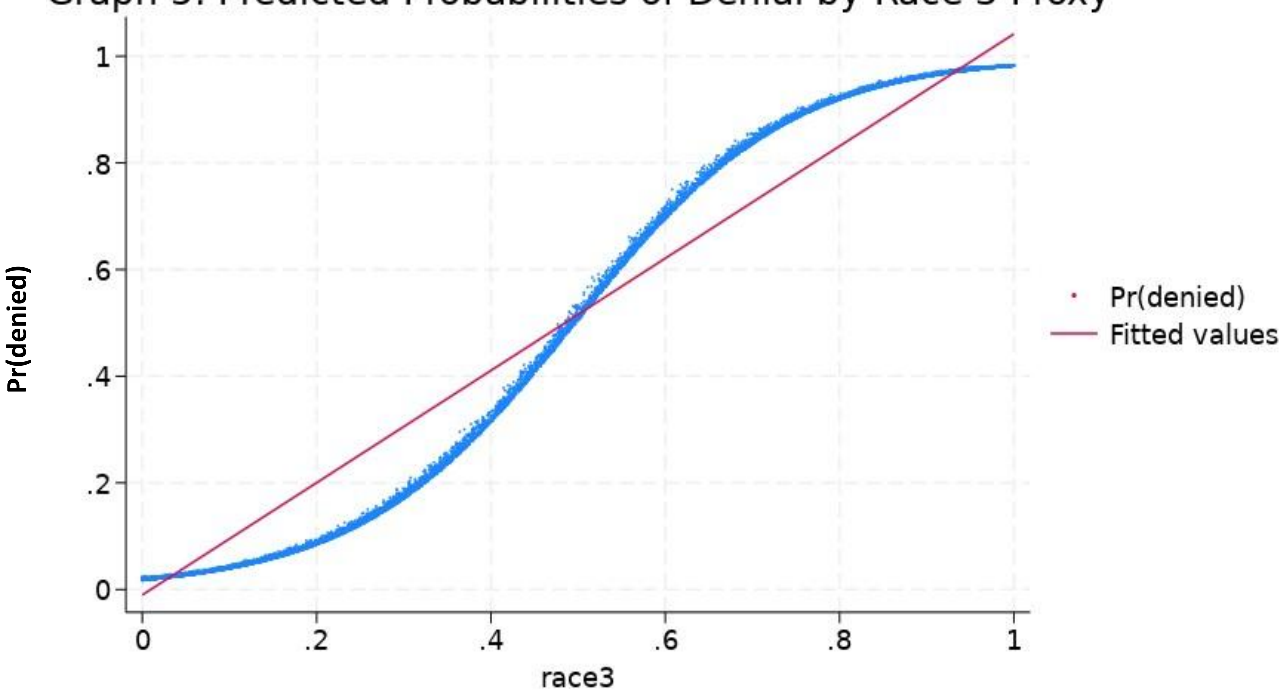
A third item of note in Graph 3 is the mass of applications on the far left of the graph where R_3 is near 0. Following the logic from the previous point, the vertical variation in the predicted probabilities of denial in this part of the graph is driven by the values for the other four race proxies, as well as the coefficient estimates on these race proxies in the logistic regression. The reason there is such a distinct mass of predicted probabilities of denial just at small values of R_3 and not at slightly larger values is because of the proxy's bimodal distribution with a significant mass near 0, a type of distribution that is common for BISG proxies.

Finally, and most importantly, the values of the other race proxy variables, and the coefficient estimates on these proxy variables, impact the overall estimated relationship between R_3 and the predicted probabilities of denial, and the estimated marginal effects for R_3. Graph 4 reproduces Graph 3, but with a fitted regression line to illustrate the overall general relationship between R_3 and the predicted probabilities of denial. Graph 5 presents a similar graph, but only using applications where R_3 and R_6 sum to greater than 0.95. Graph 5 minimizes the impact of the other race proxies on the predicted probabilities of denials, isolating the results to a comparison primarily of race 3 to race 6. As discussed above, Graph 5 just shows the standard s-shaped curve from a logistic regression. Comparing Graphs 4 and 5 shows that the values of the other race proxies and their coefficient estimates from the logistic regression rotate the fitted line clockwise, thereby altering the overall general estimated relationship between R_3 and the predicted probabilities of denial. The other race proxies also impact the aggregate estimated marginal effect for R_3. As noted above, the AME for R_3 using all applications is 0.7351. Using just the applications in Graph 5, the AME is 0.4362. The direction of this difference is somewhat surprising given the fitted lines from Graphs 4 and 5. However, as shown in the Theoretical Framework section above, the application-specific marginal effects estimates are at a maximum when the predicted probability of denial is 0.5. Therefore, whether the AME based on the applications used for Graph 5 is higher or lower than the AME using all applications depends on whether the applications excluded from Graph 5 have predicted probabilities of denial near 0.5. In our example here, the applications excluded in Graph 5 have predicted probabilities of denial disproportionately near 0.5. Overall, the main takeaway for this final point is that the other race proxy variables and their coefficient estimates impact the estimated relationship between R_3 and the predicted probabilities of denial, as well as the estimated marginal effects for R_3.

Graph 4: Predicted Probabilities of Denial by Race 3 Proxy



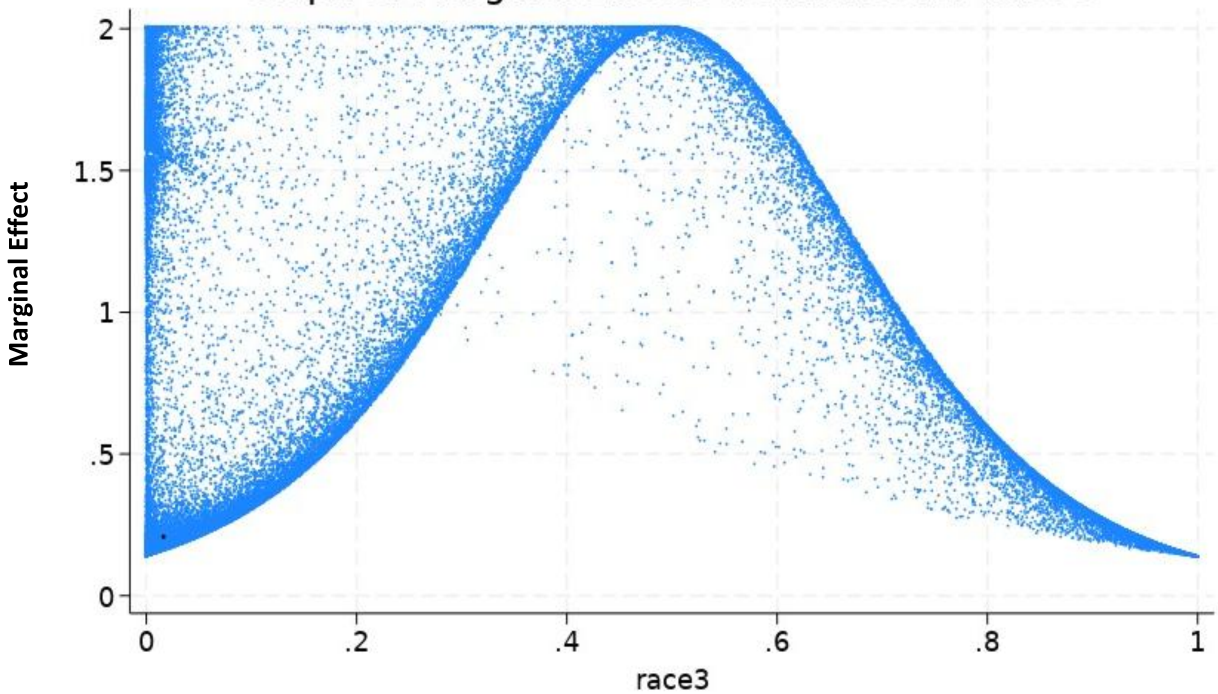
Graph 5: Predicted Probabilities of Denial by Race 3 Proxy



We now pivot to the application-specific marginal effects estimates using equation (7). Graph 6 presents a scatterplot of R_3 and the application-specific estimated marginal effects for R_3. The AME is just the average of these application-specific marginal effects estimates, so understanding the patterns in this scatterplot helps to understand the AME disparity measure. There are three results of note in Graph 6. First, following the s-shaped curve in Graph 3, there is a corresponding distinct normal-looking curve to the marginal effects estimates. The maximum of this normal-looking curve is the inflection point of the s-shaped curve in Graph 3. For the applications that make up this normal-looking curve, the other race proxies have minimal impact on the estimated marginal effects. Second, the estimated marginal effects off of this normal-looking curve are the applications from Graph 3 that deviate vertically from the distinct s-shaped curve. For these applications, the other race proxies impact the estimated marginal effects. Finally, following the theoretical result that the maximum possible application-specific estimated marginal effect is bounded by the coefficient estimate divided by 4, we see a cap in Graph 6 at 2.01 (= 8.03107 / 4). Per equation (10), this maximum occurs for applications with predicted probabilities of denial equal to 0.5.

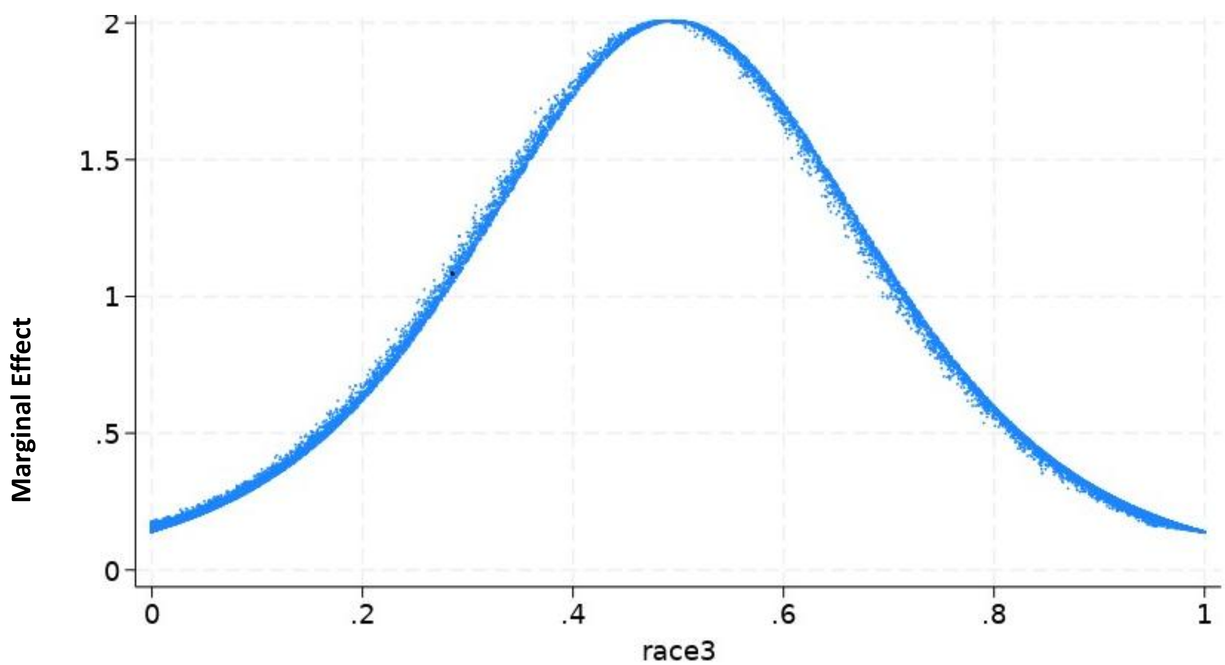
To further illustrate the impact of the other race proxies on the estimated marginal effects, we generate Graph 7, which uses only applications where R_3 and R_6 sum to close to 1 and shows only the distinct normal-shaped curve portion of Graph 6. Since the AME for R_3 is the average of the application-specific estimated marginal effects for all applications, the applications that are included in Graph 6 but not Graph 7 drive the impact of the other race proxies on the AME. For fair lending analyses, it is therefore important to analyze and understand the volume and characteristics of these applications.

Graph 6: Marginal Effects Estimates for Race 3

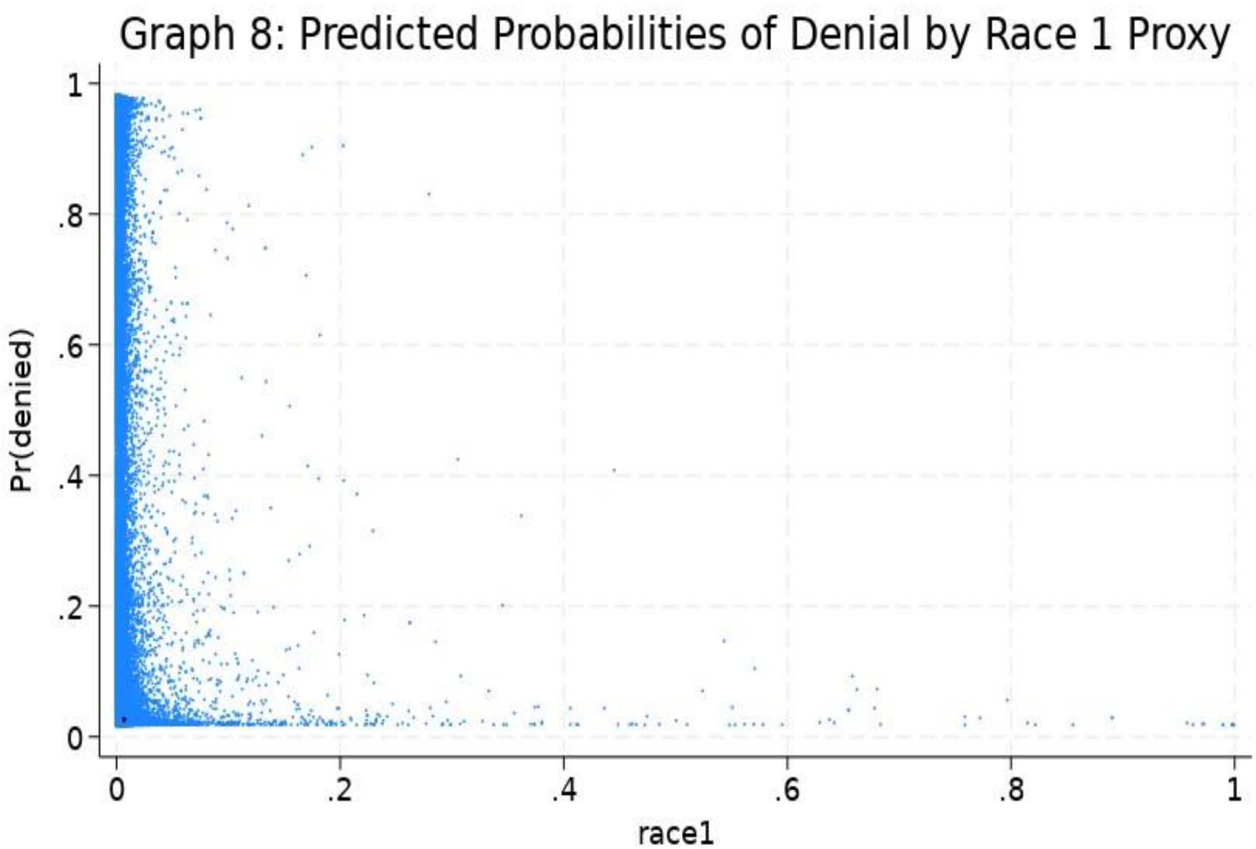


Graph 7: Marginal Effects Estimates for Race 3

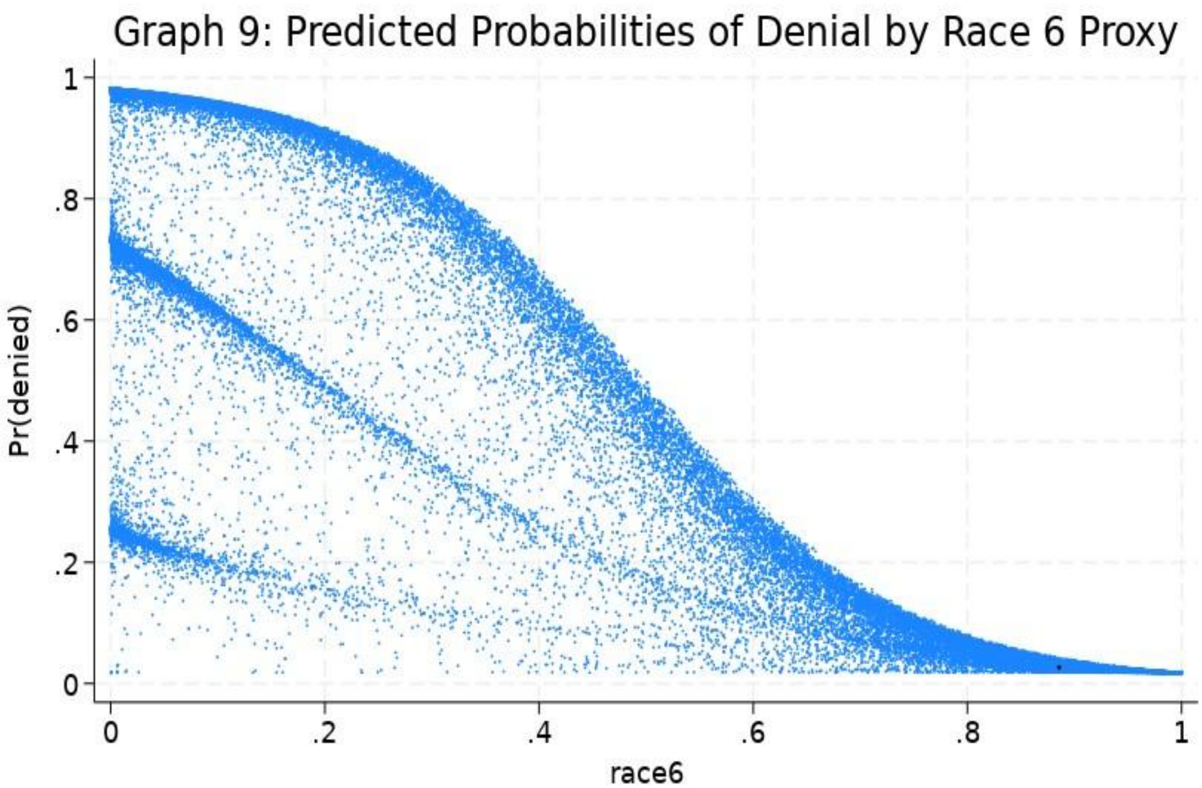
Using Only Applications Where $R_3 + R_6 \sim= 1$



To this point, we have focused on the results for R_3. Per our DGP, R_3 has a large and significant impact on the probability of denial, which allowed for a clear and pronounced illustration of the characteristics and patterns identified in the Theoretical Framework section. It is also instructive to briefly explore racial groups, such as R_1, which have little effect on the probability of denial per the DGP. Graph 8 presents a scatterplot of the predicted probabilities of denial and R_1. As expected given the DGP and the logistic regression results, the graph shows little relationship between the two variables. The only characteristic from the Theoretical Framework section that is present is the distinct mass of data points where R_1 nears 0, similar to what occurred for R_3.



As a final exercise, we explore patterns between the predicted probabilities of denial and R_6 (the omitted race). Graph 9 presents these results. The first item of note here is that the slope of the data is now downward, unlike the graphs for the other race proxies above. Given that R_6 is the omitted race in the logistic regression in Table 2, the coefficient estimates for the included race proxies are all relative to race 6. All of the coefficient estimates for the included race proxies are positive, suggesting that relative to race 6, higher values for the included race proxies increase the probability of denial. This suggests that relative to the included race proxies, higher values for R_6 reduce the probability of denial, consistent with the downward sloping data in Graph 9.



Continuing the findings from above, Graph 9 shows clear relationships for different combinations of race proxy values. The top-most, inverted s-shaped curve reflects applications where R_3, which has the largest coefficient estimate in the logistic regression, and R_6 sum to near 1. The middle, slightly less clear mass of data points reflects applications where the proxy for R_4, which has the second largest coefficient estimate, makes up the majority of the non-race 6 proxy percentage. Finally, the faint mass of points near the bottom reflects applications where the proxies for races 1, 2, and 5 make up the majority of the non-race 6 proxy percentage. All of the predicted probabilities of denial between these three distinct curves are applications with a mix of percentages across several of the race proxy variables. The results in Graph 9 are of interest, because they show that different subsets of applications can have different and distinct impacts on the predicted probability of denial. Again, being aware of these types of patterns helps understand the predicted probabilities of denial and estimated marginal effects for race during fair lending analyses.

<u>Question 2: How do non-proxy variables impact marginal effects estimates for race?</u>

In addition to being a function of all race proxy variables and their coefficients, the predicted probabilities of denial, as well as the marginal effects estimates for R_i, will also be a function of all other variables in the model and their coefficient estimates. This is important for fair lending analyses, because regression models of the probability that a lender denies applications for credit also include the factors the lender considered when making underwriting decisions. As a result, these policy factors will drive a portion of the predicted probability of denial and the marginal effects estimates for R_i.

To explore these potential indirect effects, we extend the DGP from Section III by including one additional variable (FICO score). Specifically, we now define Z as,

$$Z = \beta_0 + (0.01 * R_1) + (0.01 * R_2) + (\beta_3 * R_3) + (0.01 * R_4) + (0.01 * R_5) - \theta * FICO \qquad (14)$$

where R_1 through R_5 are the same race proxy variables for the same sample of applications used to generate the original simulated data. Our general analytical approach here is to generate Z for a variety of combinations of β_3, θ, and correlations between FICO score and R_3, and then estimate the indirect impact FICO score has on the marginal effects estimates for R_3 under all different combinations. To simplify this analysis, we set the parameter values for all other race proxy variables in equation (14) to a small number to minimize their impact on the estimates of the indirect impacts caused by FICO score. Relaxing this assumption will affect the estimates of these impacts, but significantly complicate the analysis, so we leave this to others to explore.

Overall, the analysis consists of 13 steps,

Step 1: Set β_3 to 1

Step 2: Generate FICO scores using random draws from a normal distribution with mean 750 and standard deviation 40 for applications with a value of 0.5 or less for R_3 and random draws from a normal distribution with mean 660 and standard deviation 40 for applications with a value of greater than 0.5 for R_3. These means and variances generally reflect the FICO score distributions found during fair lending analyses.

Step 3: Set θ to 0.01. To center the effect of FICO score, we set β_0 equal to θ times 735, which is the mid-point between FICO scores of 620 and 850.

Step 4: Generate Z for each application in the entire sample using equation (14).

Step 5: Using the race proxy variables, Z from step 4, and F(Z) from equation (5), generate a 0/1 denied flag using random draws from a binomial distribution with mean F(Z).

Step 6: Generate the actual correlation coefficient between R_3 and the simulated FICO score from step 2.

Step 7: Estimate a logistic regression of the probability of denial as a function of five race proxy variables (R_1- R_5) and FICO score.

Step 8: Calculate marginal effects estimates for R_3 for each application using equation (7). These estimates represent the actual marginal effects estimates for R_3, including any indirect impact from FICO score.

Step 9: Generate marginal effects estimates for R_3 for each application excluding the indirect impact from FICO score. We do this by setting FICO score to its mean for all applications and then re-calculating the marginal effects estimates for each application using equation (7).[6]

Step 10: Take the difference of the two marginal effects estimates from steps 8 and 9 for each application, and then compute the mean of these differences across all applications. This will be our estimate of the indirect impact of FICO score on the marginal effects estimates of R_3.

Step 11: Re-run steps 3 through 10, iterating over values of θ from 0.01 to 0.05 by 0.001. This step consists of 41 iterations.

Step 12: Re-run steps 2 through 11, iterating over mean FICO scores from 660 to 750 by 1 for applications with a value greater than 0.5 for R_3. This step consists of 91 iterations.

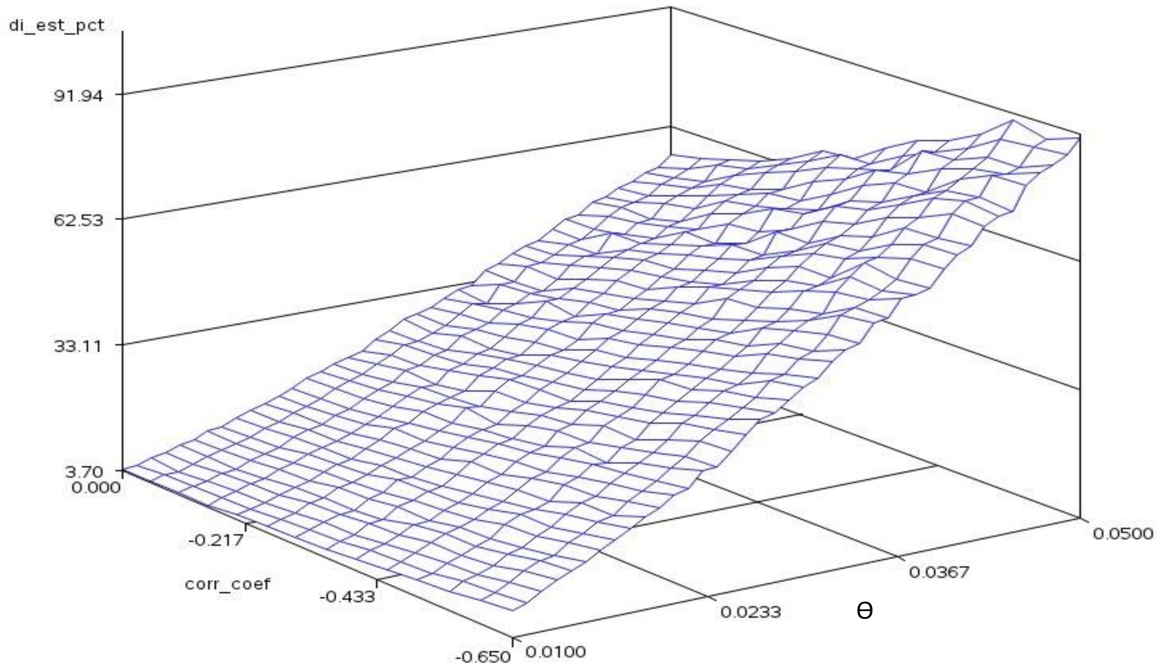
Step 13: Re-run steps 1 through 12, iterating over values of β_3 equal to 4 and 8. Including the initial β_3 value of 1 in step 1 above, this yields 3 total iterations.

Forty-one iterations from step 11 and 91 iterations from step 12 yields 3,731 total sets of results for each of the three values of β_3 in step 13.

Graph 10 presents the estimates of the indirect impact of FICO score on marginal effects estimates of R_3 when β_3 is equal to 1. The DGP with β_3 equal to 1 is the weakest relationship between R_3 and the probability of denial that we analyze here. The x-axis measures the correlation between R_3 and FICO score, running from 0 to -0.65. The y-axis measures the effect of FICO score on the probability of denial (θ), running from 0.01 to 0.05. The z- or vertical-axis measures the percentage of the estimated marginal effect of R_3 that is driven by the indirect impact from FICO score. As an example of how to interpret the indirect impact estimates, suppose the estimated impact is 10% for a given θ and a given correlation between R_3 and FICO score. This means that, for this modeling scenario, 10% of the actual AME for R_3 is due to indirect impacts from FICO score. As a frame of reference for the results in Graph 10, across the 3,731 sets of results, the actual

[6] The indirect impacts from FICO score will vary by the choice of the fixed FICO score value we use for this step. Specifically, choosing a value lower than the mean will increase the estimate and choosing a value above the mean will decrease the estimate.

Graph 10: Estimated Indirect Impact of FICO Score on the Estimated Marginal Effect of R_3
By Corr(R_3, FICO) and FICO Parameter, with β_3 = 1



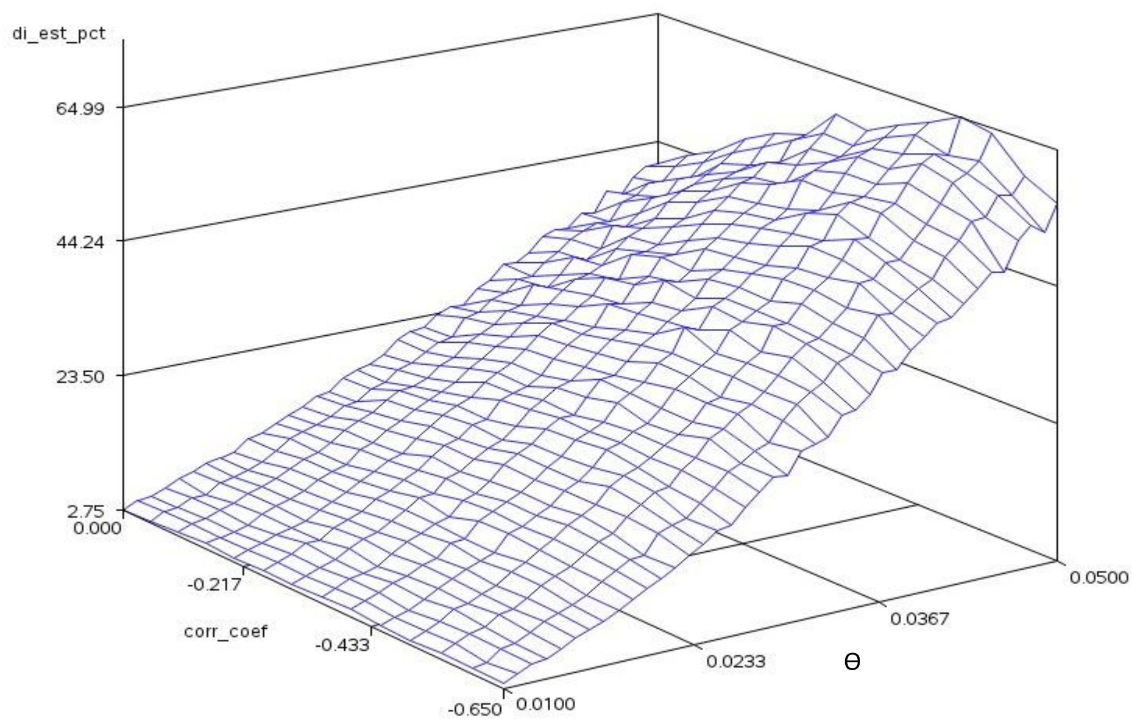
AMEs for R_3 range from 0.11 to 0.25 with an average AME of 0.18.[7] As Graph 10 shows, when R_3 and FICO score are not correlated, and FICO score has a weaker effect on the probability of denial (value of Θ near 0.01), FICO score accounts for a relatively small portion of the estimated marginal effects of R_3, approximately 3.70%. As the correlation increases, this impact increases, but only slightly. As the strength of the effect of FICO score on the probability of denial increases, the indirect impact of FICO score on the estimated marginal effects of R_3 increases much more quickly with the largest impacts overall (approximately 92%) occurring when the correlation and the FICO score parameter value are both at their highest, -0.65 and 0.05, respectively. It makes sense that the indirect impact of FICO score on the estimated marginal effects of R_3 is highest

[7] As a reminder, the AME conveys how a small change in a race proxy variable impacts the predicted probability of denial on average. For example, if the AME for R_3 is 0.7351, if R_3 changes by 0.001 (for example, from 10.0% to 10.1%), the predicted probability of denial would change by 0.0007351 (for example from 30.0% to 30.07351%) on average.

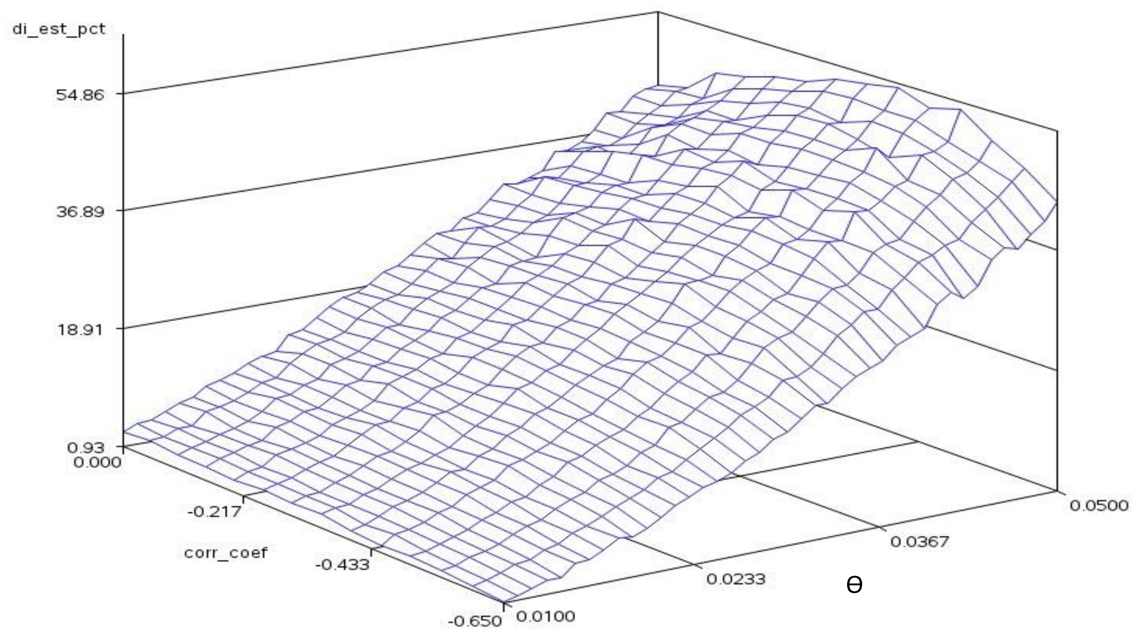
when the correlation between FICO score and R_3 is highest and the impact of FICO score on the probability of denial is strongest.

Graphs 11 and 12 present the results for β_3 values of 4 and 8, respectively. As a frame of reference for the results in Graph 11, across the 3,731 sets of results, the actual AMEs range from 0.44 to 0.76 with an average AME of 0.61. For Graph 12, the actual AMEs range from 0.83 to 0.133 with an average AME of 1.08. In general, the patterns in Graphs 11 and 12 are similar to those in Graph 10, but with smaller estimated indirect impacts for higher values of β_3. This makes sense since higher values of β_3 suggest a stronger relationship between R_3 and the probability of denial, which would suggest it would be more difficult for other factors in the model to impact the estimated marginal effects of R_3. One additional interesting item of note in

Graph 11: Estimated Indirect Impact of FICO Score on the Estimated Marginal Effect of R_3
By Corr(R_3, FICO) and FICO Parameter, with β_3 = 4



Graph 12: Estimated Indirect Impact of FICO Score on the Estimated Marginal Effect of R_3
By Corr(R_3, FICO) and FICO Parameter, with β_3 = 8



Graphs 10 – 12 is the clockwise rotation of the plane as the values of β_3 increase. Since the objective here is just to show the types of indirect impacts that might occur, we leave this specific result for others to explore.

V. Potential Best Practices

In this section we use the concepts from the theoretical framework, together with the results from the empirical examples based on simulated data, to develop a list of potential best practices for real-world fair lending analyses focused on using a logistic estimator to generate marginal effects estimates for continuous race proxies in a limited dependent variable model.

1. Characterize Race Proxy Distributions

   Using graphs similar to Graphs 1 and 2, or other appropriate statistical approaches, develop an understanding of the distributions of race proxy values across applications.

These analyses will provide insight into the potential impacts the other race proxy variables might have on estimated marginal effects for the race of interest.

2. Generate Scatterplots

   Generate scatterplots of each race proxy variable against both the predicted probabilities of denial and the application-specific marginal effects estimates. These scatterplots will provide insight into the patterns of the underlying application-specific estimates that are used to generate aggregate disparity measures.

3. Calculate Bounds on Marginal Effects

   Use equation (12) to bound the application-specific marginal effects estimates. Again, this will inform the construction and interpretation of the aggregate marginal effects estimates.

4. Estimate Indirect Impacts of Non-proxy Variables

   For each non-proxy variable in the regression model, separately estimate the amount of the marginal effect estimate for each race proxy that is driven by indirect impacts from that variable.

5. Generate a Hybrid Disparity Measure

   Construct a hybrid version of the AME and MEM where application-specific marginal effects estimates are generated at actual values for all other race proxy variables and at mean values for all non-proxy control variables. This hybrid measure does not violate the property of BISG proxies that the set of proxy values for each application must sum to 1, and it excludes all indirect impacts from the other control variables.

6. Filtered Disparity Measure

   Generate AME, MEM, and hybrid disparity measures using only applications where the sum of the race proxy variable of interest and the omitted race proxy variable sum to 0.95 or higher. This check excludes the impacts of the other race proxy variables.

VI. Conclusion

A common approach to testing for discrimination in underwriting decisions when application-level race data are unavailable is to include continuous BISG proxy variables in a limited dependent variable model, estimate the model using a logistic estimator, and then generate marginal effects estimates of the impact each race proxy has on the probability of denial. It can be difficult to determine whether results from this approach indicate evidence of discrimination since the marginal effects estimates will be application-specific and will be a function of every variable in the regression model as well as the coefficient estimates for those variables. Therefore, it is important to understand the underlying components and patterns of the marginal effects estimates, as well as the impacts that the other variables in the regression model have on these estimates.

In this paper we developed the theoretical characteristics of marginal effects estimates for this specific use case. We then used simulated data to construct empirical examples illustrating these theoretical characteristics. Throughout, our main objective was to highlight the analytical challenges that arise when conducting this type of analysis and interpreting the results, as well as the importance of fully understanding the underlying drivers of aggregate marginal effects estimates of race. To that end, we ended by providing a set of potential best practices for how to conduct additional exploratory analyses of the underlying drivers of marginal effects estimates. These best practices will aid regulators and industry when conducting analyses of underwriting decisions for discrimination.

All of the theoretical and empirical results in this paper focused on one specific approach to testing for discrimination in underwriting decisions. Since alternative estimators and alternative approaches to incorporating race proxies into fair lending analyses are available, the results and best practices presented here will also be useful when considering the tradeoffs of alternative analytical approaches and which approach to use for a given analysis.

References

Berkson, Joseph. 1944. "Application of the Logistic Function to Bio-Assay," Journal of the American Statistical Association, Volume 39, No. 227, September 1944, pp. 357-65.

Greenwald, Daniel, Sabrina Howell, Cangyuan Li, and Emmanuel Yimfor. 2023. "Regulatory Arbitrage or Random Errors? Implications of Race Prediction Algorithms in Fair Lending Analysis," SSRN Working Paper, April 12, 2023, available at http://dx.doi.org/10.2139/ssrn.4417513

Voicu, Ioan. 2018. "Using First Name Information to Improve Race and Ethnicity Classification," Statistics and Public Policy, Volume 5, Issue 1, pp. 1-13.

Zhang, Yan. 2016. "Assessing Fair Lending Risks Using Race/Ethnicity Proxies," Management Science, Volume 64, Issue 1, January 2016.