Confounded? This example shows how to use SAS[®] chi-square tests, correlations and logistic regression to unconfound a result.

MIKE GRIERSON IS A FEDERAL EMPLOYEE CURRENTLY WORKING ON SAS, DATA WAREHOUSE, AND PROCESS AUTOMATION ISSUES. HE RECEIVED A BSEE DEGREE FROM THE UNIVERSITY OF CENTRAL FLORIDA AND AN MSSE DEGREE (WITH AN OPERATIONS RESEARCH CONCENTRATION) FROM THE GEORGE WASHINGTON UNIVERSITY. AFTER WORKING FOR MANY YEARS ON INFORMATION TECHNOLOGY SYSTEMS, THE OPERATIONS RESEARCH STUDY PROVIDED INSPIRATION FOR HIS FEDERAL SERVICE WORK AS AN OPERATIONS RESEARCHER AND BUSINESS ANALYST (FOR THE DEPARTMENT OF THE ARMY, THE COAST GUARD, THE DEPARTMENT OF VETERANS AFFAIRS AND NOW THE DEPARTMENT OF EDUCATION). HE CURRENTLY LIVES IN THE WASHINGTON DC AREA AND CAN BE CONTACTED AT <u>MCG@MCG-</u> <u>CT.COM</u>.

THIS PRESENTATION IS MIKE'S OWN WORK AND DOES NOT REPRESENT OPINIONS/CONCLUSIONS OF HIS EMPLOYER.



1

Overview

Confounding > Example article

Methodology:

Get more data (that Includes suspected confounding attributes) and rank that data

> Apply Chi Square tests, Correlations, and Logistic Regression

Why this matters

Thinking Fast (Intuitively/Instinctually) vs. Slow (Analytically/Rationally)

Evidence based analytic conclusions

CONFOUNDING

Center for American Progress

New Federal Data Show a Student Loan Crisis for African American Borrowers

References Department of Education data resources

Example article

12/18/2017 New Federal Data Show a Student Loan Crisis for African American Borrowers - Center for American Progress
Center for American Progress

EDUCATION, POSTSECONDARY

New Federal Data Show a Student Loan Crisis for African American Borrowers

By Ben Miller | Posted on October 16, 2017, 9:00 am



https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018410

Two weeks ago, the U.S. Department of Education provided the first-ever look at long-term outcomes for student loan borrowers, including results by race and ethnicity.

https://www.americanprogress.org/issues/education-postsecondary/news/2017/10/16/440711/new-federal-data-show-student-loan-crisis-african-americ... 1/6

The data show that 12 years after entering college, the typical African American* student who started in the 2003-04 school year and took on debt for their undergraduate education owed more on their federal student loans than they originally borrowed. This holds true even for students who finished a bachelor's degree at a public institution. One reason they might not be paying down their loans? Nearly half of African American borrowers defaulted, including 75 percent of those who dropped out of for-profit colleges.

These results show that the U.S. Department of Education cannot ignore the interaction of race and student loans. Traditionally, the agency has not collected any data on the race of borrowers, except in irregular sample surveys conducted by its quasi-independent statistical arm. Unfortunately, not collecting this information has allowed for the disparate outcomes by race to go unnoticed.

https://www.americanprogress.org/issues/education-postsecondary/news/2017/10/16/440711/newfederal-data-show-student-loan-crisis-african-american-borrowers/ Specific data used in reference 2 is found by using NCES's PowerStat tool in table id cembhag3e.

Example follow-on articles

12/18/2017

Black Grads Struggle to Repay Student Loans, New Data Show | Money

Money

There's a Massive Racial Gap in Student Loan Defaults, New Data Show

By KAITLIN MULHERE October 17, 2017

African-American students who borrowed to earn their bachelor's degrees are four times more likely to default on their student loans than white peers.

That's one of the most striking findings in an analysis of new data on close to a decade of loan repayment rates. Many of the findings in the analysis, by the Center for American Progress, add to previous research showing student debt is far more burdensome for black borrowers than for white borrowers.

The CAP analysis, which covered students who entered college in 2003-04, found that African-Americans, on average, had made no progress paying off their debt 12 years after enrolling. In fact, their median student loan balance had actually grown, because of interest, to 113% of the amount originally borrowed.

For white borrowers, the median balance owed 12 years later was 65% of the original balance, and it was 83% for Latino borrowers.

The statistics on default rates were just as bleak. Overall, nearly half of black borrowers defaulted, more than double the rate for white borrowers. And while completing a degree is supposed to be the best protection against defaulting on your student loans—the research covered not only four-year graduates but also dropouts, associate's degree earners and those still in school—earning a bachelor's degree didn't shield minority borrowers nearly as well as it did whites. While just 6% of white borrowers who earned a bachelor's degree defaulted on their loans, 14% of Latino borrowers did so, as did 23% of black borrowers. Half of black student loan borrowers default, new federal data show





1/3

(https://www.insidehighered.com)

12/18/2017

Half of black student loan borrowers default, new federal data show

Submitted by Paul Fain on October 17, 2017 - 3:00am

Two analyses of newly released federal data on student loans reveal serious default problems for African-American borrowers.

Earlier this month the U.S. Department of Education's National Center for Education Statistics <u>published a report</u> in on patterns of student loan repayment for two groups of borrowers who first enrolled in college in 1995-1996 and in 2003-2004.

Historically the department has not collected much data on student debt that can be broken out by the race or ethnic background of borrowers. The new report, however, included tools that researchers can use to compare how various groups are faring.

Two resulting analyses found a troubling picture for black students who take out loans.

Nearly half (49 percent) of all black borrowers in the 2004 group defaulted on at least one loan within 12 years, <u>wrote Robert Kelchen</u> R, an assistant professor of higher education at Seton Hall University. That default rate was more than twice that of white students (20 percent) and more than four times the rate of Asian students (11 percent).

"The differentials are still present across sector, with more than one-third of black students defaulting across all sectors while a relatively small percentage of Asian students defaulted across all nonprofit sectors," Kelchen said, "Default rates at for-profit colleges are high for all racial/ethnic." groups, with almost half of white students defaulting alongside nearly two-thirds of black students.

The Center for American Progress on Monday released a report [3] on the new numbers,

The federal data show that the typical black student who enrolled in 2004 and took on debt for an undergraduate education owed more on their student loans after 12 years than the amount originally borrowed, found Ben Miller, the senior director for postsecondary education at the center

ADVERTISING

1/4

NCES article and PowerStats reference (cembhag3e)



ELES NATIONAL CENTER FOR EDUCATION STATISTICS

Repayment of Student Loans as of 2015 Among 1995–96 and 2003–04 First-Time Beginning Students

First Look



This reference does not mention race. It does mention <u>median</u> income and profit status of college attended.

National Center for Education Statistics PowerStats

Ever defaulted on a federal loan - 12 years by Race/ethnicity, First institution sector (level and control) 2003-04, Cumulative persistence and attainment anywhere 2008-09, Ever received a Pell Grant - 12 years and Income group in 2003-04, for Cumulative amount of undergraduate federal loans - 12 years ($X \ge 1$).

	No (%)	Yes (%)	Total
Estimates			
Total	71.2	28.8	100%
Race/ethnicity			
White	78.5	21.5	100%
Black or African American	50.9	49.1	100%
Hispanic or Latino	63.9	36.1	100%
Asian	87.7	12.3	100%
American Indian or Alaska Native	58.8	41.2	100%
Native Hawaiian / other Pacific Islander	84.5	15.5 !!	100%
Other	74.7	25.3	100%
More than one race	60.3	39.7	100%

PowerStats data is accurately reported.

National Center for Education Statistics PowerStats

Ever defaulted on a federal loan - 12 years by Race/ethnicity, First institution sector (level and control) 2003-04, Cumulative persistence and attainment anywhere 2008-09, Ever received a Pell Grant - 12 years and Income group in 2003-04, for Cumulative amount of undergraduate federal loans - 12 years ($X \ge 1$).

		\sim	
	No (%)	Yes (%)	Total
Estimates	(14)		
Total	71.2	28.8	100%
Race/ethnicity			
White	78.5	21.5	100%
Black or African American	50.9	49.1	100%
Hispanic or Latino	63.9	36.1	100%
Asian	87.7	12.3	100%
American Indian or Alaska Native	58.8	41.2	100%
Native Hawaiian / other Pacific slander	84.5	15.5 !!	100%
Other	74.7	25.3	100%
More than one race	60.3	39.7	100%
First institution sector (level and control) 2	003-04		
Public 4-year	81.0	19.0	100%
Private not-for-profit 4-year	81.0	19.0	100%
Public 2-year	73.6	26.4	100%
Private For Profit	46.5	53.5	100%
Other	68.0	32.0	100%
Cumulative persistence and attainment an	ywhere 2008-09		
Attained bachelor's degree	91.2	8.8	100%
Attained associate's degree	78.5	21.5	100%
Attained certificate	53.7	46.3	100%
No degree, still enrolled	70.7	29.3	100%
No degree, left without return	54.2	45.8	100%
Ever received a Pell Grant - 12 years			
No	87.2	12.8	100%
Yes	64.5	35.5	100%
ncome group in 2003-04			
Low	57.5	42.5	100%
Low middle	70.7	29.3	100%
High middle	81.0	19.0	100%
High	86.9	13.1	100%

Interpret data with caution. Estimate is unstable because the standard error represents more than 50 percent of the estimate.

The names of the variables used in this table are: FSECTOR, PROUT6, S15FEDCUM1_12Y, RACE, S15EVRDEF_12Y, INCGRP and S15PELLEVR_12Y. The weight variable used in this table is WTB000.

Source: U.S. Department of Education, National Center for Education Statistics, 2003-04 Beginning Postsecondary Students Longitudinal Study, Second Follow-up (BPS:04/09).

NCES article and PowerStats reference (cembhag3e)

PowerStats data is accurately reported.

MWSUG2018 AA-030

Computation by NCES PowerStats on 7/30/2018.

Confounding

Center for American Progress article asserts that race and student loan defaults are associated.

"These results show that the U.S. Department of Education cannot ignore the interaction of race and student loans".

Suspect this result is confounded by median income, because the mechanism of not having money and defaulting on loans is well understood, whereas the mechanism of being a member of a racial group and defaulting on loans is not well understood.

Other examples of possible confounding?

- >You will live longer if you drink 4 to 8 cups of coffee a day? (Well off people can afford all that coffee?)
- > People with graduate degrees have 11% more brain gliomas? (Graduate degree holders live in cities?)

METHODOLOGY

Find Student Debt Data that Includes Confounders and Apply

Chi Square,

Spearman Correlations and

Logistic Regression

Relevant Data mappingstudentdebt.org

DATA ORIGINS

"This geographic analysis uses two primary datasets: credit reporting data on student debt from Experian and income data from the American Community Survey. The Experian data includes eight key student debt variables (see Figure 2 (@ reference 4)) aggregated from household-level microdata to the zip code level. The underlying household data are a snapshot of the entire U.S. population at a single point in time—in this case, the autumn of 2015."

SUMMARIZED

- At the Zip Code level
- Experian for student debt data
- American Community Survey for income data

As of 2015

https://mappingstudentdebt.org/

Ś	MAPPING S			Γ	(f			
	how borrowing for college More than 42 million Americar			lebt, making it the second	d-largest liability or	n the		
	national balance sheet. A gener	ration ago, stude	ent debt was a relative ra	arity, but for today's stud	ents and recent gra	aduates,		
	it's a central fact of economic li show how borrowing for colleg						zip	Delinqu
	way in which student debt rela	tes to economic	cinequality.				64854	Extreme
VIEW MAP	Map 1: An Introduction	WITH LAYER	Delinguency	 NEAR City or 	zip code		48843	Low
					purpus	North Bay	85743	Modera
	- Billings		The second se	the second se				
a trai			Pierre	inneapolis			04971	Somew
	oise Idaho Falls GPocatello Ca	sper in the second	A REAL PROPERTY AND A REAL	inneapolis Milwauk	Colorador Constra	Toronto-Roch	04971 78705	
	Pocatello Ca		A REAL PROPERTY AND A REAL		Detroit	Roch		Extreme
	- Pocatello Ca	Cheyenne	Pierre Omaha	Circope 50524 Delinquency	Detroit	Extremely Very Low	78705	Extreme
	Pocatello Ca	Cheyenne Denver	Pierre	Milwauk ZIPCODE 50524	Colorador Constra	Extremely Very Low Low Moderatel Somewhat	78705 29056 37871 ^{y Low}	Somew Extreme Modera Very Hig
	Elio Salt Lake City	Cheyenne Denver	Pierre Omaha V United States	ZIPCODE 50524 Delinquency Very Low Average Loan Balance	anapolis Gincinnati	Extremely Very Low Low Moderatel	78705 29056 37871 y Low Llow	Extreme

Take a sample of 60 randomly selected zip codes and build data table below

maps below ctive on the	zip	Delinquency	dn	Average_loan_balance	albn	Median_income	Aamerican	Latino
	64854	Extremely High	10	Moderately Low	2	33333	4.2	31.5
	48843	Low	3	Slightly High	4	67477	0.7	2.6
North Bay	85743	Moderately Low	4	Slightly High	4	69577	3.7	19.9
Toronto—	04971	Somewhat Low	5	Moderately Low	2	43393	0.2	
Roch	78705	Extremely Low	1	Slightly High	4	12143	4.1	17.3
Extremely	29056	Moderately High	7	Moderately High	6	23023	80.6	0.2
Very Low	37871	Very High	9	Average	3	46565	2.4	1.3
Low	vlow 🎽							

Test the data sample

proc univariate data=work.sd plot normaltest;

var dn median_income aamerican;

run;

	Moments									
N	60	Sum Weights	60							
Mean	5.08333333	Sum Observations	305							
Std Deviation	2.75121955	Variance	7.56920904							
Skewness	0.22844224	Kurtosis	-1.1262885							
Uncorrected SS	1997	Corrected SS	446.583333							
Coeff Variation	54.1223519	Std Error Mean	0.35518092							

Variable: dn (Delinguency Category)

Variable: median_income (Median Income)

Moments								
Ν	60	Sum Weights	60					
Mean	50155.8833	Sum Observations	3009353					
Std Deviation	20258.8013	Variance	410419028					
Skewness	0.68105405	Kurtosis	0.14196054					
Uncorrected SS	1.75151E11	Corrected SS	2.42147E10					
Coeff Variation	40.3916747	Std Error Mean	2615.4					

Variable: aamerican (% African American)

Moments								
Ν	60	Sum Weights	60					
Mean	13.3633333	Sum Observations	801.8					
Std Deviation	21.1044052	Variance	445.395921					
Skewness	2.00025002	Kurtosis	3.01953587					
Uncorrected SS	36993.08	Corrected SS	26278.3593					
Coeff Variation	157.927702	Std Error Mean	2.724567					

Rank the data

/* rank into two groups for chisq and logistic regression calculations */

proc rank groups=2 data=work.sd1 out=work.sd2 ties=low;

var median income aamerican latino dn;

ranks rank_median_income2 rank_aamerican2 rank_latino2 rank_dn2 ;

run;

/* rank into ten groups for spearman correlations */

proc rank groups=10 data=work.sd2 out=work.sd3 ties=low;

var median income aamerican latino;

ranks rank median income10 rank aamerican10 rank latino10;

run;

zip	delinquency	dn	Average_loan_balance	albn	Median_income	Aamerican	Latino	rank_median_income2	rank_aamerican2	rank_latino2	rank_dn2	rank_median_income10	rank_aamerican10	rank_latino10
64854	Extremely High	10	Moderately Low	2	33333	4.2	31.5	low median income category	low % African American category	high % latino category	high student loan default category	2	4	8
48843	Low	3	Slightly High	4	67477	0.7	2.6	high median income category	low % African American category	low % latino category	low student loan default category	8	1	3
85743	Moderately Low	4	Slightly High	4	69577	3.7	19.9	high median income category	low % African American category	high % latino category	low student loan default category	8	4	7
4971	Somewhat Low	5	Moderately Low	2	43393	0.2		low median income category	low % African American category		low student loan default category	3	0	
78705	Extremely Low	1	Slightly High	4	12143	4.1	17.3	low median income category	low % African American category	high % latino category	low student loan default category	0	4	7
29056	Moderately High	7	Moderately High	6	23023	80.6	0.2	low median income category	high % African American category	low % latino category	high student loan default category	0	9	0
37871	Very High	9	Average	3	46565	2.4	1.3	low median income category	low % African American category	low % latino category	high student loan default category	4	3	1
85338	Moderately Low	4	Slightly High	4	67132	7.3	34	high median income category	high % African American category	high % latino category	low student loan default category	7	6	9
62959	Moderately High	7	Average	3	45947	6.9	2.4	low median income category	high % African American category	low % latino category	high student loan default category	4	6	2

Chi-Square

title "Chisq output for Student Loan default by African American group";
proc freq data=work.sd;

tables rank_aamerican2*rank_dn2 / cmh chisq expected norow nocol nopercent;
run;

title "Chisq output for Student Loan default group by median income group";
proc freq data=work.sd;

tables rank_median_income2*rank_dn2 / cmh chisq expected norow nocol nopercent;
run;

Chi-Square



Chi	sq output	for Student Loan defa	-		iedian ii	ncome group					
Frequency		Table of rank_dn2 by rank_median_income2									
Expected			rank_n	nedian_inc	ome2(Med	lian Income 2 Cate	gories)				
	rank_dn2	(Delinquency 2 Categories)	low m	edian inco	me hig	h median income	Total				
	low stude	ent Ioan default			7 17	27 17	34				
	high stud	lent loan default			23 13	3 13	26				
	Total		30 30			30	60				
		Statistics for Table of rank_d	n2 by ra	nk_median	_income2						
		Statistic	DF	Value	Prob						
	(Chi-Square	1	Value 27.1493	Prob <.0001						
			1								
		Chi-Square	1 re 1	27.1493	<.0001						
		Chi-Square Likelihood Ratio Chi-Squar	1 e 1	27.1493 30.0065	<.0001 <.0001						

0.5581

Contingency Coefficient

Cramer's V

These results show that the U.S. Department of Education SHOULD ignore the interaction of race and student loans because chi-square results show median income levels are far more associated with student loan defaults

Spearman Correlation

/* spearman correlations between delinquency categories, median_income and race */
proc corr data=work.sd spearman plots=scatter;

var dn;

with rank_median_income10 rank_aamerican10 rank_latino10;

run;

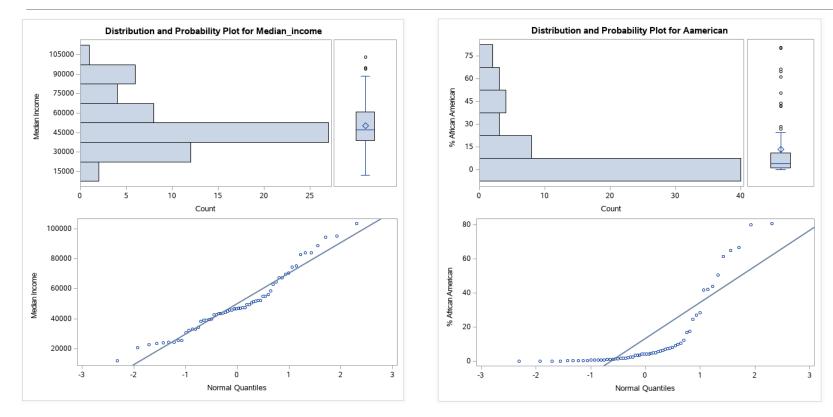
Spearman Correlation Coefficients, N = 60 Prob > r under H0: Rho=0						
	dn					
rank_median_income10	-0.69092					
Median Income 10 Categories	<.0001					
rank_aamerican10	0.24955					
% African American 10 Categories	0.0545					
rank_latino10	0.04261					
% Latino 10 Categories	0.7465					

from http://www.statstutor.ac.uk/resources/uploaded/spearmans.pdf

.00-.19 "very weak" .20-.39 "weak" .40-.59 "moderate" .60-.79 "strong" .80-1.0 "very strong"

These results show that the U.S. Department of Education SHOULD ignore the interaction of race and student loans because spearman correlation results show median income levels are far more associated with student loan defaults

Spearman Correlation



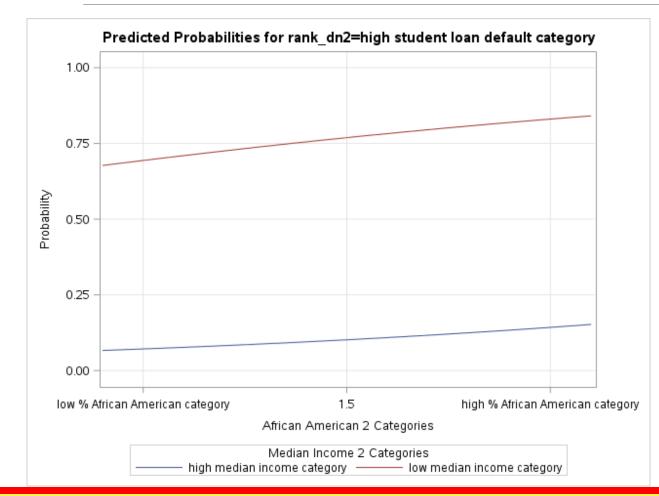
Linearity assumptions for percent African American are suspect

Percent African American data not particularly normally distributed

Logistic Regression

Use Logistic regression effects plots to get some sense of the magnitude of the relative associations (not provided by the earlier tests and not provided by Odds Ratios)

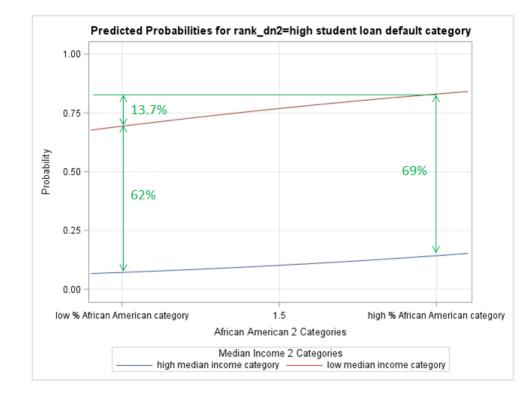
Logistic Regression



Effects plots use Maximum Likelihood Estimators (MLE) to provide probabilities for all 60 of the sample cases in the sample.

Analysis of Maximum Likelihood Estimates								
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq		
Intercept		1	0.8178	0.5357	2.3303	0.1269		
rank_aamerican2		1	0.7710	0.7178	1.1538	0.2828		
rank_median_income2	high median income category	1	-3.3815	0.7578	19.9140	<.0001		

Logistic Regression



Changes in median income category change the probability estimate of student loan default by greater than (>) 62%

Changes in African American category change the probability estimate of student loan default by less than (<) 14%

These results show that the U.S. Department of Education SHOULD ignore the interaction of race and student loans because logistic regression effects plot results show median income levels are far more associated with student loan defaults

WHY THIS MATTERS

Instinctual/Intuitive thinking vs Analytic/Rational thinking

See: Kahneman, D. (2011): Thinking, Fast and Slow Penguin, 496 pp., ISBN 978-0141033570

> Lewis, M. (2016): The Undoing Project W. W. Norton & Company, 368 pp., ISBN 978-0393254594

Consequences of incorrect analytic conclusions

Analytic conclusions

Consequences of incorrect associations

> Approximately 40 million people in poverty in the US (2016 ACS)

- > Approximately 8 or 9 million of those in poverty are African American.
- > African Americans are over represented in poverty roles (by more than 2 times), but most impoverished people are NOT African American.
- ➤A 'solution' to student loan defaults that has a focus on race (in particular African Americans) would leave out most impoverished people (perhaps 75% of the folks needing help due to poverty).



Name: Michael Grierson

E-Mail: mcg@mcg-ct.com

Web: https://www.mcg-ct.com/saspaper4/

Organization: Federal Employee

Logistic Regression MLE

 $\operatorname{logit}(p_i) = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots \hat{\beta}_k x_{ki}$

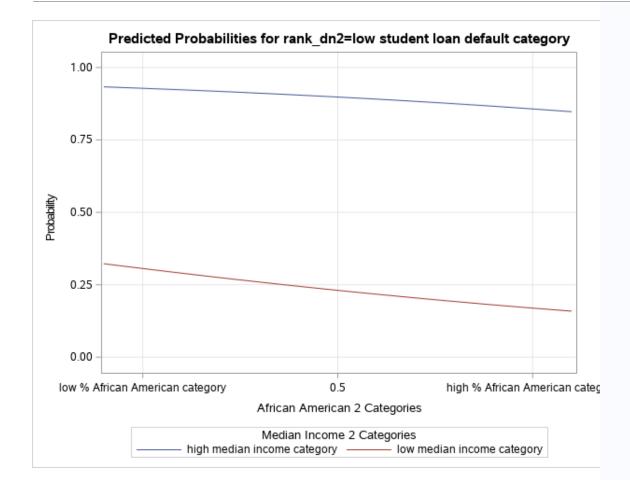
From SAS® online course "Predictive Modeling Using Logistic Regression (v14.2)"

Beta hats are MLE coefficients that provide the maximum probability (likelihood) of the equation above using all 60 data points as constraints on an optimization to maximize likelihood

Intuitive explanation of maximum likelihood estimation

Maximum likelihood estimation is a method that determines values for the parameters of a model. The parameter values are found such that they maximise the likelihood that the process described by the model produced the data that were actually observed.

From https://towardsdatascience.com/probability-concepts-explained-maximum-likelihood-estimation-c7b4342fdbb1



Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

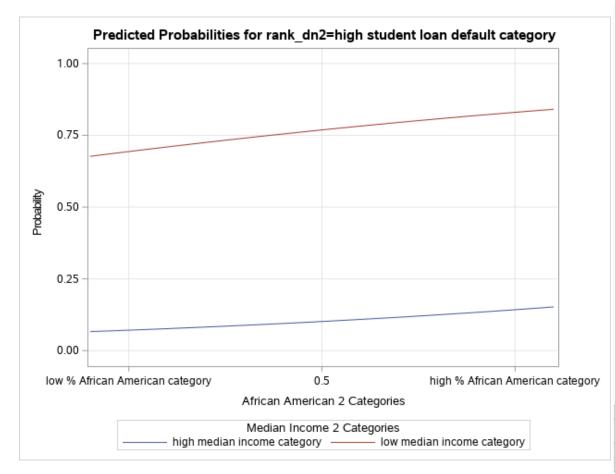
	Model Fit Statistics								
Criterion Intercept Only Intercept and Covariate									
AIC	84.108	56.923							
SC	86.202	63.206							
-2 Log L	82.108	50.923							

Testing Global Null Hypothesis: BETA=0										
Test	Chi-Square	DF	Pr > ChiSq							
Likelihood Ratio	31.1844	2	<.0001							
Score	27.8021	2	<.0001							
Wald	20.2380	2	<.0001							

Type 3 Analysis of Effects									
Effect DF Chi-Square Pr > Chi									
rank_aamerican2	1	1.1536	0.2828						
rank_median_income2	1	19.9140	<.0001						

Analysis of Maximum Likelihood Estimates									
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq			
Intercept		1	-0.8178	0.5357	2.3303	0.1269			
rank_aamerican2		1	-0.7710	0.7178	1.1536	0.2828			
rank_median_income2	high median income category	1	3.3815	0.7578	19.9140	<.0001			

Odds Ratio Estimates								
Effect	Point Estimate		6 Wald nce Limits					
rank_aamerican2	0.463	0.113	1.889					
rank_median_income2 high median income category vs low median income category	29.416	6.661	129.896					



Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics								
Criterion	Intercept Only	Intercept and Covariates						
AIC	84.108	56.923						
SC	86.202	63.206						
-2 Log L	82.108	50.923						

Testing Global Null Hypothesis: BETA=0										
Test	Chi-Square	DF	Pr > ChiSq							
Likelihood Ratio	31.1844	2	<.0001							
Score	27.8021	2	<.0001							
Wald	20.2380	2	<.0001							

Type 3 Analysis of Effects									
Effect DF Chi-Square Pr > ChiSq									
rank_aamerican2	1	1.1536	0.2828						
rank_median_income2	1	19.9140	<.0001						

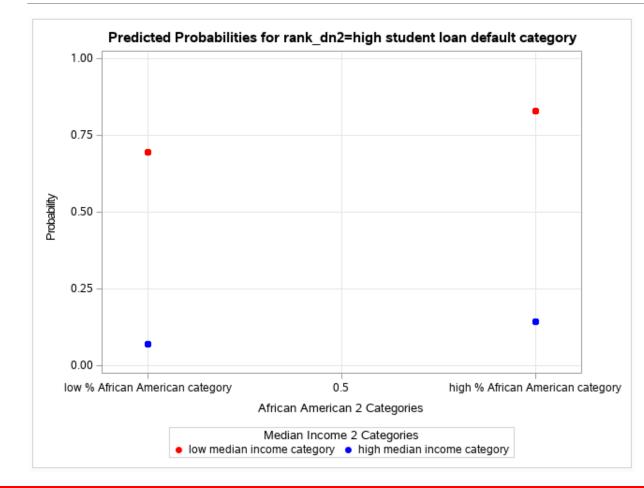
Analysis of Maximum Likelihood Estimates										
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq				
Intercept		1	0.8178	0.5357	2.3303	0.1269				
rank_aamerican2		1	0.7710	0.7178	1.1536	0.2828				
rank_median_income2	high median income category	1	-3.3815	0.7578	19.9140	<.0001				

Odds Ratio Estimates			
Effect	Point Estimate	95% Confiden	
rank_aamerican2	2.162	0.529	8.827
$rank_median_income2\ high\ median\ income\ category\ vs\ low\ median\ income\ category$	0.034	0.008	0.150

```
proc logistic data=work.sd plots(only)=effect;
    class rank_median_income2 / param=ref;
    model rank_dn2=rank_aamerican2 rank_median_income2 ;
    code file='/folders/myfolders/effectscode.sas';
    quit;
    title ' ';
    data predprob;
    set work.sd;
    %include '/folders/myfolders/effectscode.sas';
    run;
```

I_rank_dn2	U_rank_dn2	P_rank_dn20	P_rank_dn21	_LMR_BAD	_st12	_0_0	_LPO	_TEMP	_P0	_P1	_MAXP	_IY
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
0	0	0.928489406	0.071510594	0	1	0	0	2.563713366	0.928489406	0.071510594	0.928489406	1
0	0	0.928489406	0.071510594	0	1	0	0	2.563713366	0.928489406	0.071510594	0.928489406	1
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
1	1	0.169552694	0.830447306	0	0	1	-4.152514091	0.204170321	0.169552694	0.830447306	0.830447306	2
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
0	0	0.857262458	0.142737542	0	1	0	-0.770976819	1.792736547	0.857262458	0.142737542	0.857262458	1
1	1	0.169552694	0.830447306	0	0	1	-4.152514091	0.204170321	0.169552694	0.830447306	0.830447306	2
1	1	0.169552694	0.830447306	0	0	1	-4.152514091	0.204170321	0.169552694	0.830447306	0.830447306	2
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
0	0	0.928489406	0.071510594	0	1	0	0	2.563713366	0.928489406	0.071510594	0.928489406	1
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
0	0	0.928489406	0.071510594	0	1	0	0	2.563713366	0.928489406	0.071510594	0.928489406	1
0	0	0.928489406	0.071510594	0	1	0	0	2.563713366	0.928489406	0.071510594	0.928489406	1
0	0	0.857262458	0.142737542	0	1	0	-0.770976819	1.792736547	0.857262458	0.142737542	0.857262458	1
1	1	0.169552694	0.830447306	0	0	1	-4.152514091	0.204170321	0.169552694	0.830447306	0.830447306	2
1	1	0.306225779	0.693774221	0	0	1	-3.381537272	0.441391119	0.306225779	0.693774221	0.693774221	2
0	0	0.857262458	0.142737542	0	1	0	-0.770976819	1.792736547	0.857262458	0.142737542	0.857262458	1

Logistic Regression



Effects plots use Maximum Likelihood Estimators (MLE) to provide probabilities for all 60 of the sample cases in the sample.

The effects plot just connects the dots