

Lending Mortgage Analysis

Exploratory Data Analysis

Marjorie Blanco, Joe Thompson, Haodi Tu

The data set contains 1989 records. The overall descriptive statistics:

```
##      Married      Meet credit history guidelines
## No      : 678    0 : 171
## Unknown:   3    1 :1816
## Yes      :1308   666:   2
##
##
##
## Other obligations as a percent of total income non-Hispanic Black
## Min.      : 0.00                                No :1792
## 1st Qu.:28.00                                Yes: 197
## Median :33.00
## Mean      :32.39
## 3rd Qu.:37.00
## Max.      :95.00
## Hispanic      Male      Mortgage loan approved
## No :1878      No      : 369      No : 244
## Yes: 111      Unknown:  15      Yes:1745
##                      Yes      :1605
##
##
##
## Loan amount/purchase price      Race
## Min.      : 2.105                Hispanic      : 111
## 1st Qu.: 70.000                non-Hispanic Black: 197
## Median : 80.000                non-Hispanic White:1681
## Mean      : 77.064
## 3rd Qu.: 89.894
## Max.      :257.143
```

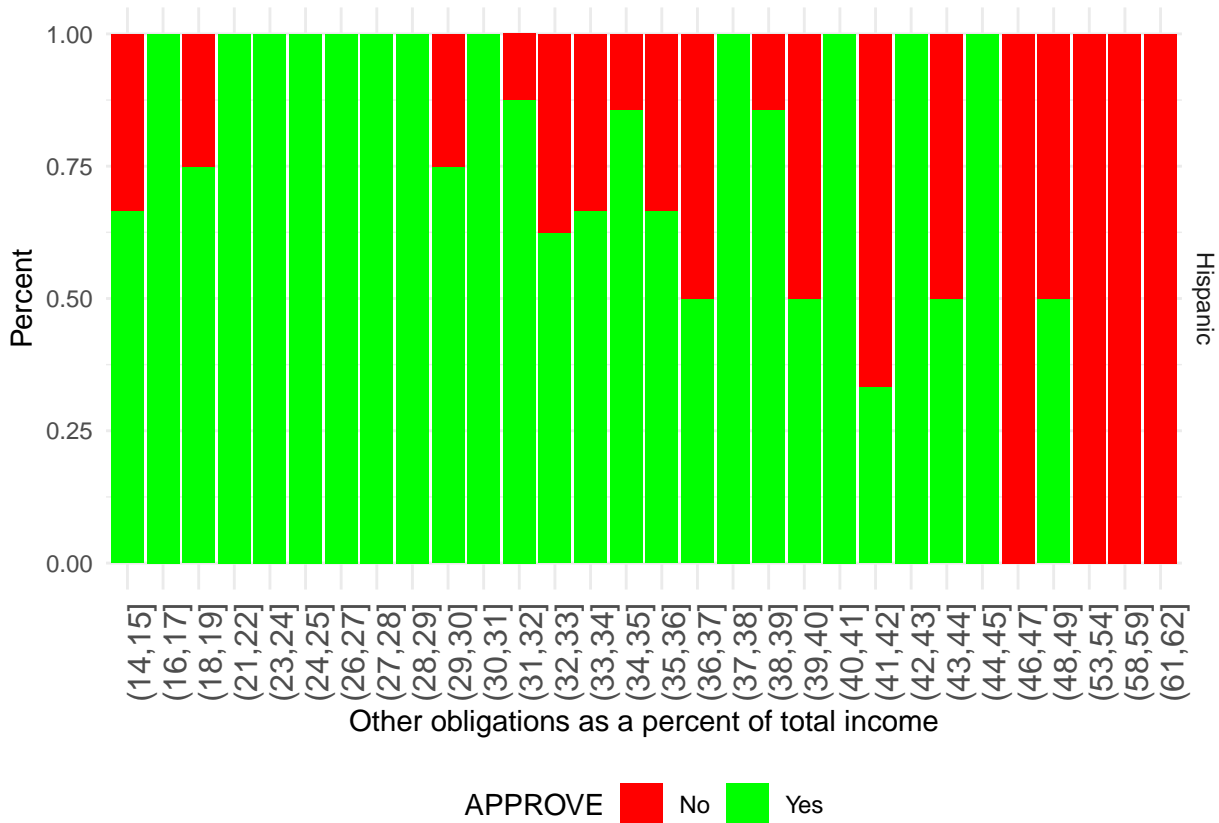
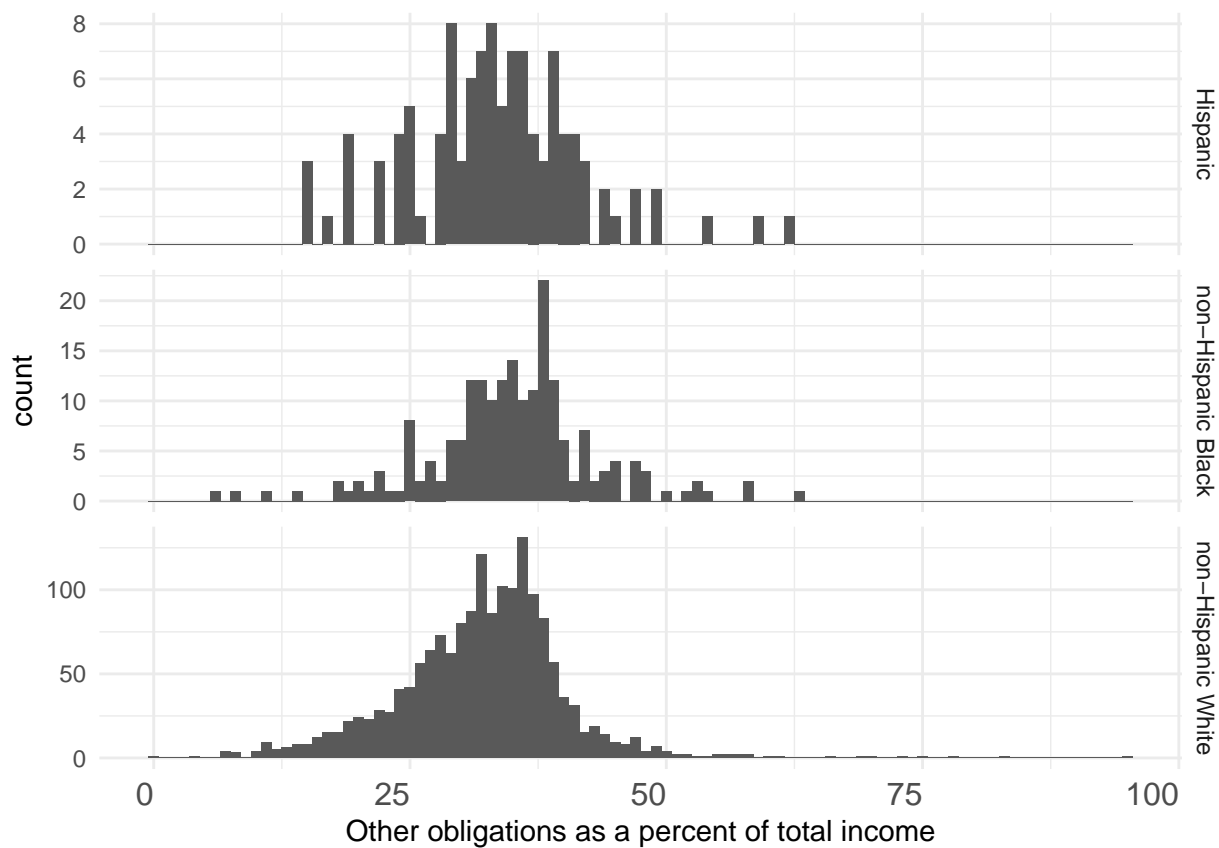
Descriptive statistics by Race:

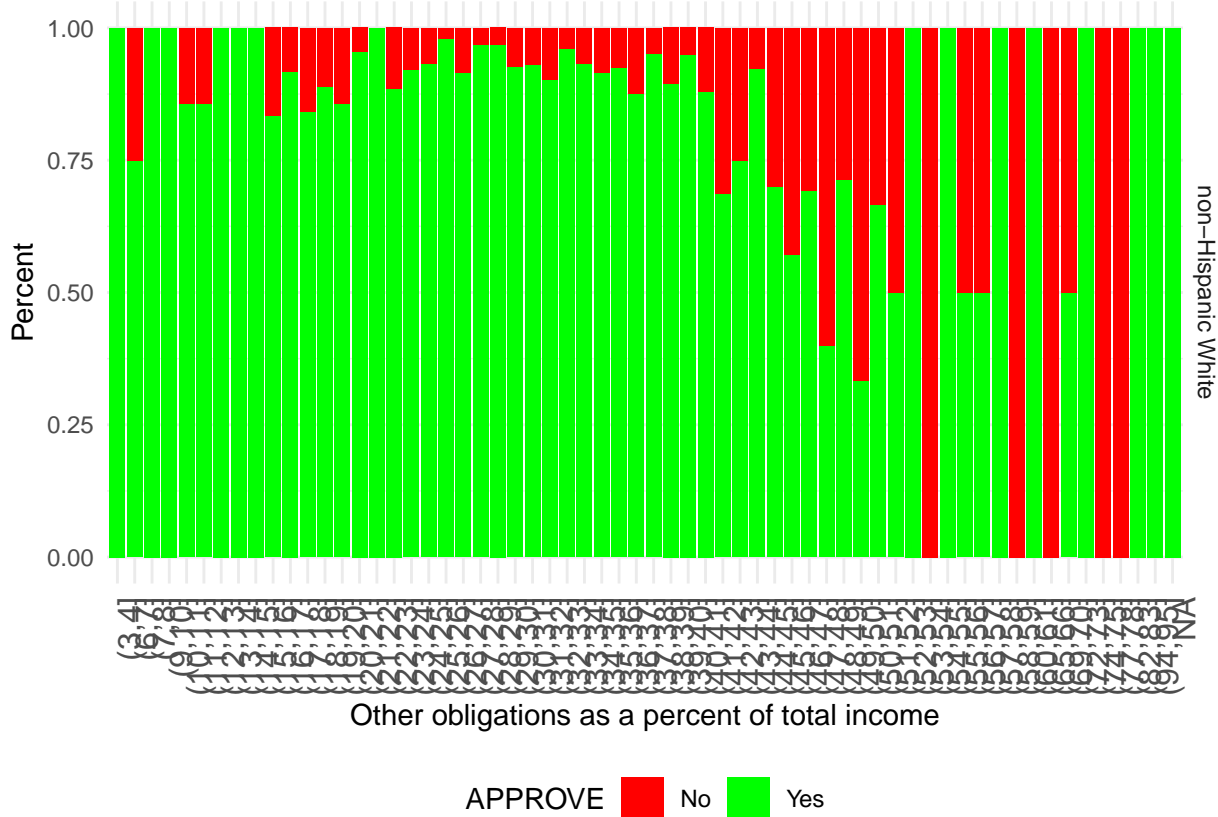
```
## $Hispanic
##      MARRIED      GDLIN      OBRAT      MALE      APPROVE
## No      :31    0 :16    Min.      :14.60    No      :22    No :26
## Unknown:  1    1 :95    1st Qu.:29.00    Unknown:  2    Yes:85
## Yes      :79   666:  0    Median   :33.00    Yes      :87
##
##                      Mean      :33.46
##                      3rd Qu.:38.45
##                      Max.      :62.00
##      LOANPRC      RACE
## Min.      : 39.39    Hispanic      :111
## 1st Qu.: 80.00    non-Hispanic Black:  0
## Median : 89.39    non-Hispanic White:  0
## Mean      : 85.17
## 3rd Qu.: 90.42
## Max.      :162.63
```

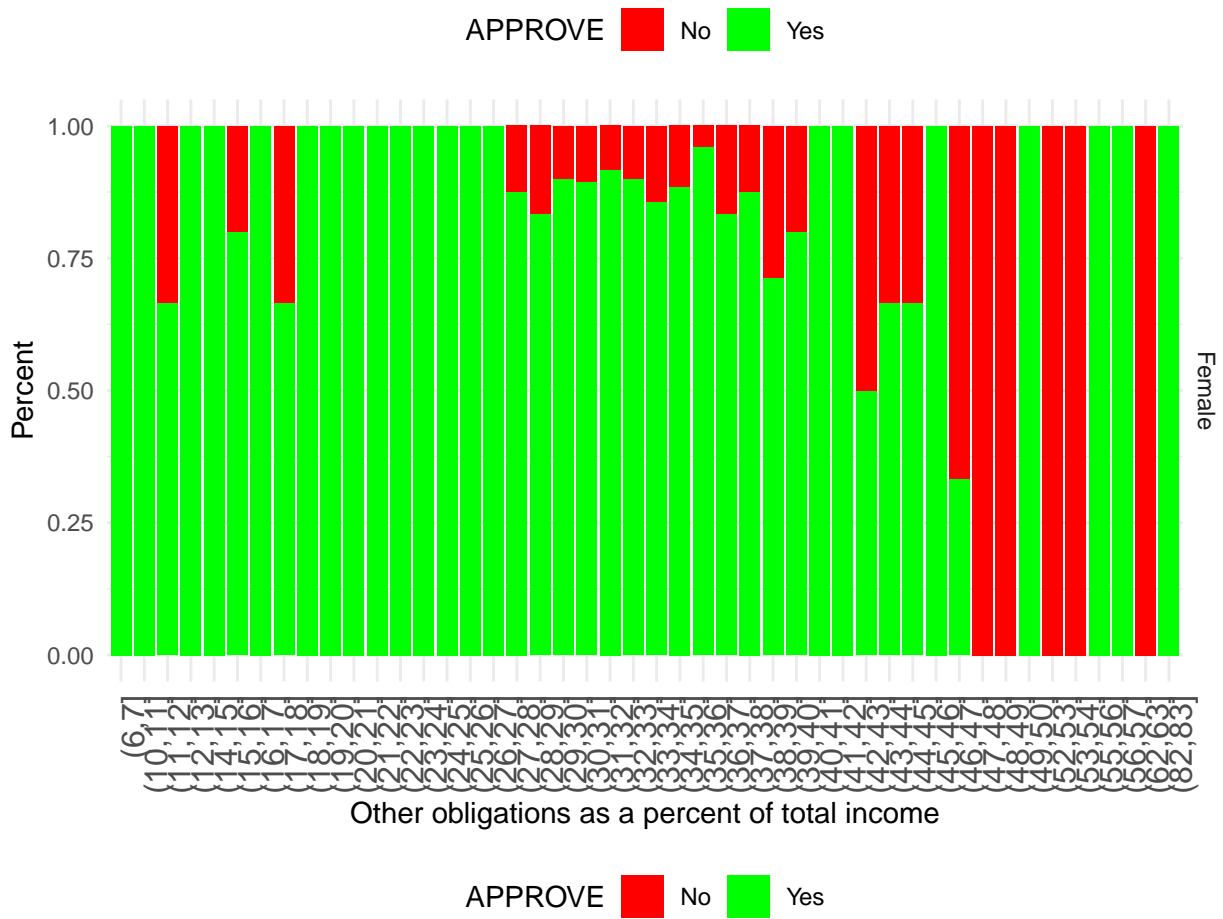
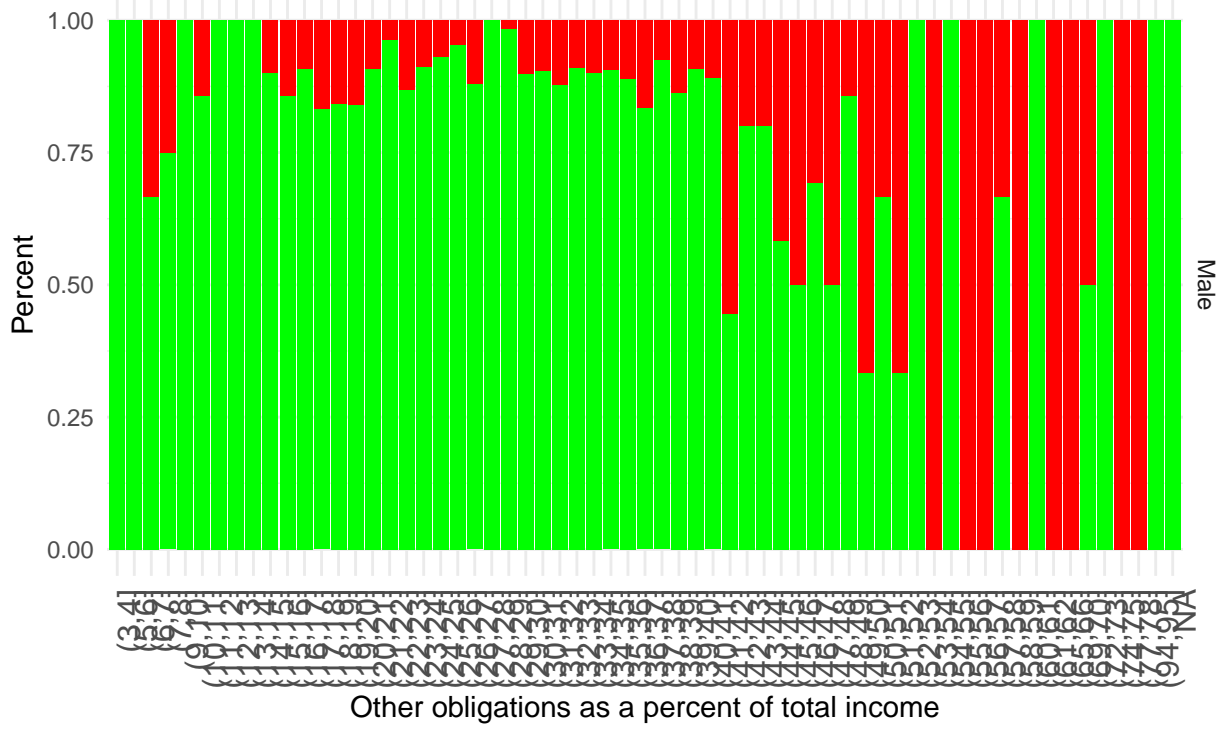
```

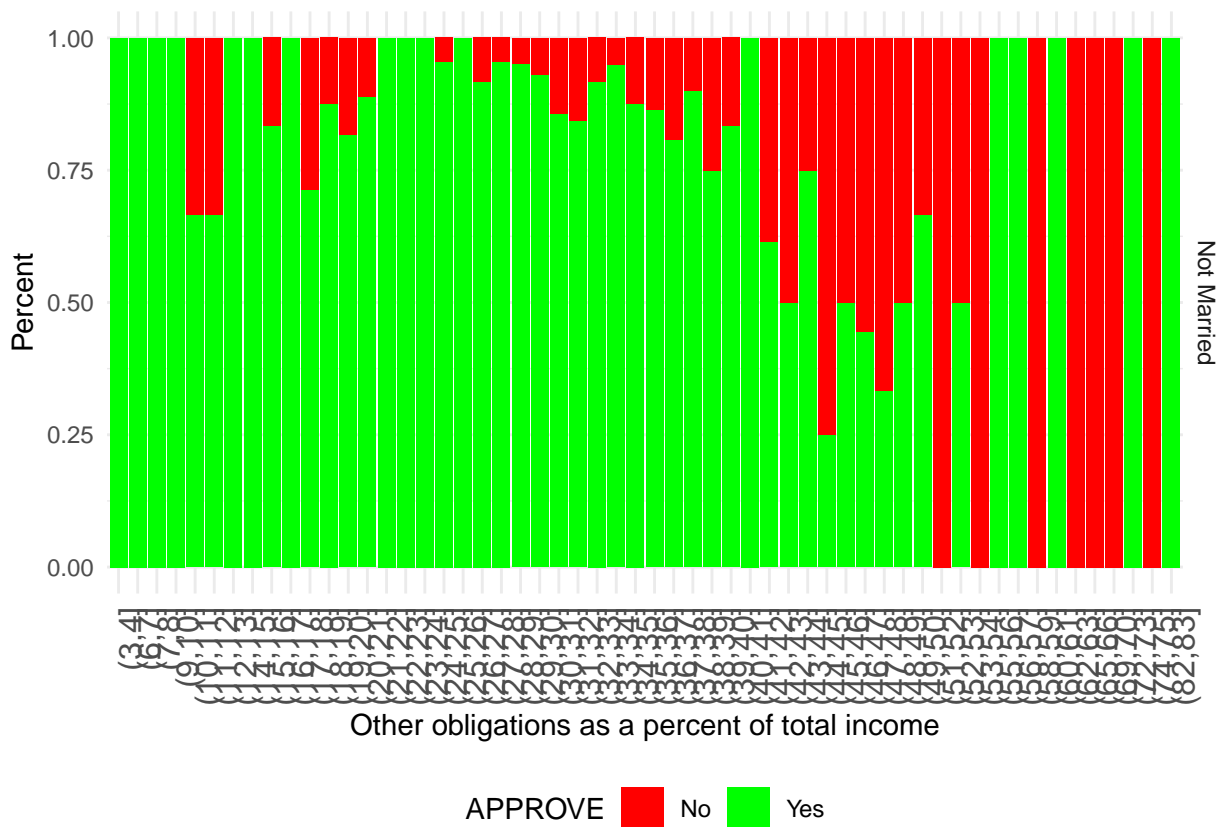
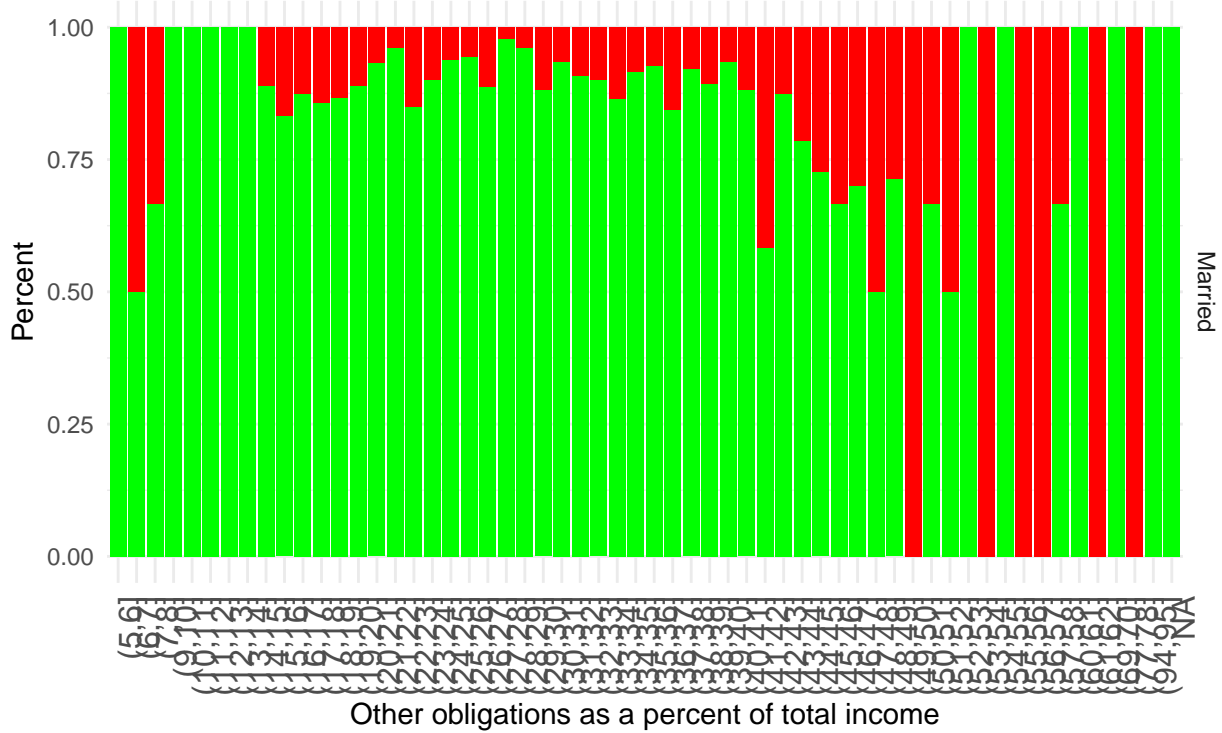
##
## $`non-Hispanic Black`
##      MARRIED      GDLIN      OBRAT      MALE      APPROVE
## No      : 76      0 : 53      Min.    : 5.60      No      : 51      No : 64
## Unknown:  0      1 :144      1st Qu.:31.00      Unknown:  2      Yes:133
## Yes      :121      666:  0      Median :35.00      Yes       :144
##
##                               Mean    :34.94
##                               3rd Qu.:38.90
##                               Max.    :63.00
##      LOANPRC      RACE
## Min.    : 28.99      Hispanic      :  0
## 1st Qu.: 80.00      non-Hispanic Black:197
## Median : 87.02      non-Hispanic White:  0
## Mean    : 83.97
## 3rd Qu.: 90.24
## Max.    :255.52
##
## $`non-Hispanic White`
##      MARRIED      GDLIN      OBRAT      MALE      APPROVE
## No      : 571      0 : 102      Min.    : 0.00      No      : 296      No : 154
## Unknown:  2      1 :1577      1st Qu.:27.60      Unknown: 11      Yes:1527
## Yes      :1108      666:  2      Median :32.50      Yes       :1374
##
##                               Mean    :32.02
##                               3rd Qu.:36.50
##                               Max.    :95.00
##      LOANPRC      RACE
## Min.    :  2.105      Hispanic      :  0
## 1st Qu.: 68.182      non-Hispanic Black:  0
## Median : 79.888      non-Hispanic White:1681
## Mean    : 75.719
## 3rd Qu.: 89.623
## Max.    :257.143

```



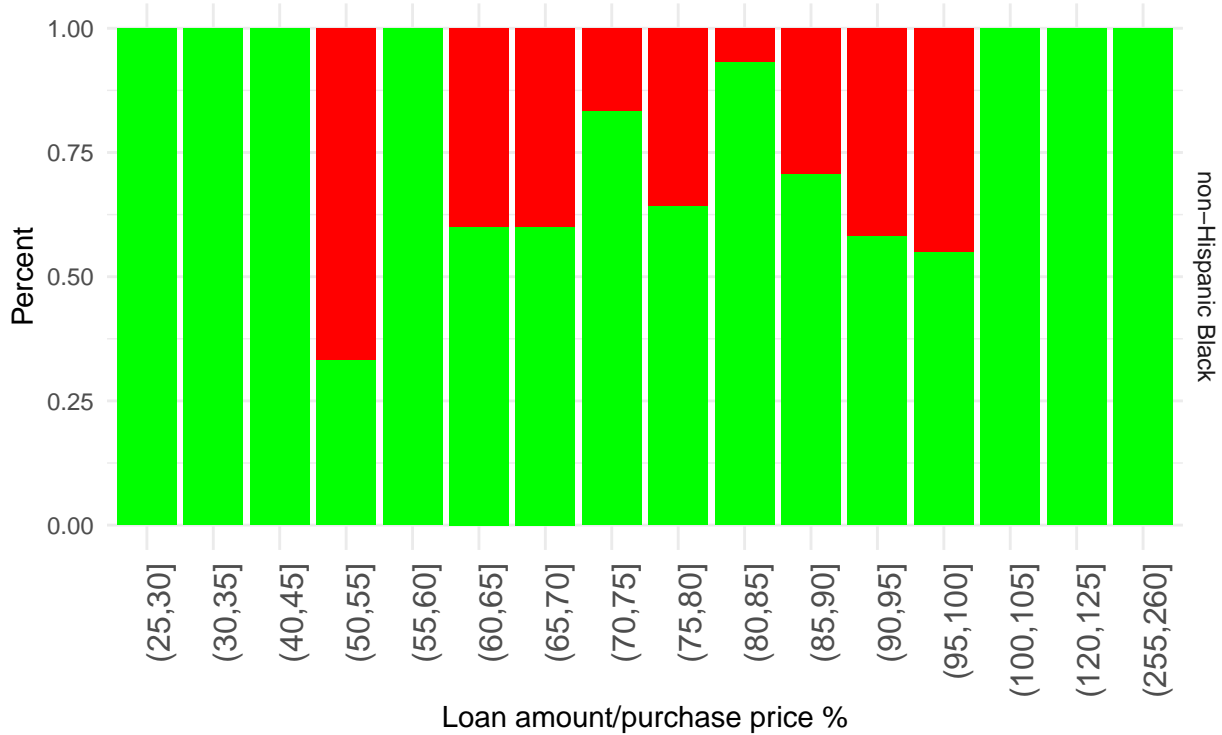




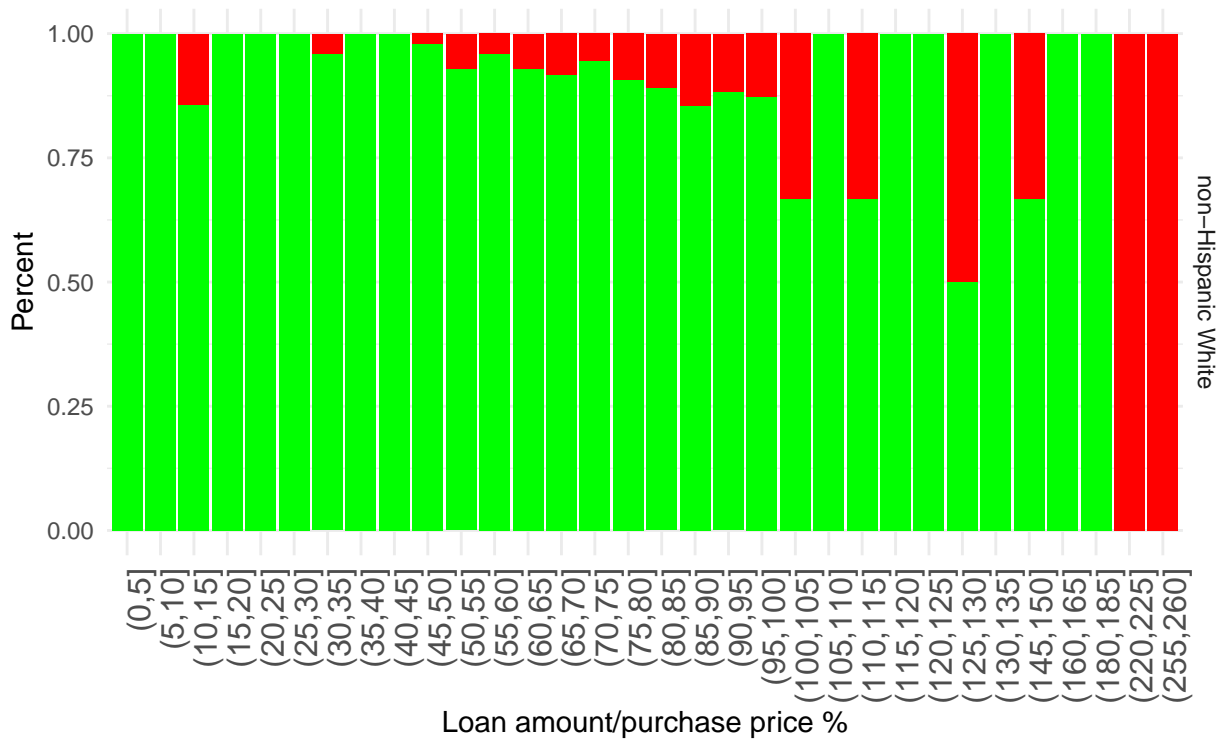




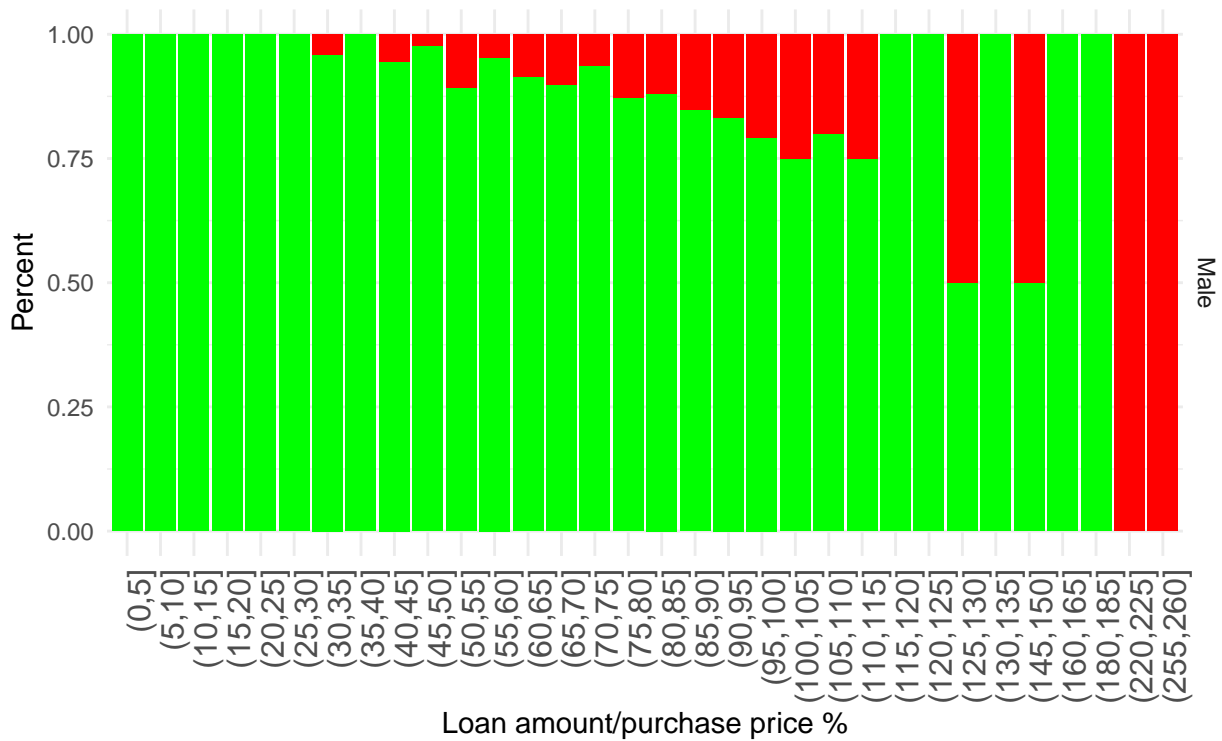
APPROVE No Yes



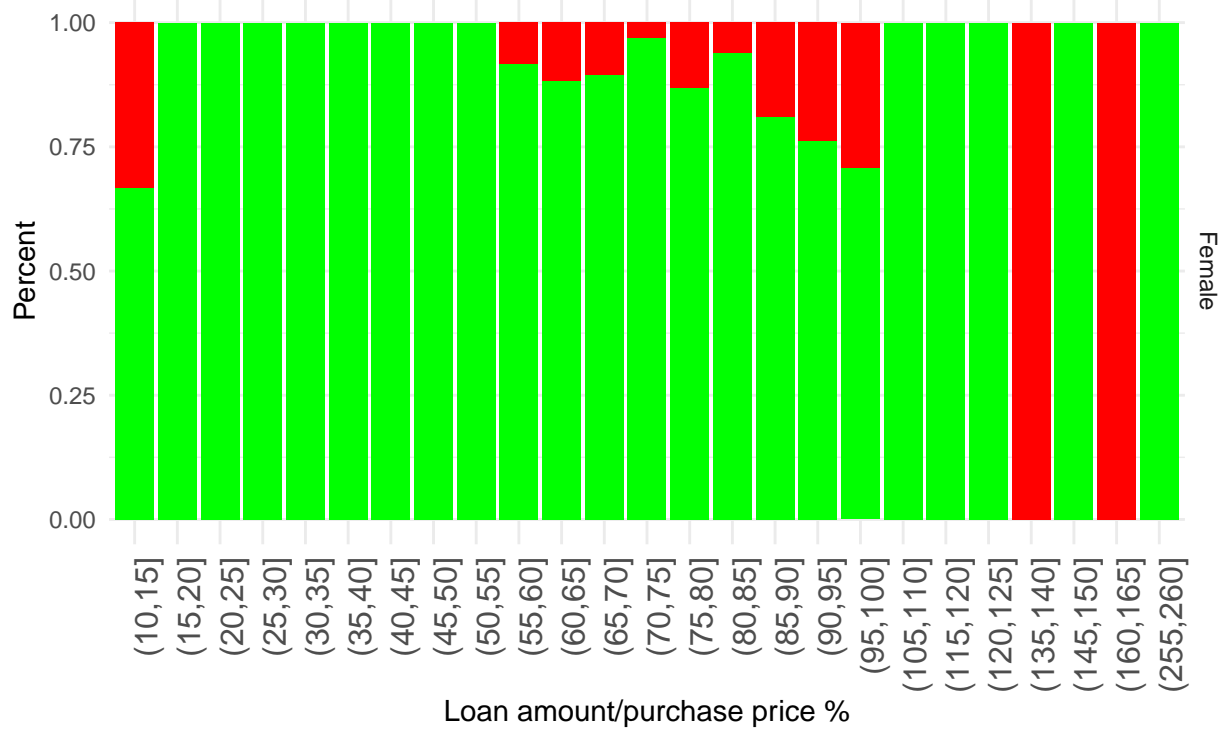
APPROVE No Yes

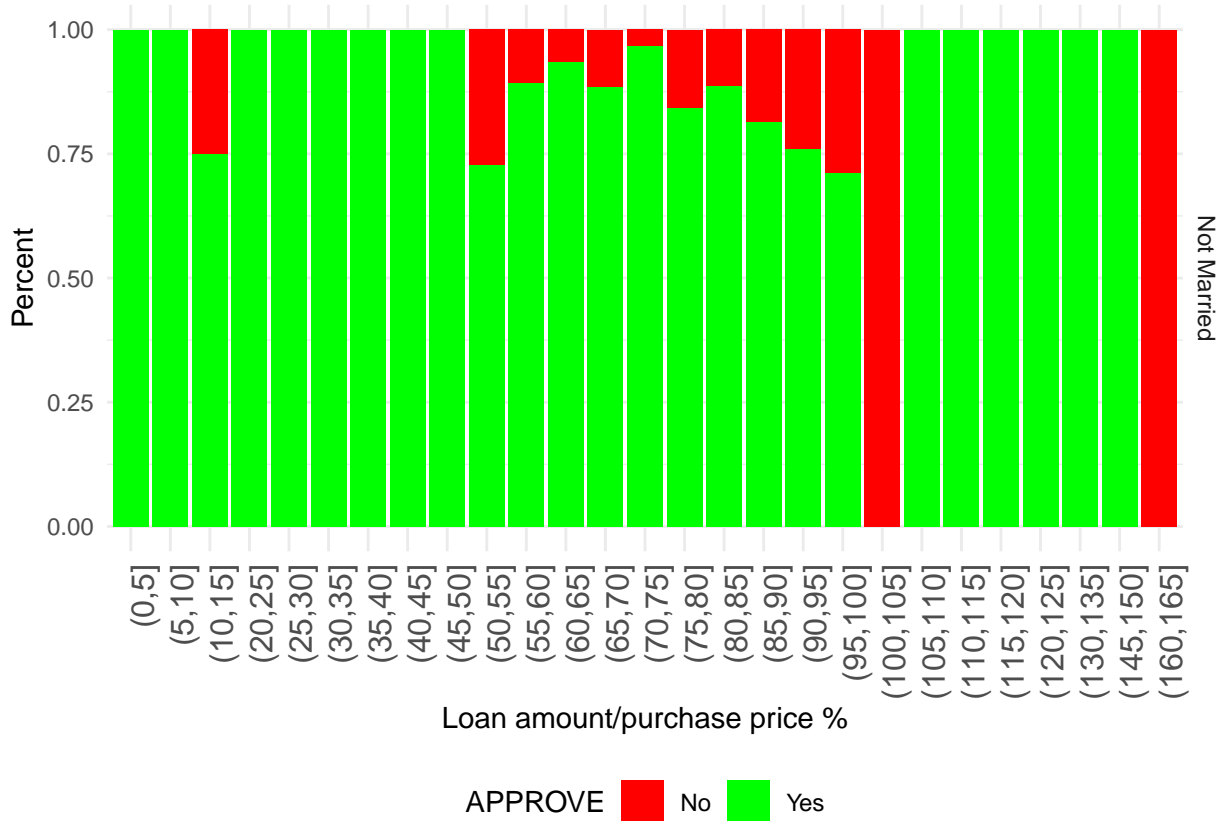


APPROVE No Yes



APPROVE No Yes





Descriptive statistics by Marital Status:

```
## $No
##      MARRIED      GDLIN      OBRAT      MALE      APPROVE
## No      :678      0 : 64      Min.    : 4.00      No      :252      No :102
## Unknown:  0      1 :614      1st Qu.:28.00      Unknown:  7      Yes:576
## Yes      :  0      666:  0      Median :33.00      Yes      :419
##
##                      Mean    :32.74
##                      3rd Qu.:37.00
##                      Max.    :83.00
##      LOANPRC      RACE
## Min.    :  2.105      Hispanic      : 31
## 1st Qu.: 72.426      non-Hispanic Black: 76
## Median  : 80.000      non-Hispanic White:571
## Mean    : 77.967
## 3rd Qu.: 89.978
## Max.    :162.626
##
## $Unknown
##      MARRIED      GDLIN      OBRAT      MALE      APPROVE      LOANPRC
## No      :0      0 :0      Min.    :13.0      No      :1      No :0      Min.    : 86.96
## Unknown:3      1 :3      1st Qu.:23.3      Unknown:0      Yes:3      1st Qu.: 88.62
## Yes      :0      666:0      Median :33.6      Yes      :2      Median : 90.29
##
##                      Mean    :27.2      Mean    : 98.16
##                      3rd Qu.:34.3      3rd Qu.:103.76
##                      Max.    :35.0      Max.    :117.24
##
##      RACE
## Hispanic      :1
## non-Hispanic Black:0
```

```

## non-Hispanic White:2
##
##
##
##
## $Yes
##      MARRIED      GDLIN      OBRAT      MALE      APPROVE
## No      :    0    0 : 107    Min.    : 0.00    No      : 116    No : 142
## Unknown:    0    1 :1199    1st Qu.:28.00    Unknown:    8    Yes:1166
## Yes     :1308    666:    2    Median :33.00    Yes      :1184
##                                     Mean   :32.22
##                                     3rd Qu.:37.00
##                                     Max.   :95.00
##      LOANPRC      RACE
## Min.    : 8.772    Hispanic      : 79
## 1st Qu.: 68.857    non-Hispanic Black: 121
## Median : 80.000    non-Hispanic White:1108
## Mean    : 76.547
## 3rd Qu.: 89.866
## Max.    :257.143

```

Descriptive statistics by Gender:

```

## $No
##      MARRIED      GDLIN      OBRAT      MALE      APPROVE
## No      :252    0 : 31    Min.    : 6.99    No      :369    No : 50
## Unknown:    1    1 :338    1st Qu.:28.00    Unknown:    0    Yes:319
## Yes     :116    666:    0    Median :33.00    Yes      :    0
##                                     Mean   :32.64
##                                     3rd Qu.:37.00
##                                     Max.   :83.00
##      LOANPRC      RACE
## Min.    : 11.01    Hispanic      : 22
## 1st Qu.: 70.83    non-Hispanic Black: 51
## Median : 80.00    non-Hispanic White:296
## Mean    : 77.66
## 3rd Qu.: 90.00
## Max.    :255.52
##
## $Unknown
##      MARRIED      GDLIN      OBRAT      MALE      APPROVE
## No      :7    0 : 0    Min.    :24.00    No      : 0    No : 0
## Unknown:0    1 :15    1st Qu.:29.95    Unknown:15    Yes:15
## Yes     :8    666:    0    Median :34.50    Yes      : 0
##                                     Mean   :33.33
##                                     3rd Qu.:37.65
##                                     Max.   :40.30
##      LOANPRC      RACE
## Min.    :39.39    Hispanic      : 2
## 1st Qu.:74.93    non-Hispanic Black: 2
## Median :75.42    non-Hispanic White:11
## Mean    :75.59
## 3rd Qu.:80.43
## Max.    :92.90
##
## $Yes

```

```

##      MARRIED      GDLIN      OBRAT      MALE      APPROVE
## No      : 419      0 : 140      Min.   : 0.00      No      :   0      No : 194
## Unknown:   2      1 :1463      1st Qu.:28.00      Unknown:   0      Yes:1411
## Yes     :1184      666:   2      Median :33.00      Yes      :1605
##
##                               Mean    :32.32
##                               3rd Qu.:37.00
##                               Max.    :95.00
##      LOANPRC      RACE
## Min.   :   2.105      Hispanic      :   87
## 1st Qu.: 69.655      non-Hispanic Black: 144
## Median : 80.000      non-Hispanic White:1374
## Mean    : 76.942
## 3rd Qu.: 89.881
## Max.    :257.143

```

There are 3 records are missing married (MARRIED) field.

ID	MARRIED	GDLIN	OBRAT	BLACK	HISPAN	MALE	APPROVE	LOANPRC	RACE
356	Unknown	1	35.0	No	Yes	Yes	Yes	86.95652	Hispanic
759	Unknown	1	33.6	No	No	Yes	Yes	90.28571	non-Hispanic White
1392	Unknown	1	13.0	No	No	No	Yes	117.24140	non-Hispanic White

There are 3 records are missing married (GDLIN) field.

ID	MARRIED	GDLIN	OBRAT	BLACK	HISPAN	MALE	APPROVE	LOANPRC	RACE
881	Yes	666	35	No	No	Yes	Yes	75.82939	non-Hispanic White
1229	Yes	666	26	No	No	Yes	Yes	100.00000	non-Hispanic White

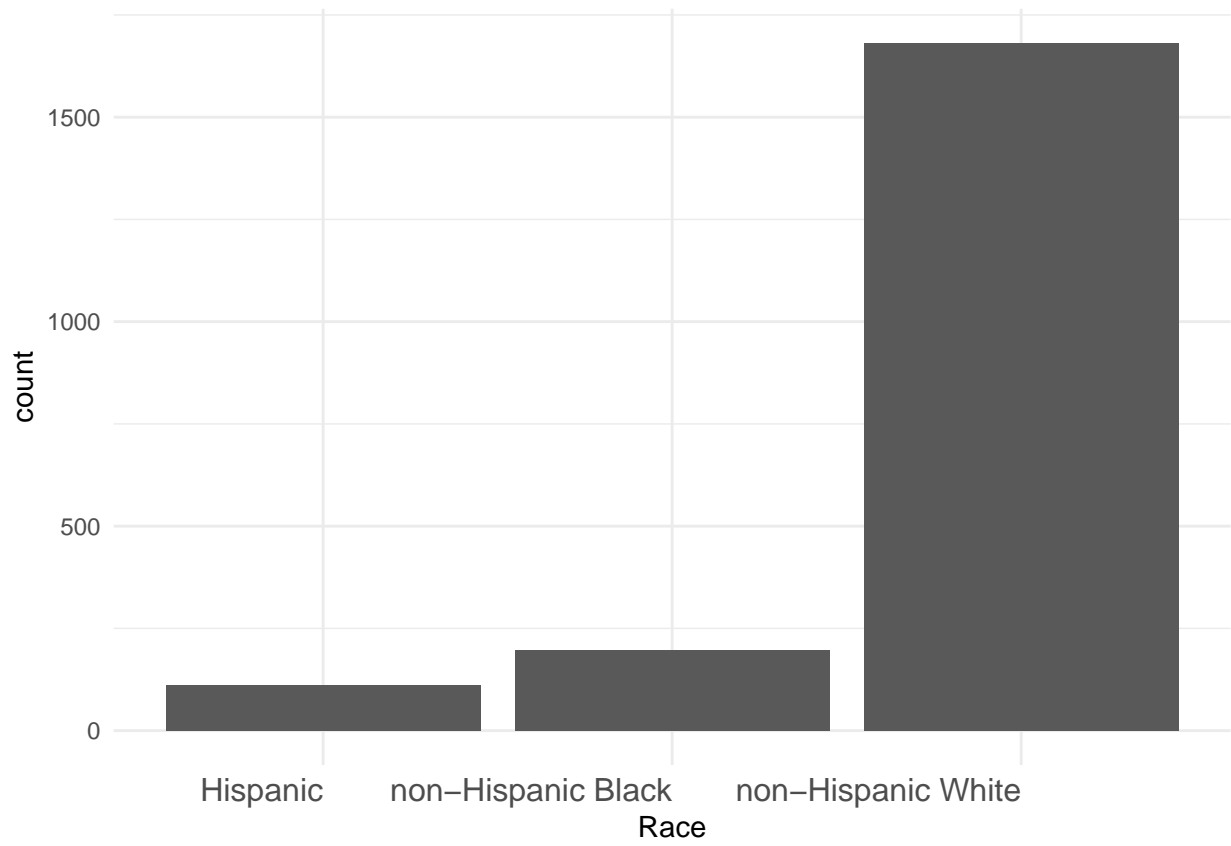
There are 15 records are missing gender (MALE) field.

ID	MARRIED	GDLIN	OBRAT	BLACK	HISPAN	MALE	APPROVE	LOANPRC	RACE
1	No	1	34.5	No	No	Unknown	Yes	75.42373	non-Hispanic White
127	No	1	31.6	No	No	Unknown	Yes	80.80000	non-Hispanic White
286	Yes	1	37.3	No	No	Unknown	Yes	80.05337	non-Hispanic White
452	Yes	1	40.1	Yes	No	Unknown	Yes	75.00000	non-Hispanic Black
618	Yes	1	38.5	No	No	Unknown	Yes	92.90323	non-Hispanic White
695	Yes	1	25.0	No	No	Unknown	Yes	64.48276	non-Hispanic White
762	Yes	1	27.6	No	No	Unknown	Yes	75.55556	non-Hispanic White
768	No	1	35.6	No	No	Unknown	Yes	64.74397	non-Hispanic White
833	Yes	1	24.0	No	Yes	Unknown	Yes	79.80769	Hispanic
979	No	1	31.7	No	No	Unknown	Yes	74.86033	non-Hispanic White
1040	No	1	38.0	Yes	No	Unknown	Yes	75.38462	non-Hispanic Black
1070	Yes	1	40.3	No	Yes	Unknown	Yes	39.39394	Hispanic
1092	Yes	1	29.7	No	No	Unknown	Yes	90.10239	non-Hispanic White
1613	No	1	30.2	No	No	Unknown	Yes	90.00000	non-Hispanic White
1924	No	1	35.8	No	No	Unknown	Yes	75.32051	non-Hispanic White

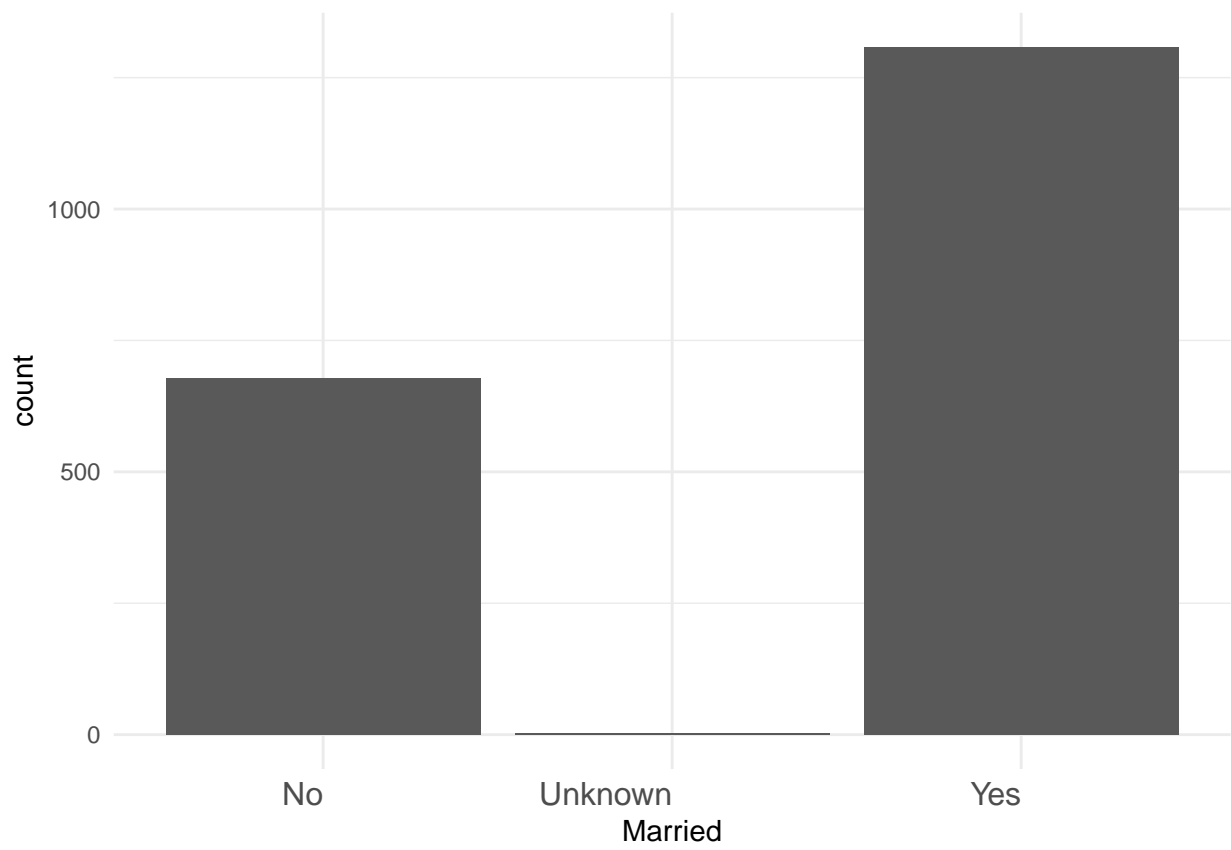
```

##      APPROVE
## RACE      No  Yes
## Hispanic      26  85
## non-Hispanic Black  64 133
## non-Hispanic White 154 1527

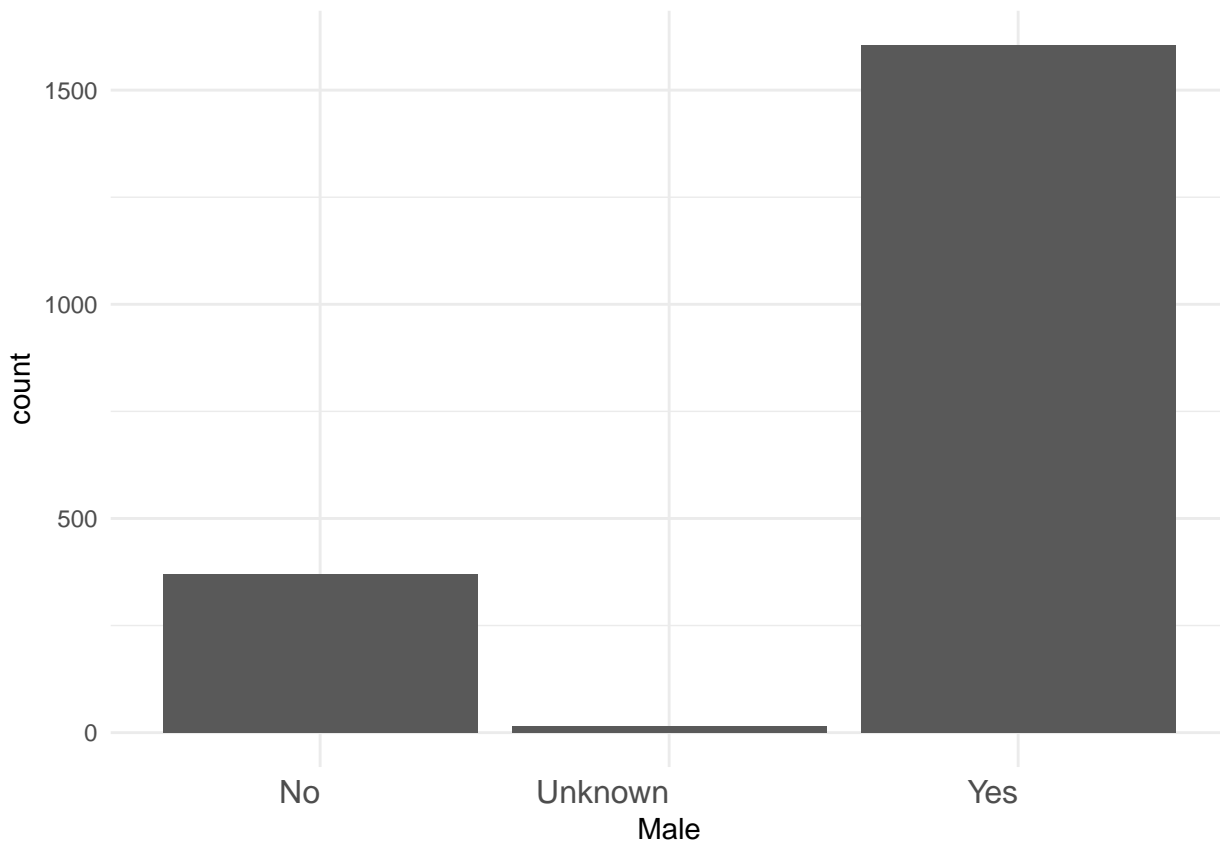
```



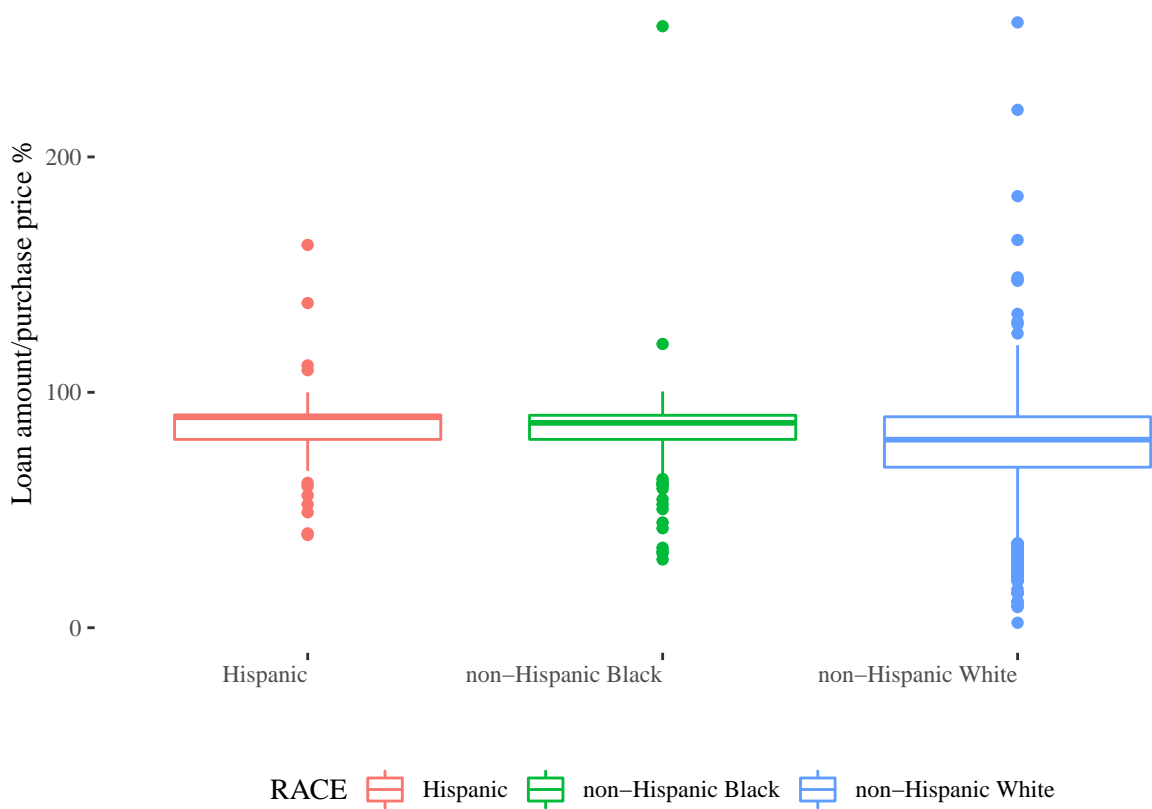
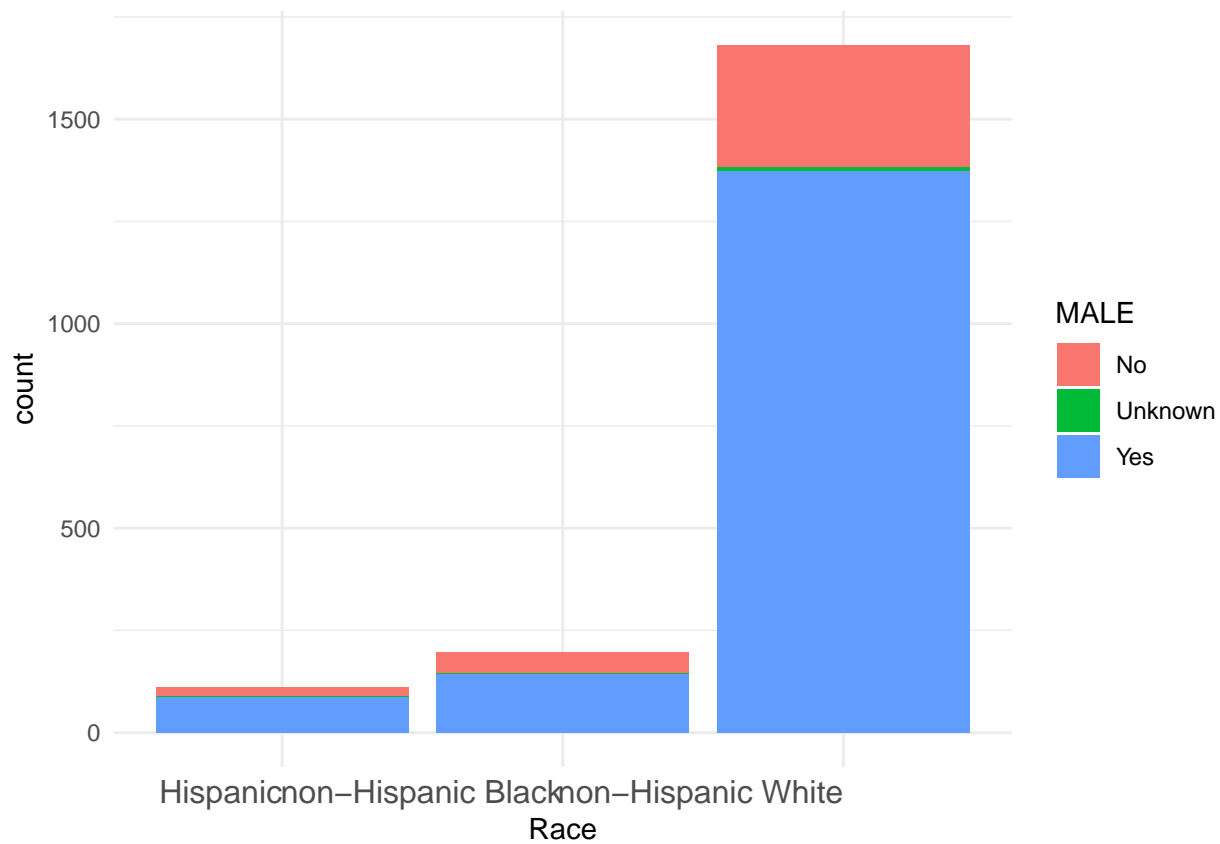
## APPROVE			
## MARRIED	No	Yes	
## No	102	576	
## Unknown	0	3	
## Yes	142	1166	

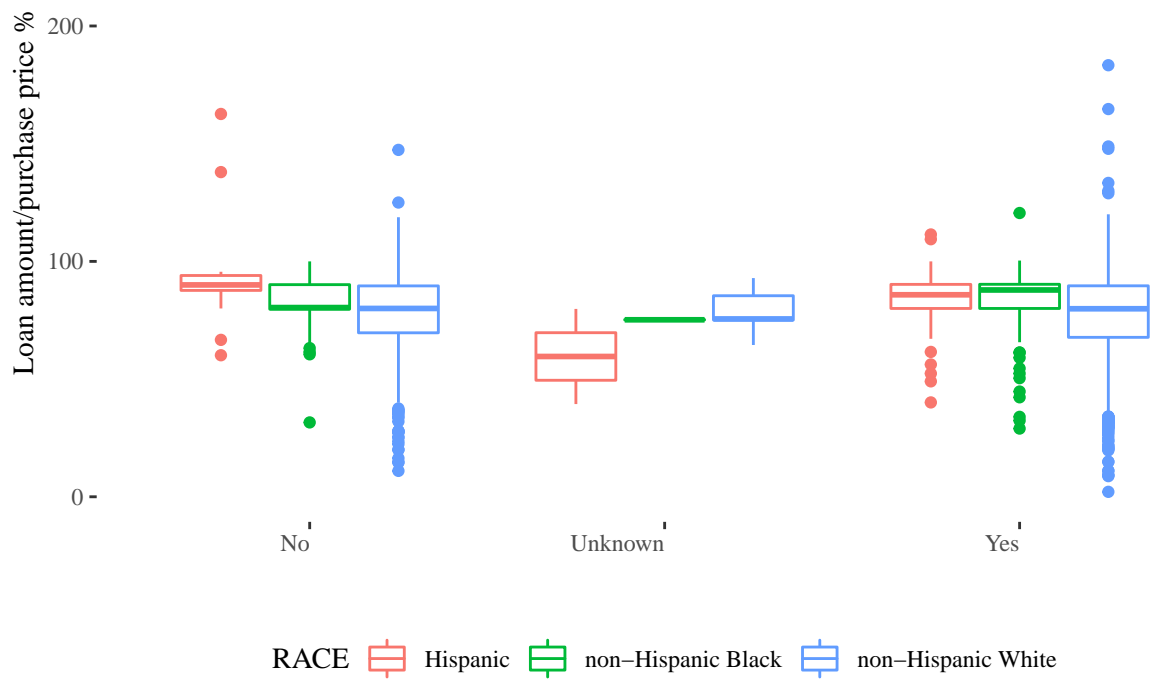
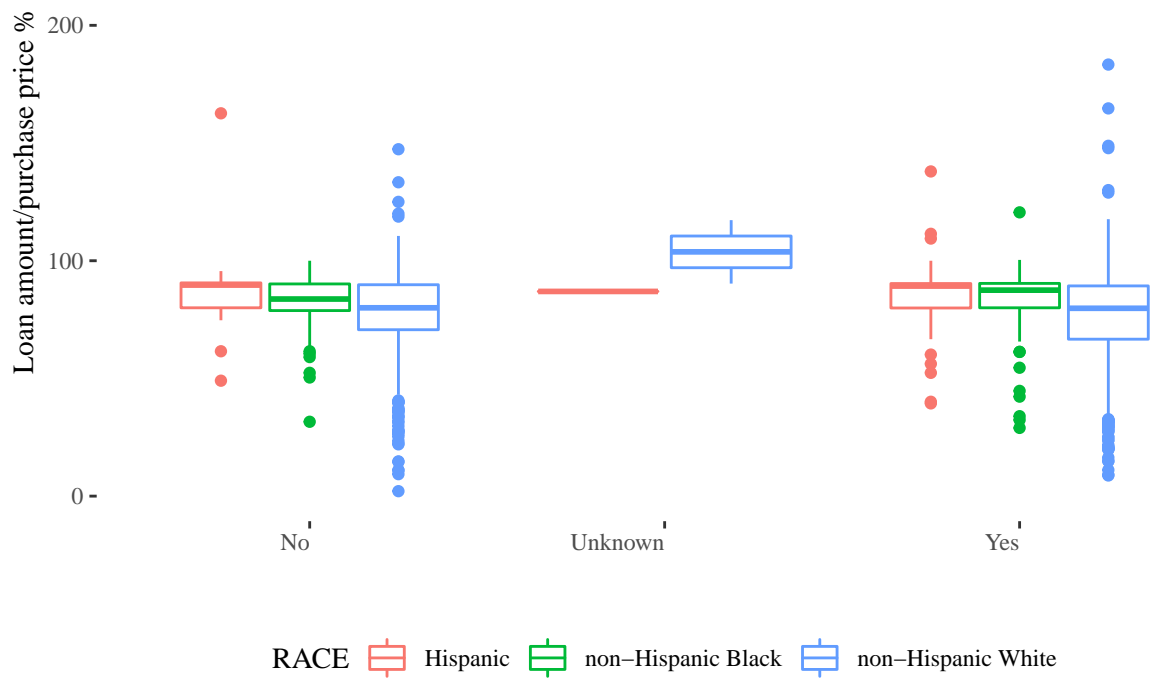


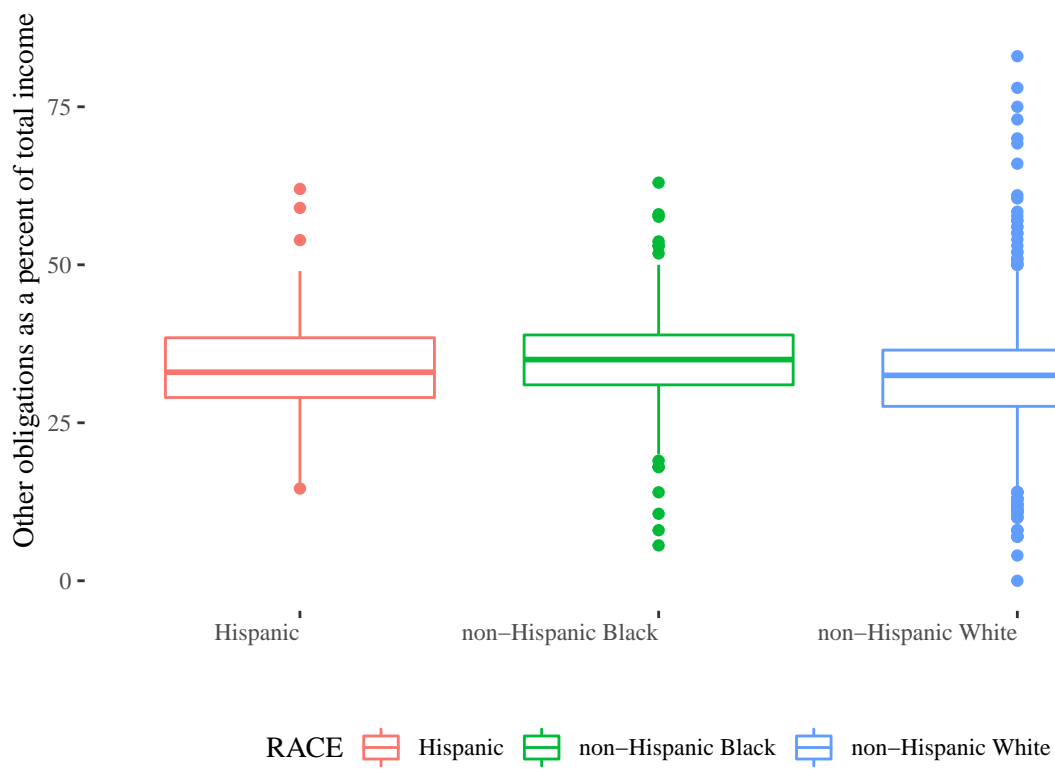
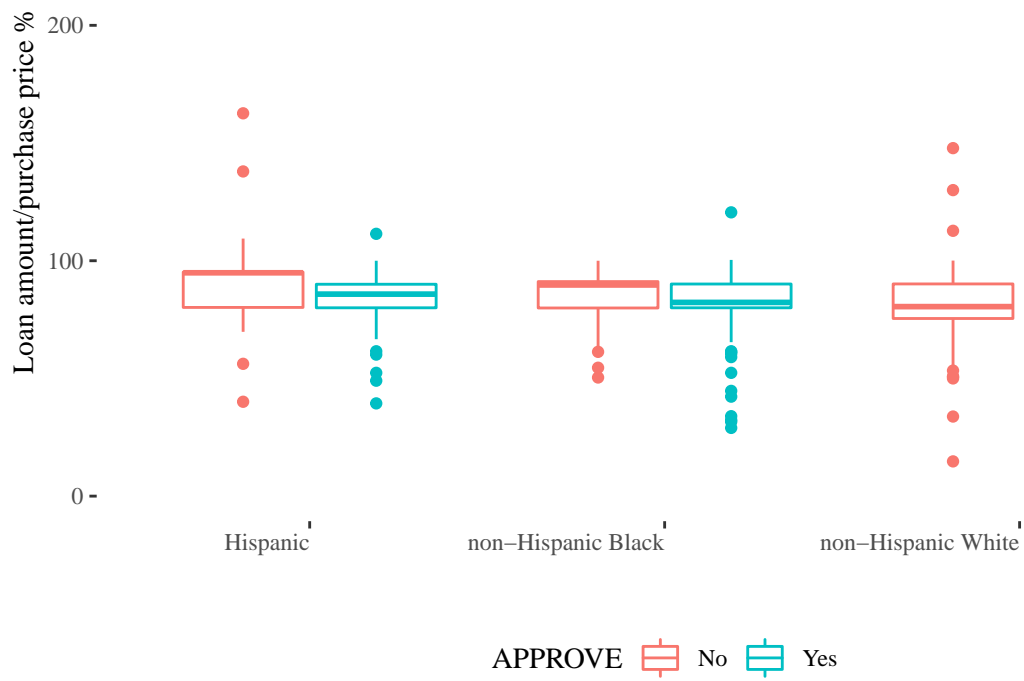
```
##          APPROVE
## MALE          No  Yes
##   No           50 319
##  Unknown         0  15
##   Yes          194 1411
```



```
## , , RACE = Hispanic
##
##      APPROVE
## MARRIED    No  Yes
##   No         7  24
##   Unknown     0   1
##   Yes        19  60
##
## , , RACE = non-Hispanic Black
##
##      APPROVE
## MARRIED    No  Yes
##   No        27  49
##   Unknown     0   0
##   Yes        37  84
##
## , , RACE = non-Hispanic White
##
##      APPROVE
## MARRIED    No  Yes
##   No        68 503
##   Unknown     0   2
##   Yes        86 1022
```







Other obligations as a percent of total income

75 -
50 -
25 -
0 -

No

Unknown

Yes

RACE ▢ Hispanic ▢ non-Hispanic Black ▢ non-Hispanic White

Other obligations as a percent of total income

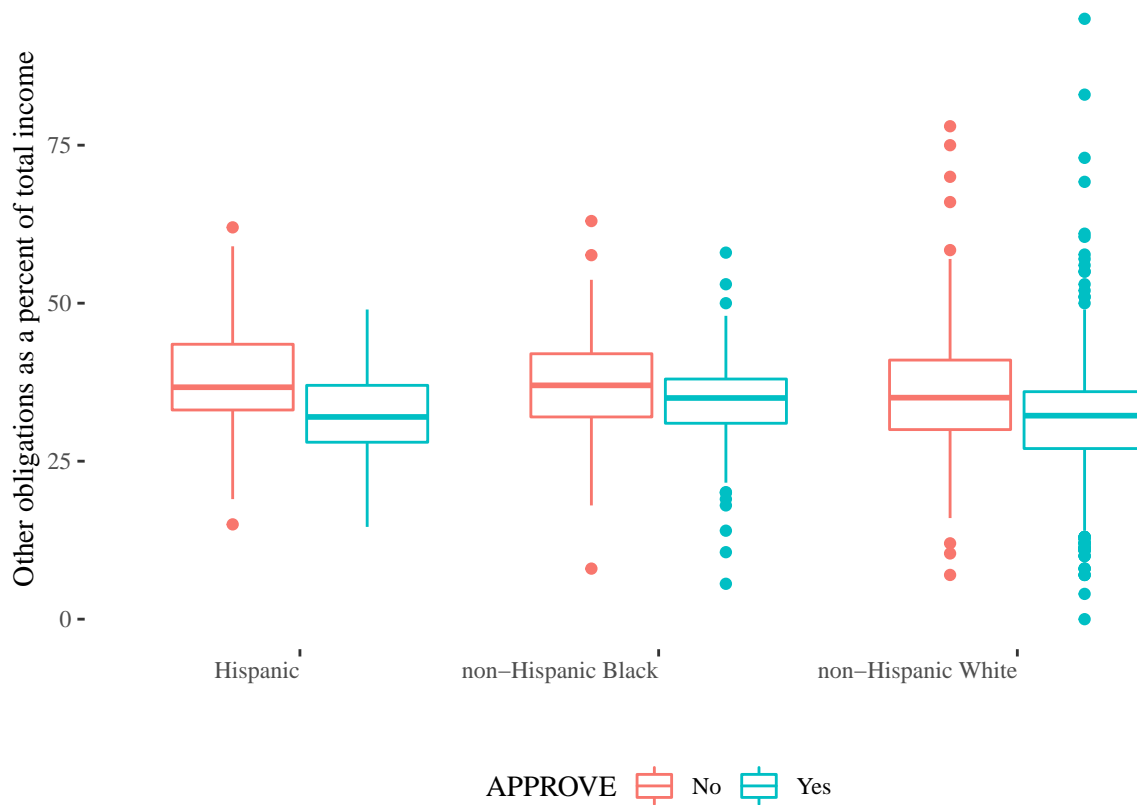
75 -
50 -
25 -
0 -

No

Unknown

Yes

RACE ▢ Hispanic ▢ non-Hispanic Black ▢ non-Hispanic White



Logistic Regression Full Model

$$\log(p/1-p) = b_0 + b_1 * GDLIN + b_2 * OBRAT + b_3 * BLACK + b_4 * HISPAN + b_5 * MALE + b_6 * LOANPRC + b_7 * MARRIED$$

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED + MALE, family = "binomial", data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
##   1.38153      3.71927     -0.03407     -0.81569     -0.90001
##   LOANPRC  MARRIED1      MALE1
##  -0.01681      0.47574     -0.05395
##
## Degrees of Freedom: 1968 Total (i.e. Null);  1961 Residual
## Null Deviance:      1475
## Residual Deviance: 959.4    AIC: 975.4
```

For every one unit change in **OBRAT**, the log odds of loan approval (versus non loan approval) decreases by 0.0340739.

For every one unit change in **LOANPRC**, the log odds of loan approval (versus non loan approval) decreases by 0.0168119.

The log odds of loan approval for applicants that meet credit guidelines increases by 3.719269.

The log odds of loan approval for married applicants increases by 0.4757419.

The log odds of loan approval for Black applicants decreases by 0.8156932.

The log odds of loan approval for Hispanic applicants decreases by 0.9000102.

Statistical Tests for Individual Predictors

```
##           Overall
## GDLIN1    17.126161
## OBRAT      3.305071
## BLACK1     3.396220
## HISPAN1    2.897795
## LOANPRC    3.313389
## MARRIED1   2.477753
## MALE1      0.229981

## Wald test for GDLIN
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = "binomial", data = data1)
## F = 293.3054 on 1 and 1961 df: p= < 0.000000000000000222

## Wald test for OBRAT
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = "binomial", data = data1)
## F = 10.9235 on 1 and 1961 df: p= 0.00096664

## Wald test for BLACK
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = "binomial", data = data1)
## F = 11.53431 on 1 and 1961 df: p= 0.00069686

## Wald test for HISPAN
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = "binomial", data = data1)
## F = 8.397215 on 1 and 1961 df: p= 0.0037997

## Wald test for MALE
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = "binomial", data = data1)
## F = 0.05289126 on 1 and 1961 df: p= 0.81813

## Wald test for LOANPRC
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = "binomial", data = data1)
## F = 10.97855 on 1 and 1961 df: p= 0.00093849

## Wald test for MARRIED
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = "binomial", data = data1)
## F = 6.13926 on 1 and 1961 df: p= 0.013305

##           llh           llhNull           G2           McFadden           r2ML
## -479.7016510 -737.7148785 516.0264551 0.3497465 0.2305470
##           r2CU
## 0.4372079
```

Logistic Regression Reduced Model 1

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
##      1.34239      3.72139     -0.03410     -0.81143     -0.89733
##      LOANPRC      MARRIED1
##     -0.01677      0.46093
##
## Degrees of Freedom: 1968 Total (i.e. Null);  1962 Residual
## Null Deviance:      1475
## Residual Deviance: 959.5      AIC: 973.5

## Wald test for GDLIN
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 294.1283 on 1 and 1962 df: p= < 0.000000000000000222

## Wald test for OBRAT
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 10.93867 on 1 and 1962 df: p= 0.00095878

## Wald test for BLACK
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 11.47424 on 1 and 1962 df: p= 0.00071961

## Wald test for HISPAN
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 8.358631 on 1 and 1962 df: p= 0.0038808

## Wald test for MARRIED
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 6.482227 on 1 and 1962 df: p= 0.010972

## Wald test for LOANPRC
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 10.92337 on 1 and 1962 df: p= 0.00096669
```

For every one unit change in OBRAT, the log odds of loan approval (versus non loan approval) decreases by 0.0340979.

For every one unit change in LOANPRC, the log odds of loan approval (versus non loan approval) decreases by 0.0167734.

The log odds of loan approval for applicants that meet credit guidelines increase by 3.721387.

The log odds of loan approval for married applicants decreases by 0.4757419.

The log odds of loan approval for non married applicants decreases by NA.

The log odds of loan approval for Black applicants decreases by 0.8114263.

The log odds of loan approval for Hispanic applicants decreases by 0.8973309.

CIs using profiled log-likelihood

```
##              2.5 %      97.5 %
## (Intercept)  0.23664368  2.465735226
## GDLIN1      3.30547346  4.157682544
## OBRAT       -0.05434377 -0.013920887
## BLACK1      -1.27292218 -0.332543498
## HISPAN1     -1.48462721 -0.264593691
## LOANPRC     -0.02697832 -0.006937324
## MARRIED1     0.10487116  0.815796783
```

CIs using standard errors

```
##              2.5 %      97.5 %
## (Intercept)  0.23158635  2.453187301
## GDLIN1      3.29609767  4.146676358
## OBRAT       -0.05430447 -0.013891283
## BLACK1      -1.28092587 -0.341926697
## HISPAN1     -1.50565235 -0.289009441
## LOANPRC     -0.02672039 -0.006826433
## MARRIED1     0.10609951  0.815760700
```

Odds ratios only

```
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1      LOANPRC
##   3.8281698  41.3216682   0.9664769   0.4442240   0.4076563   0.9833665
##   MARRIED1
##   1.5855480
```

Odds ratios and 95% CI

```
##              OR      2.5 %      97.5 %
## (Intercept)  3.8281698  1.2669896  11.7721341
## GDLIN1      41.3216682  27.2614457  63.9232116
## OBRAT       0.9664769   0.9471065   0.9861756
## BLACK1      0.4442240   0.2800122   0.7170975
## HISPAN1     0.4076563   0.2265868   0.7675177
## LOANPRC     0.9833665   0.9733823   0.9930867
## MARRIED1    1.5855480   1.1105675   2.2609765
```

Logistic Regression Reduced Model 2

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC,
##   family = "binomial", data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
##   1.66805      3.69461     -0.03512     -0.81768     -0.85828
##   LOANPRC
##   -0.01659
##
## Degrees of Freedom: 1968 Total (i.e. Null);  1963 Residual
## Null Deviance:      1475
```

```
## Residual Deviance: 965.9      AIC: 977.9

## Wald test for GDLIN
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 294.1283 on 1 and 1962 df: p= < 0.000000000000000222

## Wald test for OBRAT
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 10.93867 on 1 and 1962 df: p= 0.00095878

## Wald test for BLACK
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 11.47424 on 1 and 1962 df: p= 0.00071961

## Wald test for HISPAN
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 8.358631 on 1 and 1962 df: p= 0.0038808

## Wald test for LOANPRC
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
## F = 10.92337 on 1 and 1962 df: p= 0.00096669
```

For every one unit change in OBRAT, the log odds of loan approval (versus non loan approval) decreases by 0.0351171.

For every one unit change in LOANPRC, the log odds of loan approval (versus non loan approval) decreases by 0.0165852.

The log odds of loan approval for applicants that meet credit guidelines increases by 3.6946116.

The log odds of loan approval for Black applicants decreases by 0.8176845.

The log odds of loan approval for Hispanic applicants decreases by 0.8582767.

CIs using profiled log-likelihood

```
##           2.5 %      97.5 %
## (Intercept) 0.59360475 2.764561689
## GDLIN1      3.28231679 4.126731408
## OBRAT       -0.05537654 -0.014888611
## BLACK1     -1.27852143 -0.339452086
## HISPAN1     -1.44328269 -0.227316096
## LOANPRC     -0.02676923 -0.006824773
```

CIs using standard errors

```
##           2.5 %      97.5 %
## (Intercept) 0.58607300 2.750024092
## GDLIN1      3.27321875 4.116004514
## OBRAT       -0.05536176 -0.014872430
## BLACK1     -1.28654690 -0.348822144
## HISPAN1     -1.46462838 -0.251925037
## LOANPRC     -0.02649064 -0.006679743
```


Odds ratios only

```
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1      LOANPRC
##   5.3018115  40.2299456   0.9654924   0.4414526   0.4238919   0.9835516
```

Odds ratios and 95% CI

```
##           OR      2.5 %      97.5 %
## (Intercept) 5.3018115  1.8105031 15.8720816
## GDLIN1      40.2299456 26.6374146 61.9750205
## OBRAT       0.9654924  0.9461288  0.9852217
## BLACK1      0.4414526  0.2784487  0.7121604
## HISPAN1     0.4238919  0.2361513  0.7966689
## LOANPRC     0.9835516  0.9735859  0.9931985
```

Model comparison

```
## Analysis of Deviance Table
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC + MARRIED +
##      MALE
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC + MARRIED
## Model 3: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      1961      959.40
## 2      1962      959.46 -1   -0.0531  0.81770
## 3      1963      965.87 -1   -6.4163  0.01131 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Likelihood ratio test
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC + MARRIED +
##      MALE
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC + MARRIED
## Model 3: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC
##   #Df LogLik Df  Chisq Pr(>Chisq)
## 1     8 -479.70
## 2     7 -479.73 -1  0.0531  0.81770
## 3     6 -482.94 -1  6.4163  0.01131 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Probit Regression Full Model

```
##
## Call: glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##      MARRIED + MALE, family = binomial(link = "probit"), data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
##   0.566210    2.142459   -0.016400   -0.425865   -0.463475
##      LOANPRC    MARRIED1      MALE1
##  -0.008409    0.237894   -0.033267
```

```
##
## Degrees of Freedom: 1968 Total (i.e. Null); 1961 Residual
## Null Deviance: 1475
## Residual Deviance: 958.8 AIC: 974.8
```

Statistical Tests for Individual Predictors

```
## Wald test for GDLIN
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = binomial(link = "probit"), data = data1)
## F = 312.1492 on 1 and 1961 df: p= < 0.000000000000000222
```

```
## Wald test for OBRAT
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = binomial(link = "probit"), data = data1)
## F = 9.41885 on 1 and 1961 df: p= 0.0021771
```

```
## Wald test for BLACK
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = binomial(link = "probit"), data = data1)
## F = 11.25188 on 1 and 1961 df: p= 0.00081057
```

```
## Wald test for HISPAN
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = binomial(link = "probit"), data = data1)
## F = 8.034928 on 1 and 1961 df: p= 0.0046354
```

```
## Wald test for MALE
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = binomial(link = "probit"), data = data1)
## F = 0.0795587 on 1 and 1961 df: p= 0.77793
```

```
## Wald test for LOANPRC
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = binomial(link = "probit"), data = data1)
## F = 10.54337 on 1 and 1961 df: p= 0.0011857
```

```
## Wald test for MARRIED
## in glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED + MALE, family = binomial(link = "probit"), data = data1)
## F = 6.123722 on 1 and 1961 df: p= 0.013422
```

Probit Regression Reduced Model 1

```
##
## Call: glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
## MARRIED, family = binomial(link = "probit"), data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
## 0.541687    2.143873   -0.016411   -0.422677   -0.461722
## LOANPRC      MARRIED1
## -0.008386    0.228926
```

```
##
## Degrees of Freedom: 1968 Total (i.e. Null); 1962 Residual
## Null Deviance: 1475
## Residual Deviance: 958.9 AIC: 972.9

## Overall
## GDLIN1 17.691026
## OBRAT 3.070619
## BLACK1 3.337594
## HISPAN1 2.825314
## LOANPRC 3.239942
## MARRIED1 2.530275

## 11h 11hNull G2 McFadden r2ML
## -479.4603569 -737.7148785 516.5090432 0.3500736 0.2307356
## r2CU
## 0.4375655
```

Probit Regression Reduced Model 2

```
##
## Call: glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC,
## family = binomial(link = "probit"), data = data1)
##
## Coefficients:
## (Intercept) GDLIN1 OBRAT BLACK1 HISPAN1
## 0.708006 2.135031 -0.017003 -0.426583 -0.438093
## LOANPRC
## -0.008356
##
## Degrees of Freedom: 1968 Total (i.e. Null); 1963 Residual
## Null Deviance: 1475
## Residual Deviance: 965.3 AIC: 977.3

## Overall
## GDLIN1 17.692269
## OBRAT 3.183979
## BLACK1 3.379077
## HISPAN1 2.687345
## LOANPRC 3.247739

## 11h 11hNull G2 McFadden r2ML
## -482.6350057 -737.7148785 510.1597458 0.3457703 0.2282510
## r2CU
## 0.4328537
```

Model comparison

```
## Analysis of Deviance Table
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC + MARRIED +
## MALE
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC + MARRIED
## Model 3: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC
```

```
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      1961      958.84
## 2      1962      958.92 -1  -0.0799  0.77742
## 3      1963      965.27 -1  -6.3493  0.01174 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Likelihood ratio test
##
## Model 1: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC + MARRIED +
##       MALE
## Model 2: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC + MARRIED
## Model 3: APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1    8 -479.42
## 2    7 -479.46 -1  0.0799   0.77742
## 3    6 -482.64 -1  6.3493   0.01174 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

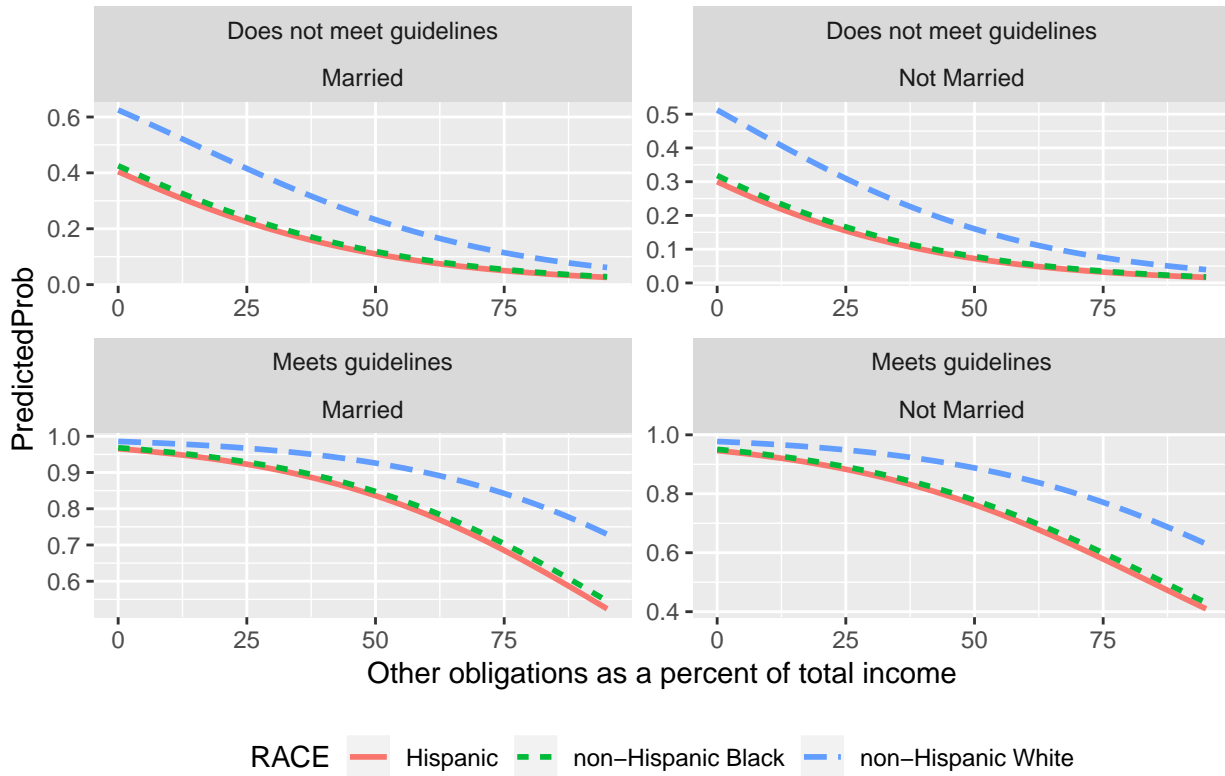
Logit Prediction

Logit Model 1

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = "binomial", data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
##    1.34239      3.72139     -0.03410     -0.81143     -0.89733
##    LOANPRC      MARRIED1
##   -0.01677      0.46093
##
## Degrees of Freedom: 1968 Total (i.e. Null);  1962 Residual
## Null Deviance:      1475
## Residual Deviance: 959.5      AIC: 973.5
```

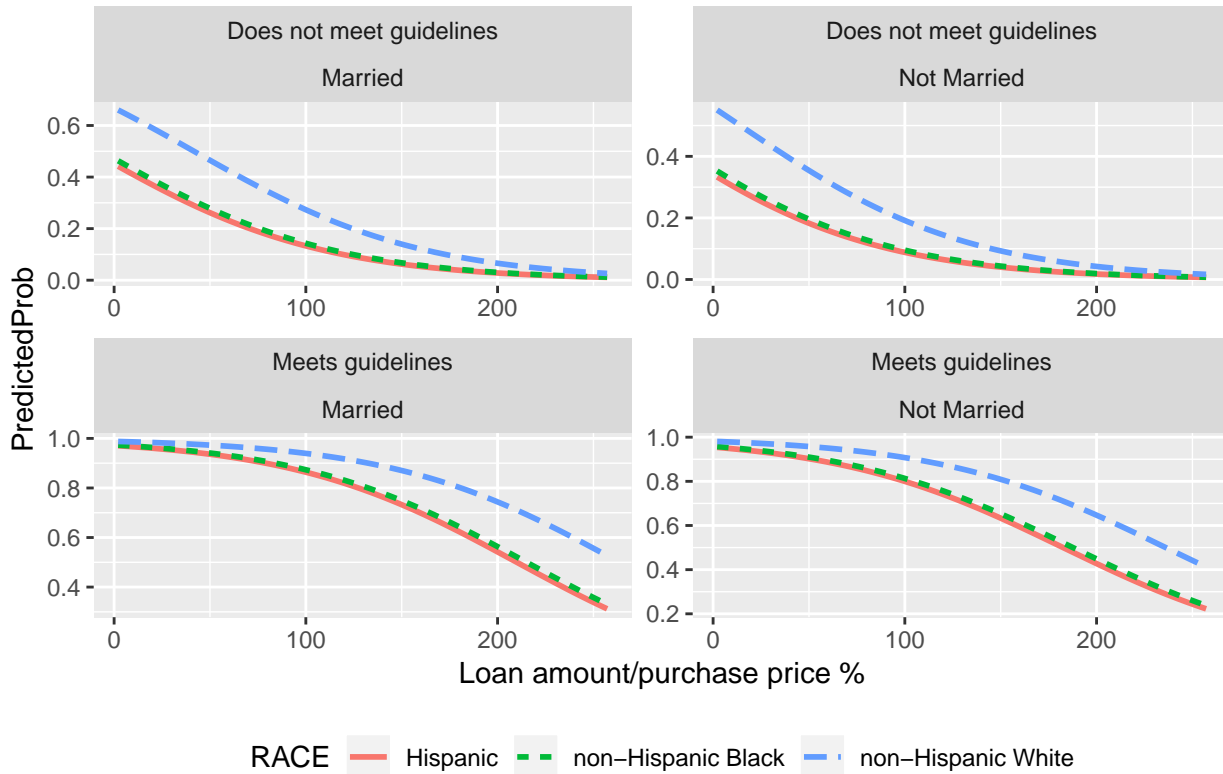
GDLIN	OBRAT	BLACK	HISPAN	MARRIED	LOANPRC	fit	PredictedProb
0	0	1	0	0	77.06418	-0.7616686	0.3182841
0	1	1	0	0	77.06418	-0.7957664	0.3109318
0	2	1	0	0	77.06418	-0.8298643	0.3036738
0	3	1	0	0	77.06418	-0.8639622	0.2965122
0	4	1	0	0	77.06418	-0.8980601	0.2894493
0	5	1	0	0	77.06418	-0.9321579	0.2824871

Predicted probabilities



GDLIN	OBRAT	BLACK	HISPAN	MARRIED	LOANPRC	fit	PredictedProb
0	32.38382	1	0	0	2.105000	-0.6085669	0.3523862
0	32.38382	1	0	0	4.681141	-0.6517776	0.3425891
0	32.38382	1	0	0	7.257283	-0.6949883	0.3329243
0	32.38382	1	0	0	9.833424	-0.7381990	0.3233981
0	32.38382	1	0	0	12.409566	-0.7814096	0.3140162
0	32.38382	1	0	0	14.985707	-0.8246203	0.3047838

Predicted probabilities

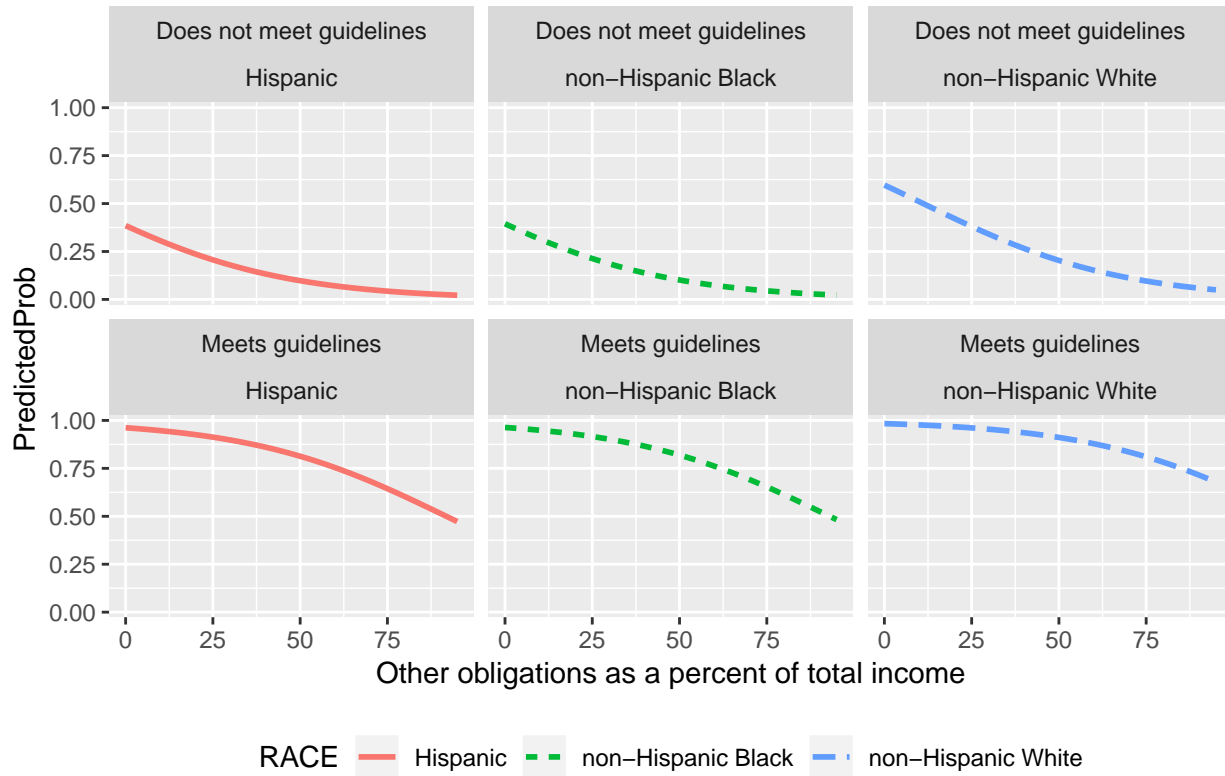


Logit Model 2

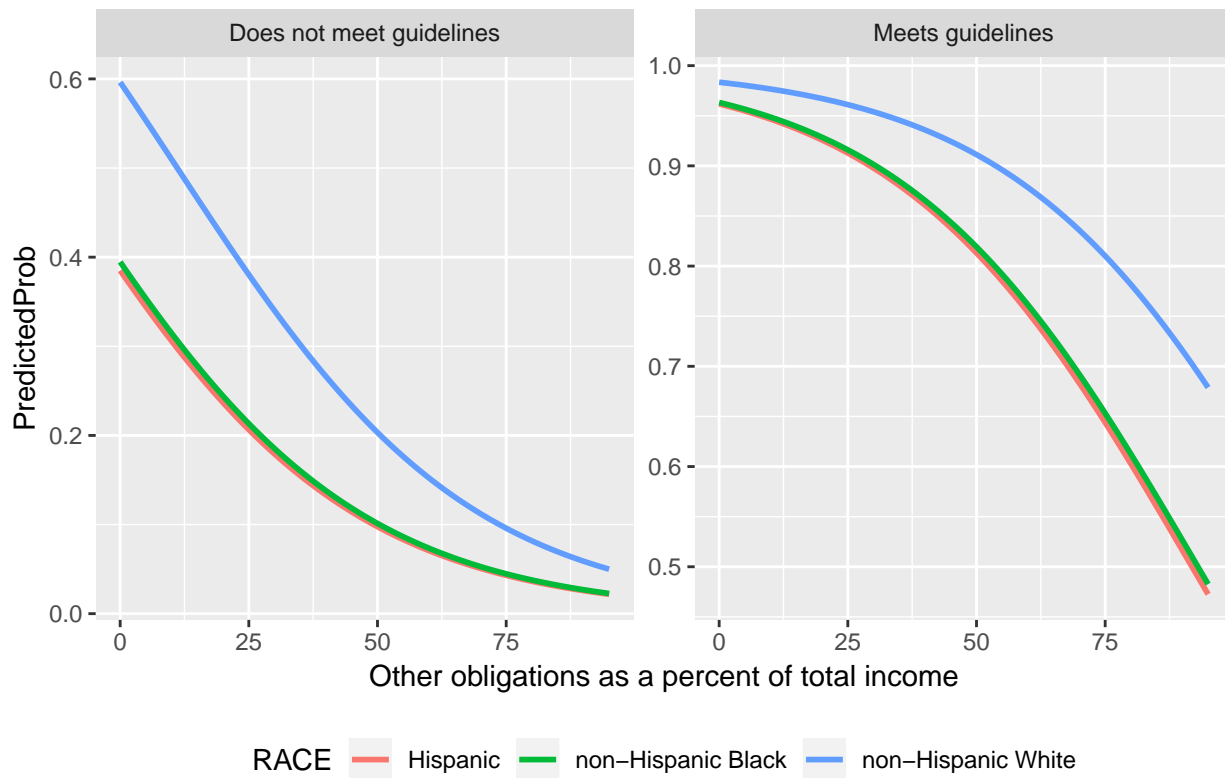
```
##
## Call: glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC,
##   family = "binomial", data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
##    1.66805     3.69461    -0.03512    -0.81768    -0.85828
##   LOANPRC
##   -0.01659
##
## Degrees of Freedom: 1968 Total (i.e. Null);  1963 Residual
## Null Deviance:      1475
## Residual Deviance: 965.9    AIC: 977.9
```

GDLIN	OBRAT	BLACK	HISPAN	LOANPRC	fit	PredictedProb
0	0	1	0	77.06418	-0.4277603	0.3946613
0	1	1	0	77.06418	-0.4628774	0.3863034
0	2	1	0	77.06418	-0.4979945	0.3780121
0	3	1	0	77.06418	-0.5331116	0.3697914
0	4	1	0	77.06418	-0.5682287	0.3616456
0	5	1	0	77.06418	-0.6033458	0.3535786

Predicted probabilities

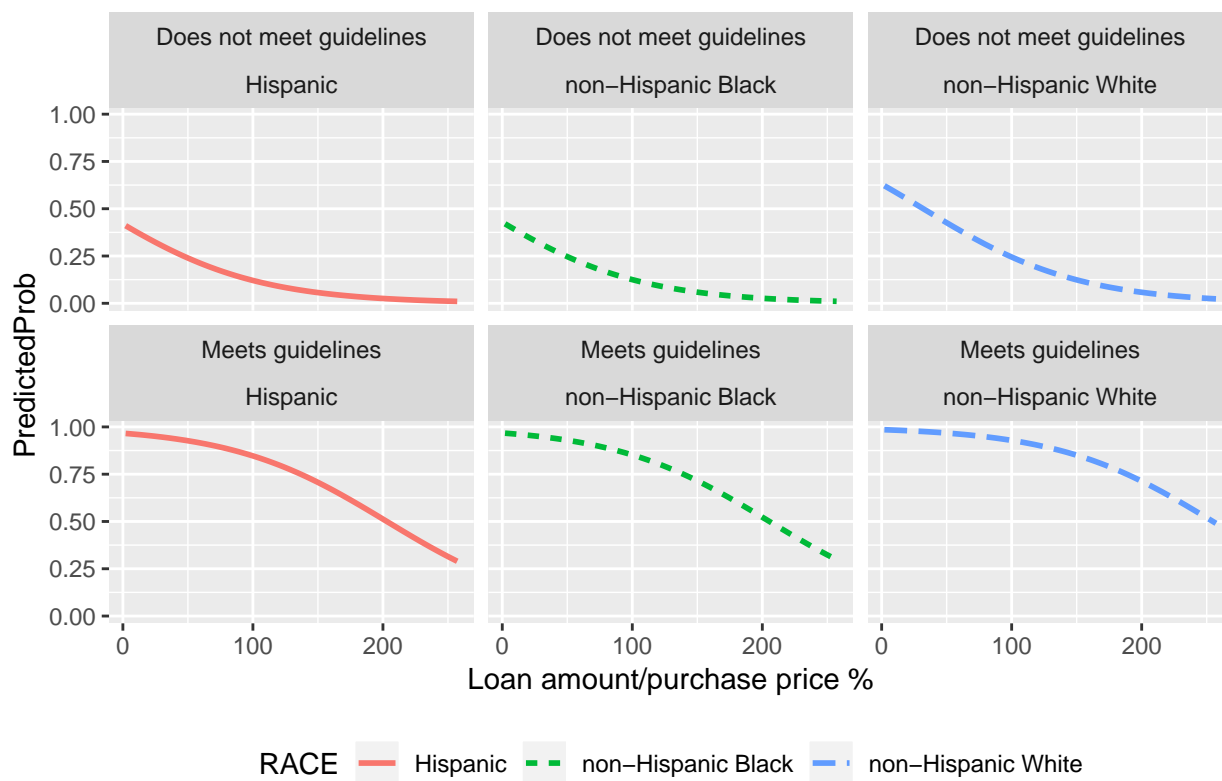


Predicted probabilities

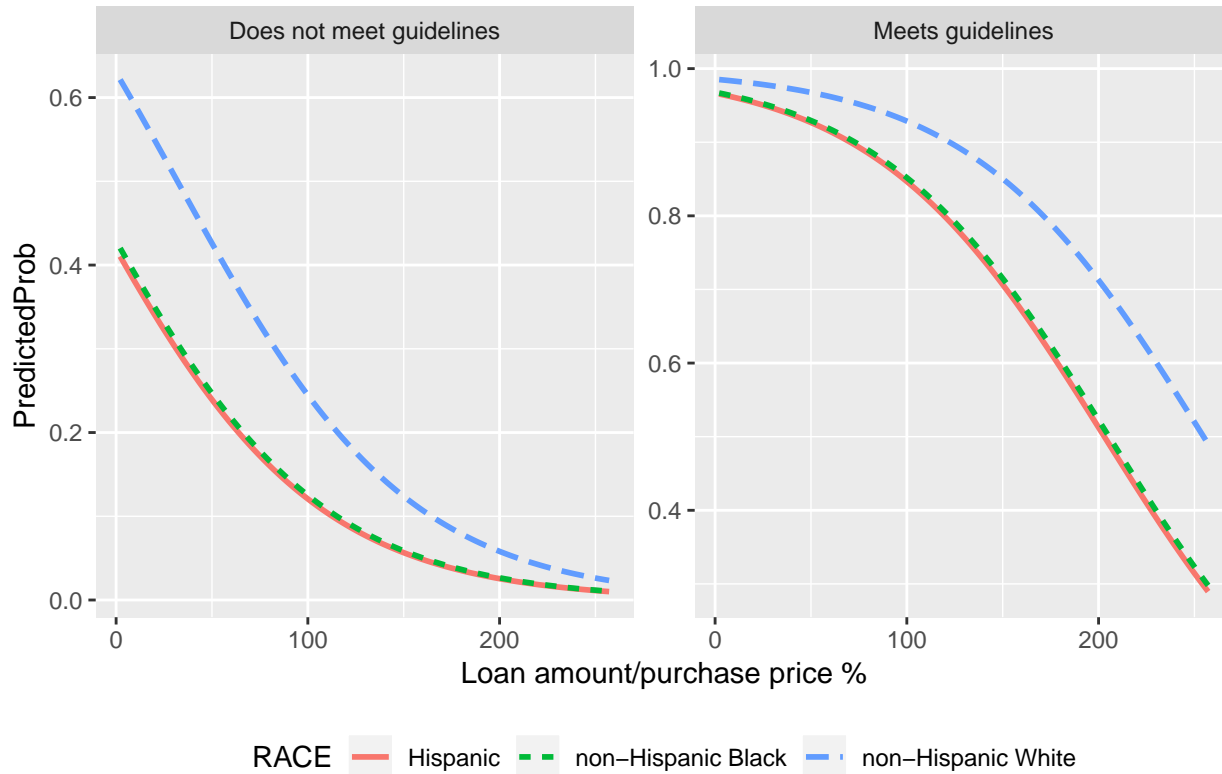


GDLIN	OBRAT	BLACK	HISPAN	LOANPRC	fit	PredictedProb
0	32.38382	1	0	2.105000	-0.3217734	0.4202436
0	32.38382	1	0	4.681141	-0.3644992	0.4098709
0	32.38382	1	0	7.257283	-0.4072250	0.3995777
0	32.38382	1	0	9.833424	-0.4499508	0.3893725
0	32.38382	1	0	12.409566	-0.4926766	0.3792632
0	32.38382	1	0	14.985707	-0.5354024	0.3692578

Predicted probabilities



Predicted probabilities



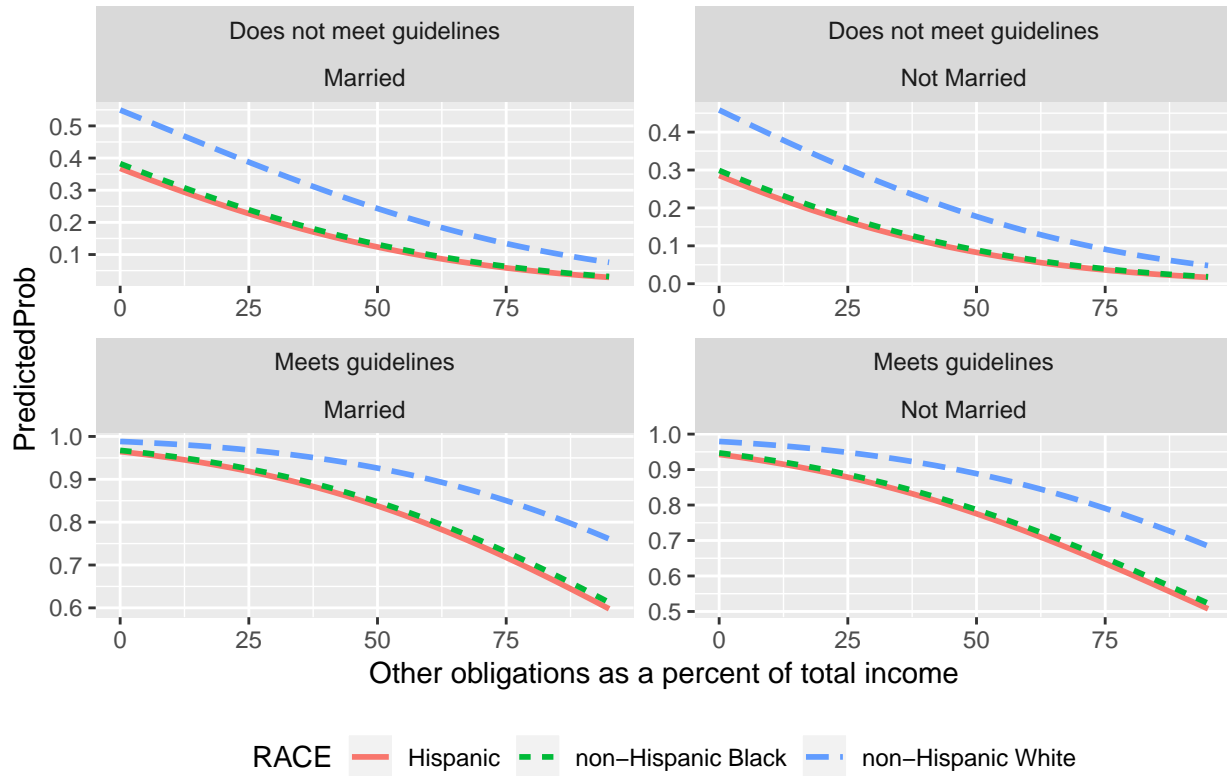
Probit Prediction

Probit Model 1

```
##
## Call:  glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC +
##       MARRIED, family = binomial(link = "probit"), data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
##   0.541687    2.143873   -0.016411   -0.422677   -0.461722
##   LOANPRC    MARRIED1
##  -0.008386    0.228926
##
## Degrees of Freedom: 1968 Total (i.e. Null);  1962 Residual
## Null Deviance:      1475
## Residual Deviance: 958.9    AIC: 972.9
```

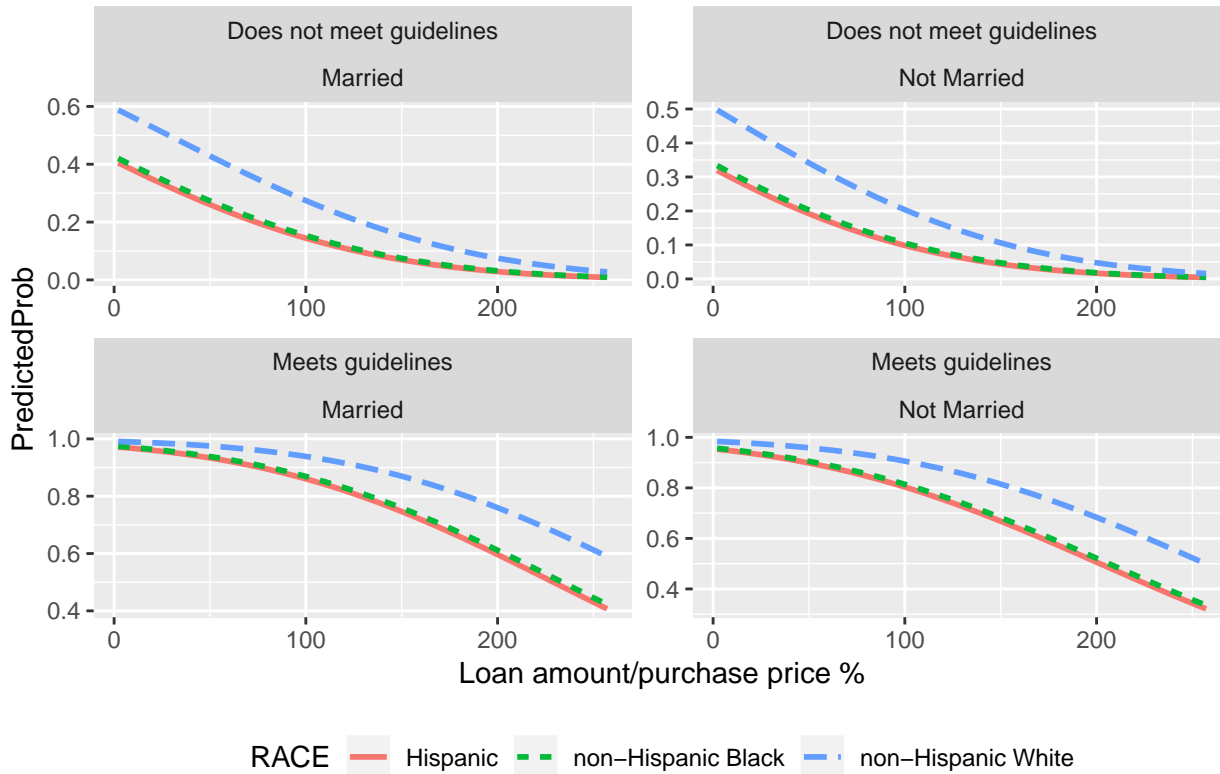
GDLIN	OBRAT	BLACK	HISPAN	MARRIED	LOANPRC	fit	se.fit
0	0	1	0	0	77.06418	0.2989992	0.0861630
0	1	1	0	0	77.06418	0.2933267	0.0839904
0	2	1	0	0	77.06418	0.2877046	0.0818437
0	3	1	0	0	77.06418	0.2821340	0.0797245
0	4	1	0	0	77.06418	0.2766158	0.0776346
0	5	1	0	0	77.06418	0.2711510	0.0755755

Predicted probabilities



GDLIN	OBRAT	BLACK	HISPAN	MARRIED	LOANPRC
0	32.38382	1	0	0	2.105000
0	32.38382	1	0	0	4.681141
0	32.38382	1	0	0	7.257283
0	32.38382	1	0	0	9.833424
0	32.38382	1	0	0	12.409566
0	32.38382	1	0	0	14.985707

Predicted probabilities

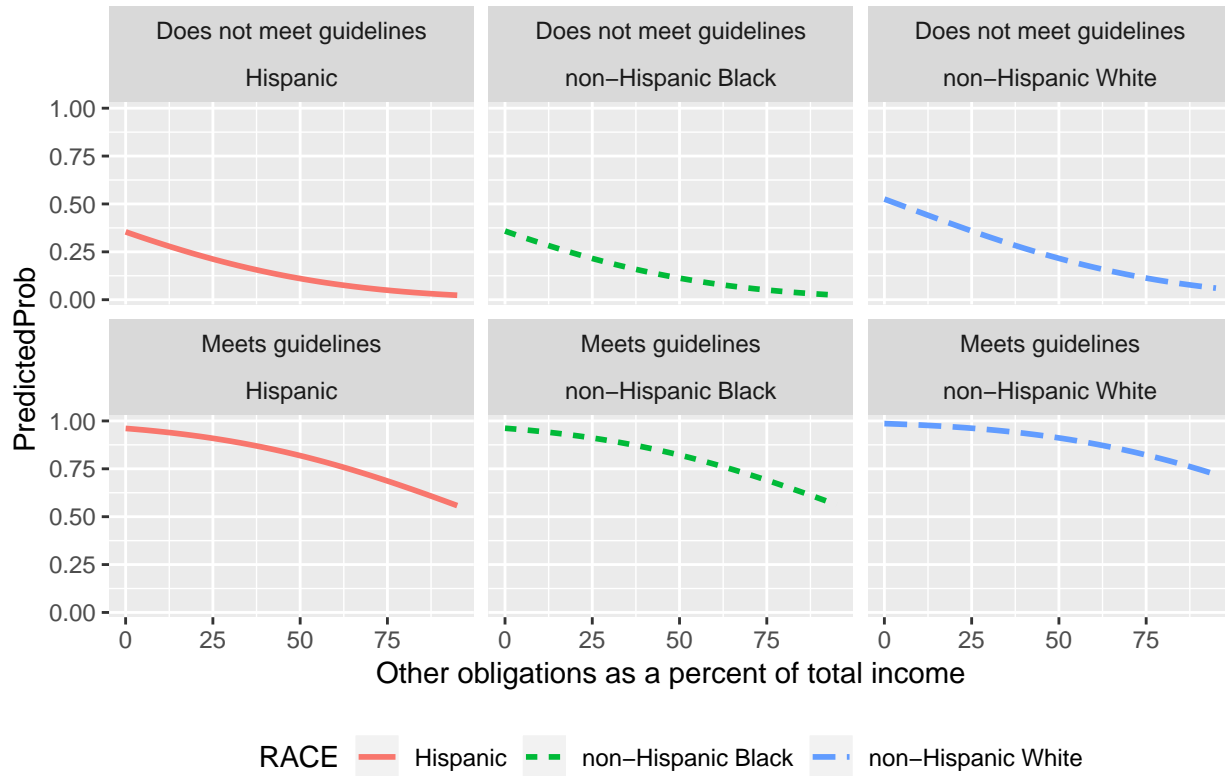


Probit Model 2

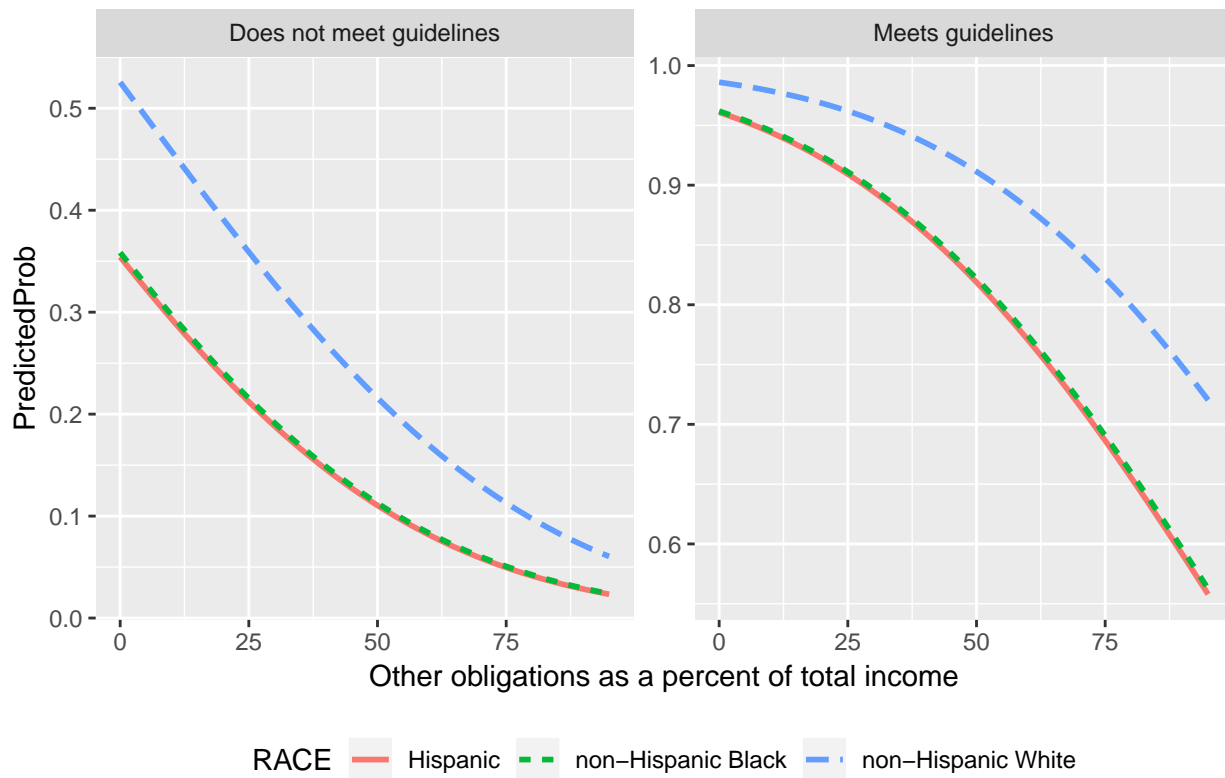
```
##
## Call: glm(formula = APPROVE ~ GDLIN + OBRAT + BLACK + HISPAN + LOANPRC,
##   family = binomial(link = "probit"), data = data1)
##
## Coefficients:
## (Intercept)      GDLIN1      OBRAT      BLACK1      HISPAN1
##   0.708006    2.135031   -0.017003   -0.426583   -0.438093
##   LOANPRC
##  -0.008356
##
## Degrees of Freedom: 1968 Total (i.e. Null);  1963 Residual
## Null Deviance:      1475
## Residual Deviance: 965.3    AIC: 977.3
```

GDLIN	OBRAT	BLACK	HISPAN	LOANPRC	fit
0	0	1	0	77.06418	0.3584733
0	1	1	0	77.06418	0.3521414
0	2	1	0	77.06418	0.3458502
0	3	1	0	77.06418	0.3396013
0	4	1	0	77.06418	0.3333962
0	5	1	0	77.06418	0.3272363

Predicted probabilities

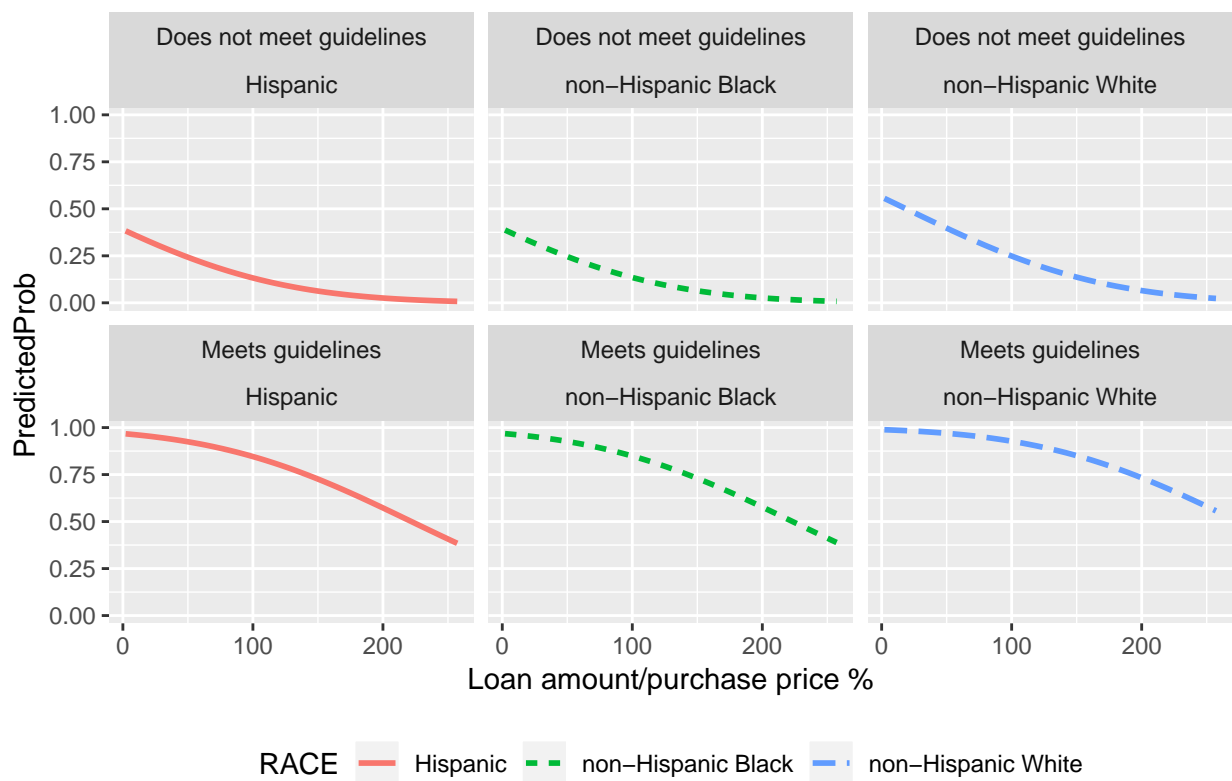


Predicted probabilities



GDLIN	OBRAT	BLACK	HISPAN	LOANPRC	fit
0	32.38382	1	0	2.105000	0.3871373
0	32.38382	1	0	4.681141	0.3789214
0	32.38382	1	0	7.257283	0.3707598
0	32.38382	1	0	9.833424	0.3626560
0	32.38382	1	0	12.409566	0.3546132
0	32.38382	1	0	14.985707	0.3466347

Predicted probabilities



Predicted probabilities

