```
  ___  ____  ____  ____  ____ ®
 /__    /   ____/   /   ____/      18.0
___/   /   /___/   /   /___/       SE—Standard Edition

 Statistics and Data Science       Copyright 1985-2023 StataCorp LLC
                                   StataCorp
                                   4905 Lakeway Drive
                                   College Station, Texas 77845 USA
                                   800-STATA-PC        https://www.stata.com
                                   979-696-4600        stata@stata.com

Stata license: Unlimited-user network, expiring 20 Mar 2025
Serial number: 401809310349
  Licensed to: Center for Information Technology
               University of Groningen

Notes:
      1. Unicode is supported; see help unicode_advice.
      2. Maximum number of variables is set to 5,000 but can be increased; see help set_maxvar.
      3. New update available; type -update all-

 1 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

 2 . /* Stata Code for Data Cleaning, Exploration, and Econometric Analysis */
 3 .
 4 .
 5 . /* Step 1: Load Data */
 6 . clear all

 7 . set more off

 8 .
 9 . use "/Users/uli/Documents/Stata/Project/Data/micro_idn.dta", clear

10 . summarize
```

Variable	Obs	Mean	Std. dev.	Min	Max
economy	0				
economycode	0				
wpid_random	1,062	1.61e+08	2.85e+07	1.11e+08	2.11e+08
wgt	1,062	1	.650534	.3013038	2.775282
female	1,062	1.433145	.4957437	1	2
age	1,062	38.39077	14.44704	15	85
educ	1,062	1.735405	.5251891	1	3
inc_q	1,062	3.200565	1.435473	1	5
emp_in	1,062	1.369115	.4827925	1	2
urbanicit~2f	1,062	1.637476	.4809553	1	2
account	1,062	.547081	.498013	0	1
account_fin	1,062	.5376648	.4988142	0	1
account_mob	1,062	.0856874	.2800337	0	1
fin1_1a	502	1.613546	.4874224	1	2
fin1_1b	502	1.663347	.4730366	1	2
fin2	1,062	1.611111	.4877277	1	2
fin4	413	1.615012	.4871826	1	2
fin4a	159	1.553459	.4987046	1	2
fin5	523	1.759082	.4325022	1	3
fin6	523	1.814532	.3939425	1	3
fin7	523	1.948375	.229968	1	3
fin8	28	1.5	.6938887	1	4
fin8a	16	1.5	.5163978	1	2
fin8b	28	1.714286	.7126966	1	4
fin9	523	1.525813	.5074192	1	3
fin9a	250	1.668	.4718757	1	2

fin10	523	1.470363	.5034186	1	3
fin10a	278	1.456835	.4990316	1	2
fin10b	523	1.397706	.4898926	1	2
fin11_1	481	1.72973	.5100344	1	3
fin11a	539	1.615955	.5055506	1	3
fin11b	539	1.61039	.5422345	1	3
fin11c	539	1.717996	.4585727	1	3
fin11d	539	1.90538	.3507146	1	3
fin11e	539	1.935065	.2947095	1	3
fin11f	539	1.28757	.4691738	1	3
fin11g	539	1.619666	.5046835	1	3
fin11h	539	1.6141	.5132679	1	3
fin13a	74	1.351351	.4806512	1	2
fin13b	74	1.621622	.4882932	1	2
fin13c	74	1.945946	.2276679	1	2
fin13d	75	1.346667	.4791133	1	2
fin14_1	1,062	1.93597	.2487404	1	3
fin14_2	118	1.288136	.4732448	1	3
fin14_2_Ch~a	0				
fin14a	1,062	1.923729	.2655566	1	2
fin14a1	1,062	1.926554	.2681159	1	3
fin14b	1,062	1.814501	.388885	1	2
fin14c	197	1.852792	.5284313	1	3
fin14c_2	197	1.284264	.4522124	1	2
fin14c_2_C~a	0				
fin16	1,062	1.723164	.4476456	1	2
fin17a	1,062	1.777778	.4204431	1	3
fin17a1	103	1.699029	.460923	1	2
fin17b	1,062	1.8258	.3794597	1	2
fin20	1,062	1.897363	.3036266	1	2
fin22a	1,062	1.881356	.3264218	1	3
fin22b	1,062	1.710923	.4658491	1	4
fin22c	185	1.524324	.5007632	1	2
fin24	1,062	2.870998	1.801488	1	9
fin24a	988	1.742915	.6712359	1	4
fin24b	988	1.381579	.6230499	1	5
fin26	1,062	1.84275	.366787	1	3
fin27_1	168	1.52381	.5009259	1	2
fin27c1	88	1.443182	.499608	1	2
fin27c2	88	1.693182	.4638161	1	2
fin28	1,062	1.780603	.4163035	1	3
fin29_1	234	1.58547	.4936967	1	2
fin29c1	137	1.416058	.4947123	1	2
fin29c2	137	1.781022	.4150714	1	2
fin30	1,062	1.236347	.4250376	1	2
fin31a	811	1.866831	.3399669	1	2
fin31b	811	1.92725	.2598855	1	2
fin31b1	127	1.299213	.4597266	1	2
fin31b1_Ch~a	0				
fin31c	684	1.033626	.1803957	1	2
fin32	1,062	1.724105	.4471745	1	2
fin33	293	1.866894	.3402702	1	2
fin34a	293	1.737201	.440907	1	2
fin34b	293	1.94198	.2483759	1	3
fin34d	210	1.085714	.2806106	1	2
fin34e	18	1.777778	.4277926	1	2
fin35	87	1.770115	.4231979	1	2
fin37	1,062	1.731638	.4517397	1	3
fin38	1,062	1.990584	.1059308	1	3

fin39a	297	1.612795	.4948088	1	3
fin39b	297	1.962963	.1891713	1	2
fin39d	179	1.318436	.4671763	1	2
fin39e	57	1.701754	.4615545	1	2
fin42	1,062	1.811676	.3959446	1	3
fin42a	202	1.09901	.3155963	1	3
fin43a	202	1.940594	.2369702	1	2
fin43b	202	1.980198	.1396654	1	2
fin43d	188	1.058511	.2353332	1	2
fin43e	11	2	0	2	2
fin44a	1,062	1.606403	.759158	1	5
fin44b	1,062	1.517891	.7663764	1	5
fin44c	1,062	1.746704	.8356249	1	5
fin44d	1,062	1.869115	.9183474	1	5
fin45	928	2.474138	1.202918	1	5
fin45_1	1,062	1.386064	.6553727	1	5
fin45_1_Ch~a	0				
saved	1,062	.5357815	.498953	0	1
borrowed	1,062	.4246704	.4945258	0	1
receive_wa~s	1,062	3.379473	1.043842	1	4
receive_tr~s	1,062	3.372881	1.098512	1	5
receive_pe~n	1,062	3.975518	.2593464	1	5
receive_ag~e	1,062	3.618644	.8079196	1	5
pay_utilit~s	1,062	2.374765	.9731764	1	4
remittances	1,062	4.149718	1.45986	1	6
mobileowner	1,062	1.248588	.4536717	1	4
internetac~s	1,062	1.474576	.5014715	1	3
anydigpaym~t	1,062	.3907721	.4881533	0	1
merchantpa~g	1,062	.1299435	.3363997	0	1

```
11 .
12 .
13 . /* Step 2: Keep Only Relevant Columns */
14 . keep female age educ inc_q emp_in urbanicity_f2f /* Demographic Variables: */ account_fin account_mob fin5 fin7 fin8 fin8a fin8b f
   > in13c fin20 fin22a fin22b fin22c borrowed fin24 fin24a /* Financial Access & Usage */ mobileowner internetaccess anydigpayment mer
   > chantpay_dig /* Technology & Digital Finance */

15 .
16 .
17 . /* Step 3: Data Cleaning - Renaming Columns */
18 . rename female gender /* 1:female, 3:male */

19 . rename educ education_level

20 . rename inc_q income_quintile

21 . rename emp_in employment_status

22 . rename urbanicity_f2f rural_residence

23 . rename fin5 mobile_access_account

24 . rename fin7 has_credit_card

25 . rename fin8 used_credit_card

26 . rename fin8a used_credit_card_instore

27 . rename fin8b paid_credit_card_full

28 . rename fin13c borrowed_mobile_money

29 . rename fin20 borrowed_medical_purpose
```

```
. rename fin22a borrowed_financial_institution

. rename fin22b borrowed_family_friends

. rename fin22c borrowed_savings_club

. rename fin24 main_source_emergency_funds

. rename fin24a difficulty_emergency_funds /*in 30 days*/

. rename mobileowner owns_mobile_phone

. rename internetaccess internet_access

. rename anydigpayment made_digital_payment

. rename merchantpay_dig digital_merchant_payment

.
. /* Save new dataset */
. save "borrow_behavior_clean.dta", replace
file borrow_behavior_clean.dta saved

.
.
. /* Step 4: Summary Statistics */
. summarize
```

Variable	Obs	Mean	Std. dev.	Min	Max
gender	1,062	1.433145	.4957437	1	2
age	1,062	38.39077	14.44704	15	85
education_~l	1,062	1.735405	.5251891	1	3
income_qui~e	1,062	3.200565	1.435473	1	5
employment~s	1,062	1.369115	.4827925	1	2
rural_resi~e	1,062	1.637476	.4809553	1	2
account_fin	1,062	.5376648	.4988142	0	1
account_mob	1,062	.0856874	.2800337	0	1
mobile_acc~t	523	1.759082	.4325022	1	3
has_credit~d	523	1.948375	.229968	1	3
used_credi~d	28	1.5	.6938887	1	4
used_credi~e	16	1.5	.5163978	1	2
paid_credi~l	28	1.714286	.7126966	1	4
borrowed_m~y	74	1.945946	.2276679	1	2
borrowed_m~e	1,062	1.897363	.3036266	1	2
borrowed_f~n	1,062	1.881356	.3264218	1	3
borrowed_f~s	1,062	1.710923	.4658491	1	4
borrowed_s~b	185	1.524324	.5007632	1	2
main_sourc~s	1,062	2.870998	1.801488	1	9
difficulty~s	988	1.742915	.6712359	1	4
borrowed	1,062	.4246704	.4945258	0	1
owns_mobil~e	1,062	1.248588	.4536717	1	4
internet_a~s	1,062	1.474576	.5014715	1	3
made_digit~t	1,062	.3907721	.4881533	0	1
digital_me~t	1,062	.1299435	.3363997	0	1

```
.
.
. /* Step 5: Check missing values */
. misstable summarize
```

				Obs<.		
Variable	Obs=.	Obs>.	Obs<.	Unique values	Min	Max
mobile_acc~t	539		523	3	1	3

has_credit~d	**539**	**523**	3	1	3
used_credi~d	**1,034**	**28**	3	1	4
used_credi~e	**1,046**	**16**	2	1	2
paid_credi~l	**1,034**	**28**	4	1	4
borrowed_m~y	**988**	**74**	2	1	2
borrowed_s~b	**877**	**185**	2	1	2
difficulty~s	**74**	**988**	4	1	4

```
50 .
51 .
52 . /* Step 6: Handle Missing Values */
53 . // replace borrowed = 0 if missing(borrowed)
54 . foreach var of varlist gender age education_level income_quintile employment_status rural_residence account_fin account_mob mobile
   > _access_account has_credit_card used_credit_card used_credit_card_instore paid_credit_card_full borrowed_mobile_money borrowed_med
   > ical_purpose borrowed_financial_institution borrowed_family_friends borrowed_savings_club owns_mobile_phone internet_access made_d
   > igital_payment digital_merchant_payment {
  2.     capture confirm numeric variable `var'
  3.     if _rc == 0 {
  4.         replace `var' = . if `var' == -999   // Handle missing values coded as -999
  5.     }
  6.     else {
  7.         replace `var' = "" if `var' == " "   // Handle missing values for string variables
  8.     }
  9. }
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)
(0 real changes made)

55 .
56 . /* Step 7: Generate Key Variables */
57 . * Create a binary variable for borrowing (1 if borrowed from any source, 0 otherwise)
58 . gen borrowed_binary = (borrowed_financial_institution == 1 | borrowed_family_friends == 1 | borrowed_savings_club == 1 | borrowed_
   > mobile_money == 1)

59 .
60 .
61 . /* Step 8: Exploratory Data Analysis */
62 . tabulate owns_mobile_phone borrowed_binary, chi2
```

Owns a mobile phone	borrowed_binary 0	1	Total
yes	**493**	**312**	**805**
no	**174**	**79**	**253**
(dk)	**1**	**0**	**1**
(ref)	**3**	**0**	**3**
Total	**671**	**391**	**1,062**

```
           Pearson chi2(3) =   7.0349   Pr = 0.071

63 . tabulate borrowed_binary internet_access, chi2

borrowed_b |         Internet access
     inary |       yes         no       (dk) |     Total
-----------+---------------------------------+----------
         0 |       335        335          1 |       671
         1 |       224        167          0 |       391
-----------+---------------------------------+----------
     Total |       559        502          1 |     1,062

           Pearson chi2(2) =   5.8478   Pr = 0.054

64 . tabstat borrowed owns_mobile_phone internet_access, statistics(mean sd)

   Stats |  borrowed  owns_m~e  intern~s
---------+------------------------------
    Mean |  .4246704  1.248588  1.474576
      SD |  .4945258  .4536717  .5014715
----------------------------------------

65 .
66 . graph bar (mean) borrowed_binary, over(owns_mobile_phone) title("Borrowing Rate by Mobile Ownership")

67 . graph bar (mean) borrowed_binary, over(internet_access) title("Borrowing Rate by Internet Access")

68 .
69 . /* Step 9: Correlation Matrix */
70 . pwcorr borrowed_binary owns_mobile_phone internet_access, sig

             | borro~ry owns_m~e intern~s
-------------+---------------------------
borrowed_b~y |   1.0000
             |
             |
owns_mobil~e |  -0.0783   1.0000
             |   0.0106
             |
internet_a~s |  -0.0723   0.5208   1.0000
             |   0.0185   0.0000
             |

71 .
72 . * Correlation matrix
73 . correlate owns_mobile_phone internet_access account_mob
(obs=1,062)

             | owns_m~e intern~s accoun~b
-------------+---------------------------
owns_mobil~e |   1.0000
internet_a~s |   0.5208   1.0000
 account_mob |  -0.1381  -0.2563   1.0000


74 .
75 . /* Step 10: OLS Regression: Borrowing behavior as a function of mobile phone ownership and internet access */
76 . reg borrowed_binary owns_mobile_phone internet_access education_level income_quintile employment_status rural_residence

      Source |       SS           df       MS      Number of obs   =     1,062
-------------+----------------------------------   F(6, 1055)      =      3.99
       Model |  5.48141951         6  .913569918   Prob > F        =    0.0006
    Residual |  241.562837     1,055  .228969513   R-squared       =    0.0222
-------------+----------------------------------   Adj R-squared   =    0.0166
       Total |  247.044256     1,061  .232840958   Root MSE        =    .47851

------------------------------------------------------------------------------
  borrowed_binary | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
------------------+---------------------------------------------------------------
owns_mobile_phone |  -.0556913   .0387393    -1.44   0.151    -.1317061    .0203235
```

```
      internet_access |  -.0740118   .0369278    -2.00   0.045    -.1464722   -.0015514
      education_level |  -.0348962   .0323173    -1.08   0.280    -.0983098    .0285174
      income_quintile |  -.0205866   .0108906    -1.89   0.059    -.0419563    .0007832
    employment_status |  -.1028953   .0308128    -3.34   0.001    -.1633567   -.0424338
      rural_residence |  -.0226604   .0311521    -0.73   0.467    -.0837876    .0384667
                _cons |   .8512738   .1213018     7.02   0.000     .6132535    1.089294
------------------------------------------------------------------------------

77 .
   end of do-file

78 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

79 . /* Step 11: Logistic Regression: Probability of borrowing */
80 . logit borrowed_binary owns_mobile_phone internet_access age education_level income_quintile employment_status rural_residence

   Iteration 0:  Log likelihood = -698.77083
   Iteration 1:  Log likelihood = -686.21465
   Iteration 2:  Log likelihood = -686.17481
   Iteration 3:  Log likelihood =  -686.1748

   Logistic regression                                     Number of obs =  1,062
                                                           LR chi2(7)    =  25.19
                                                           Prob > chi2   = 0.0007
   Log likelihood = -686.1748                              Pseudo R2     = 0.0180

------------------------------------------------------------------------------
  borrowed_binary | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
------------------------------------------------------------------------------
owns_mobile_phone |  -.2921518   .1810272    -1.61   0.107    -.6469586    .0626549
  internet_access |  -.3724162   .1698976    -2.19   0.028    -.7054095    -.039423
              age |   .0060166   .0055043     1.09   0.274    -.0047717    .0168048
  education_level |  -.1108203   .1484026    -0.75   0.455    -.4016841    .1800435
  income_quintile |  -.0955784   .0479694    -1.99   0.046    -.1895967   -.0015601
employment_status |  -.4319033   .1391053    -3.10   0.002    -.7045447   -.1592619
  rural_residence |    -.12141   .1376071    -0.88   0.378     -.391115     .148295
            _cons |   1.416825   .5600531     2.53   0.011     .3191413    2.514509
------------------------------------------------------------------------------

81 .
   end of do-file

82 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

83 . /*
   > ROC Curve (Receiver Operating Characteristic Curve) : evaluates the model's classification performance,
   > to plots the True Positive Rate (Sensitivity) against the False Positive Rate (1 - Specificity).
   > */
84 .
85 . * Predict probabilities
86 . predict pred_probs
   (option pr assumed; Pr(borrowed_binary))

87 .
88 . * Generate ROC Curve:
89 . roctab borrowed pred_probs, graph

90 .
   end of do-file

91 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

92 .
93 . * Alternative command for a smoother ROC curve
94 . lroc, title("ROC Curve for Logistic Regression Model")

   Logistic model for borrowed_binary

   Number of observations =     1062
   Area under ROC curve   =   0.5909
```

```
 95 .
    end of do-file

 96 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph ROC Curve for Logistic Regression Model.gph", replace
    file /Users/uli/Documents/Stata/Project/Graph/Graph ROC Curve for Logistic Regression Model.gph saved

 97 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

 98 . * Create an odds ratio plot
 99 . coefplot, ///
    >     drop(_cons) ///
    >     xlabel(, angle(45)) ///
    >     title("Odds Ratios with 95% Confidence Intervals") ///
    >     xline(1, lcolor(red)) ///
    >     ytitle("Predictors") ///
    >     mcolor(blue) msymbol(O) ///
    >     ciopts(lcolor(black))

100 .
    end of do-file

101 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

102 . * Create an odds ratio plot
103 . coefplot, ///
    >     drop(_cons) ///
    >     xlabel(, angle(45)) ///
    >     title("Odds Ratios with 95% Confidence Intervals") ///
    >     xline(1, lcolor(red)) ///
    >     ytitle("Predictors") ///
    >     mcolor(blue) msymbol(O) ///
    >     ciopts(lcolor(black))

104 .
    end of do-file

105 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph Odds Ratios with 95% Confidence Intervals.gph", replace
    file /Users/uli/Documents/Stata/Project/Graph/Graph Odds Ratios with 95% Confidence Intervals.gph saved

106 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

107 . /* Step 12: OLS Regression Model */
108 . reg borrowed main_source_emergency_funds difficulty_emergency_funds owns_mobile_phone internet_access age education_level income_q
    > uintile employment_status rural_residence

      Source |       SS           df       MS      Number of obs   =       988
-------------+----------------------------------   F(9, 978)       =      3.90
       Model |  8.42952372         9  .936613747   Prob > F        =    0.0001
    Residual |  235.165618       978  .240455642   R-squared       =    0.0346
-------------+----------------------------------   Adj R-squared   =    0.0257
       Total |  243.595142       987  .246803588   Root MSE        =    .49036

------------------------------------------------------------------------------
                    borrowed | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-----------------------------+------------------------------------------------
main_source_emergency_funds |   .0368633   .0123775     2.98   0.003     .0125737    .0611529
 difficulty_emergency_funds |  -.0484668   .0247223    -1.96   0.050    -.0969816    .0000481
          owns_mobile_phone |  -.0691158    .042325    -1.63   0.103    -.1521742    .0139425
            internet_access |  -.0659068   .0408757    -1.61   0.107     -.146121    .0143075
                        age |   .0015126   .0013473     1.12   0.262    -.0011314    .0041565
            education_level |  -.0323657   .0358955    -0.90   0.367    -.1028068    .0380754
            income_quintile |  -.0197237   .0120669    -1.63   0.102    -.0434038    .0039563
          employment_status |   -.086078   .0335502    -2.57   0.010    -.1519167   -.0202392
            rural_residence |   .0021148   .0335685     0.06   0.950    -.0637598    .0679895
                      _cons |   .7911925   .1476902     5.36   0.000     .5013663    1.081019
------------------------------------------------------------------------------

109 .
    end of do-file
```

```
110 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

111 . /* Step 13: Logistic Regression Model */
112 . logit borrowed main_source_emergency_funds difficulty_emergency_funds owns_mobile_phone internet_access age education_level income
    > _quintile employment_status rural_residence

    Iteration 0:  Log likelihood = -678.00397
    Iteration 1:  Log likelihood =  -660.7036
    Iteration 2:  Log likelihood = -660.68375
    Iteration 3:  Log likelihood = -660.68375

    Logistic regression                                     Number of obs =    988
                                                            LR chi2(9)    =  34.64
                                                            Prob > chi2   = 0.0001
    Log likelihood = -660.68375                             Pseudo R2     = 0.0255

    ------------------------------------------------------------------------------
                 borrowed | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
    -------------------------+----------------------------------------------------------------
    main_source_emergency_funds |   .1521726   .0516621     2.95   0.003     .0509167    .2534285
     difficulty_emergency_funds |  -.2036518   .1040303    -1.96   0.050    -.4075475    .0002439
              owns_mobile_phone |  -.2946874   .1823321    -1.62   0.106    -.6520517    .0626769
                internet_access |  -.2753216   .1721246    -1.60   0.110    -.6126796    .0620364
                            age |   .0063472   .0056466     1.12   0.261    -.0047199    .0174142
                education_level |  -.1366553   .1501019    -0.91   0.363    -.4308496    .1575391
                 income_quintile |  -.0818109   .0503004    -1.63   0.104     -.180398    .0167761
              employment_status |  -.3599176   .1410378    -2.55   0.011    -.6363466   -.0834887
                rural_residence |   .0087366   .1403807     0.06   0.950    -.2664045    .2838777
                          _cons |   1.232539   .6196634     1.99   0.047     .0180211    2.447057
    ------------------------------------------------------------------------------

113 .
    end of do-file

114 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

115 . /*
    > ROC Curve (Receiver Operating Characteristic Curve) : evaluates the model's classification performance,
    > to plots the True Positive Rate (Sensitivity) against the False Positive Rate (1 - Specificity).
    > */
116 .
117 . * Predict probabilities
118 . predict pred_probs
    variable pred_probs already defined
    r(110);

    end of do-file

    r(110);

119 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

120 . * Predict probabilities
121 . predict pred_probs_em
    (option pr assumed; Pr(borrowed))
    (74 missing values generated)

122 .
123 . * Generate ROC Curve:
124 . roctab borrowed pred_probs_em, graph

125 .
126 . * Alternative command for a smoother ROC curve
127 . lroc, title("ROC Curve for Logistic Regression Model")

    Logistic model for borrowed

    Number of observations =      988
    Area under ROC curve   =   0.6060
```

```
128 .
    end of do-file

129 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph ROC Curve for Logistic Regression Model with Emergency Fund.gph
    > "
    file /Users/uli/Documents/Stata/Project/Graph/Graph ROC Curve for Logistic Regression Model with Emergency Fund.gph saved

130 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

131 . /* Odds Ratio Plot: visualizes the effect size of each predictor in the logistic regression model. */
132 . * Run Logistic Regression again to store results
133 . logit borrowed main_source_emergency_funds difficulty_emergency_funds owns_mobile_phone internet_access age education_level income
    > _quintile employment_status rural_residence

    Iteration 0:  Log likelihood = -678.00397
    Iteration 1:  Log likelihood =  -660.7036
    Iteration 2:  Log likelihood = -660.68375
    Iteration 3:  Log likelihood = -660.68375

    Logistic regression                                     Number of obs =    988
                                                            LR chi2(9)    =  34.64
                                                            Prob > chi2   = 0.0001
    Log likelihood = -660.68375                             Pseudo R2     = 0.0255

    ------------------------------------------------------------------------------
                borrowed | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
    -------------------------+----------------------------------------------------
    main_source_emergency_funds |   .1521726   .0516621     2.95   0.003     .0509167    .2534285
     difficulty_emergency_funds |  -.2036518   .1040303    -1.96   0.050    -.4075475    .0002439
              owns_mobile_phone |  -.2946874   .1823321    -1.62   0.106    -.6520517    .0626769
                internet_access |  -.2753216   .1721246    -1.60   0.110    -.6126796    .0620364
                            age |   .0063472   .0056466     1.12   0.261    -.0047199    .0174142
                education_level |  -.1366553   .1501019    -0.91   0.363    -.4308496    .1575391
                income_quintile |  -.0818109   .0503004    -1.63   0.104     -.180398    .0167761
              employment_status |  -.3599176   .1410378    -2.55   0.011    -.6363466   -.0834887
                rural_residence |   .0087366   .1403807     0.06   0.950    -.2664045    .2838777
                          _cons |   1.232539   .6196634     1.99   0.047     .0180211    2.447057
    ------------------------------------------------------------------------------

134 .
    end of do-file

135 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

136 . * Create an odds ratio plot
137 . coefplot, ///
    >     drop(_cons) ///
    >     xlabel(, angle(45)) ///
    >     title("Odds Ratios with 95% Confidence Intervals") ///
    >     xline(1, lcolor(red)) ///
    >     ytitle("Predictors") ///
    >     mcolor(blue) msymbol(O) ///
    >     ciopts(lcolor(black))

138 .
139 .
    end of do-file

140 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph Odds Ratios with 95% Confidence Intervals with Emergency Fund.g
    > ph"
    file /Users/uli/Documents/Stata/Project/Graph/Graph Odds Ratios with 95% Confidence Intervals with Emergency Fund.gph saved

141 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

142 . /* Step 14: Instrumental Variable Regression (IV) */
143 . * Assuming 4G network availability as an instrument for mobile phone ownership
144 . generate network_4G_Cov = (owns_mobile_phone == 1 & internet_access == 1)

145 . // Check instrument validity:
```

```
146 . correlate owns_mobile_phone network_4G_Cov
    (obs=1,062)

                 | owns_m~e networ~v
-----------------+------------------
owns_mobil~e     |   1.0000
network_4G~v     |  -0.5682   1.0000


147 .
    end of do-file

148 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

149 . ivregress 2sls borrowed_binary (owns_mobile_phone = network_4G_Cov) internet_access age education_level income_quintile employment
    > _status rural_residence

    Instrumental variables 2SLS regression            Number of obs   =      1,062
                                                      Wald chi2(7)    =      24.83
                                                      Prob > chi2     =     0.0008
                                                      R-squared       =     0.0161
                                                      Root MSE        =     .47841

------------------------------------------------------------------------------
  borrowed_binary | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
------------------+-----------------------------------------------------------
owns_mobile_phone |  -.1714302   .1150442    -1.49   0.136    -.3969127    .0540523
  internet_access |  -.0464133   .0547767    -0.85   0.397    -.1537736     .060947
              age |   .0018886   .0013565     1.39   0.164      -.00077    .0045472
  education_level |  -.0350836   .0353295    -0.99   0.321    -.1043281     .034161
  income_quintile |  -.0224814   .0109751    -2.05   0.041    -.0439922   -.0009706
employment_status |  -.0888149   .0323646    -2.74   0.006    -.1522485   -.0253814
  rural_residence |  -.0227971   .0319045    -0.71   0.475    -.0853287    .0397346
            _cons |   .8699181   .1389064     6.26   0.000     .5976665     1.14217
------------------------------------------------------------------------------
    Endogenous: owns_mobile_phone
    Exogenous:  internet_access age education_level income_quintile employment_status
                rural_residence network_4G_Cov

150 .
    end of do-file

151 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

152 . /* Scatter Plot with Fitted Regression Line */
153 .
154 . twoway (scatter borrowed_binary owns_mobile_phone, mcolor(blue) msymbol(O)) ///
    >        (lfitci borrowed_binary owns_mobile_phone, lcolor(red%50) fintensity(50)), ///
    >        title("Instrumented Mobile Phone Ownership vs Borrowing Behavior") ///
    >        xtitle("Instrumented Mobile Phone Ownership") ///
    >        ytitle("Borrowing Behavior (Binary)") ///
    >        legend(order(1 "Observed Data" 2 "Fitted Regression Line with CI"))

155 .
    end of do-file

156 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph Instrumented Mobile Phone Ownership.gph", replace
    file /Users/uli/Documents/Stata/Project/Graph/Graph Instrumented Mobile Phone Ownership.gph saved

157 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

158 . /* Use jitter to reduce overplotting (useful for binary variables) */
159 . twoway (scatter borrowed_binary owns_mobile_phone, jitter(2) mcolor(blue) msymbol(O)) ///
    >        (lfit borrowed_binary owns_mobile_phone, lcolor(red) lwidth(medium)), ///
    >        title("Instrumented Mobile Phone Ownership vs Borrowing Behavior") ///
    >        xtitle("Instrumented Mobile Phone Ownership") ///
    >        ytitle("Borrowing Behavior (Binary)") ///
    >        legend(order(1 "Observed Data" 2 "Fitted Regression Line"))

160 .
```

```
    end of do-file

161 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph Instrumented Mobile Phone Ownership using Jitter.gph", replace
    file /Users/uli/Documents/Stata/Project/Graph/Graph Instrumented Mobile Phone Ownership using Jitter.gph saved

162 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

163 . /* Step 15: Endogeneity using IV Regression */
164 . * Checking if mobile phone ownership is endogenous
165 . estat endogenous

      Tests of endogeneity
      H0: Variables are exogenous

      Durbin (score) chi2(1)          =   1.0273  (p = 0.3108)
      Wu-Hausman F(1,1053)            =  1.01958  (p = 0.3129)

166 .
    end of do-file

167 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

168 . /* Perform IV Regression and Save Residuals */
169 . * IV Regression: Using 4G network availability as an instrument
170 . ivregress 2sls borrowed_binary (owns_mobile_phone = network_4G_Cov) ///
>           internet_access age education_level income_quintile ///
>           employment_status rural_residence

Instrumental variables 2SLS regression          Number of obs   =      1,062
                                                Wald chi2(7)    =      24.83
                                                Prob > chi2     =     0.0008
                                                R-squared       =     0.0161
                                                Root MSE        =     .47841

------------------------------------------------------------------------------
 borrowed_binary | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
owns_mobile_phone |  -.1714302   .1150442    -1.49   0.136    -.3969127    .0540523
  internet_access |  -.0464133   .0547767    -0.85   0.397    -.1537736     .060947
              age |   .0018886   .0013565     1.39   0.164      -.00077    .0045472
  education_level |  -.0350836   .0353295    -0.99   0.321    -.1043281     .034161
  income_quintile |  -.0224814   .0109751    -2.05   0.041    -.0439922   -.0009706
employment_status |  -.0888149   .0323646    -2.74   0.006    -.1522485   -.0253814
  rural_residence |  -.0227971   .0319045    -0.71   0.475    -.0853287    .0397346
            _cons |   .8699181   .1389064     6.26   0.000     .5976665     1.14217
------------------------------------------------------------------------------
Endogenous: owns_mobile_phone
Exogenous:  internet_access age education_level income_quintile employment_status
            rural_residence network_4G_Cov

171 .
    end of do-file

172 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

173 . * Generate residuals from the IV regression
174 . predict iv_residuals, resid

175 .
    end of do-file

176 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

177 . /* Plot Histogram of Residuals */
178 . * Histogram of IV regression residuals
179 . histogram iv_residuals, normal ///
>     title("Histogram of Residuals from IV Regression") ///
>     xtitle("Residuals") ytitle("Frequency") ///
>     color(blue%60) width(0.1)
    (bin=15, start=-.60565788, width=.1)
```

```
180 .
    end of do-file

181 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph Histogram of Residuals from IV Regression.gph"
    file /Users/uli/Documents/Stata/Project/Graph/Graph Histogram of Residuals from IV Regression.gph saved

182 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

183 . /* (a) Overlay Kernel Density for Smoothness */
184 . histogram iv_residuals, normal kdensity ///
    >     title("Histogram of Residuals from IV Regression") ///
    >     xtitle("Residuals") ytitle("Frequency") ///
    >     color(blue%60) width(0.1) ///
    >     legend(order(2 "Kernel Density" 3 "Normal Distribution"))
    (bin=15, start=-.60565788, width=.1)

185 .
    end of do-file

186 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Overlay Kernel Density for Smoothness.gph"
    file /Users/uli/Documents/Stata/Project/Graph/Overlay Kernel Density for Smoothness.gph saved

187 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

188 . /* (b) Perform Skewness and Kurtosis Tests for Normality */
189 . sktest iv_residuals

    Skewness and kurtosis tests for normality
                                                              ------ Joint test ------
        Variable |       Obs   Pr(skewness)   Pr(kurtosis)   Adj chi2(2)  Prob>chi2
    -------------+-------------------------------------------------------------------
    iv_residuals |     1,062         0.0000             .             .          .

190 .
    end of do-file

191 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

192 . /* (c) Q-Q Plot for Residuals */
193 . qnorm iv_residuals

194 .
    end of do-file

195 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph Q-Q Plot for Residuals.gph", replace
    file /Users/uli/Documents/Stata/Project/Graph/Graph Q-Q Plot for Residuals.gph saved

196 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

197 . * Run OLS regression
198 . reg borrowed_binary owns_mobile_phone internet_access age education_level income_quintile employment_status rural_residence

          Source |       SS           df       MS      Number of obs   =     1,062
    -------------+----------------------------------   F(7, 1054)      =      3.59
           Model |  5.75073114         7   .82153302   Prob > F        =    0.0008
        Residual |  241.293525     1,054  .228931238   R-squared       =    0.0233
    -------------+----------------------------------   Adj R-squared   =    0.0168
           Total |  247.044256     1,061  .232840958   Root MSE        =    .47847

    ------------------------------------------------------------------------------
      borrowed_binary | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
    ------------------+-----------------------------------------------------------
    owns_mobile_phone |  -.0622036   .0391986    -1.59   0.113    -.1391199    .0147126
      internet_access |  -.0857798   .0384858    -2.23   0.026    -.1612974   -.0102622
                  age |   .0013553   .0012496     1.08   0.278    -.0010966    .0038072
      education_level |  -.0244358   .0337231    -0.72   0.469    -.0906079    .0417364
      income_quintile |  -.0219777    .010965    -2.00   0.045    -.0434933    -.000462
    employment_status |  -.0974617   .0312149    -3.12   0.002    -.1587121   -.0362113
      rural_residence |  -.0278444    .031514    -0.88   0.377    -.0896818     .033993
```

```
            _cons |   .8120761   .1265613     6.42   0.000     .5637354    1.060417
------------------------------------------------------------------------------

199 .
    end of do-file

200 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

201 . estimates store ols_model

202 .
    end of do-file

203 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

204 . * Run logistic regression
205 . logit borrowed_binary owns_mobile_phone internet_access age education_level income_quintile employment_status rural_residence

    Iteration 0:  Log likelihood = -698.77083
    Iteration 1:  Log likelihood = -686.21465
    Iteration 2:  Log likelihood = -686.17481
    Iteration 3:  Log likelihood =  -686.1748

    Logistic regression                                     Number of obs =  1,062
                                                            LR chi2(7)    =  25.19
                                                            Prob > chi2   = 0.0007
    Log likelihood = -686.1748                              Pseudo R2     = 0.0180

    ------------------------------------------------------------------------------
      borrowed_binary | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
    ------------------+-----------------------------------------------------------
    owns_mobile_phone |  -.2921518   .1810272    -1.61   0.107    -.6469586    .0626549
      internet_access |  -.3724162   .1698976    -2.19   0.028    -.7054095    -.039423
                  age |   .0060166   .0055043     1.09   0.274    -.0047717    .0168048
      education_level |  -.1108203   .1484026    -0.75   0.455    -.4016841    .1800435
      income_quintile |  -.0955784   .0479694    -1.99   0.046    -.1895967   -.0015601
    employment_status |  -.4319033   .1391053    -3.10   0.002    -.7045447   -.1592619
      rural_residence |    -.12141   .1376071    -0.88   0.378     -.391115     .148295
                _cons |   1.416825   .5600531     2.53   0.011     .3191413    2.514509
    ------------------------------------------------------------------------------

206 .
207 . estimates store logit_model

208 .
    end of do-file

209 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

210 . * Run IV regression
211 . ivregress 2sls borrowed_binary (owns_mobile_phone = network_4G_Cov) ///
    >           internet_access age education_level income_quintile employment_status rural_residence

    Instrumental variables 2SLS regression          Number of obs   =      1,062
                                                    Wald chi2(7)    =      24.83
                                                    Prob > chi2     =     0.0008
                                                    R-squared       =     0.0161
                                                    Root MSE        =     .47841

    ------------------------------------------------------------------------------
      borrowed_binary | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
    ------------------+-----------------------------------------------------------
    owns_mobile_phone |  -.1714302   .1150442    -1.49   0.136    -.3969127    .0540523
      internet_access |  -.0464133   .0547767    -0.85   0.397    -.1537736     .060947
                  age |   .0018886   .0013565     1.39   0.164      -.00077    .0045472
      education_level |  -.0350836   .0353295    -0.99   0.321    -.1043281     .034161
      income_quintile |  -.0224814   .0109751    -2.05   0.041    -.0439922   -.0009706
    employment_status |  -.0888149   .0323646    -2.74   0.006    -.1522485   -.0253814
      rural_residence |  -.0227971   .0319045    -0.71   0.475    -.0853287    .0397346
                _cons |   .8699181   .1389064     6.26   0.000     .5976665     1.14217
```

```
------------------+---------------------------------------------------------------
Endogenous: owns_mobile_phone
Exogenous:  internet_access age education_level income_quintile employment_status
            rural_residence network_4G_Cov

212 .
213 . estimates store iv_model

214 .
end of do-file

215 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

216 . * Coefficient plot with confidence intervals
217 . coefplot ols_model logit_model iv_model, ///
>     drop(_cons) ///
>     xlabel(,angle(45)) ///
>     title("Coefficient Plot with 95% Confidence Intervals") ///
>     legend(order(1 "OLS" 2 "Logit" 3 "IV Regression"))

218 .
end of do-file

219 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph Coefficient Plot with 95% Confidence Intervals.gph"
file /Users/uli/Documents/Stata/Project/Graph/Graph Coefficient Plot with 95% Confidence Intervals.gph saved

220 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

221 . /* 2. ROC Curve for Logistic Regression Performance */
222 . * Generate predicted probabilities from logistic regression
223 . logit borrowed_binary owns_mobile_phone internet_access age education_level income_quintile employment_status rural_residence

Iteration 0:  Log likelihood = -698.77083
Iteration 1:  Log likelihood = -686.21465
Iteration 2:  Log likelihood = -686.17481
Iteration 3:  Log likelihood =  -686.1748

Logistic regression                                     Number of obs =  1,062
                                                        LR chi2(7)    =  25.19
                                                        Prob > chi2   = 0.0007
Log likelihood = -686.1748                              Pseudo R2     = 0.0180

------------------------------------------------------------------------------
  borrowed_binary | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
------------------+-----------------------------------------------------------
owns_mobile_phone |  -.2921518   .1810272    -1.61   0.107    -.6469586    .0626549
  internet_access |  -.3724162   .1698976    -2.19   0.028    -.7054095    -.039423
              age |   .0060166   .0055043     1.09   0.274    -.0047717    .0168048
  education_level |  -.1108203   .1484026    -0.75   0.455    -.4016841    .1800435
  income_quintile |  -.0955784   .0479694    -1.99   0.046    -.1895967   -.0015601
employment_status |  -.4319033   .1391053    -3.10   0.002    -.7045447   -.1592619
  rural_residence |    -.12141   .1376071    -0.88   0.378     -.391115     .148295
            _cons |   1.416825   .5600531     2.53   0.011     .3191413    2.514509
------------------------------------------------------------------------------

224 .
225 . predict logit_pred, pr

226 .
227 . * Generate ROC curve
228 . roctab borrowed_binary logit_pred, graph

229 . graph display, title("ROC Curve for Logistic Regression")
option title() not allowed
r(198);

end of do-file

r(198);
```

```
230 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph ROC Curve for Logistic Regression Summary.gph"
    file /Users/uli/Documents/Stata/Project/Graph/Graph ROC Curve for Logistic Regression Summary.gph saved

231 . do "/var/folders/_4/04hd4cwx7px1kj0zhv7lp03m0000gn/T//SD01179.000000"

232 . /* 3. Scatter Plot of Predicted vs. Actual Borrowing Behavior */
233 . * Generate predicted values from OLS regression
234 . logit borrowed_binary owns_mobile_phone internet_access age education_level income_quintile employment_status rural_residence

    Iteration 0:  Log likelihood = -698.77083
    Iteration 1:  Log likelihood = -686.21465
    Iteration 2:  Log likelihood = -686.17481
    Iteration 3:  Log likelihood =  -686.1748

    Logistic regression                                     Number of obs =  1,062
                                                            LR chi2(7)    =  25.19
                                                            Prob > chi2   = 0.0007
    Log likelihood = -686.1748                              Pseudo R2     = 0.0180

    ------------------------------------------------------------------------------
      borrowed_binary | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
    ------------------+-----------------------------------------------------------
    owns_mobile_phone |  -.2921518   .1810272    -1.61   0.107    -.6469586    .0626549
      internet_access |  -.3724162   .1698976    -2.19   0.028    -.7054095    -.039423
                  age |   .0060166   .0055043     1.09   0.274    -.0047717    .0168048
      education_level |  -.1108203   .1484026    -0.75   0.455    -.4016841    .1800435
      income_quintile |  -.0955784   .0479694    -1.99   0.046    -.1895967   -.0015601
    employment_status |  -.4319033   .1391053    -3.10   0.002    -.7045447   -.1592619
      rural_residence |    -.12141   .1376071    -0.88   0.378     -.391115     .148295
                _cons |   1.416825   .5600531     2.53   0.011     .3191413    2.514509
    ------------------------------------------------------------------------------

235 .
236 . predict ols_pred
    (option pr assumed; Pr(borrowed_binary))

237 .
238 . * Scatter plot of actual vs predicted borrowing behavior
239 . scatter borrowed_binary ols_pred, ///
    >     mcolor(blue) msize(small) ///
    >     title("Scatter Plot: Predicted vs Actual Borrowing") ///
    >     xlabel(0(0.2)1) ylabel(0(0.2)1) ///
    >     xline(0.5, lcolor(red)) yline(0.5, lcolor(red)) ///
    >     legend(off)

240 .
    end of do-file

241 . graph save "Graph" "/Users/uli/Documents/Stata/Project/Graph/Graph Scatter Plot- Predicted vs Actual Borrowing Summary.gph"
    file /Users/uli/Documents/Stata/Project/Graph/Graph Scatter Plot- Predicted vs Actual Borrowing Summary.gph saved

242 .
```