```
                                                     ___  ____  ____  ____  ____(R)
                                                    /__    /   ____/   /   ____/
                                                   ___/   /   /___/   /   /___/
                                                     Statistics/Data Analysis

                                                       User: Assignment 1 log file
------------------------------------------------------------------------------------------
      name:  <unnamed>
       log:  H:\HAD5744\Assignment #3 2019\HAD5744 A3.smcl
  log type:  smcl
 opened on:   4 Dec 2019, 16:19:56

1 .
2 . **** use Dataset2 (imported from excel) *****
3 . use "H:\HAD5744\Assignment #3 2019\Dataset2rev.dta"

4 .
5 . ***** describe the dataset ***
6 . describe

Contains data from H:\HAD5744\Assignment #3 2019\Dataset2rev.dta
  obs:           322
 vars:            12                          4 Dec 2019 15:52
 size:        21,896
------------------------------------------------------------------------------------------
              storage   display     value
variable name   type    format      label      variable label
------------------------------------------------------------------------------------------
Exper           byte    %10.0g                 Exper
HP              byte    %10.0g                 HP
Age             byte    %10.0g                 Age
YSRF            double  %10.0g                 YSRF
HSRF            double  %10.0g                 HSRF
F               byte    %10.0g
G               double  %10.0g
H               double  %10.0g
I               double  %10.0g
J               double  %10.0g
K               double  %10.0g
L               str8    %9s
------------------------------------------------------------------------------------------
Sorted by:

7 . summarize

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       Exper |        320    18.95625    4.136915         12         25
          HP |        320    75.52813     14.9215         50        100
         Age |        320      50.125    11.59356         30         70
        YSRF |        320    97.84796    13.73888   59.68127   125.6781
        HSRF |        320    76.17073    19.92223   18.33078   120.2732
-------------+---------------------------------------------------------
           F |          0
           G |          1    18.95625           .   18.95625   18.95625
           H |          1    75.52813           .   75.52813   75.52813
           I |          1      50.125           .     50.125     50.125
           J |          1    97.84796           .   97.84796   97.84796
-------------+---------------------------------------------------------
           K |          1    76.17073           .   76.17073   76.17073
           L |          0

8 .
```

```
. * no missing value, all variables are numerical. Good
.
.
.
. * generate summary stats
. summarize,detail
```

Exper

	Percentiles	Smallest		
1%	12	12		
5%	12	12		
10%	13	12	Obs	320
25%	15	12	Sum of Wgt.	320
50%	19		Mean	18.95625
		Largest	Std. Dev.	4.136915
75%	23	25		
90%	24	25	Variance	17.11407
95%	25	25	Skewness	-.1516396
99%	25	25	Kurtosis	1.721133

HP

	Percentiles	Smallest		
1%	50	50		
5%	52	50		
10%	55	50	Obs	320
25%	63	50	Sum of Wgt.	320
50%	76		Mean	75.52813
		Largest	Std. Dev.	14.9215
75%	89	100		
90%	96	100	Variance	222.6512
95%	98	100	Skewness	-.0527046
99%	100	100	Kurtosis	1.773009

Age

	Percentiles	Smallest		
1%	30	30		
5%	32	30		
10%	34.5	30	Obs	320
25%	39	30	Sum of Wgt.	320
50%	51		Mean	50.125
		Largest	Std. Dev.	11.59356
75%	59.5	70		
90%	67	70	Variance	134.4107
95%	69	70	Skewness	.0161358
99%	70	70	Kurtosis	1.823681

YSRF

	Percentiles	Smallest		
1%	63.22226	59.68127		
5%	74.5025	60.80004		
10%	81.2437	61.47729	Obs	320
25%	88.15823	63.22226	Sum of Wgt.	320
50%	98.6814		Mean	97.84796
		Largest	Std. Dev.	13.73888
75%	108.1888	122.6407		
90%	115.5274	124.8386	Variance	188.7568
95%	119.3006	125.2264	Skewness	-.2526107
99%	122.6407	125.6781	Kurtosis	2.572784

HSRF

	Percentiles	Smallest		
1%	22.53686	18.33078		
5%	44.48184	18.76664		
10%	51.48454	19.68741	Obs	320
25%	61.80416	22.53686	Sum of Wgt.	320
50%	75.57711		Mean	76.17073
		Largest	Std. Dev.	19.92223
75%	92.79999	112.1047		
90%	101.5063	114.1884	Variance	396.8951
95%	104.0921	115.8988	Skewness	-.3129038
99%	112.1047	120.2732	Kurtosis	2.731626

F

no observations

G

	Percentiles	Smallest		
1%	18.95625	18.95625		
5%	18.95625	.		
10%	18.95625	.	Obs	1
25%	18.95625	.	Sum of Wgt.	1
50%	18.95625		Mean	18.95625
		Largest	Std. Dev.	.
75%	18.95625	.		
90%	18.95625	.	Variance	.
95%	18.95625	.	Skewness	.
99%	18.95625	18.95625	Kurtosis	.

H

	Percentiles	Smallest		
1%	75.52813	75.52813		
5%	75.52813	.		
10%	75.52813	.	Obs	1
25%	75.52813	.	Sum of Wgt.	1
50%	75.52813		Mean	75.52813
		Largest	Std. Dev.	.
75%	75.52813	.		
90%	75.52813	.	Variance	.
95%	75.52813	.	Skewness	.
99%	75.52813	75.52813	Kurtosis	.

I

	Percentiles	Smallest		
1%	50.125	50.125		
5%	50.125	.		
10%	50.125	.	Obs	1
25%	50.125	.	Sum of Wgt.	1
50%	50.125		Mean	50.125
		Largest	Std. Dev.	.
75%	50.125	.		
90%	50.125	.	Variance	.
95%	50.125	.	Skewness	.
99%	50.125	50.125	Kurtosis	.

```
                                   J
-------------------------------------------------------------
      Percentiles      Smallest
 1%     97.84796       97.84796
 5%     97.84796              .
10%     97.84796              .       Obs                   1
25%     97.84796              .       Sum of Wgt.           1

50%     97.84796                      Mean           97.84796
                        Largest       Std. Dev.             .
75%     97.84796              .
90%     97.84796              .       Variance              .
95%     97.84796              .       Skewness              .
99%     97.84796       97.84796       Kurtosis              .

                                   K
-------------------------------------------------------------
      Percentiles      Smallest
 1%     76.17073       76.17073
 5%     76.17073              .
10%     76.17073              .       Obs                   1
25%     76.17073              .       Sum of Wgt.           1

50%     76.17073                      Mean           76.17073
                        Largest       Std. Dev.             .
75%     76.17073              .
90%     76.17073              .       Variance              .
95%     76.17073              .       Skewness              .
99%     76.17073       76.17073       Kurtosis              .

                                   L
-------------------------------------------------------------
no observations

15 .
16 . * Determination of functional form
17 . * Health equation: inherited health and age
18 .
19 . scatter HSRF HP, saving(scatterHealthInheri)
   (file scatterHealthInheri.gph saved)

20 . scatter HSRF Age, saving(scatterHealthAge)
   (file scatterHealthAge.gph saved)

21 . gr combine scatterHealthInheri.gph scatterHealthAge.gph

22 .
23 . * Inherited health is ok; age-health is non-linear
24 . * plot health vs. logage and age^2
25 . gen logage = log(Age)
   (2 missing values generated)

26 . gen agesq = Age^2
   (2 missing values generated)

27 .
28 . scatter HSRF logage, saving(scatterHealthLogage)
   (file scatterHealthLogage.gph saved)
```

```
. scatter HSRF agesq, saving(scatterHealthSqage)
  (file scatterHealthSqage.gph saved)

. gr combine scatterHealthLogage.gph scatterHealthSqage.gph

. * age^2 is better
.
. * Income equation: working experience and age
.
. scatter YSRF Exper, saving(scatterIncomeExp)
  (file scatterIncomeExp.gph saved)

. scatter YSRF Age, saving(scatterIncomeAge)
  (file scatterIncomeAge.gph saved)

. gr combine scatterIncomeExp.gph scatterIncomeAge.gph

.
. * working exp is ok exp for some potential outliers
. * age needs to be transformed
. * plot income vs. logage and age^2
. scatter YSRF logage, saving(scatterIncomeLogage)
  (file scatterIncomeLogage.gph saved)

. scatter YSRF agesq, saving(scatterIncomeSqage)
  (file scatterIncomeSqage.gph saved)

. gr combine scatterIncomeLogage.gph scatterIncomeSqage.gph

.
. * log(age) is still non-linear; use age^2
.
.
. * Note:  without HP, the YSRF model is not identified in our system of M=2 equations
. * is that at least M-1 variables (in this case, one variable) from the system be left out of each equ
. * In health equation, work experience is omitted; therefore, this equation is identified
. * However, income equation is not identified as it includes all variables in the system
. * Therefore, need to include HP in HSRF model in order to identify this equation
. * Estimating the final 2SLS model
. * use inherited health as an IV for health
. ivregress 2sls YSRF  Exper Age agesq (HSRF=HP)

  Instrumental variables (2SLS) regression              Number of obs =      320
                                                         Wald chi2(4)  = 2061.27
                                                         Prob > chi2   =  0.0000
                                                         R-squared     =  0.8739
                                                         Root MSE      =  4.8717
```

YSRF	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
HSRF	.4187065	.0222697	18.80	0.000	.3750586	.4623544
Exper	2.069323	.0691817	29.91	0.000	1.93373	2.204917
Age	1.258022	.2281341	5.51	0.000	.8108879	1.705157
agesq	-.0109084	.0023068	-4.73	0.000	-.0154296	-.0063872
_cons	-7.460969	5.73362	-1.30	0.193	-18.69866	3.776719

```
  Instrumented:  HSRF
  Instruments:   Exper Age agesq HP
```

```
57 . predict resIncome,residuals
   (2 missing values generated)

58 . predict hatIncome
   (option xb assumed; fitted values)
   (2 missing values generated)

59 . estat endog

     Tests of endogeneity
     Ho: variables are exogenous

     Durbin (score) chi2(1)          =  18.3141  (p = 0.0000)
     Wu-Hausman F(1,314)             =  19.0617  (p = 0.0000)

60 .
61 . * use work exp as an IV for income
62 . ivregress 2sls HSRF HP Age agesq (YSRF=Exper)

   Instrumental variables (2SLS) regression          Number of obs =      320
                                                     Wald chi2(4)  =  5120.38
                                                     Prob > chi2   =   0.0000
                                                     R-squared     =   0.9425
                                                     Root MSE      =   4.7695
```

HSRF	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
YSRF	.2911361	.0275131	10.58	0.000	.2372113	.3450609
HP	.7237591	.0210054	34.46	0.000	.6825892	.764929
Age	1.056241	.2305769	4.58	0.000	.604319	1.508164
agesq	-.0211947	.0023099	-9.18	0.000	-.0257219	-.0166675
_cons	-3.832772	5.613893	-0.68	0.495	-14.8358	7.170257

```
   Instrumented:  YSRF
   Instruments:   HP Age agesq Exper

63 . predict resHealth, residuals
   (2 missing values generated)

64 . predict hatHealth
   (option xb assumed; fitted values)
   (2 missing values generated)

65 . estat endog

     Tests of endogeneity
     Ho: variables are exogenous

     Durbin (score) chi2(1)          =   30.949  (p = 0.0000)
     Wu-Hausman F(1,314)             =  33.6203  (p = 0.0000)

66 .
67 . * sensitvity analysis
68 . * check the strength of IV by examing the repsective first-stage t-statistics
69 . ivregress 2sls YSRF  Exper Age agesq (HSRF=HP),first

   First-stage regressions
   -----------------------

                                               Number of obs     =        320
                                               F(   4,     315)  =     905.66
                                               Prob > F          =     0.0000
                                               R-squared         =     0.9200
                                               Adj R-squared     =     0.9190
                                               Root MSE          =     5.6704
```

HSRF	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Exper	.6860895	.0770838	8.90	0.000	.5344253	.8377537
Age	1.619973	.2637108	6.14	0.000	1.101116	2.13883
agesq	-.0277537	.0026111	-10.63	0.000	-.0328911	-.0226163
HP	.8242337	.0213649	38.58	0.000	.7821977	.8662696
_cons	-6.838553	6.717289	-1.02	0.309	-20.05498	6.377871

Instrumental variables (2SLS) regression

Number of obs	=	320
Wald chi2(4)	=	2061.27
Prob > chi2	=	0.0000
R-squared	=	0.8739
Root MSE	=	4.8717

YSRF	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
HSRF	.4187065	.0222697	18.80	0.000	.3750586	.4623544
Exper	2.069323	.0691817	29.91	0.000	1.93373	2.204917
Age	1.258022	.2281341	5.51	0.000	.8108879	1.705157
agesq	-.0109084	.0023068	-4.73	0.000	-.0154296	-.0063872
_cons	-7.460969	5.73362	-1.30	0.193	-18.69866	3.776719

Instrumented: HSRF
Instruments: Exper Age agesq HP

70 . ivregress 2sls HSRF Age agesq (YSRF=Exper),first

First-stage regressions

Number of obs	=	320
F(3, 316)	=	216.91
Prob > F	=	0.0000
R-squared	=	0.6731
Adj R-squared	=	0.6700
Root MSE	=	7.8922

YSRF	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Age	1.783129	.3667994	4.86	0.000	1.061451	2.504806
agesq	-.0211107	.0036321	-5.81	0.000	-.0282569	-.0139644
Exper	2.459222	.1069214	23.00	0.000	2.248854	2.66959
_cons	17.72046	9.031548	1.96	0.051	-.049106	35.49003

Instrumental variables (2SLS) regression

Number of obs	=	320
Wald chi2(3)	=	539.82
Prob > chi2	=	0.0000
R-squared	=	0.6787
Root MSE	=	11.275

HSRF	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
YSRF	.3786555	.0621142	6.10	0.000	.2569139	.5003971
Age	.5789243	.540442	1.07	0.284	-.4803226	1.638171
agesq	-.0163724	.0054081	-3.03	0.002	-.0269721	-.0057727
_cons	53.43106	13.14587	4.06	0.000	27.66563	79.19649

Instrumented: YSRF
Instruments: Age agesq Exper

```
71 .
72 . * very big t- and F-score.
73 .
74 . * check the residual of the two equations
75 . hist resIncome,saving(hisresIncome)
   (bin=17, start=-14.334509, width=1.673621)
   (file hisresIncome.gph saved)

76 . hist resHealth,saving(hisresHealth)
   (bin=17, start=-11.971336, width=1.8924575)
   (file hisresHealth.gph saved)

77 . gr combine hisresIncome.gph hisresHealth.gph

78 .
79 . * Shapiro-Wilk
80 . swilk resIncome
```

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
resIncome	320	0.99641	0.810	-0.495	0.68962

```
81 . swilk resHealth
```

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
resHealth	320	0.99338	1.494	0.945	0.17236

```
82 .
83 . * cannot reject normality
84 .
85 . * check heteroskedasticity: plot fitted income/health against res
86 . scatter hatIncome resIncome, saving(hatresIncome)
   (file hatresIncome.gph saved)

87 . scatter hatHealth resHealth, saving(hatresHealth)
   (file hatresHealth.gph saved)

88 . gr combine hatresIncome.gph hatresHealth.gph

89 .
90 .
91 . * check heteroskedasticity and outliers: plot fitted value vs. predictor
92 . * income equation
93 . scatter resIncome Age,saving(rIncomeAge)
   (file rIncomeAge.gph saved)

94 . scatter resIncome agesq, saving(rIncomeagesq)
   (file rIncomeagesq.gph saved)

95 . scatter resIncome Exper, saving(rIncomeExp)
   (file rIncomeExp.gph saved)
```

```
 96 . scatter resIncome YSRF, saving(rIncomeHealth)
    (file rIncomeHealth.gph saved)

 97 .
 98 . * health equation
 99 . scatter resHealth Age,saving(rHealthAge)
    (file rHealthAge.gph saved)

100 . scatter resHealth agesq, saving(rHealthagesq)
    (file rHealthagesq.gph saved)

101 . scatter resHealth HP, saving(rHealthInheri)
    (file rHealthInheri.gph saved)

102 . scatter resHealth HSRF, saving(rHealthIncome)
    (file rHealthIncome.gph saved)

103 .
104 . gr combine rHealthAge.gph rHealthagesq.gph rHealthInheri.gph rHealthIncome.gph

105 . * no cluster, no pattern -> heteroskedastic
106 . * also no obv outliers
107 .
108 . * Model comparison: logage vs. age^2
109 . * first we report the regression results if age -> log(age)
110 . ivregress 2sls YSRF  Exper logage (HSRF=HP)

    Instrumental variables (2SLS) regression          Number of obs =       320
                                                       Wald chi2(3)  =   1827.64
                                                       Prob > chi2   =    0.0000
                                                       R-squared     =    0.8668
                                                       Root MSE      =    5.0067

    ------------------------------------------------------------------------------
            YSRF |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
    -------------+----------------------------------------------------------------
            HSRF |   .4158628   .0231588    17.96   0.000     .3704725    .4612532
           Exper |   2.084258   .0714223    29.18   0.000     1.944273    2.224243
          logage |   8.204055   1.743682     4.71   0.000     4.786501    11.62161
           _cons |  -5.223639   7.922289    -0.66   0.510    -20.75104    10.30376
    ------------------------------------------------------------------------------
    Instrumented:  HSRF
    Instruments:   Exper logage HP

111 . estat endog

      Tests of endogeneity
      Ho: variables are exogenous

      Durbin (score) chi2(1)          =  28.1297  (p = 0.0000)
      Wu-Hausman F(1,315)             =  30.3589  (p = 0.0000)

112 .
113 . ivregress 2sls HSRF HP logage (YSRF=Exper)

    Instrumental variables (2SLS) regression          Number of obs =       320
                                                       Wald chi2(3)  =   2948.42
                                                       Prob > chi2   =    0.0000
                                                       R-squared     =    0.9064
                                                       Root MSE      =    6.0849
```

HSRF	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
YSRF	.307922	.034569	8.91	0.000	.240168	.3756761
HP	.7099119	.026564	26.72	0.000	.6578474	.7619764
logage	-50.0952	1.497665	-33.45	0.000	-53.03056	-47.15983
_cons	187.1193	7.200878	25.99	0.000	173.0058	201.2328

```
Instrumented:  YSRF
Instruments:   HP logage Exper
```

```
114 . estat endog

  Tests of endogeneity
  Ho: variables are exogenous

  Durbin (score) chi2(1)          =  66.0694  (p = 0.0000)
  Wu-Hausman F(1,315)             =  81.9589  (p = 0.0000)

115 .
116 .
117 . * RESET test is conducted to examine omitted var
118 . * original model
119 . reg YSRF Exper Age agesq
```

Source	SS	df	MS
Model	40530.9427	3	13510.3142
Residual	19682.4905	316	62.2863622
Total	60213.4332	319	188.756844

```
Number of obs =       320
F(  3,   316) =    216.91
Prob > F      =    0.0000
R-squared     =    0.6731
Adj R-squared =    0.6700
Root MSE      =    7.8922
```

YSRF	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Exper	2.459222	.1069214	23.00	0.000	2.248854	2.66959
Age	1.783129	.3667994	4.86	0.000	1.061451	2.504806
agesq	-.0211107	.0036321	-5.81	0.000	-.0282569	-.0139644
_cons	17.72046	9.031548	1.96	0.051	-.049106	35.49003

```
120 . predict Incomehat
    (option xb assumed; fitted values)
    (2 missing values generated)

121 . reg HSRF Incomehat Age agesq
```

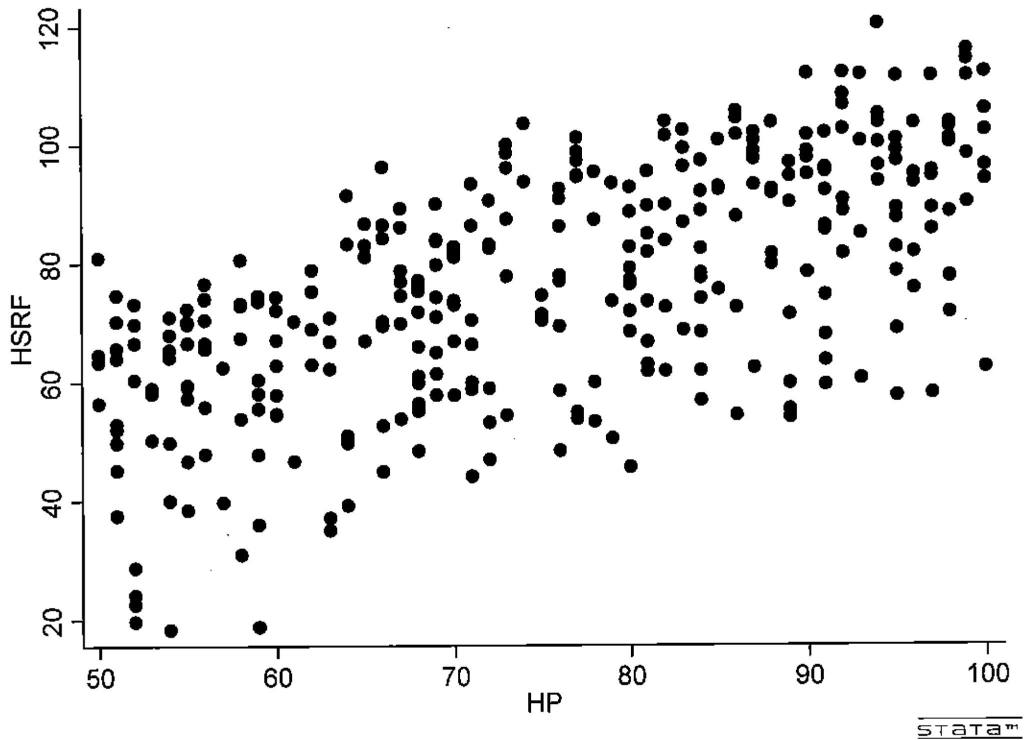
Source	SS	df	MS
Model	68626.0485	3	22875.3495
Residual	57983.496	316	183.492076
Total	126609.544	319	396.895124

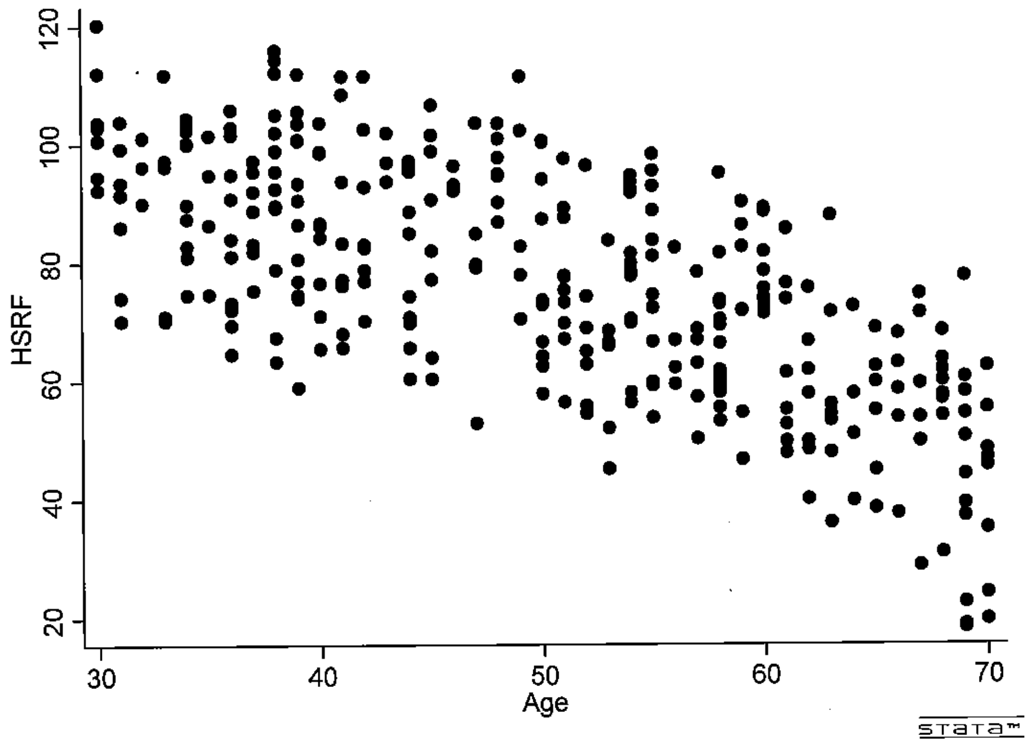
```
Number of obs =       320
F(  3,   316) =    124.67
Prob > F      =    0.0000
R-squared     =    0.5420
Adj R-squared =    0.5377
Root MSE      =    13.546
```

HSRF	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Incomehat	.3786555	.0746241	5.07	0.000	.2318327	.5254784
Age	.5789242	.6492877	0.89	0.373	-.698549	1.856397
agesq	-.0163724	.0064973	-2.52	0.012	-.0291558	-.0035889
_cons	53.43106	15.79346	3.38	0.001	22.35743	84.50469

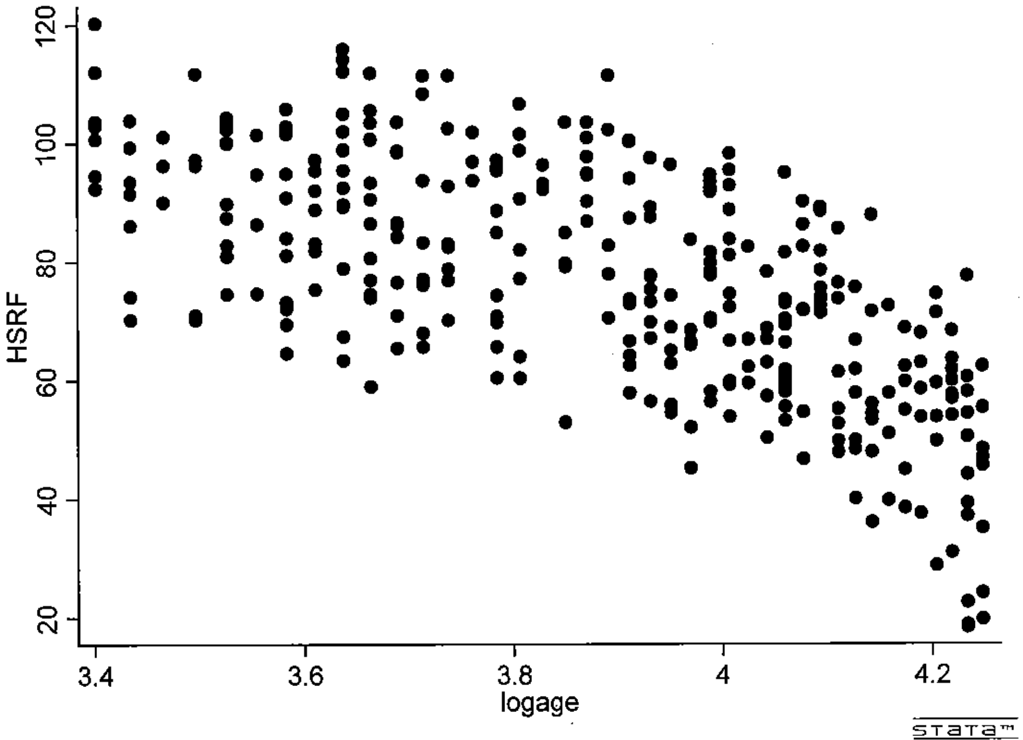
```
122 . estat ovtest

    Ramsey RESET test using powers of the fitted values of HSRF
           Ho:  model has no omitted variables
                    F(3, 313) =      1.47
                     Prob > F =      0.2220

123 .
124 . * fail to reject absence of omitted variables
125 . * original model is adequate
126 .
127 .
128 . * RESET test is conducted to examine omitted var bias if we now
129 . * include HP in HSRF equation
130 .
131 . reg YSRF Exper Age agesq

          Source |       SS           df       MS      Number of obs   =       320
    -------------+----------------------------------   F(3, 316)       =    216.91
           Model |  40530.9427         3  13510.3142   Prob > F        =    0.0000
        Residual |  19682.4905       316  62.2863622   R-squared       =    0.6731
    -------------+----------------------------------   Adj R-squared   =    0.6700
           Total |  60213.4332       319  188.756844   Root MSE        =    7.8922

    ------------------------------------------------------------------------------
            YSRF |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
    -------------+----------------------------------------------------------------
           Exper |   2.459222   .1069214    23.00   0.000     2.248854     2.66959
             Age |   1.783129   .3667994     4.86   0.000     1.061451    2.504806
           agesq |  -.0211107   .0036321    -5.81   0.000    -.0282569   -.0139644
           _cons |   17.72046   9.031548     1.96   0.051     -.049106    35.49003
    ------------------------------------------------------------------------------

132 . predict Incomehat2
    (option xb assumed; fitted values)
    (2 missing values generated)

133 . reg HSRF Incomehat2 HP Age agesq

          Source |       SS           df       MS      Number of obs   =       320
    -------------+----------------------------------   F(4, 315)       =    905.66
           Model |  116481.157         4  29120.2892   Prob > F        =    0.0000
        Residual |  10128.3877       315  32.1536116   R-squared       =    0.9200
    -------------+----------------------------------   Adj R-squared   =    0.9190
           Total |  126609.544       319  396.895124   Root MSE        =    5.6704

    ------------------------------------------------------------------------------
            HSRF |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
    -------------+----------------------------------------------------------------
      Incomehat2 |   .2789864   .0313448     8.90   0.000     .2173148     .340658
              HP |   .8242337   .0213649    38.58   0.000     .7821977    .8662696
             Age |   1.122504   .2721612     4.12   0.000     .5870206    1.657988
           agesq |  -.0218641   .0027235    -8.03   0.000    -.0272228   -.0165055
           _cons |  -11.78232   6.823932    -1.73   0.085    -25.20857    1.643927
    ------------------------------------------------------------------------------

134 . estat ovtest

    Ramsey RESET test using powers of the fitted values of HSRF
           Ho:  model has no omitted variables
                    F(3, 312) =      1.72
                     Prob > F =      0.1629
```

```
135 .
136 . * Fail to reject absence of omitted variables
137 . * New model is now adequate
138 . * This shows that RESET test does not always show that a variable
139 . * is statistically omitted even though conceptually it is necessary to include
140 .
141 .
142 . reg YSRF Exper logage

      Source |       SS       df       MS              Number of obs =     320
-------------+------------------------------           F(  2,   317) =  265.98
       Model |  37730.0772     2  18865.0386           Prob > F      =  0.0000
    Residual |   22483.356   317  70.9254131           R-squared     =  0.6266
-------------+------------------------------           Adj R-squared =  0.6242
       Total |  60213.4332   319  188.756844           Root MSE      =  8.4217

------------------------------------------------------------------------------
        YSRF |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       Exper |   2.489385   .1139881    21.84   0.000     2.265116    2.713654
      logage |  -15.06513   1.962559    -7.68   0.000    -18.92641   -11.20384
       _cons |   109.2096   7.917034    13.79   0.000     93.63303    124.7862
------------------------------------------------------------------------------

143 . predict Incomehat3
    (option xb assumed; fitted values)
    (2 missing values generated)

144 . reg HSRF Incomehat3 HP logage

      Source |       SS       df       MS              Number of obs =     320
-------------+------------------------------           F(  3,   316) =  659.37
       Model |  109169.923     3  36389.9742           Prob > F      =  0.0000
    Residual |  17439.6219   316  55.1886769           R-squared     =  0.8623
-------------+------------------------------           Adj R-squared =  0.8609
       Total |  126609.544   319  396.895124           Root MSE      =  7.4289

------------------------------------------------------------------------------
        HSRF |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  Incomehat3 |   .2956721   .0405252     7.30   0.000     .2159388    .3754054
          HP |    .814169   .0279773    29.10   0.000     .7591237    .8692143
      logage |  -50.10059   1.828213   -27.40   0.000     -53.6976   -46.50358
       _cons |   180.4645   9.355858    19.29   0.000     162.0569    198.8722
------------------------------------------------------------------------------

145 . estat ovtest

    Ramsey RESET test using powers of the fitted values of HSRF
           Ho:  model has no omitted variables
                    F(3, 313) =     14.13
                     Prob > F =      0.0000

146 . * log(age) model is not adequate
147 . * choose our unlogged model for discussion
148 .
149 . * END
150 . log close
        name:  <unnamed>
         log:  H:\HAD5744\Assignment #3 2019\HAD5744 A3.smcl
    log type:  smcl
   closed on:   4 Dec 2019, 16:20:25
```

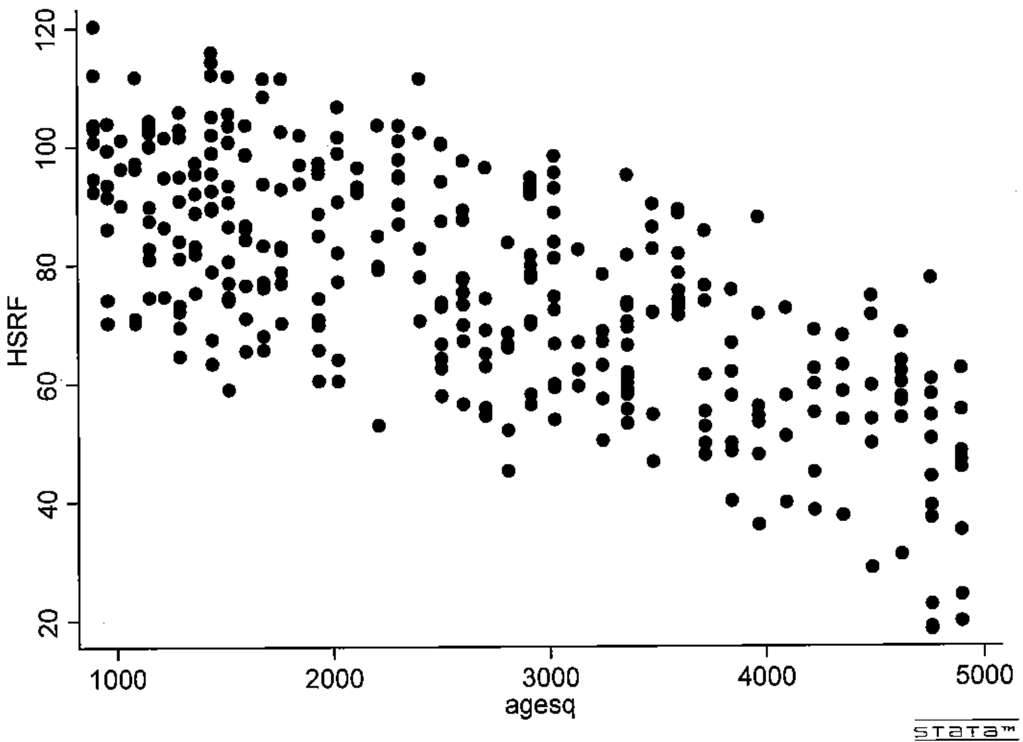

HSRF
120
100
80
60
40
20
50
60
70
80
90
100
HP
STATA™

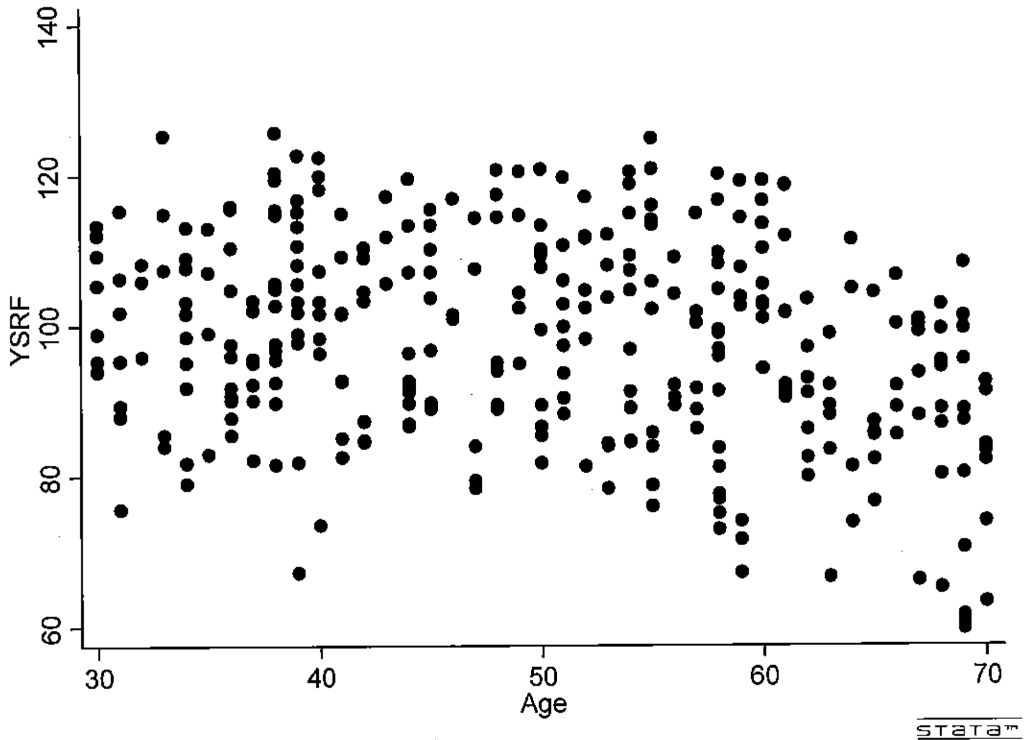


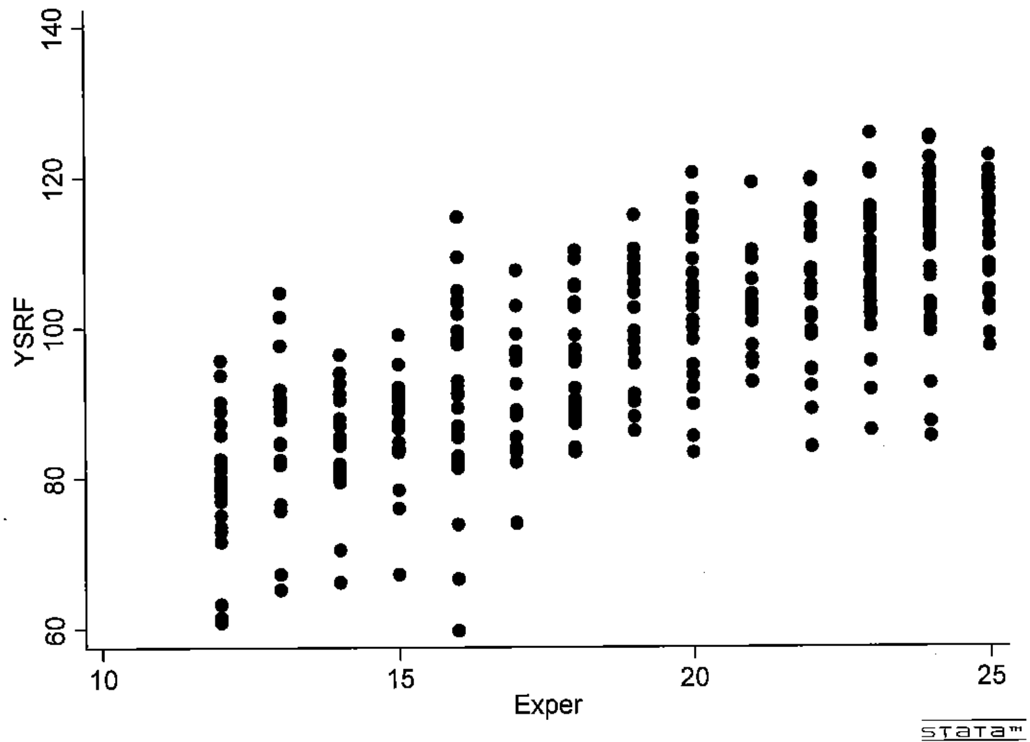
HSRF
120
100
80
60
40
20
30
40
50
60
70
Age



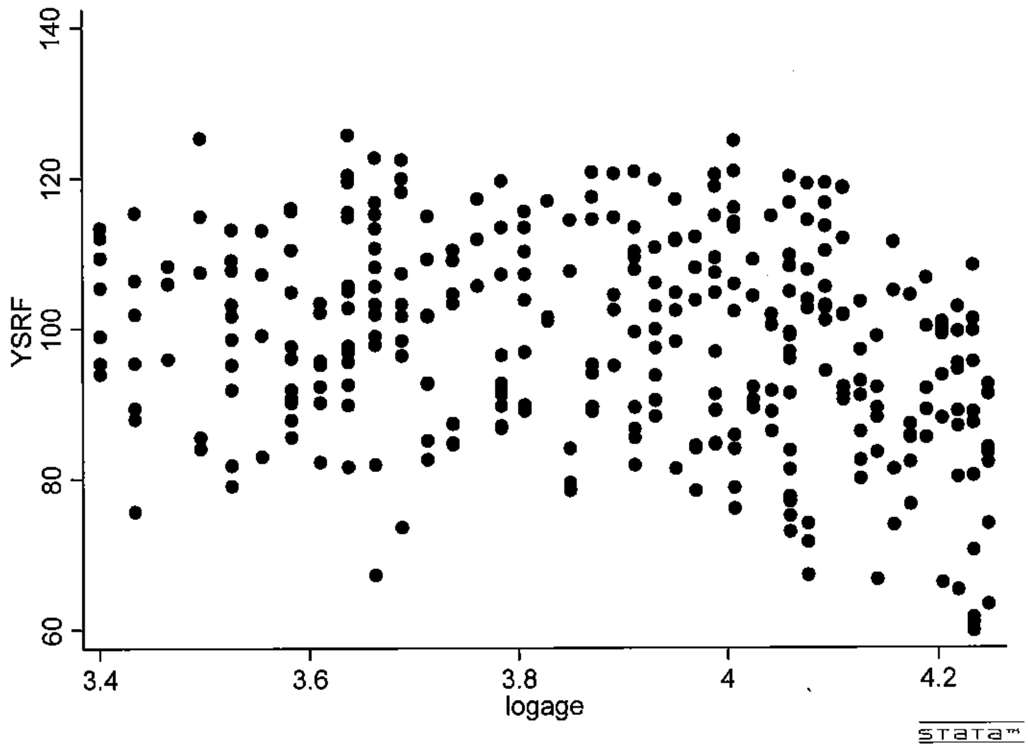
120
100
80
60
40
20
HSRF
3.4
3.6
3.8
4
4.2
logage


120
100
80
60
40
20
HSRF
1000
2000
3000
4000
5000
agesq
stata™

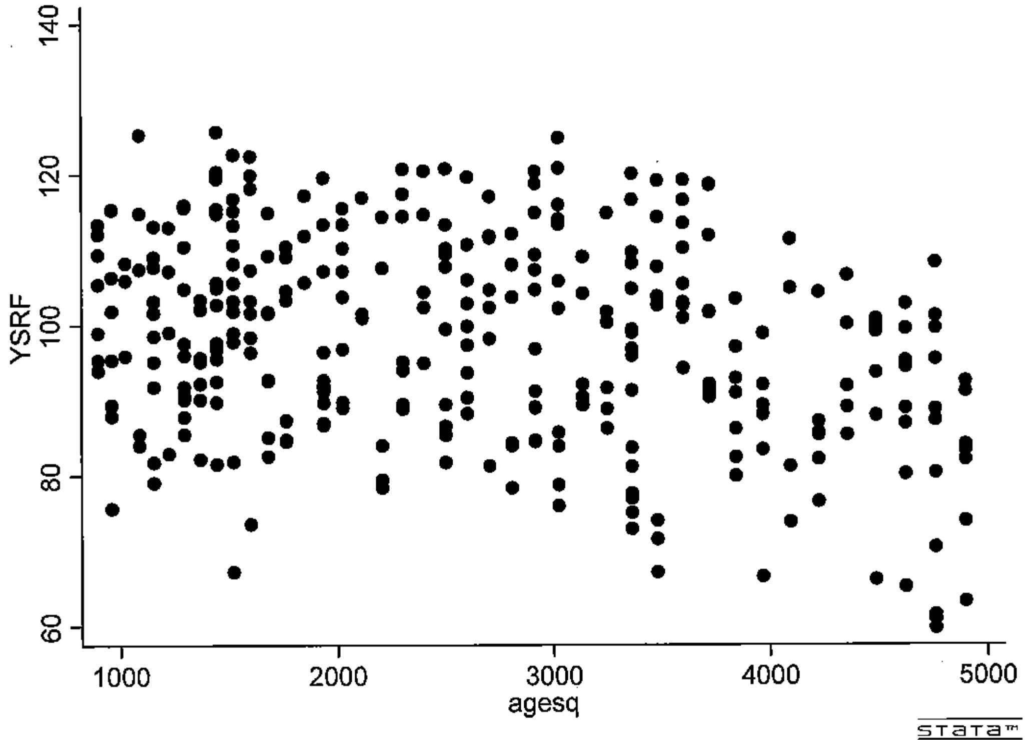

YSRF
140
120
100
80
60
30
40
50
60
70
Age
STATA™

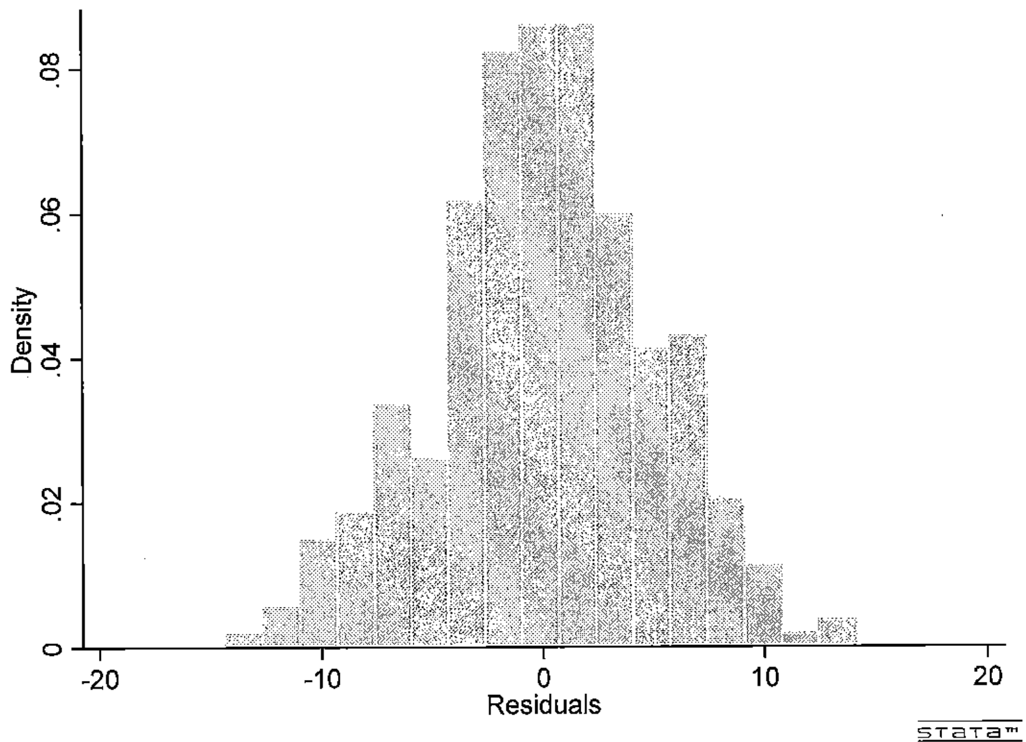


YSRF
140
120
100
80
60
10
15
20
25
Exper
stata™

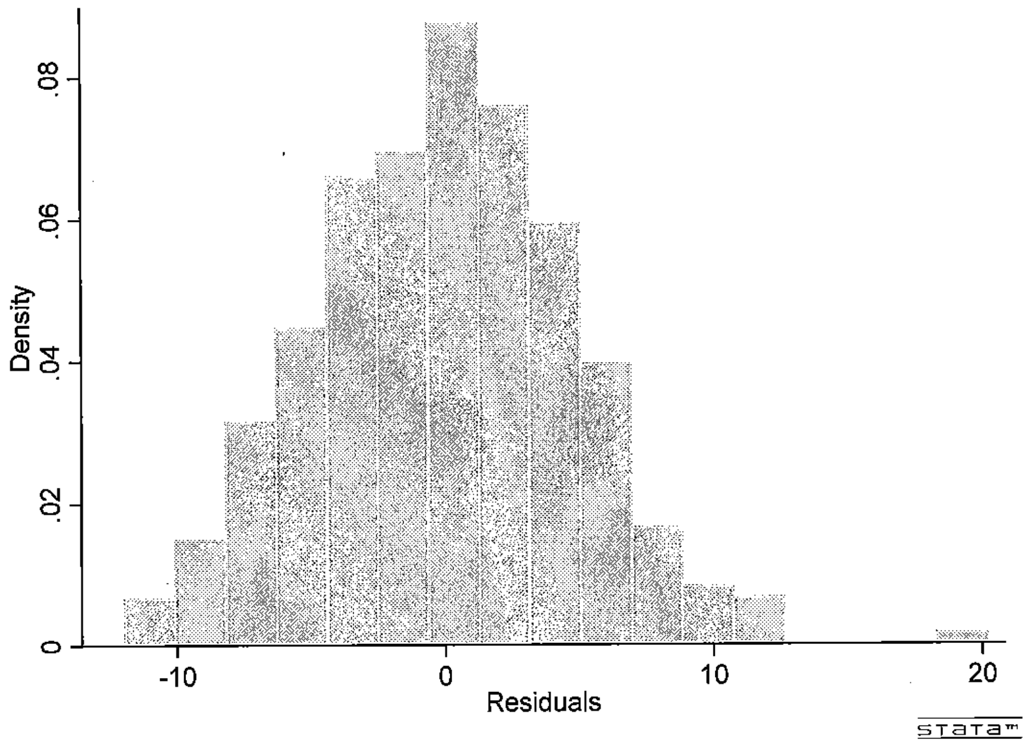


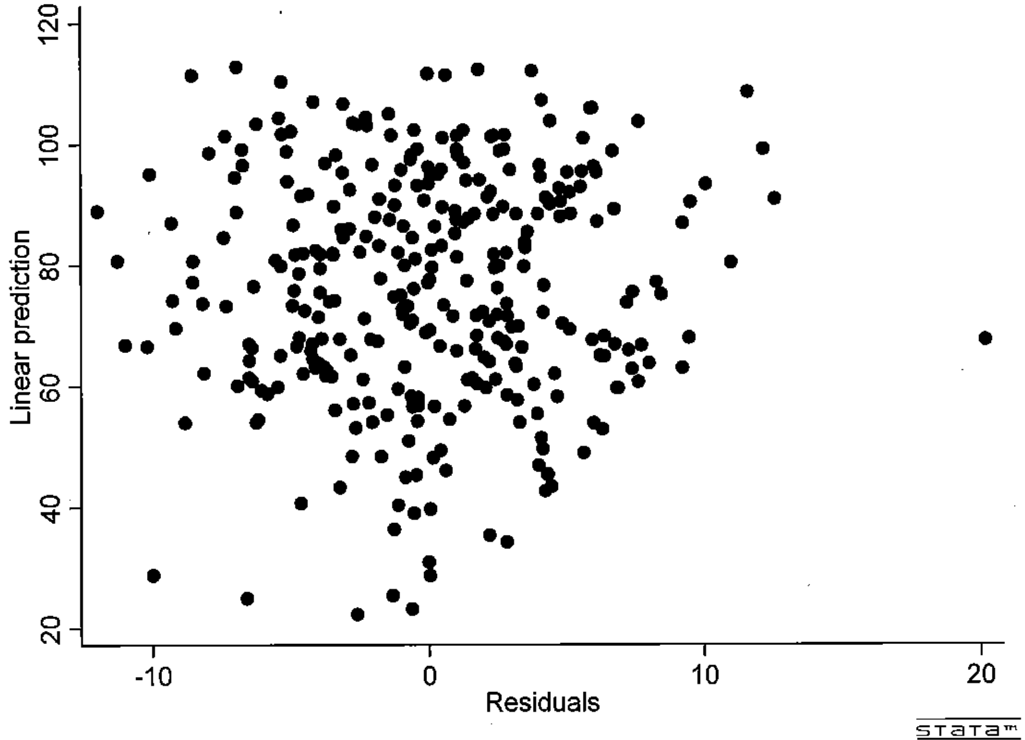
YSRF
140
120
100
80
60
3.4
3.6
3.8
4
4.2
logage
stata™


140
120
100
80
60
YSRF
1000
2000
3000
4000
5000
agesq
stata™

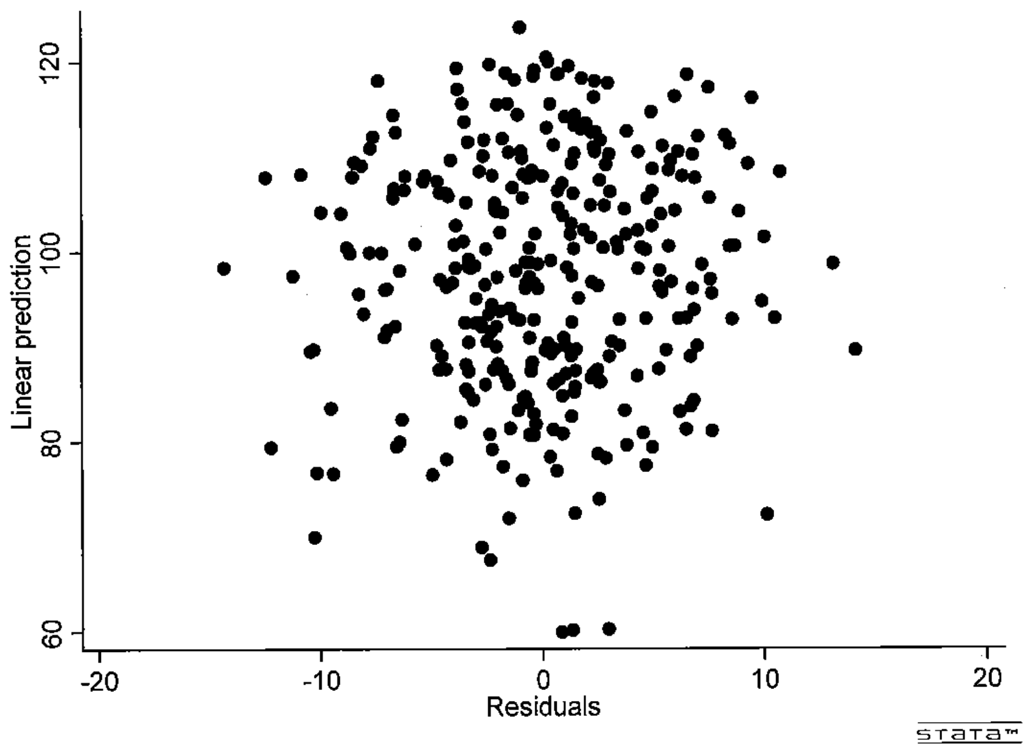


Density
0
.02
.04
.06
.08
-20
-10
0
10
20
Residuals
stata™

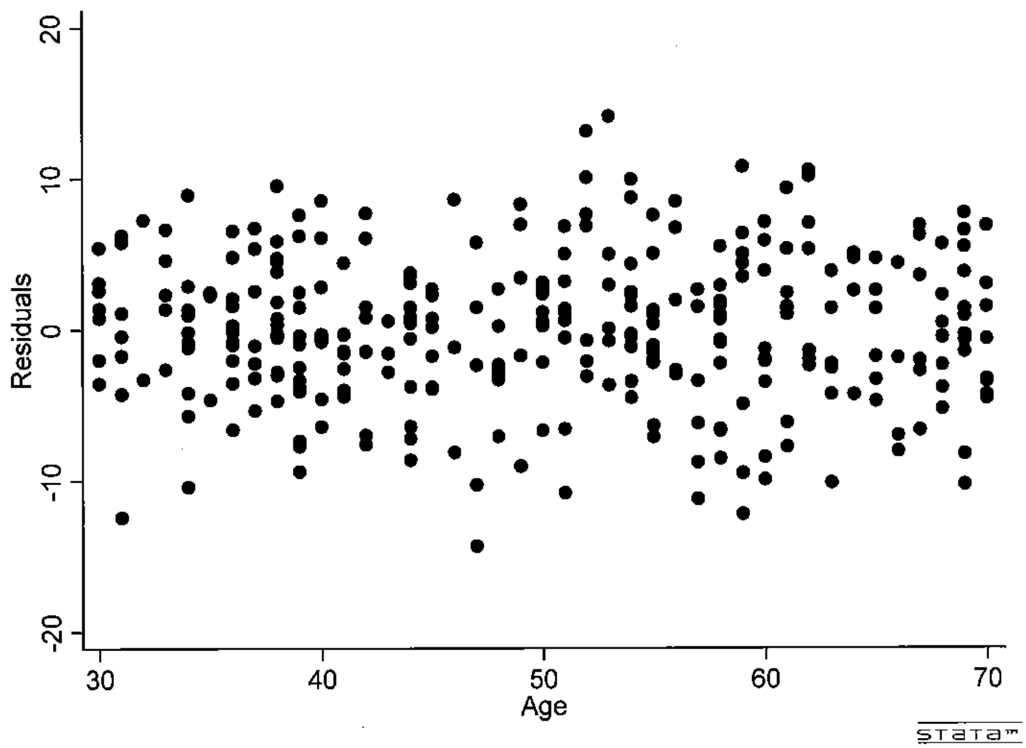

Density
0
.02
.04
.06
.08
-10
0
10
20
Residuals
stata™



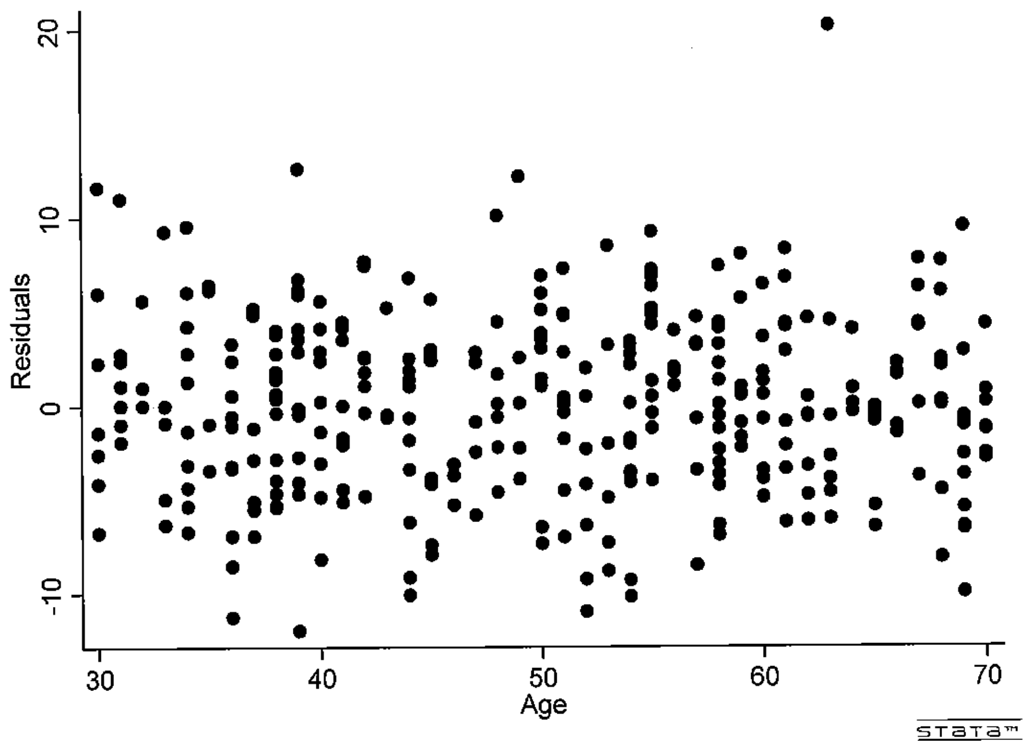
Linear prediction
120
100
80
60
40
20
-10
0
10
20
Residuals
stata™



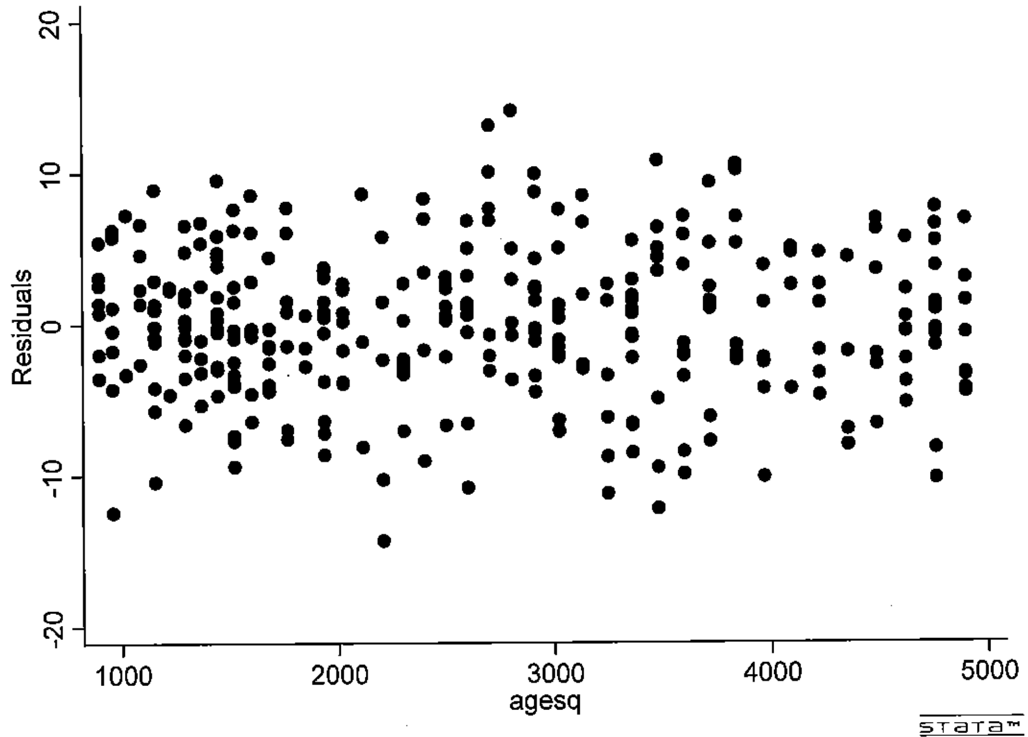
Linear prediction
60
80
100
120
-20
-10
0
10
20
Residuals
STATA™

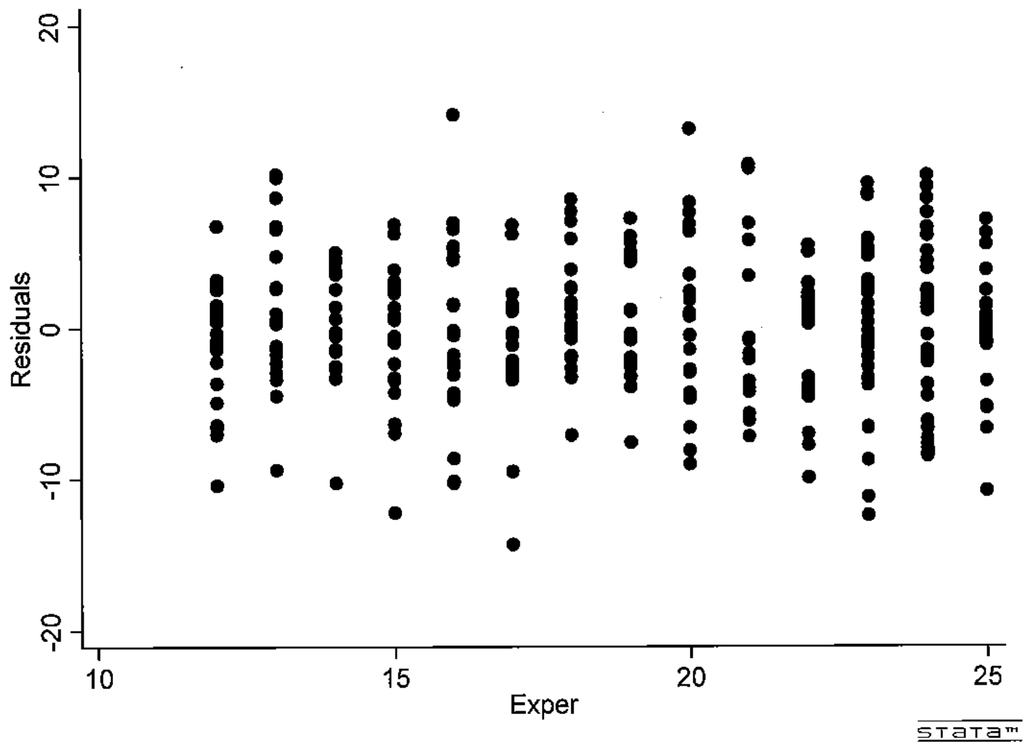


Residuals
20
10
0
-10
-20
30
40
50
60
70
Age
stata™

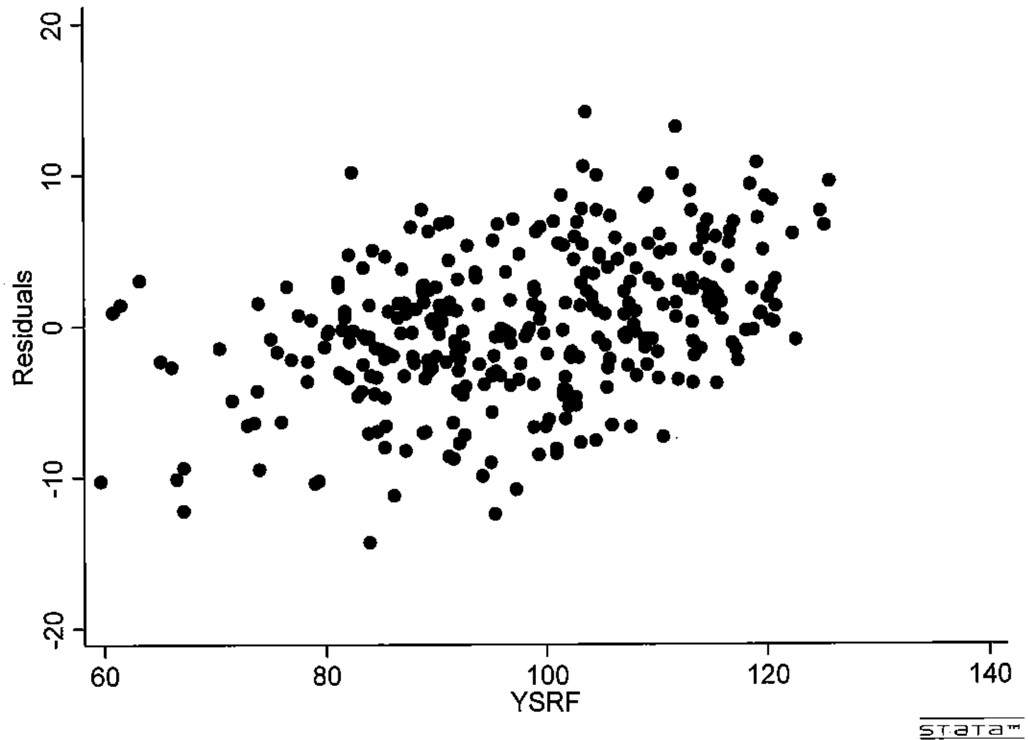


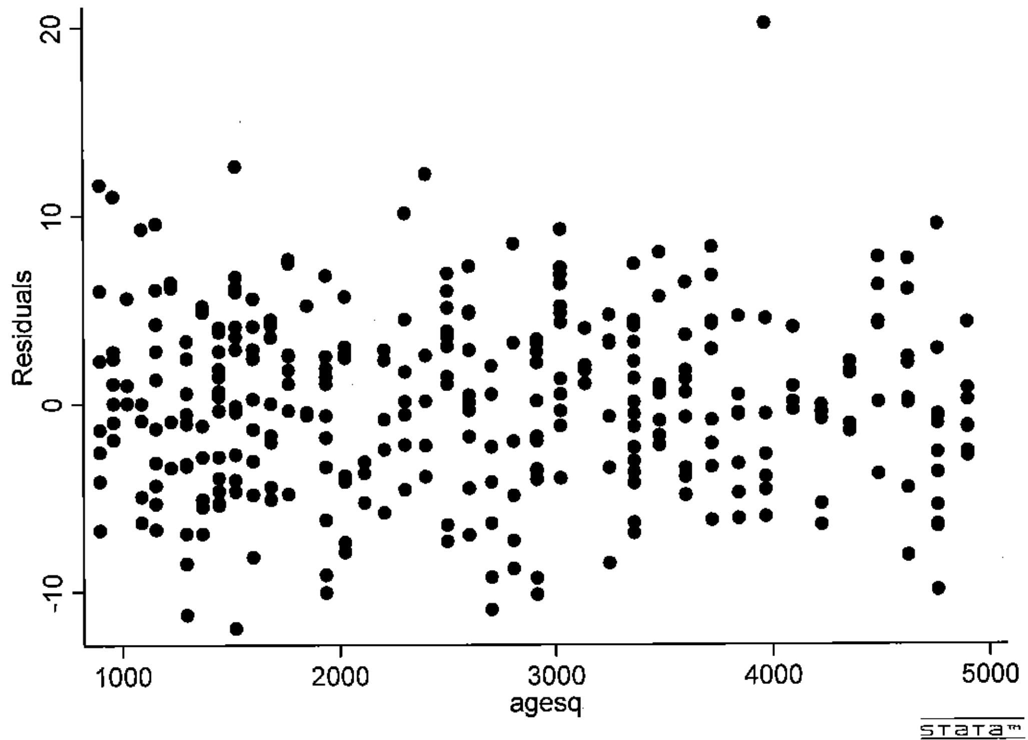
20
10
0
-10
Residuals
30
40
50
60
70
Age
stata™


20
10
0
-10
-20
Residuals
1000
2000
3000
4000
5000
agesq
STATA™

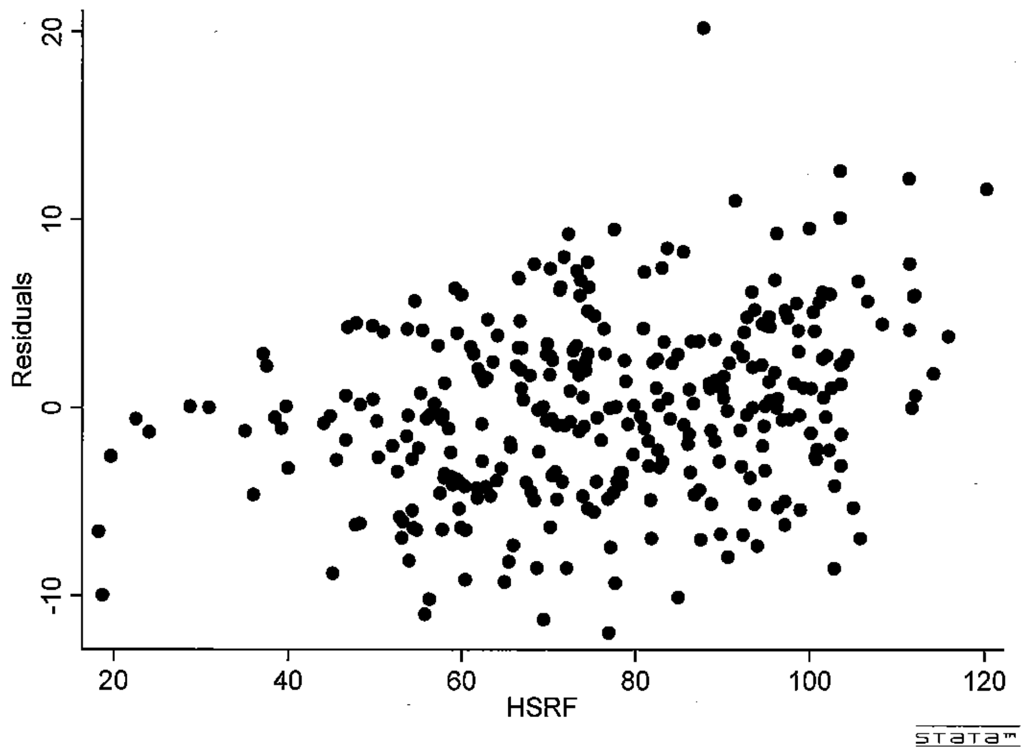


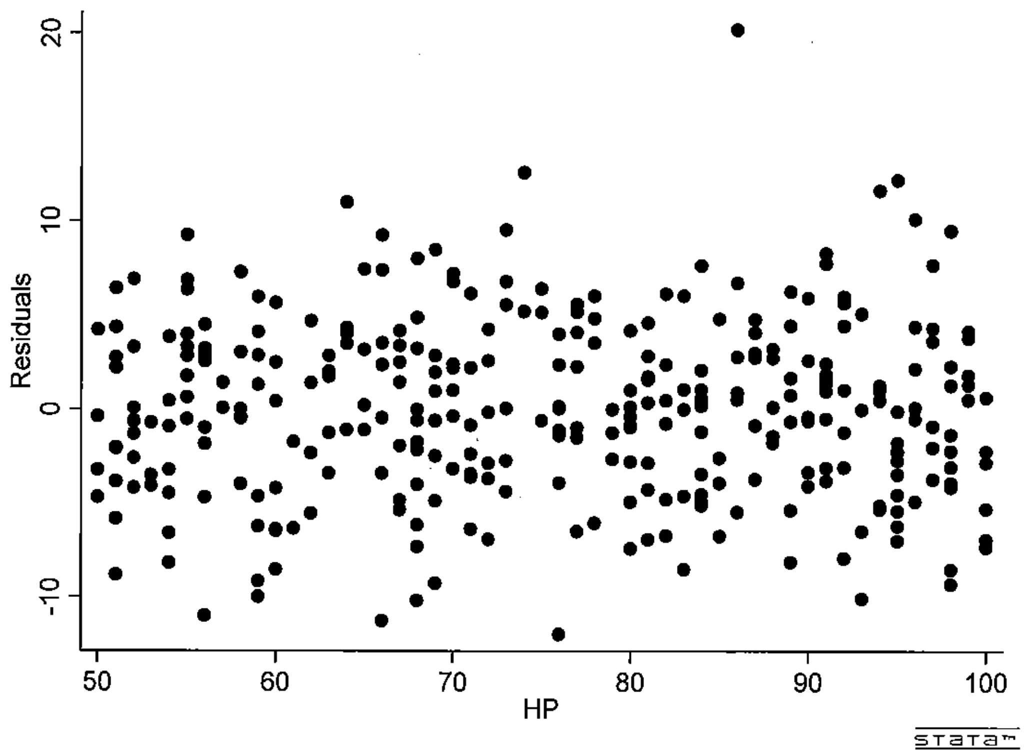
Residuals
20
10
0
-10
-20
10
15
20
25
Exper
STATA™


20
10
0
-10
-20
Residuals
60
80
100
120
140
YSRF
stata



Residuals
20
10
0
-10
1000
2000
3000
4000
5000
agesq
stata™


Residuals
20
10
0
-10
20
40
60
80
100
120
HSRF
stata™



Residuals
20
10
0
-10
50
60
70
80
90
100
HP
stata™