Professor Brad Jones
University of Arizona
POL 681
Stata Tutorial Companion for Lecture Notes on Functional Form

```
. set text 250
. use funcform

. gr y x1
```

(SEE FIGURE 1)

```
. reg y x1

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,    18) =   60.05
       Model |  .179153697     1  .179153697           Prob > F      =  0.0000
    Residual |  .053701274    18  .002983404           R-squared     =  0.7694
-------------+------------------------------           Adj R-squared =  0.7566
       Total |  .232854971    19  .012255525           Root MSE      =  .05462

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0164135   .0021181     7.75   0.000     .0119636    .0208635
       _cons |   .5423647     .02445    22.18   0.000     .4909971    .5937322
------------------------------------------------------------------------------

. predict xb
(option xb assumed; fitted values)
(3 missing values generated)

. gr xb y x1, ylab xlab c(l) t1("Standard Model")
```

(**SEE FIGURE 2)**

Illustrations of Nonlinear Functional Forms

```
. gr y x1_1, ylab xlab  t1("Marginality 1") l1(" ") b2("x") saving(nonlin1, replace)
. gr y x1_2, ylab xlab  t1("Marginality 2") l1(" ") b2("x") saving(nonlin2, replace)
. gr ynew2 x1, ylab xlab  t1("Nonmonotonicity 1") l1(" ") b2("x") saving(nonlin3, replace)
. gr ynew2 x1, ylab xlab  t1("Nonmonotonicity 2") l1(" ") b2("x") saving(nonlin4, replace)

. gr using nonlin1 nonlin2 nonlin3 nonlin4, b2(Examples of Nonlinear Relationships) saving(nonlin,
replace)
```

(SEE FIGURE 3)

Illustration of OLS applied to Nonlinear Data

Example 1 (Top Left, Figure 3)

```
. reg y x1_1

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,    18) =   25.06
       Model |  .135527668     1  .135527668           Prob > F      =  0.0001
    Residual |  .097327303    18  .005407072           R-squared     =  0.5820
-------------+------------------------------           Adj R-squared =  0.5588
       Total |  .232854971    19  .012255525           Root MSE      =  .07353

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        x1_1 |  -.1925604   .0384622    -5.01   0.000    -.2733664   -.1117543
       _cons |   .7542487   .0190083    39.68   0.000     .7143137    .7941836
------------------------------------------------------------------------------
. predict xb1
```

```
(option xb assumed; fitted values)
(3 missing values generated)

. rvfplot, yline(0) t1("Marginality 1") b2("Y-Hat") ylab xlab saving(nonlin1xbres, replace)
```

Example 2 (Top Right, Figure 3)

```
. reg y x1_2

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,    18) =   22.58
       Model |  .129559535     1  .129559535           Prob > F      =  0.0002
    Residual |  .103295436    18  .005738635           R-squared     =  0.5564
-------------+------------------------------           Adj R-squared =  0.5318
       Total |  .232854971    19  .012255525           Root MSE      =  .07575

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        x1_2 |    .000676   .0001423     4.75   0.000     .0003771    .0009748
       _cons |   .6164295   .0254219    24.25   0.000     .5630202    .6698388
------------------------------------------------------------------------------

. predict xb2
(option xb assumed; fitted values)
(3 missing values generated)


. rvfplot, yline(0) t1("Marginality 2") b2("Y-Hat") ylab xlab saving(nonlin2xbres, replace)
```

Example 3 (Bottom Left, Figure 3)

```
. reg ynew x1

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,    18) =    0.01
       Model |  .000066316     1  .000066316           Prob > F      =  0.9293
    Residual |  .147413688    18  .008189649           R-squared     =  0.0004
-------------+------------------------------           Adj R-squared = -0.0551
       Total |  .147480004    19  .007762105           Root MSE      =   .0905

------------------------------------------------------------------------------
        ynew |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.0003158   .0035093    -0.09   0.929    -.0076886     .007057
       _cons |   .6171579   .0405094    15.23   0.000     .5320509    .7022649
------------------------------------------------------------------------------

. predict xb3
(option xb assumed; fitted values)
(3 missing values generated)

. rvfplot, yline(0) t1("Nonmonotonicity 1") b2("Y-Hat") ylab xlab saving(nonlin3xbres, replace)
```

Example 4 (Bottom Right, Figure 3)

```
. reg ynew2 x1

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,    18) =    0.00
       Model |  4.9414e-06     1  4.9414e-06           Prob > F      =  0.9935
    Residual |  1.31043676    18  .072802042           R-squared     =  0.0000
-------------+------------------------------           Adj R-squared = -0.0556
       Total |   1.3104417    19  .068970616           Root MSE      =  .26982

------------------------------------------------------------------------------
       ynew2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.0000862   .0104631    -0.01   0.994    -.0220684     .021896
```

```
       _cons |   1.664984   .1207799    13.79   0.000     1.411235    1.918733
------------------------------------------------------------------------------

. predict xb4
(option xb assumed; fitted values)
(3 missing values generated)


. rvfplot, yline(0) t1("Nonmonotonicity 2") b2("Y-Hat") ylab xlab saving(nonlin4xbres, replace)
```

Graphing Regression Functions

```
. gr y xb1 x1_1, ylab xlab c(.l)  t1("Marginality 1") l1(" ") b2("x") saving(nonlin1xb, replace)
. gr y xb2 x1_2, ylab xlab c(.l)  t1("Marginality 2") l1(" ") b2("x") saving(nonlin2xb, replace)
. gr ynew xb3 x1, ylab xlab c(.l)  t1("Nonmonotonicity 1") l1(" ") b2("x") saving(nonlin3xb,
replace)
. gr ynew2 xb4 x1, ylab xlab c(.l)  t1("Nonmonotonicity 2") l1(" ") b2("x") saving(nonlin4xb,
replace)
```

GRAPHING PREDICTED REGRESSION FUNCTIONS (Composite Graph)

```
. gr using nonlin1xb nonlin2xb nonlin3xb nonlin4xb, b2(Nonlinear Relationships w/ OLS)
saving(nonlinear2, replace)
```

(SEE FIGURE 4)

GRAPHING OLS RESIDUALS FROM 4 MODELS

```
gr using nonlin1xbres nonlin2xbres nonlin3xbres nonlin4xbres, b2(Nonlinear Relationships w/ OLS
Residuals) saving(nonlinear3, replace)
```

Log Transformation

Rerun Following Model

```
. reg y x1_1

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,    18) =   25.06
       Model |  .135527668     1  .135527668           Prob > F      =  0.0001
    Residual |  .097327303    18  .005407072           R-squared     =  0.5820
-------------+------------------------------           Adj R-squared =  0.5588
       Total |  .232854971    19  .012255525           Root MSE      =  .07353

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        x1_1 |  -.1925604   .0384622    -5.01   0.000    -.2733664   -.1117543
       _cons |   .7542487   .0190083    39.68   0.000     .7143137    .7941836
------------------------------------------------------------------------------
```

Generate Log-Transformation

```
. gen logx1_1=log(x1_1)
(3 missing values generated)
```

Estimate Regression Model with Log(X)

```
. reg y logx1_1

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,    18) =  189.92
       Model |  .212696422     1  .212696422           Prob > F      =  0.0000
    Residual |  .020158549    18  .001119919           R-squared     =  0.9134
```

```
-------------+------------------------------           Adj R-squared =  0.9086
       Total |  .232854971    19  .012255525           Root MSE      =  .03347

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     logx1_1 |  -.1115523   .0080945   -13.78   0.000    -.1285582   -.0945463
       _cons |   .4819501   .0179301    26.88   0.000     .4442804    .5196199
------------------------------------------------------------------------------
```

```
. predict xblog
(option xb assumed; fitted values)
(3 missing values generated)
```

Graphing Regression Function From Model with Log(x)

```
. gr y xblog logx1_1, ylab xlab b2("log x") t1("Regression with Log x")  c(.l) saving(logx1,
replace)
```

(See FIGURE 6)

Graphing Regression Function with respect to X (i.e. not log(x))

```
. gr y xblog x1_1, ylab xlab b2("x") t1("Regression with Log x")  c(.l) saving(logx1_2)
```

(SEE FIGURE 7)

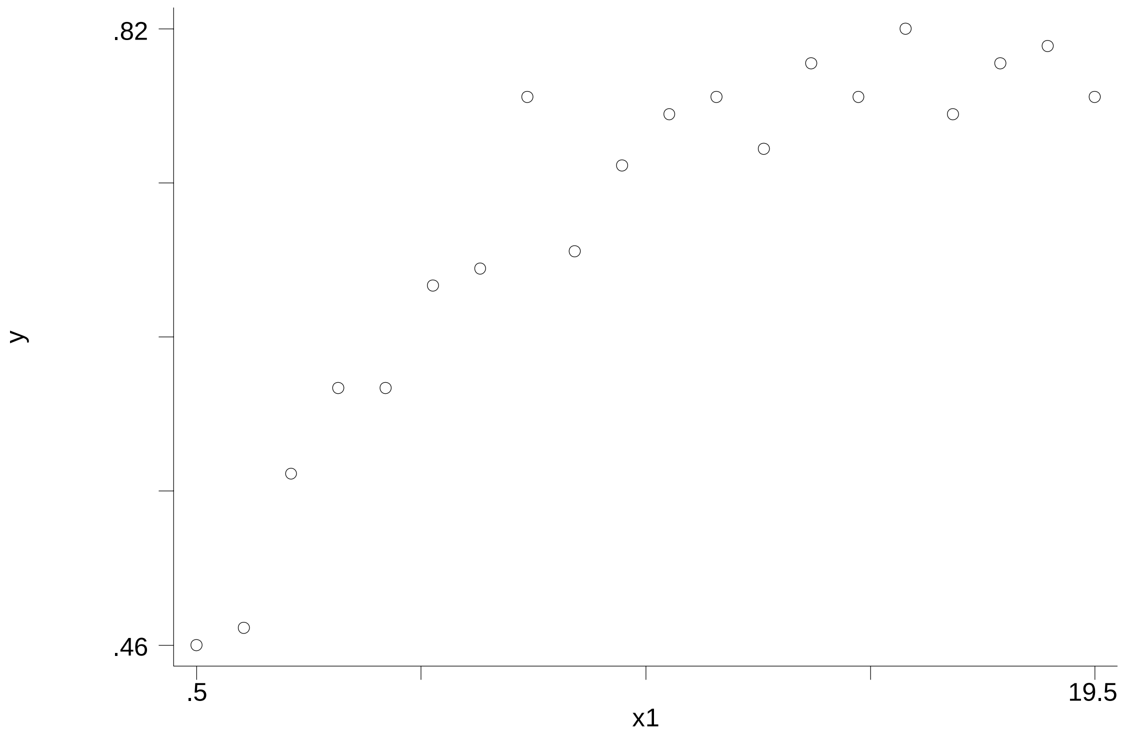


Figure 1: x1 and y

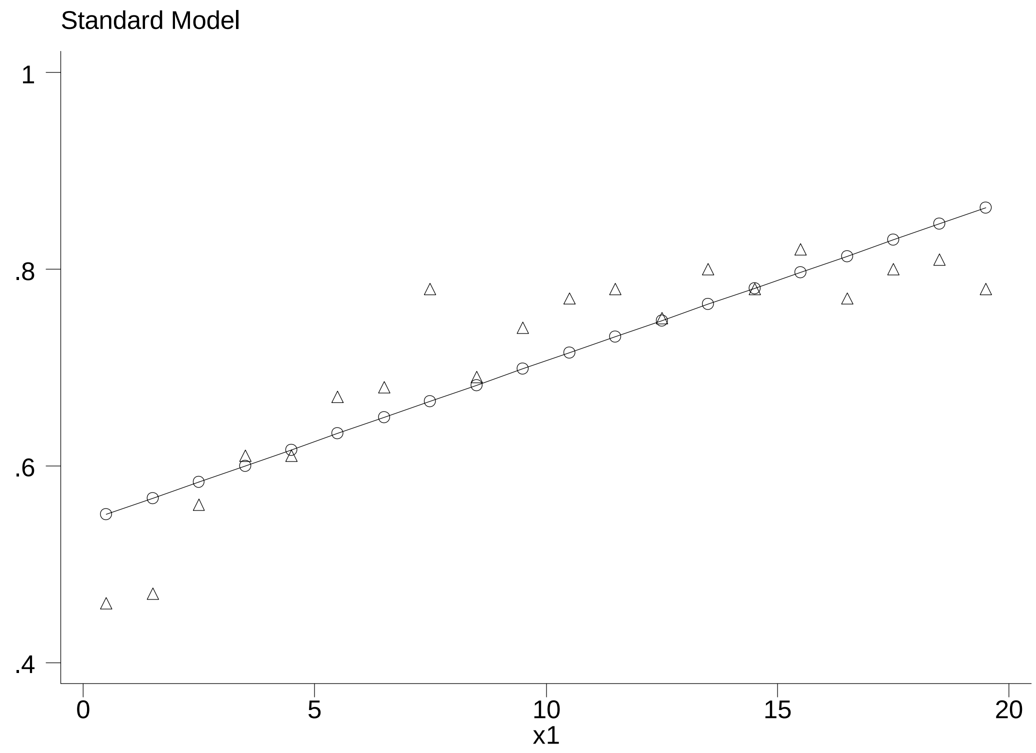


Figure 2: x1 and y with predicted regression function

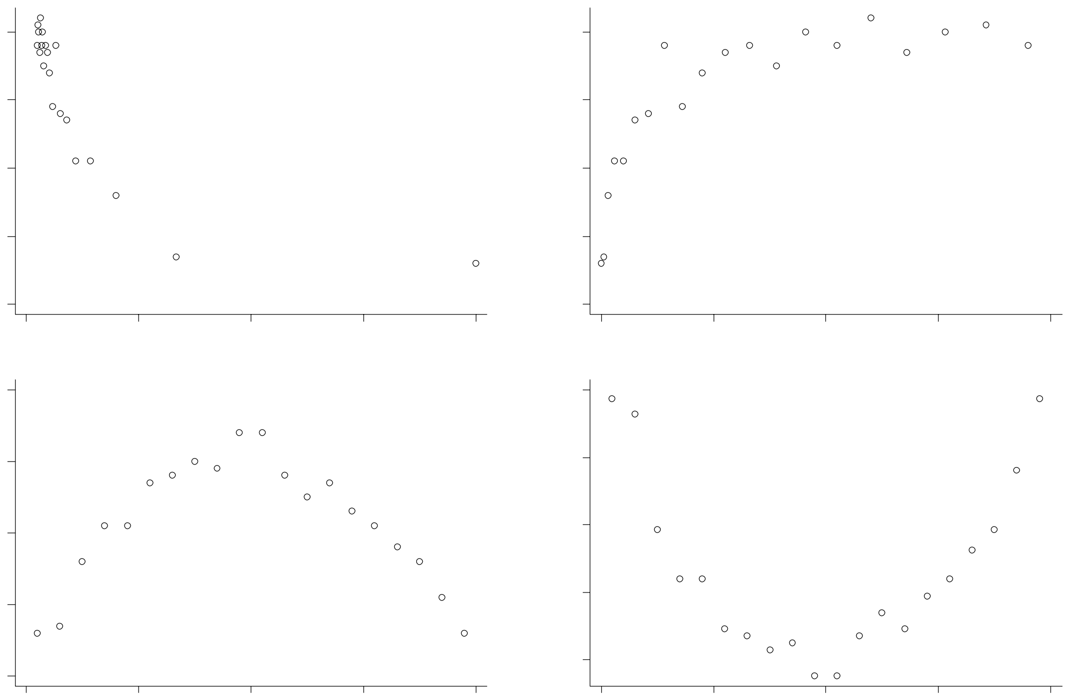


Figure 3: Nonlinear Relationships

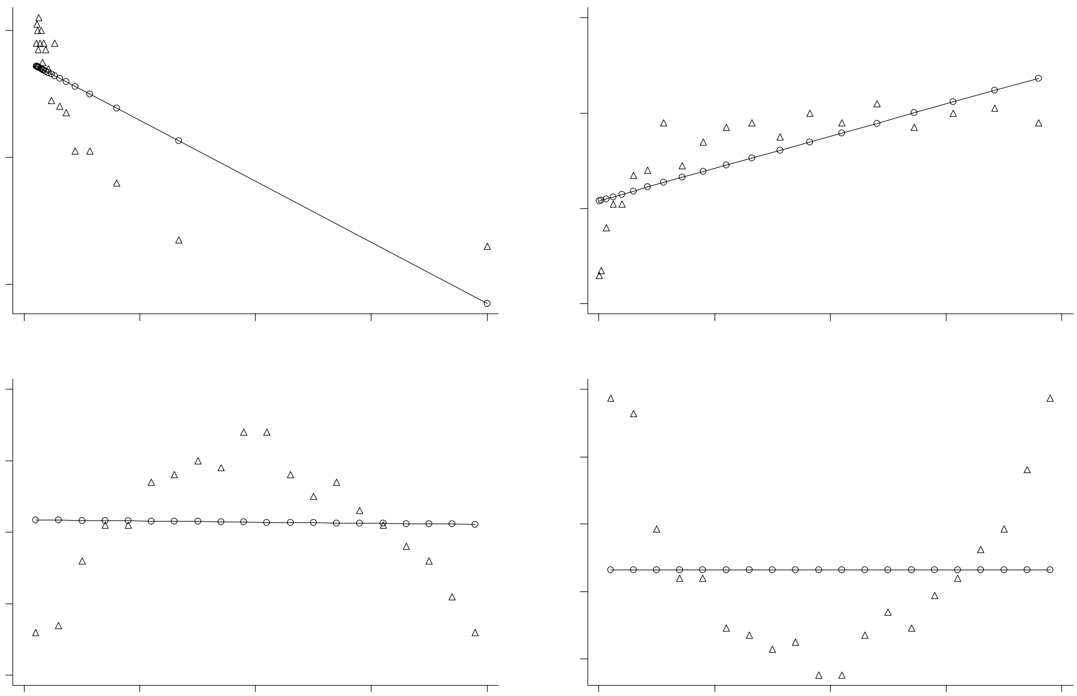


Figure 4: Nonlinear Relationships w/OLS Regression

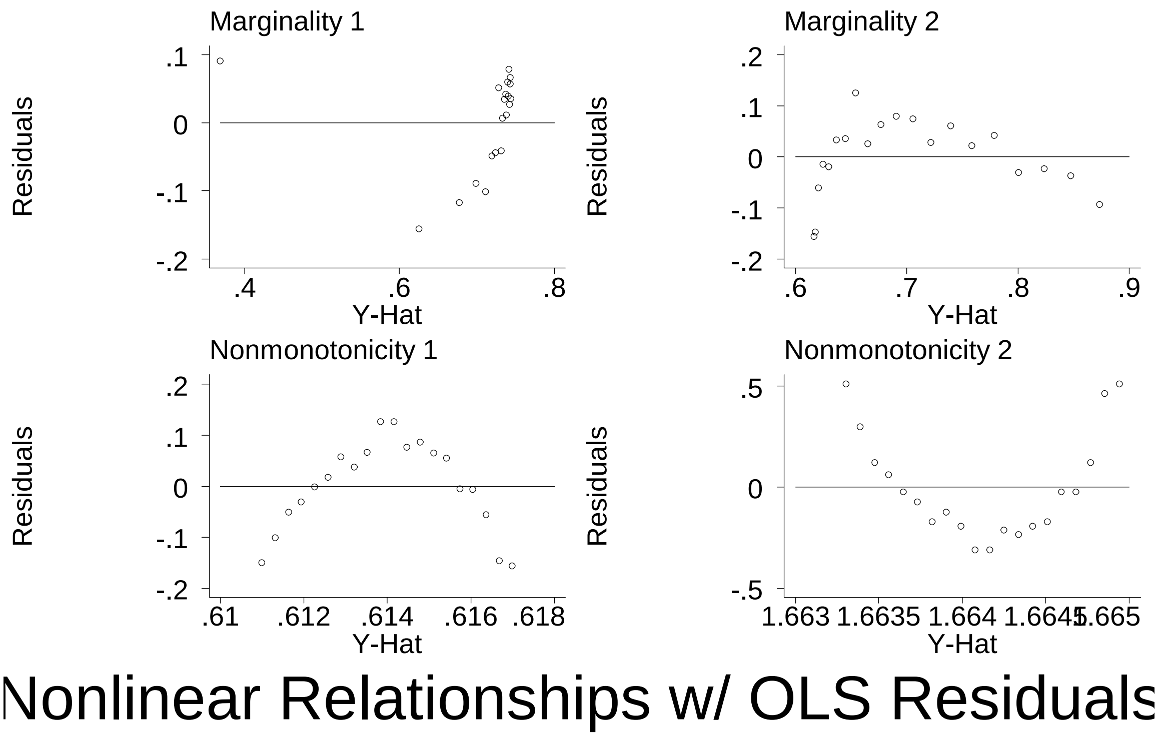


Figure 5: Nonlinear Relationship w/OLS Residuals

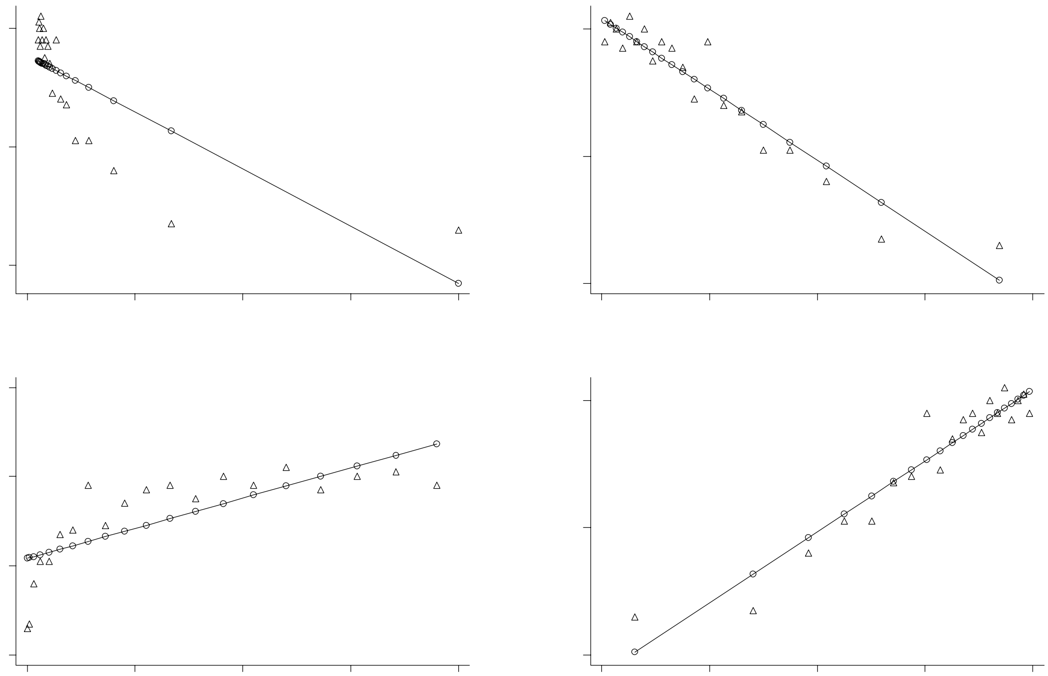


Figure 6: Regression with Unlogged and Logged X.
Top Left: Unlogged X; Top Right: Logged X; Bottom Left: Unlogged X; Bottom Right: Logged X

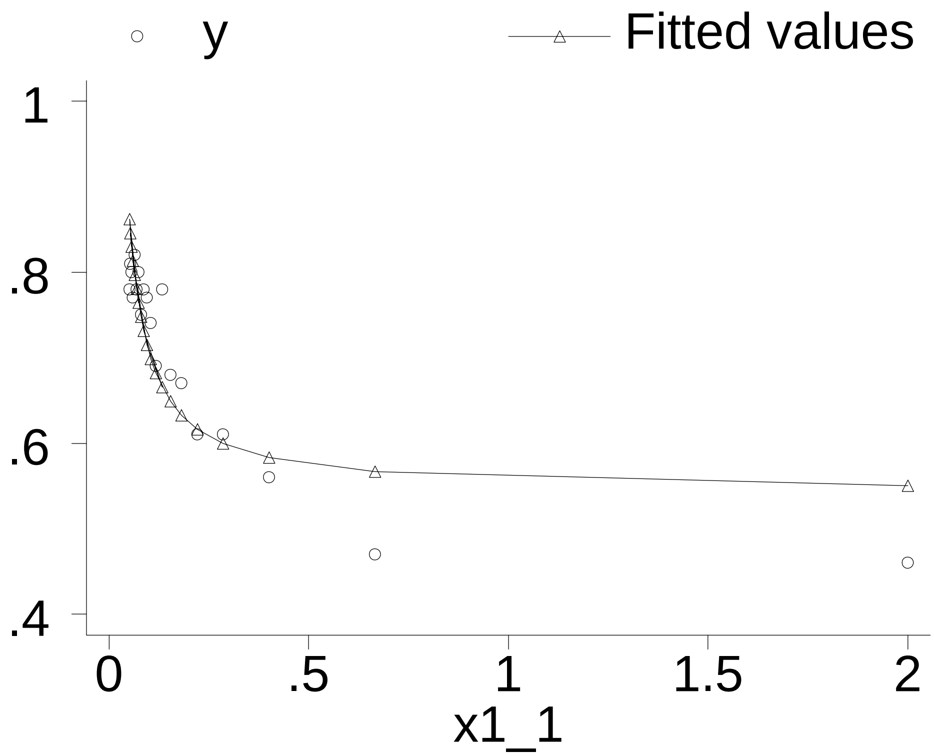


Figure 7: Predicted Regression Function graphed with Unlogged X

STATA CODE AND EXAMPLES OF QUADRATIC FORM

Begin with standard regression model of form:

reg ynew x1

```
      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  1,    18) =    0.01
       Model |  .000066316     1  .000066316           Prob > F      =  0.9293
    Residual |  .147413688    18  .008189649           R-squared     =  0.0004
-------------+------------------------------           Adj R-squared = -0.0551
       Total |  .147480004    19  .007762105           Root MSE      =   .0905

------------------------------------------------------------------------------
        ynew |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |  -.0003158   .0035093    -0.09   0.929    -.0076886     .007057
       _cons |   .6171579   .0405094    15.23   0.000     .5320509    .7022649
------------------------------------------------------------------------------
```

Now, consider quadratic.

Generate Quadratic Term

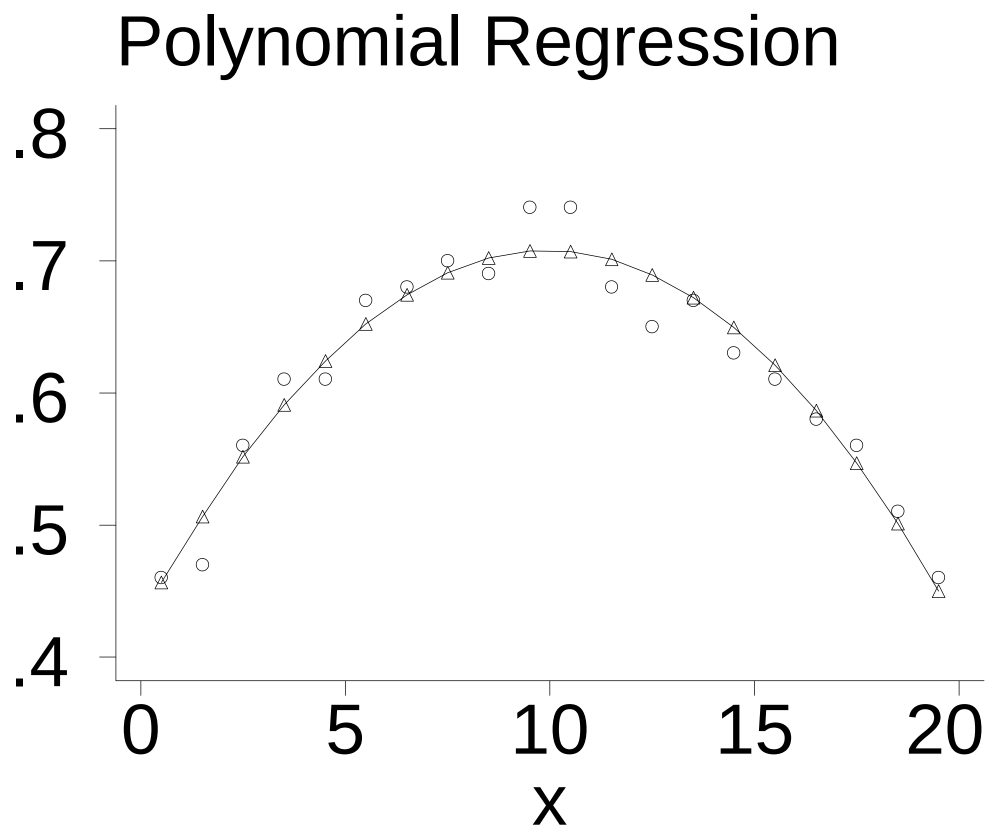
```
. gen x1squared=x1^2
```

Estimate quadratic regression

```
. reg ynew x1 x1squared

      Source |       SS       df       MS              Number of obs =      20
-------------+------------------------------           F(  2,    17) =  156.00
       Model |   .13985964     2   .06992982           Prob > F      =  0.0000
    Residual |  .007620364    17  .000448257           R-squared     =  0.9483
-------------+------------------------------           Adj R-squared =  0.9423
       Total |  .147480004    19  .007762105           Root MSE      =  .02117

------------------------------------------------------------------------------
        ynew |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   .0561208   .0032996    17.01   0.000     .0491592    .0630823
   x1squared |  -.0028218   .0001598   -17.66   0.000     -.003159   -.0024847
       _cons |   .4288009   .0142683    30.05   0.000     .3986975    .4589044
------------------------------------------------------------------------------

. predict xbquad
(option xb assumed; fitted values)
(3 missing values generated)
```

Graphing Regression Function

```
. gr ynew xbquad x1, ylab xlab b2("x") t1("Polynomial Regression ")  c(.l) saving(poly, replace)
```



Generating the conditional slope coefficients:

```
. gen condslope=_b[x1]+ 2*(_b[x1squared]*x1)
(3 missing values generated)
```

Generating the var-cov matrix of the parameters:

```
.  vce

             |       x1 x1squa~d    _cons
-------------+---------------------------
          x1 |  .000011
   x1squared | -5.1e-07  2.6e-08
       _cons | -.000041  1.7e-06  .000204
```

Generating Standard Errors on Conditional Slopes

```
. gen se_cs=sqrt(.000011 + (4*x1squared*.000000026) + (4*x1*-.00000051))
(3 missing values generated)
```

Generating t-ratios for conditional slopes:

```
.  gen tratio=condslope/se_cs
(3 missing values generated)
```

Obtaining Critical *t* for graphing purposes:

```
display invttail(17, .975)
```

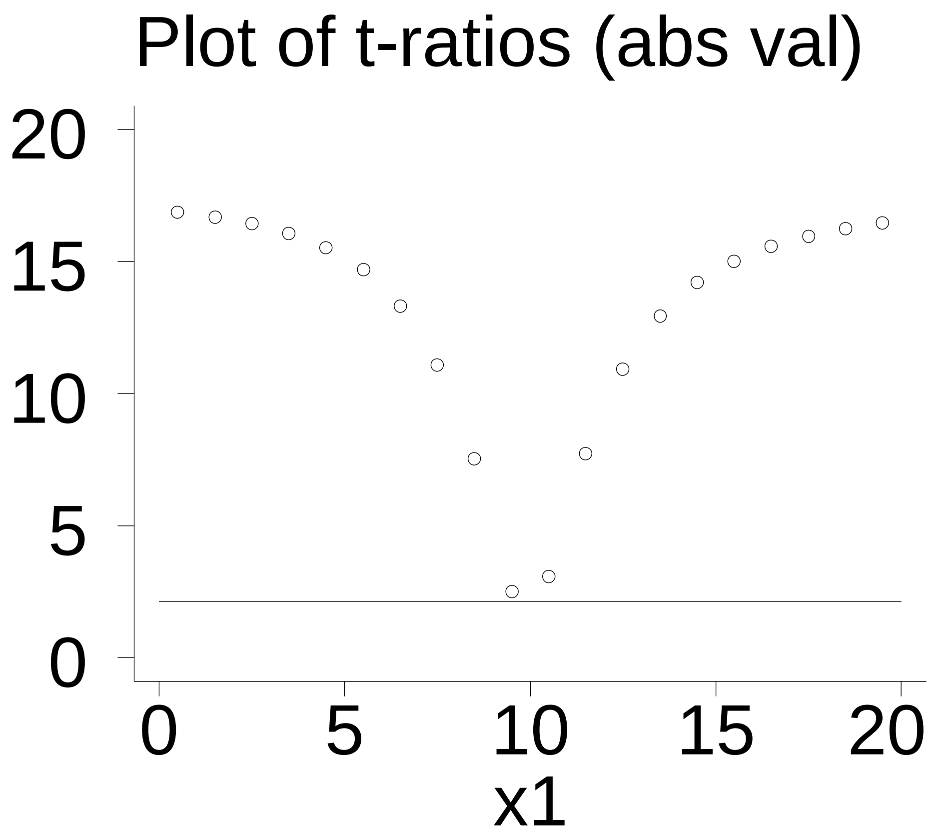
(This command returns the value -2.11)

Graphing t-ratios

```
. gr tratio x1, yline(2.11, -2.11) ylab xlab t1("Plot of t-ratios") l1(" ") saving(tratios, replace)
```

Graphing t-ratios (absolute values)

```
gr abst x1, yline(2.11) ylab xlab t1("Plot of t-ratios (abs val)") l1(" ") saving(tratios, replace)
```

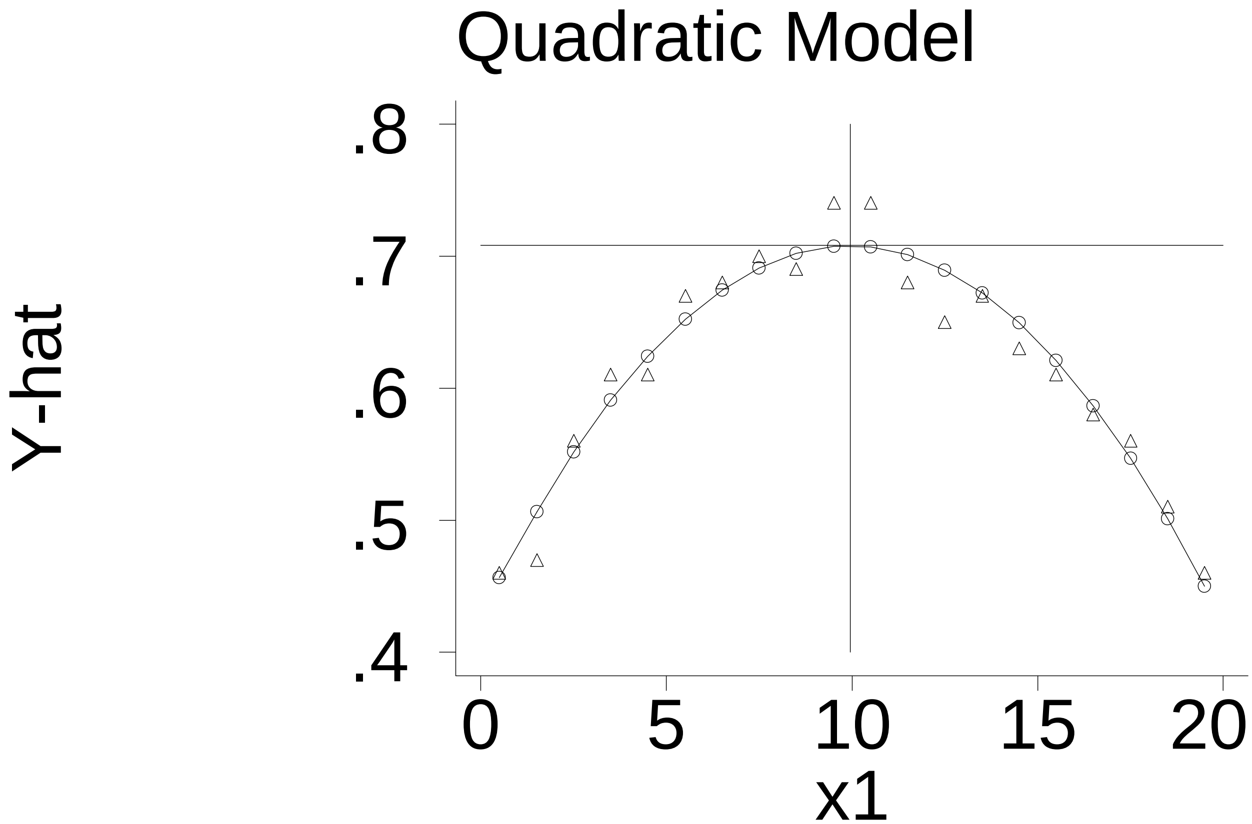


Displaying Inflection Point

```
. display -_b[x1]/(2*_b[x1squared])
9.9440451
```

Graphing Regression Function w/Inflection Point Noted

```
. gr xbquad ynew x1, c(l) ylab xlab t1("Quadratic Model") xline(9.944) yline(.708) l1("Y-hat")
saving(inflection, replace)
```


Quadratic Model
Y-hat
.8
.7
.6
.5
.4
0
5
10
15
20
x1