```r
#
#   Introduction to Econometrics
#
#   Final Project
#
#   Due Tuesday, August 2, 2017
#
#   Code Written by Michael Barrett
#
###################################################################
#
# Start with a clean slate!
closeAllConnections()
rm(list = ls())
#
#   Estimating Keynesian Consumption Function
#
# Load .csv
cons <- read.csv("W:/2017 Spring-Summer/Introduction to Statistics and Econometrics/Final Project/Source Data/cons.csv")
# Make the regression
cons_HAT <- lm(C ~ Yd, data = cons)
# Make a plot with the fitted line
plot(x = cons$Yd, y = cons$C, xlab = "Disposable Income", ylab = "Consumption", main = "Keynesian Consumption Function, Actual and Fitted")
abline(cons_HAT)
#
#   Estimating Cobb-Douglas Function
#
# Load .csv
cobb.douglas <- read.csv("W:/2017 Spring-Summer/Introduction to Statistics and Econometrics/Final Project/Source Data/cobb-douglas.csv")
# Make the regression
Y_HAT <- lm(log(GDP) ~ log(Labor) + log(Capital), data = cobb.douglas)
#
#   Determinants of Cross Country GDP Growth Rates
#
#  Load .csv
growth <- read.csv("W:/2017 Spring-Summer/Introduction to Statistics and Econometrics/Final Project/Source Data/growth.csv")
# Make the regression
GDPgrowth_HAT <- lm(GDPgrowth ~ initGDP + MSE + FSE + MHE + FHE + life_exp + eduGDP + invGDP + govGDP + pol, data = growth)
# Create a list for variables
variables <- list()
variables[[1]] <- variable.names(GDPgrowth_HAT)[2:GDPgrowth_HAT$rank]
# Number of variables will be useful to have on-hand
k <- GDPgrowth_HAT$rank - 1
# Make a blank list to store info for models with fewer variables
model_info <- list()
# This algorithm will create models omitting the variable with the highest p-value from the next model
```

```r
for (i in 1:k) {
    # Write the model out from the given variables
    model_info$trial_model[i] <- paste("growth$GDPgrowth ~ ",
        paste("growth$", variables[[i]], sep = "", collapse = " + "), sep = "")
    # Estimate values for the model
    GDPgrowth_HAT_trial <- lm(as.formula(model_info$trial_model[i]))
    # Store the adjusted R^2
    model_info$adj_R_sq[i] <- summary(GDPgrowth_HAT_trial)$adj.r.squared
    # Omit the variable with the highest p-value from the next model
    variables[[i + 1]] <- variables[[i]][-(which.max(summary(GDPgrowth_HAT_trial)
      $coefficients[2:(k-i+2), 4]))]
}
#
#   Estimating Crime Model
#
# Load .csv
crime <- read.csv("W:/2017 Spring-Summer/Introduction to Statistics and Econometrics/
  Final Project/Source Data/crime.csv")
# Make the regression
crime_HAT <- lm(crime ~ pov + metro + popdens, data = crime)
# Regress against log(crime)
log_crime_HAT <- lm(log(crime) ~ pov + metro + popdens, data = crime)
#
#   Estimating Wage Model
#
# Load .csv
wage <- read.csv("W:/2017 Spring-Summer/Introduction to Statistics and Econometrics/
  Final Project/Source Data/wage.csv")
# Add variables to the data
wage$exper_2 <- wage$exper ^ 2
wage$tenure_2 <- wage$tenure^2
# Make the regression
log_wage_HAT <- lm(log(wage) ~ educ + exper + exper_2 + tenure + tenure_2 + married +
  black + south + urban, data = wage)
# Marginal effect of each
ME_exper <- log_wage_HAT$coefficients["exper"] + 2 * log_wage_HAT$coefficients
  ["exper_2"] * c(0:49)
ME_tenure <- log_wage_HAT$coefficients["tenure"] + 2 * log_wage_HAT$coefficients
  ["tenure_2"] * c(0:49)
#
#######################################################################################
  ##################
#
#   Time to answer questions and write them to an output file
#
# Open the .txt
sink("Final_Project_Output.txt")
# Print responses for #1
cat("1. Estimating Keynesian Consumption Function\n\n")
# Print the model with info
model <- cons_HAT
cat("Fitted Model:\n")
```

```
cat(all.vars(model$call)[1], " = ", round(coefficients(model)[1], 4), " + ",
    paste(sprintf("%.4f * %s", coefficients(model)[-1], names(coefficients(model)
      [-1])),
          collapse = " + "))
cat("\nCoefficients:\n")
print(summary(model)$coefficients)
cat("R^2: ", summary(model)$r.squared, "  |  Adjusted R^2: ", summary(model)
  $adj.r.squared)
cat("\nNumber of observations: ", model$df.residual + model$rank)
cat("\nGraph is attached.")
# Print responses for #2
cat("\n\n\n2. Estimating Cobb-Douglas Function\n\n")
# Print the model with info
model <- Y_HAT
cat("Fitted Model:\n")
cat(all.vars(model$call)[1], " = ", round(coefficients(model)[1], 4), " + ",
    paste(sprintf("%.4f * %s", coefficients(model)[-1], names(coefficients(model)
      [-1])),
          collapse = " + "))
cat("\nCoefficients:\n")
print(summary(model)$coefficients)
cat("R^2: ", summary(model)$r.squared, "  |  Adjusted R^2: ", summary(model)
  $adj.r.squared)
cat("\nNumber of observations: ", model$df.residual + model$rank)
cat("\n\nA = ", exp(Y_HAT$coefficients[1]), ", alpha = ", Y_HAT$coefficients[2], ",
  beta = ", Y_HAT$coefficients[3])
# Print responses for #3
cat("\n\n\n3. Determinants of Cross-Country GDP Growth Rates\n\n")
# Print the model with info
model <- GDPgrowth_HAT
cat("Fitted Model:\n")
cat(all.vars(model$call)[1], " = ", round(coefficients(model)[1], 4), " + ",
    paste(sprintf("%.4f * %s", coefficients(model)[-1], names(coefficients(model)
      [-1])),
          collapse = " + "))
cat("\nCoefficients:\n")
print(summary(model)$coefficients)
cat("R^2: ", summary(model)$r.squared, "  |  Adjusted R^2: ", summary(model)
  $adj.r.squared)
cat("\nNumber of observations: ", model$df.residual + model$rank)
cat("\n\nThe following models have each have eliminated the variable from the previous
   with the highest P-value.")
cat("\nThe Adjusted R^2 has been given in each case for comparison, the highest value
  represents the best model.\n\n")
cat(paste(model_info$trial_model, "\nAdjusted R^2 = ", model_info$adj_R_sq, "\n"))
# Print responses for #4
cat("\n\n\n4. Estimating Crime Model\n\n")
# Print the model with info
model <- crime_HAT
cat("Fitted Model 1:\n")
cat(all.vars(model$call)[1], " = ", round(coefficients(model)[1], 4), " + ",
    paste(sprintf("%.4f * %s", coefficients(model)[-1], names(coefficients(model)
```

```
      [-1])),
          collapse = " + "))
cat("\nCoefficients:\n")
print(summary(model)$coefficients)
cat("R^2: ", summary(model)$r.squared, "  |  Adjusted R^2: ", summary(model)
  $adj.r.squared)
cat("\nNumber of observations: ", model$df.residual + model$rank)
model <- log_crime_HAT
cat("\n\nFitted Model 2:\n")
cat("log(crime) = ", round(coefficients(model)[1], 4), " + ",
    paste(sprintf("%.4f * %s", coefficients(model)[-1], names(coefficients(model)
      [-1])),
          collapse = " + "))
cat("\nCoefficients:\n")
print(summary(model)$coefficients)
cat("R^2: ", summary(model)$r.squared, "  |  Adjusted R^2: ", summary(model)
  $adj.r.squared)
cat("\nNumber of observations: ", model$df.residual + model$rank)
cat("\n\nInterpretations:")
cat("\nA 1 point ceterus paribus increase in the percentage impoverished would yeild a
   ",
    100 * coefficients(model)[2], " percent increase in crime.")
cat("\nA 1 point ceterus paribus increase in the percentage metropolitan would yeild a
   ",
    100 * coefficients(model)[3], " percent increase in crime.")
cat("\nA 1 person per square mile ceterus paribus increase in population density would
   yeild a ",
    100 * coefficients(model)[3], " percent increase in crime.")
# Print responses for extra credit question
cat("\n\n\nE.C. Estimating Wage Model\n\n")
# Print the model with info
model <- log_wage_HAT
cat("Fitted Model 1:\n")
cat("log(wage) = ", round(coefficients(model)[1], 4), " + ",
    paste(sprintf("%.4f * %s", coefficients(model)[-1], names(coefficients(model)
      [-1])),
          collapse = " + "))
cat("\nCoefficients:\n")
print(summary(model)$coefficients)
cat("R^2: ", summary(model)$r.squared, "  |  Adjusted R^2: ", summary(model)
  $adj.r.squared)
cat("\nNumber of observations: ", model$df.residual + model$rank)
cat("\n\nThe marginal effect of the 10th year of experience on log(wage) is ",
  ME_exper[10])
cat("\nThe marginal effect of the 10th year of tenure on log(wage) is ", ME_tenure
  [10])
cat("\n\nBeing black vs. non-black changes wage by ", 100 * coefficients(model)[8], "
  percent.")
cat("\nLiving in the South vs. elsewhere changes wage by ", 100 * coefficients(model)
  [9], " percent.")
cat("\nLiving in an urban area vs. a rural area changes wage by ", 100 * coefficients
  (model)[10], " percent.")
```

```
cat("\n\nWe can observe interactions of dummy variables by adding them.")
cat("\nMarried blacks can be expected to earn ", 100 * (coefficients(model)[7] *
  coefficients(model)[8]),
    " percent different wages than single non-blacks.")
sink()
```

1. Estimating Keynesian Consumption Function

```
Fitted Model:
C  =  -49.7522  +  0.9174 * Yd
Coefficients:
               Estimate   Std. Error    t value     Pr(>|t|)
(Intercept) -49.7522318 17.876691035  -2.783078 6.872126e-03
Yd            0.9173958  0.005411752 169.519190 1.965090e-95
R^2:  0.9975008   |  Adjusted R^2:  0.997466
Number of observations:  74
Graph is attached.
```

2. Estimating Cobb-Douglas Function

```
Fitted Model:
GDP  =  -3.3385  +  1.4988 * log(Labor) + 0.4899 * log(Capital)
Coefficients:
               Estimate Std. Error   t value     Pr(>|t|)
(Intercept)  -3.3384555  2.4495082 -1.362908 0.1979394665
log(Labor)    1.4987669  0.5398026  2.776509 0.0167584828
log(Capital)  0.4898585  0.1020435  4.800487 0.0004331776
R^2:  0.8890304   |  Adjusted R^2:  0.8705354
Number of observations:  15

A =  0.03549173 , alpha =  1.498767 , beta =  0.4898585
```

3. Determinants of Cross-Country GDP Growth Rates

```
Fitted Model:
GDPgrowth  =  -0.0378  +  -0.0184 * initGDP + 0.0010 * MSE + -0.0083 * FSE + 0.1231 * MHE + -0.1245 * FHE + 0.0483 * life_exp + 0.2716
* eduGDP + 0.0562 * invGDP + -0.1225 * govGDP + -0.0005 * pol
Coefficients:
                 Estimate  Std. Error     t value     Pr(>|t|)
(Intercept) -0.0378375650 0.061778059 -0.61247578 0.5425374067
initGDP     -0.0184378783 0.004558853 -4.04441184 0.0001520132
MSE          0.0010498612 0.010548703  0.09952514 0.9210531137
FSE         -0.0083014020 0.010762527 -0.77132465 0.4435416441
MHE          0.1231189845 0.049398947  2.49234027 0.0154697443
FHE         -0.1244888172 0.057359790 -2.17031507 0.0339518267
life_exp     0.0482844763 0.019064781  2.53265305 0.0139539716
eduGDP       0.2715559171 0.201546079  1.34736393 0.1829293631
invGDP       0.0562215097 0.030366125  1.85145484 0.0690270026
govGDP      -0.1224601668 0.043683064 -2.80337863 0.0068019612
pol         -0.0005337157 0.016221240 -0.03290227 0.9738617301
R^2:  0.4222773   |  Adjusted R^2:  0.3259902
Number of observations:  71
```

The following models have each have eliminated the variable from the previous with the highest P-value.
The Adjusted R^2 has been given in each case for comparison, the highest value represents the best model.

```
growth$GDPgrowth ~ growth$initGDP + growth$MSE + growth$FSE + growth$MHE + growth$FHE + growth$life_exp + growth$eduGDP + growth$invGDP
+ growth$govGDP + growth$pol
Adjusted R^2 =  0.325990179422733
 growth$GDPgrowth ~ growth$initGDP + growth$MSE + growth$FSE + growth$MHE + growth$FHE + growth$life_exp + growth$eduGDP +
growth$invGDP + growth$govGDP
Adjusted R^2 =  0.337027559172956
 growth$GDPgrowth ~ growth$initGDP + growth$FSE + growth$MHE + growth$FHE + growth$life_exp + growth$eduGDP + growth$invGDP +
growth$govGDP
Adjusted R^2 =  0.347624749898788
 growth$GDPgrowth ~ growth$initGDP + growth$FSE + growth$MHE + growth$FHE + growth$life_exp + growth$invGDP + growth$govGDP
Adjusted R^2 =  0.338527921273906
 growth$GDPgrowth ~ growth$initGDP + growth$MHE + growth$FHE + growth$life_exp + growth$invGDP + growth$govGDP
Adjusted R^2 =  0.331538687357407
 growth$GDPgrowth ~ growth$initGDP + growth$MHE + growth$FHE + growth$life_exp + growth$govGDP
Adjusted R^2 =  0.294285981714504
 growth$GDPgrowth ~ growth$initGDP + growth$MHE + growth$FHE + growth$life_exp
Adjusted R^2 =  0.239567152287585
 growth$GDPgrowth ~ growth$initGDP + growth$MHE + growth$FHE
Adjusted R^2 =  0.140543070354936
 growth$GDPgrowth ~ growth$MHE + growth$FHE
Adjusted R^2 =  0.132852598911581
 growth$GDPgrowth ~ growth$MHE
Adjusted R^2 =  0.00558044106156197
```

4. Estimating Crime Model

```
Fitted Model 1:
crime  =  1638.805  +  83.2998 * pov + 31.1747 * metro + 0.4563 * popdens
Coefficients:
                Estimate  Std. Error  t value     Pr(>|t|)
(Intercept) 1638.8046250 774.6802653 2.115459 0.0397188728
pov           83.2997846  35.4091183 2.352495 0.0228868083
metro         31.1746603   7.3552889 4.238401 0.0001043325
popdens        0.4563047   0.1279502 3.566269 0.0008451603
R^2:  0.562504   |  Adjusted R^2:  0.5345787
Number of observations:  51

Fitted Model 2:
log(crime) =  7.8125  +  0.0142 * pov + 0.0069 * metro + 0.0000 * popdens
Coefficients:
                Estimate   Std. Error   t value     Pr(>|t|)
(Intercept) 7.812496e+00 1.622168e-01 48.160845 1.176785e-41
```

pov	1.415772e-02	7.414610e-03	1.909435	6.232082e-02
metro	6.890744e-03	1.540185e-03	4.473971	4.854949e-05
popdens	4.344114e-05	2.679255e-05	1.621389	1.116247e-01

R^2: 0.443232 | Adjusted R^2: 0.4076936
Number of observations: 51

Interpretations:
A 1 point ceterus paribus increase in the percentage impoverished would yeild a 1.415772 percent increase in crime.
A 1 point ceterus paribus increase in the percentage metropolitan would yeild a 0.6890744 percent increase in crime.
A 1 person per square mile ceterus paribus increase in population density would yeild a 0.6890744 percent increase in crime.

E.C. Estimating Wage Model

Fitted Model 1:
log(wage) = 5.3587 + 0.0643 * educ + 0.0172 * exper + -0.0001 * exper_2 + 0.0249 * tenure + -0.0008 * tenure_2 + 0.1985 * married + -0.1907 * black + -0.0912 * south + 0.1854 * urban
Coefficients:

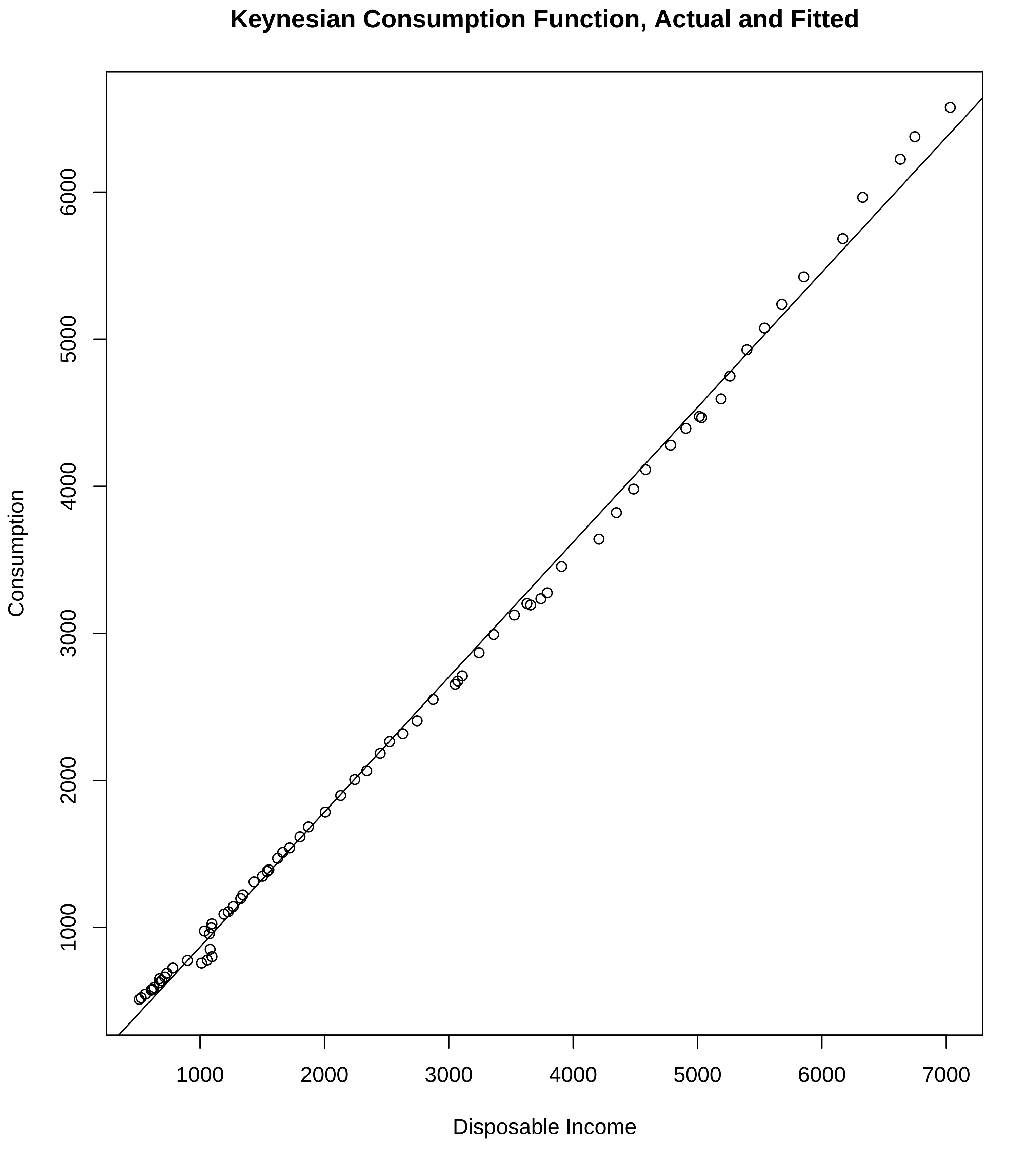
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.3586755928	0.1259142936	42.5581198	4.658568e-220
educ	0.0642760588	0.0063114760	10.1839979	3.691563e-23
exper	0.0172145971	0.0126137688	1.3647465	1.726646e-01
exper_2	-0.0001138015	0.0005318714	-0.2139643	8.306220e-01
tenure	0.0249290622	0.0081296615	3.0664330	2.229260e-03
tenure_2	-0.0007964474	0.0004710134	-1.6909229	9.118840e-02
married	0.1985469819	0.0391103437	5.0765849	4.645856e-07
black	-0.1906636159	0.0377011200	-5.0572401	5.128487e-07
south	-0.0912153107	0.0262356205	-3.4767735	5.311015e-04
urban	0.1854240938	0.0269585368	6.8781216	1.115870e-11

R^2: 0.2549576 | Adjusted R^2: 0.2477085
Number of observations: 935

The marginal effect of the 10th year of experience on log(wage) is 0.01516617
The marginal effect of the 10th year of tenure on log(wage) is 0.01059301

Being black vs. non-black changes wage by -19.06636 percent.
Living in the South vs. elsewhere changes wage by -9.121531 percent.
Living in an urban area vs. a rural area changes wage by 18.54241 percent.

We can observe interactions of dummy variables by adding them.
Married blacks can be expected to earn -3.785569 percent different wages than single non-blacks.


Keynesian Consumption Function, Actual and Fitted
Consumption
1000
2000
3000
4000
5000
6000
1000
2000
3000
4000
5000
6000
7000
Disposable Income