

```

#
#   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

```

```

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$rank)
cat("\nGraph is attached.")
# Print responses for #2
cat("\n\nn2. Estimating Cobb-Douglas Function\n\n")
# Print the model with info
model <- Y_HAT
cat("Fitted Model:\n")
cat(all.vars(model$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\n $\alpha$  = ", exp(Y_HAT$coefficients[1]), ",  $\alpha$  = ", Y_HAT$coefficients[2], ",
    beta = ", Y_HAT$coefficients[3])
# Print responses for #3
cat("\n\nn3. 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$rank)
cat("\n\nThe following models have each have eliminated the variable from the previous
    with the highest P-value.")
cat("\n\nThe Adjusted R^2 has been given in each case for comparison, the highest value
    represents the best model.\n\n")
cat(paste(model_info$trial_model, "\nAdjusted R^2 = ", model_info$adj_R_sq, "\n"))
# Print responses for #4
cat("\n\nn4. Estimating Crime Model\n\n")
# Print the model with info
model <- crime_HAT
cat("Fitted Model 1:\n")
cat(all.vars(model$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$rank)
cat("\n\nInterpretations:")
cat("\nA 1 point ceterus paribus increase in the percentage impoverished would yeild a",
    "
    100 * coefficients(model)[2], " percent increase in crime.")
cat("\nA 1 point ceterus paribus increase in the percentage metropolitan would yeild a",
    "
    100 * coefficients(model)[3], " percent increase in crime.")
cat("\nA 1 person per square mile ceterus paribus increase in population density would
    yeild a ",
    100 * coefficients(model)[3], " percent increase in crime.")
# Print responses for extra credit question
cat("\n\nE.C. Estimating Wage Model\n\n")
# Print the model with info
model <- log_wage_HAT
cat("Fitted Model 1:\n")
cat("log(wage) = ", round(coefficients(model)[1], 4), " + ",
    paste(sprintf("%.4f * %s", coefficients(model)[-1], names(coefficients(model)
        [-1])),
        collapse = " + "))
cat("\nCoefficients:\n")
print(summary(model)$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("\n\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("\n\nLiving in the South vs. elsewhere changes wage by ", 100 * coefficients(model)
    [9], " percent.")
cat("\n\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()
```