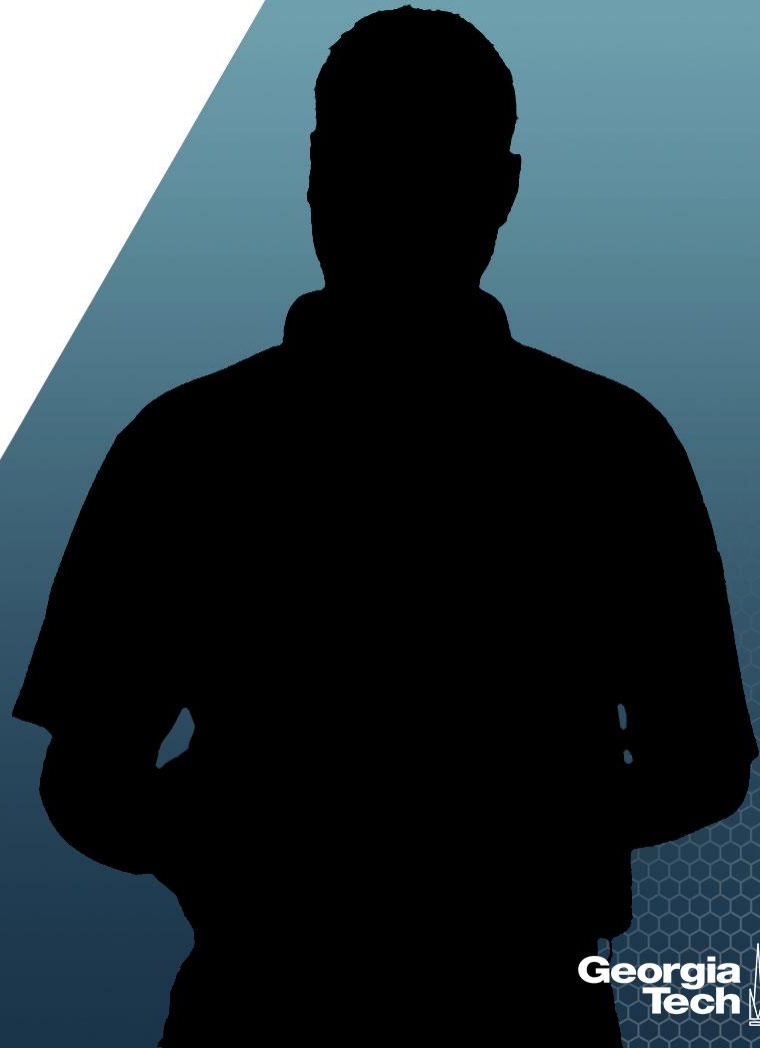# Regression Analysis
## Simple Linear Regression

**Nicoleta Serban, Ph.D.**
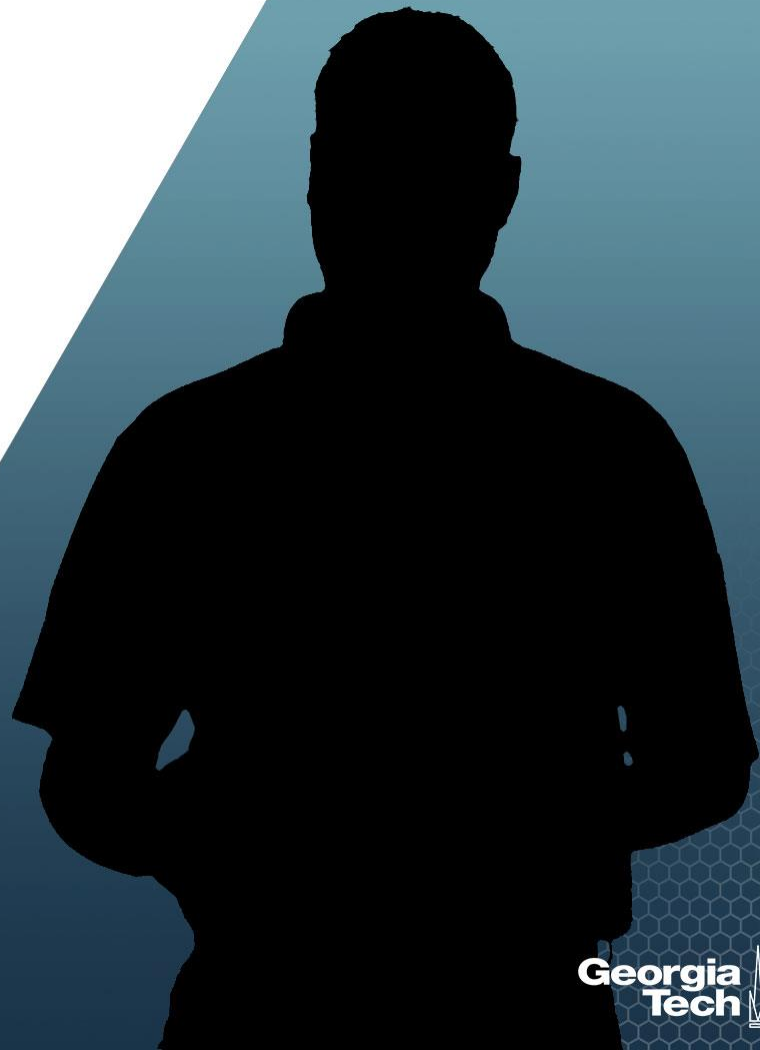*Professor*

School of Industrial and Systems Engineering

## Testing the Theory of Purchasing Power Parity (Part 1)

# About This Lesson

# Testing the Theory of Purchasing Power Parity

**Georgia Tech**

# Regression Variables

**_Response Variable:_** Average annual change in the exchange rate

$$\frac{\ln(\text{Exclange Rate for 2012}) - \ln(\text{Exchange Rate for 1975})}{\text{no. years}} \% = \textit{Annualized Percentage Change}$$

**_Predicting Variable:_** Average of the _difference in annual inflation rates_ for a country vs U.S.

$$\frac{1}{\text{no. years}} \sum_{y=1975}^{2012} \left(\text{Inflation}_y(\text{U.S.}) - \text{Inflation}_y(\text{Country})\right)$$

| Country | Inflation.difference | Exchange.rate.change | Developed |
|---------|---------------------|---------------------|-----------|
| Australia | -1.2351 | -3.1870 | 1 |
| Austria | 1.5508 | 1.4781 | 1 |
| Belgium | 1.0371 | 0.0395 | 1 |
| Canada | 0.0461 | -1.6416 | 1 |
| Chile | -18.4126 | -20.6329 | 0 |

Georgia Tech

# Read the Data in R

**# Use read.table R command: pay attention to the file type to use the correct read file!#**

*ppp = read.table("ppp.dat",sep="\t", header=T, row.names=NULL)*

**## How many countries?**

*dim(ppp)*

[1] 40  4

**##  Brazil is an outlier and it was not included in the data set initially; I am adding it back as follows**

*Addp = data.frame("Brazil",-76,-73,0)*

*names(addp) = names(ppp)*

**## Save the data variables to be recognized by R as separate variables**

*ppp = data.frame(rbind(ppp,addp))*

*attach(ppp)*

**## Re-label the 'Developed' column to differentiate between Developed and Developing countries**

*Developed[Developed==1] = "Developed"*

*Developed[Developed==0] = "Developing"*

# Exploratory Data Analysis in R

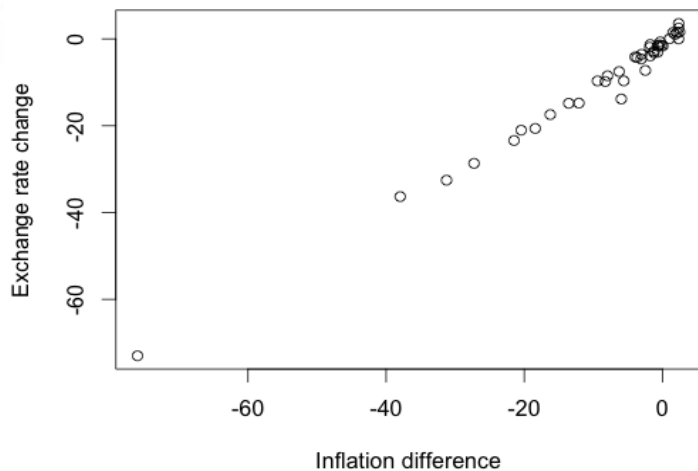**# Evaluate the Linear Relationship: Perform a scatter plot of the two variables**

```
plot(Inflation.difference,Exchange.rate.change,  main="Scatterplot of
Exchange rate change vs Inflation difference", xlab="Inflation
difference",ylab="Exchange rate change")
```

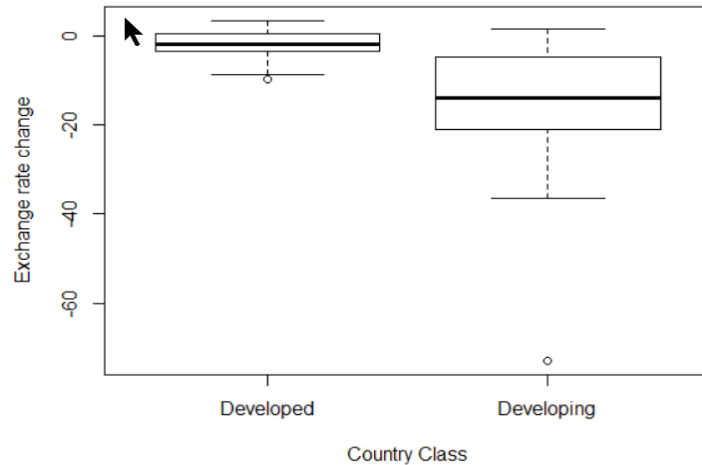**# Evaluate differences between developed and developing countries**

```
boxplot(Exchange.rate.change~as.factor(Developed), main="Boxplot of
Exchange rate change by Developed vs Developing
Countries",xlab="Country Class",ylab="Exchange rate change")
```

# Exploratory Data Analysis in R

# Fitting Linear Regression in R

*pppa = lm(Exchange.rate.change ~ Inflation.difference)  ## regression model*
*summary(pppa)*

Coefficients:

|  | Estimate | Std. Error | t value | Pr(>|t|) |
|---|---|---|---|---|
| (Intercept) | -1.51930 | 0.29415 | -5.165 | 7.43e-06 |
| Inflation.difference | 0.96185 | 0.01781 | 53.991 | < 2e-16 |
| --- | | | | |

$\hat{\beta}_0 = -1.5193$, $se(\hat{\beta}_0) = 0.2941$
$\hat{\beta}_1 = 0.9618$, $se(\hat{\beta}_1) = 0.0178$
Test for statistical significance:
$\beta_0$: t-value= -5.165, p-value ≈ 0
$\beta_1$: t-value= 53.991, p-value ≈ 0

Residual standard error: 1.646 on 39 degrees of freedom
Multiple R-squared: 0.9868,    Adjusted R-squared:  0.9865
F-statistic:  2915 on 1 and 39 DF,  p-value: < 2.2e-16

$\hat{\sigma}$ = 1.646, n-2 = 39
$R^2$ = 98.7% variability explained

Georgia
Tech

# Does the Theory Hold?

The principle of purchasing power parity (PPP) states:

Average annual change in the exchange rate $=$ Difference in average annual inflation rates $+$ Random error

The economic theory says that $\beta_0 = 0,\ \beta_1 = 1$.
The estimates for these coefficients are: $\hat{\beta}_0 = -1.519, \hat{\beta}_1 = 0.9618$

**Violations of PPP theory with respect to both the intercept and the slope.**

**Testing the theory:**
$\beta_0 = 0$: Based on the t-test of statistical significance we find that $\beta_0$ is statistically different from zero.

$\beta_1 = 1$: We need to perform a t-test with this as the null hypothesis:
T-value = $\dfrac{\hat{\beta}_1 - 1}{se(\hat{\beta}_1)} = \dfrac{0.9618 - 1}{0.0178} = $ -2.1448
p-value = $2(1 - P(T_{39} < \text{-}2.1448)) = 0.038$

Georgia
Tech

# Hypothesis Testing in R

**# Perform the hypothesis test for slope coefficient H0: slope=1**

**# use the library 'car' available in R (you need to install this library first then download it)**

 *install.packages("car")*
 *library(car)*
 *linearHypothesis(pppa,c(0,1),rhs=1)*

**## Alternatively, you can compute the t-value and p-value as follows:**

 *tvalue = (0.9618−1)/0.01781*
 *pvalue = 2\*(1-pt(abs(tvalue),39))*

**## Use the help menu to learn more about the functions used above:**

 *help(pt)*
 *help(linearHypothesis)*

P-value = $2P(T_{n-2} > |t\text{-value}|)$
where
$$t-value = \frac{\hat{\beta}_1 - 1}{\sqrt{\hat{\sigma}^2 / S_{XX}}}$$

Georgia Tech

# Summary