V. Extensions of Multiple Regression

A. Dummy Variables

B. Scaling Variables

C. Non-Linear Models

1. **Basic Idea for Estimation**.
2. **Models that don't require logs**
3. **Models that require log transform**.

B. Scaling Variables

1. Scaling the Dependent Variable:

- The effects of scaling by a known constant are perfectly predictable:

  Suppose we multiply Y by 'a'

$$\hat{\beta}_1^* = \frac{\sum x_i (a \cdot Y_i)}{\sum x_i^2} = (a)\frac{\sum x_i Y_i}{\sum x_i^2} = \hat{\beta}_1 (a)$$

Scaling by a constant has no real consequence.

2. Scaling the Independent Variable:

- The effects of scaling by a known constant are again perfectly predictable:

  Suppose we multiply X by 'b'

$$\hat{\beta}_1^* = \frac{\sum (b \cdot x_i) Y_i}{\sum (b \cdot x_i)^2} = \left(\frac{1}{b}\right)\frac{\sum x_i Y_i}{\sum x_i^2} = \hat{\beta}_1 \left(\frac{1}{b}\right)$$

Scaling by a constant, no real consequence.

3. Scaling both Variables – combine the two results – again, no real consequences.

4. Indexes – the case of an unknown scale.

- The independent variable, and possibly the dependent variable, may be "indexes."
- Examples: Consumer Price Indexes.
- You have no idea about the "base year" value.
- Indexes – pure numbers, no "units."
- Solution – ***calculate elasticities***, pure numbers.
- A 1% change in X causes an ε% change in Y.

C. Non-Linear Models

1. ***Basic idea***: Make the model look linear in the parameters. Then estimate using OLS.
2. ***Models that don't require a log transform***.

Quadratic form:

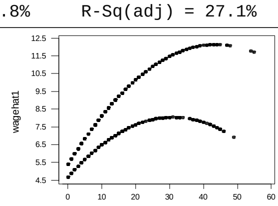
Examples:

Interpretation:

```
wage = - 5.49 + 0.905 ed + 0.309 ex - 0.00352 exsq
      - 0.726 fe - 0.0903 feex

Predictor        Coef       StDev          T         P
Constant       -5.494      1.2140      -4.53     0.000
ed             0.9046      0.0791      11.43     0.000
ex             0.3093      0.0554       5.59     0.000
exsq          -0.0035      0.0012      -2.98     0.003
fe            -0.7264      0.6725      -1.08     0.281
feex          -0.0903      0.0309      -2.93     0.004

S = 4.388       R-Sq = 27.8%     R-Sq(adj) = 27.1%
```



C. Non-Linear Models

Inverse form: $Y = \beta_0 + \beta_1 (1/X) + u$

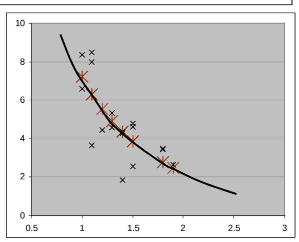
Example:

Interpretation:

```
The regression equation is
w-dot = - 1.04 + 7.91 INV(U)

Predictor        Coef     SE Coef          T         P
Constant       -1.044       2.373      -0.44     0.666
INV(U)          7.907       3.164       2.50     0.025

S = 1.734       R-Sq = 29.4%     R-Sq(adj) = 24.7%
```



C. Non-Linear Models

3. Models that require a log transform.

- ***Cobb-Douglas***: log-log models.

- Transform using natural logarithm.

- Interpretation:

Regression Analysis: lnsales versus lnprose, lnpcarn, lndinc

```
The regression equation is
lnsales = 6.29 - 1.86 lnprose + 1.45 lnpcarn + 0.560
lndinc

Predictor      Coef  SE Coef      T      P
Constant      6.288    4.875   1.29  0.221
lnprose     -1.8562   0.3438  -5.40  0.000
lnpcarn      1.4541   0.5724   2.54  0.026
lndinc       0.5596   0.9211   0.61  0.555

S = 0.175868   R-Sq = 73.7%   R-Sq(adj) = 67.2%

Analysis of Variance

Source          DF       SS       MS      F      P
Regression       3  1.04145  0.34715  11.22  0.001
Residual Error  12  0.37115  0.03093
Total           15  1.41260
```

3. Models that require a log transform.

- ***Semi-Log models***: log-linear

$$Y = \exp^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + u}$$

- Transform using natural logarithm.
- Parameters:

3. Models that require a log transform.

- ***Semi-Log models***: linear-log.

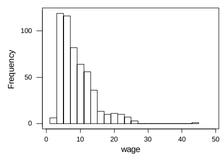
$$\exp^{Y} = A X_1^{\beta_1} X_2^{\beta_2} \exp^{u}$$

- Transform using natural logarithm.
- Parameters:

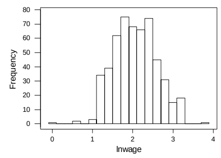
Normality Assumption

- What did we assume was normally distributed?
- What part of our analysis relies on this assumption?
- Some data are not normally distributed – wages.
- Log transformation – the natural log is normal.

Distribution of Wages



Distribution of Ln(Wage)




```
lnwage = - 0.921 + 0.988 lned + 0.235 lnex + 0.047 fe
        - 0.124 lnexfe

Predictor       Coef       StDev          T        P
Constant     -0.9212      0.2524      -3.65    0.000
lned         0.98774     0.08731      11.31    0.000
lnex         0.23505     0.03111       7.56    0.000
fe            0.0475      0.1256       0.38    0.706
lnexfe      -0.12383     0.04555      -2.72    0.007

S = 0.4471      R-Sq = 27.4%     R-Sq(adj) = 26.8%
```