# Multicollinearity



## The nature of Multicollinearity

Multiple regression

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + ... + \beta_k X_{ki} + u_i$$

$X_1 = 1$ for all observations to allow for the intercept term, an exact linear relationship is said to exist if the following condition is satisfied:

$$\lambda_1 X_{1i} + \lambda_2 X_{2i} + ... + \lambda_k X_{ki} = 0$$

where $\lambda_1, \lambda_2, ..., \lambda_k$ are constants such that not all of them are zero simultaneously



The case where the X variables are intercorrelated but not perfectly so, as follows:

$$\lambda_1 X_{1i} + \lambda_2 X_{2i} + ... + \lambda_k X_{ki} + \nu_i = 0$$

where $\nu_i$ is a stochastic error term



The difference between perfect and less than perfect multicollinearity

$$X_{2i} = \frac{\lambda_1}{\lambda_2} X_{1i} - \frac{\lambda_3}{\lambda_2} X_{3i} - ... - \frac{\lambda_k}{\lambda_2} X_{ki}$$

$$X_{2i} = \frac{\lambda_1}{\lambda_2} X_{1i} - \frac{\lambda_3}{\lambda_2} X_{3i} - ... - \frac{\lambda_k}{\lambda_2} X_{ki} - \frac{1}{\lambda_2} \nu_i$$

$if\ \lambda_2 \neq 0$



## Perfect Multicollinearity

| $X_2$ | $X_3$ |
|---|---|
| 10 | 50 |
| 15 | 75 |
| 18 | 90 |
| 24 | 120 |
| 30 | 150 |

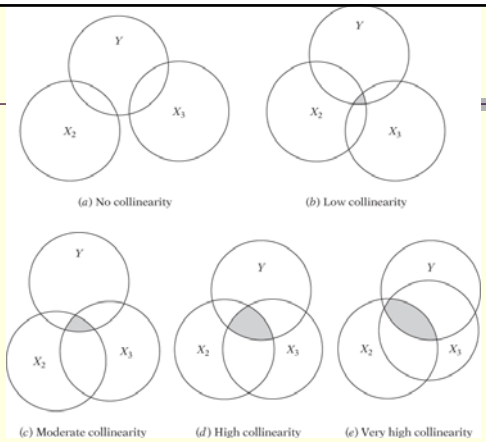
$$-5X_{2i} + X_{3i} = 0$$



## Not Perfect Multicollinearity

| $X_{2i}$ | $X_{3i}$ | $\nu_i$ |
|---|---|---|
| 10 | 52 | 2 |
| 15 | 75 | 0 |
| 18 | 97 | 7 |
| 24 | 129 | 9 |
| 30 | 152 | 2 |

$$-5X_{2i} + X_{3i} + \nu_i = 0$$




Multicollinearity refers only to linear relationships among the X variables. It does not rule out nonlinear relationships among them.

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^3 + u_i$$

This model is nonlinear, therefore, it does not violate the assumption of no multicollinearity



If multicollinearity is perfect, the regression coefficients of the X variables are indeterminate and their standard errors are infinite.

$$X_{2i} = \frac{\lambda_1}{\lambda_2} X_{1i} - \frac{\lambda_3}{\lambda_2} X_{3i} - \ldots - \frac{\lambda_k}{\lambda_2} X_{ki}$$



If multicollinearity is less than perfect, the regression coefficients, although determinate, possess large standard errors, which means the coefficients cannot be estimated with great precision or accuracy

$$X_{2i} = \frac{\lambda_1}{\lambda_2} X_{1i} - \frac{\lambda_3}{\lambda_2} X_{3i} - \ldots - \frac{\lambda_k}{\lambda_2} X_{ki} - \frac{1}{\lambda_2} \nu_i$$



## Several sources of multicollinearity

- The data collection method employed
- Constraints on the model or in the population being sampled
- Model specification
- An overdetermined model



- Estimation in the presence of perfect multicollinearity
- Estimation in the presence of “High” but “Imperfect” Multicollinearity

## Estimation in the presence of perfect multicollinearity

In the case of perfect multicollinearity the **regression coefficients remain indeterminate** and their **standard errors are infinite**



$$y_i = Y_i - \bar{Y} \quad x_{2i} = X_{2i} - \bar{X}_2 \quad x_{3i} = X_{3i} - \bar{X}_3$$

$$y_i = \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{u}_i$$

$$\hat{\beta}_2 = \frac{(\sum y_i x_{2i})(\sum x_{3i}^2) - (\sum y_i x_{3i})(\sum x_{2i} x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2}$$

$$\hat{\beta}_3 = \frac{(\sum y_i x_{3i})(\sum x_{2i}^2) - (\sum y_i x_{2i})(\sum x_{2i} x_{3i})}{(\sum x_{2i}^2)(\sum x_{3i}^2) - (\sum x_{2i} x_{3i})^2}$$



## Assume that

$$X_{3i} = \lambda X_{2i}$$
$$\bar{X}_{3i} = \lambda \bar{X}_{2i}$$
$$(X_{3i} - \bar{X}_{3i}) = \lambda (X_{2i} - \bar{X}_{2i})$$
$$x_{3i} = \lambda x_{2i}$$



## $\hat{\beta}_2$ and $\hat{\beta}_3$ are indeterminate

$$\hat{\beta}_2 = \frac{\lambda^2\{(\sum y_i x_{2i})(\sum x_{3i}^2) - (\sum y_i x_{2i})(\sum x_{2i}^2)\}}{\lambda^2\{(\sum x_{2i}^2)(\sum x_{2i}^2) - (\sum x_{2i}^2)^2\}} = \frac{0}{0}$$

$$\hat{\beta}_3 = \frac{\lambda\{(\sum y_i x_{2i})(\sum x_{2i}^2) - (\sum y_i x_{2i})(\sum x_{2i}^2)\}}{\lambda^2\{(\sum x_{2i}^2)(\sum x_{2i}^2) - (\sum x_{2i}^2)^2\}} = \frac{0}{0}$$



In the case of perfect multicollinearity one cannot get a unique solution for the individual regression coefficients

The variances and standard errors of $\hat{\beta}_2$ and $\hat{\beta}_3$ individually are infinite



## Estimation in the presence of "High" but "Imperfect" Multicollinearity

$$X_{3i} = \lambda X_{2i} + \nu_i$$
$$\bar{X}_{3i} = \lambda \bar{X}_{2i}$$
$$X_{3i} - \bar{X}_{3i} = \lambda (X_{2i} - \bar{X}_{2i}) + \nu_i$$
$$x_{3i} = \lambda x_{2i} + \nu_i,$$

where $\lambda \neq 0$

$$\sum x_i \nu_i = 0$$

$$\hat{\beta}_2 = \frac{\sum (y_i x_{2i})(\lambda^2 \sum x_{2i}^2 + \sum v_i^2) - (\lambda \sum y_i x_{2i} + \sum y_i v_i)(\lambda \sum x_{2i}^2)}{\sum x_{2i}^2 (\lambda^2 \sum x_{2i}^2 + \sum v_i^2) - (\lambda \sum x_{2i}^2)^2}$$



## Theoretical Consequences of Multicollinearity

- The OLS estimators are **unbiased**. But unbiasedness is a multisample or repeated sampling property

Keeping the values of the variables X fixed, if one obtains repeated samples and computes the OLS estimators for each of these samples, the average of the sample values will converge to the true population values of the estimators as the number of sample increases



- Collinearity does not destroy the property of minimum variance
- In the class of all linear unbiased estimators, the OLS estimators have minimum variance- they are **efficient**
- But this does not mean that the variance of an OLS estimator will necessarily be small



- Multicollinearity is essentially a sample (regression) phenomenon, even if the X variables are not linearly related in the population, they may be so related in the particular sample
- When we postulate the theoretical or population regression function (PRF), we believe that all the X variables included in the model have a separate or independent influence on the dependent variable Y.



$$Consumption_i = \beta_1 + \beta_2 Income_i + \beta_3 Wealth_i + u_i$$

Two variables may be highly , if not perfectly, correlated: Wealthier people generally tend to have higher incomes.

To assess the individual effects of wealth and income on consumption expenditure we need a sufficient number of sample observations of wealthy individuals with low income, and high income individuals with low wealth



- OLS estimators are **BLUE** despite multicollinearity

## Practical Consequences of Multicollinearity



## Practical Consequences of Multicollinearity

1. OLS estimators have large variance and covariance
2. The confidence intervals tend to be much wider, leading to the acceptance of the "zero null hypothesis"
3. t ratio of one or more coefficients tends to be statistically insignificant
4. Although the t ratio of one or more coefficients is statistically insignificant, R-Squared can be very high
5. The OLS estimators and their standard errors can be sensitive to small changes in the data



## OLS estimators have large variance and covariance

$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2i}^2 (1 - r_{23}^2)}$$

$$\text{var}(\hat{\beta}_3) = \frac{\sigma^2}{\sum x_{3i}^2 (1 - r_{23}^2)}$$

$$\text{cov}(\hat{\beta}_2, \hat{\beta}_3) = \frac{-r_{23}\sigma^2}{(1 - r_{23}^2)\sqrt{\sum x_{2i}^2 \sum x_{3i}^2}}$$

$r_{23}$ is the coefficient of correlation between $X_2$ and $X_3$



## variance-inflating factor (VIF)

VIF shows how the variance of an estimator is inflated by the presence of multicollinearity

$$VIF = \frac{1}{(1 - r_{23}^2)}$$

$$r_{23}^2 \to 1, VIF \to \infty$$



$$\text{var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_{2i}^2} VIF$$

$$\text{var}(\hat{\beta}_3) = \frac{\sigma^2}{\sum x_{3i}^2} VIF$$



## Example

TABLE 10.1 The Effect of Increasing $r_{23}$ on var $(\hat{\beta}_2)$ and cov $(\hat{\beta}_2, \hat{\beta}_3)$

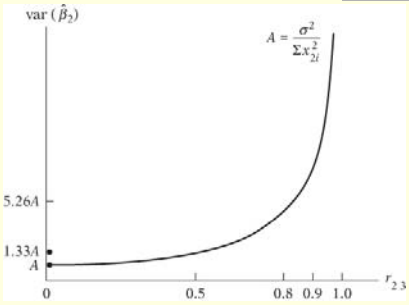
| Value of $r_{23}$ (1) | VIF (2) | var $(\hat{\beta}_2)$ (3)* | $\frac{\text{var}(\hat{\beta}_2)(r_{23} \neq 0)}{\text{var}(\hat{\beta}_2)(r_{23} = 0)}$ (4) | cov $(\hat{\beta}_2, \hat{\beta}_3)$ (5) |
|---|---|---|---|---|
| 0.00 | 1.00 | $\frac{\sigma^2}{\sum x_{2i}^2} = A$ | — | 0 |
| 0.50 | 1.33 | $1.33 \times A$ | 1.33 | $0.67 \times B$ |
| 0.70 | 1.96 | $1.96 \times A$ | 1.96 | $1.37 \times B$ |
| 0.80 | 2.78 | $2.78 \times A$ | 2.78 | $2.22 \times B$ |
| 0.90 | 5.76 | $5.26 \times A$ | 5.26 | $4.73 \times B$ |
| 0.95 | 10.26 | $10.26 \times A$ | 10.26 | $9.74 \times B$ |
| 0.97 | 16.92 | $16.92 \times A$ | 16.92 | $16.41 \times B$ |
| 0.99 | 50.25 | $50.25 \times A$ | 50.25 | $49.75 \times B$ |
| 0.995 | 100.00 | $100.00 \times A$ | 100.00 | $99.50 \times B$ |
| 0.999 | 500.00 | $500.00 \times A$ | 500.00 | $499.50 \times B$ |

Note: $A = \frac{\sigma^2}{\sum x_{2i}^2}$

$B = \frac{-\sigma^2}{\sqrt{\sum x_{2i}^2 \sum x_{3i}^2}}$

$\times$ = times

*To find out the effect of increasing $r_{23}$ on var $(\hat{\beta}_2)$, note that $A = \sigma^2 / \sum x_{2i}^2$ when $r_{23} = 0$, but the variance and covariance magnifying factors remain the same.






$$\text{var}(\hat{\beta}_j) = \frac{\sigma^2}{\sum x_j^2} VIF$$

var($\hat{\beta}_j$) is large or small will depend on the three ingredients: (1) $\sigma^2$

(2) *VIF*

(3) $\sum x_j^2$



**TABLE 10.2**
**The Effect of Increasing Collinearity on the 95% Confidence Interval for $\beta_2$: $\hat{\beta}_2 \pm 1.96\,\text{se}(\hat{\beta}_2)$**

| Value of $r_{23}$ | 95% Confidence Interval for $\beta_2$ |
|---|---|
| 0.00 | $\hat{\beta}_2 \pm 1.96\sqrt{\frac{\sigma^2}{\sum x_{2i}^2}}$ |
| 0.50 | $\hat{\beta}_2 \pm 1.96\sqrt{(1.33)}\sqrt{\frac{\sigma^2}{\sum x_{2i}^2}}$ |
| 0.95 | $\hat{\beta}_2 \pm 1.96\sqrt{(10.26)}\sqrt{\frac{\sigma^2}{\sum x_{2i}^2}}$ |
| 0.995 | $\hat{\beta}_2 \pm 1.96\sqrt{(100)}\sqrt{\frac{\sigma^2}{\sum x_{2i}^2}}$ |
| 0.999 | $\hat{\beta}_2 \pm 1.96\sqrt{(500)}\sqrt{\frac{\sigma^2}{\sum x_{2i}^2}}$ |

*Note:* We are using the normal distribution because $\sigma^2$ is assumed for convenience to be known. Hence the use of 1.96, the 95% confidence factor for the normal distribution.
The standard errors corresponding to the various $r_{23}$ values are obtained from Table 10.1.



## Wider Confidence Intervals

Because of the large standard errors, the confidence intervals for the relevant population parameters tend to be larger



## "Insignificant" t Ratios

In cases of high collinearity the estimated standard errors increase dramatically, thereby making the t values smaller



## A High $R^2$ but Few Significant t Ratios

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} + \mu_i$$

In cases of high collinearity, it is possible to find, the partial slope coefficients are individually statistically insignificant on the basis of the t test

## Sensitivity of OLS Estimators and Their Standard Errors to Small Changes in Data

As long as multicollinearity is not perfect, estimation of the regression coefficients is possible but the estimates and their standard errors become very sensitive to even the slightest change in the data



## Example

**TABLE 10.3** Hypothetical Data on $Y$, $X_2$, and $X_3$

| $Y$ | $X_2$ | $X_3$ |
|---|---|---|
| 1 | 2 | 4 |
| 2 | 0 | 2 |
| 3 | 4 | 12 |
| 4 | 6 | 0 |
| 5 | 8 | 16 |

$$\hat{Y}_i = 1.1939 + 0.4463X_{2i} + 0.0030X_{3i}$$
$$(0.7737)\ (0.1848)\quad (0.0851)$$
$$t = (1.5431)\ (2.4151)\quad (0.0358)$$
$$R^2 = 0.8101$$
$$r_{23} = 0.5523$$
$$\text{cov}(\hat{\beta}_2, \hat{\beta}_3) = -0.00868$$
$$df = 5 - 3 = 2$$



```
      Source |       SS       df       MS              Number of obs =       5
-------------+------------------------------           F(  2,     2) =    4.27
       Model |  8.10121951     2  4.05060976           Prob > F      =  0.1899
    Residual |  1.89878049     2  .949390244           R-squared     =  0.8101
-------------+------------------------------           Adj R-squared =  0.6202
       Total |        10     4         2.5           Root MSE      =  .97437

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |   .4463415   .1848104     2.42   0.137    -.3488336    1.241517
          x3 |   .0030488   .0850659     0.04   0.975    -.3629602    .3690578
       _cons |   1.193902   .7736789     1.54   0.263    -2.134969    4.522774
------------------------------------------------------------------------------
```



**TABLE 10.4** Hypothetical Data on $Y$, $X_2$, and $X_3$

| $Y$ | $X_2$ | $X_3$ |
|---|---|---|
| 1 | 2 | 4 |
| 2 | 0 | 2 |
| 3 | 4 | 0 |
| 4 | 6 | 12 |
| 5 | 8 | 16 |

$$\hat{Y}_i = 1.2108 + 0.4014X_{2i} + 0.0270X_{3i}$$
$$(0.7480)\ (0.2721)\quad (0.1252)$$
$$t = (1.6187)\ (1.4752)\quad (0.2158)$$
$$R^2 = 0.8143$$
$$r_{23} = 0.8285$$
$$\text{cov}(\hat{\beta}_2, \hat{\beta}_3) = -0.0282$$
$$df = 5 - 3 = 2$$



```
      Source |       SS       df       MS              Number of obs =       5
-------------+------------------------------           F(  2,     2) =    4.39
       Model |  8.14324324     2  4.07162162           Prob > F      =  0.1857
    Residual |  1.85675676     2  .928378378           R-squared     =  0.8143
-------------+------------------------------           Adj R-squared =  0.6286
       Total |        10     4         2.5           Root MSE      =  .96352

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x2 |   .4013514   .272965     1.48   0.278    -.7692498    1.571953
          x3 |   .027027    .1252281    0.22   0.849    -.5117858    .5658399
       _cons |   1.210811   .7480215    1.62   0.247    -2.007666    4.429288
------------------------------------------------------------------------------
```



## Detection of Multicollinearity

Kmenta (1986)

1. Multicollinearity is a question of degree and not of kind. The meaning distinction is not between the presence and the absence of multicollinearity, but between its various degrees
2. Since multicollinearity refers to the condition of the explanatory variables that are assumed to be nonstochastic, it is a feature of the sample an not of the population

Therefore, we do not “test for multicollinearity” but can, if we wish, measure its degree in any particular sample



## Detection of Multicollinearity

- High R-Squared but few significant t-ratios
- High pair-wise correlations among regressors
- Examination of partial correlations
- Auxiliary regressions
- VIF
- Scatter plot



## High R-Squared but few significant t-ratios

- If R-Squared is high, say, in excess of 0.8, the F-test in most cases will reject the hypothesis that the partial slope coefficients are simultaneously equal to zero, but the individual t tests will show that none or very few of the partial slope coefficients are statistically different from zero.



## Example

| Source | SS | df | MS |
|---|---|---|---|
| Model | 8565.55487 | 2 | 4282.77704 |
| Residual | 324.445926 | 7 | 46.3493418 |
| Total | 8890 | 9 | 987.777778 |

Number of obs = 10
F( 2, 7) = 92.40
Prob > F = 0.0000
R-squared = 0.9635
Adj R-squared = 0.9531
Root MSE = 6.808

| y | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| x2 | .9415373 | .8228983 | 1.14 | 0.290 | -1.004308 | 2.887383 |
| x3 | -.0424345 | .0806645 | -0.53 | 0.615 | -.2331757 | .1483067 |
| _cons | 24.77473 | 6.7525 | 3.67 | 0.008 | 8.807609 | 40.74186 |



## High pair-wise correlations among regressors

- The pair-wise or zero-order correlation coefficient between two regressors is high, say, in excess of 0.8



## Example

Correlation between income (x2) and wealth (x3)

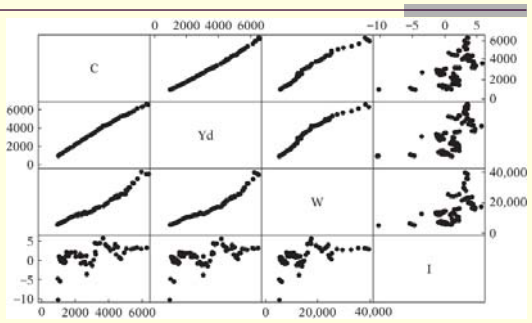
| | x2 | x3 |
|---|---|---|
| x2 | 1.0000 | |
| x3 | 0.9990 | 1.0000 |

## Variance inflation factor

- **As a rule of thumb**, if the VIF of a variable exceeds 10, which will happen if $R_j^2$ exceeds 0.90, that variable is said be highly collinear



## Scatter plot






*C – Consumption*

*$Y_d$ – Real disposable personal income*

*W – Real wealth*

*I – Real Interest Rate*



## Remedial Measures



## Remedial Measures

- Do Nothing
- Rule-Thumb Procedures

1. A priori information
2. Combining cross-sectional and time series data
3. Dropping a variable (s) and specification bias
4. Adding or new data
5. Transformation of variables



## A priori information

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

$$Y = consumption, X_2 = income, X_3 = wealth$$

$$\beta_3 = 0.10\beta_2$$

$$Y_i = \beta_1 + \beta_2 X_{2i} + 0.10\beta_2 X_{3i} + u_i$$

$$= \beta_1 + \beta_2 X_i + u_i$$

$$X_i = X_{2i} + 0.10 X_{3i}$$

## Combining cross-sectional and time series data

A variant of the extraneous or a priori information technique is the combination of cross-sectional and time series data known as pooling the data



## Dropping a variable (s) and specification bias

But in dropping a variable from the model we may be committing a **specification bias** or **specification error**.



## Adding or new data

As the sample size increases, $\sum_i (X_{ji} - \bar{X}_j)^2$ will generally increases. Therefore, for any given $r_{23}$ , the variance of $\hat{\beta}_2$ will decrease, thus decreasing the standard error, which will enable us to estimate $\beta_2$ more precisely



## Transformation of variables

- First difference form
- Ratio transformation



- First difference form

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$$
$$Y_{t-1} = \beta_1 + \beta_2 X_{2,t-1} + \beta_3 X_{3,t-1} + u_{t-1}$$
$$Y_t - Y_{t-1} = \beta_2 (X_{2t} - X_{2,t-1}) + \beta_3 (X_{3t} - X_{3,t-1}) + v_t$$

*where* $v_t = u_t - u_{t-1}$



First difference form may not satisfy one of the assumptions of the CLRM – the disturbances are serially uncorrelated (We will see in Autocorrelation chapter)

First differencing – may not appropriate in cross-sectional data where there no logical ordering of the observations

- Ratio transformation

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t$$

$$\frac{Y_t}{X_{3t}} = \beta_1\left(\frac{1}{X_{3t}}\right) + \beta_2\left(\frac{X_{2t}}{X_{3t}}\right) + \left(\frac{u_t}{X_{3t}}\right)$$



- Ratio transformation, the error term $\left(\frac{u_t}{X_{3t}}\right)$ will be **heteroscedastic**, if the original error term is homoscedastic



Multicollinearity may not pose a serious problem

- When R-squared is high and the regression coefficients are individually significant as revealed by the higher t values



# Example

Multicollinearity



## Consumption Expenditure in Relation to Income and Wealth

**TABLE 10.5**
**Hypothetical Data on Consumption Expenditure $Y$, Income $X_2$, and Wealth $X_3$**

| $Y$, $ | $X_2$, $ | $X_3$, $ |
|---|---|---|
| 70 | 80 | 810 |
| 65 | 100 | 1009 |
| 90 | 120 | 1273 |
| 95 | 140 | 1425 |
| 110 | 160 | 1633 |
| 115 | 180 | 1876 |
| 120 | 200 | 2052 |
| 140 | 220 | 2201 |
| 155 | 240 | 2435 |
| 150 | 260 | 2686 |



| Source | SS | df | MS | Number of obs = | 10 |
|---|---|---|---|---|---|
| | | | | F( 2, 7) = | 92.40 |
| Model | 8565.55407 | 2 | 4282.77704 | Prob > F = | 0.0000 |
| Residual | 324.445926 | 7 | 46.349418 | R-squared = | 0.9635 |
| | | | | Adj R-squared = | 0.9531 |
| Total | 8890 | 9 | 987.777778 | Root MSE = | 6.808 |

| y | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| x2 | .9415373 | .8228983 | 1.14 | 0.290 | -1.004308 | 2.887383 |
| x3 | -.0424345 | .0806645 | -0.53 | 0.615 | -.2331757 | .1483067 |
| _cons | 24.77473 | 6.7525 | 3.67 | 0.008 | 8.807609 | 40.74186 |

$$\hat{Y}_i = 24.7747 + 0.9415X_{2i} - 0.0424X_{3i}$$

$$(6.7525) \quad (0.8229) \quad (0.0807)$$

$$t = (3.6690) \quad (1.1442) \quad (-0.5261)$$

$$R^2 = 0.9635 \quad \bar{R}^2 = 0.9531 \quad df = 10 - 3 = 7$$



Regression shows that income and wealth together explain about 96 % of the variation in consumption expenditure, and yet **neither of the slope coefficients is individually statistically significant**. Moreover, not only is the wealth variable statistically insignificant but also it has the wrong sign



TABLE 10.6
ANOVA Table for the Consumption–Income–Wealth Example

| Source of Variation | SS | df | MSS |
|---|---|---|---|
| Due to regression | 8,565.5541 | 2 | 4,282.7770 |
| Due to residual | 324.4459 | 7 | 46.3494 |



$$H_0 = \beta_2 = \beta_3 = 0$$

$$F = \frac{4282.7770}{46.3494} = 92.4019$$

Reject the null hypothesis
(92.4019 > Critical F-value)



This example shows dramatically what multicollinearity does. The fact that the F test is significant but the t values of $X_2$ and $X_3$ are individually insignificant means that the two variables are so highly correlated that it is impossible to isolate the individual impact of either income and wealth on consumption



$$\hat{X}_{3i} = 7.5454 + 10.1909X_{2i}$$

| Source | SS | df | MS |
|---|---|---|---|
| Model | 3427202.73 | 1 | 3427202.73 |
| Residual | 7123.27273 | 8 | 890.409091 |
| Total | 3434326 | 9 | 381591.778 |

Number of obs = 10
F( 1, 8) = 3849.02
Prob > F = 0.0000
R-squared = 0.9979
Adj R-squared = 0.9977
Root MSE = 29.84

| x3 | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| x2 | 10.19091 | .1642623 | 62.04 | 0.000 | 9.81212 | 10.5697 |
| _cons | 7.545455 | 29.47581 | 0.26 | 0.804 | -60.42589 | 75.5168 |

$$\hat{Y}_i = 24.4545 + 0.5091X_{2i}$$

| Source | SS | df | MS | Number of obs = | 10 |
|---|---|---|---|---|---|
| Model | 8552.72727 | 1 | 8552.72727 | F( 1, 8) = | 202.87 |
| Residual | 337.272727 | 8 | 42.1590909 | Prob > F = | 0.0000 |
| | | | | R-squared = | 0.9621 |
| Total | 8890 | 9 | 987.777778 | Adj R-squared = | 0.9573 |
| | | | | Root MSE = | 6.493 |

| y | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| x2 | .5090909 | .0357428 | 14.24 | 0.000 | .4266678 | .591514 |
| _cons | 24.45455 | 6.413817 | 3.81 | 0.005 | 9.664256 | 39.24483 |



$$\hat{Y}_i = 24.411 + 0.0498X_{3i}$$

| Source | SS | df | MS | Number of obs = | 10 |
|---|---|---|---|---|---|
| Model | 8504.87666 | 1 | 8504.87666 | F( 1, 8) = | 176.67 |
| Residual | 385.123344 | 8 | 48.1404181 | Prob > F = | 0.0000 |
| | | | | R-squared = | 0.9567 |
| Total | 8890 | 9 | 987.777778 | Adj R-squared = | 0.9513 |
| | | | | Root MSE = | 6.9383 |

| y | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| x3 | .0497638 | .003744 | 13.29 | 0.000 | .0411301 | .0583974 |
| _cons | 24.41104 | 6.874097 | 3.55 | 0.007 | 8.559349 | 40.26274 |



Regressions show very clearly that in situations of extreme multicollinearity dropping the highly collinear variable will often make the other X variable statistically significant.

This result would suggest that a way out of extreme collinearity is to drop the collinearity variable.