

ACE 564
Spring 2006

Lecture 8

*Violations of Basic Assumptions I: Multicollinearity
and Non-Sample Information*

by
Professor Scott H. Irwin

Readings:

Griffiths, Hill and Judge. "Collinear Economic Variables," Chapter 13 and "Combining Sample and Nonsample Information and Further Applications of the General Linear Statistical Model." Chapter 11 in *Learning and Practicing Econometrics*

Kennedy. "Violating Assumption Five: Multicollinearity," Chapter 11; "Incorporating Extraneous Information," Chapter 12 in *A Guide to Econometrics*

Violations of Regression Model Assumptions

To begin, let's review the assumptions multiple regression model

$$\text{MR1. } y_t = \beta_1 + \beta_2 x_{2,t} + \beta_3 x_{3,t} + e_t, \quad t = 1, \dots, T$$

$$\text{MR2. } E(e_t) = 0 \Leftrightarrow E(y_t) = \beta_1 + \beta_2 x_{2,t} + \beta_3 x_{3,t}$$

$$\text{MR3. } \text{var}(e_t) = \text{var}(y_t) = \sigma^2$$

$$\text{MR4. } \text{cov}[e_t, e_s] = \text{cov}[y_t, y_s] = 0 \quad t \neq s$$

MR5. The values of $x_{2,t}$ and $x_{3,t}$ are not random or exact linear functions of one another

$$\text{MR6. } e_t \sim N(0, \sigma^2) \Leftrightarrow y_t \sim N(\beta_1 + \beta_2 x_{2,t} + \beta_3 x_{3,t} + e_t, \sigma^2)$$

The LS estimator for a regression model is valid so long as all of the “classical” assumptions hold

There are many situations faced by an applied researcher where the “classical” assumptions do not hold and a different estimator is required

We can think of an econometrics textbook as a “catalog” that tells a researcher what to “buy” in different situations

- If a situation differs from the classical assumptions, a researcher can turn to the catalog to see if LS can be used, and if not, what alternative estimator should be used
- Kennedy, p.42-43
- Why you should always have one or two good econometric textbooks on your shelf if you will be using regression on the job!

Table 3.1 The assumptions of the CLR model

Assumption	Mathematical expression			Chapter in which discussed
	Bivariate	Multivariate	Violations	
(1) Dependent variable a linear function of a specific set of independent variables, plus a disturbance	$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t, t = 1, \dots, T$	$Y = X\beta + \varepsilon$	Wrong regressors Nonlinearity Changing parameters	6
(2) Expected value of disturbance term is zero	$E\varepsilon_t = 0, \text{ for all } t$	$E\varepsilon = 0$	Biased intercept	7
(3) Disturbances have uniform variance and are uncorrelated	$E\varepsilon_t \varepsilon_r = 0, t \neq r$ $= \sigma^2, t = r$	$E\varepsilon\varepsilon' = \sigma^2 I$	Heteroskedasticity Autocorrelated errors	8
(4) Observations on independent variables can be considered fixed in repeated samples	x_t fixed in repeated samples	X fixed in repeated samples	Errors in variables Autoregression Simultaneous equations	9 10
(5) No exact linear relationships between independent variables and more observations than independent variables	$\sum_{i=1}^r (x_i - \bar{x})^2 \neq 0$	Rank of $X = K \leq T$	Perfect multicollinearity	11

The mathematical terminology is explained in the technical notes to this section. The notation is as follows: Y is a vector of observations on the dependent variable; X is a matrix of observations on the independent variables; ε is a vector of disturbances; σ^2 is the variance of the disturbances; I is the identity matrix; K is the number of independent variables; T is the number of observations.

Kennedy, P. *Guide to Econometrics*, Fourth Edition. The MIT Press, Cambridge, Mass. 1998.

We will examine four of the most common violations of the assumptions of the multiple regression model

- **Multicollinearity**
- **Heteroskedasticity**
- **Autocorrelation**
- **Specification Errors**

As the previous chart shows, other violations also are important, but due to time limitations we will not have time to cover them in this class

A highly valuable and readable presentation of econometric issues can be found in the “Symposium on Econometric Tools” found in the Fall 2001 issue of the Journal of Economic Perspectives

The Problem

Charged with designing and estimating a model to explain the annual demand for beer in the US

Want to know,

- Sensitivity of beer demand to the price of beer
- Sensitivity of beer demand to the price of competing beverages
- Sensitivity of beer demand to general economic conditions

Economic Model

Economic theory of consumer behavior suggests that demand for a good depends on the price of that good, on the prices of substitute and complement goods and income

For beer it is reasonable to assume that the quantity demanded for beer (y) depends on the price of beer (x_1), the price of other liquor (x_2), the price of all other remaining goods and services (x_3) and income (x_4)

Mathematically, we can express this relationship as,

$$y = f(x_2, x_3, x_4, x_5)$$

Let's assume a double-log functional form for the demand relationship,

$$\ln y = \beta_1 + \beta_2 \ln x_2 + \beta_3 \ln x_3 + \beta_4 \ln x_4 + \beta_5 \ln x_5$$

The double-log functional form is convenient in this case because,

- It precludes infeasible negative prices and quantities
- Slope coefficients are elasticities

We can note that,

- $\beta_2 =$ direct price elasticity of demand for beer (-)
- $\beta_3 =$ cross-price elasticity of demand for beer with respect to a change in the price of other liquor (+)

- β_4 = cross-price elasticity of demand for beer with respect to a change in the price of remaining goods and services (+)
- β_5 = income elasticity of demand for beer (- or +)

The Statistical Model

The double-log economic model predicts that the log of beer consumption for given levels of the independent variables will be the same for all years

Recognize that the log of actual beer consumption for given levels of the independent variables will not be the same for all years

$$\ln y_t = \beta_1 + \beta_2 \ln x_{2,t} + \beta_3 \ln x_{3,t} + \beta_4 \ln x_{4,t} + \beta_5 \ln x_{5,t} + e_t$$

Note that:

$$E(\ln y_t) = \beta_1 + \beta_2 \ln x_{2,t} + \beta_3 \ln x_{3,t} + \beta_4 \ln x_{4,t} + \beta_5 \ln x_{5,t}$$

forms a hyper-plane and the data points scatter in five dimensions around this hyper-plane!

Table 11.1 Price, Quantity and Income Data for Beer Demand Model

q (liters)	p_B (\$)	p_L (\$)	p_R (\$)	m (\$)
81.7	1.78	6.95	1.11	25088
56.9	2.27	7.32	0.67	26561
64.1	2.21	6.96	0.83	25510
65.4	2.15	7.18	0.75	27158
64.1	2.26	7.46	1.06	27162
58.1	2.49	7.47	1.10	27583
61.7	2.52	7.88	1.09	28235
65.3	2.46	7.88	1.18	29413
57.8	2.54	7.97	0.88	28713
63.5	2.72	7.96	1.30	30000
65.9	2.60	8.09	1.17	30533
48.3	2.87	8.24	0.94	30373
55.6	3.00	7.96	0.91	31107
47.9	3.23	8.34	1.10	31126
57.0	3.11	8.10	1.50	32506
51.6	3.11	8.43	1.17	32408
54.2	3.09	8.72	1.18	33423
51.7	3.34	8.87	1.37	33904
55.9	3.31	8.82	1.52	34528
52.1	3.42	8.59	1.15	36019
52.5	3.61	8.83	1.39	34807
44.3	3.55	8.86	1.60	35943
57.7	3.72	8.97	1.73	37323
51.6	3.72	9.13	1.35	36682
53.8	3.70	8.98	1.37	38054
50.0	3.81	9.25	1.41	36707
46.3	3.86	9.33	1.62	38411
46.8	3.99	9.47	1.69	38823
51.7	3.89	9.49	1.71	38361
49.9	4.07	9.52	1.69	41593

Griffiths, William E., R. Carter Hill, George G. Judge. Learning and Practicing Econometrics. John Wiley & Sons, Inc. New York. 1993.

Data and Estimation

Sample information consists of thirty years of annual data

- Beer consumption is stated on a per capita basis
- Price series are constructed as weighted averages of the nominal prices for different kinds of beer, other liquor, and remaining goods and services

The estimation results are,

$$\ln \hat{y}_t = -3.243 - 1.020 \ln x_{2,t} - 0.583 \ln x_{3,t} + 0.210 \ln x_{4,t} + 0.923 \ln x_{5,t}$$

(3.743) (0.239) (0.560) (0.080) (0.416) (s.e.)

$$R^2 = 0.825$$

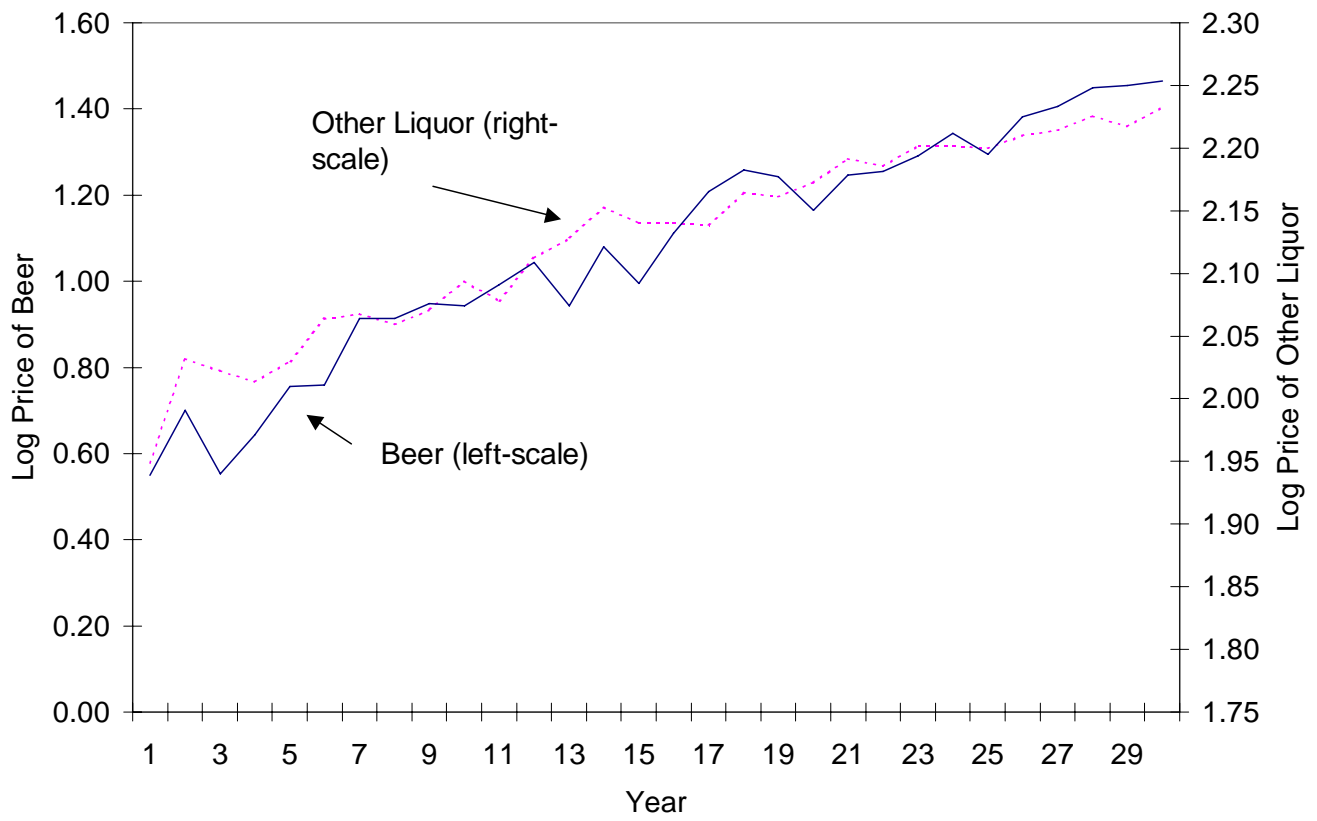
95% CI estimates,

$$\beta_2 : [-1.512, -0.528]$$

$$\beta_3 : [-1.737, 0.571]$$

$$\beta_4 : [0.045, 0.375]$$

$$\beta_5 : [0.066, 1.780]$$



Nature of Multicollinearity

To a large extent, observed economic data are non-experimental or passively generated

When economic data are the result of uncontrolled experiments carried about by nature and society, many of the economic variables move together in systematic ways

- Termed the problem of collinearity or multicollinearity
- In this environment, there is no guarantee that data will be “rich in information”
- It is difficult to precisely isolate and estimate economic relationships

Two types of multicollinearity

- Exact linear dependence among x 's, or perfect multicollinearity
- Inexact linear dependence among x 's, or less than perfect multicollinearity

Case 1: Exact linear dependence

Formally, for the $K-1$ independent variable multiple regression model, an exact linear relation exists if

$$\lambda_2 x_2 + \lambda_3 x_3 + \dots + \lambda_K x_K = 0$$

where $\lambda_2, \lambda_3, \dots, \lambda_K$ are constants and at least one is non-zero

Example for $K=5$:

$$\text{Set } \lambda_2 = 1, \lambda_3 = -1, \lambda_4 = -2, \lambda_5 = 0$$

Substituting,

$$1x_2 - 1x_3 - 2x_4 + 0x_5 = 0$$

or,

$$x_2 = x_3 + 2x_4$$

To see the severe implications of perfect multicollinearity, it is helpful to refer back the formula for the sampling variance of b_k in the two-independent variable case,

$$\text{var}(b_k) = \frac{\sigma^2}{\sum_{t=1}^T x_{k,t}'^2 (1 - r_{23}^2)}$$

where r_{23} is the simple correlation coefficient between the two independent variables

In the case of perfect multicollinearity, r_{23} will equal one, and,

- Least squares estimates of standard errors of slope parameters are infinite
- Least squares estimates of slope parameters are indeterminate

⇒ Least squares estimation is impossible

In other words, least squares estimation is not possible because the estimation procedure cannot identify the separate influences of the x 's on y because you really only have one x

- Fortunately, this situation is rarely encountered in practice
- One example is the “dummy variable trap”

Case 2: Inexact linear dependence

Formally, for the $K-1$ independent variable multiple regression model, an inexact linear relation exists if

$$\lambda_2 x_2 + \lambda_3 x_3 + \dots + \lambda_K x_K + v_t = 0$$

where $\lambda_2, \lambda_3, \dots, \lambda_K$ are constants and at least one is non-zero and v_t is a random error term

Example for $K=5$:

$$\text{Set } \lambda_2 = 1, \lambda_3 = -1, \lambda_4 = -2, \lambda_5 = 0$$

Substituting,

$$1x_2 - 1x_3 - 2x_4 + 0x_5 + v_t = 0$$

or,

$$x_2 = x_3 + 2x_4 + v_t$$

While there is no longer a perfect relationship among the x 's, the R^2 of a regression of x_2 on x_3 and x_4 would be positive

Hence, we have the situation of LS estimation in the presence of “high” but “imperfect” correlation

As noted earlier in this section, this type of “imperfect” multicollinearity is quite common in economic time series

- It is not surprising that many economic variables move together, as they all “share” a common trend related to the economy as a whole

More specific causes may also be found

- x 's are lagged values of a particular variable
- Similar economic relationship to y
- Example: print advertising versus television advertising for a food product

Consequences of “Inexact” Multicollinearity

Classical good news, bad news story

Good news:

BLUE property of LS estimators is preserved

- Parameter estimates are unbiased
- Variance of parameter estimates is the minimum possible for a linear estimator

Bad news:

1. Standard errors may be quite large

This can be seen by referring again to the formula for the sampling variance of b_k in the two-independent variable case,

$$\text{var}(b_k) = \frac{\sigma^2}{\sum_{t=1}^T x_{k,t}'^2 (1 - r_{23}^2)}$$

As the correlation, r_{23} , increases the sampling variance gets larger

- Illustrates the difficulty least squares has in precisely estimating parameters
- Least squares has difficulty identifying which x to “credit” with relationship to y

2. Confidence intervals will tend to be much wider

3. Low t -ratios, but high R^2 s and regression F -statistics possible

4. LS estimators and standard errors may be quite sensitive to small changes in the data

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

Value of r_{23} (1)	VIF (2)	$\text{var}(\hat{\beta}_2)$ (3)*	$\text{var}(\hat{\beta}_2) (r_{23} \neq 0)$	$\text{cov}(\hat{\beta}_2, \hat{\beta}_3)$
			$\text{var}(\hat{\beta}_2) (r_{23} = 0)$ (4)	(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

*To find out the effect of increasing r_{23} on $\text{var}(\hat{\beta}_3)$, note that $A = \sigma^2 / \sum x_{3i}^2$ when $r_{23} = 0$, but the variance and covariance magnifying factors remain the same.

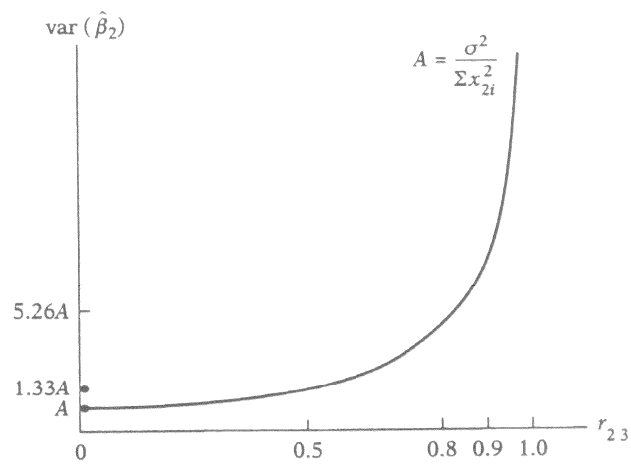


FIGURE 10.2
The behavior of $\text{var}(\hat{\beta}_2)$ as a function of r_{23} .

Gujarati, D.N. Basic Econometrics, Third Edition. McGraw-Hill, New York, NY, 1995.

TABLE 10.2
The effect of increasing collinearity on
the 95 percent confidence interval for
 $\beta_2 : \hat{\beta}_2 \pm 1.96 \text{ se}(\hat{\beta}_2)$

Value of r_{23}	95% confidence interval for β_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	$\hat{\beta}_2 \pm 1.96 \sqrt{(10.26)} \sqrt{\frac{\sigma^2}{\sum x_{2i}^2}}$
0.99	$\hat{\beta}_2 \pm 1.96 \sqrt{(100)} \sqrt{\frac{\sigma^2}{\sum x_{2i}^2}}$
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 σ^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.

Gujarati, D.N. Basic Econometrics, Third Edition. McGraw-Hill, New York, NY, 1995.

Methods for Detecting Multicollinearity

1. High R^2 but few significant t -ratios on individual coefficients
2. Sample correlation coefficients between pairs of explanatory variables
 - Rule-of-thumb: multicollinearity is a serious problem if correlation coefficient exceeds about 0.8 (in absolute value)
 - Limitation: May not detect more complex patterns of correlation among sets of three or more independent variables
3. Auxiliary regressions
 - Regress each independent variable on the remaining independent variables under consideration and test for “significance” of each regression using an F -test
 - We call such regressions “auxiliary regressions” since they have the form of a regression model but are only descriptive devices

- Klein's rule-of-thumb: multicollinearity is not a problem unless the highest R^2 from an auxiliary regression is higher than the R^2 of the original, estimated regression model

In practice, it is a good idea to examine each of the above methods.

Detecting Multicollinearity in the Beer Demand Regression

Pairwise Correlation Matrix:

	$x_{2,t}$	$x_{3,t}$	$x_{4,t}$	$x_{5,t}$
$x_{2,t}$	1.00			
$x_{3,t}$	0.97	1.00		
$x_{4,t}$	0.77	0.81	1.00	
$x_{5,t}$	0.97	0.97	0.82	1

Auxiliary Regressions:

$$x_{2,t} = f(x_{3,t}, x_{4,t}, x_{5,t}) \quad R^2=0.95$$

$$x_{3,t} = f(x_{2,t}, x_{4,t}, x_{5,t}) \quad R^2=0.95$$

$$x_{4,t} = f(x_{2,t}, x_{3,t}, x_{5,t}) \quad R^2=0.69$$

$$x_{5,t} = f(x_{2,t}, x_{3,t}, x_{4,t}) \quad R^2=0.96$$

Remedies for Multicollinearity

It is worth repeating that correlation among the independent variables does not destroy the fundamental BLUE property of least squares estimators

In this sense, there is no “problem” to correct

While this is true, it may make our sample estimates so imprecise that they are of little practical use

1. Do Nothing

- Existence of multicollinearity in a data set does not necessarily mean that parameter estimates have unacceptably high sampling variances
- Researcher may be interested in linear combinations of the parameters, instead of the individual parameters
- If forecasting is the purpose of research, then as long as multicollinearity is stable in the future, the estimated model may perform well in prediction

2. Drop a variable

The simplest solution is to drop one of the correlated independent variables from the model

- However, this may create a [different](#), and potentially, worse problem
- To understand the problem we need to introduce the problem known as “omitting a relevant variable” (covered thoroughly in Lecture 11)

Suppose that the true regression model is,

$$y_t = \beta_1 + \beta_2 x_{2,t} + \beta_3 x_{3,t} + e_t$$

but we omit $x_{3,t}$ and estimate the following model,

$$y_t = \beta_1 + \beta_2 x_{2,t} + e_t$$

then, in general, the LS estimator b_2 is [biased](#),

$$E(b_2) = \beta_2 + \beta_3 b_{23}$$

where b_{23} is the slope from a regression of $x_{2,t}$ on $x_{3,t}$

Dropping a correlated independent variable from a regression involves a tradeoff between bias and efficiency

- Dropping the variable may increase efficiency because multicollinearity is avoided, but at the cost of obtaining a biased estimate
- Hence, dropping variables that are clearly important based on economic theory generally is considered bad practice

3. Additional information

The problem of multicollinearity is really that the data do not contain enough information about the individual relationships of the independent variables relative to the dependent variable

Therefore, a solution is to obtain more information and include it in the analysis

- Expand the sample
- Add structure to the problem by introducing non-sample (“extra”) information in the form of restrictions on the parameters

Non-Sample Information and the Beer Demand Regression

Recall the original economic model,

$$\ln y = \beta_1 + \beta_2 \ln x_2 + \beta_3 \ln x_3 + \beta_4 \ln x_4 + \beta_5 \ln x_5$$

Demand theory suggests a relevant piece of non-sample information is that if all prices and income go up by the same proportion, quantity demanded should not change (no “money illusion”)

We can impose this restriction by multiplying each independent variable by a constant,

$$\ln y = \beta_1 + \beta_2 \ln(\lambda x_2) + \beta_3 \ln(\lambda x_3) + \beta_4 \ln(\lambda x_4) + \beta_5 \ln(\lambda x_5)$$

$$\ln y = \beta_1 + \beta_2 \ln \lambda + \beta_2 \ln x_2 + \beta_3 \ln \lambda + \beta_3 \ln x_3 + \beta_4 \ln \lambda + \beta_4 \ln x_4 + \beta_5 \ln \lambda + \beta_5 \ln x_5$$

$$\ln y = \beta_1 + \beta_2 \ln x_2 + \beta_3 \ln x_3 + \beta_4 \ln x_4 + \beta_5 \ln x_5 + (\beta_2 + \beta_3 + \beta_4 + \beta_5) \ln \lambda$$

In order for $\ln y$ to be unchanged,

$$\beta_2 + \beta_3 + \beta_4 + \beta_5 = 0$$

There are two ways to introduce non-sample information in the form of parameter restrictions

- One is to estimate a re-parameterized version of the regression model by imposing the restrictions on the model
- The other is to develop a general restricted estimator

Both yield exactly the same estimation results

We will use the re-parameterization approach as it is the most direct approach

To derive the re-parameterized beer demand regression, re-write the parameter restriction as,

$$\beta_4 = -\beta_2 - \beta_3 - \beta_5$$

Substitute this restriction into the original regression model,

$$\ln y_t = \beta_1 + \beta_2 \ln x_{2,t} + \beta_3 \ln x_{3,t} + (-\beta_2 - \beta_3 - \beta_5) \ln x_{4,t} + \beta_5 \ln x_{5,t} + e_t$$

Re-arranging,

$$\ln y_t = \beta_1 + \beta_2 (\ln x_{2,t} - \ln x_{4,t}) + \beta_3 (\ln x_{3,t} - \ln x_{4,t}) + \beta_5 (\ln x_{5,t} - \ln x_{4,t}) + e_t$$

or,

$$\ln y_t = \beta_1 + \beta_2 \ln \left(\frac{x_{2,t}}{x_{4,t}} \right) + \beta_3 \ln \left(\frac{x_{3,t}}{x_{4,t}} \right) + \beta_5 \ln \left(\frac{x_{5,t}}{x_{4,t}} \right) + e_t$$

To get restricted least squares estimates that satisfy the restrictions, we can simply apply least squares to the above model

The estimation results are,

$$\hat{\ln y}_t = -4.798 - 1.299 \ln \left(\frac{x_{2,t}}{x_{4,t}} \right) + 0.187 \ln \left(\frac{x_{3,t}}{x_{4,t}} \right) + 0.946 \ln \left(\frac{x_{5,t}}{x_{4,t}} \right)$$

(3.714) (0.166) (0.284) (0.427)

$$R^2 = 0.808$$

We can obtain estimates of the eliminated parameter β_4 by simply substituting the other parameter estimates into the restriction as follows,

$$\begin{aligned} b_4 &= -(-1.2994) - 0.1868 - 0.9458 \\ &= 0.1668 \end{aligned}$$

A similar, but somewhat more complicated, procedure can be used to infer $se(b_4) = 0.077$ from the standard errors and covariances of the estimated restricted model

Leads to the question of why did we choose to eliminate β_4 ?

- Remember that $x_{4,t}$ is the price of all goods and services in the economy except beer and liquor
- Hence, dividing through by $x_{4,t}$ can be thought of as deflating nominal values to obtain real, or inflation-adjusted, values
- Whenever we deflate variables in demand analysis, we are in effect imposing the restriction examined here

Table 11.2 Unrestricted and Restricted Least Squares Estimates and Their Standard Errors (in Parentheses)

Parameter	Unrestricted		Restricted	
β_1	-3.243	(3.743)	-4.798	(3.714)
β_2	-1.020	(0.239)	-1.299	(0.166)
β_3	-0.583	(0.560)	0.187	(0.284)
β_4	0.210	(0.080)	0.167	(0.077)
β_5	0.923	(0.416)	0.946	(0.427)
R^2	0.825		0.808	

Griffiths, William E., R. Carter Hill, George G. Judge. Learning and Practicing Econometrics. John Wiley & Sons, Inc. New York. 1993.

Testing the Non-Sample Restriction in the Beer Demand Regression

If there is some [uncertainty](#) about whether the non-sample information is valid, we may want to test whether the non-sample information is consistent with the [sample data](#)

- The restriction considered for the beer demand regression is an example of a linear combination of parameters
- Examined how to test this form of restriction in the Lecture 6

1. Hypotheses

$$H_0 : \beta_2 + \beta_3 + \beta_4 + \beta_5 = 0$$

$$H_1 : \beta_2 + \beta_3 + \beta_4 + \beta_5 \neq 0$$

2. Test statistic

$$F = \frac{(SSE_R - SSE_U) / J}{SSE_U / (T - K)} = \frac{(0.098901 - 0.08992) / 1}{0.08992 / (30 - 5)} \\ = 2.27$$

where SSE_U is from the estimates for the unrestricted model,

$$\ln y_t = \beta_1 + \beta_2 \ln x_{2,t} + \beta_3 \ln x_{3,t} + \beta_4 \ln x_{4,t} + \beta_5 \ln x_{5,t} + e_t$$

and SSE_R is from the estimates for the restricted model,

$$\ln y_t = \beta_1 + \beta_2 \ln \left(\frac{x_{2,t}}{x_{4,t}} \right) + \beta_3 \ln \left(\frac{x_{3,t}}{x_{4,t}} \right) + \beta_5 \ln \left(\frac{x_{5,t}}{x_{4,t}} \right) + e_t$$

3. Rejection region

Reject the null hypothesis if $F \geq F_{\alpha}(J, T-K)$

If $\alpha = 0.05$, then $F_{\alpha}(J, T-K) = F_{0.05}(1, 30-5) = 4.24$

Reject if $F \geq 4.24$

4. Decision

- Since $2.27 < 4.24$ we fail to reject the null hypothesis that the sum of the slope parameters is zero
- Implies that beer consumers do not suffer from money illusion (at least most of the time!)

Important Points about Restricted Least Squares

Introduction of restrictions will reduce the sampling variance of parameter estimates, whether the restrictions are true or not

- Combining non-sample information with sample information reduces the variation in the estimation procedure caused by random sampling
- So, why not just impose non-sample restrictions until sampling variance is zero?

The reason we do not is that the restricted estimator is unbiased only if the constraints imposed are exactly correct

Since it is unlikely that constraints are exactly correct in practice, we usually give up unbiasedness in return for reduced variances

This leads to an important principle:

A good economist will obtain more reliable parameter estimates than a poor one, because a good economist will introduce better non-sample information

Final Perspective on Multicollinearity

Blanchard (1987, p.449)

“When students run their first ordinary least squares (OLS) regression, the first problem that they usually encounter is that of multicollinearity. Many of them conclude that there is something wrong with LS, some resort to new and often creative techniques to get around the problem. But, we tell them, this is wrong. Multicollinearity is God’s will, not a problem with OLS or statistical techniques in general. Only use of more economic theory in the form of additional restrictions may help alleviate the multicollinearity problem. One should not, however, expect miracles; *multicollinearity is likely to prevent the data from speaking loudly on some issues, even when all of the resources of economic theory have been exhausted.*” (Emphasis added)