2. Moreover, let $\hat{R}^2$ and $\tilde{R}^2$ be the unrestricted R^2 (i.e., the coefficient of determination) and the restricted R^2.

   Note that $\hat{R}^2 = 1 - \dfrac{e'e}{y'My}$ and $\tilde{R}^2 = 1 - \dfrac{\tilde{u}'\tilde{u}}{y'My}$, where $M = I_n - \dfrac{1}{n}ii'$ and $i = (1, 1, \cdots, 1)'$.

$$\frac{(\tilde{u}'\tilde{u} - e'e)/G}{e'e/(n-k)} = \frac{(\hat{R}^2 - \tilde{R}^2)/G}{(1 - \hat{R}^2)/(n-k)} \sim F(G, n-k).$$

   $e'e = (1 - \hat{R}^2)y'My$ and $\tilde{u}'\tilde{u} = (1 - \tilde{R}^2)y'My$ are substituted.

7 Example: F Distribution (Restricted OLS and Unrestricted OLS)

Date file $\Longrightarrow$ cons99.txt (Next slide)

Each column denotes year, nominal household expenditures (家計消費，10 billion yen), household disposable income (家計可処分所得，10 billion yen) and household expenditure deflator (家計消費デフレータ，1990=100) from the left.

1955	5430.1	6135.0	18.1	1970	37784.1	45913.2	35.2	1985	185335.1	220655.6	93.9
1956	5974.2	6828.4	18.3	1971	42571.6	51944.3	37.5	1986	193069.6	229938.8	94.8
1957	6686.3	7619.5	19.0	1972	49124.1	60245.4	39.7	1987	202072.8	235924.0	95.3
1958	7169.7	8153.3	19.1	1973	59366.1	74924.8	44.1	1988	212939.9	247159.7	95.8
1959	8019.3	9274.3	19.7	1974	71782.1	93833.2	53.3	1989	227122.2	263940.5	97.7
1960	9234.9	10776.5	20.5	1975	83591.1	108712.8	59.4	1990	243035.7	280133.0	100.0
1961	10836.2	12869.4	21.8	1976	94443.7	123540.9	65.2	1991	255531.8	297512.9	102.5
1962	12430.8	14701.4	23.2	1977	105397.8	135318.4	70.1	1992	265701.6	309256.6	104.5
1963	14506.6	17042.7	24.9	1978	115960.3	147244.2	73.5	1993	272075.3	317021.6	105.9
1964	16674.9	19709.9	26.0	1979	127600.9	157071.1	76.0	1994	279538.7	325655.7	106.7
1965	18820.5	22337.4	27.8	1980	138585.0	169931.5	81.6	1995	283245.4	331967.5	106.2
1966	21680.6	25514.5	29.0	1981	147103.4	181349.2	85.4	1996	291458.5	340619.1	106.0
1967	24914.0	29012.6	30.1	1982	157994.0	190611.5	87.7	1997	298475.2	345522.7	107.3
1968	28452.7	34233.6	31.6	1983	166631.6	199587.8	89.5				
1969	32705.2	39486.3	32.9	1984	175383.4	209451.9	91.8				

Estimate using TSP 5.0.

```
LINE  ************************************************
|       1  freq a;
|       2  smpl 1955 1997;
|       3  read(file='cons99.txt') year cons yd price;
|       4  rcons=cons/(price/100);
|       5  ryd=yd/(price/100);
|       6  d1=0.0;
|       7  smpl 1974 1997;
|       8  d1=1.0;
|       9  smpl 1956 1997;
|      10  d1ryd=d1*ryd;
|      11  olsq rcons c ryd;
|      12  olsq rcons c d1 ryd d1ryd;
|      13  end;
******************************************************
```

```
                 Equation   1
                 ============

   Method of estimation = Ordinary Least Squares


 Dependent variable: RCONS
 Current sample:  1956 to 1997
 Number of observations:  42

  Mean of dependent variable = 149038.
 Std. dev. of dependent var. = 78147.9
    Sum of squared residuals = .127951E+10
       Variance of residuals = .319878E+08
    Std. error of regression = 5655.77
                   R-squared = .994890
          Adjusted R-squared = .994762
     Durbin-Watson statistic = .116873
   F-statistic (zero slopes) = 7787.70
  Schwarz Bayes. Info. Crit. = 17.4101
  Log of likelihood function = -421.469

           Estimated    Standard
```

Variable	Coefficient	Error	t-statistic
C	-3317.80	1934.49	-1.71508
RYD	.854577	.968382E-02	88.2480

Read `E+10`, `E+08` and `E-02` as $\times 10^{10}$, $\times 10^{8}$ and $\times 10^{-2}$, respectively.

That is, `.127951E+10` should be $.127951 \times 10^{10}$.

```
                Equation   2
                ============

  Method of estimation = Ordinary Least Squares


Dependent variable: RCONS
Current sample:  1956 to 1997
Number of observations:  42

 Mean of dependent variable = 149038.
Std. dev. of dependent var. = 78147.9
   Sum of squared residuals = .244501E+09
      Variance of residuals = .643423E+07
   Std. error of regression = 2536.58
                  R-squared = .999024
         Adjusted R-squared = .998946
    Durbin-Watson statistic = .420979
  F-statistic (zero slopes) = 12959.1
 Schwarz Bayes. Info. Crit. = 15.9330
 Log of likelihood function = -386.714

           Estimated    Standard
```

Variable	Coefficient	Error	t-statistic
C	4204.11	1440.45	2.91861
D1	-39915.3	3154.24	-12.6545
RYD	.786609	.015024	52.3561
D1RYD	.194495	.018731	10.3839

1. `Equation 1`

   Significance test:

   `Equation 1` is:

$$\mathtt{RCONS} = \beta_1 + \beta_2 \mathtt{RYD}$$

   $H_0: \ \beta_2 = 0$

   (No.1) t Test $\Longrightarrow$ Compare 88.2480 and $t(42-2)$.

   (No.2) F Test $\Longrightarrow$ Compare $\dfrac{R^2/G}{(1-R^2)/(n-k)} = \dfrac{.994890/1}{(1-.994890)/(42-2)} =$ 7787.8 and $F(1,40)$. Note that $\sqrt{7787.8} = 88.2485$.

   1% point of $F(1,40) = 7.31$

   $H_0: \ \beta_2 = 0$ is rejected.

2. `Equation 2:`

$$\mathtt{RCONS} = \beta_1 + \beta_2\mathtt{D1} + \beta_3\mathtt{RYD} + \beta_4\mathtt{RYD} \times \mathtt{D1}$$

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

F Test $\Longrightarrow$ Compare $\dfrac{R^2/G}{(1-R^2)/(n-k)} = \dfrac{.999024/3}{(1-.999024)/(42-4)} = 12965.5$ and $F(3,38)$.

1% point of $F(3,38) = 4.34$

$H_0: \ \beta_2 = \beta_3 = \beta_4 = 0$ is rejected.

3. `Equation 1` vs. `Equation 2`

Test the structural change between 1973 and 1974.

`Equation 2` is:

$$\texttt{RCONS} = \beta_1 + \beta_2\texttt{D1} + \beta_3\texttt{RYD} + \beta_4\texttt{RYD} \times \texttt{D1}$$

$H_0:\ \beta_2 = \beta_4 = 0$

Restricted OLS $\Longrightarrow$ `Equation 1`

Unrestricted OLS $\Longrightarrow$ `Equation 2`

$$\frac{(\tilde{u}'\tilde{u} - e'e)/G}{e'e/(n-k)} = \frac{(.127951\texttt{E}+10 - .244501\texttt{E}+09)/2}{.244501\texttt{E}+09/(42-4)} = 80.43$$

which should be compared with $F_{0.01}(2, 38)$.

1% point of F distribution with (2,38) degrees of freedom

$\Longrightarrow\ F_{0.01}(2, 38) = 5.211 < 80.43$

$H_0:\ \beta_2 = \beta_4 = 0$ is rejected. $\Longrightarrow$ The structure was changed in 1974.

Or, using the coefficients of determination,

$$\begin{aligned}\frac{(\tilde{u}'\tilde{u} - e'e)/G}{e'e/(n-k)} &= \frac{(\hat{R}^2 - \tilde{R}^2)/G}{(1-\hat{R}^2)/(n-k)} \\ &= \frac{(.999024 - .994890)/2}{(1 - .999024)/(42-4))} \\ &= 80.48 > 5.211 = F_{0.01}(2, 38)\end{aligned}$$

$\hat{R}^2$ and $\tilde{R}^2$ are the unrestricted R^2 and the restricted R^2.

8 Generalized Least Squares Method (GLS, 一般化最小自乗法)

1. Regression model: $y = X\beta + u, \qquad u \sim N(0, \sigma^2\Omega)$

2. **Heteroscedasticity (不等分散，不均一分散)**

$$\sigma^2\Omega = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n^2 \end{pmatrix}$$

First-Order Autocorrelation (一階の自己相関，系列相関)

In the case of time series data, the subscript is conventionally given by t, not i .

$u_t = \rho u_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma_\epsilon^2)$

$$\sigma^2\Omega = \frac{\sigma_\epsilon^2}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix}$$

$$\mathrm{V}(u_t) = \sigma^2 = \frac{\sigma_\epsilon^2}{1-\rho^2}$$

3. The Generalized Least Squares (GLS，一般化最小二乗法) estimator of β,

denoted by b, solves the following minimization problem:

$$\min_b \ (y - Xb)'\Omega^{-1}(y - Xb)$$

The GLSE of β is:

$$b = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$$

4. In general, when Ω is symmetric, Ω is decomposed as follows.

$$\Omega = A'\Lambda A$$

Λ is a diagonal matrix, where the diagonal elements of Λ are given by the eigen values.

A is a matrix consisting of eigen vectors.

When Ω is a positive definite matrix, all the diagonal elements of Λ are positive.

5. There exists P such that $\Omega = PP'$ (i.e., take $P = A'\Lambda^{1/2}$). $\Longrightarrow$ $P^{-1}\Omega P'^{-1} = I_n$

Multiply P^{-1} on both sides of $y = X\beta + u$.

We have:

$$y^\star = X^\star\beta + u^\star,$$

where $\quad y^\star = P^{-1}y, \quad X^\star = P^{-1}X, \quad$ and $\quad u^\star = P^{-1}u.$

The variance of $u^\star$ is:

$$\mathrm{V}(u^\star) = \mathrm{V}(P^{-1}u) = P^{-1}\mathrm{V}(u)P'^{-1} = \sigma^2 P^{-1}\Omega P'^{-1} = \sigma^2 I_n.$$

because $\Omega = PP'$, i.e., $P^{-1}\Omega P'^{-1} = I_n$.

Accordingly, the regression model is rewritten as:

$$y^\star = X^\star\beta + u^\star, \qquad u^\star \sim (0, \sigma^2 I_n)$$

Apply OLS to the above model.

Let b be as estimator of β from the above model.

That is, the minimization problem is given by:

$$\min_b \ (y^\star - X^\star b)'(y^\star - X^\star b),$$

which is equivalent to:

$$\min_b \ (y - Xb)'\Omega^{-1}(y - Xb).$$

Solving the minimization problem above, we have the following estimator:

$$\begin{aligned} b &= (X^{\star\prime}X^\star)^{-1}X^{\star\prime}y^\star \\ &= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y, \end{aligned}$$

which is called GLS (Generalized Least Squares) estimator.

b is rewritten as follows:

$$b = \beta + (X^{\star\prime} X^{\star})^{-1} X^{\star\prime} u^{\star} = \beta + (X'\Omega^{-1}X)^{-1} X'\Omega^{-1} u$$

The mean and variance of b are given by:

$$\mathrm{E}(b) = \beta,$$

$$\mathrm{V}(b) = \sigma^2 (X^{\star\prime} X^{\star})^{-1} = \sigma^2 (X'\Omega^{-1}X)^{-1}.$$

6. Suppose that the regression model is given by:

$$y = X\beta + u, \qquad u \sim N(0, \sigma^2\Omega).$$

In this case, when we use OLS, what happens?

$$\hat{\beta} = (X'X)^{-1}X'y = \beta + (X'X)^{-1}X'u$$

$$\mathrm{V}(\hat{\beta}) = \sigma^2 (X'X)^{-1} X'\Omega X (X'X)^{-1}$$

Compare GLS and OLS.

(a) Expectation:

$$\mathrm{E}(\hat{\beta}) = \beta, \quad \text{and} \quad \mathrm{E}(b) = \beta$$

Thus, both $\hat{\beta}$ and b are unbiased estimator.

(b) Variance:

$$\mathrm{V}(\hat{\beta}) = \sigma^2 (X'X)^{-1} X'\Omega X (X'X)^{-1}$$

$$\mathrm{V}(b) = \sigma^2 (X'\Omega^{-1} X)^{-1}$$

Which is more efficient, OLS or GLS?.

$$
\begin{aligned}
\mathrm{V}(\hat{\beta}) - \mathrm{V}(b) &= \sigma^2 (X'X)^{-1} X'\Omega X (X'X)^{-1} - \sigma^2 (X'\Omega^{-1}X)^{-1} \\
&= \sigma^2 \Big((X'X)^{-1}X' - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1} \Big) \Omega \\
&\qquad \times \Big((X'X)^{-1}X' - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1} \Big)' \\
&= \sigma^2 A\Omega A'
\end{aligned}
$$

Note that A is $k \times n$ and Ω is $n \times n$.

Ω is the variance-covariance matrix of u, which is a positive definite matrix.

Therefore, except for $\Omega = I_n$, $A\Omega A'$ is also a positive definite matrix.

(From $\Omega = PP'$ and $A\Omega A' = AP(AP)'$, we have $xAP(xAP)' = \sum_{i=1}^{k} z_i^2 > 0$ for $x \neq 0$, where x is $1 \times k$, $z = xAP$ is $1 \times k$ and $z = (z_1, z_2, \cdots, z_k)$.)

This implies that $\mathrm{V}(\hat{\beta}_i) - \mathrm{V}(b_i) > 0$ for the ith element of β.

Accordingly, b is more efficient than $\hat{\beta}$.

7. If $u \sim N(0, \sigma^2\Omega)$, then $b \sim N(\beta, \sigma^2(X'\Omega^{-1}X)^{-1})$.

   Consider testing the hypothesis $H_0 : R\beta = r$.

   $$R : \ G \times k, \quad \text{rank}(R) = G \leq k.$$

   $$Rb \sim N(R\beta, \sigma^2 R(X'\Omega^{-1}X)^{-1}R').$$

   Therefore, the following quadratic form is distributed as:

   $$\frac{(Rb - r)'(R(X'\Omega^{-1}X)^{-1}R')^{-1}(Rb - r)}{\sigma^2} \sim \chi^2(G)$$

8. Because $(y^\star - X^\star b)'(y^\star - X^\star b)/\sigma^2 \sim \chi^2(n - k)$, we obtain:

   $$\frac{(y - Xb)'\Omega^{-1}(y - Xb)}{\sigma^2} \sim \chi^2(n - k)$$

9. Furthermore, from the fact that b is independent of $y - Xb$, the following F distribution can be derived:

$$\frac{(Rb - r)'(R(X'\Omega^{-1}X)^{-1}R')^{-1}(Rb - r)/G}{(y - Xb)'\Omega^{-1}(y - Xb)/(n - k)} \sim F(G, n - k)$$

10. Let b be the unrestricted GLSE and $\tilde{b}$ be the restricted GLSE.

    Their residuals are given by e and $\tilde{u}$, respectively.

$$e = y - Xb, \qquad \tilde{u} = y - X\tilde{b}$$

    Then, the F test statistic is written as follows:

$$\frac{(\tilde{u}'\Omega^{-1}\tilde{u} - e'\Omega^{-1}e)/G}{e'\Omega^{-1}e/(n - k)} \sim F(G, n - k)$$