6.1 Instrumental variables

In the linear regression model
\[ y_i = x'_i \beta + \varepsilon_i \]  
(6.1)
we have been assuming that \( x_i \) and \( \varepsilon_i \) are uncorrelated. This happens to be a critical assumption because without it none of the proofs of consistency and unbiasedness of OLS or GLS would remain valid.

An alternative method of estimation to deal with this problem is the method of Instrumental Variables (IV). The idea in IV estimation is the existence of an additional set of variables \( z_i \) that are correlated with \( x_i \), but uncorrelated with \( \varepsilon_i \).

6.1.1 Assumptions

Recall the assumptions of the classical regression model: (1) linearity, (2) full rank, (3) exogeneity of the independent variables, (4) homoscedasticity and nonautocorrelation, (5) stochastic or nonstochastic data, and (6) normally distributed disturbances. All but assumption (3) will hold here. Our new assumption is:

**Assumption 3IV:**
\[ E[\varepsilon_i | x_i] = \eta_i \]  
(6.2)

The implication from this assumption is that the disturbances and the regressors are now correlated:

**Assumption 3IV:**
\[ E[x_i \varepsilon_i] = \gamma \]  
(6.3)
where \( \gamma \neq 0 \). We say that the regressors \( X \) are no longer exogenous. Now, assume that there exist an additional set of variables \( Z \) such that the two following properties hold:

1. **Exogeneity:** \( Z \) is uncorrelated with the disturbance \( \varepsilon \).
2. **Relevance:** \( Z \) is correlated with the independent variables \( X \).

\( Z \) is called the set of instrumental variables.

### 6.1.2 Ordinary Least Squares

Under **Assumption 3IV** the OLS estimator, \( b \), is no longer unbiased:

\[
E[b|X] = \beta + (X'X)^{-1}X'\eta \neq \beta.
\]  

(6.4)

Hence, the Gauss-Markov theorem does not longer hold. Not only that OLS is biased, it is also inconsistent.

### 6.1.3 Instrumental Variables estimator

The derivation of the IV estimator is the following:

\[
y = X\beta + \varepsilon
\]  

(6.5)

\[
Z'y = Z'X\beta + Z'\varepsilon
\]

\[
Z'y = Z'X\beta
\]

\[
b_{IV} = (Z'X)^{-1}Z'y
\]

The estimate of the asymptotic variance of \( b_{IV} \) is:

\[
\text{Est.Asy.Var}[b_{IV}] = \hat{\sigma}^2(Z'X)^{-1}(Z'Z)(X'X)^{-1},
\]  

(6.6)

where \( \hat{\sigma}^2 \) is the consistent estimator of \( \sigma^2 \) and it is given by:

\[
\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n},
\]  

(6.7)

with \( \hat{\varepsilon} \) being just the estimated residuals from the IV regression.

### 6.1.4 Two-Stage Least Squares

The IV approach works fine when the number of columns in \( Z \) is the same as the number of columns in \( X \). However, what if we have more instruments than regressors in the model? Then \( b_{IV} \) cannot be estimated because \( Z'X \) will not be full rank and cannot be inverted. Because every linear combination of \( Z \) is uncorrelated with \( \varepsilon \), one naive approach would be simply to drop columns in \( Z \) until we have the same
number of columns as in $X$. Clearly this means that we are loosing some informations. A better approach is to use the projection of the columns of $X$ on the column space of $Z$:

$$
\hat{X} = Z(Z'Z)^{-1}Z'X
$$

(6.8)

Then, we just use $\hat{X}$ as the set of instruments $Z$ in the IV estimation.

$$
b_{IV} = (\hat{X}'\hat{X})^{-1}\hat{X}'y
$$

(6.9)

$\mathbf{M}_Z$ is the residual maker matrix based on $Z$, hence $(\mathbf{I} - \mathbf{M}_Z)$ is the (idempotent) projection matrix based on $Z$. When $\hat{X}$ is the set of instruments, the IV estimator is computed by OLS of $y$ on $\hat{X}$. Therefore, we can see $b_{IV}$ as a two step approach; first computing $\hat{X}$ and then using $\hat{X}$ in Equation 6.9. This is why this is called the two-stage least squares (2SLS) estimator. A key point here is that the asymptotic covariance matrix in Equation 6.6 should not use the $\hat{\sigma}^2$ presented in Equation 6.7 because this last one is inconsistent for $\sigma^2$.

### 6.1.5 Hausman and Wu Test

If $X$ is uncorrelated with $\varepsilon$, the asymptotic covariance of the OLS estimator is never larger than that of the IV estimator. Hence, if $X$ is not endogenous (not correlated with $\varepsilon$), OLS is the preferred estimator. Obviously, if $X$ is endogenous we have to use IV. The problem in choosing between OLS and IV is that sometimes it is not clear whether $X$ is correlated with $\varepsilon$. Recall that $\varepsilon$ is unobserved, so the question is not that trivial. Fortunately Hausman (1978) and Wu (1973) developed a test.

The idea behind the Hausman’s test is as follows. The null hypothesis is that $X$ is uncorrelated with $\varepsilon$. Under the null we have two consistent estimators of $\beta$, $b_{LS}$ and $b_{IV}$. Under the alternative only $b_{IV}$ is consistent. Then, if we examine:

$$
d = b_{IV} - b_{LS},
$$

(6.10)

under the null, plim $d = 0$, whereas under the alternative, plim $d \neq 0$. $d$ is given by:

$$
d = (\hat{X}'\hat{X})^{-1}\hat{X}'e,
$$

(6.11)

and the Wald statistic used to test this hypothesis is:

$$
H = \frac{d'(\hat{X}'\hat{X})^{-1} - (X'X)^{-1}]^{-1}d}{\hat{\sigma}^2}.
$$

(6.12)
Based on data from the 1980 census, we want to estimate the following equation:

\[ rent_i = \beta_0 + \beta_1 hsngval_i + \beta_2 pcturban_i + u_i, \]  

(6.13)

where \( rent \) is the median monthly gross rent, \( hsngval \) is the median dollar value of owner-occupied housing, and \( pcturban \) is the percentage of the population living in urban areas. Because random shocks that affect rental rates in a state probably also affect housing values, we treat \( hsngval \) as endogenous (correlated with \( u_i \)).

Because \( hsngval \) is potentially endogenous, we need at least one instrument that is correlated with \( hsngval \) but uncorrelated with \( u_i \). These excluded exogenous variables must not affect \( rent \), because if they do then they should be included in the regression. We have family income (\( faminc \)) and region of the country (\( region \)) that we believe are correlated with \( hsngval \) but not with \( u_i \). The set of instrumental variables is then \( pcturban, faminc, 2.region, 3.region, \) and \( 4.region \).

Let’s say that we first estimate the model using OLS:

```
regress rent pcturban hsngval
```

```
Source | SS df MS Number of obs = 50
-------------+------------------------------ F( 2, 47) = 47.54
Model | 40983.5269 2 20491.7635 Prob > F = 0.0000
Residual | 20259.5931 47 431.055172 R-squared = 0.6692
-------------+------------------------------ Adj R-squared = 0.6551
Total | 61243.12 49 1249.85959 Root MSE = 20.762

------------------------------------------------------------------------------
rent | Coef. Std. Err. t P>|t| [95% Conf. Interval]
-------------+----------------------------------------------------------------
pcturban | .5248216 .2490782 2.11 0.040 .0237408 1.025902
hsngval | .0015205 .0002276 6.68 0.000 .0010627 .0019784
_cons | 125.9033 14.18537 8.88 0.000 97.36603 154.4406
------------------------------------------------------------------------------
```

The 2SLS command is:

```
ivregress 2sls rent pcturban (hsngval = faminc i.region)
```

```
Instrumental variables (2SLS) regression Number of obs = 50
Wald chi2(2) = 90.76
Prob > chi2 = 0.0000
R-squared = 0.5989
Root MSE = 22.166

------------------------------------------------------------------------------
rent | Coef. Std. Err. z P>|z| [95% Conf. Interval]
-------------+----------------------------------------------------------------
hsngval | .0022398 .0003284 6.82 0.000 .0015961 .0028836
pcturban | .081516 .2987652 0.27 0.785 -.504053 .667085
_cons | 120.7065 15.22839 7.93 0.000 90.85942 150.5536
------------------------------------------------------------------------------
Instrumented: hsngval
Instruments: pcturban faminc 2.region 3.region 4.region
```

Notice the big difference between the OLS estimate and the 2SLS estimate for \( hsngval \). If in addition we are interested in the result from the first stage, we need to type The 2SLS command is:
ivregress 2sls rent pcturban (hsngval = faminc i.region), first

First-stage regressions

|              | Coef.   | Std. Err. | t     | P>|t| [95% Conf. Interval] |
|--------------|---------|-----------|-------|------------------------|
| hsngval      | 182.2201| 115.0167  | 1.58  | 0.120 [-49.58092 414.0211] |
| pcturban     | 2.731324| .6818931  | 4.01  | 0.000 1.357058 4.105589 |
| faminc       |         |           |       |                        |
| region       |         |           |       |                        |
| 1            |         |           |       |                        |
| 2            | -5095.038| 4122.112  | -1.24 | 0.223 -13402.61 3212.533 |
| 3            | -1778.05| 4072.691  | -0.44 | 0.665 -9986.019 6429.919 |
| 4            | 13413.79| 4048.141  | 3.31  | 0.002 5255.296 21572.28 |
| _cons        | -18671.87| 11995.48  | -1.56 | 0.127 -42847.17 5503.439 |

For the Hausman-Wu test we need to run

estat endogenous

Tests of endogeneity
Ho: variables are exogenous
Durbin (score) chi2(1) = 12.8473 (p = 0.0003)
Wu-Hausman F(1,46) = 15.9067 (p = 0.0002)

This command reports the Durbin and the Hausman-Wu endogeneity test. The null hypothesis in these tests is that the variable under consideration (hsngval) can be treated as exogenous. Here both test statistics are highly significant, so we reject the null of exogeneity and we must treat hsngval as endogenous.

6.2 Simultaneous equations

An illustrative example of a system of simultaneous equations is a model of market equilibrium:

\[
\begin{align*}
\text{Demand equation:} & \quad q_{d,t} = \alpha_1 p_t + \alpha_2 x_t + \epsilon_{d,t} \\
\text{Supply equation:} & \quad q_{s,t} = \beta_1 p_t + \epsilon_{s,t} \\
\text{Equilibrium condition:} & \quad q_{d,t} = q_{s,t} = q_t
\end{align*}
\]  

(6.14) (6.15) (6.16)

There are structural equations derived theoretically, where each describes a particular aspect of the economy. Because the market equilibrium jointly determined price and quantity, they are jointly dependent or endogenous variables. This is a complete system of equations because the number of endogenous variables is equal to the number of equations. Because OLS is inconsistent in this system, potential estimators include IV and 2SLS.
6.2.1 The identification problem

The identification problem should be solved before estimation. We say that two theories are observationally equivalent if both are consistent with the same ‘data.’ Then we say that the structure is unidentified. Figure 6.1 shows that the observed data that is consistent with the two structures in panels (b) and (c). We say that in panel (a) the structure is unidentified. If we have additional information, for example, that supplied was fixed, then this rules out (c) and the correct structure is (b).

Fig. 6.1 Market equilibria and observational equivalence, from [Greene (2008)].

6.2.2 Single equation estimation methods

If we are only interested in a single equation of the system, we can employ the IV or the 2SLS estimators presented previously. The Stata example in Equation 6.13 can also be viewed as a simultaneous equation problem if we write the model as:

\[
\begin{align*}
\text{hsngval}_i &= \alpha_0 + \alpha_1 \text{faminc}_i + \alpha_2 \text{region}_i \\
&\quad + \alpha_3 \text{region}_i + \alpha_4 \text{region}_i + v_i, \\
\text{rent}_i &= \beta_0 + \beta_1 \text{hsngval}_i + \beta_2 \text{pcturban}_i + u_i,
\end{align*}
\]  

(6.17) (6.18)
6.2 Simultaneous equations

The results are the same as before.

6.2.3 System methods of estimation

The system of $M$ equation can be formulated as:

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix} =
\begin{bmatrix}
Z_1 & 0 & 0 & \ldots & 0 \\
0 & Z_2 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & Z_M
\end{bmatrix}
\begin{bmatrix}
\delta_1 \\
\delta_2 \\
\vdots \\
\delta_M
\end{bmatrix} +
\begin{bmatrix}
\varepsilon_1 \\
\varepsilon_2 \\
\vdots \\
\varepsilon_M
\end{bmatrix}
\]

or

\[
y = Z\delta + \varepsilon \quad (6.19)
\]

where

\[
E[\varepsilon|X] = 0, \quad \text{and} \quad E[\varepsilon\varepsilon'|X] = \tilde{\Sigma} = \Sigma \otimes I. \quad (6.20)
\]

OLS on the system is inconsistent. IV is consistent, but is inefficient compared to estimators that make use of the cross-sectional correlations of the disturbances. Three techniques are generally used for joint estimation of the entire system of equations: three-stage least squares, GMM, and full information maximum likelihood (FIML).

6.2.4 Three-stage least squares

The three-stage least squares estimator of the system presented in Equation 6.19 is:

\[
\hat{\delta}_{2SLS} = \left[\tilde{Z}'(\Sigma^{-1} \otimes I)\tilde{Z}\right]^{-1}\tilde{Z}'(\Sigma^{-1} \otimes I)y. \quad (6.21)
\]

That has the following asymptotic covariance matrix:

\[
\text{Asy.Var}[\hat{\delta}_{2SLS}] = \left[\tilde{Z}'(\Sigma^{-1} \otimes I)\tilde{Z}\right]^{-1}, \quad (6.22)
\]

which is estimated with $[\tilde{Z}'(\Sigma^{-1} \otimes I)\tilde{Z}]^{-1}$. Under certain conditions, 3SLS satisfies the requirements for an IV estimator, hence it is consistent. Moreover, 3SLS is asymptotically efficient.

Consider the simple example with $M = 2$ that relates consumption ($\text{consump}$) to private and government wages paid ($\text{wagepriv}$ and $\text{wagegovt}$). Simultaneously, private wages depend on consumption, total government expenditures ($\text{govt}$), and the lagged stock of capital in the economy ($\text{capital1}$). The model can be written as:

1 $\otimes$ is the kronecker product, which multiplies each of the elements of the first matrix by the entire second matrix.
\[
\text{consump} = \beta_0 + \beta_1 \text{wagepriv} + \beta_2 \text{wagegovt} + \epsilon_1 \quad (6.23)
\]
\[
\text{wagepriv} = \beta_3 + \beta_4 \text{consump} + \beta_5 \text{govt} + \beta_6 \text{capital1} + \epsilon_2 \quad (6.24)
\]

We assume that \text{consump} and \text{wagepriv} are endogenous while \text{wagegovt}, \text{govt}, and \text{capital1} are exogenous. The Stata commands for this example of 3SLS are:

```stata
reg3 (consump wagepriv wagegovt) (wagepriv consump govt capital1)
```

<table>
<thead>
<tr>
<th>Equation</th>
<th>Obs</th>
<th>Parms</th>
<th>RMSE</th>
<th>&quot;R-sq&quot;</th>
<th>chi2</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>consump</td>
<td>22</td>
<td>2</td>
<td>1.77</td>
<td>0.938</td>
<td>208.02</td>
<td>0.0000</td>
</tr>
<tr>
<td>wagepriv</td>
<td>22</td>
<td>3</td>
<td>2.37</td>
<td>0.854</td>
<td>80.04</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

|                         | Coef. | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|-------------------------|-------|------------|------|-----|---------------------|
| consump                 |       |            |      |     |                     |
| wagepriv                | .8012754 | .1279329 | 6.26 | 0.000 | .5505314 1.052019   |
| wagegovt                | 1.029531 | .3048424 | 3.38 | 0.001 | .432051 1.627011    |
| _cons                   | 19.3559 | 3.583772  | 5.40 | 0.000 | 12.33184 26.37996   |
| wagepriv                |       |            |      |     |                     |
| consump                 | .4026076 | .2567312 | 1.57 | 0.117 | -.1005764 .9057916   |
| govt                   | 1.177792 | .5421253 | 2.17 | 0.030 | .1152461 2.240338    |
| capital1                | -.0281145 | .0572111 | -.49 | 0.623 | -.1402462 .0840173   |
| _cons                   | 14.63026 | 10.26693 | 1.42 | 0.154 | -5.492552 34.75306   |

Endogenous variables: consump wagepriv
Exogenous variables: wagegovt govt capital1

Because the consumption equation is just identified—the number of excluded exogenous variables from this equation (govt and capital1) is equal to the number of endogenous variables included (consump and wagepriv)—the 3SLS point estimates are identical to the 2SLS estimates.

### 6.3 Systems of Equations

A multiple equation structure can be written as

\[
\begin{align*}
  y_1 &= X_1 \beta_1 + \epsilon_1, \\
  y_2 &= X_2 \beta_2 + \epsilon_2, \\
        &\vdots \\
  y_M &= X_M \beta_M + \epsilon_M,
\end{align*}
\]

where there are \( M \) equations and \( T \) observations in the sample.
6.3 Systems of Equations

6.3.1 Seemingly Unrelated Regressions

In the *seemingly unrelated regressions* (SUR) models the regressions are related because the (contemporaneous) errors associated with the dependent variables may be correlated. Equations 6.24 can be expressed as:

\[ y_i = X_i \beta_i + \epsilon_i \quad i = 1, 2, \ldots, M, \]  
(6.26)

where

\[ \epsilon = [\epsilon_1', \epsilon_2', \ldots, \epsilon_M']. \]  
(6.27)

\( X_i \) is strictly exogenous:

\[ E[\epsilon|X_1, X_2, \ldots, X_M] = 0, \]  
(6.28)

and the disturbances are homoscedastic:

\[ E[\epsilon_m \epsilon_m'|X_1, X_2, \ldots, X_M] = \sigma_{mm} I_T. \]  
(6.29)

As before, there are \( T \) observations to estimate \( M \) equation. Each equation has \( K_i \) regressors, for a total of \( K = \sum_{i=1}^{M} K_i \). We require at least \( T > K \) observations. In addition, we assume that disturbances are uncorrelated across observations but correlated across equations:

\[ E[\epsilon_t \epsilon_s|X_1, X_2, \ldots, X_M] = \sigma_{ij}, \quad \text{if} \ t = s \text{ and 0 otherwise.} \]  
(6.30)

Therefore, the formulation of the disturbance is:

\[ E[\epsilon \epsilon'|X_1, X_2, \ldots, X_M] = \sigma_{ij} I, \]  
(6.31)

or

\[ E[\epsilon \epsilon'|X_1, X_2, \ldots, X_M] = \Omega = \begin{bmatrix} \sigma_{11} I & \sigma_{12} I & \ldots & \sigma_{1M} I \\ \sigma_{21} I & \sigma_{22} I & \ldots & \sigma_{2M} I \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} I & \sigma_{M2} I & \ldots & \sigma_{MM} I \end{bmatrix}. \]

6.3.2 GLS on Seemingly Unrelated Regressions

Each of the \( M \) equations is, by itself, a classical regression. Hence, using OLS equation by equation yields consistent, but not necessarily efficient parameters. A more efficient estimator is to use *generalized least squares* to the stacked model
\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} = \begin{bmatrix}
  X_1 & 0 & \ldots & 0 \\
  0 & X_2 & \ldots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \ldots & X_M
\end{bmatrix} \begin{bmatrix}
  \beta_1 \\
  \beta_2 \\
  \vdots \\
  \beta_M
\end{bmatrix} + \begin{bmatrix}
  \epsilon_1 \\
  \epsilon_2 \\
  \vdots \\
  \epsilon_M
\end{bmatrix} = X\beta + \epsilon.
\]

For the \( t \)th observation, the \( M \times M \) covariance matrix of the disturbances is

\[
\Sigma = \begin{bmatrix}
  \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1M} \\
  \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2M} \\
  \vdots & \vdots & \ddots & \vdots \\
  \sigma_{M1} & \sigma_{M2} & \ldots & \sigma_{MM}
\end{bmatrix}
\]

Therefore, in Equation 6.31,

\[
\Omega = \Sigma \otimes I \quad (6.32)
\]

and

\[
\Omega^{-1} = \Sigma^{-1} \otimes I \quad (6.33)
\]

Then, the GLS estimator for the system in Equations 6.26 is

\[
\hat{\beta} = \left[ X' \Omega^{-1} X \right]^{-1} X' \Omega^{-1} y \quad (6.34)
\]

The GLS results covered before apply to this model. The equations are linked only by the disturbances. How much efficiency is gained when using GLS over OLS?

1. If the equations are unrelated, GLS yields the same estimates as OLS.
2. If we have the same \( X \) for all equations, then GLS and OLS are identical.
3. If the regression in one block of equations are a subset of those in another, then GLS has no efficiency gains over OLS in the estimation of the smaller set of equations.

With unrestricted correlation of the disturbances and different regressors in the equations, two propositions that generally apply are:

1. The greater the correlation of the disturbances, the greater the efficiency gain in using GLS.
2. The less the correlation in the \( X \) matrices, the greater the efficiency gain in using GLS.

### 6.3.3 SUR in Stata

Consider the following system of two equations:

\[
\text{price} = \beta_0 + \beta_1 \text{foreign} + \beta_2 \text{length} + u_1 \quad (6.35)
\]

\[
\text{weight} = \gamma_0 + \gamma_1 \text{foreign} + \gamma_2 \text{length} + u_2 \quad (6.36)
\]
6.3 Systems of Equations

Even if the share the same $X$ and SUR is this system is the same as OLS, the benefit in using arises, if for example, we want to test $H_0: \beta_2 = \gamma_2$. The commands in Stata are:

```
sureg (price foreign length) (weight foreign length), small dfk
```

The options `small` and `dfk` are used to obtain small-sample statistics comparable with OLS.

Seemingly unrelated regression

<table>
<thead>
<tr>
<th>Equation</th>
<th>Obs</th>
<th>Parms</th>
<th>RMSE</th>
<th>&quot;R-sq&quot;</th>
<th>F-Stat</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td>74</td>
<td>2</td>
<td>2474.593</td>
<td>0.3154</td>
<td>16.35</td>
<td>0.0000</td>
</tr>
<tr>
<td>weight</td>
<td>74</td>
<td>2</td>
<td>250.2515</td>
<td>0.8992</td>
<td>316.54</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

```
| Coef.     | Std. Err. | t     | P>|t|   | [95% Conf. Interval] |
|-----------|-----------|-------|------|----------------------|
| price     |           |       |      |                      |
| foreign   | 2801.143  | 766.117 | 3.66 | 0.000 | 1286.674 | 4315.611 |
| length    | 90.21239  | 15.83368 | 5.70 | 0.000 | 58.91219 | 121.5126 |
| _cons     | -11621.35 | 3124.436 | -3.72 | 0.000 | -17797.77 | -5444.93  |
| weight    |           |       |      |                      |
| foreign   | -133.6775 | 77.47615 | -1.73 | 0.087 | -286.8332 | 19.4782 |
| length    | 31.44455  | 1.601234 | 19.64 | 0.000 | 28.27921 | 34.60989 |
| _cons     | -2850.25  | 315.9691 | -9.02 | 0.000 | -3474.861 | -2225.639 |
```

Then we can test $H_0: \beta_2 = \gamma_2$ by using:

```
test length
```

The options `small` and `dfk` are used to obtain small-sample statistics comparable with OLS.

```
( 1) [price]foreign = 0
( 2) [weight]foreign = 0
F( 2, 142) = 17.99
Prob > F = 0.0000
```

Now, consider the following system where the $X$ are different across equations:

\[
\begin{align*}
\text{price} &= \beta_0 + \beta_1 \text{foreign} + \beta_2 \text{mpg} + \beta_3 \text{displ} + u_1 \\
\text{weight} &= \gamma_0 + \gamma_1 \text{foreign} + \gamma_2 \text{lenght} + u_2
\end{align*}
\] (6.37) (6.38)

Here the command is:

```
sureg (price foreign mpg displ) (weight foreign length), corr
```

That yields the following output

Seemingly unrelated regression

<table>
<thead>
<tr>
<th>Equation</th>
<th>Obs</th>
<th>Parms</th>
<th>RMSE</th>
<th>&quot;R-sq&quot;</th>
<th>chi2</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td>74</td>
<td>3</td>
<td>2165.321</td>
<td>0.4537</td>
<td>49.64</td>
<td>0.0000</td>
</tr>
<tr>
<td>weight</td>
<td>74</td>
<td>2</td>
<td>245.2916</td>
<td>0.8990</td>
<td>661.84</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

```
| Coef.     | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-----------|-----------|-------|------|----------------------|
| price     |           |       |      |                      |
| foreign   | -265.6775 | 77.47615 | -3.42 | 0.000 | -415.1022 | -216.2531 |
| length    | 311.44455 | 1.601234 | 19.44 | 0.000 | 282.7921 | 340.0999 |
| _cons     | -2750.25  | 315.9691 | -8.69 | 0.000 | -3474.861 | -2025.639 |
| weight    |           |       |      |                      |
| foreign   | -133.6775 | 77.47615 | -1.73 | 0.087 | -286.8332 | 19.4782 |
| length    | 31.44455  | 1.601234 | 19.64 | 0.000 | 28.27921 | 34.60989 |
| _cons     | -2850.25  | 315.9691 | -9.02 | 0.000 | -3474.861 | -2225.639 |
```
The option `corr` displays the correlation between the disturbances in Equations 6.38 and 6.38. In this case the correlation is 0.3285 and we reject the null hypothesis that the correlation is zero. Hence, GLS is more efficient than OLS.