2SLS Estimates
ECON 30331

Bill Evans
Fall 2015

Two-Stage Least Squares (2SLS)

Consider a standard linear bivariate regression model $y_i = \beta_0 + x_i\beta_1 + \varepsilon_i$. In this case, begin with the assumption that $E[\varepsilon_i \mid x_i] \neq 0$ which means that OLS estimates of $\hat{\beta}_1$ will produced biased and inconsistent estimates. Suppose however that there is a variable z_i that directly impacts x but it is uncorrelated with ε_i. This can be expressed as $E[\varepsilon_i \mid z_i] = 0$. The variable z is known as an 'instrument" in the model. The implicit assumption is that z impacts x and x impacts y. Therefore, if we were to somehow "shock" x with z, we can infer the relationship between x and y.

This is done by considering the relationship between x and z as a simple linear model. This "first-stage" relationship between x and z is given by the equation

$$x_i = \theta_0 + z_i\theta_1 + u_i$$

Where we assume *z* and u are uncorrelated. Note that *x* has two components. One is a deterministic one that is a function of z, $x_i^p = \theta_0 + z_i\theta_1$. Because z is uncorrelated with ε then this predictive component of x is also uncorrelated with ε. Since x is correlated with ε and x_i^p is not, it must be that what is driving the correlation between *x* and ε is then *u*.

The suggestion then is instead of using x in the regression $y_i = \beta_0 + x_i\beta_1 + \varepsilon_i$ -- which we know would produce biased estimates -- why not use x_i^p instead. The x_i^p component of x contains information about x, but it is NOT correlated with ε so it should produce estimates with "nice" properties. That is a great idea except that we do not know θ_0 or θ_1 and therefore, we cannot use x_i^p in a regression. We can however use the next best thing: an unbiased estimate of x_i^p which is produced from a first-stage regression of x on z.

This "two-stage least squares" procedure is very straight forward. Run a regression of the form $x_i = \theta_0 + z_i\theta_1 + u_i$. It should be no surprise that the OLS estimates for the parameters $\hat{\theta}_0$ and $\hat{\theta}_1$ are

$\hat{\theta}_0 = \bar{x} - \bar{z}\hat{\theta}_1$ and $\hat{\theta}_1 = \dfrac{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}{\sum_{i=1}^{n}(z_i - \bar{z})^2}$. With these estimates we can construct a predicted value for

$\hat{x}_i = \hat{\theta}_0 + z_i\hat{\theta}_1$. Instead of using x in the estimates for β_1 -- which produces $\hat{\beta}_1^{ols} = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

use the predicted value of x, $\hat{x}_i = \hat{\theta}_0 + z_i\hat{\theta}_1$, instead. The predicted value is a linear function of z which is uncorrelated with ε so the 2SLS estimate for β_1 should have "nice" properties. The 2SLS estimate for β_1 is therefore

$$\hat{\beta}_1^{2sls} = \frac{\sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})(y_i - \bar{y})}{\sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})^2}.$$

Working with the 2SLS estimate

Given that $\hat{x}_i = \hat{\theta}_0 + z_i\hat{\theta}_1$ it is clear that $\bar{\hat{x}} = \hat{\theta}_0 + \bar{z}\hat{\theta}_1$ and hence $\hat{x}_i - \bar{\hat{x}} = (z_i - \bar{z})\hat{\theta}_1$. Substitute this into the equation for $\hat{\beta}_1^{2sls}$ and we obtain

$$\hat{\beta}_1^{2sls} = \frac{\sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})(y_i - \bar{y})}{\sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})^2} = \frac{\sum_{i=1}^{n}\hat{\theta}_1(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}\hat{\theta}_1^2(z_i - \bar{z})^2} = \frac{\hat{\theta}_1\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\hat{\theta}_1^2\sum_{i=1}^{n}(z_i - \bar{z})^2} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\hat{\theta}_1\sum_{i=1}^{n}(z_i - \bar{z})^2}$$

Substitute the definition of $\hat{\theta}_1$ into this equation and we obtain

$$\hat{\beta}_1^{2sls} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\hat{\theta}_1\sum_{i=1}^{n}(z_i - \bar{z})^2} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\left(\frac{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}{\sum_{i=1}^{n}(z_i - \bar{z})^2}\right)\sum_{i=1}^{n}(z_i - \bar{z})^2} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}$$

Divide the numerator and denominator by (n-1) and we note the 2SLS estimte is therefore the ratio of the covariate between y and z divided by the covariance between x and z.

$$\hat{\beta}_1^{2sls} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})/(n-1)}{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})/(n-1)} = \frac{\hat{\sigma}_{yz}}{\hat{\sigma}_{zx}}$$

Is $\hat{\beta}_1^{2sls}$ an unbiased estimate?

Anytime we have a new random variable, we always ask what is the expected value and the variance? This is a difficult question. In the end, we are unable to show that $\hat{\beta}_1^{2sls}$ is an unbiased estimate in this context? Why? In order to demonstrate the properties of the estimate, we always substitute the truth back into the model. Start with the estimate for $\hat{\beta}_1^{2sls}$ and note that we can drop one of the means in both the numeatror and denominator.

$$\hat{\beta}_1^{2sls} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})y_i}{\sum_{i=1}^{n}(z_i - \bar{z})x_i}$$

At this point, we usually substitute in the defintion of y into the numerator. However, x is also a random variable that is a function of the error *v* so we would need to substitute that definition in as well.

$$\hat{\beta}_1^{2sls} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})y_i}{\sum_{i=1}^{n}(z_i - \bar{z})x_i} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(\beta_0 + x_i\beta_1 + \varepsilon_i)}{\sum_{i=1}^{n}(z_i - \bar{z})(\theta_0 + z_i\theta_1 + u_i)}$$

The problem is that we now have random variables in the numerator and the denominator. If a and b are random variables E[a/b]≠E[a]/E[b]. Without knowing the joint distribution of the numerator and denominator, we cannot figure out the expected value of the estimate. Our only option is then to consider whether the estimate is consistent.

The consistency of 2SLS estimates

To show whether the 2SLS estimate is consistent, start with the definition $\hat{\beta}_1^{2sls} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z})}$ drop $\bar{y}$ and substitute in the definition of y. This produces $\hat{\beta}_1^{2sls} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(\beta_0 + x_i\beta_1 + \varepsilon_i)}{\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z})}$. Expand the numerator, and we get $\hat{\beta}_1^{2sls} = \frac{\beta_0\sum_{i=1}^{n}(z_i - \bar{z})}{\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z})} + \frac{\beta_1\sum_{i=1}^{n}(z_i - \bar{z})x_i}{\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z})} + \frac{\sum_{i=1}^{n}(z_i - \bar{z})\varepsilon_i}{\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z})}$. The first term equals zero because sums of deviation from means equal zero. The second term equals β_1 and the

whole term reduces to $\hat{\beta}_1^{2sls} = \beta_1 + \frac{\sum_{i=1}^{n}(z_i - \overline{z})\varepsilon_i}{\sum_{i=1}^{n}(x_i - \overline{x})(z_i - \overline{z})}$. Divide the numerator and denominator of the right

hand term by (n-1) and you get $\hat{\beta}_1^{2sls} = \beta_1 + \frac{\sum_{i=1}^{n}(z_i - \overline{z})\varepsilon_i / (n-1)}{\sum_{i=1}^{n}(x_i - \overline{x})(z_i - \overline{z}) / (n-1)} = \beta_1 + \frac{\hat{\sigma}_{ze}}{\hat{\sigma}_{zx}}$. Now we ask what

happens to this estimate as the sample sizes increases to infinity. Note that if a and b are estimates, then

plim(a/b)=plim(a)/plim(b). Note also that the $p\lim(\hat{\sigma}_{z\varepsilon}) = \sigma_{z\varepsilon}$ and $p\lim(\hat{\sigma}_{zx}) = \sigma_{zx}$, and therefore, it is

no surprise that $p\lim(\hat{\beta}_1^{2SLS}) = \beta_1 + \frac{\sigma_{ze}}{\sigma_{zx}}$. In this case, the estimate $\hat{\beta}_1^{2SLS}$ will only be consistent if

$p\lim(\hat{\sigma}_{z\varepsilon}) = \sigma_{z\varepsilon} = 0$. To produce consistent estimates, it must be the case that z only produces a change

in y through an initial change in z. If z has a direct effect on y then $p\lim(\hat{\sigma}_{z\varepsilon}) = \sigma_{z\varepsilon} \neq 0$ and the model will

produce inconsistent estimates.

An alternative interpretation of 2SLS estimates

The structural equation of interest is $y_i = \beta_0 + x_i\beta_1 + \varepsilon_i$. We anticipate that x is correlated with ε so we cannot estimate this model by OLS. We do however have a variable z that we believe is predictive of x but it is not directly correlated with ε. The first-stage equation that relations x to z is given by the equation $x_i = \theta_0 + z_i\theta_1 + u_i$. Substitute this equation into the equation for y and one obtains the equation

$$y_i = \beta_0 + (\theta_0 + z_i\theta_1 + u_i)\beta_1 + \varepsilon_i = (\beta_0 + \theta_0\beta_1) + z_i\theta_1\beta_1 + u_i\beta_1 + \varepsilon_i$$

In this model, define $(\beta_0 + \theta_0\beta_1) = \pi_0$ let $\theta_1\beta_1 = \pi_1$ and let $u_i\beta_1 + \varepsilon_i = v_i$. This allows us to write the equation for y as $y_i = \pi_0 + z_i\pi_1 + v_i$. This equation is referred to as the "reduced-form" and it represents the correlation between z and y. The coefficient is $\frac{\partial y}{\partial z} = \pi_1$. Note that the coefficient has an interesting interpretation. The model we are suggesting is that y is a function of x and x is a function of z. Therefore y=f(x(z)). If we take the derivative of y with respect to z, because we assume z only changes y through a change in x, we obtain $\frac{\partial y}{\partial z} = \frac{\partial y}{\partial x}\left(\frac{\partial x}{\partial z}\right)$. Note that $\pi_1 = \beta_1\theta_1$, which is exactly $\frac{\partial y}{\partial x}\left(\frac{\partial x}{\partial z}\right)$. Therefore, if we take the reduced form estimate for π_1 which equals $\frac{\partial y}{\partial x}\left(\frac{\partial x}{\partial z}\right) = \beta_1\theta_1$ and divide it by the 1st stage coefficient θ_1 we get β_1. It is also the case that $\hat{\beta}_1^{2sls} = \frac{\hat{\pi}_1}{\hat{\theta}_1}$.

$$\hat{\beta}_1^{2sls} = \frac{\hat{\pi}_1}{\hat{\theta}_1} = \frac{\dfrac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}(z_i - \bar{z})^2}}{\dfrac{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}{\sum_{i=1}^{n}(z_i - \bar{z})^2}} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}$$

which is exactly the same estimate we got above.

To recap, we call the initial equation the "structural equation of interest":

$$y_i = \beta_0 + x_i\beta_1 + \varepsilon_i$$

To estimate this via 2SLS or indirect least squares, we need a "first-stage" relationship

$$x_i = \theta_0 + z_i\theta_1 + u_i$$

The direct relationship between the instrument and the outcome of interest is called the "reduced-form".

$$y_i = \pi_0 + z_i\pi_1 + v_i$$

The variance of $\hat{\beta}_1^{2sls}$

Note that if we were to estimate the regression $y_i = \beta_0 + x_i\beta_1 + \varepsilon_i$ by OLS, the variance for $\hat{\beta}_1^{ols}$ would be $Var(\hat{\beta}_1^{ols}) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$. The variance is proportional to how much variance in x is used to produce the estimate. In the case of 2SLS, we are not use x but instead, $\hat{x}_i$. Therefore $Var(\hat{\beta}_1^{2sls}) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})^2}$.

Recall that in a regression, SST=SSM+SSE and in this case $\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})^2 + \sum_{i=1}^{n}\hat{v}_i^2$. As a result, by construction, $\sum_{i=1}^{n}(x_i - \bar{x})^2 > \sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})^2$ which means also that by construction $Var(\hat{\beta}_1^{ols}) < Var(\hat{\beta}_1^{2sls})$. The benefit of 2SLS is that we are only using a part of the variation in x – that part that is uncorrelated with ε -- so our estimates are consistent. However, because we are using smaller variation in x that in the OLS case, the cost of 2SLS is a reduction in precision – the value $Var(\hat{\beta}_1^{2sls})$ increases considerably.

What will produce smaller values for $Var(\hat{\beta}_1^{2sls})$? Start with the definition $Var(\hat{\beta}_1^{2sls}) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})^2}$ and remember that $\hat{x}_i - \bar{\hat{x}} = (z_i - \bar{z})\hat{\theta}_1$. Substitute this in and we get $Var(\hat{\beta}_1^{2sls}) = \frac{\sigma_\varepsilon^2}{\hat{\theta}_1^2 \sum_{i=1}^{n}(z_i - \bar{z})^2}$. Note that

$\hat{\theta}_1 = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}{\sum_{i=1}^{n}(z_i - \bar{z})^2}$ and substitute this into the variance equation and you get

$$Var(\hat{\beta}_1^{2sls}) = \frac{\sigma_\varepsilon^2}{\left(\frac{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}{\sum_{i=1}^{n}(z_i - \bar{z})^2}\right)^2 \sum_{i=1}^{n}(z_i - \bar{z})^2} = \frac{\sigma_\varepsilon^2 \sum_{i=1}^{n}(z_i - \bar{z})^2}{\left(\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})\right)^2}$$

The denominator in this term $\left(\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})\right)$ is nothing more that (n-1) $\hat{\sigma}_{xz}$ while the numerator can be thought of as $\sum_{i=1}^{n}(z_i - \bar{z})^2 = (n-1)\hat{\sigma}_z^2$. Therefore,

$$Var(\hat{\beta}_1^{2sls}) = \frac{\sigma_\varepsilon^2 \sum_{i=1}^{n}(z_i - \bar{z})^2}{\left(\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})\right)^2} = \frac{\sigma_\varepsilon^2 (n-1)\hat{\sigma}_z^2}{(n-1)^2 \hat{\sigma}_{xz}^2} = \frac{\sigma_\varepsilon^2 \hat{\sigma}_z^2}{(n-1)\hat{\sigma}_{xz}^2}$$

and we will reduce the $Var(\hat{\beta}_1^{2sls})$ is if we can find a variable z that is strongly correlated with x. At the other end of the spectrum, suppose $\hat{\sigma}_{xz}$ approaches zero – this means z does not explain much of x. What is the cost? In this case, if z does not explain much of x, then we cannot learn x's impact on y and the $Var(\hat{\beta}_1^{2sls})$ will explode.