Notes on the Frisch-Waugh-Lovell Theorem

16 March 2024

In this note, we state the FWL Theorem and present a proof.

1 Terminology

When we say "regress y on x", or "regress y on 1 and x", or "run a regression of y on 1 and x", we mean estimating the linear regression model $y = \beta_0 + \beta_1 x + u$. (Here we can view β_0 as $\beta_0 \cdot 1$, so 1 is another independent variable.) In other words, we calculate the OLS estimator $(\hat{\beta}_0, \hat{\beta}_1)$, such that $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{u}_i$.

2 The FWL Theorem

In the multiple linear regression model, we are interested in estimating β_1 from the following equation:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i.$$

There are three methods to estimate β_1:

1. We can use the sample analogue or the least squares idea to calculate the estimator. We denote the estimator as $\hat{\beta}_1$.

2. Now consider the following method:

    - Run a regression of x_1 on 1, x_2, $\cdots$, x_k. This basically means we estimate the following model $x_{i1} = \alpha_0 + \alpha_1 x_{i2} + \cdots + \alpha_{k-1} x_{ik} + r_{i1}$. The OLS estimators are denoted by $\hat{\alpha}_0, \cdots, \hat{\alpha}_{k-1}$, and the residual is denoted by $\hat{r}_{i1}$. We have $x_{i1} = \hat{x}_{i1} + \hat{r}_{i1}$.
    - Regress y on $\hat{r}_{i1}$. This means we estimate the following model $y_i = \delta_0 + \delta_1 \hat{r}_{i1} + \epsilon_i$. The OLS estimators are denoted by $\hat{\delta}_0$ and $\hat{\delta}_1$. Using the slope coefficient formula for the simple linear regression model, we can write out the expression of $\hat{\delta}_1$ as:

    $$\hat{\delta}_1 = \frac{\sum_{i=1}^{N} y_i(\hat{r}_{i1} - \bar{\hat{r}}_{i1})}{\sum_{i=1}^{N}(\hat{r}_{i1} - \bar{\hat{r}}_{i1})^2}.$$

Note that $\bar{\hat{r}}_{i1} = \frac{1}{N}\sum_{i=1}^{N}\hat{r}_{i1} = 0$, because the sum of residuals is zero. We can then simplify the above notation to:

$$\hat{\delta}_1 = \frac{\sum_{i=1}^{N} y_i \hat{r}_{i1}}{\sum_{i=1}^{N} \hat{r}_{i1}^2}. \tag{1}$$

- In the last step, whether we estimate $y_i = \delta_0 + \delta_1 \hat{r}_{i1} + \epsilon_i$ or we estimate $y_i = \delta_1 \hat{r}_{i1} + \epsilon_i$ will result in the same slope estimate $\hat{\delta}_1$. This is because $\frac{1}{N}\sum_{i=1}^{N}\hat{r}_{i1} = 0$.

  Note that in general, when we regress y on x but exclude an intercept, then the OLS estimator $\tilde{\beta}$ will be
  $$\tilde{\beta} = \frac{\sum_{i=1}^{N} y_i x_i}{\sum_{i=1}^{N} x_i^2}.$$
  This can be deduced from the (single) first-order condition for the least squares problem. Therefore, the formula (1) means we can simply run a regression of y_i on $\hat{r}_{i1}$ without the intercept (setting $\delta_0 = 0$ directly).

3. Finally, consider the following method:

- Run a regression of x_1 on $1, x_2, \cdots, x_k$. We get $x_{i1} = \hat{\alpha}_0 + \hat{\alpha}_1 x_{i2} + \cdots + \hat{\alpha}_{k-1} x_{ik} + \hat{r}_{i1}$. This is the same as the first step of bullet point 3.

- Regress y on $1, x_2, \cdots, x_k$. This means we estimate the following model $y_i = \omega_0 + \omega_1 x_{i2} + \cdots + \omega_{k-1} x_{ik} + r_{iy}$. The OLS estimators are denoted by $\hat{\omega}_0, \cdots, \hat{\omega}_{k-1}$ and the residual is denoted by $\hat{r}_{iy}$.

- Regress $\hat{r}_{iy}$ on $\hat{r}_{i1}$. In other words, we estimate the following model $\hat{r}_{iy} = \theta_0 + \theta_1 \hat{r}_{i1} + e_i$. The OLS estimators are denoted by $\hat{\theta}_0$ and $\hat{\theta}_1$. Using the slope coefficient formula for the simple linear regression model, we can write out the expression of $\hat{\theta}_1$ as:
  $$\hat{\theta}_1 = \frac{\sum_{i=1}^{N} \hat{r}_{iy}(\hat{r}_{i1} - \bar{\hat{r}}_{i1})}{\sum_{i=1}^{N} (\hat{r}_{i1} - \bar{\hat{r}}_{i1})^2}.$$
  Again, we have $\bar{\hat{r}}_{i1} = \frac{1}{N}\sum_{i=1}^{N}\hat{r}_{i1} = 0$, because the sum of residuals is zero. We can then simplify the above notation to:
  $$\hat{\theta}_1 = \frac{\sum_{i=1}^{N} \hat{r}_{iy}\hat{r}_{i1}}{\sum_{i=1}^{N} \hat{r}_{i1}^2}.$$

- Again, in the last step, whether or not we include an intercept into regression does not matter. We always get the same $\hat{\theta}_1$.

The Frisch-Waugh-Lovell (FWL) Theorem states that all three methods results in the same estimator.

3 Proof

The proof has two parts. In the first part, we show that $\hat{\delta}_1 = \hat{\theta}_1$, i.e., the last two estimators are equivalent. In the second part, we then show that either is the same as $\hat{\beta}_1$.

3.1 Proof of $\hat{\delta}_1 = \hat{\theta}_1$

We first show that $\hat{\delta}_1 = \hat{\theta}_1$. This is equivalent to proving:

$$\frac{\sum_{i=1}^N y_i \hat{r}_{i1}}{\sum_{i=1}^N \hat{r}_{i1}^2} = \frac{\sum_{i=1}^N \hat{r}_{iy} \hat{r}_{i1}}{\sum_{i=1}^N \hat{r}_{i1}^2}.$$

Note that the denominators are the same. So all we need to prove is that the numerators are the same, or equivalently,

$$\sum_{i=1}^N y_i \hat{r}_{i1} = \sum_{i=1}^N \hat{r}_{iy} \hat{r}_{i1}.$$

We can plug in $y_i = \hat{\omega}_0 + \hat{\omega}_1 x_{i2} + \cdots + \hat{\omega}_{k-1} x_{ik} + \hat{r}_{iy}$:

$$\begin{aligned}
\sum_{i=1}^N y_i \hat{r}_{i1} &= \sum_{i=1}^N (\hat{\omega}_0 + \hat{\omega}_1 x_{i2} + \cdots + \hat{\omega}_{k-1} x_{ik} + \hat{r}_{iy}) \hat{r}_{i1} \\
&= \hat{\omega}_0 \sum_{i=1}^N \hat{r}_{i1} + \hat{\omega}_1 \sum_{i=1}^N x_{i2} \hat{r}_{i1} + \cdots + \hat{\omega}_{k-1} \sum_{i=1}^N x_{ik} \hat{r}_{i1} + \sum_{i=1}^N \hat{r}_{iy} \hat{r}_{i1} \\
&= \sum_{i=1}^N \hat{r}_{iy} \hat{r}_{i1}.
\end{aligned}$$

The last step comes from the fact that the sum of residuals is zero, and the sum of the independent variables times the residuals is zero (These properties are held by the construction of the OLS estimators.)

3.2 Proof of $\hat{\delta}_1 = \hat{\beta}_1$

In the second part, we show $\hat{\delta}_1 = \hat{\beta}_1$. $\hat{\delta}_1$ is from the two-step estimation method. $\hat{\beta}_1$ is from the sample analogue or least squares method when there are two independent variables. One way to prove is to derive explicitly the expression for $\hat{\beta}_1$, and then prove they are the same. We demonstrate this method when there are two independent variables in the solution of problem set 1. Another way is to use the properties of the OLS estimator, which we consider now.

Here is the sketch of the proof:

- Start with the condition that $\sum_{i=1}^N x_{i1} \hat{u}_i = 0$, where $\hat{u}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - ... - \hat{\beta}_k x_{ik}$. This is the first-order condition we derived for the OLS estimator, so it holds for $\hat{\beta}_1$.

- We also know that $x_{i1} = \hat{\alpha}_0 + \hat{\alpha}_1 x_{i2} + ... + \hat{\alpha}_{k-1} x_{ik} + \hat{r}_{i1} = \hat{x}_{i1} + \hat{r}_{i1}$.

- First, plug $x_{i1} = \hat{\alpha}_0 + \hat{\alpha}_1 x_{i2} + ... + \hat{\alpha}_{k-1} x_{ik} + \hat{r}_{i1}$ in $\sum_{i=1}^N x_{i1}\hat{u}_i = 0$. Simplify the equation. We have:

$$\begin{aligned} 0 &= \sum_{i=1}^N x_{i1}\hat{u}_i \\ &= \sum_{i=1}^N (\hat{\alpha}_0 + \hat{\alpha}_1 x_{i2} + ... + \hat{\alpha}_{k-1} x_{ik} + \hat{r}_{i1})\hat{u}_i \\ &= \sum_{i=1}^N \hat{r}_{i1}\hat{u}_i. \end{aligned}$$

  Note that $\sum_{i=1}^N \hat{\alpha}_0\hat{u}_i = \hat{\alpha}_0 \sum_{i=1}^N \hat{u}_i = 0$ because the sum of residuals is zero, and $\sum_{i=1}^N \hat{\alpha}_{j-1} x_{ij}\hat{u}_i = \hat{\alpha}_{j-1} \sum_{i=1}^N x_{ij}\hat{u}_i = 0, \forall j = 2, ..., k$ because the sum of the independent variable and the residuals is zero.

- Then, plug $\hat{u}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - ... - \hat{\beta}_k x_{ik}$ in the simplified term.

$$\begin{aligned} 0 &= \sum_{i=1}^N \hat{r}_{i1}\hat{u}_i \\ &= \sum_{i=1}^N \hat{r}_{i1}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - ... - \hat{\beta}_k x_{ik}) \\ &= \sum_{i=1}^N (\hat{r}_{i1} y_i - \hat{r}_{i1}\hat{\beta}_1 x_{i1}). \end{aligned}$$

  Note that $\sum_{i=1}^N \hat{r}_{i1}\hat{\beta}_0 = \hat{\beta}_0 \sum_{i=1}^N \hat{r}_{i1} = 0$, because the sum of residuals is zero, and $\sum_{i=1}^N \hat{r}_{i1} x_{ij} = 0, \forall j = 2, 3, ..., k$, because the sum of the independent variable and the residuals is zero.

- After rearranging things, we find that

$$\hat{\beta}_1 = \frac{\sum_{i=1}^N y_i \hat{r}_{i1}}{\sum_{i=1}^N \hat{r}_{i1} x_{i1}}.$$

  Note the denominator, $\sum_{i=1}^N \hat{r}_{i1} x_{i1} = \sum_{i=1}^N \hat{r}_{i1}(\hat{x}_{i1} + \hat{r}_{i1}) = \sum_{i=1}^N \hat{r}_{i1}^2 + \sum_{i=1}^N \hat{r}_{i1}\hat{x}_{i1} = \sum_{i=1}^N \hat{r}_{i1}^2$, where $\sum_{i=1}^N \hat{r}_{i1}\hat{x}_{i1} = 0$ because the sum of residuals times the fitted value is zero. Finally, we have

$$\hat{\beta}_1 = \frac{\sum_{i=1}^N y_i \hat{r}_{i1}}{\sum_{i=1}^N \hat{r}_{i1}^2}.$$

  The right hand side is exactly the expression for $\hat{\delta}_1$, so $\hat{\delta}_1 = \hat{\beta}_1$.