Multiple variable regression model

1. Introduction

Consider Figure 1a, which plots *e* (earnings) against *s* (school) and as can be seen there is no apparent relationship between earnings and schooling. Estimating a 2-variable model of earnings on schooling by OLS yields an estimate of a slope coefficient of 0.01, with an estimated standard error on this coefficient of 0.016 (suggesting that an extra year of schooling increases does not increase earnings at all) – why is this the case when we all know that extra schooling increases earnings? In actual fact what is unclear from Figure 1a is there are two types of individuals in the sample (males and females). In Figure 1b, we replot Figure 1a, but identifying males (diamonds) and females (crosses) separately and here there is clearly a strong positive relationship between earnings and schooling for the two groups separately, The problem was that in Figure 1a we did not take into account the gender of the person. What we really want to know is the effect of an additional year of schooling, for a given gender.

Looking only at the diamonds (males) – we get can get an estimate of the effect of an increase in the schooling (*s*), given the individual is male and this is represented by the thick solid line (estimate = 0.010). Looking at the crosses (females) – we can get an estimate of the effect of an increase in schooling (*s*), given the individual is female and this is represented by the thin solid line (estimate = 0.11) – so why are we getting an overall estimate of 0.01? This answer is that in this case females have more schooling than males, but females are generally paid less than males in the same job – so when we look at a simply plot of two variables the higher schooled individuals are receiving lower pay. Therefore we need to potentially look at including the influence of all variables that might affect earnings in order to get the “true” effect of schooling on earnings.

We can do this analysis more formally we regress earnings on an intercept and gender and take the residuals – these residuals are then that part of wages not explained by gender differences. Similarly regress schooling on an intercept and gender and take the residuals – these residuals are that part of schooling not explained by gender differences. Now if we regress the wage residuals on the school residuals then we are explicitly estimating the effect of schooling on wages having extracted (holding constant) the effect of gender. This process of estimating effects having extracted the influence of other variables is known as ***partitioned regression.***

2. Deriving the Multiple-Variable OLS Estimators

We now extend the analysis in the previous handout to the more realistic case in which we believe the random variable of interest, y, depends upon a number of different factors, x_{1i}, x_{2i}, ..., x_{ki}, and these are exogenous variables. The true but unknown relationship is defined as being

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{1}$$

where we maintain the Classical Linear Regression Model (CLRM) assumptions of the previous handout:

1 $E(\varepsilon_i \mid x_{1i}, x_{2i}, \ldots x_{ki}) = E(\varepsilon_i) = 0$.

2 $Var(\varepsilon_i \mid x_{1i}, x_{2i}, \ldots x_{ki}) = \sigma^2$

3 $\text{cov}(\varepsilon_i, \varepsilon_j \mid x_{1i}, x_{2i}, \ldots x_{ki}) = 0 \quad i \neq j$

4. $\varepsilon_i \sim N(0, \sigma^2)$

As before for some values $(a, b_1, \ldots, b_k)$ of the unknown coefficients $(\alpha, \beta_1, \ldots, \beta_k)$ define the Residual Sum of Squares (*RSS*) as

$$RSS = \sum_{i=1}^{n} (y_i - a - b_1 x_{1i} - b_2 x_{2i} - \ldots - b_k x_{ki})^2 \tag{2}$$

and we wish to find those values of a, b_1, b_2, ..,b_k which minimise this expression. Consequently, we must differentiate the *RSS* with respect to each of the parameters, a, b_1, b_2, ..,b_k in turn and set the resultant derivatives equal to zero.

Differentiating equation (2) with respect to a and setting the expression to zero

$$\frac{\partial (\sum_{i=1}^{n} (y_i - a - b_1 x_{1i} - b_2 x_{2i} - \ldots - b_k x_{ki})^2}{\partial a} = 0 \tag{3a}$$

$$= -2 \sum_{i=1}^{n} (y_i - a - b_1 x_{1i} - b_2 x_{2i} - \ldots - b_k x_{ki}) = -2 \sum_{i=1}^{n} e_i = 0 \Rightarrow \sum_{i=1}^{n} e_i = 0 \tag{4a}$$

And this is the same expression as equation (3) from Handout 3.

Differentiating equation (2) with respect to b_1 and setting the expression to zero

$$\frac{\partial (\sum_{i=1}^{n} (y_i - a - b_1 x_{1i} - b_2 x_{2i} - \ldots - b_k x_{ki})^2}{\partial b_1} = 0 \tag{3b}$$

$$= -2 \sum_{i=1}^{n} x_{1i} (y_i - a - b_1 x_{1i} - b_2 x_{2i} - \ldots - b_k x_{ki}) = -2 \sum_{t=1}^{n} x_{1i} e_i = 0 \Rightarrow \sum_{t=1}^{n} x_{1i} e_i = 0 \tag{4b}$$

And this is similar to the expression in equation (5) in Handout 3.

Differentiating equation (2) with respect to b_2 and setting the expression to zero

$$\frac{\partial(\sum_{i=1}^{n}(y_i - a - b_1x_{1i} - b_2x_{2i} - \ldots - b_kx_{ki})^2}{\partial b_2} = 0 \qquad (3c)$$

$$= -2\sum_{i=1}^{n} x_{2i}(y_i - a - b_1x_{1i} - b_2x_{2i} - \ldots - b_kx_{ki}) = -2\sum_{t=1}^{n} x_{2i}e_i = 0 \Rightarrow \sum_{t=1}^{n} x_{2i}e_i = 0 \qquad (4c)$$

Finally differentiating equation (2) with respect to b_k and setting the expression to zero

$$\frac{\partial(\sum_{i=1}^{n}(y_i - a - b_1x_{1i} - b_2x_{2i} - \ldots - b_kx_{ki})^2}{\partial b_k} = 0 \qquad (3d)$$

$$= -2\sum_{i=1}^{n} x_{ki}(y_i - a - b_1x_{1i} - b_2x_{2i} - \ldots - b_kx_{ki}) = -2\sum_{t=1}^{n} x_{ki}e_i = 0 \Rightarrow \sum_{t=1}^{n} x_{ki}e_i = 0 \qquad (4d)$$

NOTE

(i) Equation (4a) implies that the residuals always sum to zero (providing there is an intercept in the model)

(ii) Equation (4b) implies that the covariance (and hence correlation) between the residuals and the explanatory variable x_1 is zero, that is, they are orthogonal.

(iii) Equation (4c) implies that the covariance (and hence correlation) between the residuals and the explanatory variable x_2 is zero, that is, they are orthogonal.

(iv) Equation (4d) implies that the covariance (and hence correlation) between the residuals and the explanatory variable x_k is zero, that is, they are orthogonal.

The degrees of freedom (*DoF*) for this model is the number of observation, *n*, less the number of restrictions imposed on the residuals, which are:

$$\overbrace{\sum_{i=1}^{n} e_i = 0,\quad \sum_{i=1}^{n} x_{1i}e_i = 0,\quad \sum_{i=1}^{n} x_{2i}e_i = 0,\quad \sum_{i=1}^{n} x_{3i}e_i = 0\ldots,\quad \sum_{i=1}^{n} x_{ki}e_i = 0}^{k+1}$$

That is, *DoF* is, (*n*-*k*-1).

Beyond the two-variable case it is not easy to write out an explicit formula for the **Ordinary Least Squares (OLS)** estimators, $a, b_1, b_2, \ldots, b_k$ (without using matrix algebra) - although solving the (k+1) equations (4a), (4b), (4c) and (4d) simultaneously yields your OLS estimates.

One can use the idea of partition regression in order to get an expression for these estimates, suppose, for example, one wants the OLS coefficient estimate on x_1. Then we can obtain an expression for the OLS estimate of β_1 as:

1. Run the regression $y_i = \delta_0 + \delta_2 x_{2i} + \ldots + \delta_k x_{ki} + \varepsilon_{yi}$ by OLS and save the OLS residuals $\tilde{y}_i = y_i - (d_0 + d_2 x_{2i} + \ldots + d_k x_{ki})$
2. Run the regression $x_{1i} = \gamma_0 + \gamma_2 x_{2i} + \ldots + \gamma_k x_{ki} + \varepsilon_{x_1 i}$, by OLS and save the OLS residuals $\tilde{x}_{1i} = y_i - (\mathrm{g}_0 + g_2 x_{2i} + \ldots + g_k x_{ki})$
3. Run the regression: $e_{yi} = \alpha + \beta_1 e_{x_1 i} + \varepsilon_i$ then as this is a 2-variable regression we know that $b_1 = \frac{\sum_{i=1}^{n}(\tilde{y}_i - \bar{\tilde{y}})(\tilde{x}_{1i} - \bar{\tilde{x}}_1)}{\sum_{i=1}^{n}(\tilde{x}_{1i} - \bar{\tilde{x}}_1)^2} = \frac{\sum_{i=1}^{n}\tilde{y}_i \tilde{x}_{1i}}{\sum_{i=1}^{n}\tilde{x}_{1i}^2}$ as $\bar{\tilde{y}} = \bar{\tilde{x}}_1 = 0$. $V(b_1) = \frac{\sigma^2}{\sum_{i=1}^{n}\tilde{x}_{1i}^2}$ where $\sum_{i=1}^{n}\tilde{x}_{1i}^2$ is the variation in x_1, not accounted for by the other variables i.e. $(1-R_{x_1}^2)\sum_{i=1}(x_{1i} - \bar{x}_1)^2 = \sum_{i=1}\tilde{x}_{1i}{}^2$, where $R_{x_1}^2$ is the usual R^2 from a regression of x_1 on $x_2, x_3 \ldots x_k$ and so $\sum_{i=1}(x_{1i} - \bar{x}_1)^2 \geq \sum_{i=1}(\tilde{x}_{1i} - \bar{\tilde{x}}_1)^2$. The estimator for σ^2 (the variance of the disturbance term, ε_i) is estimated as:

$$s^2 = \frac{\sum_{i=1}^{n} e_i^2}{DoF} = \frac{RSS}{DoF} \qquad (5)$$

where *DoF* are the degrees of freedom for the residuals.

In general, the variance-covariance matrix for the OLS coefficient estimators would be:

$$\begin{bmatrix} V(a) & \mathrm{Cov}(a,b_1) & \ldots & \mathrm{Cov}(a,b_k) \\ \mathrm{Cov}(a,b_1) & V(b_1) & \ldots & \mathrm{Cov}(b_1,b_k) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{Cov}(a,b_k) & \mathrm{Cov}(b_1,b_k) & \ldots & V(b_k) \end{bmatrix}$$

3. Properties of OLS estimators

These **OLS** estimators have a number of desirable properties (Gauss Markov Theorem) in that the estimators are **BLUE**:

(i) **Best** - have the minimum variance, such that $V(b_j) \leq V(b_j^*),\ j = 1, \ldots, k$, where b_j^* is any alternative unbiased estimator.

(ii) **Linear –** linear function of the error term.

(iii) **Unbiased** - $E(a) = \alpha$ and $E(b_j) = \beta_j,\ j = 1 \ldots k$

(iv) **Estimators**.

We will not prove these conditions, but the proof of unbiasedness is similar to that in Handout 2. Based on this we could show that:

$b_j \sim N(\beta_j, V(b_j))$ and $\dfrac{DoF\ s^2}{\sigma^2} \sim \chi^2_{DoF}$.

4. Interpreting coefficients

In the regression model

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i, \quad i = 1, 2, \ldots, n$$

we can interpret the coefficients by partially differentiating the dependent variable with respect to each explanatory variable

$$\frac{\partial y_i}{\partial x_{1i}} = \beta_1 = \frac{\text{Expected change in y}_i}{1 \uparrow \text{increase in } x_1}\bigg|_{\text{Holding all else constant}}$$

We can also think of examples where the model is clearly nonlinear in the underlying variables, but is still linear in parameters enabling us to use OLS:

1. $y_i = \alpha + \beta_1 \ln(x_{1i}) + \beta_1 x_{2i} + \ldots \beta_k x_{ki} + \varepsilon_i$

$$\frac{\partial y_i}{\partial \ln(x_{1i})} = 0.01\beta_1 = \frac{\text{Expected change in y}_i}{1\% \uparrow \text{ in } x_{1i}}\bigg|_{\text{Holding all else constant}}$$

2. $\ln(y_i) = \alpha + \beta_1 x_{1i} + \beta_1 x_{2i} + \ldots \beta_k x_{ki} + \varepsilon_i$

$$\frac{\partial \ln(y_i)}{\partial x_{1i}} = 100\beta_1 = \frac{\text{Expected \% change in y}_i}{1 \uparrow \text{ in } x_{1i}}\bigg|_{\text{Holding all else constant}}$$

3. $\ln(y_i) = \alpha + \beta_1 \ln(x_{1i}) + \beta_1 x_{2i} + \ldots \beta_k x_{ki} + \varepsilon_i$

$$\frac{\partial \ln(y_i)}{\partial \ln(x_{1i})} = \beta_1 = \frac{\text{\% change in y}_i}{1\% \uparrow \text{ in } x_{1i}}\bigg|_{\text{Holding all else constant}}$$

4. $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{1i}^2 + \beta_3 x_{3i} \ldots + \beta_k x_{ki} + \varepsilon_i$ (and the relationship is quadratic x_{1i})

$$\frac{\partial y_i}{\partial x_{1i}} = \beta_1 + 2\beta_2 x_{1i} = \frac{\text{Expected change in } y_i}{1\uparrow \text{ increase in } x_1}\Bigg|_{\text{Holding all else constant}}$$ (response varies linearly in x_{1i}).

5. $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{1i} z_i + \beta_3 x_{3i} \ldots + \beta_k x_{ki} + \varepsilon_i$

$$\frac{\partial y_i}{\partial x_{1i}} = \beta_1 + \beta_2 z_i = \frac{\text{Expected change in } y_i}{1\uparrow \text{ increase in } x_1}\Bigg|_{\text{Holding all else constant}}$$ (response varies linearly in z_i).

(For more examples on interpreting coefficients see handout 4).

5. Explanatory Power of a Multiple Regression Equation

For the regression model

$$y_i = a + b_1 x_{1i} + b_2 x_{2i} + \ldots + b_k x_{ki} + e_i$$

where $a, b_j \; j = 1, \ldots, k$ are the OLS parameter estimates and e_i the residuals, then the fitted Predicted) values for y_i are written as

$$\hat{y}_i = a + b_1 x_{1i} + b_2 x_{2i} + \ldots + b_k x_{ki}.$$

Therefore

$$y_i = \hat{y}_i + e_i$$

subtracting $\bar{y}$ from both sides of this equation we have:

$$(y_i - \bar{y}) = (\hat{y}_i - \bar{y}) + e_i$$

Squaring this equation we have

$$(y_i - \bar{y})^2 = (\hat{y}_i - \bar{y})^2 + e_i^2 + 2(\hat{y}_i - \bar{y})e_i$$

Taking sums we write

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n} e_i^2 + 2\sum_{i=1}^{n}(\hat{y}_i - \bar{y})e_i$$

But the last term is zero as

$$\sum_{i=1}(\hat{y}_i - \bar{y})e_i = b_1 \sum_{i=1}(x_{1i} - \bar{x}_1)e_i + b_2 \sum_{i=1}(x_{2i} - \bar{x}_2)e_i + \ldots + b_k \sum_{i=1}(x_{ki} - \bar{x}_k)e_i$$

and we know $\operatorname{cov}(x_{ji}, e_t) = \sum_t (x_{ji} - \bar{x}_j)e_i = 0$ (from equations 4b, 4c, 4d)

Therefore

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n} e_i^2 \Rightarrow TSS = ESS + RSS \qquad (6)$$

and hence

$$0 \le R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \le 1 \qquad (7)$$

However, R^2 suffers from the problem that as you add more variables into the equation R^2 will never decrease (as *RSS* must fall or remain unchanged as more variables are added into the model) and therefore a policy of maximising R^2 will entail over-parameterising the model.

Consequently, people sometime use other criteria for choosing models:

$$\bar{R}^2 = 1 - \frac{RSS / DoF}{TSS / (n-1)}$$

in this case, while *RSS* will fall as more variables are added, *DOF* will also fall as more variables are added: so $\bar{R}^2$ may either increase or decrease. $\bar{R}^2$ will only increase if an F-test on the joint significance on the new variables is in excess of unity.

6. Hypothesis testing

6.1 The t-test

We follow a 5-step procedure:

1. $H_0 : \beta_j = \beta_{j0}$
2. $H_0 : \beta_j \ne \beta_{j0}$
3. Choose some appropriate significance level of α, and find the corresponding value from the t-distribution, denoted $-t_{\alpha/2,DoF}$ and $t_{\alpha/2,DoF}$, where *DoF* are the degree of freedom from the model, that is, *n* minus the number of restrictions on the residuals, that is, $n-(k+1)$.
4. $t = \dfrac{b_j - \beta_{j0}}{s_{jb}} \sim t_{\alpha/2,DoF}$, where s_b = standard error of b_j (in replacing σ^2 by s^2 in the standard error of b_j, we are essentially scaling a N(0,1) by a $\sqrt{\chi^2_{DoF} / DoF}$ and this yields a t_{DoF} distribution.
5. If t is either less than $-t_{\alpha/2,DoF}$, or greater than $t_{\alpha/2,DoF}$, then we have observed an event which occurs with a probability of less than α and should therefore reject H_0. The

decision rule is: Reject H_0 if $t = \left|\frac{b_j - \beta_{j0}}{s_{jb}}\right| > t_{\alpha/2,DoF}$; Do not reject H_0 if

$t = \left|\frac{b_j - \beta_{j0}}{s_{jb}}\right| < t_{\alpha/2,DoF}$.

Consider now the more complex problem where we wish to test the null hypothesis $H_0 : \beta_j - \beta_l + 2\beta_m = D_0$, we again follow the 5-step procedure outlined above, but in this case the actual test statistics is:

$$t = \frac{b_j - b_l + 2b_m - D_0}{\sqrt{V(b_j) + V(b_l) + 4V(b_m) - 2\operatorname{cov}(b_j, b_l) + 4\operatorname{cov}(b_j, b_m) - 4\operatorname{cov}(b_l, b_m)}} \sim t_{\alpha/2,DoF}.$$

6.2 The F-test

We can also test multiple restrictions:

1. $H_0 : \beta_1 = \beta_1^0, \ldots, \beta_d = \beta_d^0$
2. $H_0 : \beta_j \neq \beta_j^0,\ j = 1, \ldots, d$
3. Choose some appropriate significance level of α, and find the corresponding value from the F-distribution, denoted $F_{d,DoF}^{\alpha}$.
4. $$F = \frac{(RSS^R - RSS^U)/d}{RSS^u / DoF} \sim F_{d,DoF}^{\alpha} \qquad (8)$$

   where $RSS^R = \sum_{i=1} (y_i - a - b_1^0 x_{1i} \ldots - b_d^0 x_{di} - b_{d+1} x_{d+1i} \ldots - b_k x_{ki})^2$, is the RSS from the restricted model, i.e. the model evaluated at the restricted parameters (under H_0), and $RSS^U = \sum_{i=1} (y_i - a - b_1 x_{1i} \ldots - b_d x_{di} \ldots - b_k x_{ki})^2$, is the *RSS* from the model which imposes no restrictions on the parameter estimates, *d* is the number of restrictions under the null hypothesis, i.e. the number of equality signs in H_0 (and MUST also be the difference in the number of parameters to be estimated in the restricted and unrestricted regression models), *DoF* is the degrees of freedom associated with the unrestricted regression model (and is *n* minus the number of parameters estimated in the unrestricted model (*k*+1)).

5. If F is greater than $F^{\alpha}_{d,DoF}$, then we have observed an event which occurs with a probability of less than α and should therefore reject H_0. The decision rule is: Reject H_0 if $F > F^{\alpha}_{d,DoF}$; Do not reject H_0 if $F < F^{\alpha}_{d,DoF}$.

<u>NOTE</u>

Providing the dependent variable is the same in the restricted and unrestricted equations then one can always construct the F-test in terms of the R^2 from the restricted and unrestricted models.

$$F = \frac{(RSS^R - RSS^U)/d}{RSS^u / DoF} = \frac{\left(\frac{RSS^R}{TSS} - \frac{RSS^U}{TSS}\right)/d}{\frac{RSS^u}{TSS}/DoF}$$

$$F = \frac{((1-R_R^2)-(1-R_U^2))/d}{(1-R_U^2)/DoF} = \frac{(R_U^2 - R_R^2)/d}{(1-R_U^2)/DoF} \qquad (9)$$

and so equations (8) and (9) must yield exactly the same value for the F-test.

As an example where equation (9) is not the appropriate form for undertaking an F-test, consider the unrestricted model:

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \varepsilon_i, \quad i = 1,2,\ldots,n \qquad (10a)$$

then for testing $H_0: \beta_2 = 0.5, \beta_3 = 0$, the restricted model is:

$$y_i = \alpha + \beta_1 x_{1i} + 0.5 x_{2i} + \beta_4 x_{4i} + \varepsilon_i, \quad i = 1,2,\ldots,n$$

However, the only way of imposing a known (non-zero) coefficient on any variable is to take the variable over to the left hand side as coefficients are only estimated on variables on the right hand side. Consequently, the restricted equation is:

$$y_i - 0.5 x_{2i} = \alpha + \beta_1 x_{1i} + \beta_4 x_{4i} + \varepsilon_i, \quad i = 1,2,\ldots,n \qquad (10b)$$

The restricted equation (10b) has a different dependent variable to the unrestricted equation (10a), in which case, while one can test the null hypothesis on the basis of the F-test in equation (8), an F-test based on equation (9) will not be correct.

6.3 F-test of overall significance

1. $H_0: \beta_1 = \beta_2 = \ldots = \beta_k = 0$
2. $H_1:$ Any $\beta_j \neq 0$

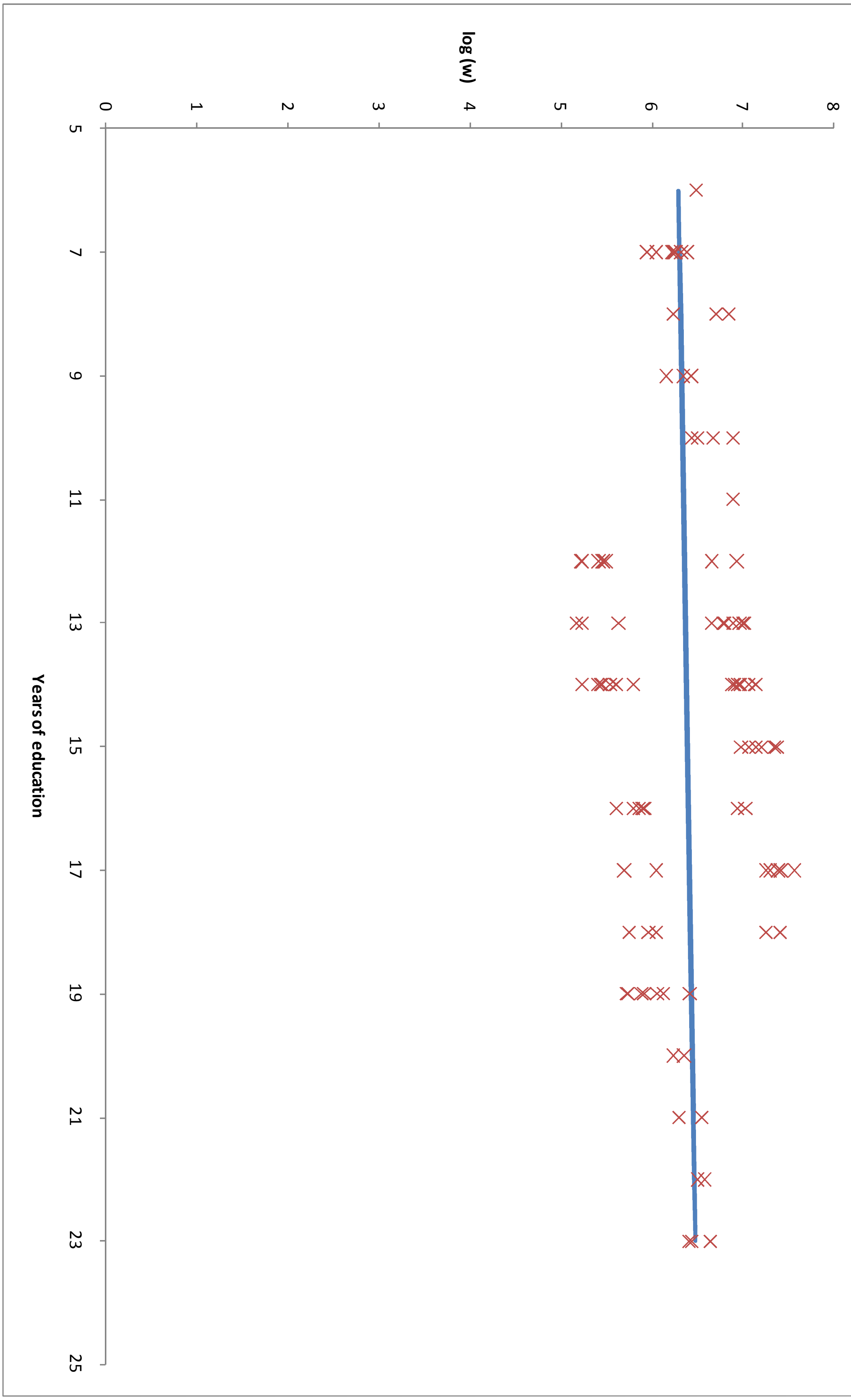
3. Choose some appropriate significance level of α , and find the corresponding value from the F-distribution, denoted $F^{\alpha}_{k,DoF}$.

4. $F = \dfrac{(RSS^R - RSS^U)/k}{RSS^u / DoF} \sim F^{\alpha}_{k,DoF}$, where $RSS^R = \sum_{i=1}(y_i - a)^2$, (so $RSS^R = TSS$) and $RSS^U = \sum_{i=1}^{n}(y_i - a - b_1 x_{1i} \ldots - b_k x_{ki})^2$ **,** k is the number of coefficients excluding the intercept and DoF is n-(k+1)**.** Therefore, in this case, $F = \dfrac{(TSS - RSS^U)/k}{RSS^u / DoF}$ and dividing top and bottom by TSS we get $F = \dfrac{\left(1 - \dfrac{RSS^U}{TSS}\right)/k}{\left(\dfrac{RSS^u}{TSS}\right)/DoF} = \dfrac{R_U^2 / k}{(1 - R_U^2)/DoF}$.

5. If F is greater than $F^{\alpha}_{k,DoF}$, then we have observed an event which occurs with a probability of less than α and should therefore reject H_0. The decision rule is: Reject H_0 if $F > F^{\alpha}_{k,DoF}$; Do not reject H_0 if $F < F^{\alpha}_{k,DoF}$.

<u>Note</u>: Appendix 1 has a labelled Stata output for a multiple regression.

Figure 1a. Cross plot of log earnings against schooling



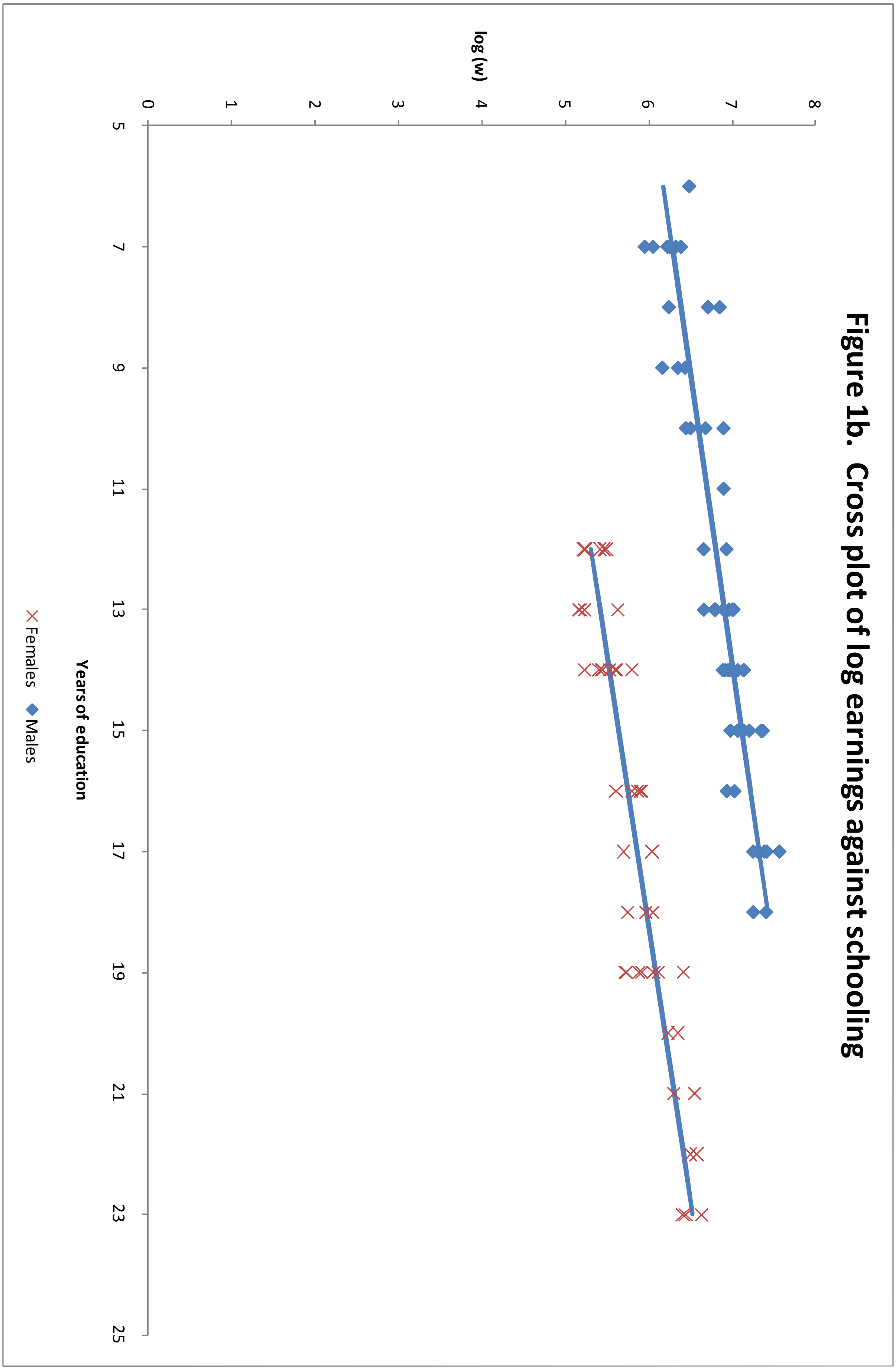


Figure 1b. Cross plot of log earnings against schooling

Appendix 1: Stata output for multiple regression

reg lhourpay age agesq edage if(edage>=0 & edage<=95)

$RSS = \sum_{i=1}^{7334} e_i^2$

ESS

$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$

DoF=n-4

Pr(F>695.10)

s^2=*RSS/DoF*

$TSS = \sum_{i=1}^{7334} (y_i - \bar{y})^2$

```
      Source |       SS       df       MS              Number of obs =    7334
-------------+------------------------------           F(  3,  7330) =  695.10
       Model |  482.463299     3  160.8211             Prob > F      =  0.0000
    Residual |  1695.90095  7330  .23136486            R-squared     =  0.2215
-------------+------------------------------           Adj R-squared =  0.2212
       Total |  2178.36425  7333  .297063173           Root MSE      =    .481

------------------------------------------------------------------------------
    lhourpay |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |    .069112   .0028355    24.37   0.000     .0635535    .0746705
       agesq |  -.0007238   .0000338   -21.41   0.000    -.0007901   -.0006575
       edage |   .0782183   .0021262    36.79   0.000     .0740503    .0823864
       _cons |  -.7217424   .0676627   -10.67   0.000    -.8543808   -.589104
------------------------------------------------------------------------------
```

n

$\bar{R}^2$

Explanatory variables

intercept

$se(b_1)$

$R^2 = 1 - RSS / TSS$

0.069112±1.96×0.0028355

$t = \frac{0.069112-0}{0.0028355}$

b_1=Expected proportionate increase in wages for an additional year of schooling for a given age.

Pr(t>24.37)=0.000