1. $Y = X_1\hat\beta_1 + X_2\hat\beta_2 + \hat\varepsilon$, then we know that: $X_1^T\hat\varepsilon = 0$, $X_2^T\hat\varepsilon = 0$, $(I_n - H_1)\hat\varepsilon = \hat\varepsilon$

$\Rightarrow (I_n - H_1)Y = \underbrace{(I_n - H_1)X_1}_{0}\hat\beta_1 + (I_n - H_1)X_2\hat\beta_2 + (I_n - H_1)\hat\varepsilon$

$\Rightarrow (I_n - H_1)Y = (I_n - H_1)X_2\hat\beta_2 + \hat\varepsilon$

$\Rightarrow X_2^T(I_n - H_1)Y = X_2^T(I_n - H_1)X_2\hat\beta_2 + X_2^T\hat\varepsilon = X_2^T(I_n - H_1)X_2\hat\beta_2$

$X_2^T(I_n - H_1)X_2 = [(I_n - H_1)X_2]^T[(I_n - H_1)X_2]$

Because $\hat\beta_2$ exists and is unique, which mean $[X_2^T(I_n - H_1)X_2]b = X_2^T[I_n - H_1)Y$ have unique solution, then we know $X_2^T(I_n - H_1)X_2$ is full rank.

Therefore $X_2^T(I_n - H_1)X_2$ is Invertible.

$\Rightarrow \hat\beta_2 = [X_2^T(I_n - H_1)X_2]^{-1}X_2^T(I_n - H_1)Y = (\tilde X_2^T\tilde X_2)^{-1}\tilde X_2^T Y$

2 $Y = x_1\beta_1 + x_2\beta_2 + \cdots + x_p\beta_p + \varepsilon$, $X = \left[\begin{array}{c|c|c|c} x_1 & x_2 & \cdots & x_p \end{array}\right]$

① First we compute OLS fit of y on x_1

we get: $\hat\beta_1 = y^T x_1 / x_1^T x_1$

②. Then we compute OLS fit of y on $\tilde X_2 := (I_n - H_{(1)})X_2$, where $H_{(1)} = X_1(X_1'X_1)^{-1}X_1'$

we get: $\hat\beta_2 = y^T\tilde X_2 / \tilde X_2^T\tilde X_2 = y^T(I_n - H_{(1)})X_2 / X_2'(I_n - H_{(1)})X_2$

... ...

③ Then we compute OLS fit of y on $\tilde X_k := (I_n - H_{(k-1)})X_k$,

where $H_{(k-1)} = X_{(k-1)}(X_{(k-1)}'X_{(k-1)})^{-1}X_{(k-1)}'$, $X_{(k-1)} = \left[\begin{array}{c|c|c|c} X_1 & X_2 & \cdots & X_{k-1} \end{array}\right]$

we get: $\hat\beta_k = y^T\tilde X_k / \tilde X_k^T\tilde X_k = y^T(I_n - H_{(k-1)})X_k / X_k'(I_n - H_{(k-1)})X_k$

3 $\hat{cov}(\hat{\varepsilon}_y, \hat{\varepsilon}_x) = \frac{1}{n}\sum_{i=1}^{n} \hat{\varepsilon}_y \hat{\varepsilon}_x = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{\beta}_{y|w} w_i)(x_i - \hat{\beta}_{x|w} w_i)$

$= \frac{1}{n}\sum_{i=1}^{n} y_i (x_i - \hat{\beta}_{x|w} w_i) - \frac{1}{n}\sum_{i=1}^{n} \hat{\beta}_{y|w} w_i (x_i - \hat{\beta}_{x|w} w_i)$

since $\sum_{i=1}^{n} w_i (x_i - \hat{\beta}_{x|w} w_i) = 0$, then it equals to:

$= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \hat{\beta}_{x|w} \sum_{i=1}^{n} y_i w_i$, since $n\hat{\beta}_{x|w}\bar{w}\bar{y} = n\bar{x}\bar{y}$, then we get:

$= \frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}\right) - \hat{\beta}_{x|w}\left(\sum_{i=1}^{n} y_i w_i - n\bar{y}\bar{w}\right)$

$= \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) - \hat{\beta}_{x|w} \sum_{i=1}^{n} (y_i - \bar{y})(w_i - \bar{w})$

$= \hat{\rho}_{xy}\hat{\sigma}_x\hat{\sigma}_y - \hat{\rho}_{xw}\frac{\hat{\sigma}_x}{\hat{\sigma}_w}\hat{\rho}_{yw}\hat{\sigma}_y\hat{\sigma}_w = \hat{\rho}_{xy}\hat{\sigma}_x\hat{\sigma}_y - \hat{\rho}_{xw}\hat{\rho}_{yw}\hat{\sigma}_x\hat{\sigma}_y$

$\hat{var}(\hat{\varepsilon}_y) = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{\beta}_{y|w} w_i)^2 = \frac{1}{n}\sum_{i=1}^{n} y_i (y_i - \hat{\beta}_{y|w} w_i)$

$= \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\sum_{i=1}^{n} y_i w_i \cdot \hat{\beta}_{y|w} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^2 - \frac{1}{n}\hat{\beta}_{y|w}\sum_{i=1}^{n}(y_i - \bar{y})(w_i - \bar{w})$

$= \hat{\sigma}_y^2 - \hat{\rho}_{yw}^2 \frac{\hat{\sigma}_y}{\hat{\sigma}_w}\hat{\sigma}_y\hat{\sigma}_w = \hat{\sigma}_y^2 - \hat{\rho}_{yw}^2\sigma_y^2 = \hat{\sigma}_y^2(1 - \hat{\rho}_{yw}^2)$, $\hat{var}(\hat{\varepsilon}_x) = \hat{\sigma}_x^2(1 - \hat{\rho}_{xw}^2)$

$$\therefore \hat{\rho}_{yx|w} = \frac{\hat{cov}(\hat{\varepsilon}_y, \hat{\varepsilon}_x)}{\sqrt{\hat{var}(\hat{\varepsilon}_x)\cdot \hat{var}(\hat{\varepsilon}_y)}} = \frac{\cancel{\hat{\sigma}_x}\cancel{\hat{\sigma}_y}(\hat{\rho}_{xy} - \hat{\rho}_{xw}\hat{\rho}_{yw})}{\cancel{\hat{\sigma}_x}\cancel{\hat{\sigma}_y}\sqrt{(1-\hat{\rho}_{yw}^2)(1-\hat{\rho}_{xw}^2)}} = \frac{\hat{\rho}_{xy} - \hat{\rho}_{xw}\hat{\rho}_{yw}}{\sqrt{1-\hat{\rho}_{yw}^2}\sqrt{1-\hat{\rho}_{xw}^2}}$$

4. $H = X(X'X)^{-1}X' = (X_1 \ \ X_2)\begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1}\begin{pmatrix} X_1' \\ X_2' \end{pmatrix}$

We know the inverse of block matrix is:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D-CA^{-1}B)^{-1} \\ -(D-CA^{-1}B)^{-1}CA^{-1} & (D-CA^{-1}B)^{-1} \end{pmatrix}$$

$D - CA^{-1}B = X_2'X_2 - X_2'X_1(X_1'X_1)^{-1}X_1'X_2 = X_2'(I_n - H_1)X_2 = \tilde{X}_2'\tilde{X}_2$

$A^{-1} + A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1} = (X_1'X_1)^{-1} + (X_1'X_1)^{-1}X_1'X_2(\tilde{X}_2'\tilde{X}_2)^{-1}X_2'X_1(X_1'X_1)^{-1}$

$A^{-1}B(D-CA^{-1}B)^{-1} = (X_1'X_1)^{-1}X_1'X_2(\tilde{X}_2'\tilde{X}_2)^{-1}$

$(D-CA^{-1}B)^{-1}CA^{-1} = (\tilde{X}_2'\tilde{X}_2)^{-1}X_2'X_1(X_1'X_1)^{-1}$

$$H = (X_1 \ \ X_2)\begin{pmatrix} (X_1'X_1)^{-1} + (X_1'X_1)^{-1}X_1'X_2(\tilde{X}_2'\tilde{X}_2)^{-1}X_2'X_1(X_1'X_1)^{-1} & -(X_1'X_1)^{-1}X_1'X_2(\tilde{X}_2'\tilde{X}_2)^{-1} \\ -(\tilde{X}_2'\tilde{X}_2)^{-1}X_2'X_1(X_1'X_1)^{-1} & (\tilde{X}_2'\tilde{X}_2)^{-1} \end{pmatrix}\begin{pmatrix} X_1' \\ X_2' \end{pmatrix}$$

$= \left(X_1(X_1'X_1)^{-1} - \tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}X_2'X_1(X_1'X_1)^{-1} \quad \tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}\right)\begin{pmatrix} X_1' \\ X_2' \end{pmatrix}$

$= H_1 - \tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}X_2'H_1 + \tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}X_2' = H_1 + \tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}X_2'(I_n - H_1)$

$= H_1 + \tilde{X}_2(\tilde{X}_2'\tilde{X}_2)^{-1}\tilde{X}_2' = H_1 + \tilde{H}_2$

5. under heteroskedasticity, we know $\hat{\beta} \overset{a}{\sim} N(\beta, \tilde{V})$

where $\tilde{V} = \hat{\Sigma}_{HC3} = (X'X)^{-1} X' \operatorname{diag}\left\{\frac{\hat{\varepsilon}_i^2}{(1-h_{ii})^2}\right\} X (X'X)^{-1}$

$H_0: C^T\beta = 0,\ C \in \mathbb{R}^p \Rightarrow C^T\hat{\beta} - C^T\beta \overset{a}{\sim} N(0, C^T\tilde{V}C)$

$\therefore$ Under H_0, $W_c = \dfrac{C^T\hat{\beta}}{\sqrt{C^T\tilde{V}C}} \overset{a}{\sim} N(0,1)$

Then we construct rejection region $\{C^T\hat{\beta} \pm Z_{1-\alpha/2} \cdot \sqrt{C^T\tilde{V}C}\}$

$H_0: C\beta = 0,\ C \in \mathbb{R}^{l\times p} \Rightarrow C\hat{\beta} - C\beta \overset{a}{\sim} N(0, C\tilde{V}C')$

under H_0, $(C\hat{\beta})'(C\tilde{V}C')^{-1}(C\hat{\beta}) \overset{a}{\sim} \chi^2_l$

Then we construct rejection region $\{(C\hat{\beta})'(C\tilde{V}C')^{-1}(C\hat{\beta}) > \chi^2_{l,1-\alpha}\}$