ECONOMETRICS I, SPRING 2018.

A "projection based" proof of the Frisch-Waugh theorem.

Consider regression

$$Y = X_1\beta_1 + X_2\beta_2 + e. \tag{1}$$

We will use three usefull facts:

1. The best fit to the least squares problem is unique (except, of course, if there is perfect collinarity).
2. Any vector or matrix of variables can be split into its projections. In particular $X_2 = P_1X_2 + M_1X_2$. This is something you can ALWAYS do. Here P_1 and M_1 are the projection and residual makers for X_1. (This is not so surprising because $M_1 = I - P_1$.)
3. A regression on orthogonal (sets of) regressors can be done on each (set) at a time while still getting the coefficients from the joint regression.

Now

$$\hat{Y} = X_1\hat{\beta}_1 + X_2\hat{\beta}_2 \ , \tag{2}$$

and, using fact 2, we have

$$\hat{Y} = X_1\hat{\beta}_1 + (P_1X_2 + M_1X_2)\hat{\beta}_2 \ , \tag{3}$$

or, because $P_1 = X_1(X_1'X_1)^{-1}X_1'$, we have

$$\hat{Y} = X_1\hat{\gamma}_1 + (M_1X_2)\hat{\beta}_2 \ , \tag{4}$$

where $\hat{\gamma}_1 = (\hat{\beta}_1 + (X_1'X_1)^{-1}X_1'X_2\hat{\beta}_2)$. Look at this expression. The fit is the same, so we are still looking at the OLS fit. The fitted coefficients are $\hat{\gamma}_1$ (the formula for which we don't often make use of) and $\hat{\beta}_2$ (the original coefficient) to M_1X_2. Now, use fact 3. Because—by construction—M_1X_2 and X_2 are orthogonal, we have that the regression

$$Y = (M_1X_2)\beta_2 + u \tag{5}$$

delivers the same $\hat{\beta}_2$ as does (5) which is the same $\hat{\beta}_2$ that comes from the full equaiton (1). Finally, notice that if you instead regress

$$(M_1Y) = (M_1X_2)\beta_2 + u \ , \tag{6}$$

you also get the same $\hat{\beta}_2$. This comes from fact 2, because $Y = M_1Y + P_1Y$ and P_1Y is orthogonal to M_1X_2, and will drop out in the regression. You may want to use the form (d) to get the same residuals as in (a).