

PLSC 503: Problem Set 6

Thad Dunning

Department of Political Science

Yale University

Spring 2010

Due February 23, 2010

1 Theoretical/Conceptual Exercises

Question 1:

See the answers to questions 8 and 9 in the discussion questions in Chapter 4 of Freedman (2009).

Question 2:

First row: F F F T. The first is false; even if we had $\epsilon \perp X$, this wouldn't usually translate to exact orthogonality on the data (due to random error). (Here, of course, X may be endogenous; then, we don't have $\epsilon \perp X$, which answers the third item). $\epsilon \perp Z$ is false for the same reason. Finally, $\epsilon \perp Z$ is true by assumption: Z is exogenous.

Second row: T F F F. $e_{OLS} \perp X$ always follows from the OLS fit: $X\hat{\beta}_{OLS}$ is the projection of Y into the column space of X , so $e_{OLS} = Y - X\hat{\beta}_{OLS}$ is orthogonal to X . However, the second is false because $\hat{\beta}_{OLS}$ is the projection of Y into the column space of X , not Z . As for the third item, e_{OLS} is calculated from X ; the fourth term is silly.

Third row: F F F F. For the first term, e_{IVLS} is not the projection of Y orthogonal to X ; for the second, e_{IVLS} is the projection of Y orthogonal to the projection of X onto the column space of Z —not the projection of Y orthogonal to Z . For the third and fourth terms, e_{IVLS} is calculated from both X and Z .

Fourth row: F F T. Notice that the IVLS estimator is just OLS applied to the transformed model (6) on p. 182 of Freedman (2009). Thus, the third term, $f \perp (Z'Z)^{-1/2}Z'X$, is true; the other two are therefore false.

Fifth row: F F F. The residual vector f is calculated from X , Z , and $(Z'Z)^{-1/2}Z'X$, so it can't be independent of these matrices.

Sixth row: F T F F. Here, Δ is the residual from the column-by-column regression of X on Z , and \hat{X} is the fitted value from this regression. The residual vector must be orthogonal to Z (the design matrix in this regression), as usual. But the first term is false (X is the response variable in the regression). Independence is also false, for the usual reasons.

The final three rows are false: orthogonality of the residuals and the design matrix is a property of the regression fit and is always true; it doesn't say anything about independence or bias.