

Problem Set #3 (Measurement Error, IV) ... Answer Key

1. Wooldridge, Problem 9.4...

- (a) For the CEV assumptions to hold, we must be able to write $tvhours = tvhours^* + e$, where e is a mean zero measurement error that is uncorrelated with $tvhours^*$ and the other explanatory variables in the model.
- (b) The CEV assumptions are unlikely to hold in this example. For children who do not watch TV, $tvhours^* = 0$ and it is likely that reported TV hours will also be zero ($e = 0$). If $tvhours^* > 0$, then measurement error can be positive or negative, but since $tvhours \geq 0$, this implies that $e \geq -tvhours^*$. So, e and $tvhours^*$ are likely to be correlated. Moreover, it is likely that e and the explanatory variables will be correlated since e and $tvhours^*$ are correlated and $tvhours^*$ and the control variables are correlated.

2. Wooldridge, Problem 15.2...

- (a) It seems reasonable to assume that $dist$ and u are uncorrelated since classrooms are not usually assigned with convenience for particular students in mind.
- (b) The variable $dist$ must be partially correlated with $atndrte$. More precisely, in the reduced form

$$atndrte = \pi_0 + \pi_1 priGPA + \pi_2 ACT + \pi_3 dist + v$$

we must have $\pi_3 \neq 0$. This can be tested using the t -statistic for π_3 .

- (c) We now need an IV for $atndrte$ and the interaction term, $priGPA * atndrte$. Under the assumption that $E[u|priGPA, ACT, dist] = 0$, any function of $priGPA$, ACT , and $dist$ is uncorrelated with u . So, the interaction between $priGPA$ and $dist$ is also an IV. Furthermore, if $dist$ is partially correlated with $atndrte$, then this interaction will also be partially correlated with the interaction, $priGPA * atndrte$.

3. Wooldridge, Problem 15.4...

- (a) The state may set the minimum wage based on past or expected future economic activity, and this may lead to correlation between $gMIN$ and u . Thus, OLS will be biased and inconsistent.
- (b) Because $gGDP$ controls for overall performance of the US economy, it seems reasonable to assume that $gUSMIN$ will be uncorrelated with the error term for state-level employment.
- (c) If a state always sets its minimum wage equal to the federal minimum wage, then $gMIN$ will be exogenous if $gUSMIN$ is exogenous, and OLS will be unbiased and consistent. However, if states deviate from the federal minimum wage, $gMIN$ may be endogenous, but it should be correlated with $gUSMIN$ since state minimum wages must always be at least as great as the federal minimum wage.

4. Wooldridge, Problem 15.6...

- (a) Plugging (15.26) into (15.22) yields

$$\begin{aligned} y_1 &= \beta_0 + \beta_1(\pi_0 + \pi_1 z_1 + \pi_2 z_2 + v_2) + \beta_2 z_1 + u_1 \\ &= (\beta_0 + \beta_1 \pi_0) + (\beta_1 \pi_1 + \beta_2) z_1 + \beta_1 \pi_2 z_1 + (\beta_1 v_2 + u_1) \end{aligned}$$

So, $\alpha_0 = \beta_0 + \beta_1 \pi_0$, $\alpha_1 = \beta_1 \pi_1 + \beta_2$, and $\alpha_2 = \beta_1 \pi_2$.

- (b) $v_1 = \beta_1 v_2 + u_1$.
- (c) u_1 and v_2 are both mean zero and uncorrelated with z_1 and z_2 . So, v_2 has these same properties. Thus, OLS consistently estimates the α 's using the reduced form equation.

5. Wooldridge, Problem C15.3...

- (a) IQ scores as well as the availability of four year colleges vary by geographic region. For whatever reason, it may be that individuals with higher ability grow up in areas with four year colleges nearby.
- (b) The results suggest that IQ scores are about 2.6 points higher on average for those who grew up near a four year college. The result is statistically significant.
- (c) Now, the coefficient on *nearc4* is statistically insignificant. So, once we control for region and the environment where one is raised, there is no apparent link between IQ and proximity to a four year college.
- (d) It appears important to include these variables in the $\ln(wage)$ equation to control for differences in access to four year colleges that may also be correlated with ability.

6. Wooldridge, Problem C15.5...

- (a) The coefficient on \hat{v}_2 is statistically insignificant. Thus, despite important differences in the results, one does not reject the null hypothesis that education is exogenous.
- (b) The return to education increases even further to 0.157.
- (c) Stata displays the overidentification test for us automatically. However, it is still instructive to be sure we know how this is calculated. See .do file. We fail to reject the null that the IVs are valid.

7. Wooldridge, Problem C15.8...

- (a) The coefficient implies that participation in a 401(k) plan is associated with a 0.054 higher probability of having an individual retirement account, holding income and age fixed.
- (b) Even holding income and age fixed, different individuals may have different preferences for saving (due to, for example, different discount rates of future utility). People who have high discount rates will save less through both types of instruments, while others with low discount rates might tend to utilize both saving instruments.
- (c) First, $e401k$ must be partially correlated with $p401k$. This is guaranteed since eligibility is a prerequisite for participation. Second, $e401k$ must be uncorrelated with the error term. Using the story in (a), if individuals with low future discount rates seek employers that provide 401(k) plans, then $e401k$ may be endogenous as well. Finally, $e401k$ must not have a direct effect on *pira*. This seems reasonable; conditional on $p401k$, $e401k$ may not affect whether one has an IRA.
- (d) $e401k$ is highly significant in the first-stage.
- (e) The IV estimate indicates a statistically insignificant effect of $p401k$ on *pira*. This is consistent with the discount rate story above biasing the coefficient up. However, interestingly, the IV result does not suggest a trade-off between 401(k) plans and IRAs. Thus, 401(k) plans do not seem to crowd out other forms of personal savings.
- (f) The coefficient on the residual, \hat{v} , is statistically significant, confirming the endogeneity of $p401k$. However, this result hinges on $e401k$ being a valid IV.

8. Wooldridge, Problem C15.9...

- (a) See .do file
- (b) The coefficients are identical, but the standard errors are slightly larger.
- (c) The coefficient on education and the standard error both decline, and the drop in the return to education is not trivial.

9. Wooldridge, Problem C15.10...

- (a) The 95% confidence interval is roughly 8.8% to 11.4%.
- (b) The correlation is -0.0169, essentially zero.
- (c) The return is now 13.7%.
- (d) The coefficient is negative and highly statistically significant.
- (e) The return to education is 25.0%, but the confidence interval is very wide. Other than excluding zero, the results are not very informative.
- (f) The IV procedure is not very useful, as stated in (e). In addition, one might worry that since *educ* and *ctuit* are only partially correlated (and not correlated in general), that the results would be very sensitive to changes in the set of control variables.