Economics Honors Exam 2008 Solutions Question 9

(a) (3 points) Δ*(LowEd share)* $= .01$, so the predicted change in the corruption rate is $.01 \times 18.4 = .184$ and the standard error of this predicted change is $.01 \times 8.7 = .087$, so the 95% confidence interval is $.184 \pm 1.96 \times .087 = .184 \pm .171 = (.013, .355)$.

Partial credit awards

1 point: for getting the mean correct (0.5 points for the correct method)
1 point: for getting the standard error correct (0.5 points for the correct method)
1 point: for getting the confidence interval correct (0.5 points for the correct method)

(b) (1) (3 points) The t-statistic is $t = 47.7/94.8 = 0.50$, which is < 1.96, so we do not reject the null at the 5% significance level.

Partial credit awards

1 point: for using t-statistic
1 point: for calculating the t-statistic correctly
1 point: for not rejecting the null

(2) (3 points) Under the null hypothesis, *Voting share* does not enter regression (3), which means that the coefficients on *Voting share* and *LowEd share*$\times$*Voting share* must both be zero. The F-statistic testing this hypothesis is $.52$ with $p = .60 > .05$, so we do not reject the null hypothesis at the 5% significance level.

Partial credit awards

1 point: for realizing that under the null hypothesis, the coefficients on *Voting share* and *LowEd share*$\times$*Voting share* must both be zero
1 point: for getting the F-statistic and p-value correct
1 point: for not rejecting the null

(c) (1) (4 points) Disagree. The coefficient on Foreign-born share is positive so the sign of the estimated coefficient indicates that more foreign-born citizens is associated with more corruption, however the t-statistic is $21.3/14.3 = 1.49 < 1.645$ so the coefficient is not significant at the 10% level, so there is no statistically significant support for this claim at conventional levels of significance.

Partial credit awards

2 points: for disagreeing

0-2 points: depending on the quality of reasoning

(2) (4 points) Disagree. The low R^2 indicates that there might be determinants of corruption omitted from the regression, but that alone does not mean there is omitted variable bias. For omitted variable bias to exist, the omitted variables (1) need to be determinants of Y and (2) need to be correlated with the included regressor(s). The low R^2 indicates that (1) is probably true (although not necessarily the error term could just be measurement error) but the R^2 is silent on point (2).

Partial credit awards

2 points: for disagreeing
0-2 points: depending on the quality of reasoning

(3) (4 points) Disagree. Heteroskedasticity-robust standard errors are valid whether the errors are heteroskedastic <u>or</u> homoskedastic.

Partial credit awards

2 points: for disagreeing
0-2 points: depending on the quality of reasoning

(d) (1) (3 points) This means that there is simultaneous causality: low education causes corruption and vice versa. As a result the OLS estimator is biased.

Partial credit awards

1 point: for answering simultaneous causality
1 point: for explaining simultaneous causality
1 point: for realizing the OLS estimator is biased

(2) (6 points) The two conditions for a valid instrument Z are (1) it is relevant, i.e. Z is correlated with X and (2) it is exogeneous, i.e. Z is uncorrelated with the error term. So:

(i) Newspapers: (1) relevance: maybe higher levels of education might mean more newspaper readership; (2) exogeneity: probably not one hopes that more newspapers would help to uncover corruption and thereby limit corruption.

(ii) Alphabet: (1) relevance: no no reason for alphabetical listing to be correlated with levels of education. (2) exogeneity: yes no reason for alphabetical listing to be correlated with anything!

Partial credit awards

2 points: 1 point each for recognizing the two conditions for a valid instrument

2 **points:** for part (i), 1 point each for the two conditions
2 **points:** for part (ii), 1 point each for the two conditions

(e) (18 points) The regressions differ in the instruments that are used, and in whether the manufacturing share is included as a regressor. The instruments should be selected based on relevance and exogeneity. Relevance is measured by the first-stage F-statistic, which should exceed 10 for the two stage least squares results to be statistically reliable. Applying this criterion, we are left with regressions (1), (3), (4), and (6). When the coefficient is overidentified (here, having at least two instruments), the hypothesis that both instruments are exogenous can be tested using the J-statistic. The null of exogeneity is rejected in regression (3) but not regression (6). This leaves us with (1), (4), and (6).

To make a further distinction, we must exercise judgment about the specifications. Are the instruments arguably exogeneous based on our judgment? It seems like they should be, after all they measure conditions in the distant past and in this sense they should not be proximate determinants of corruption in the 1990s. On the other hand, if corruption is related to overall state values and culture that varies slowly over time, it is possible that these instruments still could be correlated with these slowly-varying omitted variables. This suggests that it is sensible to control for more state conditions, for example controlling for the level of manufacturing (treating it as a control variable, for slowly-varying state conditions not as a causal variable for corruption) is warranted. This reasoning leads to preferring regressions (4) or (6). As a practical matter, there is very little difference between the two, however regression (6) includes an instrument that is basically irrelevant (first-stage $F = 2.6$) so it is warranted to drop that instrument, which leaves us with regression (4). The fact that adding $LnInc1940$ as an instrument in regression (6) does not change the results or reject exogeneity is a reassuring robustness check of regression (4).

Partial credit awards

2 **points:** for choosing regression (4)
4 **points:** for using the relevance criterion to exclude regressions (2) and (5)
4 **points:** for using the exogeneity criterion to exclude regression (3)
0-4 **points:** depending on the quality of reasoning for excluding regression (1)
4 **points:** for using the irrelevant-instrument and/or no-change-in-result argument to exclude regression (6)

(f) (6 points) Based on regression (4), low levels of education (low shares of a high school degree) are statistically significant at the 5% level. The magnitude of the effect is substantial: a one standard deviation move in LowEd share is associated with a $32.5 \times .07 = 2.3$ change in the corruption rate, which is approximately a one standard deviation change in the corruption rate (and a change

of more than one-half of the mean corruption rate). Assuming that the findings from the regression are internally valid, according to this regression, increasing the level of education in the population in particular, reducing the fraction of the population with low levels of education has not only the usual direct benefits, but the statistically significant side benefit of substantially reducing corruption.

Partial credit awards

2 points: for realizing that corruption has a significantly negative effect on the level of education
0-4 point: depending on the quality of reasoning

(g) (6 points) Here are two:

1. Are the other regressors good control variables? In particular, voting share could be subject to simultaneous casuality ("vote early and often"), and if so it would introduce simultaneous causality bias and not be a suitable control.

2. Are the instruments really exogeneous? Although they reflect things that happened in the distant past, political culture in a state changes very slowly. The identifying assumption is, in effect, that things that happened long ago are correlated with the level of education today but are not correlated with omitted determinants of corruption today. The specification of the equation for corruption today omits things that are plausibly strongly serially correlated, such as the vigor of investigative journalism, the way that prosecutors are appointed (political appointments? elected officials? rising through the bureaucrac?) It is particularly troubling that the results on instrument exogeneity hinge on whether the manufacturing share is included as a regressor (regression (3) v. (6)), when the manufacturing share itself is hard to understand as a proximate cause of corruption. So the case for instrument exogeneity is not compelling.

Partial credit awards

3 points each for the two threats, and depending on the quality of reasoning