Public Policy 529
Fall 2023: Problem Set #9

Due on Wednesday, November 29

1. Use the `world` dataset for the following:

    (a) Using your software, find the Pearson's r correlation coefficients between democracy (`dem_score14`), life expectancy (`lifeex_total`), and the level of poverty (`unpovnpl`). Make sure you have your software report statistical significance (see lecture slide 24 for commands). Report the resulting correlation matrix in your answers.

    The variable `dem_score` is a rating of a country's degree of democracy on a scale that goes from 1 to 10; `lifeex_total` is life expectancy at birth in a country; `unpovnpl` is the percentage of a country's population that is below the national poverty line.

    (b) Interpret the correlation coefficients for each pairing of the three variables (not counting each variable with itself).

    (c) If the sample size for the correlation between `dem_score14` and `unpovnpl` were 82, what would be the t-statistic and p-value for the correlation coefficient?

2. This question uses the `world` dataset. Let's use bivariate linear regression to estimate the linear relationship between `educ_f_none` (dependent variable) and `democ11` (independent variable). The variable `educ_f_none` is the percentage of females in a country with no schooling. `democ11` is an 11-point scale of democracy, which runs from 0 ("not at all democratic") to 10 ("highly democratic"). Note: we are treating `democ11` as an interval-level variable for this analysis.

    (a) Estimate the regression and report the output. See lecture slides or help documents for appropriate commands.

    (b) What is the substantive meaning of the estimated intercept? In other words, what does it tell us about the predicted percentage of females with no schooling?

    (c) What is the coefficient on `democ11`? Is this coefficient statistically significant at the .05 level of significance? How do you know?

    (d) Interpret the substantive meaning of the coefficient on `democ11`. In other words, what does it tell us about the relationship between the democracy scale and fe-

male schooling? Remember that regression coefficients give us information about both the direction and magnitude of this relationship.

(e) If a country were a 4 on the democracy scale, what would we predict to be the percentage of females with no schooling?

(f) Country A is a 7 on the democracy scale; Country B is a 2. What is the predicted difference in female schooling?

(g) Interpret the R^2 statistic for this regression.

(h) What is the size of the typical difference between the *predicted* percentage of females with no schooling and the *actual* percentage of females with no schooling (i.e. the standard error of the estimate)? This Stata calls this the Root MSE; R calls it the residual standard error.