1 Note on the Interpretation of the R^2 in the Estimation of Leader Effects

[Ancillary material accompanying Humphreys, Masters and Sandbu. The Role of Leaders in Democratic Deliberations: Results from a Field Experiment in São Tomé and Príncipe. *World Politics*. 2007.]

In our paper we use the R^2 as a measure of leader influence. As suggested in footnote 36 of the paper, the R^2, can, under some conditions, be considered a lower bound on the degree of leader influence. In particular:

- Suppose that without manipulation a share q of groups would select "1"
- Suppose that share α of leaders support "1"
- Suppose that leaders can influence groups they disagree with β of the time (more complex influencing rules are possible!)

Then, the probability that we observe a "1" is:

$$\begin{aligned} p &= \Pr(y = 1) \\ &= q \times [\alpha + (1-\alpha)(1-\beta)] + (1-q) \times \alpha\beta \\ &= \beta\alpha + (1-\beta)\,q \end{aligned}$$

Note that p is just a weighted average of α and q with weight given by β.

SST (Sum of squared errors) is then given by:

$$\begin{aligned} SST &= p[1-p]^2 + [1-p]\,[p]^2 \\ &= p[1-p][1-p+p] \\ &= p[1-p] \end{aligned}$$

So:

$$\begin{aligned} SST &= [\alpha\beta + q\,[1-\beta]]\,[1 - \alpha\beta - q\,[1-\beta]] \\ &= \alpha\beta(1-\alpha\beta) + q(1-\beta)(1-2\alpha\beta) - (1-\beta)^2 q^2 \end{aligned}$$

How does this change once you use information about the characteristics of the leader?

- If a leader favors "1" then the expected result when he leads is: $q+(1-q)\beta$
- If a leader favors "0" then the expected result when he leads is: $q(1-\beta)$

Using this information our errors would then be:

$$\begin{aligned} SSR &= \alpha[q+(1-q)\beta][1-q-(1-q)\beta]+(1-\alpha)[q(1-\beta)(1-q(1-\beta))] \\ &= \left[(1-2\alpha\beta)q+\alpha\beta-q^2(1-\beta)\right](1-\beta) \end{aligned}$$

So:

$$\begin{aligned} SST-SSR &= \alpha\beta(1-\alpha\beta)+q(1-\beta)(1-2\alpha\beta)-(1-\beta)^2q^2 \\ &\quad -\left[(1-2\alpha\beta)q+\alpha\beta-q^2(1-\beta)\right](1-\beta) \\ &= \alpha(1-\alpha)\beta^2 \end{aligned}$$

Thus:

$$R^2 = \frac{\alpha(1-\alpha)}{\left[\alpha\beta+q\left[1-\beta\right]\right]\left[1-\alpha\beta-q\left[1-\beta\right]\right]}\beta^2$$

Note that implicitely in generating this measure we have assumed an infinite number of observations for each facilitator. In practice however we do not have these and so, in the paper, we report the more conservative adjusted R^2.

Now, we claim in the text that "it can be shown" that:

$$\frac{\alpha(1-\alpha)}{\left[\alpha\beta+q\left[1-\beta\right]\right]\left[1-\alpha\beta-q\left[1-\beta\right]\right]}\beta \leq 1$$

To see this note that:

$$\begin{aligned} \alpha(1-\alpha) &\leq \left[\alpha\beta+q\left[1-\beta\right]\right]\left[1-\alpha\beta-q\left[1-\beta\right]\right] \\ &\leftrightarrow \\ \alpha(1-\alpha)\beta &\leq \alpha\beta(1-\alpha\beta)+q(1-\beta)(1-2\alpha\beta)-(1-\beta)^2q^2 \\ &\leftrightarrow \\ 0 &\leq \alpha^2\beta(1-\beta)+q(1-\beta)(1-2\alpha\beta)-(1-\beta)^2q^2 \\ &\leftrightarrow \\ 0 &\leq \alpha^2\beta+q(1-2\alpha\beta)-(1-\beta)q^2 \\ &\leftrightarrow \\ 0 &\leq \beta(\alpha-q)^2+q(1-q) \end{aligned}$$