

## **Re-Replication of BOWLES, BULLOCK, and MAZUMDER**

**By Elizabeth Rosenblatt**

My first comment is a quick but important adjustment to be made. Many of the presented replication tables label the models and independent variables incorrectly. For example, the model (column) labels in tables 1, 2 and 4 are incorrect (likely just due to an error inputting into stargazer – easy fix!). And the independent variable (row) labels in Table 5 and 6 are all off by one with their corresponding coefficients.

Secondly, I think you should pay close attention to the form of each of the dependent variables in the linear regression models. What form are these variables in and what exactly do they represent? Consider whether a linear model is the best way to calculate relationships between the variables. For example:

- Table 2.2: Civil War Incidence is the number of years a country has been in engaged in civil war since its independence. The observations are discrete counts of the number of years. As the model exists now, it helps to show the average number of years nations spend in civil wars. By altering the model, you might be able to extract other interesting characteristics from this relationship.
- Table 3.5: This model conducts a linear regression on binary data (xconst5). You could try to use a generalized linear model instead (like a probit or logit). Binary data, by nature, violates many of the assumptions of the basic linear regression and using a generalized linear regression model will address these violations. Note in this case that the dependent xconst5 variable (which represents the extent of checks and balances) was generated by the authors from a categorical scale from 1-10 where they let observations  $\geq 5$  equal 1 and observations  $< 5$  equal 0. This may influence the generalized model that you choose to use. In this table 3.5, the authors find no correlation between historical conflict and modern measures of institutional capacity. This is a surprising result. Maybe a generalized model would produce different qualitative conclusions.
- Table 4: The three dependent variables used in table 4 (inter\_group\_trust, identitydumeth, identitydumnat) are either categorical or binary. Inter\_group\_trust is a categorical variable from a survey that asks “how much the respondent trusts people from other groups. Answers are given on a four-point scale where 0 is ‘not at all,’ 1 is ‘just a little,’ 2 is ‘somewhat,’ and 3 is ‘a lot.’”<sup>1</sup>
- Table 6: The dependent variable ConflictGrid is binary. May want to try using a generalized linear model.

Lastly, the majority of the models in this paper use heteroskedasticity corrected covariance matrix in order to calculate robust standard errors. The standard errors presented in the paper are the robust standard errors. It might be revealing to

---

<sup>1</sup> Page 328

compare the RSEs to the original SEs and to try to transform your dependent variables so that the RSEs approximately equal the SEs. As mentioned above, many of the models could be improved by converting them to generalized linear models. However, the OLS assumptions should be reexamined for the remaining linear models. It is possible that variable transformations or model refinements may improve these fits. For example:

- Table 2.2: Purges is a continuous variable so the linear regression is definitely appropriate here. However, plotting a histogram of purges shows that it is not normal at all. It is highly skewed to the left. As a result, the robust standard errors are different than the original standard errors. Can you transform the purges variable in a way that makes a linear regression more accurate and returns RSE's close to the original SE's?
- Table 3: Note that the lrgdp2632000 histogram looks much more normal than either the purges variable or civil war incidence variable found in table 2. Most of the RSE's in this model are quite close to the SEs.

I know that your replication PDF suggests that you might want to use updated Afrobarometer survey data to test these same effects. I am slightly weary of this idea. The Afrobarometer data used in the paper was recorded in 2008. It seems to me that because the models are testing effects of events that happened long ago on recent metrics, it might be ineffective to simply update the Afrobarometer data to the 2016 numbers. In theory, the relationships shouldn't change between models conducted on 2008 data and 2016 data. If the relationships did change, I think it would do more to call into question the validity of the datasets and the models than to further substantiate your qualitative points. If you decide to do this, it will be important to justify why you made that choice and what factors could be contributing to the change in qualitative results. For example, the original paper states that the Afrobarometer dataset does not include data points for countries that were engaged in civil war when the data was collected in 2008. Filling in this missing data could be a good justification for updating the dataset.