Beyond Zeroes and Ones

Tables and Supporting Information

August 26, 2014

Over the process of writing and revising this paper, we subjected our original findings to a battery of robustness checks. This appendix describes and motivates each robustness check and then presents the results. Many of these suggestions came from reviewers, and we owe them our appreciation for this advice.

A1 Different Dependent Variables

The Uppsala/PRIO dataset codes a “high,” “low,” and “best” estimate of the number of battle deaths, since there can be disagreements among sources. Our manuscript used the “best” estimates. This section presents alternate specifications of the dependent variable.

We first replicated Table 3 using the high and low estimates. The results are in Tables A1 and A2. In general, the results are similar and the differences are as expected. For similarities, the coefficient on income is always negative and significant, using the low, high, and categorial dependent variables. The AR(1) coefficients are also always positive, significant and less than 1. The estimated $AR(1)$ parameter is also larger in the low series models, suggesting that prior estimates of serial dependence are conservative.

For differences, the coefficient on income is smaller in the low estimate models and higher in the high estimate models. The magnitude of the differences in estimates for the coefficient on income seem reasonable given the magnitude in differences for the low and high dependent variables. For example, the full-sample coefficient on income from Table 3 in the main paper is -321. For the low-estimate models, this coefficient in -52 and for the high-estimate models, this coefficient is -511. (For reference, the mean of the best estimates is 335 compared to 122 and 599 for the low and high estimates respectively.

We then replicated Table 3 using an often-used categorical coding of the intensity of civil conflicts/wars. For this categorical variable, “0” indicates 0-24 battle deaths. “1” indicates 25-999 battle deaths. “2” indicates battle deaths of 1,000 or higher. For the categorical dependent variable, 537 observations are coded as a “1” and 469 are coded as a “2.”Table A3 shows these results.

The Uppsala/PRIO battle deaths data also sometimes requires interpolation across conflict years. When deaths data are unavailable for particular years, the Uppsala/PRIO dataset does not report a “best” estimate. Interpolation using adjacent years of data is used to fill in missing observations in these cases. Sub-saharan Africa is the region with the most missing data. There are 193 conflict-years that include a best estimate in the Uppsala/PRIO dataset for sub-saharan Africa, but there are 121 observations missing when conflicts are occurring in the same country in adjacent years. Interpolation thus provides an additional 121 country-years of data for sub-saharan Africa. For other regions, the discrepancy is much smaller. There are 511 conflict-years outside of sub-saharan Africa with an available best estimate in Uppsala/PRIO, and interpolation fills in another 182 conflict-years. Appendix Table 4 shows the results that exclude these observations and only uses observations for which distinct, yearly deaths data were available. Estimates differ, especially in sub-Saharan Africa, for two reasons: first, the number of observations with data on battle deaths falls–reducing statistical power; second, the conflicts that remain are, on average, less severe than the excluded conflicts that require interpolation.

As a further robustness check regarding whether interpolation affects the results on dynamics, Figure A1 re-produces the results from Figure 2 without using the interpolated measures of battle deaths. The results suggest that interpolation does not substantively change the interpretation or estimates of γ.

Readers who are interested in comparisons with the extensive margin should exercise caution when combining results with the "low" series and estimates of the extensive margin from the text. Calculations were conducted using the moments of the battle deaths data; because the first and second moments of the "low" series and the "best" series differ, the results are not comparable when using the "low" series.

A2 First Stage Results and Sensitivity to Different Panel Instruments

The text presents estimates of the first stage regression using income as the dependent variable. Here we present regressions analogous to Dollar and Kraay (2002) to assess the strength of the first stage relationship for the lagged dependent variable in levels and differences. As mentioned in the text, panel-style instruments are weak in differences, so we added the system equation to compensate. Appendix Table 5 presents the results. It should be noted that the probability limit on the third lag of battle deaths (the first instrument used after Arellano-Bond autocorrelation tests) in the difference specifications in Panel A is expected to be negative.[1] While the F-statistics look small in some specifications, the parameter estimates in the levels equations are all significant at conventional levels on the panel-style instruments. As a further sensitivity test, Appendix Table 6 replicates the main results, but restricting the GMM-style instrument matrix to include only the 3rd lag of the process. The results are largely consistent with the results from the manuscript,

[1] Ignoring other regressors, the probability limit is $\gamma^2 - \gamma < 0$, where γ is the $\mathrm{AR}(1)$ parameter.

with the main exceptions being a positive coefficient on income in the sub-Saharan Africa subsamples, using the lagged war indicators.

A3 Fixed Effects and Levels vs. Differences

In one exchange on the relationship between income and civil conflict, Djankov and Reynal Querol (2010) and Brückner (2011) use different combinations of fixed effects strategies and specifications of variables in levels or differences. Specifically, Djankov and Reynal Querol find that regressing a binary indicator of civil war on income, in levels, does not yield significant effects when country-level fixed effects are included. Brückner finds that regressing civil war on income, in first differences, does yield significant negative effects, even when country-level fixed effects are included to capture differences in trends.

Here, a similar exercise is conducted using conflict intensity as captured by battle deaths. Appendix Table 7 presents regressions of battle deaths on income, in levels and first differences, with and without country fixed effects. Each regression also includes year fixed effects.

We find a pattern similar to that of Brückner, where the effect of income is weaker in the levels model once country fixed effects are included, but stronger in the first differences model, regardless of whether country fixed effects are included. There are large differences in the magnitude of parameter estimates in first differences compared to fixed effects regressions. Upon further inspection, these differences appear to be driven by outliers with very large residuals. (Note that this isn't a problem in the Blundell-Bond estimates in orthogonal deviations). Appendix Table 8 presents results after excluding outliers. The magnitude of the estimates for the effect of income falls and is much closer to the estimates in the main Blundell-Bond models.

A4 Bias in Static Models

The estimates from the dynamic panel data models presented in the main paper suggest that the conflict process is dependent. Many prior papers use static models, but the parameter estimates of any parameter of interest from static models are likely to be inconsistent even with an instrument. This is easiest to see using first differences, but the same logic applies to the within-transformed IV estimator because the justification for the most prevalent instruments used in the literature–rainfall shocks and the price of commodity exports–is that the instruments and the error are orthogonal conditional on the unobserved fixed effects. However, these instruments are not likely to be valid without the fixed effects–meaning that the instrument is correlated with the country effects. For example, a country's time-invariant mix of commodity exports or a country's long-run average weather patterns may influence the probability of civil war–but the within-country, time-varying instruments would likely satisfy the exclusion restriction after accounting for the fixed effects if the conflict process were static. If the process is dynamic, the fixed effects cannot be differenced out, so the instrument is correlated with the error, violating the exclusion restriction.

The bias can be signed in the case of the first-differenced IV estimator. Ignoring time fixed effects for exposition, let the true model generating the data be given by $y_{it} = \gamma y_{i,t-1} + x_{it}\beta + \alpha_i + \varepsilon_{it}$, with $E\left(x'_{it}\varepsilon_{it}\right) \neq 0$, $E\left(z'_{it}\varepsilon_{it}\right) = 0$, $E\left(\varepsilon'_{it}\varepsilon_{is}\right) = 0$ for $s \neq t$, and $E\left(x'_{it}z_{it}\right) \neq 0$. Suppose it is erroneously assumed that $\gamma = 0$, and estimation is via first-differenced instrumental variables. The estimated parameter is $\hat{\beta} = \left(\Delta z'_{it}\Delta x_{it}\right)^{-1}\Delta z'_{it}\Delta y_{it}$ and the bias is

$$E\left(\hat{\beta} - \beta\right) = E\left(\left(\Delta z'_{it}\Delta x_{it}\right)^{-1}\Delta z'_{it}\Delta y_{i,t-1}\right)\gamma. \tag{1}$$

To sign the bias analytically, further assume that the time series relationship for the instrument is $z_{it} = \gamma_z z_{i,t-1} + u_{it}$. [2] The bias is

$$E\left(\left(\Delta z'_{it}\Delta x_{it}\right)^{-1}\Delta z'_{it}\Delta y_{i,t-1}\right)\gamma = E\left(\left(\Delta z'_{it}\Delta x_{it}\right)^{-1}\left[\left(\gamma_z - 1\right)z_{it-1} + u_{it}\right]'\Delta y_{i,t-1}\right)\gamma.$$

The first stage implies that $E\ \left(\Delta z'_{it}\Delta x_{it}\right) > 0$ and γ is expected to be positive, so with these restrictions, the term $E\left(\left[\left(\gamma_z - 1\right)z_{it-1} + u_{it}\right]'\Delta y_{i,t-1}\right)$ determines the sign of the bias. After substituting in $z_{it-1} = \gamma_z z_{i,t-2} + u_{it-1}$, the relevant term becomes

$$E\left(\left[\gamma_z z_{it-1} + u_{it} - z_{it-1}\right]y_{it-1} - \left[\underbrace{\gamma_z^2 z_{it-2} + \gamma_z u_{it-1} + u_{it}}_{z_{it}} - \underbrace{\gamma_z z_{it-2} - u_{it-1}}_{-z_{it-1}}\right]y_{it-2}\right).$$

[2] Dickey-Fuller style tests reject the null that $\gamma_z = 1$ in favor of an alternative that z_{it} is a trend-stationary process.

Assuming that $E\left(u_{it}y_{it-s}\right)=0$ for $s>0$ and taking expectations, the sign of the bias is determined by

$$E\left([\gamma_z-1]z_{it-1}y_{it-1}-[\gamma_z-1]\gamma_z z_{it-2}y_{it-2}\right)$$

Suppose that the reduced form relationship $E\left(y_{it}z_{it}\right)<0$ is constant for all t. If z_{it} is stationary, then $\gamma_z<1$, which implies $\left(\gamma_z-1\right)-\left(\gamma_z-1\right)\gamma_z<0$ so that $E\left(y_{it}z_{it}\right)\left[\left(\gamma_z-1\right)-\left(\gamma_z-1\right)\gamma_z\right]>0$. Combined with $\gamma>0$ and $E\left(\Delta z'_{it}\Delta x_{it}\right)>0$, parameter estimates from static models are biased upward.

Presumably having an excluded instrument will alleviate some concern about the potential bias from a static model. However, this intuition is only true if the instrument z_{it} is orthogonal to both country fixed effects, α_i, and the error, ε_{it}. Otherwise, the instrument is only valid conditional on the procedure to remove α_i ; these procedures will suffer from Nickell (1981) bias in the case of the within-transformation or the bias derived previously in the case of the first-difference transformation.

To test whether the instrument is orthogonal to α_i, the null hypothesis is that the pooled OLS IV estimator and the within-transformed IV estimator have the same probability limit.[3] It is possible to construct over-identified estimators from moment conditions that impose $E\left(z_{it}\left[\alpha_i+\varepsilon_{it}\right]\right)=0$ or only $E\left(z_{it}\varepsilon_{it}\right)=0$. Using 2 sets of moment conditions, the first of which corresponds to pooled OLS IV and the second of which corresponds to within-transformed IV, equality of the estimates is rejected at the 5 percent level using Hansen's J-test. The results of this test confirm that the variation used to estimate the effect of income in static models is valid only conditional on fixed effects. However, if the true data generating process is dynamic, static estimates are biased.

How large is the bias? The empirical estimate of the bias term for the first-differenced IV estimator, $\overline{\left(\Delta z'_{it}\Delta x_{it}\right)^{-1}\Delta z'_{it}\Delta y_{i,t-1}}$, is $2,381$. This suggests that static models may be biased badly, and the bias is likely to be increasing in the degree of persistence.

[3]The within-transformed moment conditions are used rather than the first-difference IV moment conditions to preserve the same number of observations across specifications.

Supplementary Appendix: References

Nickell, Stephen. 1981. “Biases in Dynamic Models with Fixed Effects.” *Econometrica* 46(6):1417–1426.