The Causes, Consequences, and Management of Civil Wars
030:178, Section 1
Guide to Reading Statistical Articles

1) Don't skip the tables! Even though the numbers and statistics might be difficult to work through, it is important to figure out what their content implies. The authors should try to convey in the text what the tables/figures are showing.
2) Unit of analysis: what does each case represent in the data? In this class, typical units of analysis will include state-year, duration of peace (number of years since last war), and civil war outcomes. Figure out what years and countries are covered in the study, which will be useful for making comparisons across studies with different results.
3) Measurement: how are the key dependent and independent variables measured? Are they nominal (e.g. male/female), ordinal (e.g. low democracy/high democracy), or continuous (e.g. number of battle deaths)? Are they counts of events, such as number of war participations? Dependent variables that are nominal and dummies (1/0) will typically be analyzed using logit or probit models. Models that look at the duration of civil wars or peace between wars will count the number of time periods (months or years). These models are often called survival analyses because they are looking at the time to death (e.g. the end of a civil war). If the dependent variable was ordinal, then the authors might employ an ordered logit model. If the dependent variable was continuous, then a regression model would be appropriate.
4) Statistical significance: Scholars are typically trying to determine if the factors they have identified as independent variables (e.g. primary commodity exports, income inequality, regime type) have a significant effect on the dependent variable (e.g. civil war onset). The null hypothesis is that the effect is zero, while the alternative hypothesis is that the effect is different from zero. One-tailed hypothesis tests specify the direction of difference (positive, negative), while two-tailed tests specify only difference but not direction. Most statistical packages report two-tailed tests, so it is fairly common for authors to employ them. Social scientists typically employ a 95% confidence interval for their tests, which means that if we drew 100 random samples from the population, 95 would contain the population parameter being estimated. Articles we read will sometimes report the exact level of significance ($p=.07$), or the range (* will indicate $p<.10$). There are two basic rules of thumb here. If the estimated parameter (b or β) is twice as large as its standard error (or more), or $b/s.e. > 2$, then the effect of the variable is statistically different from zero. For one-tailed tests, the threshold would be lower for significance ($t>1.65$ for 95% significance). Authors might also report the p-value, which is the area under the standard normal curve (or student t distribution) beyond the estimated z(or t) score. If that p-value is less than .05, then the effect is statistically significant.
5) Model fit: Scholars will also report statistics about the overall goodness of fit. These give us a benchmark for how well the independent variables in the model explain the variance in the dependent variable. One common measure is R^2 which is a measure of the percentage of variance in the dependent variable explained by the dependent variables. If R^2 equals zero, then the independent variables do not help us explain variance in Y; If R^2 equals one, then the independent variables perfectly predict the variance in Y (the data would fit on a perfect line in a bivariate model with two variables). Of course, most

models in Political Science never approach R^2 equal to one, so scholars will discuss the percentage of explained variance. In the civil war context, a good model might explain 80% of the variance. Scholars might also report a chi-square or F-test for the overall fitness of the model. If the p-value for these tests is less than .05, then we can conclude that the model is a good one overall. Scholars may also report their improvement in predictions for logit & probit models. These help to determine how much better we can guess the actual value of the dependent variable (e.g. 1 for civil war onset, 0 for peace) based on knowledge of the independent variables (predicted value based on model).

6) Substantive significance: In addition to determining whether the independent variables have a statistically significant effect on the dependent variable, scholars are also interested in the size of that effect. If we estimated the effect of class attendance on grade, we could imagine you would come to class more often if each class raised your grade by 5 percentage points than if the effect was only 0.2 percent. Scholars assess substantive significance in regression models by looking directly at the size of the coefficient, while in logit or probit models, they calculate the predicted probability that the dependent variable equals one. This entails setting all independent variables at their mean or mode and then moving one variable at a time across some range of possible values (e.g. minimum to mean to maximum). You might see the actual probabilities reported, or sometimes you will see the chance in predicted probabilities. Suppose that the probability of civil war onset was 0.05 when democracy was high and 0.35 when democracy was low. The change in probability would be 0.35-0.05/0.05 = 6, or a 600% increase in the likelihood of war.
7) Multicollinearity: scholars may discuss this issue and it refers to the possible high correlation between independent variables. Suppose that democracy and wealth are correlated at 0.80. This means that typical democratic countries are rich and typical non-democratic countries are poor. The problem this creates is that we cannot determine whether civil wars occur because countries are poor or because they are non-democratic. It is hard to assess the independent effects of each variable because they share so much variance in common. What often happens is that the overall model fit looks good (R^2 might be 0.70) but the t-scores for each variable are insignificant ($t< 2$). In these cases, scholars sometimes put in one variable at a time or use statistics to assess the overall effects of the correlations. It is important to think about correlations between variables because these can have big effects on our findings in multivariate models.
8) Heteroskedasticity & Autocorrelation: scholars might discuss these issues and they create similar problems to multicollinearity by increasing the standard errors and making it more likely to think the independent variables have no significant effect. What happens in the case of heteroskedasticity is that the model fits differently across the cases. Imagine that at high levels of income, the model does a good job of explaining the incidence of civil war, while at low income levels, some countries have civil wars, while others do not. The model fits worse at the low range of income, which creates bigger variance in the standard errors. Autocorrelation refers to temporal correlation between cases over time. For example, presidential approval in one month is correlated with presidential approval in the previous month. Typical solutions to these problems include running fixed effects (putting in dummy variables for each country or each year) or using statistical fixes for the problem (Generalized Least Squares, panel corrected standard errors, cubic splines, lagged dependent variables, etc.)