Missing Uncertainty and Uncertain Missing in the Study of Civil War Onset

Bjørn Høyland[1,2] and Håvard Mokleiv Nygård[1,2]

[1]Department of Political Science, University of Oslo
[2]Centre for the Study of Civil War, PRIO

December 23, 2011

ABSTRACT

The quantitative study of armed civil conflict is plagued by considerable amount of missing data on variables theory considers to be important. The literature so far have for the most part assumed, more often than not implicitly, that the missing data is ignorable. Key variables, such as indicators of governance and regime type, are in addition measured without any uncertainty, and time series of such variables are used without considering how much new information they actually represent. This is highly problematic. The most robust findings in the literature on civil war onset to date are as follows: large populations, low levels of GDP, political instability, inconsistent democratic regimes, rough and mountainous terrain and being situated in a war prone neighborhood increases the risk of civil war onset. Using Mmultiple imputation of missing data we evaluate to what extent these findings are driven by ambitious missingness assumptions or inflated certainty level in the classification process. We find that some of the key results in the literature is not entirely robust to a loosening of these rather stringent assumptions.

1. INTRODUCTION

Missing data is endemic in the literature on civil armed conflict. Many of the factors the literature considers to be of importance for explaining the incidence, onset and duration of civil war are demographic and socio-economic factors for which the data coverage for many countries is scarce or almost non-existent. Of course, these countries tend to be the very same countries which sees the most conflict in the first place. Variables such as gross domestic product, energy consumption, infant mortality rates and regime type are missing for many of the country years an analyst of civil war would be most interested in.

In addition, the quality of the data that does exist is often extremely poor. It is therefore not uncommon at all to read quantitative articles on civil war dynamics with 50, 60 or even 70 percent missing data.

To a very large extent this missing data is either ignored and dealt with through listwise deletion, or is handled in a haphazard and ad-hoc fashion. This is extremely problematic and completely unnecessary. A vast literature exist on how to handle missing data, and a several easy-to-implement tools and methods can be implemented in most modern statistical software packages, or are available for free. We know that listwise deletion can induce severe bias, and that it generally leads to too narrow confidence intervals, even when data is missing at random. We also now that methods such as especially multiple imputation will outperform listwise deletion, as well as often used methods such as mean and conditional mean imputation, under very general conditions. Nevertheless, listwise deletion continues to be the staple method for handling missing data.

This paper analyzes the consequences of missing data in the literature on armed conflict, and discusses ways to improve how missing data is generally handled. We replicate the seminal works of Fearon and Laitin (2003), Collier and Hoeffler (2004) and Hegre, Ellingsen, Gates and Gleditsch (2001). We check how robust these analysis are to a more statistically sound handling of missing data. We find that the more missing data an analysis is plagued by, the more severe the potential bias of listwise deletion is.

In the following we first discuss what missing data actually is, and how one can differentiate between different types of missing data before presenting the principle behind multiple imputation. We review the main findings of the three papers before we evaluate the performance of the mulitiple imputation algorithm. The final section concludes and discuss extentions to the approach to missing data problems taken in this paper.

2. STATISTICAL ANALYSIS WITH MISSING DATA

Little and Rubin (2002, 12–13) distinguish between three types of missing data. Data missing completely at random (MCAR), data missing at random (MAR) and data not missing at random (NMAR). Assume an unknown parameter ϕ which can be thought of as the missing value generating mechanism. Then MCAR and MAR are given by equation 1 and 2.

$$MCAR : Pr(M|Y, \phi) = Pr(M|\phi) \tag{1}$$

$$MAR : Pr(M|Y, \phi) = Pr(M|Y_{obs}, \phi) \, \forall Y_{mis}, \phi \tag{2}$$

If the data is missing completely at random, then missingness depends *only* on the parameter ϕ and not at on the dependent variable. The missingness is uncorrelated with the dependent variable., With respect to the articles discussed here it would imply that knowing whether or not the country was in conflict or not would not tell you anything about the probability of an observation being missing. Missing at random entails that the missingness only depends on the observed part of the dataset and the parameter ϕ. In essence it implies that the probability of a given covariate being missing, depends, and therefore can be explained, by the other non-missing covariates. According to Gelman and Hill (2007, 530) it means that "sufficient information has been collected that we can 'ignore' the [missingness] assignment mechanism". This is a more general assumption than the pure Missing Completely at Random assumption, and in the literature under focus here it is an often universally used - but seldom discussed or justified - assumption. It should be stressed though that random is used here in a statistical sense. That the missingness is random does not imply therefore that it is unpredictable or exogenous to the model under study. Rather, it means that the missingness is probabilistic instead of deterministic (Schafer and Graham, 2002)

$$NMAR : Pr(Y, M|\theta, \phi) = Pr(Y|\theta)Pr(M|Y, \phi) = \prod_{i=1}^{n} Pr(y_i|\theta) \prod_{i=1}^{n} Pr(M|y_i, \phi) \tag{3}$$

Observations are Not Missing at Random if the the missingness depends either on predictors which are not included in the model, or if the missingness depends on the missing value itself. An example of the first sort could be a situation were GDP missingness depends on the quality of the bureaucracy in a country, so that countries with low quality bureaucracies do not report GDP numbers. An example of the second would be a situation were countries with very high infant mortality rates simply hides them or does not report them. Data that is deliberately hidden by governments, e.g. rates of repression, would also be an example of the second kind. NMAR missingness that depend on unobserved predictors can be turned into MAR if these predictors are included. Non-random missingness that depends on the value of the missingness can also modeled by using predictors of e.g. repression. Gelman and Hill (2007, 531) add the cautionary note that "while it can be possible to predict missing values based on the other variables in your dataset (...) this situation can be more complicated in that the nature of the missing-data mechanism may force these predictive models to extrapolate beyond the range of the observed data".

Analysis under MCAR will not produce systematically biased results under a likelihood estimation framework Agresti (2002, 475), failure to deal with NMAR, of the first or second kind, on the other hand will result in biased estimates, and dealing inappropriately with MAR data will as well. The core problem of course is that in general we "cannot be sure whether data really are missing at random, or whether the missingness depends on the unobserved predictors or the missing data themselves" (Gelman and Hill, 2007, 531). In most cases we do not know enough about the missing data mechanism, and analysis under substantial amounts of missing data should therefore be done with the utmost caution. Given that the literature on armed conflict at the moment seem to, at least to some degree, be pushing authors to make more policy recommendations based on their findings this issue becomes even more important.

[Table 1 about here.]

3. SINGLE IMPUTATION METHODS

3.1. *Listwise Deletion*

By far the most common way of dealing with missing data in the quantitative conflict literature is to carry out a "Complete Case" analysis (Little and Rubin, 2002), what is most often referred to as "listwise deletion". It entails analyzing only those units for which there is complete data coverage, and discarding everyone else. The number of units discarded depends then on the "missing data pattern" (Little and Rubin, 2002, 4). Under a correct MCAR assumption the only consequence of listwise deletion is a loss of efficiency, meaning that the probability of making a type 2 error increases somewhat. To avoid discarding precious data, some researchers have opted for "listwise-deletion light": imputing the missing values in the $\mathbf{X}$ vector with mean values, or with zeroes and a new dummy variable which marks all of these original missing variables (see e.g. Collier, Hoeffler and Söderbom, 2008). This method is called zero-order imputation by e.g. (Greene, 2008). These two procedures are analogous, do not add any new information at all, but will reduce the models R^2 (Greene, 2008, 62). It should be noted though that Little and Rubin (2002, 62) end their section on 'Mean Imputation' by simply stating: "This method cannot be recommended".

In many situations, listwise discarding of observations is completely harmless and not inappropriate. The problem, however, is that in the three articles replicated here, and for the most part in the literature at large, this is done without any justification for why we should believe in the MCAR assumption. To be clear: in general listwise deletion is only appropriate under an MCAR assumption (Schafer and Graham, 2002, 155). King, Honaker, Joseph and Scheve (2001, 52) argue that because of listwise deletion "the point estimate in the average political science article is about one standard error farther away from the truth". This leads them to argue that "omitted variable bias is often preferable, if only it and listwise deletion are the options" (King et al., 2001, 52).

In general, data may be MAR but listwise deletion will still induce bias if the discarded sample is not representative of the full sample. This is an important point which is often overlooked in the empirical literature. It implies that even if the missingness can be predicted by the other variables in the model - i.e. the MAR assumption - results will still be biased, as well as inefficient, if the subsample is skewed in any way. In many situation this of course will be the case.

3.2. *Linear Interpolation*

After listwise deletion linear imputation techniques are the most used tools for handling missing data in political science. Linear interpolation entails linearly filling in missing information between observed time points in a time series. Although commonly used, assumptions behind it is seldom discussed. Fearon and Laitin (2003) as an example mention in footnote 36 that data from the World Bank has been linearly interpolated. The assumption one has to buy into is that there is a strictly linear time trend in the data. This might be unproblematic for single missing time points, as when you have data for 1990 and 1992, but not for 1991. When the period between observed data points increases, however, this becomes increasingly more problematic. World Bank development data, which is a very common data source in both political science and economics, have measures for most of their variables for every five years (see e.g. World Bank, 2010).

An extreme example of how problematic this method can be is shown in figure 1. The graph shows education attainment across years in Liberia. The data has been collected and backdated through the use of a survey, and the data series for Liberia covers every year from 1960 to 2009 (on the data see Hegre, Karlsen, Nygård, Strand and Urdal, forthcoming). For most years a linear interpolation for any potentially missing years would be unproblematic, in the extreme case, however, one could end up linearly interpolating the upwards trend which ends in 1994. This is exactly when the most dramatic consequences of the conflict in Liberia kick in, and therefore are likely to be the years with the highest probability of being missing. These years are also likely to be data points of special interest to analysts of both conflict and development.

Linear imputation is usually done either by mean imputation or regression imputation. Under mean imputation a missing value is simply filled in with a full or subsample mean. Under regression imputation, a missing value is imputed with the conditional mean given the other variables. Schafer and Graham (2002) find that mean imputation produces biased estimated for any type of missingness, while regression (conditional) imputation performs somewhat better but still induces bias generally. The only single-imputation model which is unbiased under a MAR or MCAR assumption is what Schafer and Graham (2002) calls a predictive distribution imputation. The methods entails first producing a $\hat{Y_{mis}}$, a conditional mean estimate of the missing value - identical to regression imputation - and then adding a residual error $\hat{X}$. $\hat{X}$ is drawn from a normal distribution with mean 0 and variance estimated by the residual mean square error. For a logistic model Schafer and Graham (2002, 159) recommends calculating the fitted probability $\hat{p}$, then draw u from a random uniform distribution, and set $Y = 1$ if $\hat{u} \leq \hat{p}$ and $Y = 0$ if $\hat{u} > \hat{p}$. Predictive distribution imputation is easily implemented and unbiased under MAR and MCAR[1] - in contrast to the other methods. Predictive imputation, as well as the other methods discussed above, suffer however fatally from undercoverage, which is discussed below.

3.3. *Imputation Uncertainty*

An added problem with use of mean imputation, zero-order imputation and linear imputation techniques is that authors seldom if ever estimate the uncertainty of their imputed values. By and large imputed values are simply treated as known values, without any associated uncertainty. Obviously imputed values are not known, and since they are imputed they are not in any meaningful sense observed either. Treating these imputations with some degree of caution is advisable, and the most systematic way of doing this is simply to derive estimates of the added uncertainty stemming from the imputations. Little and Rubin (2002, ch. 5) suggest getting standard errors through resampling using either Bootstrapping or Jackknife. Both of these techniques are relatively straight forward to implement in most modern statistical packages, as described for example by Good (2006).

Under non-general circumstances single imputation models often yield valid results for means and point estimates, but they severely underestimates the coverage, i.e. they shrink the confidence intervals (Little and Rubin, 2002). Imputed values are by definition guesses, but under single imputation techniques they are not treated as such, instead imputed values are treated as data points with the same precision as observed values. Missing data uncertainty is hard to take into account under single imputation methods, and this leads to undercoverage - in contrast to multiple imputation techniques which explicitly take imputation uncertainty into account. Consider the often used mean imputation or regression imputation techniques. Mean imputation shrinks the data around the mean, while regression imputation shrinks it around the regression line. In both cases point estimates may be correctly estimated, but the uncertainty of the model will almost always have been underestimated - leading to too narrow confidence intervals.

[Figure 1 about here.]

4. MULTIPLE IMPUTATION METHODS

The imputation methods described so far are all single imputation methods. They take one missing observation and replace it with one imputed observation. In specific applications this might be all one really needs, but for most application it is extremely valuable to also have a measure of the imputation *uncertainty*. For multiple imputation this is straight forward. Multiple imputation procedures involve taking one missing observation and replacing it with a vector $D \geq 2$. The variance of D is directly interpretable as the imputation uncertainty. Analysis is run on all the D imputed datasets - so one would estimate a model on two data sets if a missing value was imputed twice - and the parameters of interest are simply averaged over the D estimations (Little and Rubin, 2002). Alternatively, the imputed dataset can be combined and analyzed as one. The combined estimate of a parameter of interest θ is then simply $\bar{\theta} = \frac{1}{D}\sum_{d=1}^{D}\hat{\theta}_d$ (Little and Rubin, 2002, 86). An attractive aspect of multiple imputation techniques, is that the analysis done on the imputed data sets are simply the same analysis one would use in analyzing only one set.

In order to fully use multiple imputation of missing values one has to buy into some core bayesian assumptions, most fundamentally the idea that parameters are random, while data is fixed (see e.g. Gill, 2008, 64). Averaging over parameters would simply not made any sense without this assumption. Multiple imputation works through an algorithm which closely resembles an Expectation Maximization (EM) algorithm. Two steps are repeated until convergence: an imputation step (I) and a posterior predictive step (P), conducting what is called data augmentation (Little and Rubin, 2002; Enders, 2010, ch. 10). The algorithm starts with an imputation step: A missing value is drawn from the conditional distribution of this variable given the other variables and the parameter estimates from the preceding posterior predictive step. In the first run of the algorithm, these parameter estimates are supplied as starting values to the model, e.g. from a maximum likelihood model. Formally: $Y_t^* \sim (Y_{mis}|Y_{obs}, \theta_{t-1}^*)$, where Y_t^* is the value to be imputed, Y_{mis} and Y_{obs} is the missing and observed portion of the data and θ_{t-1}^* are the estimates - mean and covariance matrix to be precise - from the P step.

The posterior predictive step uses Empirical Bayesian to get an estimate of the posterior distribution for a model and its parameters. The full model, including the filled in values from the last Imputation step, is estimated using Monte Carlo methods and a new covariance matrix and a mean vector is drawn. The step is finalized by drawing a new set of parameter estimates, formally: $\theta_t^* \sim p(\theta|Y_{obs}, Y_t^*)$. θ_t^* is the posterior parameters and Y_t^* are the imputed values. These new parameter values are then sent back to the imputation step, and new set of values are drawn from for missing data points conditional on the observed data and the new parameters. This loop is continued until it converges, i.e. until small if any changes occur from one run to the next. In dealing with multivariate missing data - which of course is most common - the imputation step requires the specification of a conditional model for each of the variables with missing data. The posterior predictive step remains the same regardless of the structure of the missingness.

In the imputation step, missing values are filled in from a conditional multivariate normal distribution. It has been shown that this works for both continuous and discreet - including categorical - variables (Little and Rubin, 2002; Schafer and Graham, 2002; Schafer and Olsen, 1998). Even though a multivariate normal is used in the imputation face of a project, this does not in any way constrain which estimators and distributions are used in the analysis / estimation face of the project (For a presentation of potential problems with the multivariate normal approximation, see Cranmer and Gill, 2012). Any type of analysis technique can be used on the imputed data. As the discussion above has implicitly touched on already, in order to perform Multiple imputation one has to assume that missing observations are MAR. This has to be the case, since the model conditions on the observed data when drawing values for the missing cells. In cases where one knows the data is NMAR, the solution would be to specify full bayesian priors for these observations.

5. MISSING DATA IN CONFLICT RESEARCH

Fearon and Laitin (2003), Collier and Hoeffler (2004) and Hegre et al. (2001) are the most cited articles in quantitative literature on the *onset* of civil armed conflict. Although the literature of course has evolved a great deal since these three articles were published, they still provide the backbone of the civil war literature. Hegre et al. (2001) find that "intermediate regimes", regimes falling in between democracies and autocracies, are the most conflict prone, but also that democracies are more stable than autocracies. Fearon and Laitin (2003) argue that the causes of civil war are primarily to be found in factors associated with insurgency, and not with dividing cleavages such as ethnicity. Situations which favor insurgency increase the probability of conflict onset, and countries with weak state institutions, weak militaries or which are generally less developed are therefore more conflict prone. Collier and Hoeffler (2004) argue along similar lines. They formulate two models: one estimating the probability of conflict from a set of grievance factors, such inequality and ethnicity, and the other from a set of greed factors, such opportunity cost for joining a rebelion and opportunities financing rebelion. They find that the greed model outperform the grievance model, and thus conclude that civil armed are fueled by greed and not grievances. Gross Domestic Product (GDP) is a core predictors in both Fearon and Laitin (2003) and in Collier and Hoeffler (2004). The predictors is highly significant in both models, but is given two different interpretations. Fearon and Laitin (2003) argue it proxies state strength, while in Collier and Hoeffler (2004) proxies opportunity cost.

These articles have of course been criticized, extended, elaborated upon and bypassed by an evolving literature, a review of this literature is however far outside the scope of this article so it will largely be ignored.

5.1. *Missing Data*

The literature on armed conflict, and then especially civil armed conflict, is plagued by considerable amounts of missing data. This of course is not unique to conflict research, but is an issue affecting most areas of study to a greater or lesser extent. In general, though, it is to be expected that any literature which depends on country - or sub-country - level data on and from developing countries will experience both missing data and data of questionable character. This is not necessarily problematic. If, however, the questions to be studied are thought to be connected with development, which e.g. civil armed conflict is theorized to be, then it is highly problematic that the probability of missingness seems to be connected with a country's developmental level.

As noted above Fearon and Laitin (2003), Collier and Hoeffler (2004) and Hegre et al. (2001) are the three most cited articles on the onset of civil armed conflict. The core time periods under study in the three articles are, respectively, 1945–1999, 1960–1999 and 1946–1992[2]. In general it has become conventional in the literature on armed conflict to study the period from 1946 onwards. This is probably due to the fact that data on a wide range of issues is to a large extent not available before 1946, and that widely used lists of armed conflicts, such as Gleditsch, Wallensteen, Eriksson, Sollenberg and Strand (2002), and the list compiled by Fearon and Laitin (2003) for their article, only cover the period after the second world war. In limiting the period under study to this period the literature is implicitly - since this is almost never discussed - implying either that their finding are valid for only the period under study, or that time periods are essentially interchangeable. Arguing that findings are only valid for the period under study, might be defensible, but it should nevertheless be noted. Assuming interchangeability on the other hand should much rather be tested, using e.g. change point analysis, than simply assumed (see for example Pang, 2010). Neither implying limited validity nor exchangeability, would mean that every year before and after the period included is missing data.

Using the Gleditsch and Ward (1999) list of independent states, there are 7606 eligible country-years between 1946 and 1999, and 6401 between 1960 and 1999. This a priori excludes the so-called microstates - countries with population below 500.000 - which all three articles appear to do. Essentially this of course is simply creating more missing data without a strong theoretical justification. Of the 7606 eligible country-year between 1946 and 1999 Fearon and Laitin (2003) analyze 6327, 5186 or 5387 depending on which model they estimate. Because of the data structure in the Hegre et al. (2001) article a direct comparison is difficult, but at least 537 observations are discarded due to missingness in the paper. Collier and Hoeffler (2004) do not utilize a more conventional strait country-year approach, instead they shrink their dataset into 5-year intervals, so that the unit of analysis becomes a country-5-year period. In effect this means tossing out 80 % of the population before the analysis has begun. In a replication of that paper Fearon (2005) finds that some of their results are indeed driven by this self-induced missingness.

The tables below give and overview of missingness across the variables used in the three articles under study. Off the cuff it should be noted that there actually is comparably little missingness in these articles. As is clear, many of the variables the authors use are brute geographical factors - whether or not the country is contiguous; essential demographic variables, such as total population, which any state not utterly collapsed collects or has a UN estimate for; or structural factors, such whether or not the country is an oil producer. These are all factors that it is comparably easier to collect data on. Nevertheless, the amount of missingness is substantial.

6. REPLICATION

[Table 2 about here.]

[Table 3 about here.]

[Table 4 about here.]

6.1. *Fearon and Laitin after Imputation*

The amount of missing data across the different covariates in Fearon and Laitin (2003) is given in table 2. They have no missing data on the dependent variable, but there is missing information on the variables measuring regime type, population and development. These are all key explanatory variables in their analysis. Table (sett inn referanse) gives the result of the two models Fearon and Laitin focus the most on. Columns 1 and 3 replicates respectively models 1 and 3 in their results. Columns 2 and 4 gives the results after performing multiple imputation for their missing data.

We build an imputation model were we allow ordinal variables to be treated as continues, if the authors treat them as continues in their analysis. One such variable is the Polity regime classification variable which can take on any integer from -10 to 10. In the imputed data sets therefore this variable can, and do, take on values falling in between the integers from -10 to 10. A regime score of 3.45 is however just as meaningful as a regime score of 3 or 4 – it simply means that the the country would be more democratic than one scoring 3 but less than one scoring 4. The added bonus of treating it as continuous is that this includes more uncertainty in the model. Dichotomous variables are treated as dichotomous however.

Trends over time are very likely to differ across cross sectional units, i.e. across countries, so polynomials of time are included in the model to let trends vary across countries. For variables which are likely to exhibit high degrees of autocorrelation across time, such as GDP levels, lagged and leading values of the variables are included to get a better fit. In order to speed up convergence of the imputation a ridge prior is also included. This prior adds 1 promille of observations with mean and variance equal to the overall mean and variance[check this]. The efficiency gains from this are substantial, but since the amount of added data is so small it will not - or rather is highly unlikely to - induce any bias. Following Graham, Olchowski and Gilreath (2007) we run 20 multiple imputations chains for this as well as for the two following imputations.

By and large the Fearon and Laitin (2003) results are robust to the imputation, which reflects the fact that the degree of missingness in their data is modest. No variables change direction, or lose significance. The size of the effect does change somewhat however for some of the variables. The effect of democracy e.g. is reduces to almost a tenth, from .126 to .015, and the effect of religious fractionalization increases with a little over 40 percent, from .285 to .410. In neither case though are the variables significant in the original or imputed model.

For the Fearon and Laitin (2003) case then the analysis simply underscores the known fact that as the degree of missing data decreases, the possible biased induces by the missingness - ceteris paribus - as well decreases. It should be noted however that some missing data has already been filled in by Fearon and Laitin (2003). They linearly interpolate GDP data from the World Bank to increase the time covered. As discussed above such interpolation is not always advisable.

[Table 5 about here.]

[Table 6 about here.]

[Table 7 about here.]

[Table 8 about here.]

[Table 9 about here.]

[Table 10 about here.]

6.2. *Collier and Hoeffler after Imputation*

The amount of missing in Collier and Hoeffler (2004) is shown in table 5. Comparing this table to table 2 it is clear that there is substantially more missing data in the Collier and Hoeffler than in the Fearon and Laitin article, thus potentially increasing missing data problems.

The imputation model used for Collier and Hoeffler (2004) closely resembles the one used for the Fearon and Laitin analysis. Trends over time for variables are allowed to vary between countries, and for variables with a high degree of autocorrelation both lagged and leading variables are included. To speed up convergence, a ridge prior similar to the one used for the Fearon and Laitin data is also included.

As expected, as the amount of missingness increases, potential bias also increases. The analysis run on the imputed data sets is significantly and substantively different from the original results presented in the Collier and Hoeffler (2004). Almost all variables see some to a lot of change in the estimated effect. In addition specific variables change sign completely, some gain significance and some loose significance.

Tables 6 to 10 shows the replicated and imputed results for a subset of the analysis run by Collier and Hoeffler (2004)[3]. Table 6 replicates what Collier and Hoeffler (2004) label their 'Opportunity model'. We are able to replicate their original results within a small margin. It is unclear why we are unable to do a complete replication, but the differences are ignorable. One of the key results reported by Collier and Hoeffler (2004) is the effect of primary commodities exports on the probability of civil war onset. They argue that in countries that are rich in natural resources it is easier to finance rebelion, and therefore that civil war is more likely. This finding has already been disputed by Fearon (2005), and these results point in that direction as well. The size of the effect of primary commodities export as a share of GDP, drops by a factor of 6.7, and looses all significance. The same is true for the square of this variable which Collier and Hoeffler (2004) include the check for non-linearity of its effect. The effect of male secondary schooling, a measure Collier and Hoeffler (2004) argue for the opportunity cost of joining a rebelion, drops to basically 0, and turns insignificant. In addition to secondary schooling the GDP growth is also included as a proxy for opportunity cost. In the imputed dataset, this variable remains significant but its effect is reduced by a little above 30 percent. Geographic dispersion, a measure of military advantage over a possible rebelion, changes sign, and loses significance. There are changes for all of the other variables as well, but these changes are either not significant or not theoretically important.

Table 7 reports the replicated and imputed results of Collier and Hoeffler (2004) 'Grievance model'.In contrast to the 'Opportunity model', for this model the results stay more or less substantively the same. The only major change is again geographic dispersion, which shifts sign and becomes insignificant. The number of observations included in this model in the original article is 850, for the 'Opportunity cost' model in contrast only 688 observations are kept after listwise deletion. This underscores how the probability that missingness becomes problematic increases dramatically as the number of missing cells increases. Table 8 shows the first set of replicated and imputed results from Collier and Hoeffler (2004)'s combined model. The only different between this model, and the one reported in table 6 is the inclusion of the variable giving the predicted values from the grievance model. One of the key arguments made by Collier and Hoeffler (2004) is that opportunity - or greed - is much more important for understanding civil war onset than grievances. This variable then controls for grievances in the original opportunity model. The effect of primary commodity exports is again substantially decreased, but in this model it remains a significant predictor of civil war. Secondary male education loses significance, as does geographic dispersion. Interestingly, total population, one of the most robust predictors of civil war (Hegre and Sambanis, 2006), also becomes insignificant. Table 9 analogously reports the same set of predictors as table 8 but includes the predicted value from the opportunity model. This model, as expected, as well stays by and large the same.

Lastly, table 10 reports the results from model 3 in their combined model, one of the largest models they run. This specification only retains 479 observations after listwise deletion, and in this subset there are only 32 wars. This model therefore runs a serious risk of having exhausted its degrees of freedom. There are substantial differences between their original results and the results after multiple imputation. Key variables such as male secondary education and primary commodities exports have their effect substantially reduced and both of the become insignificant. The democracy measure, which is negative and insignificant in the original model, remains negative but becomes a significant predictor in the new estimation.

In contrast to the re-estimation of Fearon and Laitin (2003) then, after multiple imputation we see that a number of the key results reported by Collier and Hoeffler (2004) are weakened and many their core theoretical claims are supported any more.

Collier and Hoeffler (2004) argue forcefully that to understand the outbreak of civil war, one needs to understand how rebellions can be financed, and why people chose to join them. After performing multiple imputation however, in many of their specification we no longer find primary commodities exports - their proxy for how difficult it is to finance rebelion - or male secondary schooling - their proxy for opportunity cost - to be significant.

[Figure 2 about here.]

[Figure 3 about here.]

[Table 11 about here.]

[Table 12 about here.]

[Table 13 about here.]

[Table 14 about here.]

6.3. *Hegre er al. after Imputation*

Tables 12 to 14 gives the original and imputed results for Hegre et al. (2001). We use an imputation model similar to the two used above. Trends are allowed to vary between countries. A ridge prior is included to increase the speed of convergence, but in this case the prior is also necessary for dealing high multicollinearity between some of the covariates in the dataset. The ridge prior consists of 9 observations with mean and variance equal to the overall mean and variance of the dataset. Without this prior the imputation algorithm does not converge, the amount of added data is however so small that it is highly unlikely to induce any serious bias. Variables that change slowly over time, such as energy consumption, are included with lagged and leading values in the imputation model.

By and large, there are only minor changes in the estimates after performing multiple imputation. Some of the estimates get slightly weaker - such as proximity to change, and some get slightly stronger - such as energy consumption. None of the covariates change signs, and there are no substantial changes in significance. The only slight exception is the variable measuring time since the country became independent, proximity to independence. This variable significant in one of the Hegre er at specifications, but after performing multiple imputation its effect is reduced by 1.5, and it becomes insignificant. As in Fearon and Laitin (2003) the amount of missing data in Hegre et al. (2001) is modest.

6.4. *Imputation Diagnostics*

Since the original and imputed results did not differ dramatically for the Fearon and Laitin (2003) and Hegre et al. (2001) models we did not perform any checks of the quality of our imputation models. For Collier and Hoeffler (2004) in contrast the results are markedly different, and the plausibility of the imputed data therefore have to be assessed. Figure 2 shows time series plot for observed (black dots) and imputed (red dots) values for male secondary education (left panel) and income inequality (right panel) for Liberia. Education data is missing for the period after 1980 – the same year as the conflict in Liberia for the first time incurs enough battle deaths to be counted as a minor conflict in the UCDP-PRIO conflict database (Gleditsch et al., 2002). The figure shows a steady, almost monotone, increase in education levels in Liberia up until 1980, after that the imputation model picks up the effect of the conflict, as well as incorporating information from other socio-economic indicators in the dataset, and reports a more or less flat trend in education outcomes. The imputation uncertainty, marked by the vertical lines, shows the imputation model is quite agnostic about the absolute level of education in Liberia. This figure can usefully be compared to figure 1 which shows education data for Liberia from another data source (see Hegre et al., forthcoming) which was not used in imputing the Collier and Hoeffler data. The two sources are not directly comparable though, 1 reports the proportion of secondary male education *attainment*, while the Collier and Hoeffler gives the percent *enrolled*. That said, the imputed enrollment values match up fairly well to the observed attainment data. the Imputed value in 1990 is on average a little below 35 percent, while the observed attainment proportion is .35. In 1995 the observed proportion is .39 while the imputed value is still 35 percent.

The right panel of figure 2 shows imputed values for income inequality. In the original dataset income inequality is missing throughout the time period for Liberia so all of the dots in the figure are imputed. The mean of the imputed values falls between 42.5 and 43 for the entire period. This puts inequality levels in Liberia slightly above those in Sierra Leone which seems highly plausible. The uncertainty around this estimate runs from 38 to 46, implying that the imputation model places Liberia somewhere between Nigeria and El Salvador in terms of income inequality. In terms of the raw observations then, the imputations appear to produce highly believable estimates.

The left panel of figure 3 shows the distribution, relative frequency, of the observed and imputed education data. Plotting these two densities against each other is an efficient way of seeing if the imputations produce a radically different distributions of values then what is found in the original variable. Of course though, this can never be evidence of anything other then a *potential* problem. By definition we do not now the distribution of the actual variable, so if the imputed values differ from the observed ones we can never know if this is because the imputed model is right or wrong or if it is because the observed data is skewed in some way. “Obviously we can not expect, *a priori*, that these distributions [the observed and imputed] will be identical as the missing values may differ systematically from the observed values - this is fundamental reason to impute to begin with!”(Honaker, King and Blackwell, 2011, 25).

The density plots in figure 3 show relatively minor differences between the observed and imputed values. The the imputed density covers mean of the observed data, and they have roughly similar coverage. Clearly though the imputed data has more weight - higher density - on some values than the observed data, but this is to be expected. As stated above, there is absolutely no reason a priori to expect, or want for that matter, the densities to be identical. It should also be noted that since the data were imputed under an assumption of Missing at Random, then by definition the difference between the observed and imputed densities for the education data is *explained* by the other variables in the dataset. Abayomi, Gelman and Levy (2008) advice using a Kolmogorov-Smirnov test to assess whether the imputed and observed densities are statistically different. Abayomi, Gelman and Levy (2008, 280) note though that “p-values from these tests are approximate: the imputations are generated from the observed data; thus the empirical distributions of the imputations are not independent of the observed data”. A Kolmogorov-Smirnov test for the education data (as well as several of the other imputed variables) rejects the hypothesis that the distributions are significantly different. The density plots there also suggest that the imputation model is appropriate.

Lastly the left panel of figure 3 shows the results of an overimputation exercise. The method works by sequentially dropping observed values, and then generating several hundred - in contrast to the 10 imputations we run for the full set - imputed values for that observed value, treating as if it was missing. 10 imputations are enough for the purposes of analyzing the data, but generating several hundred imputations as the added advantage of making it possible to construct confidence intervals - represented by the lines running from the points - for the imputations (Honaker, King and Blackwell, 2011). In every case where the confidence interval overlaps the 45 degree line then, the “imputation can confidently predict the true value of the observation” (Honaker, King and Blackwell, 2011, 28). The figure shows overimputation for the education data, and it reveals a model that is extremely capable of predicting the observed values witnessed by the clustering of points narrowly around the 45 degree line. The different shadings marks the proportion of missing values “behind” each imputation, i.e. the number of missing variables for a given row. There does not seem to be any systematic differences in the proportion of missingness and the predictive ability of the model. It should be stressed again that assessing an imputation model is difficult since the true value by definition is unknown. However, all of the diagnostics discussed above indicate that the imputation model used for the Collier and Hoeffler data is both highly plausible and reliable.

7. CONCLUSION

Ignoring missing data can potentially lead to severally biased results, and it will generally lead to inefficient results and results with insufficient coverage. In replicating Fearon and Laitin (2003) and Hegre et al. (2001) we found that the modest amount of missing data did not severely affect the results, and listwise deletion is "safe" in so far that it does not induce bias. It does however affect the coverage of the results. For the Collier and Hoeffler (2004) replication in contrast we find that listwise deletion induces massive bias. Properly handling this missingness leads to substantially different results.

While Multiple Imputation is by now a fairly standard approach for handling missing data, it is neither the only nor always the best approach for handeling missing data. If the data is NMAR, multiple imputation will not be appropriate. In the cases where missing is non-ignorable, it should not be ignored. Instead, it should be modelled appropriately.

Notes

[1]Even though imputing from a predictive distribution is relatively straight forward and easily implemented in statistical software packages it is almost never used in the conflict literature.

[2]Hegre et al. (2001) also estimate a reduced model on the time period 1812–1992

[3]All in all Collier and Hoeffler (2004) run a total of 15 models, and in addition perform several robustness tests. It is not feasible to replicate all of those results, so we concentrate on a few. The ones we focus on were selected before we knew the results from the imputation - in other words, we did not cherry pick the worst cases.

REFERENCES

Abayomi, Kobi, Andrew Gelman and Marc Levy. 2008. "Diagnostics for Multivariate Imputations." *Applied Statistics* 57:273–291.

Agresti, Alan. 2002. *Categorical Data Analysis, 2nd Edition.* Hoboken, NJ: Wiley.

Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–595.

Collier, Paul, Anke Hoeffler and Måns Söderbom. 2008. "Post-conflict risks." *Journal of Peace Research* 45(4):461–478.

Cranmer, Skyler J. and Jeff Gill. 2012. "We Have to Be Discrete About This: A Non-Parametric Imputation Technique for Missing Categorical Data." *British Journal of Political Science* Forthcoming.

Enders, Craig K. 2010. *Applied Missing Data Analysis.* New York: The Guilford Press.

Fearon, James D. 2005. "Primary Commodity Export and Civil War." *Journal of Conflict Resolution* 49(4):483–507.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/hierarchical Models.* Cambridge University Press.

Gill, Jeff. 2008. *Bayesian Methods: A Social and Behavioral Sciences Approach.* Chapman & Hall/CRC.

Gleditsch, Kristian S. and Michael D. Ward. 1999. "A Revised List of Independent States since the Congress of Vienna." *International Interactions* 25(4):393–413.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand. 2002. "Armed Conflict 1946–2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.

Good, Philip I. 2006. *Resampling Methods: A Practical Guide to Data Analysis.* 3rd ed. New York: Birkhauser Boston.

Graham, John W., Allison E. Olchowski and Tamika D. Gilreath. 2007. "How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory." *Prevention Science* 8:206–213.

Greene, William H. 2008. *Econometric Analysis.* 6th ed. Upper Saddle River, NJ: Pearson.

Hegre, Håvard, Joakim Karlsen, Håvard Mokleiv Nygård, Håvard Strand and Henrik Urdal. forthcoming. "Predicting Armed Conflict 2010–2050." *International Studies Quarterly* 2013.

Hegre, Håvard and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." *Journal of Conflict Resolution* 50(4):508–535.

Hegre, Håvard, Tanja Ellingsen, Scott Gates and Nils Petter Gleditsch. 2001. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816–1992." *American Political Science Review* 95(1):33–48.

Honaker, James, Gary King and Michael Blackwell. 2011. "Amelia II: A Program for Missing Data." *Journal of Statistical Software* 45:1–47.

King, Garry, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1):49–69.

Little, Roderick, J. A. and Donald B. Rubin. 2002. *Statistical Analysis With Missing Data*. 2nd ed. Hoboken, NJ: Wiley.

Pang, Xun. 2010. "Modeling Heterogeneity and Serial Correlation in Binary Time-Series Cross-sectional Data: A Bayesian Multilevel Model with AR(p) Errors." *Political Analysis* 18(4):470 – 498.

Schafer, Joseph L. and John W. Graham. 2002. "Missing Data: Our View of the State of the Art." *Psychological Methods* 7:147–177.

Schafer, Joseph L. and Maren K. Olsen. 1998. "Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective." *Multivariate Behavioral Research* 33:545–571.

World Bank. 2010. "World Development Indicators." Washington, DC: The World Bank.

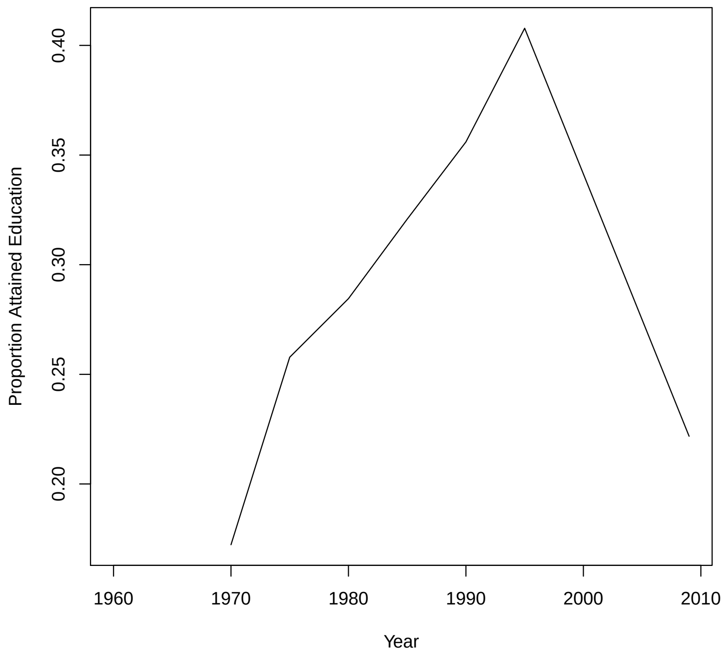

Figure 1: Liberia Education Attainment

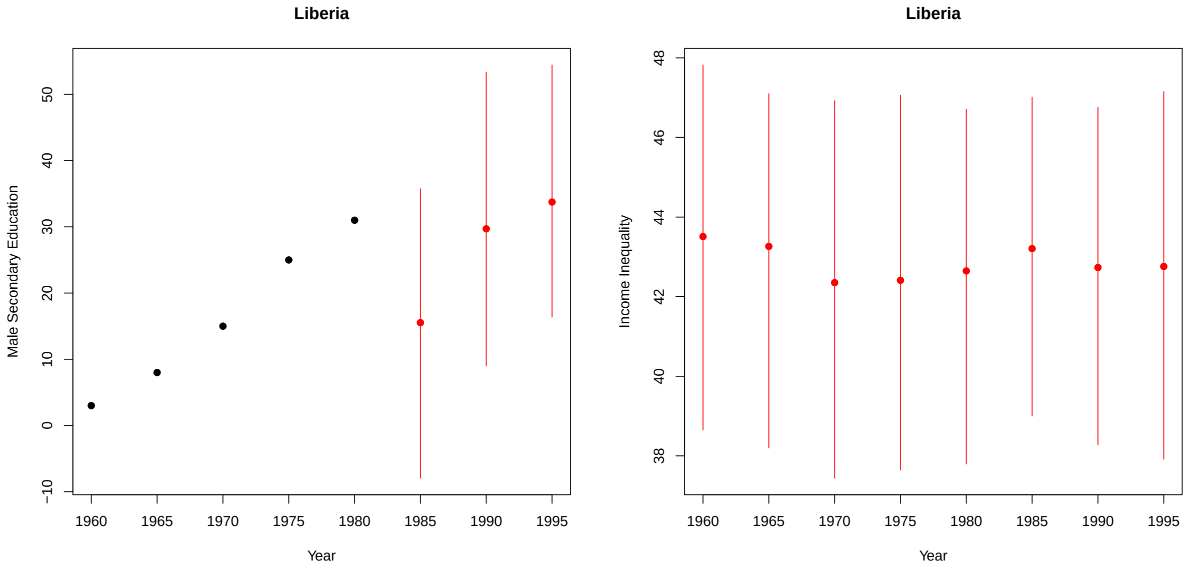

Figure 2: Time Series Plots, Education and Inequality

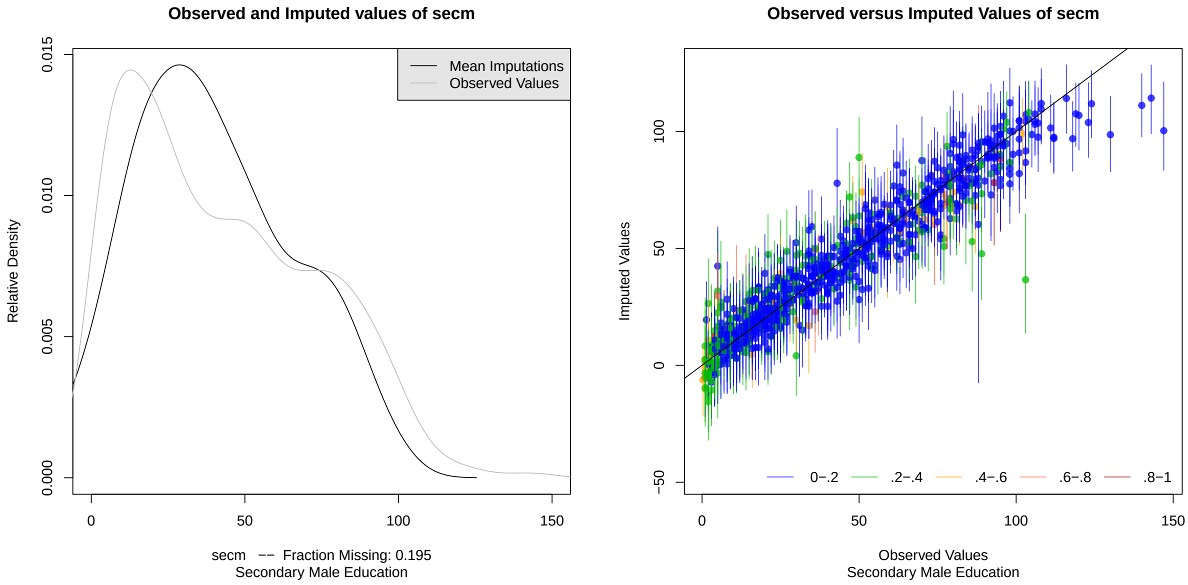

Figure 3: Imputation Diagnostics

Table 1: Missing Mechanisms and Consequences

Mechanism	**Assumption**	**Consequence**	**Solution**
MCAR	Missing is completely uncorrelated with dependent variable	Efficiency loss if discarded	
MAR	Missing explained by included covariates	Efficiency loss if discarded	Model or impute missingness
NMAR 1	Missing depends on unobserved variables	Biased results	Model or impute missingness
NMAR 2	Missing depends on value of dependent variable	Biased results	Impute

Variable	missing
Prior War	0
Per Capita Income	237
log(poulation)	25
log(% mountainous)	0
Noncontiguous state	0
Oil exporter	0
New State	0
Instability	14
Democracy, polity	69
Ethnic Fractionalization	0
Religious Fractionalization	0
Anocracy	69
Democracy	69

Table 2: Missing data in Fearon and Laitin

variables	orig	amelia
Constant	-6.731	-6.816
	(0.748)	(0.751)
Prior War	-0.954	-0.954
	(0.313)	(0.305)
Per Capita Income	-0.344	-0.313
	(0.076)	(0.07)
log(population)	0.263	0.251
	(0.069)	(0.07)
log(mountainous)	0.219	0.236
	(0.084)	(0.083)
Noncontguous state	0.443	0.456
	(0.253)	(0.254)
Oil exporter	0.858	0.851
	(0.255)	(0.248)
New state	1.709	1.688
	(0.332)	(0.321)
Instability	0.618	0.645
	(0.233)	(0.229)
Democracy	0.021	0.014
	(0.017)	(0.016)
Ethnic fractionalization	0.166	0.198
	(0.369)	(0.362)
Religious fractionalization	0.285	0.411
	(0.53)	(0.516)

Table 3: Fearon and Laitin (table 1, model 1). There is no qualitative difference between the original results and the results we obtain when imputing the missing values using MI

variables	orig	amelia
Constant	-7.019	-7.074
	(0.768)	(0.774)
Prior War	-0.916	-0.923
	(0.309)	(0.302)
Per Capita Income	-0.318	-0.289
	(0.076)	(0.07)
log(population)	0.272	0.263
	(0.071)	(0.072)
log(mountainous)	0.199	0.214
	(0.083)	(0.082)
Noncontguous state	0.426	0.432
	(0.252)	(0.251)
Oil exporter	0.751	0.748
	(0.258)	(0.25)
New state	1.658	1.637
	(0.333)	(0.319)
Instability	0.513	0.529
	(0.244)	(0.24)
Ethnic fractionalization	0.164	0.193
	(0.36)	(0.353)
Religious fractionalization	0.326	0.442
	(0.527)	(0.511)
Anocracy	0.521	0.494
	(0.241)	(0.228)
Democracy	0.127	0.015
	(0.317)	(0.312)

Table 4: Fearon and Laitin (table 1, model 3). There is no qualitative difference between the original results and the results we obtain when imputing the missing values using MI

Variables	Missing
War starts	121
Primary commodity exports/GDP	125
(Primary commodity exports/GDP)^2	125
Cold War	0
Male secondary schooling	251
Ln GDP per capita	370
Peace duration	121
Previous war	121
Mountainous terrain	0
Geographic dispersion	160
Ln population	22
Social fractionalization	128
Ethnic fractionalization	104
Religious fractionalization	56
Ethnic dominance	104
Democracy	234
Income Inequality	518
Polarization	280

Table 5: Missing data in Collier and Hoeffler

variables	orig	amelia
Constant	-12.339	-7.196
	(2.491)	(1.608)
Primary commodity exports/GDP	18.149	3.739
	(6.628)	(2.878)
(Primary commodity exports/GDP)^2	-27.445	-4.259
	(14.514)	(4.461)
Cold War	-0.326	0.182
	(0.448)	(0.324)
Male secondary schooling	-0.025	-0.009
	(0.01)	(0.009)
Ln GDP per capita	-0.117	-0.085
	(0.041)	(0.039)
Peace duration	-0.003	-0.002
	(0.001)	(0.001)
Previous war	0.464	0.398
	(0.504)	(0.508)
Mountainous terrain	0.013	0.02
	(0.009)	(0.008)
Geographic dispersion	-2.211	0.61
	(1.098)	(0.672)
Social fractionalization	0	0
	(0)	(0)
Ln population	0.669	0.276
	(0.164)	(0.107)

Table 6: Collier and Hoeffler (table 3, model 1). There are substantial differences between the original results and the results we obtain when imputing the missing values using MI.

variables	orig	amelia
Constant	-5.07	-5.683
	(1.679)	(1.356)
Ethnic fractionalization	0.01	0.012
	(0.006)	(0.006)
Religious fractionalization	-0.003	-0.003
	(0.007)	(0.006)
Polarization	-3.067	-2.33
	(7.022)	(8.171)
Ethnic dominance	0.414	0.181
	(0.482)	(0.572)
Democracy	-0.109	-0.118
	(0.04)	(0.043)
Peace duration	-0.004	-0.002
	(0.001)	(0.001)
Mountainous terrain	0.011	0.017
	(0.007)	(0.007)
Geographic dispersion	-0.509	0.46
	(0.769)	(0.708)
Ln population	0.221	0.216
	(0.09)	(0.079)

Table 7: Collier and Hoeffler (table 4, model 1). There is no qualitative difference between the original results and the results we obtain when imputing the missing values using MI.

variables	orig	amelia
Constant	-7.907	-4.768
	(3.442)	(2.775)
Primary commodity exports/GDP	18.192	10.893
	(6.8)	(3.669)
(Primary commodity exports/GDP)^2	-28.264	-16.383
	(14.854)	(6.742)
Cold War	-0.33	-0.23
	(0.467)	(0.412)
Male secondary schooling	-0.022	-0.009
	(0.011)	(0.009)
Ln GDP per capita	-0.107	-0.117
	(0.04)	(0.038)
Peace duration	0.001	0.001
	(0.002)	(0.002)
Previous war	0.502	0.529
	(0.536)	(0.39)
Mountainous terrain	0.004	0.003
	(0.01)	(0.008)
Geographic dispersion	-2.054	-0.355
	(1.154)	(0.903)
Ln population	0.468	0.227
	(0.205)	(0.157)
Social fractionalization	0	0
	(0)	(0)
Grievance predicted values	0.775	0.841
	(0.386)	(0.304)

Table 8: Collier and Hoeffler (table 5, model 1). There are substantial differences between the original results and the results we obtain when imputing the missing values using MI.

variables	orig	amelia
Constant	0.33	0.169
	(2.314)	(2.309)
Peace duration	0.001	0.001
	(0.001)	(0.001)
Mountainous terrain	0.001	0.001
	(0.008)	(0.008)
Geographic dispersion	0.053	0.127
	(1.089)	(1.079)
Ln population	-0.022	-0.014
	(0.125)	(0.125)
Ethnic fractinalization	0.008	0.008
	(0.007)	(0.007)
Religious fractionalization	-0.005	-0.005
	(0.008)	(0.008)
Polarization	-9.338	-9.638
	(9.355)	(9.391)
Ethnic dominance	1.21	1.222
	(0.677)	(0.677)
Democracy	-0.036	-0.04
	(0.048)	(0.048)
Greed, predicted values	1.044	1.051
	(0.195)	(0.196)

Table 9: Collier and Hoeffler (table 5, model 2). There is no qualitative difference between the original results and the results we obtain when imputing the missing values using MI.

variables	orig	amelia
Constant	-18.477	-7.633
	(5.57)	(2.216)
Primary commodity exports/GDP	35.314	3.024
	(9.983)	(2.866)
(Primary commodity exports/GDP)^2	-64.752	-3.979
	(20.893)	(4.371)
Cold War	-1.109	0.188
	(0.863)	(0.33)
Male secondary schooling	-0.03	-0.003
	(0.013)	(0.01)
Ln GDP per capita	-0.042	-0.078
	(0.053)	(0.039)
Peace duration	0.001	-0.002
	(0.002)	(0.001)
Previous war	0.723	0.402
	(0.705)	(0.54)
Mountainous terrain	0.004	0.018
	(0.011)	(0.007)
Geographic dispersion	-4.229	0.273
	(1.5)	(0.879)
Ln population	0.91	0.275
	(0.249)	(0.129)
Social fractionalization	-0.001	0
	(0)	(0)
Ethnic fractinalization	0.039	0.008
	(0.019)	(0.012)
Religious fractionalization	0.013	-0.006
	(0.02)	(0.014)
Polarization	-24.56	-2.521
	(16.181)	(9.857)
Ethnic dominance	2.079	0.365
	(1.167)	(0.659)
Democracy	-0.018	-0.11
	(0.058)	(0.05)
Income inequality	0.028	0.017
	(0.031)	(0.021)

Table 10: Collier and Hoeffler (table 5, model 3). There are substantial differences between the original results and the results we obtain when imputing the missing values using MI.

Variable	Missing
Proximity to regime change	126
Democracy	0
Democracy squared	0
Proximity to civil war	0
Proximity to independence	0
International war in country	0
Neighboring civil war	0
Development	267
Development squared	267
Ethnic heterogenity	166
Proximity to smal democratization	126
Proximity to large democratization	125
Proximity to small autocratization	126
Proximity to large autocratization	127
Proximity to regime change	125

Table 11: Missing data in Hegre et al

variables	orig	amelia
Proximity to regime change	1.268	1.138
	(0.406)	(0.413)
Democracy	-0.002	-0.002
	(0.024)	(0.024)
Democracy squared	-0.012	-0.013
	(0.005)	(0.005)
Proximity to civil war	1.159	1.149
	(0.341)	(0.318)
Proximity to independence	1.515	0.324
	(0.956)	(0.901)
International war in country	0.858	0.693
	(0.667)	(0.662)
Neighboring war	0.097	0.087
	(0.326)	(0.313)
Development	-0.481	-0.518
	(0.171)	(0.171)
Development squared	-0.066	-0.081
	(0.039)	(0.039)
Ethnic heterogenity	0.797	0.74
	(0.408)	(0.394)

Table 12: Hegre (table 2, 1946-92). There are some minor differences between the original results and the results we obtain when imputing the missing values using MI.

variables	orig	amelia
Proximity to regime change	1.268	1.138
	(0.406)	(0.413)
Democracy	-0.002	-0.002
	(0.024)	(0.024)
Democracy squared	-0.012	-0.013
	(0.005)	(0.005)
Proximity to civil war	1.159	1.149
	(0.341)	(0.318)
Proximity to independence	1.515	0.324
	(0.956)	(0.901)
International war in country	0.858	0.693
	(0.667)	(0.662)
Neighboring war	0.097	0.087
	(0.326)	(0.313)
Development	-0.481	-0.518
	(0.171)	(0.171)
Development squared	-0.066	-0.081
	(0.039)	(0.039)
Ethnic heterogenity	0.797	0.74
	(0.408)	(0.394)

Table 13: Hegre (table 2, 1946-92). There are some minor differences between the original results and the results we obtain when imputing the missing values using MI.

variables	orig	amelia
Proximity to small democratization	1.539	1.363
	(0.64)	(0.665)
Proximity to large democratization	1.22	0.837
	(0.9)	(0.972)
Proximity to small autocratization	1.22	0.969
	(0.713)	(0.755)
Proximity to large autocratization	2.634	2.483
	(0.701)	(0.711)
Proximity to regime change	0.288	0.515
	(0.65)	(0.633)
Democracy	0.002	0.004
	(0.027)	(0.025)
Democracy squared	-0.012	-0.013
	(0.005)	(0.005)
Proximity to civil war	1.141	1.115
	(0.332)	(0.314)
Proximity to independence	2.517	0.961
	(1.054)	(1)
International war in country	0.854	0.698
	(0.621)	(0.632)
Neighboring war	0.165	0.118
	(0.326)	(0.309)
Development	-0.477	-0.514
	(0.172)	(0.17)
Development squared	-0.066	-0.079
	(0.039)	(0.039)
Ethnic heterogenity	0.8	0.731
	(0.414)	(0.402)

Table 14: Hegre (table 3, 1946-92). There are some minor differences between the original results and the results we obtain when imputing the missing values using MI.