Who are the ‘don't knows’? The problem of missingness in surveys in conflict-affected regions, with illustrations from 10 post-Soviet settings.

Francis Naylor and John O’Loughlin

Institute of Behavioral Science
University of Colorado-Boulder
Boulder, Co. 80309-0483
Email: johno@colorado.edu
December 2019

Forthcoming in **EUROPE-ASIA STUDIES**

Researchers who carry out survey work either in obviously sensitive settings such as hospitals or prisons, in conflict zones where violence is either continuing or recently ended, or who ask penetrating questions as part of a survey are usually aware that the answers that they get might be distorted. Answers to such questions might be intentionally misleading, unintentionally misleading (such as those coming from persons in ‘polite societies’ who give an answer that they think that the researcher would like to hear, known as social desirability bias; De Maio 1984 and Javeline, 1999) or offer a response (typically ‘don’t know’) that avoids giving a truthful answer. Such possible biases have in part motivated the use of experimental designs in surveys including the use of item-lists and endorsement options (Blair et al, 2014: Fair et al. 2012, 2014; Shapiro and Fair, 2010). It is increasingly common in the social sciences to see surveys with such checks and corrections to reduce misleading answers. The possible confounding effects of patterning in the missing (‘don’t know’ and refusal) responses have generated suggestions for imputing the missing values especially after King et al (2001) introduced new procedures and evaluated them with simulated data and empirical examples.

The problem of mis-leading answers has been evident from the beginnings of polling about eighty years ago in the United States. It was highlighted by the repeated underestimation of the votes for racist candidates (e.g. George Wallace in the 1968 US presidential race) and other racially-motivated votes such as support for referendums that prohibit use by undocumented residents of governmental services, like Proposition 187 in California in 1994. Underestimation is clear when pre-election survey numbers are compared to actual vote proportions. In recent years, doubts about the high polling numbers for Vladimir Putin’s popularity at about 85-88% in Russia have been checked but have found to be quite accurate by using item-list surveys that show an inflation factor of only 5-8%

(Frye et al, 2017). The sensitivity of survey questions must be deliberately and preemptively considered at every stage of the research process. This includes survey design and testing. Measuring the sensitivity of a question, for both respondents and interviewers, can help to validate or contextualize hypothesized missing data mechanisms later (Kaplan and Yu, 2015). Furthermore, the use of pretesting practices, such as performing a small sample pilot survey, may reveal unanticipated sensitivities which can be mitigated by rewriting questions or retraining/substituting interviewers.

In this paper, we focus on one side of the biased or misleading answer problem, that of answer avoidance through the use of 'I don't know' or 'It's hard to say' responses to enumerators who pose the sensitive question, hereafter the 'don't know' (DK) problem. We leave aside in this paper the related matter of the difference between a 'don't know' response and an outright refusal to answer the question, though we plan to probe that distinction in future work. (In the post-Soviet surveys discussed in the empirical parts of this paper, the refusal rates are very small, typically under 2-3%). DK's are never prompted but always recorded by the interviewer and are entered by the interviewer into the tablet or paper questionnaire for coding as 'missing data'. These missing data are conventionally dropped through listwise or pairwise deletion in the analysis. Unless the researcher engages in some sort of imputation of the missing data, the results are potentially incorrect. This inappropriate model fitting is especially likely if the missing data are clustered in the nature of the responses of a key demographic group (e.g. by age or gender), or in our case, of one of the nationalities in surveys from a conflict-affected area, a conflict in which the group was involved as participants, victims or perpetrators. This builds upon the work of Lall (2016) by investigating three common missing data treatments simultaneously - listwise deletion, pairwise deletion (a common default in statistical programs), and multiple imputation (MI). We further show that even MI can be erroneous in missing not at random cases (elaborated upon in a further section) such as Abkhazia, emphasizing that there is no viable substitute offered for an understanding of context.

We report preliminary work on the DK problem in ten surveys in conflict-affected regions of the former Soviet Union that were directed by the second author over the decade 2005-2014 in the North Caucasus of Russia (2005), in the contested oblasts of South-east Ukraine (2014), in the republic of Crimea annexed to Russia (2014) and in de facto unrecognized republics of Transnistria (2010 and 2014), Abkhazia (2010 and 2014), South Ossetia (2010 and 2014) and Nagorno-Karabakh (2011). Each of these surveys focused on the attitudes and behaviors of representative samples of residents of these regions whose lives were dislocated by violence and who were living with its consequences in the form of closed borders, hostility and discrimination towards minorities, trauma from war experiences, damaged infrastructures and significant hostility and suspicion between nationalities.

The post-Soviet survey were conducted in the same manner. Door-to-door interviews in local languages by experienced enumerators using random route method sampling generated large and representative samples. Many questions that focused on a large range of post-war attitudes were potentially sensitive; the DK option had to be always volunteered by the respondent. For comparative purposes, we choose five potentially sensitive questions about local political figures and postwar attitudes asked in almost all of the sites; we matched the answers with 10 demographic variables (standard across all samples) in this study. The modeling results for many of these questions have been published in various outlets (see for example, Bakke et al., 2014, 2017: O'Loughlin, Toal and Kolosov, 2017. Of particular note, our analysis questions the missing data treatment undertaken in two papers using the same data, with the Abkhazia data having been unadvisedly treated with multiple imputation (Bakke et al., 2014), as well as being unadvisedly outright dropped (Bakke et al., 2018).

We present tests of the assumptions of missing data values, identify the respective mechanisms for each sample, and compare the results (coefficients) in simple models using different methods of handling the missing data. An extended analysis of the South-East Ukraine 2014 sample illustrates some possible corrections to commonly-used modeling procedures. The samples and the variables available for each survey, as well as the texts of the sensitive questions, are presented in Appendix 1. As can be seen in Table 1, some of the data sets have more than one in five respondents as DK and for some sub-samples (such as women of Russian nationality in south-east Ukraine), the ratio reaches almost 50% DK.

Table 1 about here

The Missing Data Mechanisms

When a respondent gives a 'don't know' (DK) reply to a question, the reasons for the response often cannot be definitively known even if the interviewer engages in follow-up probing of that answer. Even then, avoidance of a question can be parlayed into a vague answer or a refusal to engage with the enumerator if the topic is particularly sensitive (Bradburn et al., 1978). But there are also two legitimate reasons why a DK response could be given. A lack of information is a genuine reason for a DK answer if the question is about a topic, event or person with which the respondent is unfamiliar. The literature on missing values documents that people with a lower educational level and women (especially in traditional societies) are more likely to answer 'don't know' (Schuman and Presser, 1996). In other studies about weather patterns and climate change, a definite urban-rural difference can be seen in the ratio of DK answers since rainfall trends and seasonal changes are not as important to those whose livelihoods do not depend on agriculture. A second reason for a legitimate 'don't know' is the uncertainty produced by a volume of information about the subject on which the respondent is well informed and cannot decide between two or more plausible options. An example of such a question would be one about whether the current US government will take action to combat global climate change. Even a person familiar with the subject would find it hard to anticipate an almost unknowable outcome. For both of these reasons for giving a DK, it is generally not expected that its probability not vary significantly across demographic groups.

The growing understanding of the importance of the distribution of DK responses in social surveys has led to a sizable literature on procedures and methods for examining such responses. Some survey enumerations include a 'don't know' option as one of the possible answers to a question. A filtered option is often phrased as 'do you have an opinion about…? If so, could you tell me….'); in these filtered question formats, it has been shown that the DK ratio rises by a significant amount over the standard format. For US respondents asked about a variety of foreign policy issues, including knowledge of the topics and opinion about them, Schumann and Presser (1996) showed that the DK ratio varied markedly, from 23% to 45% about Israel-Arab relations and from 63% to 88% about a coup in Portugal. The authors distinguished between 'floaters' and others, where 'floaters' flit among positions which are often incompatible. A DK answer on the basis of a general (geo)political orientation suggests that they do not know much about the specifics of the topic; 'floaters are answering on the basis of a general underlying orientation that is tapped vaguely' (Schumann and Presser, 1996. 132; also see Sturgis et al, 2014). Importantly, floaters do not have consistent positions but jump from answer to answer in an almost random way. While the filtered question format (asking first if the respondent has an opinion on the topic) is useful for identifying floaters, it is not the preferred one for social, political, health and educational surveys (De Leeuw et al., 2016).

While not illegitimate in the sense of protecting oneself against possible reprisal or revealing a irresponsible or embarrassing attitude, a DK answer in the case of a respondent who has had a particular experience or a firm, but non-vocalized, belief is a greater worry in survey research. The subject of missing values has been widely studied in public health (Johnson and Van Vijver, 2003), in educational outcome studies (Baraldi and Enders, 2010; Enders, 2010) and in criminology (Tourangeau and Yan, 2007). In conflict studies, especially in particularly sensitive war environments such as Afghanistan (Blair et al. 2014) as well as in non-democratic contexts where corruption and criminal state roles are manifestly evident such as Nicaragua (Gonzalez-Ocantos et al, 2012), political scientists are trying both to reduce the DK responses and to gain reliable results by implementation of field experimental methods. In earlier work about contextual effects, most of the estimation and imputation methodologies were directed to electoral and political preferences. Arguably, the effort to work in difficult survey environments is both essential to understand the motivations for violence and helpful in countering it (Shapiro and Fair, 2010).

Tackling the Missing Data Problem

The use of listwise and pairwise deletion in data analysis is predicated on the assumption that the missing values are randomly distributed across the full sample. We use the terms and definitions of Little and Rubin (2002) for missing data analysis. The mechanism known as MCAR (Missing Completely at Random) sees no relationship between the values of other variables, including the outcome values (Y), and the values of X and the missingness propensity for X, the predictor of interest. In other words, the missingness pattern in the data is not systematic and can be regarded as random across the complete data set since no relationship between the missing values and the other values for a particular variable are observed. Examples of a MCAR situation are participants dropping out of a panel study due to a move or to incorrect data entry and miscoding (Baraldi and Enders, 2010). While standing as a high bar for a missing data approach and subsequent statistical modeling, the MCAR assumption underpins much of what survey researchers assume is the structure of their data. Since Little's (1988) test for violations of the MCAR assumption is now readily available, close consideration of this MCAR belief including possible violations of MCAR should be carefully considered before any statistical analysis. In the test data sets that we analyze in this paper, the sample of questions for the 'de facto' state of Transnistria in 2014 is MCAR. Our comparisons of various missing data procedures for this sample do not show much difference in the regression coefficients, as one would expect from the MCAR mechanism.

The other end of the missingness spectrum from MCAR is MNAR (Missing Not at Random), seen in a relationship between the missing values of the variable of interest and the key outcome measure. As stated by Baraldi and Enders (2010, 8), MNAR are 'data that are missing based on the would-be values of the missing scores.' This condition is 'non-ignorable' and an imputation of the missing data should be attempted. Allison (2002) indicates that there should be good *a priori* knowledge of the mechanism for imputation with the particular study's context guiding the model specifications. Obvious examples of MNAR are respondents who are poor readers giving a DK answer to questions that ask about their reading comprehension or drug users giving a DK answer to a question about risky drug consumption. In our study, we will examine the case of Abkhazia 2010 where patterns in the data follow a MNAR trend, as indicated by the statistical tests. In Abkhazia, a high ratio of Georgians (a marginalized minority) provided a DK response for the sensitive questions, especially about the right of return which particularly affects this nationality. As we show below, imputing a response for the missing values in Abkhazia on the basis of demographic or other attitudinal responses is highly problematic because the DK answer was likely to be a preference for refugee return, with listwise and pairwise deletion essentially representing a particularly vicious case of sampling on the dependent variable.

MAR (Missing At Random) is somewhat misnamed since it occurs when the pattern of missing values is related to the values of other variables in the analysis, but not to the underlying values of the incomplete variable. Sweet and Grace-Martin (2008) state that MAR is an unfortunate misnomer, and recommend that it should properly be labelled as 'conditionally missing at random.' Data for Y (dependent variable) are missing at random if the probability of missing data is unrelated to the Y value. Allison (2002) gives the example of a MAR condition when the probability of missing reported income values depends on a person's marital status but within each marital status category (single, divorced, etc.), the probability of a missing reported income value is unrelated to income. If the reported income missingness is related to the value of a person's income, then the data are not missing at random, controlling for other variables.

Data that are MAR show a systematic rather than a random pattern of missingness that can be predicted by other observed variables and does not depend on any unobserved variables. If missingness can be predicted from the observed variables, then multiple imputation (MI) is appropriate. in our study, we examine the case of another 'de facto' state, Nagorno-Karabakh, in 2010 which we identify as being MAR based on CDM analysis (see below). The advice given by Grace-Martin (n.d.) is helpful in distinguishing between MAR and MNAR. 'The first thing in diagnosing randomness of the missing data is to *use your substantive scientific knowledge of the data and your field*. The more sensitive the issue, the less likely people are to tell you.' (italics in original).

If each variable in a survey is missing just a few percentages of the observations, there is cumulatively a big drop in the cases available for analysis using listwise deletion. For a survey of 1000 people with 5% missing for 20 variables, the sample size drops to 360 in such a listwise deletion procedure (Allison, 2002). Earlier methods of imputing the missing values such as inserting the mean values for the particular variable or a single imputed value based on the coefficients from a regression have fallen out of favor since they are as likely to introduce as much bias into the data as a listwise deletion. Pairwise deletion (dropping only the cases with a missing value on the variables in a particular model) will likely yield slightly more cases but the same uncertainty about whether there is some underlying relationship between the missing values and characteristics of persons in the survey will persist. Preferences for maximum likelihood methods and multiple imputations, often based on 20-100 imputations, now dominate the literature (Little and Rubin, 2002; Enders, 2010, Young 2012). Imputation of dependent variables 'is essential for getting unbiased estimates of the regression coefficients' (Allison, 2002: 52). However, imputation might also introduce biased estimates for the dependent variable, as we will show below for Abkhazia in 2010. In the statistical packages, SPSS, SAS and Stata, the default MI method is based on Markov Chain Monte Carlo (MCMC) ('fully conditional explanation') approaches developed by Rubin (see Little & Rubin, 2002). The imputed missing response from a respondent for an item in the survey is based on the respondent's other responses and responses of other subjects similar to the respondent who gave a DK answer. As a rule of thumb, Bagheri et al. (2014) recommends that imputation be conducted if the missingness on an item is under 50% and other variables have the capacity to predict missingness. In this paper, we compare the coefficients of a simple model with pairwise and listwise deletions to the coefficients from a multiple imputation approach for three conflict-affected contexts.

Conflict-Affected Societies in the Post Soviet Union – Missingness in Survey Data.

All ten surveys examined in this paper were conducted over a ten year period after the collapse of the Soviet Union and all samples were drawn from regions which has seen conflicts around minority demands for separatism and attempts by the post-Soviet republics (Russia, Georgia, Azerbaijan, Moldova and Ukraine) to prevent such independence. The conflicts varied in length (from the North Caucasian wars of over 25 years to the short weeks-long conflict in Transnistria in 1992), intensity in displacements and numbers of victims (with Transnistria at one end of the scale and Abkhazia, Nagorno-Karabakh as well as the North Caucasus at the other), and international attention and subsequent impact on relations between major powers. In all cases, the surveys were conducted in communities that were caught up in the conflict or very close to the concentration of fighting that resulted in refugees arriving in the survey sites. Personal memories of war and the varied experiences of its impacts on respondents and their communities affected attitudes towards reconciliation efforts and preferences for more permanent ceasefire arrangements and political structures.

The ten surveys contain dozens of sensitive questions about post-conflict attitudes. Though many are individually targeted to the local context (e.g. attitudes about Islamist movements in the North Caucasus), other questions are comparative across the ten surveys. Such commonalities allow a context-sensitive analysis of key topics in post-Soviet ethnic relations, especially around ethno-territorial demands. For purposes of illustrating the vexing, but inevitable, problems of dealing with missingness in such survey responses, we chose five variables for examination here. It must be stressed that these variables do not represent the full range of indicators and it is possible that what we identify as MAR data or MCAR missingness in a data set might be different if the full set of outcome variables for each survey were examined. Two measures (trust in the local President and trust in the Government) are indicators of attitudes about state competence and legitimacy and show a wide range of responses across the demographic categories. Two other items quantify the post-Soviet dislocations in the respective survey sites. As previous work (Toal and O'Loughlin, 2016) has shown, the generic question 'Was the end of the Soviet Union a right or a wrong step?' is highly predictive of general beliefs about developments since 1991 and the respondent's outlook on national trajectories relative to the Soviet legacy. The question about the right of refugees to return to the homes from which they fled or were expelled is particularly sensitive in light of the history of displacements and the ethnocratic nature of post-conflict governance in these locales. The fifth question, about the possibility of getting a job without the right kind of ethnic background and connections, also targets the possible dissatisfaction of minority nationalities with the perceived discriminatory practices of the majority.

Similar to the selection of key survey attitudes was the choice of demographic predictors for our missingness analysis. We include nine measures with respondents' current mood (a generic post-Soviet question frequently asked to gauge psychological outlook) and expectations about material prospects (in a two year forecast) added to the usual age, gender, education, ethnicity and current material status. The full list of items in this current study, their proportions across the ten sites and the text of the outcome attitudinal questions are indicated in Appendix 1.

The overall ratio of missing values for the five key variables is indicated in Table 1. These overall ratios, though, hide significant variations across demographic sub-categories, especially across ethnic and religious ones. Since the wars were largely predicated on ethnic beliefs about territoriality and exclusive claims to homelands, it is unsurprising that such variations underpin the missingness scores. Missing ratios near 20% are common for these sensitive variables and it is expected that the missing values are disproportionately clustered in certain demographic categories. One could assume a MAR mechanism in play for all of the samples and engage in multiple imputation of the missing values. Listwise deletion, the default approach to missingness in survey analysis, would both significantly reduce the sample size and likely generate inaccurate or biased coefficients.

Figure 1 about here

Tests for deviations from the MCAR expectation are a necessary first step in this analysis of missingness. Examples of demographic sub-group trends in missingness in the samples are shown in Figures 1-3. As is clearly visible, the ratios are high (over one-third of respondents) for sub-groups that are cross-tabulated, with women generally more likely to give a DK answer. This gender gap is well documented in the literature on missingness and is larger in traditional societies where women are less likely to be occupied in waged labor outside the home and cultural expectations do not promote engagement with political topics. Plotting the DK ratios by predictor variables is a simple but valuable first step in highlighting possible missingness trends in the data that can be complemented by the computation of a dummy variable, represented as a binary value of missing vs. not missing, and a series of cross tabulations, including a computed Chi-square, with other variables in the analysis.

Figure 2 about here

The Abkhazia 2010 survey reflects a MNAR situation (as shown in Table 2) with very high missing ratios for ethnic Georgians. This large DK ratio was the subject of a multiple imputation based on cupola methods in the article by Bakke *et al* (2014) but as we discuss below, this imputation was based on statistical principles and almost certainly underestimates the negative Georgian responses, given the precarity of Georgians in Abkhazia. The difficult double-peripheralization of that group by both the state of Georgia and the *de facto* authorities of Abkhazia results in high sensitivity regarding the subject of the possible return of the majority of the Georgian minority that were displaced as a result of the wars of the early 1990s (Toal and Frichova Grono, 2011).

Table 2 about here

Based on Little's test (Table 2), the graphs for Transnistria in 2014 with a smaller range of missing values among the sub-groups and for south-east Ukraine (Figure 2) illustrate MAR distributions. South-east Ukraine has seen significant unrest and protests after the Maidan protests of late 2013-early 2014, and borders on both the Crimean peninsula and the active war zone of the Donbas, becoming a territorially- contested area with significant Russian populations. Russians are in a majority or are part of the dominant ethnic coalitions that support governments elsewhere in the study sites but in south-east Ukraine, as a minority population, their missing ratios are higher than the majority Ukrainians and also show a significant gender gap.

Figure 3 about here

Statistical Tests for Missingness and Results

A statistical test developed in the field of biostatistics for examining the missing data mechanism of a data set tests a null hypothesis that the missing data mechanism is MCAR (Little 1988). Rejection of the null hypothesis is indicative of data which is MAR or even possibly MNAR. Little's test compares all of the missing patterns and their associated outputs (respondent answers, i.e., yes/no) with each other; divergence from a normal distribution within the test leads to rejection of the null hypothesis. If Little's test indicates a rejection of this MCAR null hypothesis (a p value <.05, for example), the test can be extended to test covariate-dependent missingness (CDM). This method examines the relationship between a baseline covariate (e.g. gender or age) and missingness. The CDM null hypothesis is that the independent variable under CDM consideration is not related to the missingness of the dependent variable. Rejection of this null hypothesis affirms that the value of the independent variable is related to the occurrence of a missing record for the dependent variable. These tests are available in major statistical software packages. For this paper, the tests were performed using the command *mcartest* in the statistical software package Stata 13 (Li, 2013). A complementary approach to aid interpretation can be to run a multinomial logistic regression with missingness as an outcome.

We performed Little's test of MCAR on the sensitive questions in all ten surveys. The results in Table 2 indicate a consistent rejection of the MCAR assumption in all cases, except Transnistria in 2014. Overall, missing values in each of the five sensitive questions has some un-ignorable mechanism behind missingness. The follow-up tests of CDM indicate the covariates under consideration are significant across the sensitive questions. In the case of a variable that causes the CDM null hypothesis to be rejected, this variable must be further examined to understand the patterns of missingness. For example, in Table 2, the South Ossetia 2014 results show that it is not necessary across all five sensitive questions to account for the effect of gender, but it is necessary to account for the effect of age. On the other hand, in the Transnistria 2010 data, it is necessary to account for gender, age, and ethnicity across all five questions. CDM can be a factor in any subset of sensitive questions but not necessarily for every question considered to be sensitive.

Table 2 about here

The fact that age and gender are significant in the CDM analysis of survey data can be understood as a legitimate 'don't know' response mechanism for specific cultural and historical contexts, such as South Ossetia and Transnistria. The fall of the Soviet Union question is more likely to generate a 'don't know' answer by younger respondents as they lack personal experience of the time period of the Soviet Union. The question about connections necessary for getting a job is more likely to elicit a DK answer for older people because older respondents are more likely to be out of the workforce,.

While by no means a complete representation of surveys in conflict-affected regions, the evidence in Table 2 suggests that the MCAR assumptions by researchers that underpin most such survey work are unlikely to be met and that MAR mechanisms are more likely. In such circumstances, multiple imputation approaches are suggested to reduce the bias that would result from ignoring the patterns of missingness and engaging in the listwise or pairwise deletion that is common practice.

Modeling Options for Missing Data

In the context of varied missing mechanisms, several missing data treatments are mooted. Common treatments include listwise deletion, pairwise deletion, and Multiple Imputation (MI). The logistic MI used here involves over 100 imputations based upon distributions associated with the demographic variables of age, gender, ethnicity, mood, and education. We chose these predictors since they have been shown to be related to postwar attitudes in these post-Soviet contexts. (Since this is only a demonstration in comparison,) we do not suggest that it is a theory-driven model. Pairwise deletion uses available cases for each particular analysis, whereas listwise eliminates those cases which are not complete overall. The imputations fill in missing data based upon distributions within the demographic variables. These treatments of missingness were compared on three sample data sets that we chose from the ten surveys summarized above. The analysis of missingness showed that the respective datasets of Transnistria 2014 is MCAR, Nagorno-Karabakh 2011 is MAR and Abkhazia 2010 is MNAR. Listwise and pairwise deletion can be appropriately used for an MCAR distribution, while Multiple Imputation is typically applied when a MAR assumption is met, based on Little and CDM tests. Neither of these treatments are considered appropriate for cases of MNAR. After applying these respective treatments, a simple logistic model for the demographic variables (age, gender, ethnicity, mood, and education) allows comparison of the coefficients. Further analyses, such as regression, can be performed in the context of a multiple imputation treatment by pooling all results of the imputations of missingness in a survey. Further principal components analysis of the South-East Ukraine survey (not shown here) indicates that the consistency of answers across a wide range of questions suggests underlying latent beliefs that reveal segmented attitudes. This proves useful in providing a data driven approach to addressing an MNAR case when a particularly sensitive MNAR question has a non-sensitive expression within the same component. We extracted the first two components accounting for with 16.4% and 7.3%, respectively, for components 1 and 2 of the variance in the 83 measures. Based on the principal component loadings, we identify component 1 as a pro-Russia latent attitude and component 2 as general disillusionment with current (geo)political and economic conditions. We do not present the full table of loadings here due to space considerations. Loadings on component 1 included variables that measured trust in President Putin, support for a possible Russian intervention in the Donbas conflict, agreement with Putin's statement about the fascist nature of the Euromaidan protests, support for Russian language policies in Ukraine and approval of the Crimean annexation. High loadings for component 2, questions that measured how the situation was developing in Ukraine – right or wrong direction, actions of the Ukrainian government, expectations about the direction of the conflict, and disinterest in politics. In turn, the component scores can be analyzed using demographic predictors after imputation of missing values in a MAR data set (not shown here).

A common approach to dealing with missing data is the use of a dummy variable for missing observations. A new variable is coded, with the binary valuation of 1 for a missing response and 0 for a properly recorded response. This is a problematic approach if one appreciates the issue of the missing data mechanism. In the first instance of an MCAR variable, adding a dummy variable is simply adding a noise variable at best, which can contribute to overfitting and possibly false significance/inclusion in a resulting mode (Flack et al, 1987), At worst, it can be adding a spurious relationship to the model. The most common type of missing data mechanism, MAR, introduces a separate problem in the context of a dummy variable. When the dummy variable is included with a variable exhibiting CDM, multicollinearity is introduced into the results with the problematic effect of increasing the standard error, and hence unduly decreasing the significance of a coefficient. (Weissfeld and Sereika, 1991). When a dummy variable is used in the case of a non-ignorable missing data mechanism, the result is to give a false positive result of homoscedasticity. The principle of parsimony helps in reminding us to have a theoretical foundation to analytic and modeling decisions, rather than a kitchen sink approach that leads to problematic outcomes and difficult interpretation. As a result, we do not implement or support this approach.

Table 3 about here

The treatment of missingness in the three respective datasets (MCAR, MAR and MNAR) using listwise, pairwise and multiple imputation for the outcome variable (level of positive response, a yes answer or agreement with the statement) are shown in Table 3, a-c. While the percentages are similar, the assumptions underlying the respective approaches are very different. Special attention should be paid to the Abkhazia case (identified as MNAR) where the imputation of the answers of the Georgian respondents (the key group of interest) yields a much higher rate of agreement than seems merited by the precarious situation of this nationality. Georgians in their area of greatest concentration, the *rayon* of Gal(i) in the south of the republic across the Inguri river from Georgia, indicate that most were strongly in favor of the possibility of return for their friends and relations who had moved to or been displaced in the war of the early 1990s. As will be seen below, the MI procedure imputes a higher level of agreement (82%) with the (negative) proposition against refugee return in the survey questionnaire than seems merited by the facts on the ground.

The results of the logistic models are presented in Table 4, a-c. In the Transnistria 2014 dataset (MCAR under Little's tests), the variation of the three methods on the binary outcome across the five sensitive questions is small (see Table 3a); the main effect is that the use of listwise deletion tends to overstate the size of significant coefficients. For Nagorno-Karabakh 2011 (Table 4b), one of the MAR examples, a similar effect can be seen on the overall binary outcomes across the three methods. The regression, however, shows a slight effect that reduces the significant coefficients despite the overall low rate of missingness. The example indicates that one should not be confident with a low level of missingness but that imputation is warranted even with such a small value. While it may not always be necessary, overlooking this approach when it is needed will be problematic, especially if compounded across many sensitive questions, thus doing so serves as a source of added robustness.

The most interesting results appear in the context of the Abkhazia 2010 survey. While the three methods of handling missing data (pairwise, listwise and imputation) produce consistent coefficients, closer examination of the regression results reveals some problematic results. In the particular case of the 'right of return', probably the most sensitive question in this context, each missing data treatment leads to conclusions about Georgians that skew heavily against the reality on the ground, as indicated from the local interviews by the second author and colleagues. The MNAR mechanism in the Abkhazia 2010 dataset is that answers against the proposition (refugees have no right to return) are highly sensitive for Georgians and thus, their high rate of non-response is meaningful and noteworthy given the high sensitivity to this question. The ethnic balance in the republic is critical with the Abkhaz, who control the state apparatus generally against refugee return as it would threaten both their population status (barely a majority, 50.7%, according to the 2014 official census figures) and political control. The earlier CDM analysis showed that this MNAR mechanism is strongly related to ethnicity (the Georgian group is distinguished from the other ethnicities). Since the assumptions of the three missing data treatments explicitly require that data should not be MNAR, it is no surprise that they are deficient in this particular case.

Overall, the coefficients for listwise and pairwise deletion as an adopted strategy for missingness in these data sets are very similar in our examples. While the multiple imputation strategy is somewhat cumbersome, it is the preferred option for MAR data sets, the most common type of missingness in these post-Soviet data sets. We advocate comparison of missing data by key demographic categories as a first step in any survey data analysis and if large differences are seen, imputation will generally yield better estimates of the missing values. Though many imputation methods are available, calculation of multiply imputed values are likely to yield more accurate estimates. The MNAR case is more vexing since imputation will not generally produce coefficients in the models than are more reliable than pairwise or listwise deletion.. In these cases, the information gathered about the local contextual elements that might affect the rate of 'don't knows' is helpful in generating more precise estimates.

Conclusions

The main message of this article is that missing data in surveys in politically sensitive areas should be examined with more care than is the norm in current research. The problem is compounded when the 'don't know' responses are likely generated by the avoidance of sensitive questions due to possible perceived negative ramifications for the respondent.   Asking a filtered question ('do you have an opinion on this question?') is likely to yield a huge number of non-responses and is generally not recommended in the public opinion literature.

A careful examination of missing data in order to identify the underlying mechanisms should be seen as a robustness check, similar in purpose to fitting multiple alternative statistical models. While many conditions can contribute to the proliferation of missing data, alternative explanations behind these mechanisms should be considered and evaluated. Researchers should endeavor to preemptively identify and mitigate potential missing data mechanisms through common survey practices. This should including the use of pretesting to achieve construct validation, performing a pilot survey to identify sensitive questions, and revising the sensitive survey instruments by possibly rewording questions, or substituting interviewers with demographics less likely to cause 'don't know' responses when possible. Information about the political, social and economic context of the survey site may help to guide questions that are prone to 'don't know' responses. When the missingness of data is evident, even if there is a compelling theory to support a mechanism, comparing missing data treatments and the impact of treatments on results should be identified through statistical comparisons. Results that appear significant could be an artifact of the choice of the missing data treatment. Research undertaken in contexts like de facto states often target questions with some element of sensitivity, and the cases where this does make a difference may be overlooked without this more robust approach to missing data, allowing those overlooked cases to echo throughout the resulting analysis and conclusions, undermining the initial intent of scientific, data driven research.

The hoped-for, but mostly incorrect, assumption of MCAR (missing completely at random) in survey data is unlikely in the contexts of political tension and the posing of sensitive questions; with the widespread availability of powerful computing, except in cases with trivial amounts of missingness and no theoretical/contextual basis for an MNAR outcome alongside theoretically informed negative testing for CDM, researchers ought to default to imputation to avoid even small chances of a mistaken approach, as MI will not bias MCAR data, whereas deletion will bias MAR data. Cross-tabulations of zero responses /answered responses across other items in the survey can indicate any trend in missing items and generate insights into the patterns.   MAR (missing at random) mechanisms are more likely and in these cases, multiple imputations of the missing values or maximum likelihood procedures are advised.   The MNAR (missing not at random) circumstance poses special problems for the questions of particular sensitivity that drive this mechanism. In such circumstances, as we have shown here, the generation of imputed values based on overall distributions might be inappropriate and serious consideration should be given to using qualitative information based on detailed contextual information to fill in the missing values. While not explored here, these attitudes can be informed by finding an underlying component which is hidden behind a particularly sensitive question, but which has a non sensitive expression elsewhere in a survey. Follow up cognitive interviews with DK respondents can provide evidence for the response being concealed in MNAR cases as well. Furthermore, a theoretical model grounded in local context can also serve as the basis for calculating MNAR missing responses. Of course, this must be done with maximum transparency, so readers are fully aware that such an approach was used, how it compares to an improper MI or deletion approach, as well as to a multinomial logistic regression with missingness as one of the outcomes, to achieve optimal robustness in such difficult cases. In short, there is no single universally applicable solution. Multiple approaches depending on the particular contexts of the data are encouraged, and multiple sets of results should be presented, for these sensitive environments.

References

Allison, P. (2002) *Missing Data*, (Thousand Oaks, CA, SAGE) (Quantitative Applications in the Social Sciences, 136).

Bakke, K.M., Linke, A.M., O'Loughlin, J. & Toal, G. (2018) 'Dynamics of state-building after war: External-internal relations in Eurasian de facto states', *Political Geography*, 63.

Bakke, K.M., O'Loughlin, J., Toal, G. & Ward, M.D. (2014) 'Convincing State-Builders? Disaggregating Internal Legitimacy in Abkhazia', *International Studies Quarterly*, 58, 3.

Bakke, K. M., Andrew Linke, John O'Loughlin, and Gerard Toal. (2018) "Dynamics of State-Building after War: External-Internal Relations in Eurasian de facto states" *Political Geography,* 63, 2.

Baraldi, A.N. & Enders, C.K. (2010) 'An introduction to modern missing data analyses', *Journal of School Psychology*, 48, 1.

Blair, G., Imai, K. & Lyall, J. (2014) 'Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan', *American Journal of Political Science*, 58, 4.

Bradburn, N.M., Sudman, S., Blair, E. & Stocking, C. (1978) 'Question Threat and Response Bias', *Public Opinion Quarterly*, 42, 2.

DeMaio, T. (1984) 'Social desirability and survey measurement: A review', in *Surveying Subjective Phenomena*, Thousand Oaks, CA: SAGE.

Enders, C.K. (2010) *Applied Missing Data Analysis*, (Thousand Oaks, CA: SAGE)

Fair, C.C., Malhotra, N. & Shapiro, J.N. (2012) 'Faith or Doctrine? Religion and Support for Political Violence in Pakistan', *Public Opinion Quarterly*, 76, 4.

Fair, C.C., Malhotra, N. & Shapiro, J.N. (2014) 'Democratic Values and Support for Militant Politics: Evidence from a National Survey of Pakistan', *Journal of Conflict Resolution*, 58, 5.

Flack, Virginia F., & Potter C. Chang. (1987) 'Frequency of Selecting Noise Variables in Subset Regression Analysis: A Simulation Study.' *The American Statistician* 41, 1.

Frichova-Grono, M. (2011) '*Displacement in Georgia - IDP attitudes to conflict, return and justice - Georgia'* , London: Conciliation Resources. Available at: https://reliefweb.int/report/georgia/displacement-georgia-idp-attitudes-conflict-return-and-justice (accessed May 2019).

Frye, T., Gehlbach, S., Marquardt, K.L. & Reuter, O.J. (2017) 'Is Putin's popularity real?', *Post-Soviet Affairs*, 33, 1.

Gonzalez-Ocantos, E., Jonge, C.K. de, Meléndez, C., Osorio, J. & Nickerson, D.W. (2012) 'Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua', *American Journal of Political Science*, 56, 1.

Grace-Martin, K. (2013) 'How to Diagnose the Missing Data Mechanism', *The Analysis Factor*, 20 May. Available at: https://www.theanalysisfactor.com/missing-data-mechanism/ (accessed May 2019).

Javeline, D. (1999) 'Response Effects in Polite Cultures: A Test of Acquiescence in Kazakhstan', *The Public Opinion Quarterly*, 63, 1.

Jessee, S.A. (2017) ''Don't Know' Responses, Personality, and the Measurement of Political Knowledge*', *Political Science Research and Methods*, 5, 4.

Johnson, T. & Van de Vijver, F. (2003) 'Social desirability bias in cross-cultural research', in Harkness, J., Van de Vijver, F., and Mohler, P. (eds) *Cross-Cultural Survey Methods*, (Hoboken NJ, John Wiley & Sons).

Lall, R. (2016) "How Multiple Imputation Makes a Difference", *Political Analysis*, 24, 4.

Robin L Kaplan & Yu, E. (2015) Measuring Question Sensitivity, (Washington DC, Bureau of Labor Statistics Office of Survey Methods Research).

King, G., Honaker, J., Joseph, A. & Scheve, K. (2001) 'Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation', *American Political Science Review*, 95, 1.

de Leeuw, E.D., Hox, J.J. & Boevé, A. (2016) 'Handling Do-Not-Know Answers: Exploring New Approaches in Online and Mixed-Mode Surveys', *Social Science Computer Review*, 34, 1.

Li, C. (2013) 'Little's Test of Missing Completely at Random', *The Stata Journal*, 13, 4.

Little, R.J.A. (1988) 'A Test of Missing Completely at Random for Multivariate Data with Missing Values', *Journal of the American Statistical Association*, 83, 404.

Little, R.J.A. & Rubin, D.B. (2002) *Statistical analysis with missing data*. 2nd ed, (Hoboken NJ, Wiley).

O'Loughlin, J., Toal, G. & Kolosov, V. (2017) 'The rise and fall of 'Novorossiya': examining support for a separatist geopolitical imaginary in southeast Ukraine', *Post-Soviet Affairs*, 33, 2.

Raghunathan, T.E. (2004) 'What Do We Do with Missing Data? Some Options for Analysis of Incomplete Data', *Annual Review of Public Health*, 25, 1.

Schuman, H. & Presser, S. (1996) *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*, (Thousand Oaks CA, SAGE).

Shapiro, J.N. & Fair, C.C. (2010) 'Understanding Support for Islamist Militancy in Pakistan', *International Security*, 34, 3.

Sturgis, P., Roberts, C. & Smith, P. (2014) 'Middle Alternatives Revisited: How the neither/nor Response Acts as a Way of Saying 'I Don't Know'?', *Sociological Methods & Research*, 43, 1.

Sweet, S.A. & Grace-Martin, K. (2012) *Data analysis with SPSS: a first course in applied statistics*. 4th ed, (Boston MA, Allyn & Bacon).

Toal, G. & Grono, M.F. (2011) 'After Ethnic Violence in the Caucasus: Attitudes of Local Abkhazians and Displaced Georgians in 2010', *Eurasian Geography and Economics*, 52, 5.

Toal, G. & O'Loughlin, J. (2016) 'Frozen fragments, simmering spaces: The post-soviet de facto states', in Holland, E. C. and Derrick, M. (eds) *Questioning Post-Soviet*, (Washington DC, Woodrow Wilson Center International Center for Scholars).

Tourangeau, R. & Yan, T. (2007) 'Sensitive questions in surveys', *Psychological Bulletin*, 133, 5.

Weissfeld, LA & Sereika, SM (1991) 'A multicollinearity diagnostics for generalized linear models' *Commun Stat Theory Methods,* 20, 4.

Young, R. (2012) *Don't know responses in survey research*, Ph.D. The Pennsylvania State University.

Table 1: Missing "don't know" values by sample and by question. The samples that are indicated in ***italics and in bold*** are examined in greater detail.

Location and date of sample	Trust the Government	Trust the President	End of FSU- right or wrong	Getting a job	No right to return for refugees	Sample Size
Transnistria 2010	19%	19%	19%	N/A	N/A	976
Transnistria 2014	27%	4%	18%	20%	4%	***750***
Abkhazia 2010	***15%***	***10%***	***16.5%***	***18.5%***	***20%***	1000
Abkhazia 2014	11%	6%	9%	9%	10%	752
South Ossetia 2010	13%	12%	11%	6%	12%	506
South Ossetia 2014	13.5%	12%	12%	6%	12%	500
North Caucasus 2005	N/A	N/A	N/A	10.5%	12%	2000
Nagorno-Karabakh 2011	***4%***	***1%***	***7%***	***N/A***	***1%***	***800***
South-East Ukraine 2014	***16.5%***	***26%***	***22%***	***29%***	***6%***	***2033***
Crimea 2014	9%	8%	11%	9%	9%	752

Table 2: Significance of Little's Test of MCAR and CDM (rejection of null hypothesis) using the covariates of gender, age and ethnicity.

Survey	MCAR	P value	Gender CDM	Age CDM	Ethnicity CDM
Abkhazia 2010	NO	0.000	0.000	0.000	NS
Abkhazia 2014	NO	0.000	0.004	0.002	NS
Crimea 2014	NO	0.005	NS	NS	NS
N-Karabakh 2011	NO	0.055	NS	NS	Homogenous
N. Caucasus 2005	NO	0.069	N/A	N/A	N/A
S. Ossetia 2011	NO	0.000	0.000	0.000	NS
S. Ossetia 2014	NO	0.001	NS	0.035	NS
S. Ukraine 2014	NO	0.000	0.000	0.000	0.001
Transnistria 2010	NO	0.000	0.000	0.000	0.000
Transnistria 2014	YES	N/A	NS	NS	NS

Table 3: Percentages of the total samples giving a positive answer (yes or agreement) to the individual attitudinal question after pairwise or listwise deletion and Multiple Imputation (MI). Table 3d shows the Georgian subsample in Abkhazia (percentage of Georgians) agreeing with the right of return.

a) Transnistria 2014 -

	Pairwise	Listwise	MI
End of the Soviet Union	5.9%	7.1%	6%
Trust the Government	70.8%	70.8%	70.7%
Trust the President	65.8%	67.4%	65.8%
Ethnic connections for a job	21.8%	23.3%	23.3%
No right of return	87.9%	84.8%	87.9%

b) Nagorno-Karabakh 2011-

	Pairwise	Listwise	MI
End of the former Soviet Union	54.3%	53.7%	54.6%
Trust the Government	31.7%	31.6%	32.1%

c) Abkhazia 2010:

	Pairwise	Listwise	MI
End of the Soviet Union	34.2%	29.9%	34.7%
Trust the Government	50.0%	50.0%	49.7%
Trust the President	93.6%	94.1%	93.1%
Ethnic connections for a job	74.4%	72.7%	74.0%
No right of return	87.7%	90.3%	87%

d) Abkhazia 2010 - Percentage of Georgians giving a positive response (agreeing) to the attitudinal question on the right of return (agreeing means no right of return for refugees)

	% Missing	Pairwise	Listwise	MI
No right of return	37.7%	75.2%	80.6%	82.2%

Table 4: Comparison of Logistic Model Coefficients for 3 Different Options of Missingness

a) MCAR Transnistria 2014: Odds Ratios for 1) No right of return (return) and 2) End of the Soviet Union (FSU)

Variable	Listwise - Return (n=312)	Listwise - FSU (n=305)	Pairwise - Return (n=709)	Pairwise - FSU (n=584)	MI – Return (n=713)	MI – FSU (n=713)
Age	1.0243 (0.047)	0.9921309 (0.615)	1.026558 (0.002)	0.9997507 (0.983)	1.027665 (0.001)	0.9969849 (0.787)
Female	1.490182 (0.234)	0.564818 (0.212)	1.263193 (0.329)	0.643147 (0.213)	1.224495 (0.404)	0.7129804 (0.36)
Education 3	0.7160268 (0.797)	1.062173 (0.963)	0.3031954 (0.3)	0.4545744 (0.363)	0.3810238 (0.407)	0.5044111 (0.423)
Education 4	0.2746025 (0.234)	1.269459 (0.831)	0.1217462 (0.043)	0.7743977 (0.707)	0.1706747 (0.082)	0.7780159 (0.717)
Education 5	0.1599763 (0.12)	1.529242 (0.745)	0.08891 (0.026)	0.3493812 (0.275)	0.1250481 (0.052)	0.4043411 (0.333)
Education 6	0.2953723 (0.273)	1.355476 (0.791)	0.1450473 (0.068)	0.5606456 (0.433)	0.1843714 (0.102)	0.5982321 (0.461)
Education 7	0.3129513 (0.379)	1.433914 (0.808)	0.1237576 (0.067)	0.4696158 (0.528)	0.1874536 (0.142)	0.5193297 (0.569)
Education 8	0.2801745 (0.265)	1.139562 (0.915)	0.1282265 (0.053)	0.742062 (0.689)	0.1628623 (0.08)	0.7353391 (0.675)
Ukrainian	3.651537 (0.006)	0.8777973 (0.824)	2.400337 (0.007)	1.152831 (0.747)	2.3572 (0.01)	1.015236 (0.971)
Moldovan	2.13314 (0.079)	0.6579645 (0.521)	2.177392 (0.019)	0.6874371 (0.473)	2.095067 (0.029)	0.6658186 (0.421)
Other ethnic	1.75663 (0.226)	1.225197 (0.742)	1.362158 (0.348)	1.204637 (0.711)	1.327805 (0.39)	1.201968 (0.686)
Mood 2	0.7629893 (0.519)	1.493098 (0.497)	0.5137944 (0.049)	1.86467 (0.185)	0.5673615 (0.09)	1.712805 (0.278)
Mood 3	0.8045458 (0.751)	2.949154 (0.196)	0.6318751 (0.335)	2.006226 (0.272)	0.6940109 (0.439)	1.721465 (0.394)
Mood 4	1	1	1	1	1	1
Constant	3.777709 (0.283)	0.0899674 (0.084)	15.86707 (0.013)	0.0785232 (0.007)	10.87814 (0.03)	0.0898676 (0.007)

Odds Ratios followed by significance in ( ), .05 highlighted. The larger the education category, the more educated, with 8 meaning University. The larger the mood category, the worse the mood, with 4 being "fear and anguish." Russian is the comparator for the ethnic variable

b) MAR Nagorno-Karabakh 2011: Odds Ratios for 1) Trust in government (Govt) and 2) End of the Soviet Union (FSU)

Variable	Listwise - Govt (n=719)	Listwise - FSU (n=717)	Pairwise - Govt (n=766)	Pairwise - FSU (n=748)	MI - Govt (n=800)	MI - FSU (n=800)
Age	0.9943429 (0.304)	0.9653101 (0)	0.9938223 (0.247)	0.9658332 (0)	0.9929328 (0.187)	0.965721 (0)
Female	0.7611017 (0.1)	1.071588 (0.672)	0.7741378 (0.112)	1.144246 (0.402)	0.7920321 (0.154)	1.164721 (0.34)
Education 3	2.011602 (0.279)	2.072852 (0.269)	2.384165 (0.164)	1.946045 (0.311)	2.29994 (0.183)	1.988715 (0.295)
Education 4	0.990437 (0.978)	1.015902 (0.963)	1.044477 (0.9)	0.9798893 (0.951)	1.018176 (0.958)	1.004389 (0.989)
Education 5	0.9569104 (0.941)	1.181223 (0.768)	0.9903688 (0.987)	1.125307 (0.834)	0.95878 (0.944)	1.149211 (0.805)
Education 6	1.217631 (0.649)	1.221752 (0.632)	1.230136 (0.627)	1.099083 (0.819)	1.165796 (0.721)	1.104252 (0.808)
Education 7	0.7600582 (0.595)	1.924203 (0.212)	0.7644612 (0.584)	2.044338 (0.155)	0.7450702 (0.535)	2.115002 (0.143)
Education 8	1.385513 (0.429)	1.708413 (0.198)	1.339085 (0.472)	1.712751 (0.188)	1.31973 (0.493)	1.707244 (0.194)
Mood 2	0.5252001 (0.002)	0.4498306 (0.001)	0.5352084 (0.002)	0.4516945 (0.001)	0.5393768 (0.003)	0.4431118 (0.001)
Mood 3	0.1793019 (0)	0.214924 (0)	0.1683943 (0)	0.2106715 (0)	0.167833 (0)	0.204021 (0)
Mood 4	0.5559915 (0.236)	0.1241313 (0)	0.5570994 (0.234)	0.1063151 (0)	0.6380226 (0.363)	0.103678 (0)
Constant	1.185826 (0.705)	9.600437 (0)	1.153965 (0.745)	9.572809 (0)	1.220566 (0.65)	9.518097 (0)

Odds Ratios followed by significance in ( ), .05 highlighted. The larger the education category, the more educated, with 8 meaning University- level. The larger the mood category, the worse the mood, with 4 being “fear and anguish.”

c) MNAR Abkhazia 2010 : Odds Ratios for Agreement with the statement refugees have no right to return

	Listwise (n=473)	Pairwise (n=792)	MI (n= 976)
Age	1.003482 (0.72)	1.00931 (0.179)	1.008712 (0.163)
Female	0.8156246 (0.53)	1.009193 (0.968)	1.028121 (0.902)
Education 3	0.4610392 (0.415)	0.5972436 (0.553)	0.6592181 (0.636)
Education 4	1.271914 (0.666)	0.9189506 (0.836)	0.8413381 (0.708)
Education 5	1	2.326839 (0.438)	2.0633 (0.533)
Education 6	0.9224806 (0.883)	0.7337508 (0.438)	0.764515 (0.471)
Education 7	0.2749914 (0.098)	0.5272055 (0.301)	0.5774224 (0.393)
Education 8	0.918309 (0.877)	0.6020172 (0.202)	0.6690038 (0.346)
Armenian	0.3419913 (0.015)	0.4602781 (0.018)	0.5417527 (0.065)
Georgian	0.234201 (0.001)	0.2303192 (0)	0.2509937 (0)
Russian	0.4608689 (0.184)	0.4461544 (0.04)	0.3536827 (0.004)
Mood 2	0.4556438 (0.164)	0.592658 (0.155)	0.7276393 (0.397)
Mood 3	0.2641363 (0.034)	0.3508213 (0.012)	0.5230723 (0.1)
Mood 4	0.7935478 (0.846)	0.5121565 (0.318)	0.6447305 (0.473)
Constant	40.55426 (0)	20.75893 (0)	16.12264 (0)

Odds Ratios followed by significance in ( ), .05 highlighted. The larger the education category, the more educated, with 8 meaning University. The larger the mood category, the worse the mood, with 4 being “fear and anguish.” Abkhaz is the comparator for the ethnic variable.

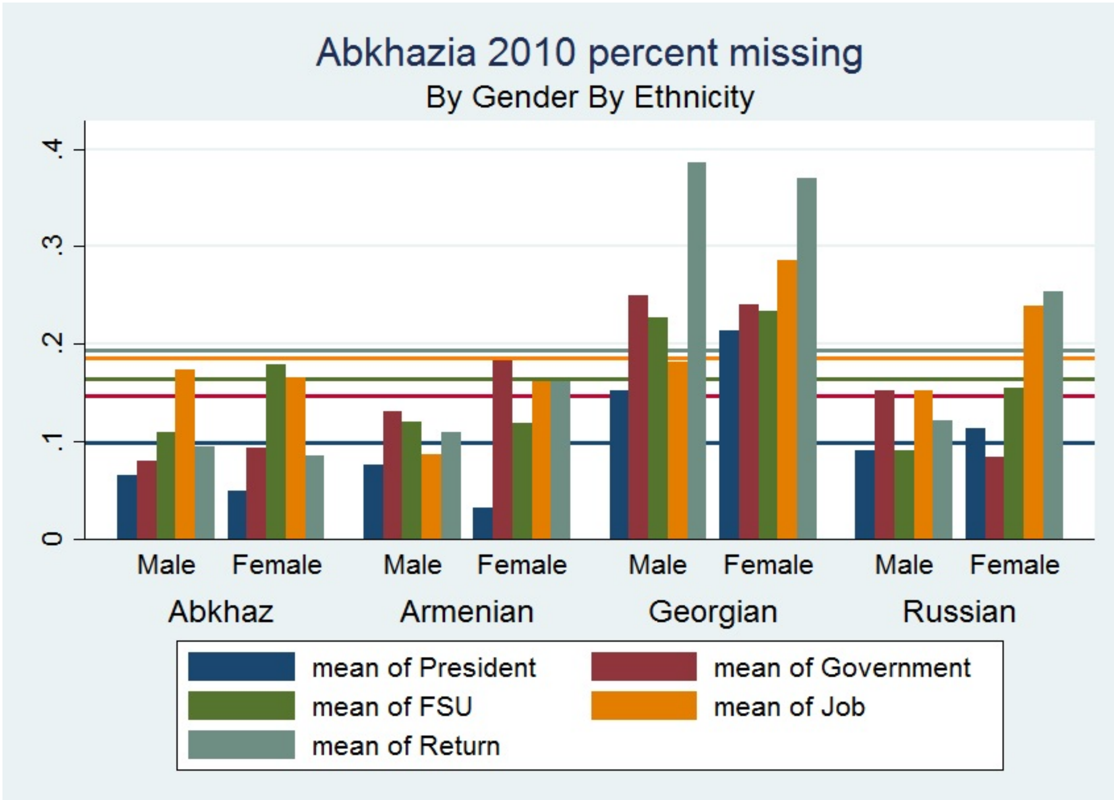

Figure 1: Example of missingness for Abkhazia 2010– MNAR example. The horizontal lines indicate the average level of missingness for each (color-coded) key variable.

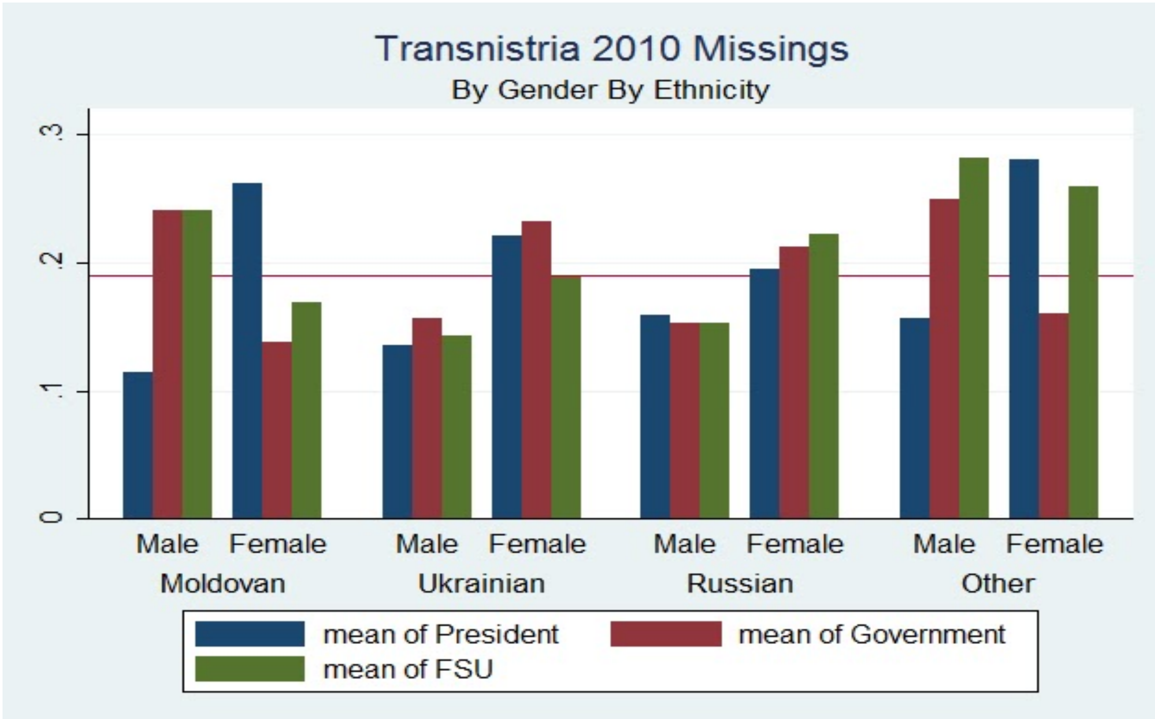

Figure 2: Example of missingness from Transnistria (MAR). The red horizontal line indicates the average rate of missingness for the 3 key variables displayed.

Figure 3 – Example of missingness in South-East Ukraine (MAR)