Online supplementary material for Frye, Margaret, and Lauren Bachan. 2017. The demography of words: the global decline of non-numeric fertility preferences, 1996-2011, *Population Studies.*

In this appendix, we provide additional information about the variables used in the analysis and some additional analyses that we were not able to fit into the main text. First, we provide a more detailed discussion of our outcome—non-numeric ideal family size—including a discussion of the change in probing protocol that occurred prior to the surveys included in our sample and a descriptive overview of the sub-categories of non-numeric response included in a minority of surveys. Next, we discuss two variables for which some of the surveys have samples that are largely homogeneous: marriage and knowledge of contraception. For each of these measures, we conducted additional robustness checks to ensure that these samples with low variance are not affecting the results presented in the paper. We then include a description of the wording of the questions and the categories used for the analysis, for all variables in the sample. Finally, we include ancillary analyses that assessed whether the association between individual- and survey-level factors and non-numeric IFS documented in the article are driven by interviewer-respondent dynamics.

C.1: Additional Information About the Non-Numeric Ideal Family Size Measure

We constructed a binary variable indicating whether a woman provided a nonnumeric response to the following question: "*If you could choose exactly the number of children to have in your whole life, how many would that be?*" For women with children, this question is prefaced by the phrase: "*If you could go back to the time when you did not have any children*…"

Change in Probing Protocol

Starting with Phase III of the DHS surveys, interviewers were instructed to probe for numeric responses in a non-suggestive manner before recording a non-numeric response (the instructions in the questionnaire simply say, "*probe for numeric response*.") If interviewers were unable to solicit a numeric response after probing, they were instructed to record the woman's exact response in the "other" category. For surveys in Phase I and Phase II, no probing instructions were given for this question.

The DHS provides no data on whether individual respondents were probed after giving a non-numeric response, so we cannot precisely estimate the effect of this change in protocol on the prevalence of non-numeric responses. However, we examined aggregate rates of non-numeric response for the 27 countries with a Phase 1 or Phase 2 survey and at least two surveys in Phase 3 or later. The prevalence of non-numeric responses declines at twice the rate per year during the period between Phase 2 (with no probing) and the subsequent survey (the first time when probing was introduced in the country), compared to the interval between the first two surveys with the probing protocol (the average rate of decline in the first period was 0.5 percent per year, compared to 0.27 percent per year for the second period). These results suggest that the change in probing protocol likely led to lower rates of non-numeric response, thus we include only surveys that were administered after this change was made.

Types of Non-Numeric Response

Due to limited data availability, we did not fully investigate the different types of non-numeric responses. Only a small subsample of DHS provide specific sub-categories of non-numeric

responses; we examine these data in Table C.1. The most common sub-category is "*Its up to God/Allah*." This option appears in more than a third of the surveys in Africa and Asia and about a quarter of the surveys in Latin America. Within surveys that include "*up to God/Allah*" as a response type, this category represents the modal type of non-numeric IFS response. A smaller number of surveys included "*I don't know*" as a specific type of non-numeric response; this option was most commonly found in surveys conducted in Latin America (39 percent of surveys) and were less common in surveys conducted in Africa (13 percent of surveys). Nine surveys in Africa also included additional specific sub-categories, including "*can't decide/never thought about it before*," "*depends on husband*," "*any number*," and "*as many as possible*." These additional sub-categories represent less than a quarter of all non-numeric responses in each survey.

While it is difficult to draw conclusions from only a sub-set of surveys, examining the types of non-numeric response suggests that non-numeric IFS is underscored to a certain extent by theology---particularly in Latin America. However, it is impossible to know whether non-numeric IFS is linked to religiosity or simply more a matter of tendencies to invoke the name of God when contemplating uncertain events.

Table C.1: Sub-Categories of Non-Numeric Responses to Ideal Family Size

	Up to God	Don't Know	Other Specific Category[1]	Total
AFRICA				
Number of Surveys with Response Category	24	8	9	59
Proportion of Surveys with Response Category	0.41	0.14	0.15	
Proportion of NNR Responses[2]	0.67	0.13	0.17	
ASIA				
Number of Surveys with Response Category	6	2	0	20
Proportion of Surveys with Response Category	0.30	0.10	--	
Proportion of NNR Responses[2]	0.64	0.18	--	
LATIN AMERICA				
Number of Surveys with Response Category	4	6	0	17
Proportion of Surveys with Response Category	0.24	0.35	--	
Proportion of NNR Responses[2]	0.87	0.39	--	
ALL COUNTRIES				
Number of Surveys with Response Category	34	16	9	96
Proportion of Surveys with Response Category	0.35	0.17	0.09	
Proportion of NNR Responses[2]	0.66	0.16	0.17	

[1]Responses include: "any number," "as many as possible," "depends on husband," and "can't decide/never thought of it before."
[2]Denominator is restricted to surveys in which the response category was included as an option for IFS. Estimates of proportions are weighted.

C.2: Variables with Low Variance in Some Surveys Included in Our Analytic Sample

Married/ In Union

The surveys vary in how many women are included who are coded as not married or in union. Five surveys contain less than 10 percent of the sample that are unmarried; these surveys are

heavily concentrated in Asia and include both of the Vietnam surveys and all surveys in Bangladesh except for 1996, in which 15% of women were unmarried.

As a robustness check, we reran all descriptive analyses with the sample restricted to currently married women for all countries, and the inter-regional comparisons do not differ notably for this limited sample. We also estimated the country-specific models with the sample restricted to married women. The only results that change in Table 2 when unmarried women are excluded are that the regional differences in the magnitude of the discrete change in predicted probability associated with experiencing a child death and knowing a modern method of contraception are no longer significant, although the estimates themselves change very little. For the multilevel models, these differences in sample characteristics may have a small effect on the random intercepts at the country and survey level for these surveys (presented in Figure 5), but should not have an impact on the other model estimates.

Knowledge of Modern Contraception
The proportion that does not know any modern method is less than 1 percent of respondents in 15 surveys (including all surveys in Bangladesh, all surveys in the Dominican Republic, and all but one of the surveys in Nepal) and less than 3 percent in an additional 15 surveys (including all surveys in Zimbabwe, Vietnam, Namibia, and Haiti). In these contexts where information about family planning has reached almost the entire population, women who claim ignorance about any modern method are likely a highly idiosyncratic group of people, and thus results related to this measure should be interpreted with caution.

This issue should not affect the estimate for knowledge of contraception in the multilevel models (Table 3), since a single coefficient for this variable is estimated for all respondents in the sample. It might, however, affect the estimates for the country-specific models presented in Table 2. As a robustness check, we estimated all country-specific models without this variable, and the results for other coefficients did not change substantively. We also estimated these models excluding the surveys for which less than 1% of women report not knowing a modern method, and the summary results for this measure do not change meaningfully from those we present in table 2.

C.3. Description of All Variables in the Analysis
Age
Current age in completed years was calculated from the date of birth of the respondent (respondents were asked, "*In what month and year were you born?")* and the date of interview. For the few respondents who do not know their birthday, current age was recorded based on their answer to the question, *How old were you at your last birthday?*

Number of Living Children
Respondents were first asked, "*Have you ever given birth?"* If they answered yes, they were next asked: "*Do you have any sons or daughters to whom you have given birth who are now living with you?"* If they answered yes to this question, they were then asked to specify how many sons and daughters were living with them. Next, respondents who indicated that they have given birth were asked, "*Do you have any sons or daughters to whom you have given birth who are alive but do not live with you?"* If they answer yes, they are asked to specify how many sons and daughters. This variable represents the sum total of responses to this series of questions.

Currently Pregnant
Respondents were asked, "*Are you pregnant now*?" If they answered yes, they were coded as currently pregnant. If they answered no or indicated that they were unsure, they were coded as not pregnant.

Muslim Religion
Questions about religious affiliation were specific to the countries in which the surveys were conducted. While we would have liked to more deeply explore the effect of religion on non-numeric responses, the DHS does not use consistent categories of Christian denominations across surveys.

Married/In Union
Due to variation in the social and legal definition of marriage across countries, the marriage variable that we used combines the "married" and "in union" status, coding as married all women who report living with their partner. The question reads: "*Are you currently married or living together with a man as if married?"* Women who answer "yes" were coded as married.

Urban Residence
This variable is assigned by the DHS survey team, based on the area where the respondent was interviewed. According to the DHS recode manual for Phase V, "*Urban areas are classified into large cities (capital cities and cities with over 1 million population), small cities (population over 50,000), and towns (other urban areas), and all rural areas are assumed to be countryside*" (DHS 2012, p. 13). We combined all urban areas into one category, with the rural as the reference category .

Experienced Child Death
This variable was created from the child roster. For each child that a woman mentioned having given birth to, the woman was asked, "*Is [name] still alive?"* This variable thus pertains to biological children and not step or foster children.

Educational Attainment
Women who say that they have ever been to school are asked: "*What is the highest level of school you attended?*" and locally specific levels of education are used in the questionnaires.

The DHS provides standardized information about educational attainment using the following categories: No education, Primary, Secondary, and Higher. We collapsed these categories into no schooling, some primary school, and completed primary or above. In some countries the educational system does not correspond to these categories. In these cases, this variable was constructed as accurately as possible from the country's own educational system by the DHS.

Knowledge of a Modern Method of Contraception:
Modern methods include: the pill, IUD, injections, diaphragm, condoms, female sterilization, male sterilization, implants, female condom, foam/jelly and lactational amenorrhea. Methods that are characterized as "traditional" and do not fall into this category include: periodic abstinence (rhythm), withdrawal, and abstinence. If a respondent knows both a traditional method and a

modern method then the modern method takes priority and she was coded as knowing a modern method.

Ever Use of Any Method of Contraception
The DHS does not consistently collect data on whether women have ever used a modern method of contraception, so the variable we used includes traditional and modern methods (see above). For this analysis, we restricted the sample to women who report having had sexual intercourse. This information is missing from all surveys in Bangladesh and Vietnam, thus we excluded these two countries from these models. We also excluded the Cambodia 2010 survey because it does not include the question asking whether respondents have ever used a method of contraception.

Country-Level Independent Variables
We used data from the World Development Indicator Databank (World Bank 2012) for the total fertility rate (TFR), under-five child mortality rate, adult HIV prevalence, percent of the population living in an urban area, and gross domestic product (GDP) per capita (converted into international dollars using purchasing power parity rates). The proportion of women that have ever attended school (any level) was aggregated for each country from the DHS survey data, using sampling weights to create nationally-representative estimates.

Language
In the multilevel regression models, we accounted for differences in language by including country-specific dummy variables, grouping together languages spoken by less than 250 respondents per country. The DHS provides three different types of information on language, and the availability of each type differs across surveys. When possible, we use language of the interview (available for 62 surveys). When this measure was not provided by the DHS, we relied on the respondent's primary language (3 surveys). In the absence of both measures, we used the language of questionnaire (1 survey). In 25 surveys in our sample no language data was collected by the DHS. In all of these cases, the language profiles for the country was homogenous and we thus assumed that interviews in those countries were conducted in a single language. Language homogeneity was determined by looking at other surveys for which language data were available, and defined as a minimum of 85% of people speaking the same language.

Onset of Fertility Transition (Used in Table A1 and Figure 4)
Onset of the fertility transition is defined as the year in which a country experienced a net decline in TFR of more than 10 percent compared to its 1960 value. 1960 was selected because it is the first year for which estimates are available for TFR. While some countries may have begun the fertility transition prior to 1960, all countries in our sample had a TFR above 5.5 in 1960, and the three countries that had a TFR of less than 6 in 1960 experienced a 10 percent net decline from this baseline value before 1980. Thus we are reasonably confident that this choice of a baseline TFR does not bias our assessment of the timing of decline for each country. Information about TFR is sourced from the World Indicators Database.

C.4. Supplementary Multilevel Models Accounting for Interviewer-Respondent Dynamics
Given what we know about survey research, it is conceivable that non-numeric IFS responses are not primarily driven by characteristics of respondents, but rather by dynamics that unfold during the survey interview (Olaleye 1993; Riley et al. 1993). Some interviewers might be more successful in garnering numeric IFS due to a friendly demeanor or more practiced probing (Hox

1994; Mensch et al. 2003; Weinreb 2006). These types of interviewer effects, known by survey methodologists as role-restricted, are related to the interviewers' behavior and conduct during the interview encounter rather than to more enduring characteristics of the interviewer (Bignami-Van Assche, et al. 2003)

Anglewicz and colleagues. (2009) found that bias resulting from role-restricted interviewer effects is more severe for sensitive topics, including gender ideologies and personal AIDS risk. In our case, the question of IFS could be perceived by some respondents as sensitive, particularly for women who desire a non-normative number of children. Additionally, because questionnaires often require interviewers to probe for a numeric answer before recording a non-numeric IFS response, differences in probing skills may have elicited more or fewer non-numeric responses. Indeed, focus group discussions conducted in Nigeria revealed that "with adequate probing, meaningful numeric responses can be elicited from women who give 'up to God' answers" (Olaleye 1993, p. 19). Recent theoretical work suggesting that fertility intentions were formed only when particular situations demanded or motivated them also highlights the importance of examining the influence of interviewers on women's likelihood of providing non-numeric IFS responses (Bachrach and Morgan 2013).

To explore whether respondent-interviewer dynamics explained the association between individual- and survey-level factors and non-numeric IFS, we adopted a slightly different multilevel modeling approach than was used in the full paper. Here we estimated models that included the same independent variables but used a different nesting scheme. In this set of models, women were nested within interviewers, who were nested within surveys. The use of interviewers as a level in hierarchical models has been advocated by statisticians as a method of ascertaining the extent to which heterogeneity in interviewers accounts for variation in outcomes of interest (Hox 1994; O'Muircheartaigh and Campanelli 1998). This method requires that respondent characteristics that influence the assignment of interviewers be included in the models, to avoid confounding of respondent and interviewer characteristics (Hox 1994; Groves et al. 2009). DHS interviewers were matched according to gender and language (Bärnighausen, et al. 2011; ICF International 2012a). Thus, we also included country-specific language fixed effects. In addition to adjusting for non-random interviewer assignment based on the language of the respondent and interviewer, this also controlled for unobserved heterogeneity between countries.

B.2.1 Results

Full tables from these supplementary analyses are available upon request. The only factor for which accounting for the interviewer notably changed the results was ever use of contraception. Accounting for interviewer effects made ever using contraception more of a distinguishing factor among women, making the original estimates presented in the main paper slightly conservative. This may be due to the fact that ever-use measure is derived from more sensitive questions about sex and thus is more likely to be prone to influences from interviewers. In all other respects, the results of the models that accounted for heterogeneity in interviewers were highly consistent with those that focused only on individual and contextual characteristics. Given this, we found no evidence to support the idea that respondent-interviewer dynamics are the primary determinant of non-numeric IFS.

While results from this analysis leave us confident that the effects of knowledge of family planning, education, and mortality-related uncertainty on non-numeric IFS were not artifacts of interviewer dynamics, we were unable to fully reject the possibility that personal characteristics of interviewers influence women's answers to the ideal family size question. This supplementary analysis went further than previous studies in disentangling elements of survey design and implementation, but explorations using smaller-scale studies that collect data on interviewer characteristics are needed in order to further investigate the influence of the interview setting.