

Teaching and Educational Methods

Undergraduate Applied Microeconomic Research with Demographic and Health Surveys

Seth R. Gitter^a, Chris M. Boyd^a, and Raissa Abbasi^a^aTowson University Department of Economics

JEL Codes: O12

Keywords: Demographic health surveys, development, undergraduate research

Abstract

Undergraduate students have demonstrated a growing demand for research opportunities, particularly concerning the world's poorest people. United States Agency for International Development's (USAID) Demographic and Health Surveys (DHSs) Program offers nationally representative data on more than 90 low-income countries that allow for the study of central issues highlighted by the United Nations Sustainable Development Goals (SDGs) such as good health, education, women's empowerment, and access to clean water (SDGs 3–6). In this paper, we provide an extensive overview of the DHS data, previous research with DHS data, and potential research ideas for undergraduate students. A detailed appendix provides instructors with a framework and resources to teach undergraduates to use DHS data as part of course assignments, course-based undergraduate research experiences (CURE), or theses. Using the DHS data and these resources, students can engage in active learning exercises that address some of the key policy issues of their generation.

1 Introduction

There are clear benefits to students in learning the methods and skills associated with applied economics research that translate to graduate school and the workforce. These skills include how to summarize the current evidence based on a research question, learning to devise new questions, gaining experience with data analysis software, and becoming better writers. Besides the benefit of learning these methods and skills, students need to find inspiration in their chosen research topics to fuel their interest in individual research. Many undergraduates ambitiously choose an area of study that would allow them to play a part in addressing the world's biggest problems. Often, this leads to a focus on the poorest countries in the world, where problems are most acute.

The objective of this paper is to show instructors and students how to use data collected as part of the Demographic and Health Surveys (DHS) Program to research some of the world's most pressing problems such as those highlighted by the United Nations Sustainable Development Goals (SDGs). The DHS are nationally representative, so research done with the DHS data can potentially show how a particular country could make progress on an SDG. The DHS have clear links to SDGs 3–6: good health and well-being, quality education, gender equality, and clean water and sanitation. The DHS Program can also be used to explore child nutrition, a dimension of hunger (SDG 2), by using anthropometric data. Finally, although the DHS Program cannot directly address the SDG 1 of no poverty, measured usually by income or consumption, it can provide asset-based measures.

The DHS can be used for both undergraduate research and traditional economics courses. The DHS are an ideal data source for undergraduate students undertaking both the semester-long course-based undergraduate research experiences (CURE) and year-long theses, as well as for certain specific assignments; using these surveys provides the benefits that undergraduate students want while addressing the cost side of the equation for instructors. In this paper, we provide in-depth detailed exercises on using DHS data with undergraduates. These exercises include teaching survey methods,

data analysis, and writing to prepare students for research. These same exercises could be used in standard courses such as international development, health economics, or econometrics to teach students specific aspects of these methods.

The DHS Program has collected data (more than 400 surveys) on more than 90 low-income countries; many countries have multiple years of associated data, since 1984. The program is supported by the United States Agency for International Development (USAID), which has been providing technical assistance to countries to administer the survey for more than thirty years.¹ The role of the DHS Program is to provide countries with reliable information and analyses to inform policy. The survey provides information on a wide range of possible research topics, including anthropometrics, child labor, contraception, fertility, HIV/AIDS, intimate partner violence, malaria, schooling, women's decision-making power, and women's employment. More importantly, the data are freely available and easy to access. To date, DHS data have been analyzed in more than 5,900 articles published in more than 1,100 journals, including leading journals in economics, health, and demography (DHS Program 2024a).

For instructors, there can be substantial time costs to overseeing undergraduate research (Fenn et al. 2010; Gitter 2021).² Helping students find suitable data that can be obtained quickly is important for a semester-long course or, at most, a year-long thesis project. Instructors need to balance the benefits of students pursuing individualized research projects while trying to address the costs of not using the economies of scale typically associated with common assignments and grading in standard courses.

DHS data are particularly well-suited to the potentially large time costs for instructors associated with undergraduate research projects. The DHS consist of mostly the same questions and data structure across countries and time. Selecting an interesting and answerable research question can be time-consuming, especially for undergraduates undertaking their first project. Instructors can guide students to build on the vast amount of existing research by replicating research questions in another country. Instructors can also benefit from economies of scale and scope. For example, to scale the process, a group of students could all examine the same research question with DHS data in different countries, allowing them to share literature and research design; given the 90 countries for which the DHS data are available, this could scale to most class sizes, even with each student choosing a unique country. Instructors can also take advantage of economies of scope and have students answer different questions for the same country, sharing literature and data cleaning work.

This paper provides a framework for instructors and students to use DHS data in the context of undergraduate research in applied economics and provides examples of specific assignments where the DHS can be used in non-research-oriented classes. We divide this paper into five more sections. In Section 2, we provide an overview of the DHS data, detail research from these surveys done in both top academic journals and by students to motivate how to help students select questions. Section 3 discusses data analysis skills, including obtaining and cleaning the data, making descriptive statistics and graphs, and econometric work from basic ordinary least squares (OLS) to more advanced techniques. At the end of this section, we also discuss the potential for linking DHS data to other datasets available, specifically satellite data, food prices, and terrorist attacks. In Section 4, we demonstrate how students can put together the work from qualitative and quantitative parts and create a paper. This section also provides information on where and how to present and publish the work. In Section 5, we offer a few concluding thoughts.

We also propose a framework for instructors using the DHS data for undergraduate research coursework in a series of appendices. The framework is summarized in Appendix A, which provides a nine-unit structure that is linked to learning goals for teaching the DHS along with accompanying readings and suggestions for assignments. Note, however, that some of these units can be used

¹ USAID also uses the surveys to monitor trends in development and inform policy.

² In this paper, we use the term instructor to be inclusive of non-tenure track, tenure track, and tenured faculty.

independently or as stand-alone in regular courses. Example assignments are shown in Appendix B. In Appendix C, we provide an example paper and in Appendix D a first-person student experience.

2 The DHS Dataset and Choosing a Research Question

2.1 Overview of the DHS

As mentioned in the introduction, the DHS mostly contain the same basic information across the world on household demographics, health, fertility, and household wealth. There are four main types of questionnaires focusing on the household, women, men, and HIV and/or other biomarkers. The DHS also collects data on GPS location and covariates of the locations such as weather, access to health care, and schools, for the most recent surveys. Some surveys have additional information on specific health outcomes such as chronic disease, disability, and mental health. The first of the DHS was conducted in 1984, and the survey has been changed roughly every five years (DHS Program 2024a), though most of the core questions remain the same.

To start understanding the survey, we recommend a simple exercise where students role-play being the interviewee or the enumerator for the DHS. This exercise could be used for students doing research or learning about survey methods in a standard course. Students can work in pairs or groups with one of them in the role of enumerator and the other (or others) in the role of the interviewee, for different modules. Before beginning this exercise, we recommend a content notice.³ This exercise allows students not only to read the questionnaire in depth and get familiar with all the questions but also to understand the codes used in the DHS and the overall structure of survey questionnaires.⁴ Moreover, it can be used to motivate related topics such as survey design, ethical concerns of data collection in the field, and varying the use of units of analysis (household, household member, etc.).

To provide an overview of a typical DHS, we discuss below the main parts of each questionnaire module, which can be shared with students who are new to survey-based research or instructors who are new to the DHS. We use Tanzania's 2022 DHS as an example because it has one of the most recent reports on a DHS with data available (MoH 2022). Reports are typically hundreds of pages long with descriptive analysis of the data. The full questionnaire is available in Appendix E of the 2022 Tanzania DHS report, and similar reports exist for all DHS datasets (MoH 2022).

The first page of the questionnaire (MoH 2022, p. 689) includes information on the location of the households including subnational units such as region, district, and ward. When showing students the information on the regions, it may be worth previewing that the DHS survey is often representative at the subnational (regional) level and that stratified sampling is used to construct the data (for Tanzania 2022 see MoH 2022, p. xxxiv), which implies the need to use survey weights when analyzing the DHS.⁵ The first page of the questionnaire also has information on whether the survey team was able to get a response. As the introduction to the report notes, in the 2022 Tanzanian DHS, "the response rate was very high (99 percent household, 97 percent women, and 91 percent men)" (MoH 2022, p. XX). This rate is consistent or better than historical response rates, which typically exceed 90 percent (Corsi et al. 2012).

³ The DHS ask women-sensitive questions, and we recommend providing a content notice before beginning work with students and again before the mock survey exercise. In particular, the DHS asks questions about intimate partner violence, including sexual violence between partners. Providing a content notice that the survey discusses sexual violence, miscarriages, and infant mortality, may help students who need it by giving them time to implement strategies to cope with triggering information (see Stanford Graduate School of Business n.d.; University of Michigan n.d.). This content notice is particularly important to reemphasize before having students go through the survey.

⁴ Furthermore, this exercise can allow a discussion about survey measurement errors. For example, variables in the DHS such as birth date have low levels of reporting or large levels of error, since they rely on mothers' recall; meanwhile, anthropometric measures have small measurement error because they are methodically collected by the enumerator.

⁵ We discuss weights in more detail in Section 3.4.

The questionnaire can also be used to discuss ethical concerns about gathering data in developing countries with the students. First, it is worth showing students the informed consent as part of the survey (MoH 2022, p. 692), and this may be an opportunity to discuss the role of institutional review boards (IRBs) in doing ethical research. Second, the DHS protect anonymity by randomly displacing the actual GPS longitude/latitude information of the cluster (i.e., a grouping of households; DHS Program 2024b) by up to 2 kilometers in urban areas and up to 10 kilometers in rural areas. Data available to the public are anonymized and do not contain names, though names are collected as part of the survey. Third, as mentioned before, the DHS asks detailed questions about sensitive information; this can be used to discuss with students how the DHS approaches these questions, and why not everyone is asked to answer these questions.

The DHS household survey questionnaire gathers comprehensive data on both usual members of the household and visitors (MoH 2022, p. 693). The data cover key demographic details of all household members such as age, sex, relationship to the household head, marital status, marriage age, education, parental survivorship, and if the person resides in the household. The questionnaire also includes household characteristics such as source of drinking water, toilet facilities, cooking fuel, assets owned by the household, and exposure to second-hand smoke. Assets such as livestock animals and durable goods are also detailed, which are used to construct a wealth index (see discussion in Section 2.2).

The household questionnaire contains two main potential units of analysis: the individuals within the household and the household itself. A short YouTube video made by the DHS Program describes that data for these two units are covered by two different data files, which we will discuss below along with other DHS data files.⁶ Each survey has a household recode (a unique identifier), which is best for looking at household-level data about household characteristics. It does include a roster with limited information such as relationship to head, sex, age, and education, but it is difficult for students to work with because these variables are contained in a single row with variables for age of person 1, person 2, etc.⁷ For individual-level information, the Household Member Recode (PR) uses the individual as the unit of analysis and provides full information on each household member's age, sex, education, and numerous other variables.

The women's questionnaire is given to a woman in the household, age 15 to 49, based on reproductive years, and it is the most extensive part of the survey (MoH 2022, p. 709). There are more detailed questions on standard demographics and employment, a particular focus on women's fertility with detailed information about each of her births (including children who died), reproductive intentions, and contraception. The survey also collects information on antenatal, delivery, and postnatal care along with breastfeeding practices, as well as child health status, including recent experience of diarrhea and fever along with immunization records. Women's empowerment is also measured through questions on decision-making and attitudes toward intimate partner violence. Sensitive information (such as the experience of intimate partner violence, HIV health-related questions, and tobacco use) is also collected, in some countries, but only for a subset of women. All this data has the female respondent as the unit of observation and is contained in the individual women's respondent recode (IR).

A brief set of questions is sometimes also given to the respondent's husband. Samples of men ages 15 to 59 are collected in some countries, and in many cases, this is a subsample of the spouses of women in the sample described above. The questions cover many of the same topics as the women's questionnaire (Vaessen et al. 2005), and data are contained in the men's recode file (MR).

The DHS in some cases collects biomarker data on a subsample of the population; in the case of Tanzania, the subsample was half of the households (see MoH 2022, p. 845). Biomarker data includes

⁶ YouTube video link: <https://youtu.be/JGRIZCGiCJw>.

⁷ The advantage of the household module containing a household roster is that using only this household-level data set it is easy to measure household-level variables such as highest education level in the household or the dependency ratio.

tests for malaria and anemia. Anthropometric data, such as weight and height (or length), are also gathered for children under five years old.

2.2 Advantages and Disadvantages of Using the DHS

Besides all the characteristics mentioned in the previous sub-section, we want to emphasize that the main reason to choose the DHS over other surveys is the large comparability across numerous countries and over time, and their easy availability and accessibility. Other surveys conducted in developing countries that can serve the purpose of teaching undergraduate research are the Living Standard Measurement Surveys (LSMS) and the Young Lives surveys. Regarding comparability, both the LSMS and the Young Lives surveys are available for fewer countries and fewer years than the DHS, and the LSMS are less comparable across countries. The LSMS have been conducted for more than fifteen years, mostly in Sub-Saharan Africa, by individual countries' statistical offices with the support of the World Bank, which makes the questionnaires not standardized. The Young Lives surveys are conducted in only four countries, and their questionnaires are standardized across countries. Regarding availability and accessibility, the DHS can be downloaded from the DHS Program website or through the Integrated Public Use Microdata Series (IPUMS), which also ensures the quality of the data. Moreover, for countries in which DHS data are available via IPUMS, other variables (e.g., rainfall) are linked. The LSMS are easy to access from the LSMS website and provide all the data available in the questionnaires, although anecdotal experience of the authors suggests they may have some data quality issues. The Young Lives surveys are also easy to access, but they do not provide important information such as household GPS location, which limits their use (see a discussion on combining DHS data with other datasets in Section 2.3).⁸

Using the DHS data implies restricting undergraduate research to their available information and data structure. Regarding data content, the DHS's largest omissions that are standard in some household development surveys are income, consumption, and time use modules, present in other surveys such as some of the LSMS. Nonetheless, wealth can be estimated using the information on assets in the DHS, and an asset-based measure of wealth is already calculated for some countries (which is not comparable across countries). In most surveys, households are grouped in wealth quantiles, which makes for a good control variable for socioeconomic status in undergraduate projects.

2.3 Examples of Research Papers with DHS Data

In this section, we highlight examples of applied microeconomics research with DHS data. The DHS maintains a database of more than 5,900 articles across more than 1,100 academic journals that have used DHS data across fields (DHS Program 2024a). As examples of papers by economists to share with students, we suggest Jayachandran and Pande (2017) and Brown, Ravallion, and Van De Walle (2019), which use mainly descriptive statistics to better understand malnutrition. To teach students about using difference-in-differences, we suggest assigning McGavock (2021), who examines child marriage bans in Ethiopia, and Gehrke and Kubitzka (2024) on the effects of palm oil expansion on fertility in Indonesia. Agüero and Marks (2008) and Sunder (2019) provide good analyses on women's employment and child marriage, while also being useful to teach about endogeneity and instrumental variables.

IPUMS gives out annual awards to researchers and students (though mostly graduate students) who use their data, including the DHS, and publishes the winning papers on their website.⁹ In the last two years, researcher award winners include Mruts et al. (2022), who look at birth spacing and anemia, and Stoebenau (2021), who examines socioeconomic status and marriage age both using DHS data from multiple African countries. Student award papers with DHS data include Heng et al. (2020) who test the

⁸ Note for the three surveys, a registration is required to prove the status of academic use of these data.

⁹ Awards and applications for submissions are available at IPUMS website, <https://www.ipums.org/impact/ipums-research-award>.

relationship between malaria and low birth weight; Chang (2019) who estimates the influence of abortion laws in African countries on marriage age, births, and schooling rates; and Källmark (2018) who measures the relationship between drought and health outcomes for children in Zimbabwe. Two of the authors have both worked with undergraduate students to find research questions using DHS data. In some cases, students did a simple project such as looking at health care access and infant mortality in Nigeria, or mother's education and infant mortality in Peru as part of course projects. These projects were expanded as part of a longer summer research project and are turning into theses examining food prices during pregnancy and infant mortality (Nigeria), and air pollution and education (Peru) by using DHS data merged with other data sources (geocoded food prices and air pollution). One of the authors was a student who merged DHS data with terrorist attack data in Afghanistan to measure the effect on children's health. This paper is included as Appendix C, and a discussion of their experience is provided in Appendix D. Other students have studied child marriage's effects on education and labor outcomes for women using instrumental variables. Specifically, to use puberty age to predict marriage age (Gitter et al. 2023), which parallels Sunder (2019) mentioned above. Another paper by one of the authors and an undergraduate, under review, combines DHS data from Nepal on women's employment with foreign direct investment at the district level collected by the Nepali government to test the link between the two (Ziard et al., Forthcoming).

2.4 Choosing Research Topics

Applied microeconomic research typically uses an outcome variable and a main variable of interest that causally affects that outcome as seen in the research discussed above. The first requirement for any project is for the data to be available, either in the DHS or in another dataset that could be linked to the DHS (see Section 3 for further discussion on other data sources). Students can build on the first two chapters from *The Effect* (Huntington-Klein 2022) to learn about creating research questions. To help students generate research ideas, first have them identify outcome variables of interest from the DHS dataset, and then ask them to make a list of DHS variables that might influence that outcome (see unit 1 in Appendix A). We also recommend having them do a researcher bio to summarize the work of one researcher to get a sense of potential topics (see Appendix B.1)

The DHS-run website Statcompiler (Statcompiler.com) is an excellent place to start to look at available outcomes and get a sense of descriptive statistics. Statcompiler contains DHS summary data on hundreds of variables in a dashboard that can be used to make within- or between-country comparisons over time. A good classroom exercise is to have students look up the same variable, such as the fertility rate, for different countries. The DHS Program also runs a YouTube channel (youtube.com/@DHSprogram) with brief introductory videos for the survey to teach students about how the survey is conducted.

The structure of semester-long, course-based projects and individualized theses requires different approaches to research question selection. In semester-long courses, choosing a research topic early is important: students will need time for data work and writing. In the sample schedule, making a proposal (see Appendix B.2) or an aspirational introduction (see Shapiro 2022), described below, is done in the second of nine units (see Appendix A), or roughly the second week of a class on the quarter system, or the third week of a semester-long class. As mentioned above, students can form groups based on research questions and share information from their literature search, use similar econometric designs, and help each other with coding and writing.

This last part is particularly important for DHS data, which uses similar variable- and value-naming conventions (e.g., v133 is the mother's years of education, and its missing values are coded as 99). Given the limited nature of time in a semester-long class, instructors may even want to create a set of research questions for students or replicate the same analysis in multiple different countries, assigning each student to a different country's DHS. For example, one of the authors worked with two undergraduates to replicate Sunder's (2019) work on the relationship between child marriage and

education in Uganda by extending the analysis to 11 additional countries (Gitter et al. 2023).

For a course-based research project, the emphasis should be on learning the skills associated with the research process rather than on generating novel ideas. In this case, instructors could assign simple, well-established research topics such as the influence of maternal education on infant mortality. An instructor could also choose a subset of DHS countries for students to work on in groups, which allows students to share resources and does not require the instructor to have as much specialized knowledge of DHS datasets.

For a year-long thesis project, students often need self-motivation, so their interest in the topic becomes more important to their success (see Gitter 2021 for further discussion). Additionally, students with a whole academic year to do research can spend up to a month selecting a research question and still have plenty of time to complete the project. Former CURE students may use their course-based projects as a springboard for a year-long project by getting additional data or extending the analysis. A final step in solidifying a research question can be having students write a proposal. Appendix B.2 provides a sample assignment for a 1-page proposal that asks students to identify the outcome variable, the main variable that influences that outcome and other potential factors, and to begin summarizing the relevant literature. Shapiro (2022) suggests writing an aspirational abstract that can be thought of as a proposal with hypothetical results, which represent the best-case scenario of the empirical analysis.

2.4.1 Pitfalls When Choosing a Research Topic

Students typically make two main errors when choosing a research project with DHS data. First, they sometimes try to choose topics that DHS does not collect data for. We do discuss in Section 3.6 how to link the DHS with other datasets, but this is better for students with more than a semester of time to work on the project. In some cases, the DHS does not always collect the same data for a given country, so certain datasets may be missing information such as anthropometric measures or ethnic categories. Using Statcompiler is a quick way to check if the main variables are available in the country and survey of interest. The second main error is trying to measure the effect with no variation, for example at the regional level, if a whole country experiences an earthquake or a national ban on child marriage. One way to get around that is to find regional variation such as in Khanal (2022), that compares regions within Nepal that experienced different magnitudes of a major earthquake, or in McGavock (2021), that tests regional variation in child marriage bans in Ethiopia.

Students will also often pick research topics with clear omitted variable bias. For example, many students will want to measure the effects of access to resources (e.g., water, electricity, or roads) for health or education outcomes. Access to these resources is likely also related to socioeconomic status, which likely is an omitted variable influencing these health or education outcomes. Students can control for household wealth or include regional fixed effects to lessen these issues somewhat. It is worth remembering that these are undergraduate research papers, and they need not pass the causal identification bar of top economics journals. Discussing omitted variable bias does serve as a learning opportunity and can be written into their papers as a limitation of the work that it is potentially noncausal.

2.5 Conceptual Framework and Existing Literature

Literature reviews in applied economics papers serve a different purpose than students may be accustomed to from other disciplines. Dudenhefer (2006) suggests four purposes for literature reviews in economics: (1) critically analyzing a body of research, (2) putting your own study in the context of others, (3) highlighting your study's contribution, and (4) establishing your scholarly credentials. Typically, for undergraduates doing research, the first purpose can be to establish a theory of change. That is what others have found in terms of the relationship between the main outcome variable and the influencing variable of interest. This previous literature can help support a hypothesis about why a change might occur. Students can then use the literature review for the second and third purposes: to

contextualize their work in light of a contribution. Having students answer how their paper will differ from previous studies is useful in learning about novel research ideas. Students will be most accustomed to the last purpose, establishing that they are familiar with the literature.

To guide students through the process of creating a literature review, it is helpful to have them first find and briefly summarize in a sentence or two a list of ten to fifteen studies they believe are relevant to their work. (The schedule in Appendix A has students summarize articles in the first two units.) Instructors can check this list of studies and their corresponding summaries before students write up their literature review. We suggest guiding students toward the DHS database of 5,900 articles first, and the IPUMS research winners.¹⁰ A good second step after finding a few initial articles is to use Google Scholar to run backward and forward searches of articles (i.e., articles that cite or are cited by the article of interest; see Romem 2012). Starting with Google Scholar is likely to overwhelm students with too many articles, many of which are less relevant; another possibility is searching for papers within the EconLit database.

Helping students understand and identify economics journals or quality journals is difficult, but students can be directed to the Research Papers in Economics (RePec) portal,¹¹ which ranks more than 3,000 economic or related journals, and the Australian's Business Deans Council, which also provides a letter grade to journals in economics and business fields.¹² This may also be a good opportunity to teach students about journal metrics such as the impact factor, although with the caveat that the measures can be manipulated. Teaching students skills to manage, document, and sort articles is also important to their long-term success as researchers and can help build a database of resources for future research students. Zotero is one of many free citation managers (another is Mendeley) that can be used to teach students how to organize their literature review and how to cite correctly.¹³

After collecting articles, writing up a literature review section that motivates a theory of change is a good exercise to help students synthesize their knowledge; an example is shown in Appendix B.3. Instructors in development or health economics courses could potentially ask students to do the same exercise without the intention of doing a full research project. This will teach students to relate what they have learned in the course to the latest research.

3 Quantitative Skills

Students can learn a broad range of quantitative skills necessary to do applied economics research through working with DHS data. The example of these quantitative skills may be included in economics, statistics, or econometric courses as stand-alone exercises. These skills include obtaining data, cleaning data, doing basic descriptive analysis, econometric analysis, and creating replicable code. Basic quizzes with sample data can be used to introduce these concepts before students begin their own analyses. This section provides an overview of these skills as part of a DHS-based research project.

3.1 Downloading the Data

There are two sources of DHS data: the DHS Program website¹⁴ and IPUMS.¹⁵ The DHS Program website provides access to all DHS data, solely for research purposes, after being granted approval by the DHS Program. Before students or the instructor requests data, they will need to provide DHS with a brief summary of the research question for approval; the approval process typically only takes 24 hours, or a

¹⁰ ipums.org/research-award/previous-award-winners

¹¹ ideas.repec.org/top/top.journals.all.html

¹² abdc.edu.au/abdc-journal-quality-list/

¹³ Note that teaching how to use a citation manager has very little cost. Teaching the most important features of Zotero to students can be done in a one-hour session, which can also be run by a librarian.

¹⁴ <https://dhsprogram.com/data/available-datasets.cfm>

¹⁵ <https://www.idhsdata.org/idhs>

few days if GPS information is also requested. This requirement is also necessary to access the IPUMS DHS data. Once access is granted in the DHS Program website, original datasets can be downloaded in several formats, including Stata, SAS, SPSS, and ASCII.

IPUMS only provides data for 45 countries (in Asia and Africa only), but the main advantage of using its data is the harmonization of DHS data between these countries. Another advantage of using IPUMS is the possibility of selecting only the relevant variables for download. Thus, we recommend using the IPUMS DHS website to download data if the objective is doing cross-country comparisons in Asian and African countries or to have all the students work with a harmonized set of variables.

While downloading data either from the DHS Program or IPUMS, students might run into some issues. First, some of the files can be quite large, with file sizes in the hundreds of megabytes, because some files have thousands of variables for tens of thousands of survey respondents. This may be beyond the computing power of some student versions of software or for students with low-end computers such as Chromebooks. A second issue is that students may be overwhelmed with the number of files (see more discussion below) and the number of variables. Third, students may require time and other skills to merge the downloaded datasets at different unit-of-analysis levels. The first two issues can be mitigated by downloading only certain variables from the IPUMS DHS website, but the third issue requires conducting lessons on merging datasets.

To limit these issues, in classes where student experience with data and course time may be limited, the instructor can create small subsets of the original DHS datasets that include only the necessary variables. When teaching a CURE using DHS data, one of the authors creates datasets by country groups, asking students which variables they need. This typically only takes a couple of hours or less for a whole class to give each student a smaller dataset with a few dozen variables. The time saved by helping students navigate the vast amount of data is likely more time than it takes to prepare data for a class. Giving students a simple quiz with a sample dataset can test students' ability to load data and find the number of observations and variable names; these topics are tested in quiz 1 in unit 2 of Appendix A.¹⁶ Downloading data and giving students experience with troubleshooting and data cleaning to keep only relevant variables, helps students when doing research in the future without instructor support.

When downloading the data from the DHS website, it is worth referring back to the previous section because datasets are created around the unit of observation. There is a household recode (HR) data file; an individual recode (IR), which is the female respondent in the household aged 15–49; a household member recode (HR); and a birth recode (BR) that has data on the children of those women. The files use a standardized naming convention that includes the country's name, the data type, the survey version, and the file format. For example, Kenya's (KE) household recode (HR) 2003 DHS phase four (42 for phase 4, version 2) Stata file (DT) is called KEHR42DT.dta.

3.2 Basic Data Cleaning

Students can also be introduced to the process of cleaning data when using the DHS either for their own research or in econometrics and statistics courses. DHS data does not require deep cleaning, but it can be used to show students what a good-quality dataset should look like. The most common data cleaning needed is dealing with missing values, which are often coded as 99. Moreover, the DHS raw data can be used to show students how to transform raw data into variables ready to use in descriptive statistics tables and econometric regressions. Although researchers are accustomed to cleaning the data, it is hard for undergraduate students to do this for the first time. Appendix B.4 provides a detailed set of steps to clean and transform raw data into variables for undergraduate students. In CURE, to ensure students learn these concepts, they can be assessed in a second quiz in unit 3 (see Appendix A). In this quiz,

¹⁶ Quizzes are not provided with this paper so that we may continue to reuse them in our courses; they are available upon request from the authors. We recommend that instructors write their own to link to datasets and topics that will be used.

students can practice concepts such as how to rename, recode, and transform variables and can be asked to make a descriptive statistics table.

3.3 Descriptive Statistics Table for Teaching

Applied economics papers often provide descriptive statistics in the first table of the paper; this table can also be a teaching tool for undergraduate researchers. Table 1 in the sample paper provided in Appendix C gives an example of a typical descriptive table, which can be used to teach students about types of variables (e.g., continuous, categorical, and binary) along with an accompanying statistics textbook such as *The Effect* (Huntington-Klein 2022). Including the maximum and minimum for each variable in the descriptive statistics can also help students with data cleaning and serve as a cross-check for the instructor, particularly to identify where the DHS uses 99 for missing values, such as the number of years of schooling. Including the mean, median, and standard deviation and having students interpret these differences can help reinforce statistics concepts as well. Finally, having students order the variables such that the outcome variable goes first, and the effect variable second will teach them how to hone the argument for the paper. Appendix B.5 provides a sample assignment for students to create a descriptive statistics table.

3.4 Survey Weights

The DHS uses sampling weights to adjust for the differences in sample size needed to be representative at regional levels. For example, in the Tanzania survey, roughly 3 percent of the population lives on the island of Zanzibar, but 15 percent of the sample of the 2022 DHS is taken from the island (MoH 2022, p. xxxiv). The DHS provides YouTube videos to show students how sampling weights work and how to incorporate them in descriptive statistics.¹⁷ The DHS constructs separate weights for each household and female respondent, and in some cases, additional weights are added to correct for sampling differences based on other criteria such as rural/urban residence. In most DHS data files, the relevant survey weights are labeled as the variable v005 or hv005 to designate the women's or household's sampling weight. Encouraging students to use sampling weights in both the descriptive statistics and the econometric work can be useful to teach proper analysis with weighted samples.

3.5 Graphs for Applied Econometrics

Graphs are an excellent way to demonstrate the relationship between the outcome of interest and the main effect variable. The sample descriptive statistics assignment in Appendix B.5 asks students to make a graph with these two variables. First, students need to learn the type of graph to make based on the variables of interest. For example, categorical variables typically work best with bar graphs. Jayachandran and Pande (2017) use a graph showing birth order (a categorical variable) and average height to support their hypothesis that Indian children are shorter, mainly because of limited resources provided to the youngest children in the household. When variables are continuous, a line graph or scatterplot can better show a trend. McGavock (2021) includes an excellent example showing the relationship between child marriage laws and age in Ethiopia using a line graph.

3.6 Econometric Work

To do econometric work, students will first need to create an econometric model, then run the regressions and interpret the results. Students can start with the most basic econometric model, which includes their outcome and their effect variable of interest. Chapter 13 of Huntington-Klein (2022) is an excellent resource to help students design and run their first econometric regressions. Instructors can start by teaching students how to write a basic econometric equation; particularly, if subscripts are used,

¹⁷ <https://www.youtube.com/watch?v=SJRVxvdIc8s>

they can help emphasize the unit of analysis (e.g., individual, household, or country). Next, students can add the three levels of complexity to the basic equation: control variables, fixed effects, and interaction terms. Section 5 of the sample paper in Appendix C provides an example of a fixed effects model using DHS data. These equations can also help the instructor explain control variables; fixed effects and why many DHS papers use them, as they include multiple surveys or use subnational units or time/cohort controls (see Chapter 16 of Huntington-Klein 2022); and interaction terms and why they are often used (e.g., when specific subpopulations may be of interest). To help students understand the last concept, students can be asked to rerun regressions for rural and urban subpopulations. Jayachandran and Pande (2017) and McGavock (2021) provide helpful examples of interaction terms.

Applied economics and particularly development economics research stresses the importance of causal identification; teaching students about endogeneity is an important part of this process. Chapters 5 through 10 in Huntington-Klein (2022) provide excellent details on these concepts, including a discussion of identification, causal diagrams, and finding front doors. A more condensed discussion of the three types of endogeneity (omitted variable bias, measurement error, and reverse causality) can be found in Bellemare (2016). It is helpful for students as they write their econometric framework to think through each of the types of endogeneity.

The next task for students is to run their regressions and write up their results. The assignment shown in Appendix B.6 has students create a results table with five regressions and interpret the effect sizes. This assignment can also be used to reinforce concepts such as the difference between statistically significant (e.g., p -value) and economically meaningful results. It is also helpful to have students review their descriptive statistics to put effect sizes in terms of a standard deviation of the variable of interest and compare the effect size with the mean. In a past CURE, one author typically assigned two quizzes related to econometrics, the first to test both the students' ability to use RStudio or Stata to run a regression and interpret it, and the second to test students' ability to run and interpret a regression with interaction terms and fixed effects.

3.7 Linking the DHS to Other Sources of Data

Linking data from other sources to the DHS can create excellent potential research questions that can be assessed with a causal inference approach. This approach is best for students doing year-long theses who have time to find, clean, and merge a second dataset, while instructors may need to provide the additional data or help with the linking if done in semester-long CUREs. The DHS data typically includes GPS coordinates and other geographic information that can be used to link other data sources. For countries included in the IPUMS-DHS, accompanying geographic data (including shapefiles at harmonized subnational levels) is also available. The DHS program maintains a Spatial Data Repository that collects freely available data from other sources to match with DHS data, including population density, weather, malaria, and livestock ownership (Mayala and Donohue 2022). Weather data can be a great complement to DHS data and allows students to generate unique research topics. Another example of data linkable to the DHS is food prices. Recently, an undergraduate student doing summer research with one of the authors linked data on local food market prices in Nigeria, collected by the World Bank (see Andree 2021), to DHS data to understand the impact of food prices on infant mortality and children's height.

Matching data at subnational unit levels (e.g., provinces or states) is another research approach. In a good example of this approach, McGavock (2021) uses regional variation in child marriage laws within Ethiopia. The sample paper in Appendix C uses another potential dataset for undergraduate research, The Global Terrorism database, to measure the influence of terrorist attacks on children's health and use of health care as measured by the DHS. This dataset provides information on terrorist attacks at the subnational unit level (in this case, Afghan provinces). This dataset can be used to examine the impact of terrorist attacks on demographic and health measures from the DHS. For example, Cahalan et al. (2020) look at the influence of terrorist attacks on women's employment in Afghanistan.

Subnational economic data may also be obtainable from individual countries. For example, another paper, co-written by one of the authors and a third student, uses district-level variation in foreign direct investment in Nepal to estimate its effects on women's employment Ziard et al. (Forthcoming).

3.8 Software

Instructors can choose among data analysis software (e.g., Stata, R, and Python) for students to conduct most applied economics research projects. In a review of submitted code for replications, Nguyen (2019) reported that two-thirds of papers published in top general-interest economics journals used Stata, followed by MatLab at one quarter; R and SAS each accounted for less than 5 percent. Instructors' knowledge of and comfort with a given software package will affect the time costs of preparing the class and likely the benefit for students, as instructors' research experience helps with teaching and debugging code. The DHS Program provides the most technical assistance in Stata, including an entire workbook with exercises (DHS Program 2019). Nonetheless, familiarity with R, SAS, and Python are also useful skills that will likely help students find employment in the private or government sector. Moreover, RStudio and Python can be used for free, while Stata's six-month student license currently costs US\$48 for students. One of the authors has taught a CURE with DHS using Stata and then later RStudio. As expected, Stata was easier for students to grasp and had fewer technical issues; however, the differences in price and potential job prospects for RStudio outweighed the reduction in technical issues.

As discussed above, the DHS can be linked to weather data, but there are also R packages that allow the use of some satellite data (e.g., see the "chirps" R package), which is an advantage of using R. Nonetheless, ArcGIS or QGIS can be used to perform features not included in the R packages for weather data and when using other satellite imagery datasets that do not offer R packages. ArcMap and QGIS are user-friendly, may not take a lot of time for students to understand the basics (we recommend the "clip" and "select" tools to start), and allow students to produce high-resolution maps without requiring coding abilities. Moreover, if students have some training in Python, the ArcPy integration in ArcMap is also a great tool to put their Python skills to good use and be able to reproduce maps and file transformations within ArcGIS using only Python scripts.

4 Writing, Presenting, and Publishing

Writing up data analysis helps students link their data work to research and hones their argument. We suggest that all students write a paper to codify their ideas and mirror the research process of the literature they read. Many opportunities exist for publishing including journals specifically for undergraduates in economics, though for many students only writing the paper as an assignment is sufficient. Students should have the opportunity to present in class or to peers. We discuss potential conferences below, but like submitting for publication, presenting outside of the classroom may not be for every student. We provide a set of goals and structure in Appendix A and a sample rubric for a final paper in Appendix B.8.

The goal of writing a research paper is to answer a question of interest to the reader. For a big-picture approach, Chaubey's (2018) book on writing research papers helps students with the RAP method: Research Question, Answer, and Position. This method requires students to ask a question, answer the question, and show how the paper positions itself in the literature. Applied economics papers have a standardized structure that can be used to help support the main goal by breaking the paper into specific parts.

Typically, applied economics papers have six sections based on the standard structure of economics papers (see Bellemare 2018; Neugeboren and Jacobson 2005). Appendix B.7 provides an example outline of a paper. This example is used in one of the author's CUREs and served as the basis for the creation of the sample paper in Appendix C. These sections are the abstract, introduction, conceptual framework, data description, econometric model, results, conclusions, and works cited.

Keith Head's (n.d.) introduction formula links well to Chaubey's (2018) RAP schema and is a good framework to give to students. The formula includes five paragraphs: the hook (why the topic matters), the question, the antecedents (literature review), the value added (or positioning), and the road map (what the rest of the paper does). In the next section, the conceptual framework, students should use previous research to set up hypotheses about why their main outcome of interest will be influenced by the variables they chose. This section will serve also as a literature review, but calling it a conceptual framework reminds students that the purpose is to set up the hypothesis.

For the data description, students need to include the five Ws: who was asked (define the unit of analysis), what they were asked (define the variables), where the survey was given (i.e., is it representative), when the survey was given (year and relationship to relevant events), and why the survey was conducted. Students can use their descriptive statistics (see Appendix B.5) produced with notes as a starter for this section. This section also helps set up the econometric model section that follows.

In writing the econometric model section, Nikolov (2022) suggests starting with the simplest regression—typically, this regression includes just the main outcome and main variable of interest—explaining why other variables might need to be added. Students will often need to learn the process of writing an equation to include subscripts for the unit of measurement of the variable (e.g., household, individual, or time). Having students write a short paragraph on the potential for endogeneity using the three main types (omitted variables, measurement error, and reverse causality) can strengthen the paper and reinforce these concepts.

In the results section, students typically expand on and revise the assignment in which they submitted their main results table. Usually, this includes one paragraph discussing the main variable, a second paragraph discussing other variables of interest, and a final paragraph discussing additional regressions, if undertaken.

To provide structure for the conclusion, Bellemare (2018) suggests a formula that mirrors the introduction formula discussed above. In this case, the four parts of the conclusion are a summary, a discussion of limitations, implications for policy, and implications for future research. Students, based on our experience, struggle the most with identifying what limitations matter; for economists, these limitations typically include issues around exogeneity. Asking students to imagine the ideal dataset for their research problem may help them analyze this issue.

4.1 Peer Learning and Review

In CUREs or instances where multiple students are working on theses at the same time, students can review each other's outlines and first drafts of papers. This exercise serves a few goals. First, students get additional feedback and ensure they are not missing any key sections. Second, instructors can use student peer feedback as a method for describing the peer review process in economics and how it functions. This is especially helpful for underperforming students because they get feedback from those who are performing better and, at the same time, higher-performing students can learn more by explaining concepts to other students.

4.2 Presenting

Students can present in class and at conferences using slides or posters. The presentation can help reinforce the concepts discussed in the student's research paper. Appendix B.9 provides a sample list of slides given to students in the DHS-based CURE for a final presentation. The slides are linked to the sixth section of the paper structure described above. Earlier in the semester, students can present the first half of these slides before they create their result tables, providing a check before students begin to outline or get too far into a paper.

These slides can also be transformed into a poster. In most semesters that one of the authors taught the DHS-based CURE, students presented their work in a university-wide sponsored poster forum. Poster sessions can also be done in class. The biggest gain of using posters is not having each student present for 10 minutes, which can take up many class periods for larger class sizes. In a class-based poster session, half of the students can be at their posters explaining their project, and the others can be consumers of the research. In either presentations or posters, having a simple assignment such as summarizing others' work can further support the peer review process.

4.3 Conferences and Publications

Students can gain further experience by presenting their work at conferences and submitting their research to journals. The regional economics associations (e.g., eastern, southern, midwestern, and western) hold meetings that often include undergraduate presentations. Of these meetings, the Eastern Economics Association has the most established undergraduate research session through the Issues in Political Economy Project, which hosted the 30th annual conference in 2024 at the Eastern Economics Association meetings (Elon University 2023). This group, based at Elon University, runs sessions that have undergraduate research in fields across economics, with many applied microeconomics sessions. The Federal Reserve Banks in Cleveland and Dallas have conferences that include undergraduate presentations as well. Moreover, specialized conferences for undergraduates underrepresented in economics can be found at Williams College (PIER) and the Sadie Collective conference for Black women interested in economics and related fields.

Students can ultimately publish their articles in both undergraduate research journals and academic journals. The American Economic Association lists seven potential economics-focused undergraduate research journals on its website. At Colorado College, undergraduate-coauthored articles have been published in well-regarded field journals such as *Energy Economics*, *Journal of Economic Geography*, and *Journal of Sports Economics* (Fenn et al. 2010). One of the authors of this paper published a work using DHS data with an undergraduate student in *Oxford Development Studies* (Cahalan et al. 2020). That said, there certainly are steep costs for instructors to publish student work in academic journals. Typically, review can take up to a year, by which point many students have moved on to careers or graduate school and have less time and interest in finishing their research papers, leaving the work of polishing a paper and responding to reviews to the instructor or mentor.

5 Conclusion

The DHS offers an excellent opportunity for undergraduate students to research under the guidance of experienced applied economists to examine some of the world's most important issues. As faculty, two of the coauthors have been able to turn this work into papers that have been published or are in the pipeline. We do acknowledge the high time cost to faculty of doing research with undergraduates and feel the DHS structure addresses many of these issues. In Appendix B, we provide numerous sample assignments for CUREs or theses to help with the structure of the paper; in Appendix C, we provide an example student paper. The third coauthor shares her own experience as a student and alumni in Appendix D, which we suggest potentially sharing with CURE or thesis students to demonstrate how to leverage undergraduate research into a career. Overall, the DHS datasets offer an excellent tool for undergraduate research.

Appendices: All appendices for the paper are downloadable as supplementary documents at aetrjournal.org.

About the Authors: Seth Gitter is a Professor at Towson University (Corresponding Author Email: sgitter@towson.edu). Chris M. Boyd is an Assistant Professor at Towson University. Raissa Abbasi is an Undergraduate Researcher at Towson University.

Acknowledgments: The authors would like to thank Towson University Office of Undergraduate Research and Creative Inquiry and the Faculty Academic Center for Excellence at Towson for Funding Support for CURE development at TU was provided by an Inclusive Excellence grant from the Howard Hughes Medical Institute. We also thank Amy Bekkerman for copy edits on an earlier draft. Finally we thank our editor and anonymous reviewers for comments to improve the paper. Seth Gitter is a Professor of Economics, Chris Boyd is an Assistant Professor of Economics and Raissa Abbasi was an undergraduate student at the time of the writing and has since graduated. Data collection, survey methods and outreach materials were reviewed and found exempt by the IRB (#IRB202200390) at Towson University.

References

- Agüero, J.M., and M.S. Marks. 2008. "Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks." *American Economic Review* 98(2):500–504. <https://doi.org/10.1257/aer.98.2.500>
- Andree, B.P.J. 2021. "Estimating Food Price Inflation from Partial Surveys." *Policy Research Working Papers 9886*. The World Bank. <https://doi.org/10.1596/1813-9450-9886>
- Bellemare, M.F. 2016. "Metrics Monday: There Is More than One Source of Endogeneity." Marc F. Bellemare (blog), January 11. Retrieved from <https://marcfbellemare.com/wordpress/11579>
- Bellemare, M.F. 2018. "Between the Introduction and the Conclusion: The 'Middle Bits' Formula for Applied Papers." Marc F. Bellemare (blog), January 17. Retrieved from <https://marcfbellemare.com/wordpress/12797>
- Brown, C., M. Ravallion, and D. Van De Walle. 2019. "Most of Africa's Nutritionally Deprived Women and Children Are Not Found in Poor Households." *The Review of Economics and Statistics* 101(4):631–644. https://doi.org/10.1162/rest_a_00800
- Cahalan, L., Gitter, S. R., & Fletcher, E. K. (2020). Terrorism and women's employment in Afghanistan. *Oxford Development Studies*, 48(2), 195-208.
- Chaubey, V. 2018. *The Little Book of Research Writing: The Structural Challenge of Communicating Knowledge + A Method to Meet It*. CreateSpace Independent Publishing.
- Chang, W. 2019. Abortion Laws and Life Choices of Young Women and Girls in Sub-Saharan Africa: A Cross-Country Analysis. Retrieved from <https://ipums.org/sites/www.ipums.org/files/chang.pdf>
- Corsi, D.J., M. Neuman, J.E. Finlay, and S. Subramanian. 2012. "Demographic and Health Surveys: A Profile." *International Journal of Epidemiology* 41(6):1602–1613. <https://doi.org/10.1093/ije/dys184>
- DHS Program. 2019. *DHS Program Data Analysis Workshop Stata Exercises (version 1.0)*. Retrieved from <https://dhsprogram.com/pubs/pdf/DHSC19/DHSC19.pdf>
- DHS Program. 2024a. "Journal Articles Search." Retrieved from <https://dhsprogram.com/Publications/Journal-Articles-by-Journal.cfm>
- DHS Program. 2024b. "Links to Data Sets." Retrieved from https://dhsprogram.com/data/Using-DataSets-for-Analysis.cfm#CP_JUMP_14036
- Dudenhefer, P. 2006. "Writing a Literature Review." Duke University EcoTeach Center. Retrieved from <https://econ.duke.edu/sites/econ.duke.edu/files/documents/Writing Literature Reviews.pdf>
- Elon University. 2023. "30th Annual Conference Information." Issues in Political Economy (blog). Retrieved from <https://blogs.elon.edu/ipe/conference/about/>
- Fenn, A.J., D.K.N. Johnson, M.G. Smith, and J.L. Stimpert. 2010. "Doing Publishable Research with Undergraduate Students." *Journal of Economic Education* 41(3):259–274.
- Gehrke, E., and C. Kubitzka. 2024. "Agricultural Productivity and Fertility Rates: Evidence from the Oil Palm Boom in Indonesia." *Journal of Human Resources* 59(1):316–347.
- Gitter, S. 2021. "A Guide for Student-Led Undergraduate Research in Empirical Micro-Economics." *Journal of Economics Teaching* 5(3):83–115. <https://doi.org/10.58311/jeconteach/64926ee263799eb22daaf874770ea78fe6d93e00>
- Gitter, S., Onyemeziem, O., & Corcoran, W. (2023). *Menarche, Marriage Age, Education, and Employment in Africa, the Middle East, and Central Asia*. Towson Working Paper No. 2023-04.

- Head, Keith. n.d. "The Introduction Formula." Keith Head Research Advice (blog). Retrieved from <https://blogs.ubc.ca/khead/research/research-advice/formula>
- Huntington-Klein, N. 2022. *The Effect: An Introduction to Research Design and Causality*. Boca Raton, FL: CRC Press, Taylor & Francis.
- Ismay, C., and A.Y. Kim. 2019. *Statistical Inference Via Data Science: A Modern Dive Into R and the Tidyverse*. Boca Raton, FL: CRC Press.
- Jayachandran, S., and R. Pande. 2017. "Why Are Indian Children So Short? The Role of Birth Order and Son Preference." *American Economic Review* 107(9):2600–2629. <https://doi.org/10.1257/aer.20151282>
- Källmark, L. (2018). How does drought affect child health outcomes in Zimbabwe?. Retrieved from <https://core.ac.uk/download/pdf/211581107.pdf>
- Khanal, B. 2022. "The Impacts of the 2015 Gorkha Earthquake on Children's Health in Nepal." *World Development* 153:105826.
- Mayala, B., and R. Donohue. 2022. *The Geospatial Covariate Datasets Manual*, 3rd ed. Washington, DC: USAID, Demographic and Health Surveys Program.
- Ministry of Health (MoH) [Tanzania Mainland], Ministry of Health (MoH) [Zanzibar], National Bureau of Statistics (NBS), Office of the Chief Government Statistician (OCGS), and ICF. 2022. Tanzania Demographic and Health Survey and Malaria Indicator Survey 2022 Final Report. Dodoma, Tanzania, and Rockville, Maryland, USA: MoH, NBS, OCGS, and ICF.
- McGavock, T. 2021. "Here Waits the Bride? The Effect of Ethiopia's Child Marriage Law." *Journal of Development Economics* 149(March):102580. <https://doi.org/10.1016/j.jdeveco.2020.102580>
- Mruts, K. B., Gebremedhin, A. T., Tessema, G. A., Scott, J. A., & Pereira, G. (2022). Interbirth interval and maternal anaemia in 21 sub-Saharan African countries: A fractional-polynomial analysis. *PLoS One*, 17(9), e0275155.
- Neugeboren, R.H., and M.E. Jacobson. 2005. *Writing Economics*. Cambridge, MA: Harvard University.
- Nguyen, A. 2019. "Econ Journals and Software." *mentalBreaks*, March. Retrieved from <https://mentalbreaks.rbind.io/posts/economics-journals-and-software/>
- Nikolov, P. 2022. "Writing Tips for Economics Research Papers: 2021–2022 Edition." *IZA Discussion Paper 15057*. <https://doi.org/10.2139/ssrn.4114601>
- Romem, I. 2012. "Econ 191: Writing a Literature Review," February 7. Retrieved from https://eml.berkeley.edu/~webfac/eichengreen/e191_sp12/romem_econ191_2-7-12.pdf
- Shapiro, J.M. 2022. "Four Steps to an Applied Micro Paper." Retrieved from <https://scholar.harvard.edu/files/shapiro/files/foursteps.pdf>
- Stanford Graduate School of Business. n.d. "Writing Content Notices for Sensitive Content." Teaching and Learning Hub. Retrieved from <https://tlhub.stanford.edu/docs/writing-content-notices-for-sensitive-content/>
- Stoebenau, K., Madhavan, S., Smith-Greenaway, E., & Jackson, H. (2021). Economic inequality and divergence in family formation in sub-Saharan Africa. *Population and development review*, 47(4), 887-912.
- Sunder, N. 2019. "Marriage Age, Social Status, and Intergenerational Effects in Uganda." *Demography* 56(6):2123–2446. <https://doi.org/10.1007/s13524-019-00829-8>
- University of Michigan. n.d. "An Introduction to Content Warnings and Trigger Warnings." Retrieved from <https://sites.lsa.umich.edu/inclusive-teaching-sandbox/wp-content/uploads/sites/853/2021/02/an-introduction-to-content-warnings-and-trigger-warnings-draft.pdf>

- Vaessen, M., Thiam, M., & Lê, T. 2005. Chapter XXII The Demographic and Health Surveys. *United Nations Statistical Division, United Nations Department of Economic and Social Affairs*.
- Varian, H.R. 2016. "How to Build an Economic Model in Your Spare Time." *The American Economist* 61(1):81-90.
- Ziard, S., S.R. Gitter, and M.A. Groves. Forthcoming. "Foreign Direct Investment and Women's Employment in Nepal" *Undergraduate Papers Monthly Labor Review*.