Trends in Female Schooling in 92 Low and Middle Income Countries: Coherence across Data Sources

Kristen Jeffers[*†], Albert Esteve[*], and Joan García[*]

Introduction and Background

Over the last several decades, substantial investments have been made towards achieving gender parity in education. Beyond the self-evident merit of gender equality, female education has been linked to reduced infant and child mortality, improved maternal and child health, economic growth, poverty alleviation, and human development (Behrman 2015; Behrman et al. 2017; Cleland and Van Ginneken 1988; Fuchs et al. 2010; Hill and King 1995; Klasen 2002; Klasen and Lamanna 2009; Wang 2003;). One mechanism linking female education and positive social and economic outcomes is that more highly educated women have fewer children than women with less schooling (Cochrane 1979; Hirschman 1994; Castro 1995; Jejeebhoy 1995; Ainsworth et al. 1996; Lloyd 2005; Shapiro 2012; Bongaarts, Mensch, and Blanc 2017). Indeed, in many developing countries fertility decline is a secondary goal of policies that support female schooling (Grant 2015; Wusu and Isiugo-Abanihe 2018). Evidence suggests the impact of female schooling on fertility and other family formation outcomes varies among countries. The magnitude, and in some cases, direction of the association at the individual level varies significantly across contexts. Likewise, individual-level associations are not consistently mirrored at the population level (Esteve and Florez-Paredes 2018). Disparate national contexts explain some of this variation, but it is also important to consider the role of measurement in variation across time and place.

For many low and middle income countries (LMICs) that lack comprehensive population registers, census and household survey data represent the only available data on educational attainment of the population. These data serve as the source for global databases on educational attainment like those maintained by the United Nations. Differences in education systems across countries, changes to education systems over time, varying data-collection and data-processing practices, and competing national, regional, and international classification standards make these data difficult to compare across time and place. Scholars considering education at the global level have grappled with and developed methods for dealing with issues of comparability across data sources, but none have systematically evaluated these issues across the main sources of demographic data and made their findings known to the research community.

This paper seeks to fill this gap by providing a systematic evaluation of trends in female educational attainment across the three primary sources of demographic data in low and middle income countries: population censuses, Demographic and Health Surveys (DHS), and UNICEF Multiple Cluster Indicator Surveys (MICS). We aim to identify coherence and inconsistencies in

[*] Centre d'Estudis Demogràfics, Universitat Autònoma de Barcelona
[†] Minnesota Population Center, University of Minnesota

the measurement of education within and across data sources and to contribute to a better understanding of the validity of a critical independent variable in demographic research.

Data and Methods

Individual-level census data and survey data available from the DHS Program and the UNICEF MICS surveys provide cross-sectional information on educational attainment and other individual and household characteristics for more than 100 countries around the world. These data are the primary sources for international estimates of educational attainment for countries lacking robust administrative data. These data are also widely used by scholars investigating the relationship between education and demographic, social, and economic outcomes.

IPUMS-International provides nationally representative, high-density (typically 10 percent of a country's population) samples of census microdata for more than 70 LMICs. Data available from IPUMS are harmonized to support comparative analysis across time and space. IPUMS also disseminates original data that preserve country-specific classifications, allowing researchers to construct customized indicators and aggregations that are comparable with other sources. DHS surveys provide nationally representative microdata for more than 90 LMICs on topics related to population, health, and nutrition. Also nationally representative, MICS surveys provide DHS-like information for more than 100 countries, many of which are not included in DHS and other global survey programs. Most IPUMS census samples and DHS and MICS surveys include variables that indicate years of schooling and highest level of education completed. Differences in question wording and sampling strategies represent potential sources of discrepancies across data sources. The final paper will include a comprehensive discussion of sampling methodology used in each source and an assessment of questionnaire structure that may influence observed discrepancies in trends across data sources.

We will include in the analysis data for the 92 countries that are represented across two or more of these data sources (summarized by region in table 1).

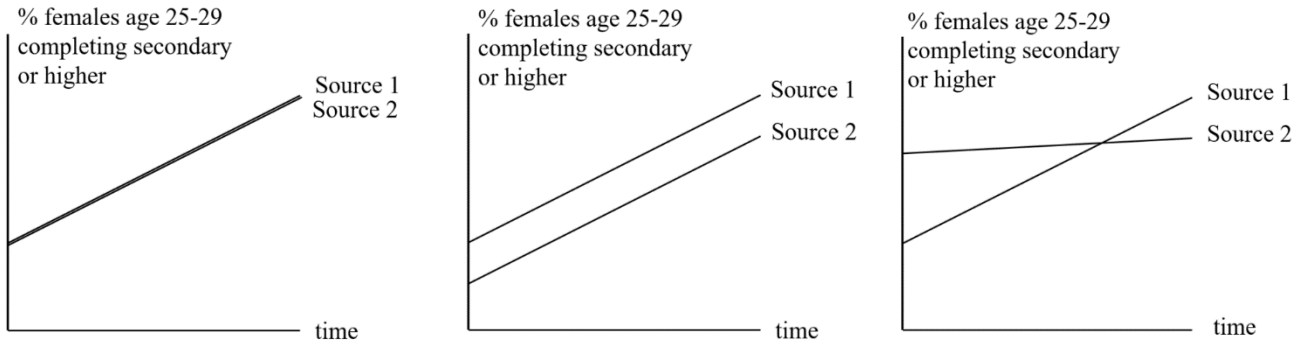
Table 1: Data availability across IPUMS, DHS, and MICS by world region (N = number of countries with available microdata)

	Africa	Asia	Europe	Latin America/ Caribbean	**Total**
DHS, MICS, IPUMS	22	10	1	7	40
IPUMS, DHS	2	2	–	6	10
IPUMS, MICS	1	3	1	8	13
DHS, MICS	17	8	2	2	29
TOTAL	42	23	4	23	**92**

The analysis will consist of two parts. The first phase will document trends in educational attainment among women age 25 to 29 in 92 countries. To account for variation in national education systems and maximize comparability across time, space, and data sources, we will construct a dichotomous measure of educational attainment that distinguishes individuals who have completed secondary school or higher from those who have not. We will focus on women age 25 to 29. This age group represents the population that recently completed formal education, more accurately reflecting the educational context in the country at the time of the survey

compared to older cohorts. Based on the results of this first phase of analysis, we will classify countries into three groups based on coherence within and across data sources, summarized in figure 1.

Figure 1: Country classification schema of coherence of trends in female educational attainment within and across data sources



Coherent trends, levels: Change over time and point estimates match across data sources

Coherent trends, incoherent levels: Change over time match across data sources, point estimates do not

Incoherent trends, incoherent levels: Neither change over time or point estimates match across data sources

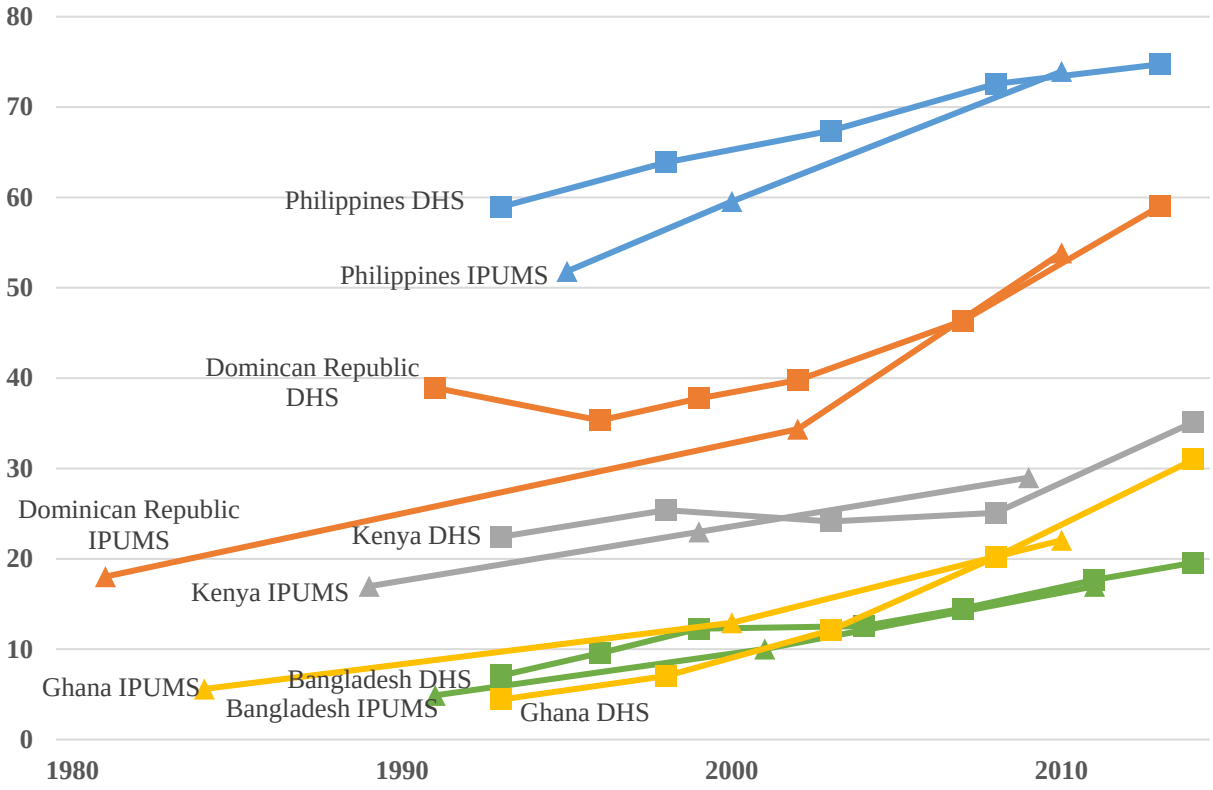
The second phase of the analysis will rely on logistic regression to identify the source of discrepancies for a subset of countries. Using data pooled across data sources, we will execute a series of logistic regression models to estimate the odds of completing secondary school or higher controlling for data source as well as other variables that may influence data quality for women such as urban-rural status, marital status, household size, and labor force participation.

Preliminary Results

Figure 2 presents preliminary results of the descriptive analysis for a subset of five countries. The final paper will present similar results for 92 countries and include observations based on MICS data where available.

Data from IPUMS and DHS indicate that the percentage of females age 25 to 29 completing secondary or higher has increased for all countries. Point estimates and trajectories across time vary within county across data sources. Estimates calculated with DHS data tend to be higher than estimates calculated with census data. Preliminary results also indicate that the period of time between the previous census and a DHS survey influences the coherence of point estimates. In the Dominican Republic, for example, the 1991 DHS survey was conducted 10 years after the previous census of the country in 1981 (and two years before the 1993 census for which microdata are not available) and provides an estimate of female secondary completion that is higher than 1996 and 1999 DHS estimates and significantly divergent from census-based trends. We will explore the role of temporal proximity of censuses and surveys in the final paper. We will also present the results of logistic regression analyses that model the relationship between data source, other individual level characteristics, and educational attainment.

Figure 2: Percent of females age 25 to 29 completing secondary education or higher, by country and data source (IPUMS and DHS)



References

Ainsworth, M., Beegle, K., & Nyamete, A. (1996). The impact of women's schooling on fertility and contraceptive use: A study of fourteen sub-Saharan African countries. *The World Bank Economic Review*, 10(1), 85-122.

Behrman, J. A. (2015). The effect of increased primary schooling on adult women's HIV status in Malawi and Uganda: Universal Primary Education as a natural experiment. *Social science & medicine*, *127*, 108-115.

Behrman, J. A., Peterman, A., & Palermo, T. (2017). Does keeping adolescent girls in school protect against sexual violence? Quasi-experimental evidence from east and southern Africa. *Journal of Adolescent Health*, *60*(2), 184-190.

Bongaarts, J., Mensch, B. S., & Blanc, A. K. (2017). Trends in the age at reproductive transitions in the developing world: the role of education. Population studies, 71(2), 139-154.

Castro, T. (1995). Women's education and fertility: results from 26 Demographic and Health Surveys. *Studies in family planning*, 187-202.

Cleland, J. G., & Van Ginneken, J. K. (1988). Maternal education and child survival in developing countries: the search for pathways of influence. *Social science & medicine*, *27*(12), 1357-1368.

Cochrane, S. H. (1979). *Fertility and education: What do we really know?* Washington, DC: Johns Hopkins University Press.

Esteve, A., & Florez-Paredes, E. (2018). The stability paradox: Why expansion of women's education has not delayed early union formation or childbearing in Latin America. *Studies in family planning*, 49(2), 127-142.

Fuchs, R., Pamuk, E., & Lutz, W. (2010). Education or wealth: which matters more for reducing child mortality in developing countries?. *Vienna Yearbook of Population Research*, 175-199.

Grant, M. J. (2015). The demographic promise of expanded female education: trends in the age at first birth in Malawi. Population and development review, 41(3), 409-438.

Hill, M. A., & King, E. (1995). Women's education and economic well-being. *Feminist Economics*, *1*(2), 21-46.

Hirschman, C. (1994). Why fertility changes. Annual review of sociology, 20(1), 203-233.

Jejeebhoy, S. J. (1995). Women's education, autonomy, and reproductive behaviour: Experience from developing countries. OUP Catalogue.

Klasen, S. (2002). Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development. *The World Bank Economic Review*, *16*(3), 345-373.

Klasen, S., & Lamanna, F. (2009) The Impact of Gender Inequality in Education and Employment on Economic Growth: New Evidence for a Panel of Countries, *Feminist Economics*,15(3), 91-132.

Lloyd, C. (2005). *Growing up global. The transition to adulthood in developing countries.* Washington DC: National Academies.

Shapiro, D. (2012). Women's education and fertility transition in sub-Saharan Africa. *Vienna Yearbook of Population Research*, 9-30.

Wang, L. (2003). Determinants of child mortality in LDCs: empirical findings from demographic and health surveys. *Health policy*, *65*(3), 277-299.

Wusu, O., & Isiugo-Abanihe, U. C. (2018). Consistency of the effects of female education on fertility across the north–south demographic divide in Nigeria, 2003–2013. *Journal of biosocial science*, 1-16.