Working out the kinks: Statistical constraints and Millennium Development Goal evaluation

Abstract: The Millennium Development Goals (MDGs) have been set up for measurability and accountability. With their deadline now passed, empirical evaluation can take advantage of the indicators and the data they provide, especially since the MDGs have been structured in such a way to compare pre- and post-treatment results. However, statistical constraints in the form of availability, quality, and predictive ability all create roadblocks. This paper explores the possibilities and challenges of evaluating the MDGs using MDG 3, gender equality and women's empowerment, as its focus. It starts by using the MDGs as a natural experiment in sub-Saharan Africa, testing for structural breaks and kinks, showing that although geared towards this purpose, the MDG architecture falls short of its goal. It then addresses the statistical roadblocks for doing so and separates the theoretical from the practical uses of the MDGs' indicator-based structure. It argues that while MDG indicators are built to be measurement tools, they are in effect framing devices that have important implications for empirical evaluation.

Introduction

      The Millennium Development Goals (MDGs), which are the eight global goals set by the United Nations meant to serve the needs of the world's poor, were meant to signify a commitment to equity and inclusion (Ki-moon, 2011). The third MDG (MDG 3) is aimed specifically at women, promoting gender equity and women's empowerment through targeting gender inequalities in education, employment, and national parliaments. This lofty and noble goal has been set for every country with the aim of completion by 2015. If achieved, it would mark a crucial starting point in the fight against discriminatory social institutions biased against women. It is particularly relevant for sub-Saharan African countries, where gender-based discrimination is a pervasive societal issue and creates barriers to women's education and participation in politics and the labour market.

      As this paper will show, however, using the MDGs as this starting point has limited discernible results. Statistical constraints hinder MDG evaluation, rendering theoretically sound quasi-experimentation far less useful in practice. Constraints like low predictive ability, stemming from poor data quality and availability, can create serious problems for indicator-based projects like the MDGs. Such barriers to evaluation are at odds with how the MDGs are set up, as the design lends itself to pre- versus post-treatment style testing. Due to these issues, this paper argues that MDG indicators have acted more as framing devices than as strictly tools of measurement and benchmarking. They frame our understanding of an issue area, such as gender equality, and thus guide policy prescriptions. As such, this paper extends the discussion brought forward by Jacob (2017), in terms of the effect of a data gap on MDG success, by exploring the impact of this alternative use of indicators.

The paper is organized as follows. First, background information is provided on the MDGs and their main critiques, followed by a closer look at MDG 3 both in general and in the context of sub-Saharan Africa. The regression discontinuity approach used in this paper is then described along with the data. Results are presented and analyzed on their own and in the context of the predictive ability of the data, which is uneven at best. A discussion follows based on the issues presented by low data availability and quality, informing the notion of an indicator's value as a framing tool over a measurement tool. Possible ways forward are also presented, in terms of evaluation for both the MDGs and the Sustainable Development Goals (SDGs).

Background

The Millennium Development Goals

The United Nations' MDGs represent a global anti-poverty campaign based on eight goals. The goals have received national-level support with state governments signing on to the associated Millennium Declaration in 2000, with the aim of achieving the goals by 2015. Each goal has one or more associated targets. These targets make the goals' aspirations more tangible. Measurability of progress towards these goals is encapsulated in the indicators. Looking at MDG 3, we see the following:

Goal: Promote gender equality and empower women

Target: Eliminate gender disparity in primary and secondary education, preferably by 2005, and in all levels of education no later than 2015.

Indicator 3.1: Ratios of girls to boys in primary, secondary and tertiary education

Indicator 3.2: Share of women in wage employment in the non-agricultural sector

Indicator 3.3: Proportion of seats held by women in national parliament (UNSD, 2008)

The creators of the MDGs took broad aspirations, like gender equality in this case, and translated them into specific, measurable numbers in the indicators. Underlying the indicators are normative viewpoints on how to best solve the problem the goal presents. The aim of MDG 3 is to promote women's empowerment, with the target and primary indicator focused on gender parity in education. Women's education is implied to be an important component of women's empowerment and gender equality. This indicator communicates gender parity in education as a priority and a preferred approach to empowering women, leading policymakers to pursue it over other approaches.

The MDGs have been constructed using goals and targets unconstrained by strict policy prescriptions. Setting targets and standards for development is a useful tool for global governance organizations. This type of tool is used to influence and steer policy at a global level while leaving states to autonomously produce their own internal policies. The way in which a global governance institution influences national policy is increasingly important given what Jakobi (2009) describes as the rise of global public policy, where education and other policy fields increasingly look to international actors for direction. Education policy in particular has been rescaled to the international level, with global policy discourses through global education policy networks influencing national policy production and forcing the state to work in new ways (Lingard & Rawolle, 2011). Verger et al. (2012) take this point further from a global governance perspective, in that global governance institutions in the field of education are redefining the relationship between the state and education policy. By operating at different scales, they argue that non-state actors are increasing their presence and authority in the field of education policy as well as in other policy fields.

Why education and gender? A closer look at MDG 3

Education and gender are presented in the MDGs as issue areas with separated but linked goals. MDG 2 is to achieve universal primary education. The associated target does bring gender into the picture by specifying that "by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling" (UNSD, 2008). MDG 3 is aimed specifically at women, promoting gender equality and women's empowerment via the sole target of eliminating gender disparities in all levels of education in terms of enrolment. By virtue of how MDG 3 is set up, much of the conversation around it relates to women's education specifically.

The focus on education in MDG 3 contains embedded assumptions about how creating equal access to education can contribute to gender equality. The idea behind it is that achieving gender parity in education breaks down gender-based norms that constrain the rights and freedoms of women (Subrahmanian, 2005). Gender equality in education can have positive ramifications for other issues impacting women, including nutrition, child mortality, and fertility, as well as economic growth. Conversely, Abu-Ghaida and Klasen (2004) found that countries that would not achieve the MDG goal of gender equity in primary and secondary education would see increases in child mortality, prevalence of underweight children under five, and fertility rates, combined with decreases in per capita growth rates.

There are also deeper historical roots tying education and gender concerns together. The MDGs were preceded by a variety of conferences and declarations that centered on one of these issues but brought in the other issue. With the World Declaration on Education for All, there was a call to improve access to education for women and girls as well as the quality of their education. It is even listed as their 'most urgent priority' (UNESCO, 1990). On the other side of

the equation, the Beijing Platform for Action, from the UN's Fourth World Conference on Women, includes 'Education and Training of Women' as one of the categories of action. Objectives include equal access to education, promoting lifelong learning, and ensuring proper resources and monitoring of educational reforms (United Nations, 1995). These declarations and conferences of the 1990s, along with other events and global development projects, helped shape the MDGs (Hulme, 2009). Since Beijing, the women's movement has moved beyond its initial focus on education to issues of disempowerment. MDG 3, however, has been accused of moving backwards. As Fukuda-Parr and Hulme (2011) put it, the MDGs "took the agenda back to the priorities of the 1970s" (pg. 27).

MDG 3 has been the subject of a broad range of critiques, from its underlying assumptions to its focus on outcomes. Connell (2010) provides three assumptions underlying MDG 3 that have been seriously challenged: gender as an unproblematic binary, formal education as an unquestionable good, and targeting girls and women without strong inclusion of boys and men. North (2010) takes issue with how MDG 3 has taken the complicated idea of gender equality and reduced it to quantifiable targets based only on gender equality in education. In terms of MDG 3 as a strategy for gender equality, Chismya et al. (2012) argue that MDG 3 alone is not enough to achieve gender equality or women's empowerment because inequitable gender norms exist within and outside the education system. Further, they posit that gender parity in the classroom does not necessarily translate directly into reducing gender-based discrimination in the community. Gender inequalities reflect societal norms affecting women throughout their lives and these norms are held by individuals at school including children themselves.

The MDGs' approach to education in general has also been called into question. MDG 2, the goal of universal primary education, puts a primary measurement focus on enrolment. None of the indicators strictly measure completion; the closest way to measure primary school completion is by accounting for the final-year enrolment rate of those enrolled at the start of primary school. Going further, there is no measure of educational quality. None of the indicators, nor the approach more broadly, address learning or competency (see Filmer, Hasan, & Pritchett, 2006). There may also be problems from a policy-making point of view. MDG 2 does not provide strict policy guidelines, only indicators of success, but education policy creation is difficult as there is no concrete list of policies or determinants of educational outcomes and few clear guidelines exist in the literature (World Bank, 2004). Bourguignon *et al.*'s report (2008) argues for a coherent policy framework between a country's national development strategy and MDG-related strategy, noting that policymakers know more about what not to do than which policies should be used. There has also been a focus on 'quick wins', which sits in tension with and distracts from medium- and long-term processes (Oya, 2011; Richard et al, 2011). These concerns compound the gendered concerns raised above.

The issues found in MDG 3 echo those within the MDGs in general; problematic underlying assumptions for indicators, narrow focus on outcomes, and oversimplification of problems. This goal also makes a strong case for the importance of data quality and statistical capacity. Data disaggregated by gender, in this case, is what will illuminate progress or gaps. As such, it provides a strong example to use for evaluating MDG evaluation possibilities. In addition, the post-2015 agenda has been significantly impacted by women's groups, who were some of the most active in the SDG formation process (Gabizon, 2016), and gender is a major

cross-cutting theme in the SDGs. Studying MDG 3 in the post-2015 context highlights how the previous goals have influenced the current set.

MDG 3 in sub-Saharan Africa

Sub-Saharan Africa has been selected as the region of study. Gender-based discrimination has proven to be a pervasive societal issue in sub-Saharan Africa, as girls and women in the region face decreased welfare in comparison with boys and men (Kevane, 2004). Mikell (1997) notes that while subordination of women has roots in traditional cultures from across Africa, colonialism has worsened gender-based discrimination. Connell's (2010) work reinforces this point, noting that imperial history is a gendered history in that imperialism in Africa destroyed traditional gender orders and created new ones.

In terms of gender-based barriers to education in the region, barriers are created by reasons existing outside of the education system, including beliefs about the 'roles' of women (Hyde, 1997). These norms and beliefs are both created and reinforced by discriminatory social structures. Chismya et al. (2012) note that gender-based expectations of women as mothers and wives are linked to barriers to schooling. They point out the argument that 'overeducated' women decrease their marriageability as the skills needed to be a wife and mother are not taught in formal education systems. Further, marriage indicates the end of schooling (Hyde, 1997) and sub-Saharan Africa has the highest rate of child marriage of any region in the world (Walker, 2012). This may explain why few countries in the region have achieved parity in secondary school (see Lewin, 2009). Breaking down these barriers would have a significant impact on education levels of future female children. For example, Glick and Sahn (2000) have shown that in West Africa, a mother's education level has a significant impact on her daughter's education level but not her son's.

While the MDGs may have been able to address some of these issues, there are many problems in their design. Easterly (2009) argues that the MDGs have been set in a manner that is unfair to Africa, given the focus on absolute changes as opposed to relative changes, and level targets as opposed to targets based on change. Based on this style of target setting, the MDG-related progress in African countries appears to indicate failure because they are unlikely to reach the extremely challenging targets set for them, based in part on initial conditions. This pass/fail view diminishes the progress that has been made, and does not recognize the advances African countries have actually realized. In addition, as Richard et al. (2011) point out, MDG implementation needs to be based on improving the partnership between wealthier and less wealthy countries, as opposed to just having money change hands. Finally, Jerven's (2013) work showcases the low statistical capacity of sub-Saharan African countries. This will prove to create a significant hurdle for such a data-driven project.

Summary statistics

Table 1: Female education outcomes

	1990	2000	2014
Gender parity: primary			
Average	0.82	0.86	0.96
Min	0.44	0.61	0.85
Max	1.25	1.05	1.09
Gender parity: secondary			
Average	0.71	0.77	0.87
Min	0.22	0.29	0.62
Max	1.57	1.36	1.37
Gender parity: tertiary			
Average	0.38	0.59	0.91
Min	0.09	0.17	0.20
Max	1.45	1.72	2.48
Adult literacy rate, female (%)			(for 2015)
Average	53.14	54.60	60.29
Min	27.52	12.80	11.04
Max	74.87	92.05	95.70
Primary completion rate, female (%)			

Average	39.26	47.14	71.89
Min	6.34	12.55	44.57
Max	113.52	104.08	108.06

All of the indicators for female education outcomes improved over the 1990 to 2014 time period, particularly during the MDG timeframe (post-2000). Gender parity at the tertiary education level made the most dramatic improvement, although the minimum level remained low. Encouragingly, the female primary completion rate also shows significant positive improvement. This metric is one of the indicators for MDG 2, to achieve universal primary education. Literacy rates are improving at a much slower pace. This may be connected to Filmer, Hasan, and Pritchett's (2006) critique that the goals do not address learning outcomes or educational quality.

Table 2: Women's rights and empowerment outcomes

	1990	2000	2014[1]
Share of women in non-agricultural wage employment: MDG 2 indicator 2 (%)[2]			
Average	20.44	32.42	35.09
Min	3.8	18.6	18.3
Max	41	42.9	47.3
Proportion of seats held by women in national parliament: MDG 2 indicator 3 (%)			
Average	8.72	9.64	21.13
Min	0	0	3
Max	20	30	63.8
Women's economic rights (0-3)[3]			
Average	0.92	1.13	0.80
Min	0	0	0
Max	2	2	3
Women's political rights (0-3)[3]			
Average	1.51	1.92	2.06
Min	0	1	1
Max	2	0	3
Women's social rights (0-3)			
Average	0.79	0.84	Data not
Min	0	0	available
Max	2	2	

The two secondary indicators for gender equality and women's empowerment also show improvements over time but significantly less dramatic than in the case of their education counterparts. The level of women in non-agricultural wage employment actually improved far more before the start of the MDGs than in the time after, although the maximum levels in the dataset have shown greater growth. Female representation in national parliaments has more than doubled since the start of the MDGs, which is encouraging given that elections are not yearly and would have fewer individual opportunities to grow. Some countries are clearly lagging behind, given the spread between the minimum and maximum levels, although no country is without female representation at that level by 2014.

In terms of rights, each indicator gives a different story. The rights indicators are from the Cingranelli-Richards (CIRI) Human Rights Dataset (Cingranelli & Richards, 2013). Each of the rights variables are indicators based on scores of 0 to 3, with 0 indicating no rights and systematic discrimination and 3 indicating either all or almost all rights are both guaranteed and enforced.[1] Women's economic rights have seen uneven growth over the time period, reaching their lowest average level in 2010 over the three years presented. Women's political rights have improved with some countries even reaching 'full' political rights for women according to the indicator. Women's social rights are the lowest, although this variable was retired in 2005 so a representation of the present reality is unavailable.

Methodology and data

In order to determine the effectiveness of MDG 3 in sub-Saharan Africa, it is useful to think of the MDGs as a form of intervention or treatment used in a natural experiment. The

---

[1] For full descriptions of all rights accounted for in these indicators, see Cingranelli and Richards (2008).

structure of the MDGs lends itself to this approach, with a potential disruption for development outcomes in the year 2000. The following model is used to test for regression discontinuities, both in terms of breaks and kinks:

$$\alpha_{it} = \beta_1 + \beta_2 year + \beta_3 X_{it} + \beta_4 MDG + \beta_5 change + \varepsilon_{it}$$

where α is the relevant MDG 3 indicator, β_1 is the constant term, *year* is the forcing variable (starting with 1985 = year 1), *X* is a set of control variables to account for existing conditions for women, *MDG* is the dummy variable indicating the treatment (testing for a break), and *change* is a variable to account for the rate of change of the slope (testing for a kink). This model is tested controlling for country fixed effects.

The control variables take a rights-based approach to measuring existing conditions for women. They have been selected based on the Social Institutions and Gender Index (SIGI) (Cerise & Francavilla, 2012), created by the Organisation for Economic Co-operation and Development, as part of Van Staveren's (2013) Women's Empowerment Model (WEP). The SIGI is presented as a measure of the root causes of gender inequalities, as they relate to development, based on social institutions. The WEP takes these root causes into account in their relation to women's agency, which leads to women's achievements. Van Staveren (2013) also uses many of the root causes, such as land rights and violence against women, in an empirical analysis to test their relationships with women's access to education and employment. While the SIGI, and the root causes it identifies, would be ideal in the model used in this current study, data is limited and no data is available for the 1990s. Thus, the SIGI was not used directly in this model but was used to aid in variable selection. The independent variables in this paper's model are comprised of indicators from the Cingranelli-Richards (CIRI) Human Rights Dataset

(Cingranelli & Richards, 2013); the same as the rights indicators provided in the summary statistics as described above.

The panel dataset includes 47 sub-Saharan African countries for years 1985 to 2015. Data has been primarily sourced from within the UN system, including the UN's official MDG database as well as UNESCO and the World Bank. The only exception is the CIRI database described above. It is important to note that the dataset is not complete and that there are many gaps. This problem will be addressed throughout the results and discussion sections as it creates a significant barrier to evaluating MDG outcomes.

Results and analysis

Regression discontinuity results

Table 3: Results[2]

	Gender parity (primary)	Gender parity (secondary)	Gender parity (tertiary)	Non-agri wage labour	Seats in parliament
Year	0.003***	0.009***	0.010***	0.438***	0.013
	(0.000)	(0.000)	(0.000)	(0.000)	(0.901)
Econ rights	0.013**	0.024***	-0.0003	-0.533	1.046*
	(0.028)	(0.004)	(0.981)	(0.214)	(0.082)
Poli rights	-0.003	-0.032***	-0.002	0.120	3.708***
	(0.548)	(0.000)	(0.887)	(0.789)	(0.000)
Social rights	-0.026***	-0.033***	0.001	0.091	1.561***
	(0.000)	(0.000)	(0.926)	(0.848)	(0.005)
MDG	-0.056*	0.061	-0.0464	4.955**	-13.739***
	(0.066)	(0.143)	(0.585)	(0.029)	(0.000)
Change	0.004**	-0.005**	0.004	-0.308**	0.922***
	(0.012)	(0.031)	(0.408)	(0.034)	(0.000)
Constant	0.798***	0.700***	0.367***	23.988***	-0.280
	(0.000)	(0.000)	(0.000)	(0.000)	(0.858)
N	681	469	376	124	354

Note: *** $p \le 0.01$, ** $p \le 0.05$, * $p \le 0.1$, standard errors in brackets

---

[2] Breaks and kinks have also been tested for years following 2000, to account for policy lag, but the strongest results are found for 2000.

The first three outcomes are all related to gender equality in education. MDG 3 put particular emphasis on women's education, as the sole target for Goal 3 was to eliminate gender disparity in all levels of education. Results are mixed: a discontinuity in the form of a kink in the year 2000 is most convincing, based on statistical significance, but results are weaker for higher levels of education. Weaker results for higher educational levels is intuitive, because it takes time for the more gender-balanced cohorts to move through the system. Finding kinks instead of breaks is also intuitive, as we would expect results to be less dramatic given the time and resources required to break down educational barriers for girls.

Female participation in the non-agricultural wage labour force tells a different story. There is both a break and a kink at 2000, with a large jump at 2000 but then a slowing of progress over time. Social norms are strongly at play with this indicator, given women's predominance in agricultural labour at home and the related norms around women's role in the family and in society. The decreasing speed of progress over time may be in reaction to MDG-based labour policies initiated early in the process. These results must be interpreted with caution, however, given the significant lack of data. Approximately 87% of the observations are missing; a disappointing statistic for an official MDG indicator.

The results for women's representation in national parliament are the most convincing, with a highly statistically significant result for both a break and a kink. Note, too, that there is no significant time trend driving the results.[3] The bounce down at the year 2000 may seem strange, but is driven by several election results. 8 of the 47 countries saw multi-percentage point drops when comparing levels in 1999 with those between 2000 and 2003. It was not until the mid-2000s that there were greater changes, with increases in both max and min values. Given that

---

[3] There is also no time trend when the rights-based control variables are omitted.

elections are not held yearly, a longer time horizon is required to see significant improvements in this category.

Overall, there is some evidence that the introduction of the MDGs had a measurably positive impact on gender equality, based on their own measures. How much, however, can we truly discern from these results? There are two ways to approach this concern. The first is by taking into account the predictive ability of the data and the model. Identifying a structural break is one way to look for predictive ability; in this case, could we have predicted the outcome for 2015 based on data and trends leading up to 2000? Put differently, did we have enough predictive ability in 2000 to now be able to run *ex post* tests on MDG achievement using the regression discontinuity approach to which MDG evaluation lends itself?

Predictive ability

Data from the first half of the dataset, 1985-2000, has been used to predict results into the future for each of the MDG 3 indicators. The associated graphs show three trend lines: actual pre-2000, actual post-2000, and predicted post-2000. At the primary school level, the actual post-2000 trend is noticeably better than predicted, and shows some indication of a break. The bottom distribution of the post-2000 data points has shifted up, showing improvement in the least equal education systems in terms of enrollment. For secondary and tertiary schooling, there is little difference between predicted and actual. As mentioned in the previous section, this is likely a problem driven by the short time horizon provided by the MDG's aim of completion by 2015.

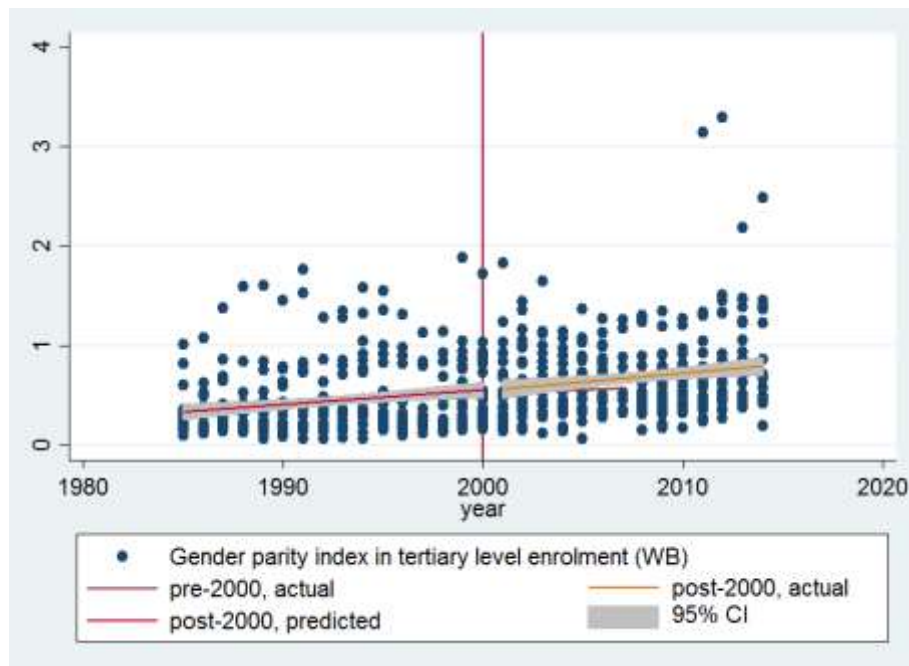Graph 1: Gender parity index in primary level enrolment
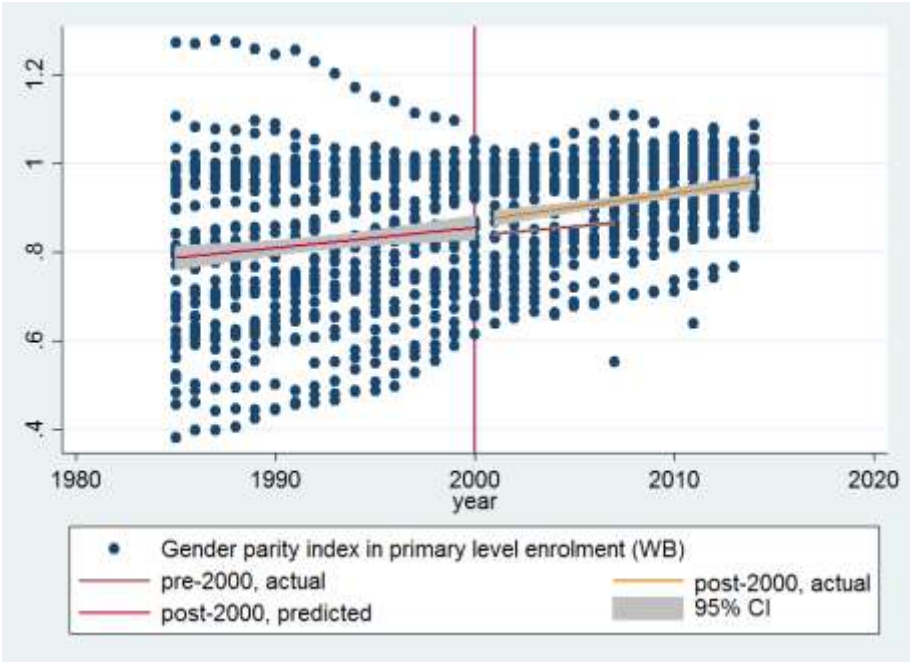
Graph 2: Gender parity in secondary level enrolment
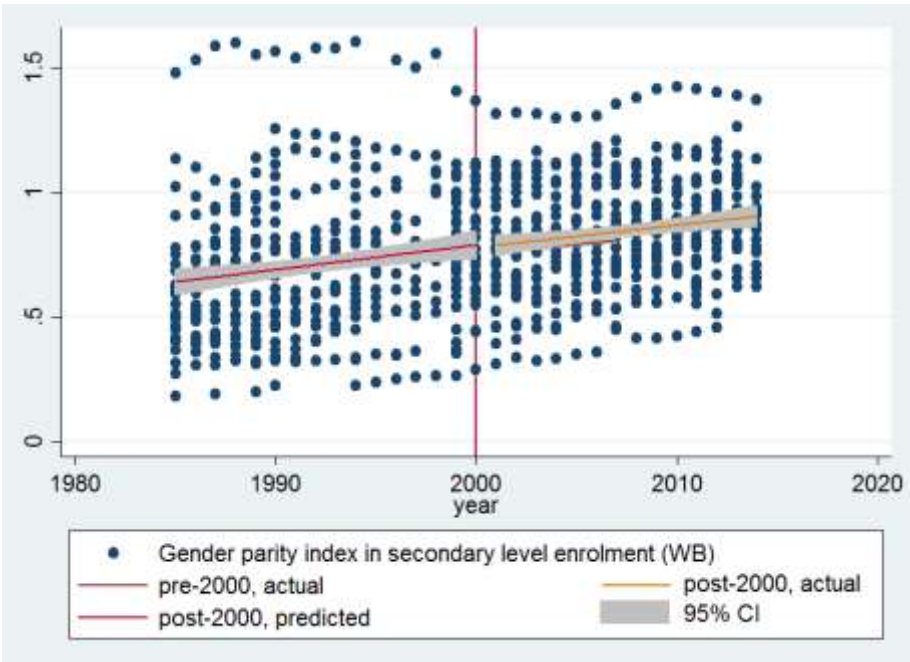
Graph 3: Gender parity index in tertiary level enrolment

The graphs for wage labour and parliamentary representation make clear the significant hurdle presented by a lack of data availability. While female representation in parliament shows significant improvement over expectations, the massive data gap in the 1990s makes it harder to make accurate predictions. The state of affairs for wage labour is much messier. Few data points and large confidence intervals do not provide a strong sense of confidence in the ability to test for real improvements in this area. Overall, the data does not provide enough information about the past to know if the MDGs have truly made a change. Noise in the dataset is a serious concern, in terms of both the lack of data and questions about measurement accuracy. The predictive ability of the data is hamstrung by data problems even though the MDGs were set up so as to be measurable and testable.

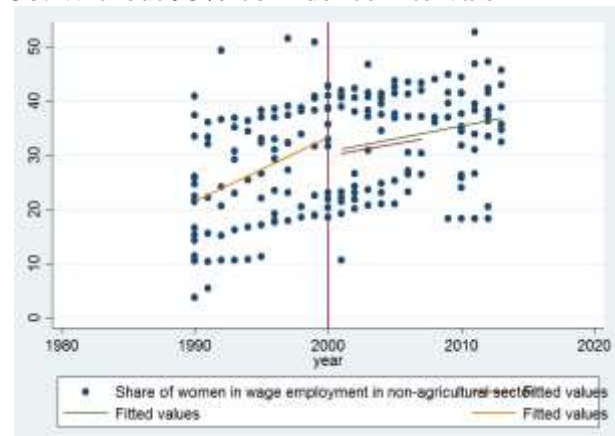Graph 4: Proportion of seats held by women in national parliament
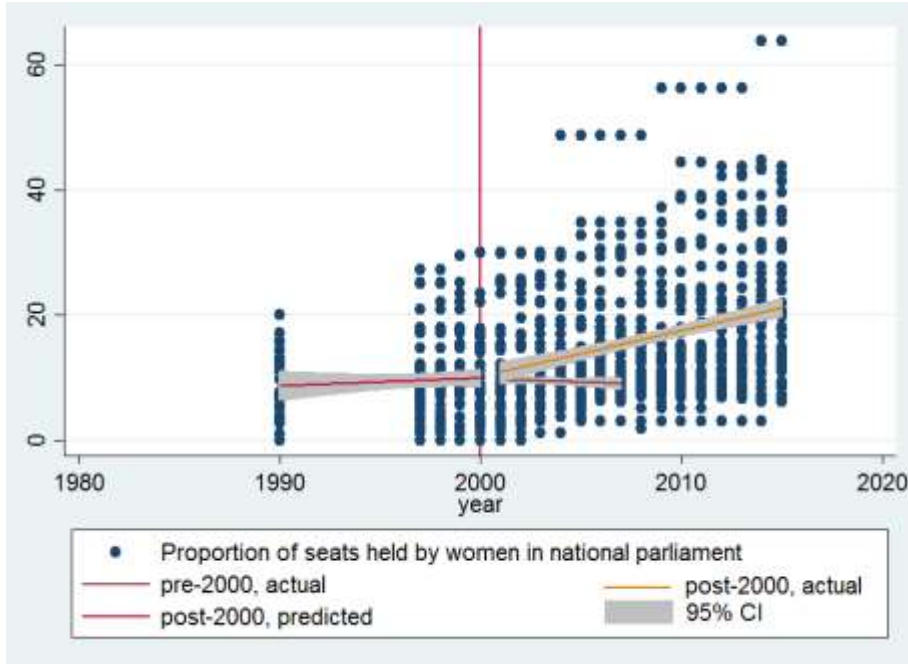
Graph 5: Share of women in wage employment in non-agricultural sector

5a. With 95% confidence intervals
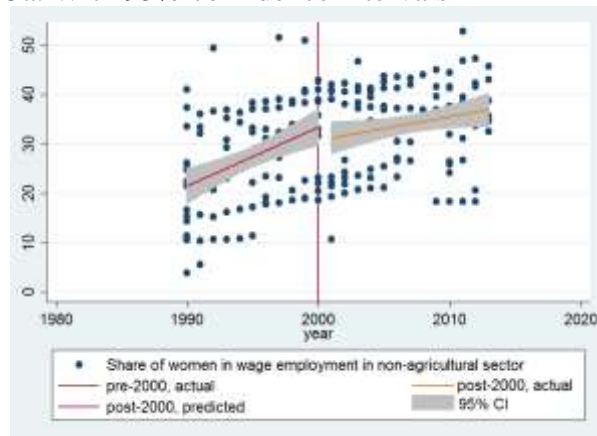
5b. Without 95% confidence intervals

Based on these results, predictive ability is unclear. There is some evidence of differences between predicted and actual outcomes, supporting the regression discontinuity results, but the data problems are so great so as to bring justifiable doubt. Low predictive ability leads to another problem in terms of incentives. Little available past data makes it hard to accurately test for

progress, providing weak incentives for quantifiable change. Low statistical capacity to measure MDG indicators compounds this incentive issue. Despite the purposefully quantitative nature of the MDGs, the reality is that we do not have the data we need to properly evaluate progress.

Discussion

Data quality and availability

Poor data quality and availability is a core concern with the MDGs. It affects evaluation and prediction, and undermines the quantitative foundation of the Goals. There are significant gaps in the datasets related to measures of women's rights and empowerment. This may be more understandable in the context of issues that are more difficult to measure, like norms and beliefs. However, there are also significant gaps for the specific MDG 3 indicators. For example, percentage of women in non-agricultural wage employment is listed as indicator 3.2. One may expect adequate documentation and data collection for an official indicator. In this dataset of 47 countries from 1985 to 2015, however, there are only 191 observations for this variable. Approximately 87% of these numbers are missing. With the MDGs making up half of the timeframe used, there is a large amount of data missing in a time when the point is to measure progress. This lack of data creates a barrier to our understanding of what women face around the world and what policies work to help.

Missing data is nothing new. Bourguignon et al. (2008) have asserted that the majority of developing countries do not produce regular or reliable data. Jerven (2013) has widely argued that the statistical capacity in Africa is low and that the data produced can be misleading. The state of statistical capacity in these countries sets them at a disadvantage for indicator-based projects like the MDGs. When setting up this style of priority setting without addressing basic problems with gathering and analyzing these numbers, global social governance actors like the

United Nations further set up developing countries for failure, or at least perceived failure. It must be pointed out that the problems of MDG measurement and evaluation have not gone unnoticed, from the forward-looking call for a data revolution to self-reflective reports like the *Lessons Learned from MDG Monitoring From A Statistical Perspective* from the United Nations Inter-Agency and Expert Group on MDG Indicators (2013). Productive awareness of the issue is critical for moving forward.

The fields of development and social policy, under the wider global governance system, have seen the rise of evidence-based policy and a reliance on indicators. This approach has been fostered in a quantitatively-minded system. Indicators and data have far-reaching influence, as the quantitatively-driven system has had a profound effect on how global social policy agendas are set. The goals of the MDGs and the SDGs, for example, are set on the foundation of predetermined indicators. While governance by indicators may be considered by some as a problematic governance approach (see, for example, Davis et al, 2012), it is clearly problematic to have data-driven projects without the data. The purpose of the quantifiable nature of the goals' indicators is the ability to measure, monitor, evaluate, and the like. Without the necessary statistical capacity, however, actors at any level in the social governance and policy sphere cannot use indicators for this purpose.

Framing versus measurement

The inherent problems with MDG data in terms of availability, quality, and predictive ability make the MDG indicators' goals of measurability and accountability unattainable for MDG 3. By using indicators to guide development policy without fixing the associated issues outlines above, we do not have the tool that we think we have.

We think we have a measurement tool. Indicators outline how to measure progress towards a given target. Using these measurements, we can assess, evaluate, rank, and the like. The MDGs were designed for this purpose. However, they do not fulfill that purpose. We think we have a measurement tool when we actually have a framing tool. Indicators frame our understanding of an issue, generating priorities and policy prescriptions in line with the given frame (Fukuda-Parr 2016). In this case, in order to achieve gender equality and to empower women, countries would prioritize policies that get women and girls in school, in non-agricultural wage employment, and sitting in parliament. Whether or not these are the most effective or efficient policy routes is removed, or sidelined, from the discussion. The indicators shape our understanding and, given their presentation as scientific and objective (see Hansen & Porter 2012), are taken as truth, the basis that guides future policy responses.

The framing versus measurement tool problem is akin to thinking you have a hammer when you actually have a wrench. You can use a wrench as if it were a hammer in order to bang a nail into a wall, but you are neither doing the job efficiently nor are you using the wrench in the way it was designed to work. With the MDGs, we are using an indicator to measure, which it has difficulty doing given current realities with statistical capacity constraints, and neglecting how it frames policy problems and solutions. When goals and indicators are not well-aligned or are overly narrow, the resultant framing can push us further away from the spirit of the goal.

Using indicators as a framing tool is not a bad thing in and of itself. It is also not an unknown or uncommon understanding of indicators, given burgeoning discussion on 'governance by indicators' and 'politics by numbers'. Much of this discussion, however, is happening outside of but adjacent to discussions of measurement issues within economic circles. Recognition of indicators as a framing, as opposed to a strict measurement, tool is what may be

missing from quantitative analysis, to which the MDGs theoretically lend themselves but practically do not. Bridges need to be built outside of dominant economic approaches in order to holistically tackle problems inherent in MDG indicators. This is in part because an indicator's value as a framing device is important but so too is its value as a measurement device. In quantitatively-based evaluation, there is a need to recognize the usefulness of an indicator as a framing device but to push statistical capacity forward so it can also be used as a measurement tool, as was its initial purpose.

Moving forward: The Sustainable Development Goals

The UN system has not overlooked the data gap in the post-2015 agenda, with the move to the SDGs. It has instead called for a data revolution, aimed at transforming data production and use. The tasks of the data revolution include improving data collection, disaggregating data by various characteristics including gender, promoting accountability, and others. Two of the SDGs' 169 targets specifically promote this data revolution. Targets 17.18 and 17.19 call for building statistical capacity in order to increase the availability of "high-quality, timely and reliable data disaggregated by income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts" (ECOSOC, 2016). The data revolution may greatly benefit SDG progress and support governance by indicators more broadly, but it is an expensive task.[4] It also cannot undo the previous data issues, making it difficult to compare SDG results with the MDG time period or before.

Efforts to improve women's rights and empowerment will benefit from greater statistical capacity, particularly when data is disaggregated not only by gender but by a variety of characteristics to address how depravity may be compounded. It will show progress, evidence,

---

[4] Jerven (2014) estimates that the cost to fully measure the MDG agenda would have been 27 billion dollars, or 1.5 billion dollars per target.

and areas of need. This, too, has been recognized within the UN system. UN Women has consistently supported a data-driven approach, in line with the data revolution. Their position paper on the gender goal of the SDGs called for "robust monitoring frameworks and timely and reliable statistics" as part of a transformative approach to development (UN Women, 2013). From the data revolution perspective, the United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG) devoted attention to gendered concerns in their report on mobilizing the data revolution for sustainable development (IEAG, 2014). It calls for more data not only on women but on gendered issues like intimate partner violence and division of household labour.

Part of the gendered data revolution involves new indicators that address discrimination. This change has been driven by taking a rights-based approach to gender equality as opposed to the strictly technical approach taken by the MDGs, which the United Nations Development Group (2010) has since recognized as too narrow. This is an important step forward for addressing foundational problems preventing gender equality and women's empowerment. However, it does complicate evaluation in terms of comparison. Some of the new indicators can be traced back to the MDG 3 indicators but many new indicators have been introduced that are not only difficult to measure but also have little if any historical data. Examples include measuring rates of child marriage and female genital mutilation. This is not to say that the indicators should not have changed over time but the change has made evaluation still more challenging than it already is.

The SDG 5 indicators may be used even more as framing tools instead of measurement tools, given the measurement challenges they present. It may be an effective policy strategy as it could counter a myopic focus on targets instead of the structures and contexts that the goals

address, as seen with MDG-related policy implementation (see, for example, Unterhalter 2012 on MDGs 1, 2, and 3). When the targets and indicators frame empowerment in terms of rights instead of parity targets in three specific areas, policy approaches to gender equality and women's empowerment could be more productive. This strategic value needs to be accommodated in work on evaluation, particularly as progress may be difficult to accurately discern.

Conclusions and future research

The creation of the MDGs meant taking aspirations of development and attaching measurable targets. It meant distilling ambitious norms into narrowly-focused indicators. For MDG 3, it meant a great focus on women's education, with secondary focus on non-agricultural wage labour and parliamentary representation. The regression discontinuity results, to which MDG evaluation is theoretically well-suited, do show some progress in gender equality in sub-Saharan Africa during the MDG timeframe…and based on the MDG indicators as measures of success. These results are less convincing when taking into account predictive ability, or lack thereof, and low data quality and availability.

In effect, the MDGs are made up of a variety of indicators that lack the strength data should provide. These indicators then act more as framing tools than as measurement tools, shaping our understanding of an issue area and which policies to pursue. In this case, the road to gender equality is equated to gender parity in classrooms, labour forces, and parliaments. This may not necessarily lead to actualized gender equality or women's empowerment. Structural barriers, for example, are neglected.

Goal 5 of the SDGs takes a broader, more rights-based approach to the goal of gender equality and women's empowerment. It brings along a significant increase in the number of

indicators; a trend echoed in the rest of the SDGs where the overwhelming number of indicators is cause for reasonable concern. Many of these indicators present significant measurement challenges. It is very difficult, for example, to measure the proportion of women and girls who have been subjected to physical, sexual, or psychological violence given disincentives to report violence, stigma, mishandling by police, etc. While the SDGs' data revolution pushes for improvements in data quality and availability, it takes great resources to do so and no quick improvements can be expected. Again, there is also the issue presented by a lack of historical data. During the research process for this present study, the author attempted to compare MDG 3 indicator outcomes with non-indicator outcomes on gender equality, many matching those found in the list of SDG indicators, and found even less data to work with.

Future research faces numerous difficulties but may build off this present study. Incorporating the idea of indicators as framing devices will benefit analysis and interpretation of MDG results. There is also room to negotiate with the Tier Classification of Global SDG Indicators, which assesses each SDG indicator based on data availability and methodological development. Applying this framework to MDG indicators may further aid in analysis, as well as integrating a longer view of historical data production and its relationship with prediction. There is also much to be gained from engaging with the governance by indicators discussions from outside of economics, bringing a pro-quantitative-methods stance to ongoing debates. Just as the SDGs attempt to break down silos between issues areas, so too must we break down disciplinary barriers in our approaches to evaluation.

References

Abu-Ghaida, D. & Klasen, S. (2004). The costs of missing the Millennium Development Goal on
  Gender Equality. *World Development, 32*(7), 1075-1107.

Bourguignon, F., Bénassy-Quéré, A., Dercon, S., Estache, A., Gunning, J.W., Kanbur, R.,
  Klasen, S., Maxwell, S., Platteau, J., & Spadaro, A. (2008). *Millennium Development
  Goals at midpoint: Where do we stand and where do we need to go?* (Rapport d'étude
  03). Paris: Centre d'Etudes Prospectives et d'Informations Internationales.

Cerise, S. & Francavilla, F. (2012). *Tackling the root causes of gender inequalities in the post
  2015 development agenda.* Paris: Organization of Economic Co-operation and
  Development, Development Centre.

Chisamya, G., DeJaeghere, J., Kendall, N., & Khan, M.A. (2012). Gender and education for all:
  Progress and problems in achieving gender equity. *International Journal of Educational
  Development, 32(*6), 743-755.

Cingranelli, D.L. & Richards, D.L. (2008). *The Cingranelli-Richards (CIRI) Human Rights Data
  Project Coding Manual*, version 7.30.08, available at http://ciri.binghamton.
  edu/documentation/ciri_coding_guide.pdf

Cingranelli, D.L. & Richards, D.L. (2013). *CIRI human rights data project* [Data file]. Available
  from: http://www.humanrightsdata.org/

Connell, R. (2010). Kartini's children: On the need for thinking gender and education together
  on a world scale. *Gender and Education, 22*(6), 603-615.

Davis, A. Fisher, B. Kingsbury, & S.E. Merry (Eds.) (2012), *Governance by indicators: Global
  power through classification and rankings*. Oxford: Oxford University
  Press.

Easterly, W. (2009). How the Millennium Development Goals are unfair to Africa. *World Development, 37*(1), 26-35.

ECOSOC, *Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators*, E/CN.3/2016/2/Rev.1 (19 February 2016).

Filmer, D., Hasan, A., & Prichett, L. (2006). *A millennium learning goal: Measuring real progress in education* (CGD Working paper No. 97).

Fukuda-Parr, S. (2016). From the Millennium Development Goals to the Sustainable Development Goals: Shifts in purpose, concept, and politics of global goal setting for development. *Gender & Development, 24*(1), 43-52.

Fukuda-Parr, S. & Hulme, D. (2011). International norm dynamics and the "end of poverty": Understanding the Millennium Development Goals. *Global Governance, 17*, 17-36.

Gabizon, S. (2016) Women's movements' engagement in the SDGs: lessons learned from the Women's Major Group. *Gender & Development, 24*(1), 99-110.

Glick, P. & Sahn, D.E. (2000). Schooling of girls and boys in a West African country: The effects of parental education, income, and household structure. *Economics of Education Review, 19,* 63-87.

Hansen, H.K. & Porter, T. (2012). What do numbers do in transnational governance? *International Political Sociology, 6*(4), 409-426.

Hulme, D. (2009). *The Millennium Development Goals (MDGs): A short history of the world's biggest promise* (Working Paper 100). Manchester, UK: Brooks World Poverty Institute.

Hyde, K.A.L. (1997). Sub-Saharan Africa. In E.M. King & M.A. Hill (Eds.), *Women's education in developing countries: Barriers, benefits, and policies* (pp. 100-135). Baltimore: Johns Hopkins University Press.

IEAG (2014). *A world that counts.* New York: United Nations.

Jacob, A. (2017). Mind the gap: Analyzing the impact of data gap in Millennium Development
    Goals' (MDGs) indicators on the progress towards MDGs. *World Development, 93*, 260-
    278.

Jakobi, A.P. (2009). Global education policy in the making: International organizations and
    lifelong learning. *Globalisation, Societies and Education, 7*(4), 473-487.

Jerven, M. (2013). *Poor numbers: How are are misled by African development statistics and
    who to do about it.* Ithaca, NY: Cornell University Press.

Jerven, M. (2014). *Benefits and cots of the data for the development targets for the post-2015
    development agenda* (Working paper). Tewksbury, MA: Copenhagen Consensus Center.

Kevane, M. (2004). *Women and development in Africa: How gender works.* Boulder, CO: Lynne
    Rienner Publishers.

Ki-moon, B. (2011). "Poorest of the poor lagging amid uneven progress on millennium
    development". United Nations Economic and Social Council. Geneva. 7 July 2011.

Lewin, K.M. (2009). "Access to education in sub-Saharan Africa: Patterns, problems and
    possibilities." *Comparative Education, 45*(2), 151-174.

Lingard, B. and Rawolle, S. (2011). New scalar politics: Implications for education policy.
    Comparative Education, 47(4), 489-502.

Mikell, G. (1997). Introduction. In G. Mikell (Ed.), *African feminism: The politics of survival in
    sub-Saharan Africa* (pp.1-52). Philadelphia: University of Pennsylvania Press.

Oya, C. (2011). Africa and the Millennium Development Goals (MDGs): What's right, what's
    wrong and what's missing. *Revista de Economia Mundias, 27*. Online.

Richard, F., Hercot, D., Ouédraogo, C., Dalvauz, T., Samaké, S., van Olmen, J., Conombo, G., Hammonds, R., & Vandemoortele, J. (2011). Sub-Saharan Africa and the health MDGs: The need to move beyond the "quick impact" model. *Reproductive Health Matter, 19*(38), 42-55.

Subrahmanian, R. (2005). Gender equality in education: Definitions and measurements. *International Journal of Educational Development, 25*(4), 395-407.

United Nations Development Group (2010). *Thematic paper on MDG 3: Promote gender equality and empower women.* New York: UNDG.

UNESCO (1990). *World declaration on education for all.* Jomtien, Thailand: United Nations Educational, Scientific and Cultural Organization.

United Nations (1995). *Platform for Action.* The United Nations Fourth World Conference on Women. Beijing, China: UN Women.

United Nations Inter-Agency and Expert Group on MDG Indicators (2013). *Lessons Learned from MDG Monitoring From A Statistical Perspective.* New York: United Nations.

UNSD (2008). *Official list of MDG indicators.* New York: United Nations Statistics Division.

Unterhalter, E. (2012). Poverty, education, gender and the Millennium Development Goals: Reflections on boundaries and intersectionality. *Theory and Research in Education, 10*(3), 253-274.

UN Women (2013). *A transformative stand-alone goal on achieving gender equality, women's rights and women's empowerment: Imperatives and key components.* New York: UN Women.

Van Staveren, I. (2013). An exploratory cross-country analysis of gendered institutions. *Journal of International Development, 25*(1), 108-121.

Verger, A., Novelli, M., & Altinyelken, H.K. (2012). Global education policy and international development: An introductory framework. In A. Verger, M. Novelli, & H.K. Altinyelken (Eds.), *Global education policy and international development: New agendas, issues and policies* (pp. 3-32). New York: Continuum.

Walker, J.A. (2012). Early marriage in Africa: Trends, harmful effects and interventions. *African Journal of Reproductive Health, 16*(2), 231-240.

World Bank. (2004). *Determinants of Primary Education Outcomes in Developing Countries.* Washington, D.C.: World Bank Operations Evaluation Department.