Technical Aspects of the Education Model

Annababette Wils

Supporting Paper to the UNFPA Project UNFPAUSA/14/037 Investment in Improving Health, Education and other Outcomes for Adolescents, Victoria University, Melbourne
2016

1. Introduction

The VISES-Education and Marriage Model (VEMM) was developed to provide answers on: a) how investments in adolescent education could improve education outcomes, b) at what costs, and c) what would be the value of those improvements.

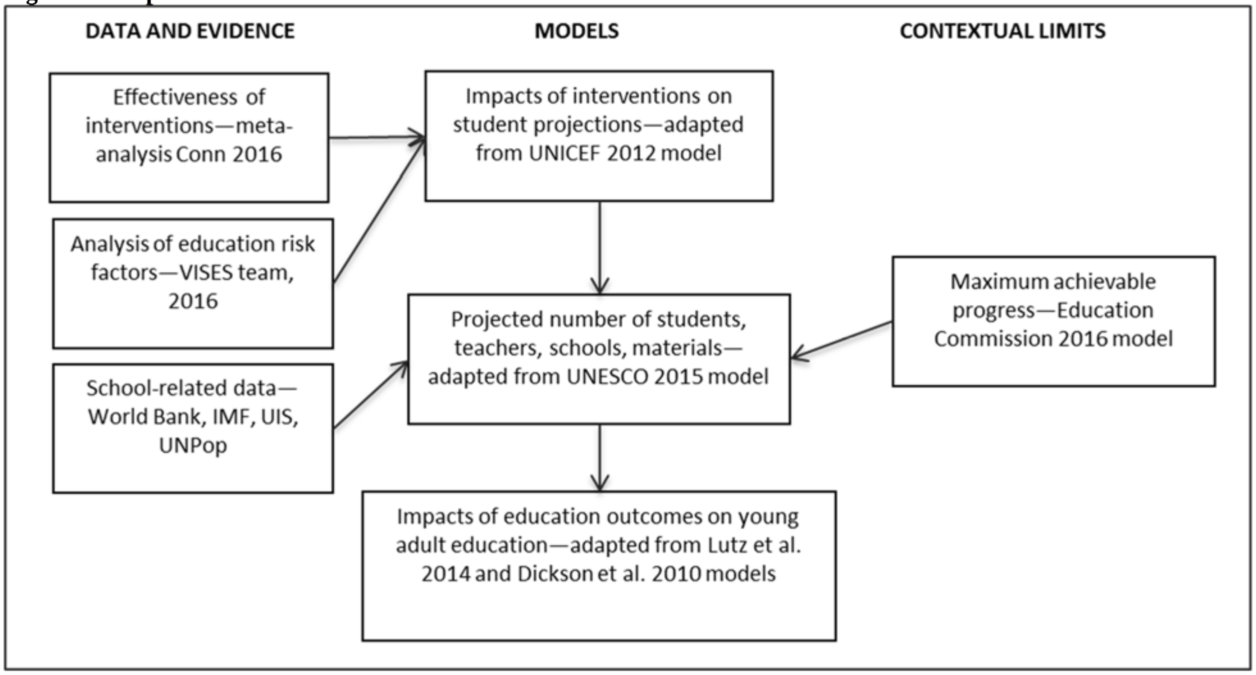
This technical paper provides the background for the new model, as well as describing—in an abbreviated form—the model and the data analysis that underlie the results described in the main report.

The question of how to improve adolescent education is not new, but, with regards to a global outlook for developing countries, models of adolescent (secondary) education until recently took a backseat to primary education. The Declaration of Education for All goals in Dakar, 2000 and the Millennium Development Goals created a need to model how much it would cost to reach the new global education goals. In response, the World Bank and UNESCO developed models to estimate the costs of attaining *primary education for all* children in developing countries by 2015, including some simple assumptions about the quality of schooling (e.g. Bruns et al. 2003; Delamonica et al. 2001; Wils et al. 2010). Later, in preparation for the new Sustainable Development Goals agreed to in Addis Ababa, 2015, UNESCO developed a costing model that includes the costs of achieving *secondary education for all* by 2030 (UNESCO 2015; Wils 2015b). These models all have a common structure: they compute the evolution of student numbers by country up to the attainment of the education goals and multiply those numbers with unit costs based on assumptions regarding teachers, salaries, materials and ongoing support, and school construction.

The limitations of this approach are that it ignores a) whether the projected rates of progress are achievable, b) how much learning improves, c) whether the assumed unit costs include the necessary interventions to improve access and school quality and d) how the outcomes affect adult skills. To answer the main question of our project, namely, which specific interventions could most (cost-) effectively improve adolescent welfare, those four gaps need to be addressed. Fortunately, other streams of research have looked at these issues. Wils and Ingram (2011) and the Education Commission (2016) analyzed feasible historical trends and applied these to education projection and costing models. The Education Commission also developed an approach to modeling learning improvements (Wils 2016). UNICEF (2012) developed a model that explicitly includes evidence from research on what types of interventions can improve education outcomes—building on the work in health that UNICEF and the WHO had used to develop the OneHealthTool and the MBB models (WHO 2013; UNICEF 2010). In parallel, researchers in Europe and the US developed demographic models that link school outcomes to improvements in adult education levels, including young adults separately (e.g. Lutz et al. 2007; Wils 2011; Lutz, Butz and KC 2014; Dickson et al. 2010; Dickson et al. 2016; Barakat 2016). The VEMM model hooks these three approaches up to the UNESCO model in a modular fashion, creating a new model out of existing, vetted parts, shown in Figure A5.1 below. Section 5.2 describes the model components and connections in more detail.

As this project is an investment case, not a theoretical exercise, the projected costs and benefits also need to be grounded in data and evidence. The VEMM model is connected to a database of more than 250,000 points, providing the education-related starting conditions for each included country. These data are a compilation from the World Bank, the UNESCO Institute for Statistics (UIS), the IMF, the United Nations Population Division (UNPop), the Education Commission, and national documents to fill in gaps left by the international databases. Regarding the effectiveness of interventions to improve education outcomes, the model includes coefficients from an analysis by Conn (2016), which brings together five meta-analyses (Conn 2014; McEwan 2014; Snilstveit et al. 2015; Nores and Barnett 2010; Rolstad et al. 2005) and produced perhaps the most comprehensive existing review of what works in education. The VEMM model also includes risk factors for adolescent education, namely poverty, rural location and female sex (an approach taken from the UNICEF 2012 model). The VISES team conducted the risk factor analysis. The data and evidence is discussed further in Section 5.3.

Figure 1 Components of the VEEM model



In summary, the VEMM model has the following features:

1. It includes 72 countries, and standard projections of basic provision of education—teachers, classrooms, and materials.
2. Projected improvements in education outcomes—access as well as learning—simulated by the assumed addition of evidence based baskets of interventions.
3. Demographic projection of students and post-school population by age, grade, and marital status for 10-19 year old adolescents and 20 to 24 year old young adults.
4. A marriage component to model the impact of increased schooling on marriage rates and of interventions to reduce child marriage on educational involvement.

2. Model Description

The **core of the VEMM model is the projected number of students** by grade and year. Although the focus is on adolescents, the model includes the entire school system starting in grade 1, to get a realistic flow-through from the end of primary into secondary school. This core is taken from the UNESCO (2015) and Wils (2015b) model. The number of students is based on the starting distribution and projected intake, promotion, repetition and dropout rates. In addition to the consideration of sex and grades, the VEMM model added age-by grade, creating a four-dimensional student matrix consisting of sex, grade, age, and time.

In mathematical terms, the projected number of students (*S*) by age and grade is determined by: the repeating students I from the same grade in the previous year and previous age; plus the promoted students (*pr*) from the prior grade in the previous year and the previous age:

$$S_{a,g,t} = S_{a-1,g,t-1} r_{g,t-1} + S_{a-1,g-1,t-1} pr_{a-1,g-1,t-1} \quad (1)$$

For the grade 1 students, intake rates replace promotion in the equation. Intake into grade 1 is projected as the gross intake rate (an exogenously assumed variable) times the population in the official entry age, distributed over age with a consideration of initial over-age entry and an exogenous assumption about the reduction of over-age entry.

Learning levels are included as a single variable, *L*, to represent the percentage of students in secondary school who attain the minimum learning benchmarks, proxied by reaching the low math score in the PISA and TIMMS (described more in Wils 2016).

The **VEMM model considers risk factors** as co-drivers of promotion, dropout, repetition and learning rates, using an adapted version of the education projection model with risk factors developed by UNICEF (2012). Risk factors such as poverty, rural location, female sex, ethnic minority status, disability all increase the probability that an adolescent will leave school prematurely, repeat a grade, or fail to reach learning benchmarks. To capture this effect, the model decomposes dropout (*d*) and repetition rateI (*r*) into a linear equation given by a conInt (*c*), plus the summed effect of risk factors, each of which is the product of the prevalence of the risk factor p_k, and the marginal impact of the risk factor β_k.

$$d/r_{a,g} = c_{a,g} + \sum_k \beta_k p_k \text{ (2),}$$

The risk factors included in the VEMM model are: poverty, female sex, and rural location. The selection, coefficients and the prevalence of the risk factors are described in the next section. The original equation (1) for student progress uses promotion rates, but these can be computed as the residual of dropout and repetition - $pr_{a,g} = 1 - d_{a,g} - r_g$. The equation for learning is similarly decomposed into a constant and risk factor effects.

With this decomposition, a projection of students will change, depending on projected changes in the prevalence of poverty, rural location, and female sex, and on the strength of the coefficients. For future prevalence of poverty, the VEMM model uses the IMF Economic Outlook (2016) and the Education Commission's (2016) projection of GDP per capita growth to 2030; future proportions of rural populations are based on the United Nations population projections (UN 2015); and the prevalence of female sex is assumed unchanged.

Changes in the coefficients for dropout, repetition, and learning assumed to be affected by **levels of investments in interventions**. For this piece, the VEMM model again uses the UNICEF (2012) education projection model. Two examples of intervention effects are: cash transfers reduce dropout by reducing the marginal impact of poverty; community programs for girls reduce the marginal impact of girl sex on dropout.

The intent of interventions is to reduce the *marginal effects* of risk factors on dropout. Other interventions, not considered in this model, could reduce the prevalence of the risk factors. The impact of interventions is the combined effect of: *coverage*—the percent of the target population receiving the intervention—and the interventions' *effectiveness*. The scale of the interventions is a user-set variable in the model, but the effectiveness is taken from the meta-meta analysis by Conn (2016) and, the VISES team's additional analysis of specific interventions.

As a simple example of how the intervention module works, say that cash transfers have the effect of reducing dropout rates of poor adolescents by 30 percent annually. If one hundred percent of adolescents is poor, and one hundred percent receive cash transfers the reduction of dropout rates will be 30 percent. But in most cases, this maximum is not reached because only a portion of adolescents is poor, and because the programs reach only a fraction of those poor adolescents. With this in mind, a general formulation for the impact of a given intervention is:

$$i_{i,k} = c_{i,k} p_k e_i \text{ (3),}$$

where *i* is the impact of the intervention, *k* is the risk group k, c is the coverage of the intended risk group (percent receiving the intervention), *p* is the prevalence of the risk group in the population, and *e*t is the effectiveness of the intervention for those who receive it.

One group of interventions has a global benefit; these can be scaled up to reach the entire population and reduce the constant in equation (2); other interventions alleviate barriers for particular risk groups and impact the β's in equation (2).

Multiple interventions are often combined in a package. If they are not complimentary, the effect is multiplicative, with each intervention's effect eroded slightly by the impact of other interventions. This effect is formulated as a chain:

$$\hat{\imath} = (1 - i_1) * (1 - i_2) * \ldots * (1 - i_n) \ (4)$$

where î is the cumulative effect of all the interventions together. The value of î changes over time, depending on the assumed path of scaling up the interventions (effectiveness is assumed to remain constant).

The education model maintains a separate $\hat{\imath}_k$ for each of the three risk groups and for the constant in the dropout equation (above). The combined effect of all four $\hat{\imath}_k$ values determines the future paths of dropout, repetition and learning.

The combined effect of all interventions is to modulate the constant and the betas in equation 2 as follows:

$$d/r_t = d_0 * (\hat{\imath}_{c,d,t} c + \hat{\imath}_{f,d,t} \beta_f p_{f,t} + \hat{\imath}_{p,d,t} \beta_p p_{p,t} + \hat{\imath}_{u,d,t} \beta_u p_{u,t}) \ (5)$$

where d_t are the projected dropout rates in time t; d_0 is the initial dropout rate; and the subscripts d, f, p, and u denote dropout, female sex, poverty and rural (non-urban) location respectively. Note that both the î effects and the prevalence rates are time-dependent, as they change in the course of the simulation, and the age and grade subscripts have been dropped to simplify the equation[1].

Finally, all of the above applies only to repetition and learning of the secondary school adolescent students; for younger students, age 5-9, there are simply target-led, user-set assumptions about changes (generally a target for lower values to be reached by a specified target year).

The projected costs are the sum of the core costs and the intervention costs. The core cost equation is from the UNESCO (2015) model. Total costs are a function of unit costs, which are determined by assumptions about class size, teacher salaries, materials, and school maintenance and construction—in basic terms this equation says that unIcosts (*c*) are equal to salaries (*s*) plus a multiplier (*m*) to consider other recurrent costs plus amortized costs per classroom (*sc*), divided by the pupil teacher ratio (*PTR*, assumed equal to the pupil classroom ratio):

$$c = \frac{s * (1 + m) + sc}{PTR} \ (6)$$

The total intervention cost (*CI*) is equal to sum of costs for all interventions assumed; and the cost of each individual intervention is equal to the number of students (*n*) receiving intervention *i* at time *t* multiplied by the unit costs of the intervention (*ci*):

$$CI = \sum_i n_{i,t} * ci_{i,t} \ (7)$$

For the **computation of the investment case,** two scenarios are considered. The first is a base case, where there are no interventions assumed, so education progress remains flat, and the only costs are the basic costs. In the intervention case, education improves along the path resulting from changes to dropout, repetition, and learning failure as a result of interventions, and the costs are both the intervention costs and the increase in base costs due to the higher number of students.

[1] Also, for repetition *and learning* the VEMM model does not consider the marginal effects of risk factors, but the scale of interventions is still limited by the prevalence of the risk factors. The equation for repetition/learning, with subscripts *r/l* for repetition or learning, is:

$$r/l_t = r/l_0 * \hat{\imath}_c \left(1 - p_{f,t}(1 - \hat{\imath}_{f,r/l,t})\right) * \left(1 - p_{p,t}(1 - \hat{\imath}_{p,r/l,t})\right) * \left(1 - p_{u,t}(1 - \hat{\imath}_{u,r/l,t})\right)$$

In addition, for the investment case and the calculations of education's impacts on worker skills and earnings, as well as on early marriage rates, the VEMM model needed to consider how the different schooling outcomes affect the **education levels of adults, in particular, young adults**. For these calculations, the VEMM model computes school leavers (L) and from those, the population by age, sex and highest completed education level. For the school leavers, the VEMM model uses equations from Wils (2011) and for the demographic component, the VEMM model uses the multi-state demographic methodology first proposed by Rogers (1980) and more recently described for education projection in Lutz et al. (2014) and Dickson et al. (2010).

From the projected students, the computation of school-leavers is straightforward as the product of students and the dropout rate:

$$L_{a,g,t} = S_{a,g,t} d_{a,g,t} \quad (8)$$

For the **adult population by education**, the VEMM model uses multi-state demographic methodology. Starting from a base situation with the population distributed by sex, age, and educational attainment, and people transition between educational attainment states as they enter and leave school. The education states included in the VEMM model are: no schooling, in school, left school with primary only, left school with incomplete secondary, and left school with completed secondary.

All people are born into the no schooling category. In the countries in the VEMM model, most transition into school, but some portion never enters school and these remain in the no schooling state even as they age into adults. The never-schooled population is:

$$P_{a,0,t} = P_{a-1,0,t-1}(1 - m_a) - I_{a,t}. \quad (9)$$

Intake is considered possible up to age 12 in the model.

Those who enter school remain in the in school category, progressing through grades according to equation (1), until they leave school at the rate computed from equation (8). The school-leavers enter into one of the three post-school states, depending on the grade they left school, and over time, the left-school population is attenuated by mortality (m):

$$P_{a,g,t} = P_{a-1,g,t-1}(1 - m_a) + L_{a,g,t}(1 - m_a). \quad (10)$$

Finally **marriage is added** through the division of each population-age-sex group into married and un-married according to proportions observed in the household surveys in the benchmark year and assumed fixed over time (thus, educational change is the only avenue modeled for reducing marriage rates)

The sum of the education-specific population groups is equal to the total population.

3. Data and Evidence

The education risks for adolescents are high, but particularly higher for married, female, poor, and rural adolescents. The study undertook a **multivariate risk factor analysis** to determine the independent contribution of each of these four risk factors to dropout rate based on household surveys. In total, 49 Demographic Health Surveys (DHS) from 2008–2014 were included in the analysis and 24 UNICEF Multiple Indicator Cluster Surveys (MICS). Five of these could not be used due to missing variables. The micro-datasets were combine into two meta-files, one containing DHS and the second containing all the MICS surveys, with the variables related to the regression—age, sex, location, marital status, relative wealth indices, educational status and derived variables for dropout, school duration, poor status, rural status, and marital-education status.

This dataset was used in a multivariate regression model following Hattori (2014), later adapted by Kan and Wils (2016) with the probability of having left school by age 19 as the dependent variable, and marital status, sex, poor and rural status as the independents. As these are all binary a logit model was used, and the results transformed into marginal effects. The model was run separately for each country. The results are the coefficients for risk-factor equation (3) above.

The study also researched the **evidence of the effectiveness of interventions**, mainly using Conn (2016), but also other earlier analyses. In health, rigorous and repeated testing of new treatments is common to identify effective interventions. In education, this rigorous approach is only recently expanding. In part, the field of education has been held back in this respect by the fact that the delivery modes of any given education intervention can be very varied (meaning that one study is never exactly comparable to another). For example, in the area of distance learning, educationalists have studies on the impacts of programs as diverse as Sesame Street for preschoolers and Telesecundaria, a program for rural secondary schools in Latin America.

To consolidate this evidence, diverse as it is, educationalists turn to meta-analyses, and account for the diversity of studies in groups of interventions in different ways. Some, like recently Glewwe and Mulharidharan (2015) present only the numbers of studies with positive, negative, or no significant results. While this is a useful first step, it is not helpful for the identification of particularly strong or cost-effective interventions. McEwan (2015), Conn (2014) and Snilstveit et al. (2015) utilize the statistical techniques of meta-analysis (e.g. Borenstein et al. 2009) to provide weighted averages of the effects of intervention-groups (the weights are determined by among other things, sample sizes, standard errors, type of study). UNICEF (2012) presented the effectiveness of interventions based on its own compilation of 300+ studies (using an un-weighted approach) to make the investment case for investing in equity in education.

Conn (2016) and the Education Commission (2016) took this intervention research one step further by consolidating five meta-analyses into one meta-meta study on education intervention impacts. These studies found the following interventions to be the most effective at increasing learning and reducing dropout rates: early childhood interventions; mother-tongue or bilingual instruction; multi-level interventions that include community, teachers, and student support; cash transfers; improved teaching methods; improved school infrastructure; schools nearby; remedial education; computer-assisted instruction with teacher support; and some health interventions such as malaria prevention.

Finally the VEMM model is grounded in **detailed country-specific data** that describe the starting position of each country with regards to students by grade, teachers, school buildings, student progression rates, learning outcomes, as well as contextual variables regarding income, poverty, and distribution of the population. The sources of these data are: the UIS (UNESCO Institute of Statistics), Word Bank Development Indicators, the UN Population Division and additional data compiled for the adolescent study. The benchmark data for the model encompasses more than 250,000 data points. The initial conditions provide a comprehensive starting point for each country including the existing student population and its education outputs (progress through school and dropout rates), risk factors for dropout in the adolescent years, and the initial learning levels.

Basic data on the pupil distribution by sex and grade, number of teachers, teacher salary, annual expenditures on materials, and estimates on classrooms, and projections of the school age population are taken from the UNESCO (2015) and the Education Commission (2016) model. Most of these data are from the UNESCO Institute of Statistics (UIS).

Repetition and dropout rates by grade were obtained from UIS data in the UNESCO (2015) model. Because we wanted to include the known effects of age on dropout rates, the grade-specific dropout divided into three groups (5–9, 10–14, and 15+). The patterns for age-specific dropout rates were taken from the same set of household surveys also use for the risk factor analysis. The age adjustments maintain the grade-specific dropout values in the aggregate. DHS and MICS household surveys were also accessed for initial age-distribution within grades, and for the age- and sex-wise initial distribution of the population over education and marital status groups—those with no schooling, those in school, those post-school with primary, post school with incomplete secondary, and post-school with secondary completed, all for never-married or ever-married.

As an estimate of the **learning levels of secondary schooling**, the model uses the proportion of secondary students who attain a minimum level of academic learning, namely the “low” benchmark for mathematics, from PISA and TIMSS. The methodology for these estimates is described in Wils (2016).

Table 1 Output of the regression for marginal effects of risk factors on early dropout probability before age 19, for 44 DHS surveys

VARIABLES	Married	s.d.	Girl sex	s.d.	Poor	s.d.	Rural location	s.d.	Observations
Albania	0.325***	(0.0574)	0.0166	(0.0354)	0.145**	(0.0607)	0.349***	(0.0302)	1,457
Bangladesh	0.398***	(0.0162)	-0.0808***	(0.0183)	0.295***	(0.0230)	-0.0585***	(0.0186)	4,940
Burkina Faso	0.426***	(0.0300)	-0.0262	(0.0323)	0.0710*	(0.0404)	0.0733*	(0.0404)	1,769
Burundi	0.578***	(0.0168)	0.0446*	(0.0263)	0.109**	(0.0473)	-0.110***	(0.0407)	2,226
Cambodia	0.311***	(0.0126)	0.0455***	(0.0157)	0.200***	(0.0162)	0.108***	(0.0200)	4,752
Cameroon	0.531***	(0.0180)	0.0206	(0.0215)	0.177***	(0.0277)	0.0214	(0.0274)	3,891
Colombia	0.357***	(0.0162)	-0.130***	(0.0108)	0.263***	(0.0201)	0.0939***	(0.0148)	11,209
Comoros	0.577***	(0.0414)	-0.123***	(0.0364)	0.0544	(0.0336)	0.0289	(0.0350)	1,368
Congo	0.396***	(0.0360)	0.0288	(0.0361)	0.283***	(0.0455)	-0.139***	(0.0480)	2,258
Congo DR	0.494***	(0.0242)	0.208***	(0.0239)	0.0619*	(0.0371)	0.0994***	(0.0336)	4,736
Cote d'Ivoire	0.383***	(0.0353)	0.137***	(0.0339)	0.169***	(0.0477)	0.252***	(0.0434)	1,647
Ethiopia	0.470***	(0.0246)	-0.0518*	(0.0278)	-0.168***	(0.0543)	0.283***	(0.0501)	3,982
Gabon	0.459***	(0.0527)	0.0503	(0.0370)			0.253***	(0.0363)	1,879
Gambia	0.507***	(0.0301)	0.00183	(0.0296)	0.157***	(0.0368)	-0.0634*	(0.0352)	2,413
Ghana	0.454***	(0.0273)	0.0934***	(0.0237)	0.0145	(0.0306)	0.110***	(0.0284)	2,407
Guinea	0.470***	(0.0397)	0.128***	(0.0356)	0.111*	(0.0625)	0.0800	(0.0632)	1,526
Guyana	0.256***	(0.0510)	-0.138***	(0.0431)	0.337***	(0.0523)	0.129***	(0.0438)	1,038
Haiti	0.603***	(0.0266)	0.0620***	(0.0212)	0.137***	(0.0291)	0.0196	(0.0293)	3,664
Kenya	0.556***	(0.0337)	0.0468	(0.0345)	-0.00265	(0.0395)	-0.0220	(0.0515)	2,072
Kyrgystan	0.232***	(0.0404)	-0.0707***	(0.0230)			0.0826***	(0.0195)	1,551
Lesotho	0.466***	(0.0193)	-0.172***	(0.0246)	0.252***	(0.0279)	0.127***	(0.0357)	2,971
Malawi	0.521***	(0.0158)	0.0245	(0.0219)	0.277***	(0.0342)	0.0598*	(0.0361)	5,515
Mozambique	0.385***	(0.0220)	0.0202	(0.0257)	0.258***	(0.0305)	0.100***	(0.0285)	2,977
Nepal	0.450***	(0.0244)	-0.0445*	(0.0263)	0.219***	(0.0263)	-0.00935	(0.0306)	2,673
Niger	0.384***	(0.0306)	-0.0919**	(0.0368)	0.143***	(0.0404)	0.183***	(0.0403)	1,188
Nigeria	0.434***	(0.0221)	0.0298**	(0.0123)	0.148***	(0.0116)	0.0702***	(0.0117)	7,180
Pakistan	0.279***	(0.0279)	0.0113	(0.0214)	0.169***	(0.0242)	0.0737***	(0.0235)	5,055
Peru	0.372***	(0.0257)	-0.0437***	(0.0139)	0.203***	(0.0279)	0.160***	(0.0179)	5,566
Rwanda	0.363***	(0.0197)	0.0411**	(0.0182)	0.0671**	(0.0309)	0.00779	(0.0292)	3,088
Senegal	0.533***	(0.0359)	-0.00174	(0.0457)	0.0485	(0.0416)	-0.000372	(0.0425)	1,484
Sierra Leone	0.530***	(0.0254)	0.161***	(0.0239)	0.0866***	(0.0332)	0.166***	(0.0271)	3,363
STP	0.392***	(0.0430)	-0.0736	(0.0549)	0.337***	(0.0467)	0.0945*	(0.0508)	731
Tajikistan	0.194***	(0.0298)	0.177***	(0.0238)	0.0905***	(0.0265)	0.103***	(0.0233)	2,230
Tanzania	0.388***	(0.0165)	0.0466**	(0.0228)	0.147***	(0.0325)	-0.00897	(0.0311)	2,683
Timor Leste	0.627***	(0.0247)	-0.0450**	(0.0190)	0.131***	(0.0199)	0.142***	(0.0206)	3,316
Togo	0.534***	(0.0222)	0.143***	(0.0275)	0.0561	(0.0430)	0.0359	(0.0399)	2,132
Uganda	0.465***	(0.0187)	0.0487*	(0.0255)	0.0945***	(0.0314)	-0.0302	(0.0312)	2,353
Zambia	0.558***	(0.0159)	0.0719***	(0.0218)	0.324***	(0.0277)	0.0904***	(0.0239)	4,597
Zimbabwe	0.287***	(0.0163)	0.0611***	(0.0186)	0.0944***	(0.0197)	0.0164	(0.0201)	2,457

References

Barakat B. Projections of educational attainment and its development impacts for scenarios of full and partial progress towards universal upper secondary schooling. Report for The International Commission on Financing Global Education Opportunity. Vienna: Wittgenstein Centre for Demography and Global Human Capital, 2016.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Chichester, UK: John Wiley, 2009.

Bruns B, Mingat A, Rakotomalala R. Achieving universal primary education by 2015: a chance for every child. Washington DC: World Bank, 2003.

Conn K. Identifying effective education interventions in Sub-Saharan Africa: a meta-analysis of rigorous impact evaluations. PhD Dissertation. Columbia University, 2014.

Conn K. The effectiveness of educations programs worldwide: evidence from a meta-analytic dataset. Background papers for the Education Commission. Washington DC: Education Commission, 2016.

Delamonica, E., Mehrotra, S. and Vandemoortele, J. Education for all is affordable: a minimum global cost estimate. Demographic and Health Survey (DHS). 2001. Datasets downloaded from https://dhsprogram.com/

Dickson JR, Hughes BB, Irfan MT. Patterns of potential human progress: advancing global education, vol. 2. Boulder, CO: Oxford University Press, 2010.

Dickson JR, Irfan MT, Hughes BB. USE 2030: exploring impacts, costs, and financing. Background Paper for the International Commission on Financing Global Education Opportunity, 2016.

Education Commission. The learning generation: analytical background paper. Background paper, Washington, DC, 2016.

Education Commission. Report to The International Commission on Financing Global Education Opportunity. Washiongton DC, forthcoming.

Glewwe P, Muralidharan K. Improving school education outcomes in developing countries: evidence, knowledge gaps, and policy implications. RISE-WP-15/001. Oxford: Blavatnik School of Government, University of Oxford, 2015. http://www.bsg.ox.ac.uk/sites/www.bsg.ox.ac.uk/files/documents/RISE_WP-001_Glewwe_Muralidharan.pdf (accessed July 12, 2016).

Hattori H. Demographic and socioeconomic determinants of school attendance: an analysis of household survey data. Paper commissioned for Fixing the Broken Promise of Education for All: Findings from the Global Initiative on Out-of-School Children (UIS/UNICEF, 2015). Montreal: UNESCO Institute for Statistics (UIS), 2014.

IMF. World economic outlook database, April. Washington DC, 2016. http://www.imf.org/external/ns/cs.aspx?id=28

Lutz W, Butz WP, KC, S. World population and human capital in the twenty-first century. Oxford: Oxford University Press, 2014.

Lutz W, Cuaresma J, Sanderson W. The demography of educational attainment and economic growth. *Science* 2008; **319**: 1047–8.

Lutz W, Goujon A, Wils A. Forecasting human capital using demographic multistate methods by age, sex,and education to show the long-term effects of investments in education. WP 07-03. Education Policy and Data Center. FHI360. Washington DC, 2007.

Kan S, Wils B. Analysis of risk factors in education in developing countries: technical report. Background papers for the Education Commission. Washington DC, 2016.

McEwan PJ. Improving learning in primary schools of developing countries: a meta-analysis of randomized experiments. *Rev Educ Res* 2015; **20**: 1–42.

Nores M, Barnett WS. Benefits of early childhood interventions across the world: (under) Investing in the very young. *Econ Educ Rev* 2010; **29**: 271–82.

OECD Universal basic skills: what countries stand to gain. Paris, 2015. http://dx.doi.org/10.1787/9789264234833-en

Rogers A. Introduction to multistate mathematical demography. *Environ Plan A* 1980; **12**: 489-498.

Rolstad K, Mahoney K, Glass G. 2005. The big picture: a meta-analysis of program effectiveness research on English language learners. *Educ Policy* 2005; **19:** 572–94.

Snilstveit B, Stevenson J, Phillips D, et al. Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review. Washington DC: International Initiative for Impact Evaluation, 2015. http://www.3ieimpact.org/en/evidence/systematic-reviews/details/259/ (accessed July 12, 2016).

UN. World population prospects: 2015 revision. 2015. http://esa.un.org/unpd/wpp/Download/Standard/Population/ (accessed April 28, 2016).
UNESCO. Pricing the right to education. 2015. http://en.unesco.org/gem-report/node/819#sthash.mDBDiEVT.dpbs (accessed June 30, 2016).
UNESCO. Reaching the marginalized: EFA global monitoring report 2010. Paris, 2010.
UNICEF (United Nations Children's Fund), 'Narrowing the Gaps to Meet the Goals', UNICEF, New York, 7 September 2010.
UNICEF. Simulations for equity in education (SEE). 2012. http://www.unicef.org/education/bege_SEE.html (accessed May 19, 2016).
UNICEF. The investment case for education and equity. 2015a. http://www.unicef.org/publications/index_78727.html# (accessed May 19, 2016).
UNICEF. Simulations for equity in education (SEE). 2015b. http://www.unicef.org/education/bege_SEE.html (accessed May 19, 2016).
UNICEF. UNICEF Data: Multiple Indicator Cluster Surveys. Multiple Survey datasets downloaded from http://mics.unicef.org/
United Nations Population Division. World population prospects, the 2015 revision. New York, 2015. https://esa.un.org/unpd/Wpp/ (accessed November 2015).
WHO (World Health Organization). OneHealth Tool: Supporting integrated strategic health planning, costing and health impact analysis. Geneva. 2013 (http://www.who.int/choice/onehealthtool/OneHealth_Tool_Supporting_integrated_strategic_health_planning.pdf (accessed December 2015)
Wils A. Workforce model: description of methodology to track output of schooling in changes of adult and workforce educational attainment. Mimeo. 2011. https://www.researchgate.net/publication/304676285_Workforce_model_-_description_of_methodology_to_track_output_of_schooling_in_changes_of_adult_and_workforce_educational_attainment (accessed July 12, 2016).
Wils A. Description of the education model for the UNFPA-VISES adolescent welfare project: interim report. Mimeo. Melbourne: VISES, 2015a.
Wils A. Reaching education targets in low and lower-middle income countries: costs and finance gaps to 2030 for pre-primary, primary, lower- and upper-secondary schooling. Background Paper for EFA Global Monitoring Report. Paris: UNESCO, 2015b.
Wils A. The learning generation analytical background paper: trends, projections, education impacts. Washington DC: The Education Commission. New York, forthcoming 2016.
Wils A, Chaluda A, Sylla B, et al. Estimating the costs of achieving education for all in low-income countries. Background paper for the Education for All Global Monitoring Report 2010. Washington DC: Education Policy and Data Center, 2010. http://unesdoc.unesco.org/images/0018/001885/188561e.pdf (accessed July 12, 2016).
Wils A, Ingram G. Universal basic education: a progress-based path to 2025. FHI360. Washington DC: Education Policy and Data Center, 2011.