WORKING PAPER · PREPRINT

Education is Health: An Integrated Cross-Sectoral Cost-Effectiveness Framework for Education Interventions Targeting Vulnerable Adolescent Girls in LMICs

Thomas R. Vargas[a,*] Ruby Emerson[a] Mahlet A. Woldetsadik[b]
Seth Garz[b] Kym Cole[b] John Firth[a]

[a] Rethink Priorities [b] Gates Foundation
[*] Corresponding author: tvargas@rethinkpriorities.org

This version: June 2026

Abstract

Applying a novel cross-sectoral cost-effectiveness analysis of education-related interventions on adolescent girls and young women (AGYW) in low- and middle-income countries (LMICs), we find traditional methods focused narrowly on economic benefits dramatically underestimate social returns, with health and social outcomes accounting for approximately 59% of total impacts. This siloed perspective mirrors the fragmented approach to budgeting and investment common across line ministries and donor strategies organized by sector, and may be particularly ill-suited to the lived experience of AGYW, for whom education and life choices — including marriage and sexual partnership — carry simultaneously economic, health, and social consequences. This study develops and applies an integrated cross-sectoral cost-effectiveness framework that captures the full range of returns to education-linked interventions for AGYW in LMICs, channeling the perspective of an archetypical ministry of finance and planning. Drawing on a structured evidence review, we model health, social, and economic outcomes using a welfare-based valuation approach that applies explicit moral weights to enable comparison across outcome domains. Using a cash benchmarking approach, we estimate the median cost-effectiveness of interventions analyzed was approximately 50% above a benchmark for high-performing unconditional cash transfers. These findings demonstrate the importance of integrating outcome measures and investment strategies across sectors for effective policy prioritization. The framework presented can be adapted beyond the specific use case shown, using local risk profiles, costs, and valuation assumptions, to support cross-sectoral planning across health, education, and finance ministries.

Keywords: low- and middle-income countries; adolescent girls and young women; sexual and reproductive health; cost-effectiveness; cross-sectoral governance; health systems

Funding. This work was supported by the Gates Foundation (INV-086182). As co-authors, MAW, SG, and KC contributed to study design, analysis, and manuscript preparation. The findings and conclusions are those of the authors and do not necessarily reflect the positions or policies of the Gates Foundation.
Conflicts of interest. MAW, SG, and KC are employees of the Gates Foundation, which funded this work.
Data availability. The data underlying this article are derived from published studies cited in the reference list. The cost-effectiveness model and simulation code are available from the corresponding author upon reasonable request.
Ethics. Ethical approval was not required, as the study exclusively uses secondary data.
Reflexivity. This study was conducted by a gender-balanced team from the United States and Ethiopia with diverse experience in global health research and at diverse career stages.
Acknowledgments. The analyses stem from an earlier report involving substantial contributions by Dylan Collins, Greer Gosnell, Jenny Kudymowa, and Aisling Leow.
ORCID. T. R. Vargas 0000-0002-9583-5160 · M. A. Woldetsadik 0000-0001-7780-2915.

Applying a novel cross-sectoral cost-effectiveness analysis of education-related interventions on adolescent girls and young women (AGYW) in low- and middle-income countries (LMICs), we find traditional methods focused narrowly on economic benefits dramatically underestimate social returns, with health and social outcomes accounting for approximately 59% of total impacts. This siloed perspective mirrors the fragmented approach to budgeting and investment common across line ministries and donor strategies organized by sector, and may be particularly ill-suited to the lived experience of AGYW, for whom education and life choices — including marriage and sexual partnership — carry simultaneously economic, health, and social consequences. This study develops and applies an integrated cross-sectoral cost-effectiveness framework that captures the full range of returns to education-linked interventions for AGYW in LMICs, channeling the perspective of an archetypical ministry of finance and planning. Drawing on a structured evidence review, we model health, social, and economic outcomes using a welfare-based valuation approach that applies explicit moral weights to enable comparison across outcome domains. Using a cash benchmarking approach, we estimate the median cost-effectiveness of interventions analyzed was approximately 50% above a benchmark for high-performing unconditional cash transfers. These findings demonstrate the importance of integrating outcome measures and investment strategies across sectors for effective policy prioritization. The framework presented can be adapted beyond the specific use case shown, using local risk profiles, costs, and valuation assumptions, to support cross-sectoral planning across health, education, and finance ministries.

- Integrated cost-effectiveness frameworks that incorporate health, social, and economic outcomes are essential for accurate valuation of AGYW interventions given that sector-specific analyses substantially underestimate total program value.
- Applying such an integrated framework to 37 intervention arms across 24 studies, we find that health and social outcomes account for approximately 59% of total program value, and that median cost-effectiveness modestly exceeds the benchmark set by high-performing unconditional cash transfers.
- The framework introduced here can be adapted to local risk profiles, costs, and stakeholder valuations to support joint resource allocation across Ministries of Health, Education, and Gender in LMICs.

1 Introduction

Adolescent girls and young women (AGYW) in low- and middle-income countries (LMICs) bear a disproportionate burden of adverse health and social outcomes, particularly in sub-Saharan Africa. Women and girls account for approximately 73% of new HIV infections in the region, reflecting intersecting vulnerabilities related to gender-based violence, inequitable norms, and limited access

to prevention and services (WHO 2021; UNAIDS 2025). AGYW in LMICs also face elevated risk of early pregnancy and child marriage: globally, one in five girls is married before age 18, with the majority of child brides residing in South Asia and sub-Saharan Africa (UNICEF 2023). Child marriage substantially increases risks of violence, early pregnancy, and poor maternal and neonatal outcomes (UNICEF 2025). In 2019, an estimated 21 million pregnancies occurred among girls aged 15–19 in LMICs, approximately half unintended (WHO 2024).

Educational disadvantage both reflects and reinforces these health risks. Despite narrowing gender gaps in enrollment, an estimated 122 million girls in LMICs remain out of school (UNESCO 2023), with gender disparities widening during adolescence (UNESCO 2023). Girls in fragile and conflict-affected settings, girls with disabilities, and girls from marginalized communities are particularly educationally disadvantaged (Bonfert and Wadhwa 2024; World Bank 2024). Sexual and reproductive health factors, including early marriage, pregnancy, and sexual harassment, are major drivers of dropout (Wodon *et al.* 2017; Bonfert and Wadhwa 2024).

The COVID-19 pandemic exacerbated these challenges. Prolonged school closures increased learning poverty in LMICs by an estimated 13 percentage points (World Bank *et al.* 2022), with losses disproportionately borne by girls, placing an estimated 11 million girls at risk of never returning to school (Hevia *et al.* 2021; Idara-e-Taleem-o-Aagahi 2021; UNESCO 2021; Gajderowicz *et al.* 2025). School closures were also associated with declines in adolescent physical and mental health (Viner *et al.* 2022; Ramaiya *et al.* 2023). Further, disruptions to reproductive and maternal health services heightened risks of gender-based violence and child marriage (Ramaiya *et al.* 2023). Simultaneously, the pandemic intensified fiscal pressures (Mawejje 2025) constraining sustained investments in adolescent health and education (Kurowski *et al.* 2021).

These pressures are magnified by rapid demographic expansion in sub-Saharan Africa, where youth comprise a growing share of the population and unmet needs among AGYW increasingly shape population-level outcomes (UNFPA ESARO n.d.). Despite decades of investment, progress in reducing HIV incidence among AGYW has stalled, suggesting existing approaches may be insufficient to address the scale of the challenge (UNAIDS 2023).

Education and health are bidirectionally linked, with schooling influencing health outcomes and health shaping educational participation and attainment (Verguet *et al.* 2015; Psaki *et al.* 2019). However, policy responses remain fragmented across sectoral ministries and donor portfolios, with AGYW investments often poorly coordinated (World Bank *et al.* 2022). In most LMICs, Ministries of Health and Education allocate budgets through separate medium-term expenditure frameworks and sector plans, with limited mechanisms for joint investment appraisal (World Bank 2013; Crawfurd *et al.* 2019; McGuire *et al.* 2019; Barroy *et al.* 2024; Filmer *et al.* 2020).

A growing literature evaluates interventions targeting AGYW, including cash transfers, safe spaces and mentorship programs, school retention support, life skills training, and expanded sexual and reproductive health services. Rigorous evaluations demonstrate meaningful gains across

domains. In Malawi, conditional cash transfers increased school enrollment, while unconditional transfers reduced adolescent pregnancy and marriage (Baird *et al.* 2011). In Sierra Leone, a safe spaces and life skills program reduced early pregnancy and supported continued school participation (Bandiera *et al.* 2018). A longitudinal evaluation of the multi-sectoral Adolescent Girls Initiative–Kenya demonstrated improvements in schooling, sexual and reproductive health knowledge, and health-seeking behaviors (Austrian *et al.* 2021).

Despite this progress, evidence generation remains largely siloed. Most evaluations measure education or health outcomes, but rarely both, even when interventions plausibly affect multiple domains. Economic evaluations frequently prioritize labor market outcomes, overlooking health, fertility, and social impacts (Bandiera *et al.* 2018; Austrian *et al.* 2021). Health-sector cost-effectiveness analyses typically emphasize disease-specific outcomes and rarely incorporate educational attainment or broader social returns (Verguet *et al.* 2015). Existing priority-setting tools used in LMICs—including WHO-CHOICE, standard health technology assessment, and Extended Cost-Effectiveness Analysis (ECEA)— have advanced equity and distributional considerations but remain anchored within single-sector outcome frameworks and do not provide a common metric for comparing educational, health, and social returns (Verguet *et al.* 2016; Turner *et al.* 2021; Hutubessy *et al.* 2002). Consequently, interventions with multi-sectoral effects are systematically disadvantaged in resource allocation (Remme *et al.* 2014).

There is currently no standardized framework for valuing the combined education, health, and social benefits of multi-sectoral AGYW interventions. This gap has practical consequences: AGYW interventions may appear less cost-effective than vertical health or education programs because substantial portions of their benefits remain unmeasured, leading governments to underinvest in interventions that generate multiple returns (Psaki *et al.* 2019; World Bank *et al.* 2022).

This analysis addresses this gap by introducing an integrated cost-effectiveness framework that adapts welfare-based evaluation methods developed by GiveWell (GW) and Coefficient Giving (CG) for use in LMIC health and education policy. A central methodological debate in cross-sectoral priority-setting concerns whether sector-specific metrics (such as DALYs or learning-adjusted years of schooling), conventional benefit-cost ratios, or welfare-based approaches offer the most defensible basis for comparing heterogeneous outcomes. We argue that welfare-based valuation—anchored in explicit moral weights and a common unit of impact— offers a useful complement to DALY-based and purely economic frameworks by making the value judgments underlying resource allocation more easily interpretable and open to scrutiny in cross-sectoral comparisons. The framework integrates health effects, explicit moral weighting, and comprehensive uncertainty analysis, expressing results in a common metric that enables comparison across educational, health, and social domains. The evidence base is concentrated in sub-Saharan Africa, with a smaller number of studies from South Asia; findings should therefore be interpreted with attention to

this geographic scope, and policy application requires adaptation to local risk profiles, costs, and valuation assumptions.

2 Methods

This study has four objectives: first, to estimate the integrated cost-effectiveness of education-related interventions targeting AGYW by incorporating their full range of health and social outcomes; second, to compare returns across intervention types, age groups, and design features; third, to assess sensitivity to key assumptions including moral weights; and fourth, to provide decision-makers in LMICs—particularly those responsible for joint planning across Ministries of Health, Education, and Gender—with a practical and adaptable framework for cross-sectoral resource allocation.

This study develops a cost-effectiveness model to estimate the total returns to education-related interventions targeting AGYW in LMICs. The primary analytic perspective is that of a funder aiming to maximize social benefit per dollar spent. Costs and outcomes are modelled under a standardized hypothetical investment of US$100 000 per intervention, enabling direct comparison across programs. For interpretability of the cost-effectiveness estimates, we adopt the "cash benchmarking" conceptual approach, comparing estimated cost-effectiveness of education-related interventions to that of unconditional cash transfers. To operationalize this conceptual approach, we set the benchmark comparator to unconditional cash transfers delivered via GiveDirectly to low-income households in sub-Saharan Africa, estimated at 135 welfare value units of impact (the "As Good As Cash" (AGAC) benchmark) (Coefficient Giving n.d.; Hickman 2025). We chose this benchmark because it is theoretically grounded, interpretable, empirically well-documented, and actively applied in real decision-making among large-scale donors, but acknowledge that it sets a high standard for cost-effectiveness and that donors and policymakers may reasonably apply different thresholds.

The modelling approach adapts the welfare-based valuation methods of GiveWell and Coefficient Giving, specifically their moral weighting system and log-utility income valuation (Coefficient Giving n.d.; Hickman 2025), while introducing an original imputation framework to estimate missing health and social outcomes from the external literature, and extending both to enable cross-sectoral cost-effectiveness comparison. Our work is based on publicly available information about their methods, and neither GiveWell nor Coefficient Giving explicitly endorse our methodology or findings. An overview of the model structure is provided in Figure 2 and each component is described below.

2.1 Study identification and selection

Eligible studies were identified through targeted keyword searches of academic and grey-literature databases, supplemented by structured consultations with implementers and researchers. Existing systematic reviews were cross-referenced to minimize omissions. The core aim was to identify

studies suitable for cost-effectiveness modelling. Although this process did not constitute a formal systematic review, it followed transparent and replicable procedures. Studies were prioritized for inclusion if they met the following criteria: 1) randomized controlled trials or quasi-experimental designs evaluating interventions targeting AGYW ages 10-24, 2) reported quantitative effects on educational outcomes and at least one additional domain (income, fertility, marriage, or health), 3) provided sufficient cost data for estimation, 4) conducted in sub-Saharan Africa and South Asia, and 5) published within the past decade. Purely descriptive, observational, and high-income-country studies were excluded.

Studies were classified as "High," "Medium-High," "Medium," "Low-Medium," or "Low" based on the rigor of their causal identification. Of 185 studies reviewed, 30 met initial eligibility criteria and 24 studies (37 intervention arms) across 14 countries provided sufficient data for full modelling (Figure 1; Table 1).

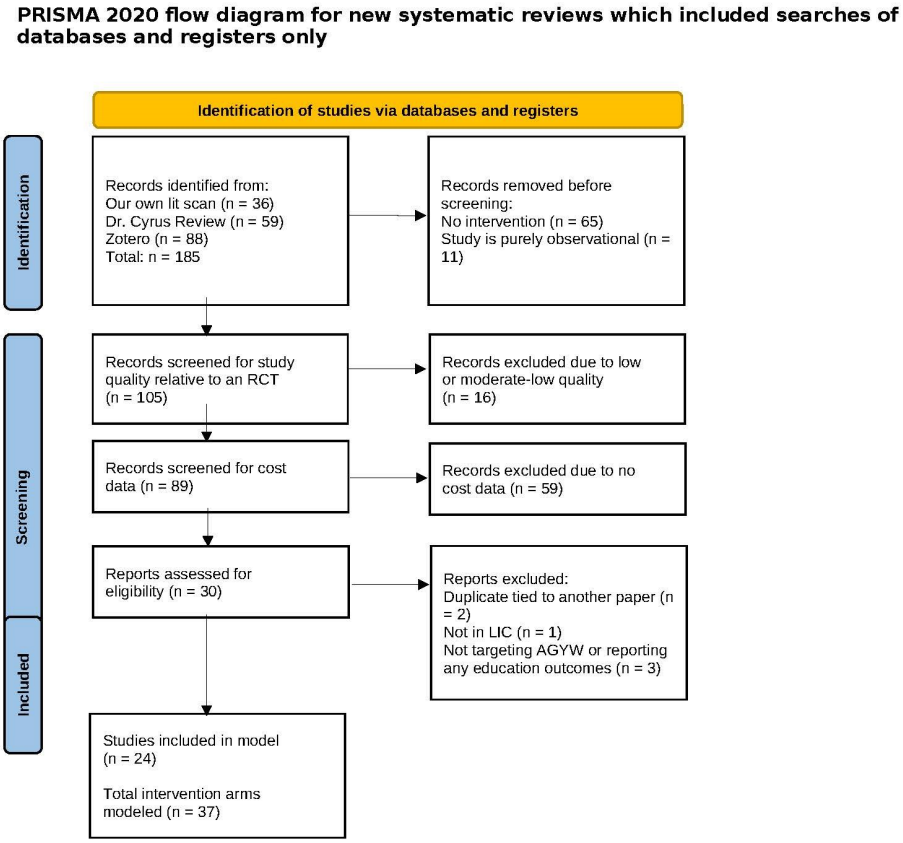


Figure 1. Study exclusion process

The modelled interventions encompassed cash transfers, scholarship programs, school meals, safe spaces, and multi-component empowerment programs, targeting youth aged 5–23, with each

study enrolling a substantial proportion of girls aged 10–19. Fifteen studies were rated high-quality; 10 were rated medium or medium-high. The evidence base was geographically concentrated in sub-Saharan Africa (n=20), with five studies in South Asia.

Table 1. Studies selected for modelling.

Authors	Program name (type)	Location	Arms
Ainul et al. (2021); Amin et al. (2018)	BALIKA (tutoring and life skills support)	Bangladesh	1
Baird et al. (2009, 2011, 2012, 2019)	Zomba (cash transfers)	Malawi	3
Bandiera et al. (2018)	BRAC ELA (package)	Uganda	1
Cohen et al. (2023)	Pathways (package)	Nigeria	1
Davis et al. (2016); Kazianga et al. (2019)	Girl-Friendly Schools (school-building)	Burkina Faso	1
Duflo et al. (2006, 2015)	Uniforms and HIV training (package)	Kenya	1
Duflo et al. (2021, 2024)	Scholarships	Ghana	1
Akresh et al. (2013)	Cash transfers	Burkina Faso	4
Amaral et al. (2024)	GBV training	Mozambique	1
Aurino et al. (2020)	School meals	Ghana	1
Austrian et al. (2022, 2024)	AGI-K (cash-plus)	Kenya	6
Blimpo et al. (2019); Gajigo (2012)	Free schooling	The Gambia	1
Erulkar and Muthengi (2007)	Berhane Hewan (package)	Ethiopia	2
Giacobino et al. (2024)	Scholarships and tutoring	Niger	1
Hahn et al. (2018)	Female Secondary School Stipends	Bangladesh	1
Hamory et al. (2016)	Multi-level interventions (package)	Kenya	1
Handa et al. (2015)	Cash Transfer	Kenya	1
Lerva et al. (2024)	Big Push (package / cash transfer)	Uganda	1
Muralidharan & Prakash (2017)	Bicycles	India	1
Palermo et al. (2021)	Cash Plus	Tanzania	1
Randall & Garcia (2020)	Scholarships and tutoring	DRC	1
Buchmann et al. (2023)	Bangladesh Oil Incentive	Bangladesh	1
Erulkar et al. (2020)	Scalable prevention (package)	Burkina Faso & Tanzania	11

continued on next page

Authors	Program name (type)	Location	Arms
Lambon-Quayefio et al. (2023)	RCT	Malawi	1

Figure 2. Cost-effectiveness modelling framework

Each intervention arm is reduced to a single, comparable metric of impact through six steps:

1. Adapt GiveWell's moral weights and spillover adjustments, and incorporate Coefficient Giving's log-utility valuation of income.
2. Apply an original approach to impute missing intervention effects from high-quality empirical studies.
3. Calculate the number of beneficiaries reachable with a standardized US$100 000 investment.
4. Estimate cost-effectiveness using two parallel models — a conservative *Reported Parameters* model and a comprehensive *Full Assumptions* model.
5. Express estimates in welfare value units of impact.
6. Benchmark results against the 135 welfare-value-unit "As Good As Cash" standard.

2.2 Analytic framework and valuation

The framework adapts GiveWell's use of moral weights and spillover adjustments, incorporates Coefficient Giving's log-utility-based valuation of income, and applies our own approach to imputing missing intervention effects from high-quality empirical studies. Cost-effectiveness is estimated using two parallel models: (1) *Reported Parameters Model*, which relies exclusively on outcomes and costs reported in the original studies, making minimal additional assumptions (likely to conservatively understate total benefits); and (2) a *Full Assumptions Model*, which imputes missing outcomes using external evidence, producing more comprehensive but less certain estimates. Results are expressed in welfare value units of impact, defined as the welfare value of transferring US$1 to an individual earning US$50 000 per year—a utility-based metric that enables comparison of education, health, and social outcomes on a common scale. The cost-effectiveness model applies a five-step process. First, per-beneficiary costs are extracted from primary studies; missing components are imputed where necessary (e.g. administrative costs), and all costs are standardized to 2024 US dollars. Second, educational impacts are standardized to additional years of schooling per beneficiary, with standard errors retained for uncertainty analysis. Third, education gains are converted into economic benefits. Short-term benefits, including direct cash or in-kind transfers, are valued using a logarithmic utility function (Coefficient Giving n.d.; Hickman 2025), implying diminishing returns to income and effectively upweighting a given dollar of income received by lower income versus higher income individuals. Long-term income gains are estimated assuming a 6.8% increase in earnings per additional year of schooling (Duflo 2001); this is consistent with, and conservative relative to, more recent global estimates of 9–10% (Psacharopoulos and Patrinos 2018;

Montenegro and Patrinos 2021). Present discounted values are calculated over a 10-year horizon, using a 4% annual discount rate (GiveWell 2023).

Fourth, health and social impacts are modelled using outcomes reported in the primary studies, including HIV, HSV-2, pregnancy, marriage, intimate partner violence, and unwanted sex. Where outcomes are not directly reported, indirect effects are imputed using pooled estimates from Psaki et al. (2019) on the causal effects of education on sexual and reproductive health (Table 3). All health and social outcomes are assigned moral weights (Table 2) and converted into welfare value units (Psaki *et al.* 2019).

Table 2. GiveWell moral weights baseline.

Moral weight	Utility
Value assigned to doubling consumption for one person for one year	1
Averting the death of an individual under five from malaria	117
Averting the death of an individual five or older from malaria	83
Averting the death of a 6- to 59-month-old child (VAS)	118
Averting the death of an individual under five from vaccine-preventable diseases	117
Averting the death of an individual five to 14 years old from vaccine-preventable diseases	134
Averting the death of an individual 15–49 years old from vaccine-preventable diseases	104
Averting the death of an individual 50–74 years old from vaccine-preventable diseases	42

Fifth, total economic, health, and social benefits are aggregated and divided by the standardized investment to generate a cost-effectiveness ratio (welfare value units per dollar). Spillover adjustments are adapted from GiveWell's assumptions: 10% of total benefits as within-household spillovers and 3% of economic benefits as community-wide spillovers (GiveWell 2024; Hickman 2025).

2.3 Moral weights

Moral weights assign relative value to non-economic outcomes and are assigned relative to the welfare value of doubling consumption for one year (Table 3). Averting an HIV infection is valued at eight times this benchmark, preventing a child's marriage at twice. Unintended pregnancy, unwanted sex, and non-sexual intimate partner violence are each valued at half; HSV-2 at one-quarter.

These weights are adapted from GiveWell's approach, which anchors valuations to disability-adjusted life years (DALYs) averted. Outcomes associated with substantial lifetime health burdens—notably HIV infection, are weighted highly, consistent with valuations for life-altering illness. Outcomes lacking well-established DALY equivalents (e.g. delayed marriage) carry provisional weights. Although moral weights make underlying value judgments explicit and contestable—a feature intended to support transparent priority-setting—they also introduce subjectivity. Alter-

native reasonable valuations could materially alter estimated program benefits. Policymakers adapting this framework to specific country contexts should consider whether the weights reflect local stakeholder valuations, and sensitivity analyses (below) assess robustness to variation in all moral weights.

Table 3. Value assigned to non-economic outcomes, relative to doubling consumption.

Parameter	Weight	Justification
One additional year of schooling	*0*	We weight additional education at zero value in itself to avoid double-counting.
Avoiding one unintended AGYW pregnancy	*0.5*	Leads to school dropout, early marriage, and poor maternal outcomes.
Avoiding one child marriage	*2.0*	Reduces schooling; increases fertility and IPV risk.
Avoiding a case of HIV	*8.0*	Approximately 6–8 DALYs per case (GBD 2019). Aligns with GiveWell valuations for life-altering illness/death.
Avoiding a case of HSV-2	*0.25*	Chronic, non-lethal condition with stigma and mild DALY burden.
Avoiding a case of unwanted sex	*0.5*	Often causes trauma and mental-health effects; DALY burden hard to quantify but plausibly high.
Avoiding a case of non-sexual intimate partner violence	*0.5*	Associated with depression, injury, reproductive harm; ≈0.3 DALYs/year on average (GBD 2019).

2.4 Imputation of missing outcomes

Under the full assumptions model, missing health and social outcomes are imputed using effect sizes from the existing literature (Table 4). Estimates for HIV, HSV-2, marriage, and pregnancy are drawn from Psaki et al. (2019), which reports pooled impacts per additional year of schooling. Intimate partner violence estimates are from Behrman et al. (2016). Imputed benefits are capped by setting-specific incidence rates to enhance contextual plausibility. All imputed parameters are explicitly identified in the results and represent a principal source of uncertainty.

2.5 Uncertainty and simulation

Uncertainty was assessed using Monte Carlo simulation with 10 000 iterations, generating probability distributions for all parameters subject to uncertainty (Table 5). In each iteration, total benefits were recalculated to produce a cost-effectiveness estimate, capturing uncertainty in effect sizes, costs, moral weights, and the discount rate across both models. Mean values for HIV and child marriage are drawn directly from GiveWell's published moral weight framework (GiveWell 2023); weights for remaining outcomes (unintended pregnancy, HSV-2, unwanted sex, and IPV) are informed estimates based on GBD 2019 DALY burden data (GBD 2019 Diseases and Injuries Collaborators 2020) and team judgment, incorporated as uniform distributions in the Monte Carlo

Table 4. Summary of imputed effects.

Imputed parameter	Input	Source
Percentage change in future income per additional year of schooling, per person	6.80%	Duflo 2001
Standard error	0.85%	Duflo 2001
Effect (negative pp) of one year schooling on pregnancy	4.80%	Psaki *et al.* 2019
Standard error	0.59%	Psaki *et al.* 2019
Effect (negative pp) of one year schooling on marriage by endline	6.60%	Psaki *et al.* 2019
Standard error	3.72%	Psaki *et al.* 2019
Effect (negative pp) of one year schooling on HIV prevalence	11.8%	Psaki *et al.* 2019
Standard error	8.52%	Psaki *et al.* 2019
Effect (negative pp) of one year schooling on HSV-2 prevalence	7.50%	Psaki *et al.* 2019
Standard error	22.50%	Psaki *et al.* 2019
Effect (negative pp) of one year schooling on unwanted sex prevalence	6.1%	Bandiera *et al.* 2018
Standard error	2.8%	Bandiera *et al.* 2018
Effect (negative pp) of one year schooling on IPV prevalence, for girls in relationships or married[a]	9.00%	Behrman *et al.* 2016
Standard error	4.00%	Behrman *et al.* 2016
Within-household spillovers to other girls in household	10%	Assumption
Within-community economic spillovers	3%	GiveWell Staff 2024

Note. [a] The imputation for intimate partner violence accounts for only the proportion of targeted AGYW in relationships — roughly approximated as the share of girls married by endline.

simulation to reflect uncertainty about their appropriate values. Range widths reflect plausible bounds given the limited empirical basis for these estimates.

Table 5. Probability distributions for the main model parameters.

Parameter	Distribution	Mean	SD	Min	Max
Model parameters					
Discount rate	Uniform	0.04		0.04	0.07
Moral weight – unintended pregnancy	Uniform	0.5		0.25	0.75
Moral weight – child marriage	Uniform	2		1	3
Moral weight – HIV	Uniform	8		6	10
Moral weight – HSV-2	Uniform	0.25		0.25	0.5
Moral weight – unwanted sex	Uniform	0.5		0.5	0.7
Moral weight – IPV	Uniform	0.5		0.1	0.6
Income per school year (mean)	Normal	0.068	0.0085		
Outcome effects					
Unintended AGYW pregnancy	Normal	0.048	0.0059		
Child marriage	Normal	0.066	0.0372		
HIV	Normal	0.118	0.0852		
HSV-2	Normal	0.075	0.225		
Unwanted sex	Normal	0.061	0.028		
Intimate partner violence	Normal	0.09	0.04		

To identify key drivers, standardized regression coefficients were estimated by regressing simulated returns on model inputs. As a robustness check, HIV and HSV-2 parameters were excluded. Results indicate substantial heterogeneity in the sensitivity of returns to moral weight and discount rate assumptions.

2.6 Heterogeneity and robustness analyses

A Bayesian hierarchical model assessed whether program impacts varied by beneficiary age group. This approach accommodates the small sample (n=37), the simulation-derived return estimates, and the hierarchical nesting of programs within age categories. Posterior distributions were compared across four primary groups: Very Young (10–14), Mid (15–18), Older (19–24), and Full Range (10–24), with two composite categories (10–18 and 15–24) also examined.

Cluster analysis assessed whether program design characteristics, such as conditionality, delivery platform, or intervention bundling, were associated with variation in cost-effectiveness. Programs were grouped using the Partitioning Around Medoids (PAM) algorithm with Gower's distance. Median returns across clusters were compared using ANOVA. Clustering inputs comprised design features only, intentionally excluding cost and return estimates to ensure clusters reflected intervention design rather than outcomes.

2.7 Limitations

Several methodological limitations should be noted. First, the imputation of missing health outcomes from pooled meta-analytic estimates (Behrman et al. 2016; Psaki et al. 2019) assumes that the relationship between schooling and health outcomes is approximately linear and generalizable across LMIC settings, which may not hold in all contexts. Second, the moral weights used to value non-economic outcomes are adapted from GiveWell's framework, which was developed for philanthropic, rather than government, priority-setting; these weights have not been validated through deliberative processes with LMIC stakeholders, and alternative valuations could materially alter the results. The limitations associated with the use of moral weights are explicitly acknowledged throughout. Third, the model does not capture implementation costs that may arise at scale, such as political economy constraints, supply-chain requirements, or capacity-building needs, which are likely to affect real-world cost-effectiveness. Fourth, the evidence base is concentrated in sub-Saharan Africa and may not be representative of other LMIC regions. Fifth, the use of a standardized US$100 000 investment as the basis for comparison abstracts from real budget allocation processes, which involve indivisibilities, fixed costs, and political constraints not captured in the model. These limitations are revisited in the discussion in relation to the interpretation of results.

This study relied exclusively on secondary data analysis and did not require ethical approval.

3 Results

Through a structured review process, 185 studies encompassing 240 intervention arms were identified and classified by study type and methodological quality. Although the process constituted a structured evidence review, it did not incorporate all elements typically associated with systematic reviews (e.g. predefined search strings, date parameters, and scoring scales). Rather, the approach was iterative in nature, drawing on resources including a collection of recent unpublished literature reviews commissioned by the Gates Foundation. Quality ratings ranged from "High," indicating well-powered randomized controlled trials with strong causal validity, to "Low," reflecting descriptive or non-comparative designs, with credibility declining as methodological rigor decreased (Table 6). Studies rated as high or medium-high quality were prioritized for inclusion, while descriptive, observational, and non-intervention studies were excluded from the model. Following screening, 24 studies comprising 37 intervention arms were included in the final model, with additional studies contributing to parameter estimation and robustness checks. The included interventions were diverse and often multi-component, encompassing cash transfers, scholarship programs, school meals, safe spaces, and empowerment initiatives, and were predominantly conducted in sub-Saharan Africa, with a smaller number from South Asia.

Table 6. Summary of reviewed literature.

			Quality assessment (relative to RCT)				
Study type	**Studies**	**Arms**	**High**	**Med.-High**	**Medium**	**Low-Med.**	**Low**
RCT	12	23	22	1			
Cluster RCT	23	39	35	2	2		
Quasi-experimental	25	41	1	10	29	1	
Natural experiment	1	2			2		
Model based on review	4	6			3	3	
Review	46	55	11	6	26	5	7
Observational	63	63			3	24	36
Informal/grey literature	11	11			1		10
Grand total	**185**	**240**	**69**	**19**	**66**	**33**	**53**

3.1 Decomposition of program value

Under the societal welfare perspective adopted in this analysis and benchmarked against the estimated impact of unconditional cash transfers (the AGAC benchmark of 135 welfare value units), the analysis demonstrated that integrating health and social outcomes substantially alters the valuation of education-related interventions. In the Full Assumptions model, health-related impacts accounted for a mean of approximately 59% of total program value (IQR: 35%–83%), while long-term economic effects contributed approximately 16% and short-term consumption gains approximately 25% (Figure 3). For many programs, particularly those with imputed HIV effects, the health share was considerably higher — over three-quarters of total value for the median program. These findings indicate that evaluations restricted to educational or economic outcomes alone substantially underestimate the total returns of AGYW interventions.

3.2 Distribution of returns

Simulated median returns across the 37 interventions exhibited substantial heterogeneity, ranging from near-zero to approximately 2 000 welfare value units per dollar invested. Figure 4 presents these estimates, with the upper panel displaying median returns for individual programs and the lower panel summarizing the distribution using a boxplot; the horizontal line at 135 welfare value units denotes the AGAC benchmark described above. The overall median return was approximately 200 welfare value units, modestly exceeding the AGAC benchmark, although a subset of approximately 10 interventions delivered returns several times higher than the benchmark. These findings indicate that while many AGYW interventions yield impacts comparable to, or only marginally better than, direct cash transfers, substantial heterogeneity exists, highlighting the potential for high value-for-money through selective investment. As mentioned above, the AGAC benchmark sets a high standard for cost-effectiveness, and decisionmakers may reasonably choose

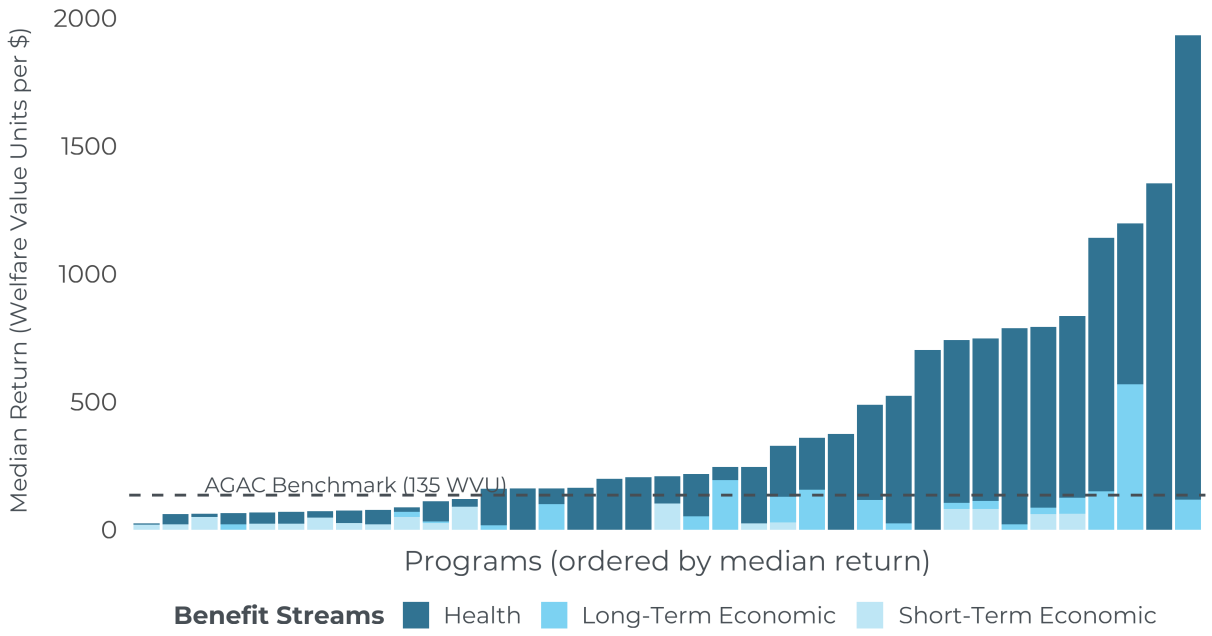


Figure 3. Health benefits account for approximately 59% of total returns from schooling programs

Note. Results from Monte Carlo simulations (with 10 000 draws per program) under the full assumptions models.

lower thresholds for investment that would make the education-related interventions studied even more attractive.

3.3 Highest- and lowest-performing interventions

A preliminary examination of the highest- and lowest-performing interventions revealed systemic patterns in cost-effectiveness. The five highest-performing programs achieved median returns of approximately 800–2 000 welfare value units —at least six times the AGAC benchmark—whereas the five lowest-performing programs fell well below this threshold, shown in Figures 5 and 6. High-performing interventions were typically characterized by low per-beneficiary costs and a narrow focus on a single, well-defined mechanism, such as community-based discussions to prompt norm change or modest financial incentives to delay marriage. In contrast, lower-performing programs were more often comprehensive, multi-component interventions employing complex "cash-plus" or "big push" models; although these approaches may generate meaningful outcomes, their higher costs reduced overall cost-effectiveness. Taken together, these findings suggest that in this sample, simpler, lower-cost program designs may yield higher returns per dollar invested, noting that real world decision-making will want to include more than just cost-effectiveness considerations.

3.4 Program design clusters

Most interventions included in the analysis comprised multiple components. Cluster analysis was conducted to assess whether specific elements within these programs were responsible for driving

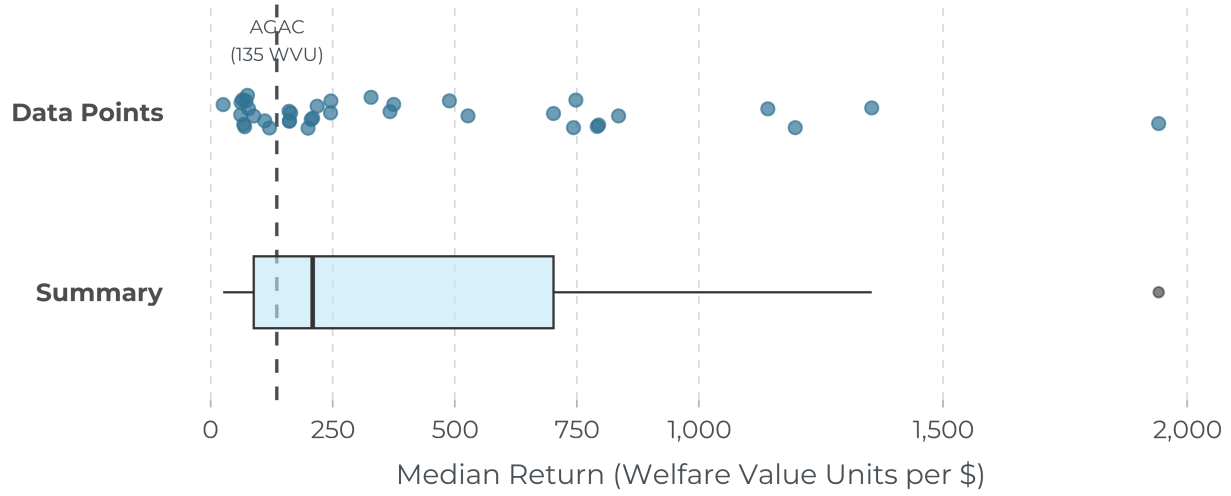


Figure 4. Most program returns exceed the cash benchmark

Note. Results are from a Monte Carlo simulation (10 000 draws per program) using the Full Assumptions model. Each data point represents the median return from a single program's simulation. The boxplot summarizes the distribution of these 37 medians.

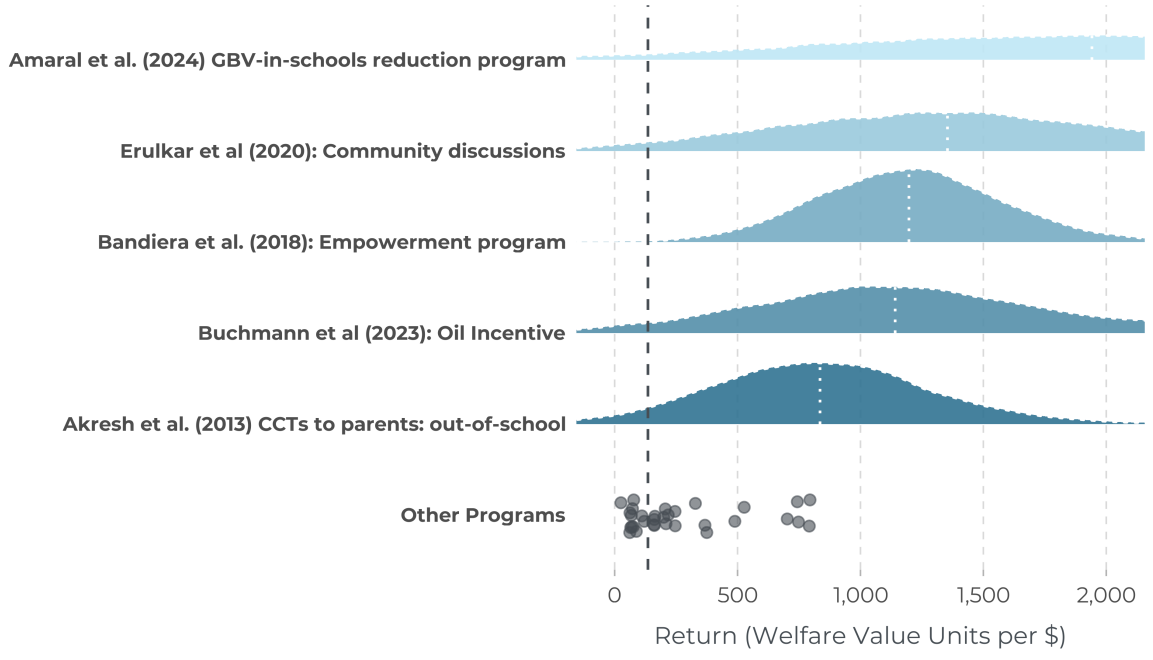


Figure 5. The top five interventions by cost-effectiveness display* *relatively large returns

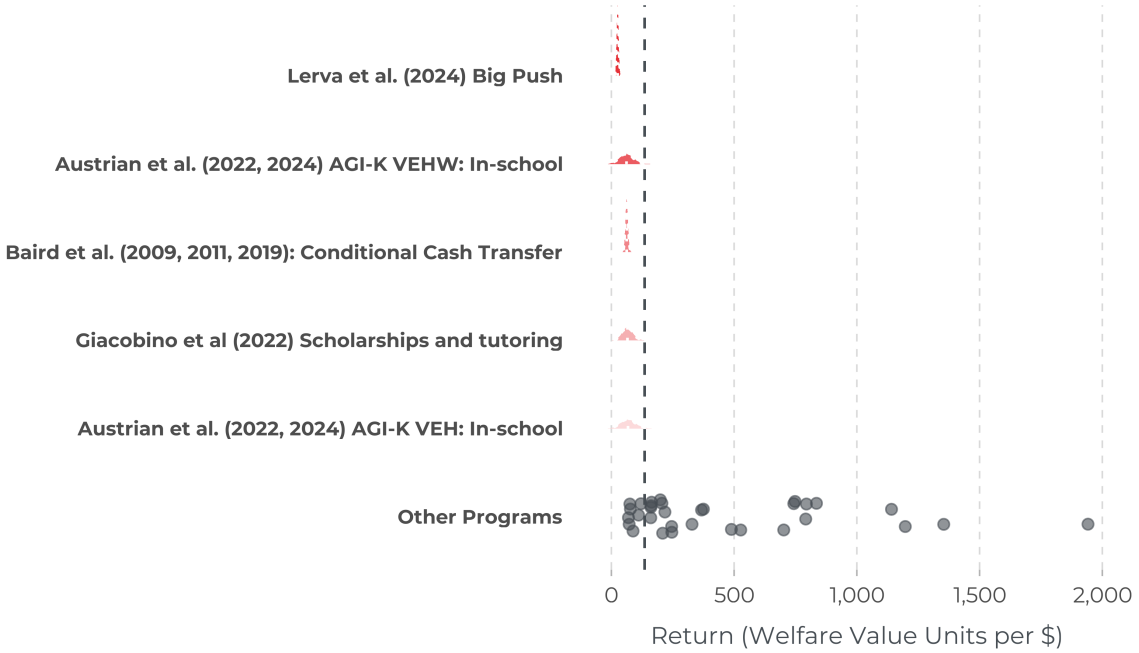


Figure 6. The bottom five interventions by return show* *very low returns and cluster together tightly

economic or health gains, and whether isolating such elements could improve impact and cost-effectiveness. Programs were grouped according to their design features. Clusters were derived from design features alone — conditionality, delivery platform, and bundling — intentionally excluding cost and return estimates to ensure groupings reflected intervention architecture rather than outcomes. This approach generated three distinct clusters:

1. Cluster 1: Multidimensional Program for the Youngest – multi-component interventions typically targeting the youngest adolescents in the sample.
2. Cluster 2: Cash Transfers for Education – predominantly cash transfer programs reporting education-related outcomes.
3. Cluster 3: Cash Transfers for Health – predominantly cash transfer programs reporting health-related outcomes.

To assess whether cost-effectiveness varies across the identified clusters, median returns were compared across groups using a one-way ANOVA. No statistically significant differences in cost-effectiveness were detected across groups ($p = 0.35$), although statistical power was limited by the small sample size.

3.5 Age-group heterogeneity

The Bayesian hierarchical model revealed meaningful variation in cost-effectiveness across beneficiary age groups, although posterior distributions overlapped substantially, and individual

program variation exceeded variation across age categories. Programs targeting Mid adolescents (15–18) exhibited the highest median posterior returns ($176; 95% CrI: $91–$426), while programs targeting only Very Young adolescents (10–14) showed the lowest ($75; 95% CrI: $42–$147). The composite Very Young plus Mid category (10–18) ranked second ($163; 95% CrI: $94–$296), suggesting that programs spanning the early-to-mid adolescent range may capture cumulative effects even where the youngest cohort alone performs less well. The probability that the Older group (19–24) outperformed the Very Young group (10–14) was approximately 91%, while comparisons among the Mid (15–18), Older (19–24), and Full Range (10–24) categories were less conclusive. These findings suggest that, where resources are constrained, programs targeting mid-adolescents—who are most accessible through lower-secondary school systems—may yield higher returns per dollar invested. The lower returns for the Very Young group may partly reflect the higher per-beneficiary costs typical of programs targeting younger girls rather than a genuine difference in impact. As with the broader results, this pattern may partly reflect the composition of the evidence base rather than a generalizable age–return gradient and should be interpreted alongside the specific programmatic mechanisms operating in each age group.

3.6 Reported vs. full-assumptions models

To assess the contribution of imputed outcomes, results from the Reported Parameters Model (which relies solely on outcomes directly measured in primary studies) were compared with the Full Assumptions Model (which incorporates imputed health and social effects). Median cost-effectiveness under the Reported Parameters Model was 60 welfare value units — approximately half the AGAC benchmark — reflecting only educational and short-term economic gains. Under the Full Assumptions Model, median cost-effectiveness rose to 200 welfare value units, a more than threefold increase driven primarily by imputed reductions in HIV, unintended pregnancy, and child marriage. The relative ranking of interventions was broadly preserved across both models: programs that performed well under reported parameters also tended to perform well under full assumptions, although the magnitude of estimated returns was substantially larger. This consistency suggests that the framework's comparative conclusions are reasonably robust, even as absolute estimates remain sensitive to imputation. The results of both models for all 37 interventions are presented in Appendix 1.

3.7 Sensitivity analysis

To identify which model parameters exert the greatest influence on cost-effectiveness estimates, sensitivity analysis was conducted using linear regression. Simulated program returns were regressed on all model inputs—including the discount rate, moral weights, and imputed health effects—to estimate the marginal association between variation in each parameter and variation in estimated returns. This approach identifies parameters with the strongest marginal predictive power but does not measure their substantive importance in an absolute sense.

Results presented in Figure 7 indicate that HIV and HSV-2 exhibit the largest standardized coefficients and are therefore the strongest predictors of overall cost-effectiveness. This finding is consistent with the core model: HIV is assigned a high moral weight, such that even modest reductions in transmission generate substantial gains in terms of cost-effectiveness. In contrast, several other parameters—including the discount rate and other moral weights— exhibit coefficients clustered near zero, indicating comparatively limited influence. Taken together, these findings underscore that imputation of health outcome effects and the moral weights assigned to those outcomes are key determinants of estimated program returns.

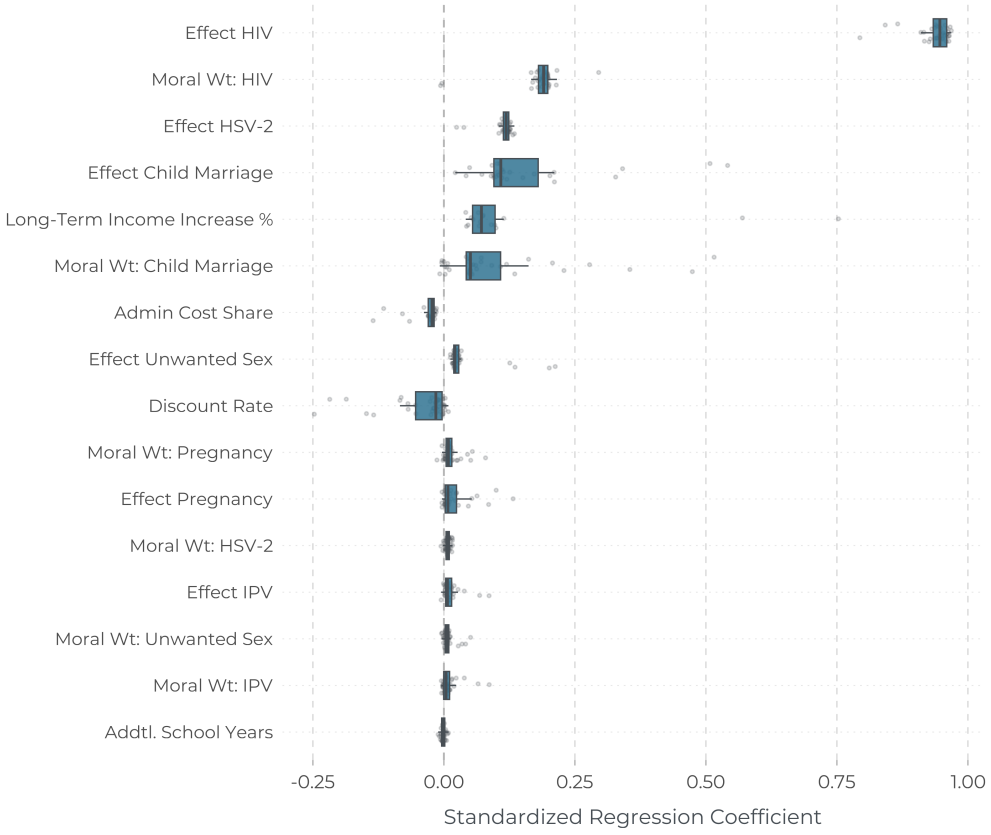


Figure 7. HIV and HSV-2 have large effect sizes on the program's overall return

The contribution of these health effects was further examined by removing HIV and HSV-2 impacts from the model. Figure 8 compares the distribution of median cost-effectiveness estimates from the Base Model with those of a model where HIV and HSV-2 are removed. Circle size denotes median cost per beneficiary for each program.

Excluding these outcomes shifts the distribution of cost-effectiveness estimates leftward, reflecting lower overall cost-effectiveness. This reduction is most pronounced among programs with the highest Base Model returns, for which HIV-related impacts account for a substantial share of total benefits. Nonetheless, the shape of the distribution remains broadly similar: a sizable proportion of

interventions continue to cluster around the AGAC benchmark, and several still exceed it, albeit by smaller margins. These results confirm that while estimates are sensitive to high-value health outcomes, the structure of the model is robust, and not driven solely by HIV and HSV-2 effects. Consistent with earlier findings, programs with lower per-beneficiary costs remain among the most cost-effective, a pattern that is robust to the exclusion of HIV and HSV-2 benefits. These findings underscore both the importance of including health measures in evaluations and the need for transparency about which health effects are directly observed versus imputed — a distinction that should inform how confidently decision-makers interpret the magnitude of estimated returns.

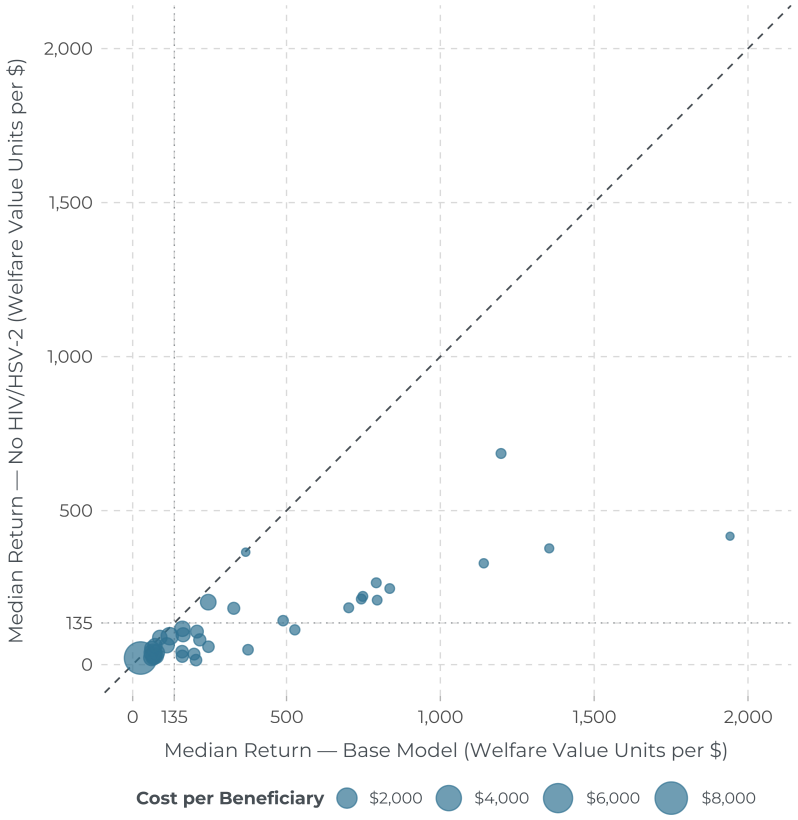


Figure 8. Removing HIV/HSV-2 effects on the models reduces return calculations substantially

4 Discussion

This analysis demonstrates that health and social outcomes account for approximately 59% of total estimated program value for education-related interventions targeting AGYW in LMICs. This finding confirms our core hypothesis that conventional education-focused evaluations substantially underestimate total returns. These findings are consistent with a wider literature documenting the established causal pathways between schooling and long-term health, as well as the economic and social costs associated with child marriage and early fertility (Wodon *et al.* 2017; Psaki *et al.* 2019).

Median cost-effectiveness of the 37 intervention arms was approximately 200 welfare value units, modestly exceeding the unconditional cash transfer benchmark of 135 welfare value units (GiveWell 2024; GiveWell Staff 2024; Oehlsen 2024). However, returns were highly heterogeneous, ranging from near-zero to approximately 2 000 welfare value units. Simpler, lower-cost programs — such as conditional transfers in rural Bangladesh that offered cooking oil to households of unmarried girls aged 15–17, provided they remained unmarried until age 18 — tended to outperform more complex, expensive interventions, such as the Adolescent Girls Initiative (AGI-K) in Kenya — in terms of cost-effectiveness (Kangwana *et al.* 2022; Buchmann *et al.* 2023; Austrian *et al.* 2024). This pattern may reflect higher operational costs rather than lower effectiveness in absolute terms and should be interpreted cautiously given that complex programs may operate in more difficult contexts or generate benefits not captured in this framework.

Sensitivity analysis indicated that cost-effectiveness estimates were most sensitive to the moral weights and imputed effect sizes assigned to HIV and HSV-2, which were the dominant drivers of health-related returns (GBD 2019; Diseases and Injuries Collaborators 2020; WHO 2020; GiveWell 2023). This finding reflects the high moral weight assigned to HIV avoidance relative to other outcomes, and the substantial uncertainty in the pooled estimates of schooling effects on HIV prevalence drawn from Psaki et al. (2019).

These findings are highly relevant to current debates regarding priority-setting tools in LMICs. Frameworks such as WHO-CHOICE enable essential comparison of health interventions using the DALY metric, but are not designed to incorporate non-health outcomes, such as education, early marriage, and income (Hutubessy *et al.* 2002; Ottersen and Norheim 2014). Extended Cost-Effectiveness Analysis (ECEA) incorporates financial risk protection and distributional equity, but similarly remains anchored within health-sector objectives without valuing the broader social or educational returns of an intervention (Verguet *et al.* 2015; Verguet *et al.* 2016; Fraser *et al.* 2024). Standard health technology assessment (HTA), while increasingly institutionalized in LMICs, typically evaluates single technologies against a cost-per-DALY threshold and does not provide a basis for comparing programs with benefits across multiple sectors (Chalkidou *et al.* 2016; Nemzoff *et al.* 2025). The welfare-based approach applied here addresses these gaps by expressing health, educational, and social outcomes on a common scale, although it does not capture the distributional effects that ECEA or distributional cost-effectiveness analysis (DCEA) are designed to measure.

DALYs remain the predominant metric in LMIC health priority-setting and are embedded in institutional processes. However, standard cost-effectiveness often misses dimensions of value, such as the equity criteria, that are critical to ministerial decision-making outside of health (Norheim *et al.* 2014). The welfare-based approach used here, rooted in the Mincer-equation tradition, is intended as a complement to existing metrics rather than a replacement, and could be used alongside DALY-based analyses in joint planning processes (Duflo 2001). Integrating equity weights from ECEA or DCEA alongside moral weights would strengthen the framework by capturing distributional as

well as aggregate impact, and this represents a priority for further methodological development (Asaria *et al.* 2016; Cookson *et al.* 2017).

The framework could be applied within existing joint health-education sector review processes. These platforms offer a natural entry point for the joint investment appraisal of programs targeting adolescent girls and young women, where interventions yield returns that are otherwise fragmented across ministry silos. Such appraisals could inform the allocation of shared budget envelopes, moving away from "fixed" benefits packages that often fail to account for non-health consequences (Glassman *et al.* 2017). This approach directly addresses the growing policy demand for integrated strategies that empower girls through education to achieve long-term economic and health outcomes (World Bank 2024). However, the leadership for such frameworks should not rest with a single ministry. To successfully institutionalize cross-sectoral resource allocation, Ministries of Finance or national planning commissions should take the lead. These central agencies are better positioned to evaluate the comprehensive investment case for health and education, ensuring that budget allocations reflect their high rate of return on human capital and economic growth (Jamison *et al.* 2013; Stenberg *et al.* 2017).

Routine application of the framework would require a coordinated data infrastructure. While many LMICs already collect the necessary baseline metrics – HIV incidence, adolescent pregnancy, and child marriage, education outcomes, and program costs – through Demographic and Health Surveys (DHS), Multiple Indicator Cluster Surveys (MICS), and Education Management Information Systems (EMIS), they are rarely synthesized into a single analytic platform (Countdown to 2030 Collaboration 2018; UNICEF 2018; UNAIDS 2023; UNESCO 2023; UNAIDS 2025; UNICEF 2025).

The observed heterogeneity in returns across programs should be interpreted with caution, as complex programs may better target hard-to-reach populations or produce benefits not captured by this framework, such as psychosocial well-being, increased agency, or the accumulation of social capital. Complex, multi-component interventions—such as the BRAC ELA model—have demonstrated high performance in specific domains despite their high delivery costs (Bandiera *et al.* 2018). Conversely, evaluations like the AGI-K study show that while multi-sectoral packages may yield lower cost-effectiveness in a narrow health-only space, they often successfully influence unmeasured psychosocial pathways that are critical for long-term resilience (Gibbs *et al.* 2012; Austrian *et al.* 2021; Austrian *et al.* 2024).

The absence of statistically significant differences across program design clusters (ANOVA p=0.35) should not be interpreted as a definitive null finding given the small sample size (n=37). This exploratory analysis suggests that while program designs differed substantially, design features alone did not reliably explain variation in cost-effectiveness; instead, contextual factors and implementation quality may have played a more influential role (Erulkar and Muthengi 2007; Erulkar *et al.* 2020). Targeting the youngest cohorts within the school system remains one of the

highest returns in human capital, preventing the compounding vulnerabilities that emerge later (Sperling and Winthrop 2015).

Several limitations warrant consideration. First, the finding that health-related outcomes account for approximately 59% of total program value depends substantially on imputed rather than directly observed effects, particularly for HIV and HSV-2 (Behrman *et al.* 2016; Psaki *et al.* 2019). This highlights a persistent evidence gap in LMICs, where the longitudinal link between educational interventions and long-term health outcomes is often assumed rather than empirically documented (Snilstveit *et al.* 2015). Second, the moral weights are adapted from GiveWell's framework, which was developed for philanthropic rather than government priority-setting, and has not been validated through LMIC-specific deliberative processes (GiveWell 2023; GiveWell Staff 2024; Hickman 2025). Consequently, the model incorporates normative choices that may not fully align with local priorities in diverse development contexts (Norheim et al . 2014). However, sensitivity analyses indicate that while alternative weighting schemes shift absolute returns, the relative rankings of interventions remain partially robust. Third, the evidence base is geographically concentrated in sub-Saharan Africa (20 out of 24 studies), which limits generalizability. The analytical framework is portable, but the parameter estimates are not, and effect sizes for school enrollment and health outcomes tend to vary significantly by region (Petrosino *et al.* 2012). Fourth, the model does not capture outcomes related to mental health, psychosocial well-being, agency, and social capital, which are foundational to adolescent development (Patel *et al.* 2018). While some studies in our model measure several psychosocial outcomes, these specific outcomes are not captured (Bandiera *et al.* 2018).

Further research should prioritize in-country validation through pilots in settings with active joint health-education planning, such as Kenya, Tanzania, or Bangladesh. These contexts already possess a rich evidence base from multi-sectoral trials like AGI-K and BALIKA, providing a robust foundation for institutionalizing priority-setting tools (Amin *et al.* 2018; Austrian et al 2021; Austrian *et al.* 2024; Ainul *et al.* 2022). International technical assistance organizations, including the International Decision Support Initiative (iDSI), could support the integration of these frameworks into national planning cycles (Baker *et al.* 2023). Methodologically, the framework should evolve by integrating equity dimensions from Extended Cost-Effectiveness Analysis (ECEA) to ensure that distributional impacts are captured alongside total welfare gains (Verguet et al . 2016; Cookson *et al.* 2017). Furthermore, expanding the outcome set to include psychosocial well-being and agency—and developing corresponding moral weights through inclusive, deliberative methods—will better reflect the holistic value of adolescent programming (Baltussen *et al.* 2017; Patel *et al.* 2018; Kapiriri *et al.* 2020). Evidence-informed investment in AGYW requires analytic tools that make multi-sectoral benefits visible, comparable, and contestable. This framework offers a starting point; its ultimate value depends on its adaptation, validation, and institutionalization within the planning processes of the governments it is designed to serve.

5 Conclusion

This study introduced an integrated cost-effectiveness framework for education-related interventions targeting AGYW in LMICs and demonstrated that health and social outcomes, systematically omitted from conventional education and economic evaluations, account for approximately three-fifths of total program value. By expressing these benefits in a common welfare-based metric alongside educational and economic returns, the framework addresses a core limitation of sector-specific priority-setting tools and makes the value judgments underlying resource allocation transparent and contestable.

Three findings carry direct implications for policy. First, median cost-effectiveness modestly exceeded the unconditional cash transfer benchmark, but a subset of approximately 10 interventions achieved returns several-fold higher, indicating substantial scope for value-for-money gains through selective investment. Second, simpler, lower-cost program designs tended to outperform more complex multi-component models, though this pattern should be interpreted cautiously given that complex programs may operate in more challenging contexts or produce unmeasured benefits. Third, the sensitivity of results to HIV-related parameters underscores that the case for investing in AGYW programming is strongest where schooling materially affects sexual and reproductive health risks, predominantly in sub-Saharan Africa, and that policy application in other regions requires re-estimation using local incidence and cost data.

These findings support four recommendations for LMIC policymakers and donors. First, health outcomes should be systematically incorporated into the evaluation of education programs for AGYW; evaluations that measure only educational or labor market returns will undervalue programs that reduce HIV, unintended pregnancy, and child marriage. Second, Ministries of Finance and national planning commissions — which already arbitrate across sectoral budget envelopes — are well positioned to institutionalize cross-sectoral cost-effectiveness frameworks within medium-term expenditure planning and joint Health–Education sector reviews. Third, routine application of this framework requires investment in linked data systems that integrate health surveillance (DHS, MICS), education management information systems (EMIS), and program cost data, platforms that exist in most LMICs but are rarely connected. Fourth, the moral weights used to value non-economic outcomes should be adapted through deliberative processes involving national stakeholders, including health and education officials, community representatives, and adolescent girls themselves — to ensure that the framework reflects local priorities rather than externally imposed valuations.

References

Ainul S, Noor F, Hossain MI *et al.* 2022. Keeping girls in schools to reduce child marriage in rural Bangladesh: endline assessment. Dhaka: Population Council.

Akresh R, de Walque D, Kazianga H. 2013. Cash transfers and child schooling: evidence from a randomized evaluation of the role of conditionality. World Bank Policy Research Working Paper 6340. Washington, DC: World Bank.

Amaral S, Garcia-Ramos A, Gulesci S *et al*. 2024. Gender-based violence in schools and girls' education: experimental evidence from Mozambique. NBER Working Paper 33203. Cambridge, MA: National Bureau of Economic Research.

Amin S, Saha JS, Ahmed JA. 2018. Skills-building programs to reduce child marriage in Bangladesh: a randomized controlled trial. *Journal of Adolescent Health* **63**: 293–300.

Asaria M, Griffin S, Cookson R. 2016. Distributional cost-effectiveness analysis: a tutorial. *Medical Decision Making* **36**: 8–19.

Aurino E, Gelli A, Adamba C, Osei-Akoto I, Alderman H. 2020. Food for thought? Experimental evidence on the learning impacts of a large-scale school feeding program. *Journal of Human Resources*. https://www.researchgate.net/publication/347661991_Food_for_thought_Experimental_Evidence_on_the_Learning_Impacts_of_a_Large-Scale_School_Feeding_Program (21 May 2026, date last accessed).

Austrian K, Maluccio JA, Soler-Hampejsek E *et al.* 2024. Long-term impacts of a cash plus program on marriage, fertility, and education after six years in pastoralist Kenya: a cluster randomized trial. *SSM – Population Health* **26**: 101663.

Austrian K, Soler-Hampejsek E, Kangwana B *et al.* 2021. Impacts of two-year multisectoral cash plus programs on young adolescent girls' education, health and economic outcomes: Adolescent Girls Initiative–Kenya (AGI-K) randomized trial. *BMC Public Health* **21**: 2159.

Austrian K, Soler-Hampejsek E, Kangwana B *et al*. 2022. Impacts of multisectoral cash plus programs on marriage and fertility after four years in pastoralist Kenya: a randomized trial. *Journal of Adolescent Health* **70**: 885–94.

Baird S, Chirwa E, McIntosh C, Özler B. 2009. *The Short-Term Impacts of a Schooling Conditional Cash Transfer Program on the Sexual Behavior of Young Women*. World Bank Policy Research Working Paper 5089. Washington, DC: World Bank.

Baird S, McIntosh C, Özler B. 2011. Cash or condition? Evidence from a cash transfer experiment. *Quarterly Journal of Economics* **126**: 1709–53.

Baird SJ, Garfein RS, McIntosh CT, Özler B. 2012. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *The Lancet* **379**: 1320–9.

Baird S, McIntosh C, Özler B. 2019. When the money runs out: do cash transfers have sustained effects on human capital accumulation? *Journal of Development Economics* **140**: 169–85.

Baker P, Barasa E, Chalkidou K *et al.* 2023. International partnerships to develop evidence-informed priority setting institutions: ten years of experience from the International Decision Support Initiative (iDSI). *Health Systems & Reform* **9**: 2330112.

Baltussen R, Jansen MPM, Bijlmakers L *et al.* 2017. Value assessment frameworks for HTA agencies: the organization of evidence-informed deliberative processes. *Value in Health* **20**: 256–60.

Bandiera O, Buehren N, Goldstein M *et al.* 2018. *The Economic Lives of Young Women in the Time of Ebola: Lessons from an Empowerment Program*. Working Paper. Abdul Latif Jameel Poverty Action Lab.

Barroy H, Yameogo P, Blecher M *et al.* 2024. Public financial management: a pathway to universal health coverage in low- and middle-income countries. *Health Systems & Reform* **10**:3.

Behrman JA, Peterman A, Palermo T. 2016. Does keeping adolescent girls in school protect against sexual violence? Quasi-experimental evidence from East and Southern Africa. *Journal of Adolescent Health* **60**: 184–90.

Bonfert A, Wadhwa D. 2024. Tracing global trends in education: a tale of old and new gender gaps. World Bank Gender Data Portal. https://genderdata.worldbank.org/en/data-stories/a-tale-of-old-and-new-gender-gaps (20 May 2026, date last accessed).

Blimpo MP, Gajigo O, Pugatch T. 2019. Financial constraints and girls' secondary education: evidence from school fee elimination in The Gambia. *World Bank Economic Review* **33**: 185–208.

Buchmann N, Field E, Glennerster R *et al.* 2023. A signal to end child marriage: theory and experimental evidence from Bangladesh. *American Economic Review* **113**: 2645–88.

Chalkidou K, Glassman A, Marten R *et al.* 2016. Priority-setting for achieving universal health coverage. *Bulletin of the World Health Organization* **94**: 462–7.

Chalkidou K, Li R, Culyer AJ *et al.* 2017. Health technology assessment: global advocacy and local realities. *International Journal of Health Policy and Management* **6**: 233–6.

Coefficient Giving. n.d. *Guide to BOTECs for Global Health and Wellbeing*. Coefficient Giving. https://docs.google.com/document/d/1cM0qtfGAE9SzmL2nZaEB0ePXM6GYqXjgV3vkv7e_XvU/edit?tab=t.0 (23 May 2026, date last accessed).

Cohen I, Abubakar M, Perlman D. 2023. *Pathways to Choice: A Bundled Intervention against Child Marriage*. CEGA Working Papers Series, WPS Paper No. 230. Berkeley: Center for Effective Global Action, University of California. https://escholarship.org/uc/item/33j1k1k4 (21 May 2026, date last accessed).

Cookson R, Mirelman AJ, Griffin S *et al.* 2017. Using cost-effectiveness analysis to address health equity concerns. *Value in Health* **20**: 206–12.

Countdown to 2030 Collaboration. 2018. Countdown to 2030: tracking progress towards universal coverage for reproductive, maternal, newborn, and child health. *The Lancet* **391**: 1538–48.

Crawfurd L, Hares S, Le Nestour A *et al.* 2019. Does education need a QALY and is LAYS it? Center for Global Development. https://www.cgdev.org/blog/does-education-need-qaly-and-lays-it (20 May 2026, date last accessed).

Davis M, Ingwersen N, Kazianga H, Linden L, Mamun A, Protik A, Sloan M. 2016. *Ten-year impacts of Burkina Faso's BRIGHT program*. Washington, DC: Mathematica Policy Research. https://mathematica.org/publications/ten-year-impacts-of-burkina-fasos-bright-program (22 May 2026, date last accessed).

Duflo E. 2001. Schooling and labor market consequences of school construction in Indonesia: evidence from an unusual policy experiment. *American Economic Review* **91**: 795–813.

Duflo E, Dupas P, Kremer M, Sinei S. 2006. Education and HIV/AIDS prevention: evidence from a randomized evaluation in western Kenya. World Bank Policy Research Working Paper 4024. Washington, DC: World Bank.

Duflo E, Dupas P, Kremer M. 2015. Education, HIV, and early fertility: experimental evidence from Kenya. *American Economic Review* **105**: 2757–97.

Duflo E, Dupas P, Kremer M. 2021. The impact of free secondary education: experimental evidence from Ghana. NBER Working Paper 28937. Cambridge, MA: National Bureau of Economic Research.

Duflo E, Dupas P, Spelke E, Walsh M. 2024. Intergenerational impacts of secondary education: experimental evidence from Ghana. Working Paper. Cambridge, MA: Massachusetts Institute of Technology / Princeton University.

Erulkar A, Medhin G, Weissman E *et al.* 2020. Designing and evaluating scalable child marriage prevention programs in Burkina Faso and Tanzania: a quasi-experiment and costing study. *Global Health, Science and Practice* **8**: 68–81.

Erulkar AS, Muthengi E. 2007. *Evaluation of Berhane Hewan: A Pilot Program to Delay Marriage in Rural Ethiopia*. Population Council.

Filmer D, Rogers H, Angrist N, Sabarwal S. 2020. Learning-adjusted years of schooling (LAYS): defining a new macro measure of education. *Economics of Education Review* **77**: 101971.

Fraser HL, Feldhaus I, Edoka IP *et al.* 2024. Extended cost-effectiveness analysis of interventions to improve uptake of diabetes services in South Africa. *Health Policy and Planning* **39**: 253–67.

Gajderowicz T, Jakubowski M, Kennedy A *et al.* 2025. *The Learning Crisis: Three Years After COVID-19*. arXiv Preprint arXiv:2501.01260.

Gajigo O. 2012. *Closing the Education Gender Gap: Estimating the Impact of Girls' Scholarship Program in The Gambia*. Working Paper Series No. 164. Tunis: African Development Bank. https://www.afdb.org/sites/default/files/documents/publications/working_paper_164_-_closing_the_education_gender_gap-_estimating_the_impact_of_girls_scholarship_program_in_the_gambia.pdf (21 May 2026, date last accessed).

GBD 2019 Diseases and Injuries Collaborators. 2020. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* **396**: 1204–22.

Giacobino H, Huillery E, Michel B, Sage M. 2024. Schoolgirls, not brides: education as a shield against child marriage. *American Economic Journal: Applied Economics* **16**: 109–43.

Gibbs A, Willan S, Misselhorn A, Mangoma J. 2012. Combined structural interventions for gender equality and livelihood security: a critical review of the evidence from southern and eastern Africa and the implications for young people. *Journal of the International AIDS Society* **15**(Suppl 1): 17362.

GiveWell. 2023. *GiveWell Cost-Effectiveness Analysis – Version 4 (Public)*. GiveWell. https://docs.google.com/spreadsheets/d/1Kq6iHSQFr3eRz1p9KclHuJTQiaJYkpViOyneSd9KCJc/edit?gid=1377543212#gid=1377543212 (20 May 2026, date last accessed).

GiveWell. 2024. *GiveDirectly's Cash for Poverty Relief Program*. GiveWell. https://www.givewell.org/international/technical/programs/givedirectly-cash-for-poverty-relief-program-november-2024-version (20 May 2026, date last accessed).

GiveWell Staff. 2024. Re-evaluating the impact of unconditional cash transfers. The GiveWell Blog. https://blog.givewell.org/2024/11/12/re-evaluating-the-impact-of-unconditional-cash-transfers/ (20 May 2026, date last accessed).

Glassman A, Giedion U, Smith PC (eds). 2017. *What's In, What's Out: Designing Benefits for Universal Health Coverage*. Center for Global Development.

Hahn Y, Islam A, Nuzhat K, Smyth R, Yang H-S. 2018. Education, marriage, and fertility: long-term evidence from a female stipend program in Bangladesh. *Economic Development and Cultural Change* **66**: 383–415.

Handa S, Peterman A, Huang C, Halpern C, Pettifor A, Thirumurthy H. 2015. Impact of the Kenya cash transfer for orphans and vulnerable children on early pregnancy and marriage of adolescent girls. *Social Science & Medicine* **141**: 36–45.

Hamory J, Kremer M, Mbiti I, Miguel E. 2016. *Evaluating the impact of vocational education vouchers on out-of-school youth in Kenya*. 3ie Impact Evaluation Report 37. New Delhi: International Initiative for Impact Evaluation (3ie). http://3ieimpact.org/evidence-hub/publications/impact-evaluations/evaluating-impact-vocational-education-vouchers-out (22 May 2026, date last accessed).

Hevia FJ, Vergara-Lope S, Velásquez-Durán A, Calderón D. 2021. Estimation of the fundamental learning loss and learning poverty related to COVID-19 pandemic in Mexico. *International Journal of Educational Development* **88**: 102515.

Hickman P. 2025. Guide to back-of-the-envelope calculations for global health and wellbeing. Open Philanthropy. https://coefficientgiving.org/research/guide-to-back-of-the-envelope-calculations-for-global-health-and-wellbeing/ (23 May 2026, date last accessed).

Hutubessy R, Baltussen R, Torres-Edejer TT, Evans DB. 2002. Generalised cost-effectiveness analysis: an aid to decision making in health. *Applied Health Economics and Health Policy* **1**: 89–95.

Idara-e-Taleem-o-Aagahi. 2021. *Measuring the Impact of COVID-19 on Education in Pakistan: ASER 2021 Final Report*. Annual Status of Education Report. https://aserpakistan.org/document/aser/2021/reports/national/ASER_report_National_2021.pdf (20 May 2026, date last accessed).

Jamison DT, Summers LH, Alleyne G *et al.* 2013. Global health 2035: a world converging within a generation. *The Lancet* **382**: 1898–1955.

Joint United Nations Programme on HIV/AIDS (UNAIDS). 2023. *The Path That Ends AIDS: Global AIDS Update 2023*. UNAIDS. https://www.unaids.org/en/resources/documents/2023/global-aids-update-2023 (20 May 2026, date last accessed).

Joint United Nations Programme on HIV/AIDS (UNAIDS). 2025. *Global HIV & AIDS Statistics — Fact Sheet*. UNAIDS. https://www.unaids.org/en/resources/fact-sheet (20 May 2026, date last accessed).

Kangwana B, Austrian K, Soler-Hampejsek E *et al.* 2022. Impacts of multisectoral cash plus programs after four years in an urban informal settlement in Kenya: a randomized trial. *PLOS ONE* 17: e0262858.

Kapiriri L, Baltussen R, Oortwijn W. 2020. Implementing evidence-informed deliberative processes in health technology assessment: a low income country perspective. *International Journal of Technology Assessment in Health Care* **36**: 29–33.

Kazianga H, Linden LL, Protik A, Sloan M. 2019. *The Medium-Term Impacts of Girl-Friendly Schools: Seven-Year Evidence from School Construction in Burkina Faso*. NBER Working Paper 26006. Cambridge, MA: National

Bureau of Economic Research. https://leighlinden.com/BRIGHT%207%202019-06.pdf (21 May 2026, date last accessed).

Kurowski C *et al*. 2021. *From double shock to double recovery: implications and options for health financing in the time of COVID-19*. Health, Nutrition and Population Discussion Paper. Washington, DC: World Bank.

Lambon-Quayefio M, Peterman A, Handa S *et al.* 2023. Government unconditional cash transfers and safe transitions to adulthood in Malawi. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4322756 (30 May 2026, date last accessed).

Lerva B, Ferris D, Fornasari M. 2024. A 'Big Push' through the finish line: evidence from a composite scholarship for Ugandan students. Working paper.

Mawejje J. 2025. *Fiscal Vulnerabilities in Low-Income Countries: Evolution, Drivers, and Policies*. Washington, DC: World Bank. https://openknowledge.worldbank.org/entities/publication/873a9c96-fcba-4e42-aadd-670e64f13156 (20 May 2026, date last accessed).

McGuire F, Vijayasingham L, Vassall A *et al*. 2019. Financing intersectoral action for health: a systematic review of co-financing models. *Globalization and Health* **15**: 86.

Montenegro CE, Patrinos HA. 2023. A data set of comparable estimates of the private rate of return to schooling in the world, 1970–2014. *International Journal of Manpower* **44**: 1248–68.

Muralidharan K, Prakash N. 2017. Cycling to school: increasing secondary school enrolment for girls in India. *American Economic Journal: Applied Economics* **9**: 321–50.

Nemzoff C, Sweeney S, Baltussen R, Vassall A. 2025. Selecting cost-effectiveness methods for health benefits package design: a systematic approach. *International Journal of Health Policy and Management* **14**: 8562.

Norheim OF, Baltussen R, Johri M *et al.* 2014. Guidance on priority setting in health care (GPS-Health): the inclusion of equity criteria not captured by cost-effectiveness analysis. *Cost Effectiveness and Resource Allocation* **12**: 18.

Oehlsen E. 2024. Philanthropic cause prioritization. *Journal of Economic Perspectives* **38**: 63–82.

Ottersen T, Norheim OF. 2014. Making fair choices on the path to universal health coverage. *Bulletin of the World Health Organization* **92**: 389–389A.

Palermo T, Prencipe L, Kajula L *et al*. 2021. Effects of government-implemented cash plus model on violence experiences and perpetration among adolescents in Tanzania, 2018–2019. *American Journal of Public Health* **111**: 2227–38.

Patel V, Saxena S, Lund C *et al.* 2018. The Lancet Commission on global mental health and sustainable development. *The Lancet* **392**: 1553–98.

Petrosino A, Morgan C, Fronius TA *et al*. 2012. Interventions in developing nations for improving primary and secondary school enrollment of children: a systematic review. *Campbell Systematic Reviews* **8**: 1–192.

Psacharopoulos G, Patrinos HA. 2018. Returns to investment in education: a decennial review of the global literature. *Education Economics* **26**: 445–58.

Psaki SR, Chuang EK, Melnikas AJ *et al*. 2019. Causal effects of education on sexual and reproductive health in low and middle-income countries: a systematic review and meta-analysis. *SSM – Population Health* **8**: 100386.

Ramaiya A, Chandra-Mouli V, Both R *et al.* 2023. Assessing the health, social, educational and economic impact of the COVID-19 pandemic on adolescents in low- and middle-income countries: a rapid review of the literature. *Sexual and Reproductive Health Matters* **31**: 2187170.

Randall J, Garcia A. 2020. Let's go girls! Evaluating the effectiveness of tutoring and scholarships on primary school girls' attendance and academic performance in the Democratic Republic of the Congo. *FIRE: Forum for International Research in Education* **7**: 20–38.

Remme M, Vassall A, Lutz B, Luna J, Watts C. 2014. Financing structural interventions: going beyond HIV-only value for money assessments. *AIDS* **28**: 425–34.

Snilstveit B, Stevenson J, Phillips D *et al*. 2015. *Interventions for Improving Learning Outcomes and Access to Education in Low- and Middle-Income Countries: A Systematic Review*. 3ie Systematic Review Summary 24. International Initiative for Impact Evaluation.

Sperling GB, Winthrop R. 2015. *What Works in Girls' Education: Evidence for the World's Best Investment*. Brookings Institution Press.

Stenberg K, Hanssen O, Edejer TTT *et al.* 2017. Financing transformative health systems towards achievement of the health Sustainable Development Goals: a model for projected resource needs in 67 low-income and middle-income countries. *The Lancet Global Health* **5**: e875–87.

Turner HC, Archer RA, Downey LE *et al.* 2021. An introduction to the main types of economic evaluations used for informing priority setting and resource allocation in healthcare: key features, uses, and limitations. *Frontiers in Public Health* **9**: 722927.

UNESCO. 2021. *#HerEducationOurFuture: Keeping Girls in the Picture During and After the COVID-19 Crisis*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000375707 (20 May 2026, date last accessed).

UNESCO. 2023. Progress on girls' access to education: what the new UNESCO data reveals. UNESCO. https://www.unesco.org/en/articles/progress-girls-access-education-what-new-unesco-data-reveals (20 May 2026, date last accessed).

United Nations Population Fund East and Southern Africa Regional Office (UNFPA ESARO). n.d. Young people. https://esaro.unfpa.org/en/topics/young-people (20 May 2026, date last accessed).

United Nations Children's Fund (UNICEF). 2018. Child marriage: latest trends and future prospects. UNICEF Data. https://data.unicef.org/resources/child-marriage-latest-trends-and-future-prospects/ (20 May 2026, date last accessed).

United Nations Children's Fund (UNICEF). 2023. Child marriage. UNICEF. https://www.unicef.org/protection/child-marriage (20 May 2026, date last accessed).

United Nations Children's Fund (UNICEF). 2025. Child marriage. UNICEF. https://data.unicef.org/topic/child-protection/child-marriage/ (20 May 2026, date last accessed).

Verguet S, Kim JJ, Jamison DT. 2016. Extended cost-effectiveness analysis for health policy assessment: a tutorial. *PharmacoEconomics* **34**: 913–23.

Verguet S, Laxminarayan R, Jamison DT, Jamison JC. 2015. Universal public finance of tuberculosis treatment in India: an extended cost-effectiveness analysis. *Health Economics* **24**: 318–32.

Viner RM, Russell SJ, Saulle R *et al.* 2022. School closures during social lockdown and mental health, health behaviors, and well-being among children and adolescents during the first COVID-19 wave: a systematic review. *JAMA Pediatrics* **176**: 400–9.

Wodon Q, Male C, Nayihouba A *et al.* 2017. *Economic Impacts of Child Marriage: Global Synthesis Report*. Conference Edition. World Bank & International Center for Research on Women.

World Bank. 2013. *Beyond the Annual Budget: Global Experience with Medium-Term Expenditure Frameworks*. World Bank. https://documents1.worldbank.org/curated/en/354601468330959258/pdf/Beyond-the-annual-budget-global-experience-with-medium-term-expenditure-frameworks.pdf (20 May 2026, date last accessed).

World Bank. 2024. Empowering girls through education. World Bank. https://www.worldbank.org/en/news/feature/2024/10/10/empowering-girls-through-education (20 May 2026, date last accessed).

World Bank, UNESCO, UNICEF *et al.* 2022. *The State of Global Learning Poverty: 2022 Update*. World Bank. https://www.worldbank.org/en/topic/education/publication/state-of-global-learning-poverty (20 May 2026, date last accessed).

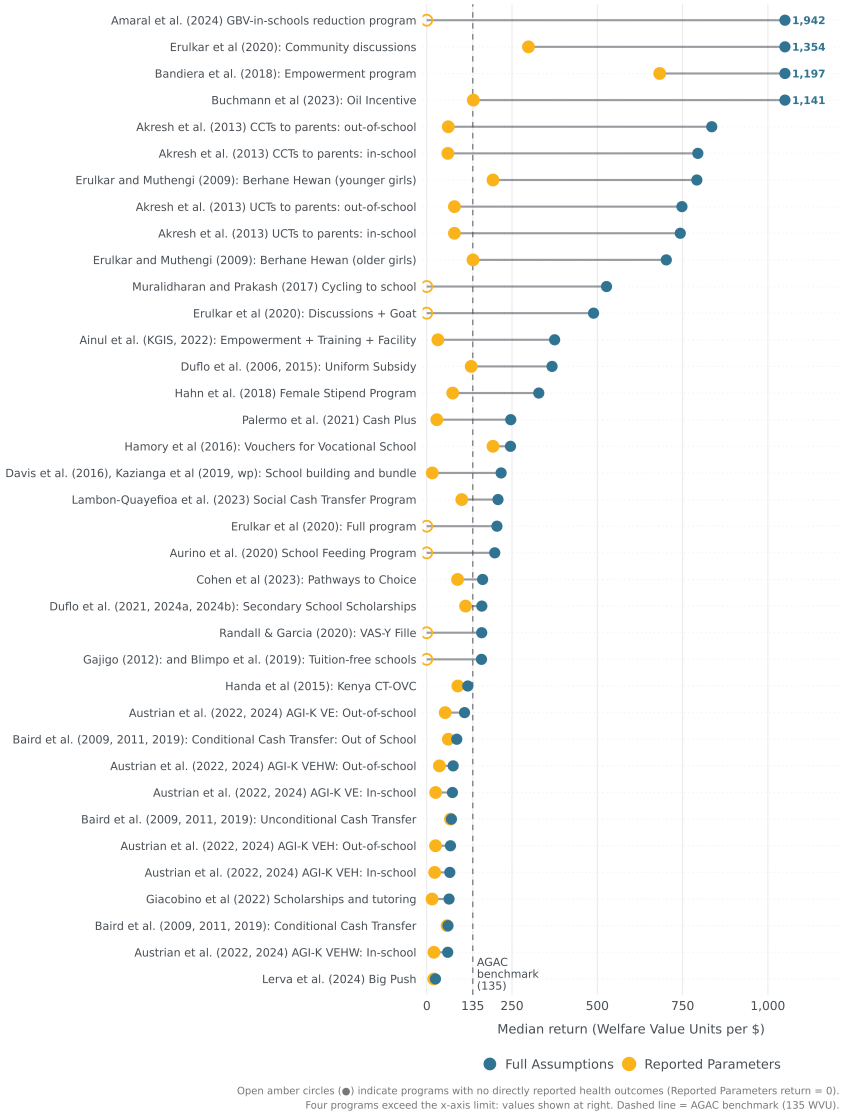
World Health Organization. 2020. *Global Health Estimates: Leading Causes of DALYs*. WHO. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-health-estimates-leading-causes-of-dalys (20 May 2026, date last accessed).

World Health Organization. 2021. *Global Progress Report on HIV, Viral Hepatitis and Sexually Transmitted Infections, 2021*. WHO. https://www.who.int/publications/i/item/9789240027077 (20 May 2026, date last accessed).

World Health Organization. 2024. Adolescent pregnancy. WHO Fact Sheet. https://www.who.int/news-room/fact-sheets/detail/adolescent-pregnancy (20 May 2026, date last accessed).

Appendix

Appendix 1: Reported Parameters vs. Full Assumptions models



Appendix Figure A1. Comparison of Reported Parameters and Full Assumptions models across all 37 intervention arms.