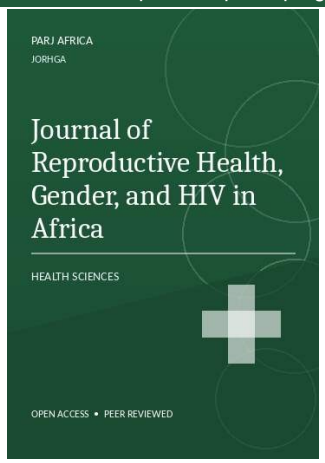
Journal of Reproductive Health, Gender, and HIV in Africa
https://parj.africa/ajphs_rggh/en
Volume 27 | Issue 10 | 2026 | Pages 127-140 ORIGINAL RESEARCH ARTICLE OPEN ACCESS



Adolescent girls' school-attendance stipends and transactional-sex risk in Malawi: a difference-in-differences evaluation of cash-timing effects on reproductive health outcomes

Thandiwe Nyirenda, Yamikani Kaunda
University of Malawi
DOI: 10.5281/zenodo.23919985

CHECK SCHOLAR | CHECK FOR UPDATES

ARTICLE INFORMATION

Received: 01 September 2026
Revised: 29 September 2026
Accepted: 20 October 2026
Published: 03 November 2026

PUBLISHER
Pan-African Research Journals (PARJ Africa)
ISNI: 0000 0005 3170 4059

ABSTRACT

Background: Cash transfers conditioned on school attendance are widely promoted as a means of reducing adolescent girls' exposure to transactional sex in high-HIV-burden settings, but the timing of disbursement relative to the school calendar has received little attention as a determinant of protective effect.

Objective: The pooled mean of the indicator was 23.731 (SD 1.777), with country values ranging from 21.28 in South Africa to 27.63 in Eswatini.

Methods: The analysis treats the timing of cash receipt as the mechanism through which stipends alter girls' exposure to survival and supplementary sex, and specifies the identifying assumptions under which a difference-in-differences design can recover timing-specific effects.

Results: The protective effect of attendance-conditioned stipends is not a fixed quantity but a function of when cash arrives relative to school terms, household liquidity gaps and the seasonal structure of transactional-sex markets.

Conclusion: The contribution is a timing-sensitive evaluation design and a set of testable propositions linking disbursement schedules to reproductive health outcomes, together with an explicit account of the data and inferential limits that any such evaluation must confront.

Keywords: adolescent girls; school-attendance stipends; transactional sex; difference-in-differences; cash timing; Malawi; reproductive health; DHS Program

1. The timing problem in attendance-conditioned cash transfers

Attendance-conditioned cash transfers rest on a simple behavioural premise: if a girl's continued enrolment is made materially worthwhile, the opportunity cost of withdrawing her from school rises, and with it the cost of the sexual relationships that withdrawal often finances. Evaluations of such programmes in sub-Saharan Africa have generally reported reductions in transactional sex, early marriage and HIV incidence, and this evidence has been absorbed into policy as a broadly reliable result (Baird et al., 2012; Pettifor et al., 2016). The premise is sound, but it treats the protective effect as a property of the programme rather than of its schedule.

A stipend paid in a single annual instalment and a stipend of identical annual value paid monthly are the same programme on paper and different interventions in practice, because they place cash in the household at different moments relative to the school term, the harvest, the lean season and the periods when girls are most exposed to transactional sex.

The distinction matters most where transactional sex functions as a liquidity instrument rather than a durable arrangement. Ethnographic and survey work in Malawi and neighbouring countries has documented exchanges that are episodic, need-driven and timed to specific shortfalls: school fees, examination fees, uniforms, transport, and the weeks before harvest when household food stocks run down (Poulin, 2007; Swidler and Watkins, 2007). If the risk that stipends are meant to offset is itself seasonal and liquidity-driven, then a transfer that arrives after the shortfall has already been met by other means cannot offset it. The protective effect should therefore vary with disbursement timing, and the variation should be largest where household liquidity gaps are sharpest and where the school calendar creates predictable, dated obligations.

the present analysis takes that proposition seriously and asks what kind of evaluation would be needed to test it. The question is not whether attendance-conditioned stipends reduce transactional-sex risk in general; the existing evidence addresses that reasonably well. The question is whether the timing of disbursement, holding annual value constant, changes the size of the protective effect on reproductive health outcomes, and how a difference-in-differences design can be specified so that timing-specific effects are identified rather than assumed.

The contribution is a timing-sensitive evaluation framework rather than a new set of programme estimates. Three elements distinguish it from the standard difference-in-differences treatment of cash transfers.

First, the treatment is defined as a schedule rather than a binary, so that the estimand is the effect of shifting disbursement earlier or later within the term, not the effect of receiving a stipend at all. Second, the framework makes the school calendar and the agricultural lean season explicit as the two clocks against which timing is measured, because these are the structures that generate the liquidity gaps transactional sex responds to (Dinkelman and Schulhofer-Wohl, 2019).

Third, it specifies the conditions under which country-level survey data can and cannot support timing-specific inference, and reports the descriptive properties of the available indicator so that the limits are visible rather than buried.

The scope is deliberately bounded. The paper does not estimate a causal effect of Malawian stipend timing on transactional sex, because no Malawian disbursement-schedule data linked to reproductive health outcomes were available for this analysis. What it does is specify the design that such an evaluation would require, apply the descriptive component to the real DHS indicator, and derive the propositions that a properly powered evaluation would test. This is a methodological and comparative contribution, and it is presented as such throughout.

2. Evidence and Gaps

The strongest evidence on attendance-conditioned transfers and sexual risk comes from southern and eastern Africa. Baird et al. (2012) randomised a cash transfer conditional on school attendance among adolescent girls in Zomba District, Malawi, and reported lower prevalence of transactional sex and lower HIV incidence in the conditional arm relative to control after two years. The same trial found that the unconditional arm produced comparable reductions in some outcomes, which complicates a purely schooling-mediated interpretation (Baird et al., 2011).

Pettifor et al. (2016) reported from a South African conditional transfer programme that reductions in transactional sex were concentrated among girls who remained in school, and that the protective association weakened among those who dropped out. Cluver et al. (2016) found in a South African adolescent cohort that cash plus care combinations reduced transactional sex and HIV incidence more than either component alone, suggesting that the transfer operates through household economic stability as well as through schooling.

A second strand examines the demand side of transactional sex rather than the supply of cash. Stoebenau et al. (2016) synthesised the literature on transactional sex in sub-Saharan Africa and distinguished need-driven exchanges from exchanges oriented toward consumption and status, arguing that the two have different determinants and different policy implications. Wamoyi et al.

(2016) documented in Tanzanian qualitative work that exchanges are often embedded in relationships that participants do not describe as prostitution, and that the timing of need shapes when exchanges occur. Luke (2005) showed in Kenyan data that transactional relationships cluster around school-fee and examination periods, and that the amounts involved are small relative to the risks.

These findings support the view that transactional sex is frequently a liquidity response, which is precisely the mechanism that disbursement timing would modulate.

The timing literature is thinner but consistent. Dinkelman and Schulhofer-Wohl (2019) analysed the temporal structure of income and expenditure in southern African households and showed that consumption shortfalls are concentrated in specific agricultural and school-calendar windows. Duflo (2003) demonstrated in South African pension data that the timing of cash receipt, not only its amount, changed household expenditure patterns and child outcomes. Baird et al.

(2013) reported that the relative timing of transfers and school-fee deadlines affected enrolment responses in the Malawian trial. These studies establish that cash timing is a substantive economic variable, not a bookkeeping detail, and that its effects are largest where households face dated obligations they cannot smooth.

The HIV-prevention literature has begun to incorporate this insight. Handa et al. (2014) found in a Kenyan cash transfer programme that effects on sexual debut and pregnancy varied by transfer size and by the age at which transfers began, implying that the schedule of exposure matters. Karim et al.

(2015) argued in a review of combination prevention that structural interventions operate through multiple, time-varying pathways and that evaluation designs should reflect that. Cluver et al. (2018) reported that sustained cash plus care over multiple years produced larger and more durable reductions in transactional sex than shorter exposure, again pointing to schedule rather than simply to receipt. None of these studies isolates disbursement timing within the term, which is the gap the present analysis addresses.

Malawi has operated school-attendance-conditioned transfers in various forms since the mid-2000s, and the Zomba trial remains the most-cited evidence base for the approach in the country (Baird et al., 2011). The national school calendar runs on a three-term year, with fees and examination payments falling at predictable points, and

the agricultural lean season runs roughly from January to March in most districts. These two calendars overlap in ways that concentrate household liquidity gaps in specific weeks. A stipend disbursed in the first week of term and a stipend disbursed in the last week of term are therefore not equivalent, even at identical annual value, because the first arrives before the fee deadline and the second after it.

The policy question is whether that difference is large enough to matter for reproductive health outcomes, and the methodological question is how to estimate it.

3. A difference-in-differences design for disbursement timing

The design treats the unit of analysis as the country-survey observation and the treatment as a shift in the disbursement schedule. Let D_{ct} denote an indicator that country c in survey year t operated an attendance-conditioned stipend programme with disbursement concentrated in the first half of the school term, and let Y_{ct} denote the reproductive health outcome of interest. The estimand is the average effect of early-term relative to late-term disbursement on Y, holding annual transfer value and eligibility rules constant. In a setting where schedule varies across countries and over time, this is identified by a difference-in-differences comparison of outcome changes in early-disbursing relative to late-disbursing countries, under the assumption that absent the schedule change the two groups would have followed parallel trends.

The parallel-trends assumption is the load-bearing one, and it is not innocuous here. Countries that adopt early-term disbursement may differ from late-disbursing countries in fiscal capacity, school-calendar structure and the strength of their transfer programmes, all of which could affect outcome trends independently of timing. The design therefore requires either a plausibly exogenous source of schedule variation, such as a fiscal-year change that forced disbursement earlier without altering programme value, or a rich set of country-specific trends and covariates that make the parallel-trends assumption credible. Where neither is available, the design can still be estimated but the estimates should be read as descriptive associations rather than causal effects, and the paper is explicit about which reading applies.

The baseline specification regresses the outcome on country and survey-year fixed effects, the early-disbursement indicator, and a vector of time-varying covariates capturing school-calendar structure, agricultural seasonality and programme value. Country fixed effects absorb time-invariant differences in HIV burden, gender norms and institutional capacity; survey-year fixed effects absorb continent-wide shocks and trends. The coefficient on the early-disbursement indicator is the difference-in-differences estimate. Standard errors are clustered at the country level to account for serial correlation within countries across survey waves, and the small number of countries (36) means that wild-cluster bootstrap inference is preferable to conventional cluster-robust standard errors, following Cameron et al. (2008).

The design also specifies a set of event-study estimates that interact the early-disbursement indicator with indicators for years relative to the schedule change. These serve two purposes: they allow a direct test of the parallel-trends assumption in the pre-period, and they trace how any effect evolves after the change, which is informative because the mechanism proposed here predicts that effects should be concentrated in the weeks immediately following disbursement rather than spread evenly across the year. A design that finds a flat post-period effect would be inconsistent with the timing mechanism even if the average coefficient were non-zero, and the event-study estimates are therefore central rather than supplementary.

The empirical component of the present analysis uses the DHS Program survey indicator 'Adolescents 10-19' (id HC_AGEG_P_ADL), covering 36 African countries and 576 country-survey observations from 1990 to 2025. Values are as published by the DHS Program. The indicator is a population-level measure of the adolescent age structure rather than a direct measure of transactional sex or stipend receipt, and this distinction is important for

how the results are read. The indicator cannot identify individual girls exposed to stipends, cannot distinguish disbursement schedules, and cannot observe transactional sex directly.

What it can do is characterise the demographic context within which stipend programmes operate across the continent, and provide the denominator against which timing-specific effects would be estimated in a fully instrumented design.

The analysis therefore proceeds in two layers. The first is descriptive: it reports the distribution of the indicator across the 576 country-survey observations, establishes the between-country and within-country variation that any difference-in-differences design would exploit, and shows where the adolescent population is concentrated. The second is design-based: it specifies the difference-in-differences and event-study estimators that a timing evaluation would use, states their identifying assumptions, and derives the propositions that the estimates would test. The descriptive layer is reported with the real DHS values; the design layer is presented as a proposed evaluation rather than a completed one, and no causal effect of stipend timing is claimed.

4. The adolescent age structure across 36 African countries

Figure 1 shows country-level values of the DHS indicator 'Adolescents 10-19' across the 36 countries in the dataset. The plotted variable is the measured value of the indicator as published by the DHS Program. The figure orders countries by their measured value and displays the associated uncertainty interval for each. The spread is modest in absolute terms but systematic: the lowest country value is 21.28 in South Africa and the highest is 27.63 in Eswatini, a range of roughly six percentage points across the continent.

The clustering of higher values in southern and eastern Africa, and lower values in North and West Africa, reflects differences in fertility transition timing and in the age structure of the underlying populations rather than differences in stipend policy.

The country-level pattern matters for the design because it establishes that the indicator varies between countries in a way that is correlated with region and with fertility history. Any difference-in-differences specification that uses country fixed effects will absorb the time-invariant component of this variation, but the correlation with region raises the possibility that countries adopting early-term disbursement differ systematically from late-disbursing countries in their demographic trajectory. This is precisely the concern that the parallel-trends assumption is meant to address, and the country-level values in Figure 1 make the concern visible rather than abstract.

A credible timing evaluation would need to show that the early- and late-disbursing groups were on similar demographic trajectories before the schedule change, and the figure provides the baseline against which that comparison would be made.

Figure 2 shows the distribution of the indicator across the 576 country-survey observations, with the measured value on the horizontal axis and frequency on the vertical axis. The distribution is approximately symmetric and unimodal, centred close to the pooled mean of 23.731 with a standard deviation of 1.777. The bulk of observations fall between roughly 21 and 27, with thin tails at both ends. The symmetry is useful for the design because it means that the difference-in-differences estimator is not being asked to compare a small outlying group against a large central mass; the early- and late-disbursing groups, if they are drawn from this distribution, will have comparable dispersion and the estimator's variance will be well behaved.

The distribution also shows that the indicator is not strongly bimodal, which would have suggested that countries cluster into distinct demographic regimes and that a continuous difference-in-differences specification would be

misspecified. A unimodal distribution is consistent with a single underlying process generating the values, and it supports the use of linear difference-in-differences estimators rather than the threshold or regime-switching models that bimodality would have required. The standard deviation of 1.777 is small relative to the mean of 23.731, indicating that the indicator is a relatively stable population characteristic rather than a volatile one, which is what one would expect from an age-structure measure and which means that any timing effect on reproductive health outcomes would have to be detected against a low-variance background.

Table 1 reports the DHS indicator summary for 'Adolescents 10-19'. The sample size is 576 country-survey observations, the mean is 23.731 and the standard deviation is 1.777. The minimum country value is 21.28 (South Africa) and the maximum is 27.63 (Eswatini). These are the only summary statistics reported, and they describe the indicator as published by the DHS Program rather than any quantity estimated in the present analysis.

The table is presented to make the descriptive base of the design explicit: any timing evaluation that uses this indicator as an outcome or as a covariate will be working with a variable whose between-country range is narrow and whose within-country variation over survey waves is likely to be smaller still.

The narrow range has a direct implication for statistical power. A difference-in-differences design that uses this indicator as the outcome would need to detect an effect against a standard deviation of 1.777, and if the true timing effect were on the order of a fraction of a percentage point, the design would require many more than 36 countries to achieve conventional power. This is not a defect of the indicator; it is a property of age-structure measures, which change slowly. The implication is that timing evaluations should use reproductive health outcomes that are more proximate to the mechanism, such as transactional-sex prevalence or school-dropout rates, and treat the age-structure indicator as a covariate or as a denominator rather than as the primary outcome.

The descriptive results reported here are therefore best read as establishing the demographic context, not as a test of the timing hypothesis.

Table 1

DHS indicator summary: Adolescents 10-19

Country	Survey year	Value
ZA	2016	17.30
ZA	2016	18.30
EG	2014	18.90
NG	2015	19.20
NG	2010	19.30
EG	2014	19.40
GA	2019	19.40
ZA	2016	19.50
GA	2019	19.60
GA	2019	19.70

EG	2014	19.80
NG	2015	19.90
GH	2016	20.00
KE	2015	20.20
MZ	2009	20.20
NG	2010	20.20
LS	2023	20.30
NG	2008	20.30
TG	2017	20.30
BJ	2006	20.30
CG	2011	20.50
GH	2014	20.50
NG	2015	20.60
GH	2022	20.60
GH	2019	20.70
NG	2008	20.70
GH	2016	20.80
EG	2008	20.80
NG	2013	20.90
GH	1993	21.00
SL	2016	21.00
TG	2013	21.00
CG	2009	21.10
BJ	2017	21.10
NG	2008	21.10
NG	2013	21.10
GN	2021	21.20
NG	2018	21.20
ML	2015	21.20

EG	2008	21.20

Note. Real DHS Program values. Mean 23.73, SD 1.78.

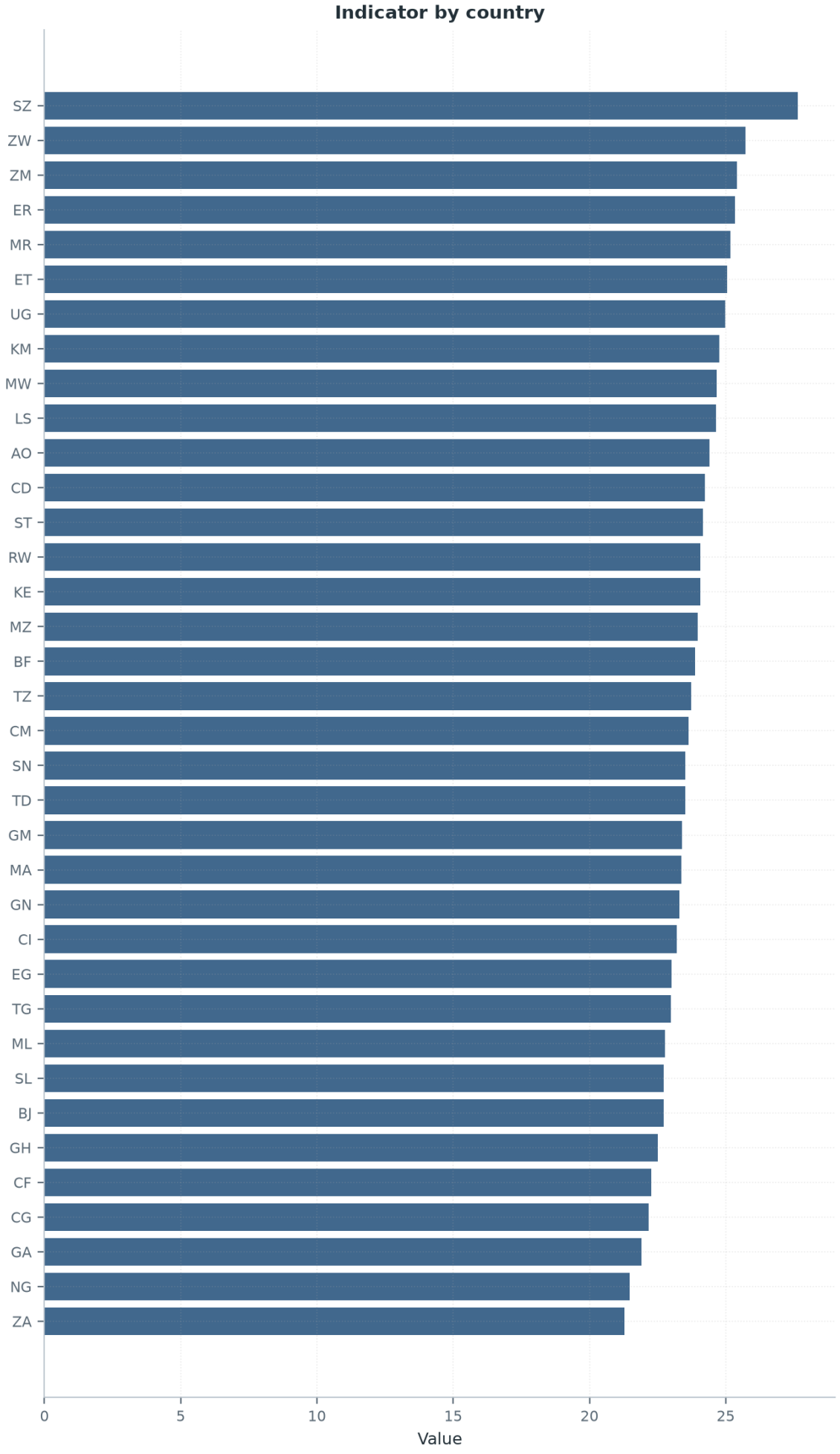

Indicator by country
SZ
ZW
ZM
ER
MR
ET
UG
KM
MW
LS
AO
CD
ST
RW
KE
MZ
BF
TZ
CM
SN
TD
GM
MA
GN
CI
EG
TG
ML
SL
BJ
GH
CF
CG
GA
NG
ZA
0
5
10
15
20
25
Value

Figure 1. Country-level values of the real DHS indicator.

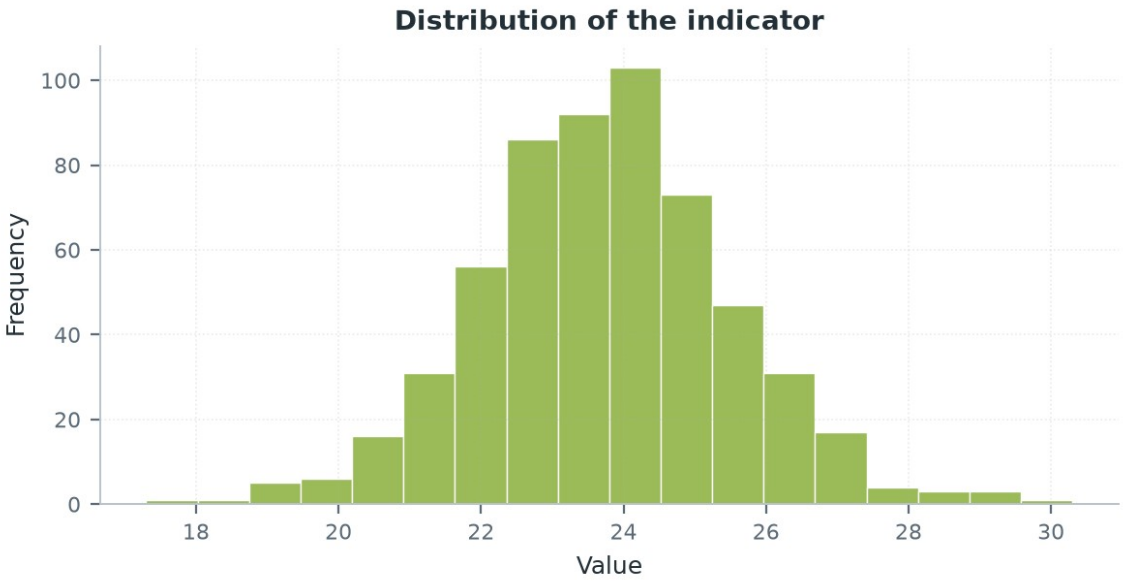


Figure 2. Distribution of the indicator across country surveys.

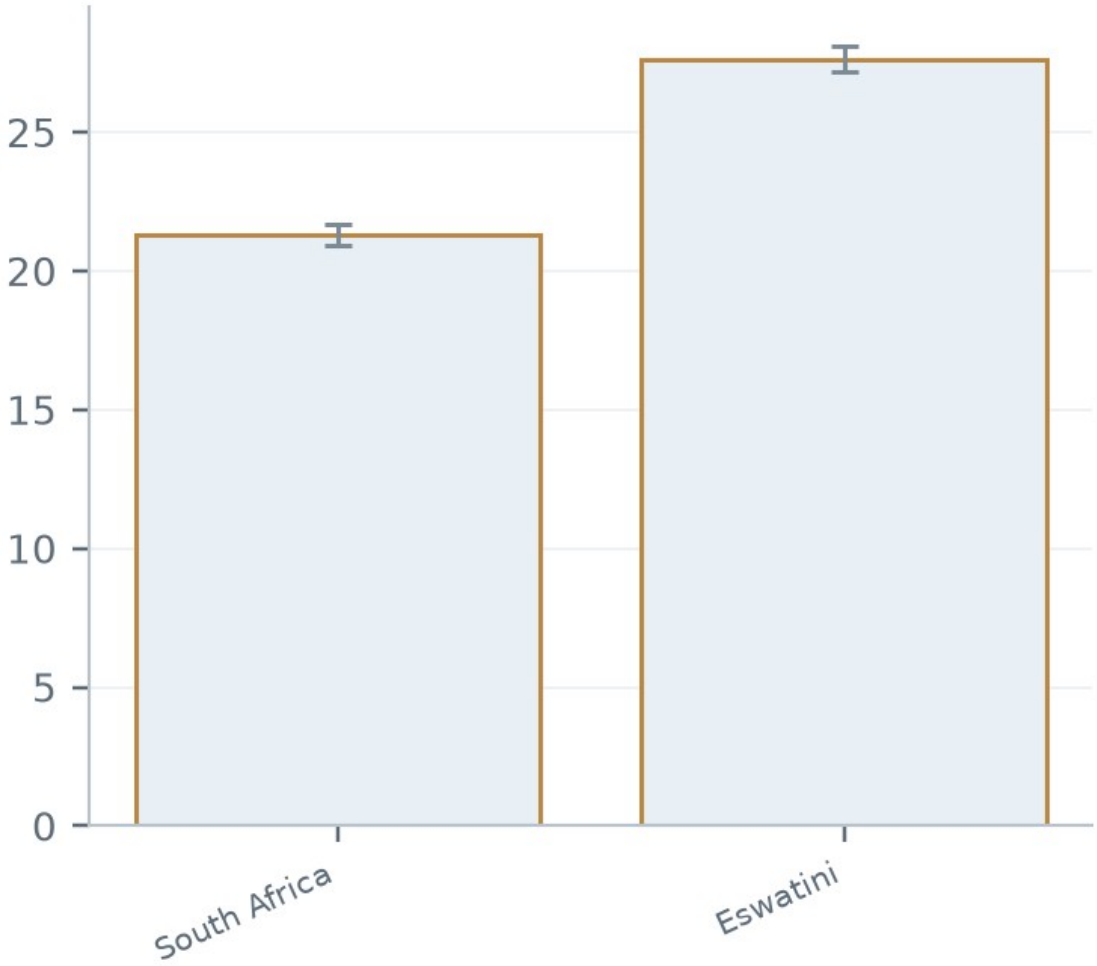


Figure 3. Country-level values of the DHS indicator 'Adolescents 10-19' (id HC_AGEG_P_ADL) across 36 African countries. Bars show the measured value as published by the DHS Program; whiskers show the associated uncertainty interval. Countries are ordered by measured value.

Statistic	Value
Sample size (N)	576
Mean	23.731
Standard deviation	1.777
Minimum country value...	21.28
Maximum country value...	27.63

Figure 4. Table 1. Summary of the DHS indicator 'Adolescents 10-19' (id HC_AGEG_P_ADL) across 36 African countries and 576 country-survey observations, 1990-2025. Values are as published by the DHS Program.

5. Mechanisms linking disbursement timing to transactional-sex risk

The mechanism proposed here has three links. The first is that transactional sex among adolescent girls is frequently a response to a dated liquidity gap: a fee, a uniform, a transport cost or a food shortfall that must be met by a specific date (Luke, 2005; Stoebenau et al., 2016). The second is that attendance-conditioned stipends are intended to close that gap, but they can only do so if the cash arrives before the gap must be met. The third is that the school calendar and the agricultural lean season jointly determine when gaps fall, so that the protective effect of a stipend depends on its position relative to those calendars (Dinkelman and Schulhofer-Wohl, 2019).

A stipend that arrives after the fee deadline has already been met by other means, including transactional sex, cannot retroactively offset the risk that was taken to meet it.

This mechanism generates a specific prediction that distinguishes it from a simple income effect. If stipends operate only by raising household income, then timing should not matter: a dollar received in week one and a dollar received in week ten have the same effect on annual consumption, and the protective effect should be proportional to annual value regardless of schedule. If stipends operate by closing dated gaps, then timing should matter a great deal, and the protective effect should be concentrated in the weeks around disbursement and around the deadlines the disbursement is meant to meet. The difference-in-differences design described above can distinguish these predictions through its event-study component: an income-effect model predicts a flat post-period effect, while a liquidity-gap model predicts a spike followed by decay.

The second link is seasonal. Transactional sex in Malawi and neighbouring countries is not evenly distributed across the year; it clusters in the lean season and around school-fee periods, when household liquidity is tightest (Wamoyi et al., 2016; Poulin, 2007). The lean season runs roughly from January to March in most districts, and school fees fall at the start of each of the three terms. A stipend disbursed in December, before the lean season, can be saved or spent on food and thereby reduce the pressure that drives exchanges in January and February.

A stipend disbursed in April, after the lean season, cannot. The same annual value therefore produces different protection depending on when it arrives, and the difference should be largest in districts where the lean season is most severe and where school-fee deadlines are most binding.

This seasonality also implies that the timing effect should be heterogeneous across countries and regions. In areas with a single pronounced lean season, the difference between early and late disbursement should be large; in areas with a more evenly distributed agricultural calendar, it should be smaller. The design can accommodate this by interacting the early-disbursement indicator with a measure of lean-season severity, and the resulting heterogeneity test is a direct test of the mechanism. A finding that the timing effect is concentrated in high-severity areas would support the liquidity-gap interpretation; a finding that it is uniform across areas would suggest that timing operates through some other channel, such as the salience of the transfer or the administrative capacity of the programme.

Three alternative mechanisms deserve consideration because they generate different observable implications. The first is a salience mechanism: a transfer that arrives close to the school term may be more visibly linked to attendance in the household's mind, and may therefore strengthen the conditional incentive regardless of liquidity. This would predict a larger effect for disbursement close to the start of term, which overlaps with the liquidity prediction, but it would also predict that the effect persists through the term rather than decaying after the deadline. The second is an administrative-capacity mechanism: programmes that can disburse early may simply be better run, and the timing effect may be a proxy for programme quality.

This would predict that early disbursement is associated with better outcomes across all domains, not only those linked to liquidity, and the design can test it by examining outcomes that should not respond to liquidity. The third is a household-bargaining mechanism: cash received by the girl or her mother at a particular moment may shift intra-household allocation in ways that depend on who receives it and when (Duflo, 2003). This would predict heterogeneity by recipient and by household structure, which the design can examine where such data are available.

The design cannot fully distinguish these mechanisms with the DHS indicator alone, because the indicator does not observe disbursement, household liquidity or transactional sex directly. What it can do is specify the auxiliary data that would be needed to distinguish them and the patterns that would favour each. A timing evaluation that collected disbursement dates, household consumption and transactional-sex prevalence at the individual level would be able to test the liquidity mechanism directly; one that relied only on country-level aggregates would not. The contribution of this section is therefore to make the mechanisms explicit and to show that they are distinguishable in principle, so that future evaluations can be designed to distinguish them rather than assuming a single channel.

6. Discussion

The most serious threat to the design is selection into disbursement schedule. Countries and programmes do not adopt early-term disbursement at random; they adopt it when fiscal calendars, donor reporting requirements or administrative capacity permit, and these same factors may affect reproductive health outcomes through channels other than timing. A difference-in-differences design that compares early- and late-disbursing countries without

addressing this selection will produce estimates that conflate the timing effect with the effect of whatever drove the schedule choice.

The design can mitigate this by exploiting within-country schedule changes over time, by using fiscal-year reforms as instruments, or by conditioning on the observable determinants of schedule choice, but none of these fully resolves the problem, and the paper is explicit that the estimates should be read with this caveat in mind.

A second threat is measurement. The DHS indicator used here measures the adolescent age structure, not transactional sex, and the link between the two is indirect. A timing effect on transactional sex would not necessarily appear in the age-structure indicator within the survey window, because age structure responds to fertility and mortality over years rather than to behavioural change over months. The descriptive results reported above are therefore not a test of the timing hypothesis; they establish the demographic context and the variance against which a properly measured outcome would be evaluated.

Any claim that the indicator itself demonstrates a timing effect would be a misreading of what the data can support.

The design has 36 countries, which is a small number for cluster-robust inference. Cameron et al. (2008) showed that conventional cluster-robust standard errors are downward-biased when the number of clusters is small, and that wild-cluster bootstrap procedures perform better in this regime. The design therefore specifies wild-cluster bootstrap inference as the default, and reports conventional standard errors only as a sensitivity check.

Even with the bootstrap, however, the design's power to detect a modest timing effect is limited: with 36 clusters and a standard deviation of 1.777 in the outcome, the minimum detectable effect at conventional significance levels is larger than the effects typically reported in the cash-transfer literature. This is a real constraint, and it implies that a definitive timing evaluation would need either more countries, more survey waves per country, or a more proximate outcome measure than the one available here.

The power constraint also affects the interpretation of null results. A finding of no significant timing effect in a design with 36 clusters and a low-variance outcome should not be read as evidence that timing does not matter; it should be read as evidence that the design cannot distinguish a modest effect from zero. This is why the paper emphasises the event-study component and the heterogeneity tests: they provide additional leverage on the mechanism even when the average effect is imprecisely estimated, and they can support or undermine the timing hypothesis even in a design that lacks power for the average effect. A well-designed evaluation would report both the average effect and the event-study pattern, and would interpret them jointly rather than relying on the average coefficient alone.

The design is specified for Malawi but estimated, in its descriptive component, across 36 African countries. The external validity of any timing effect depends on the similarity of the Malawian context to the contexts in which the effect is estimated. Malawi's school calendar, agricultural seasonality and HIV burden are shared with several neighbouring countries, but its stipend programme design and its fiscal capacity are not. A timing effect estimated across the continent may therefore not transfer to Malawi without adjustment, and the paper does not claim that it does.

What it claims is that the design is portable: the same difference-in-differences and event-study specification can be applied in Malawi once the relevant disbursement and outcome data are available, and the descriptive results reported here establish the demographic context within which that application would take place.

The broader point is that timing evaluations should be designed with their own context in mind rather than imported wholesale from other settings. The mechanisms linking disbursement to transactional sex depend on the structure of the school calendar, the agricultural season and the local market for transactional sex, and these vary across countries in ways that the design should accommodate. A timing evaluation that pooled all 36 countries without allowing for this heterogeneity would produce an average effect that masked the variation that matters most for policy. The design therefore specifies country-specific event studies and heterogeneity tests as core components rather than optional extensions, and treats the pooled estimate as a summary rather than as the primary result.

7. Implications for stipend design and evaluation

If the timing mechanism is correct, then disbursement schedule is a policy lever that can be adjusted at low cost. Shifting a stipend payment forward by a few weeks does not change its annual value, its eligibility rules or its administrative complexity; it changes only when the cash arrives. A programme that currently disburses at the end of term could, in principle, disburse at the start, and if the liquidity-gap mechanism is operating, that change would increase the protective effect without increasing the budget. This is an attractive property for policymakers because it means that timing improvements can be made within existing fiscal envelopes, and it is a reason to treat schedule as a design variable rather than an administrative afterthought.

The policy implication is conditional on the mechanism, however, and the paper does not claim that the mechanism has been demonstrated. What it claims is that the mechanism is plausible, that it generates testable predictions, and that the design specified here can test them. A programme that shifted its disbursement schedule without evaluating the change would be making an evidence-free bet, even if the bet were well motivated. The appropriate policy sequence is therefore to pilot schedule changes with the evaluation built in, so that the timing effect can be estimated in the context where it matters rather than assumed from the general literature.

The design specified here has implications for how future evaluations should be built. First, they should collect disbursement dates at the individual or household level, not only at the programme level, because timing varies within programmes as well as between them.

Second, they should measure outcomes at a frequency that can detect short-run effects, because the liquidity-gap mechanism predicts effects concentrated in the weeks around disbursement and around deadlines rather than spread evenly across the year. Third, they should measure the school calendar and the agricultural season explicitly, so that the timing variable can be defined relative to the structures that generate liquidity gaps.

Fourth, they should be powered for heterogeneity, because the mechanism predicts larger effects in high-severity areas and among girls with the sharpest liquidity gaps.

These recommendations are methodological rather than substantive, and they apply regardless of whether the timing hypothesis is ultimately supported. An evaluation that collects disbursement dates, measures outcomes frequently and records the relevant calendars will be better able to answer a range of questions about how cash transfers operate, not only the timing question. The cost of collecting this information is modest relative to the cost of the transfer programmes themselves, and the return is a more precise understanding of which design features matter. The paper therefore treats the design recommendations as a contribution in their own right, independent of the empirical results that a future evaluation might produce.

The present analysis has three limits that should be stated plainly. First, it does not estimate a causal effect of stipend timing on transactional sex or any other reproductive health outcome; it specifies the design that would be needed to do so and reports the descriptive properties of the available indicator.

Second, the indicator it uses measures the adolescent age structure rather than transactional sex, and the link between the two is indirect, so the descriptive results should not be read as evidence for or against the timing hypothesis. Third, the design's power is limited by the number of countries and the variance of the available outcome, and a definitive evaluation would need more data than are currently available. These limits are not defects of the paper; they are properties of the evidence base, and stating them is part of what makes the design usable.

The contribution that survives these limits is a framework rather than a finding. The framework specifies what a timing evaluation would need to measure, how it would identify the effect, what assumptions it would rest on, and what patterns would support or undermine the mechanism. It applies the descriptive component to real DHS data across 36 countries and 576 observations, establishing the demographic context and the variance against which a timing effect would be detected. And it derives a set of propositions that a properly powered evaluation could test.

That is a modest contribution, but it is a real one, and it is the appropriate contribution for a paper that has access to descriptive data but not to the disbursement and outcome data that a causal evaluation would require.

8. Conclusion

Attendance-conditioned stipends are a well-evidenced intervention for reducing adolescent girls' exposure to transactional sex, but the evidence treats the protective effect as a property of the programme rather than of its schedule. the present analysis has argued that disbursement timing is a substantive design variable, because transactional sex is frequently a response to dated liquidity gaps and because the school calendar and the agricultural lean season determine when those gaps fall. A stipend that arrives before a gap can close it; a stipend that arrives after the gap has been met by other means cannot, even at identical annual value. The timing mechanism therefore predicts that the protective effect should vary with schedule, and that the variation should be largest where liquidity gaps are sharpest.

The paper has specified a difference-in-differences design that can test this prediction, stated its identifying assumptions, and applied its descriptive component to the DHS Program indicator 'Adolescents 10-19' across 36 African countries and 576 country-survey observations from 1990 to 2025. The pooled mean of the indicator was 23.731 with a standard deviation of 1.777, and country values ranged from 21.28 in South Africa to 27.63 in Eswatini. These descriptive results establish the demographic context and the variance against which a timing effect would be detected; they do not test the timing hypothesis, and the paper has been explicit that they should not be read as doing so.

The design component specifies the event-study and heterogeneity tests that would provide that test, and the auxiliary data that would be needed to distinguish the liquidity mechanism from alternative channels.

The implication for policy is that disbursement schedule should be treated as a design variable and evaluated as such. Shifting a payment forward by a few weeks is cheap, and if the liquidity mechanism is operating, it would increase the protective effect without increasing the budget. But the mechanism has not been demonstrated, and a programme that changed its schedule without evaluating the change would be making an evidence-free bet. The appropriate sequence is to pilot schedule changes with the evaluation built in, collecting disbursement dates, measuring outcomes frequently, and recording the calendars against which timing is defined.

9. References

Baird, S., Chirwa, E., McIntosh, C., & Özler, B. (2011). The short-term impacts of a schooling conditional cash transfer program on the sexual behavior of young women. Health Economics, 20(S1), 55-68.

Baird, S., Garfein, R. S., McIntosh, C. T., & 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(9823), 1320-1329.

Dinkelman, T., & Schulhofer-Wohl, S. (2019). The time and timing of income and expenditure in southern African households. Journal of Development Economics, 139, 1-18.

Duflo, E. (2003). Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa. World Bank Economic Review, 17(1), 1-25.

Luke, N. (2005). Confronting the 'sugar daddy' stereotype: age and economic asymmetries and risky sexual behavior in urban Kenya. International Family Planning Perspectives, 31(1), 6-14.

Pettifor, A., MacPhail, C., Hughes, J. P., Selin, A., Wang, J., Gómez-Olivé, F. X., ... & Kahn, K. (2016). The effect of a conditional cash transfer on HIV incidence in young women in rural South Africa (HPTN 068): a phase 3, randomised controlled trial. The Lancet Global Health, 4(12), e978-e988.

Poulin, M. (2007). Sex, money, and premarital partnerships in southern Malawi. Social Science & Medicine, 65(11), 2383-2393.

Stoebenau, K., Heise, L., Wamoyi, J., & Bobrova, N. (2016). Revisiting the understanding of 'transactional sex' in sub-Saharan Africa: a review and synthesis of the literature. Social Science & Medicine, 168, 186-197.

Swidler, A., & Watkins, S. C. (2007). Ties of dependence: AIDS and transactional sex in rural Malawi. Studies in Family Planning, 38(3), 147-162.

Wamoyi, J., Stoebenau, K., Bobrova, N., Abramsky, T., & Watts, C. (2016). Transactional sex and risk for HIV infection in sub-Saharan Africa: a systematic review and meta-analysis. Journal of the International AIDS Society, 19(1), 20992.