Center for Evaluation and Development, C4ED
Alexandra Avdeenko and Markus Frölich

Improving Adolescent Lives
in Pakistan (2015-2019)
An Impact Evaluation

December 2021

Improving Adolescent Lives in Pakistan (2015-2019)
An Impact Evaluation

This evaluation is conducted

at

the Center for Evaluation and Development

by

Alexandra Avdeenko and Markus Frölich
(principal investigators and authors).

This version is from:

January 24, 2022.

Executive Summary

According to 2020 data, every fifth person of Pakistan's 220 million plus population is in the 10-19 age group. These adolescents face several of the most demanding challenges in South Asia. The Pakistan Demographic and Health Survey 2017-18 showed that, on average, 3.6% of 20- to 24-year-old women were married by the age of 15, and 18.3% by the age of 18 (National Institute of Population Studies 2018). With only every third woman having access to contraceptives in 2019, early childbirth follows early marriages. In 2019, on average, 38 out of 1,000 girls of age 15-19 were reported to have given birth. This is close to the average in South Asia (not considering high-income countries), with only Afghanistan, Bangladesh and Nepal reporting higher rates (World Bank Open Data 2021).

One of the implications of early marriages and pregnancies is early school drop-out. Compulsory education is 10 years in Pakistan, yet only 52.4% of boys and 46.3% of girls completed lower secondary education in the relevant age group in 2019. Only Afghanistan displays lower levels in South Asia and has a worse gender gap. Not surprisingly, eventually only every fifth woman (20.1%) participated in the labor force in 2019, which is the lowest rate in South Asia (Pakistan Bureau of Statistics 2020).

Against this background, the United Nations Children's Fund (UNICEF) program "Improving Adolescent Lives in Pakistan (IALP)" aimed to reduce child marriage rates, decrease the number of early pregnancies, and increase school attendance. The program aimed to empower adolescents to become "agents of change" for themselves and their communities, while voicing their rights and needs. The IALP program operated in villages of two districts each of Punjab and Sindh provinces: Bahawalpur, Rahim Yar Khan, Ghotki, and Khairpur. The program focused on three groups: Adolescents, families, and community representatives, as well as public authorities.

Scope of the Program

The Improving Adolescent Lives in Pakistan (IALP) program intended to bring change through its two interventions, a Child Protection intervention and an educational intervention, which addressed two different mechanisms, broadly defined as (1) a change in knowledge about rights and gender roles, as well as (2) education. 446 rural communities were selected for Child Protection (CP) activities, including dialogues, workshops, and street theaters. Following a set of workshops, selected adolescents were expected to become group leaders ("adolescent champions") and conduct peer-learning workshops and meetings by themselves with assistance from a social worker. Additionally, in 100 rural communities of Sindh with poor educational outcomes, new Non-Formal Basic Education (NFBE) Centers were established. The Non-Formal Basic Education (NFBE) and village education committee component was implemented to provide non-formal education to out-of-school adolescents, that is, children who have never been enrolled, who have dropped out, or who were at the risk of dropping out. The second group that the program targeted included families, communities and local decision-makers, providing them with knowledge about adolescents' rights and how to protect and advocate in their favor. Finally, for the same purpose and with similar activities, the program also involved public authorities and service providers. It addressed children's rights and communication skills by providing access to relevant knowledge, skills, and services. In a key innovation, the program also encouraged boys and men to question institutional practices and social norms.

Although the baseline survey took place in 2015, preparatory work and recruitment of implementing partners delayed implementation. Program implementation was most intense between end-2017 and December 2019. Throughout implementation, UNICEF reported that they reached over 12.5 million people, directly and indirectly through program activities and social mobilization.

About the Evaluation

Given the limited and mixed causal evidence on the impacts of empowerment programs aiming to change gender norms and reduce child marriage (without additional monetary incentives), rigorous causal evidence on innovative approaches to design, combine, and improve empowerment programs is needed. Evaluation insights can help fill the evidence gap and support policy makers as well as UNICEF and the IKEA Foundation to strengthen future programs and inform policy advocacy.

This evaluation aims to identify the causal impact of the IALP program as a whole, and to learn about selected potential mechanisms of change. A related objective is to examine the cost effectiveness of the program.

The evaluation was designed as a cluster Randomized Control Trial (RCT) with 446 of 653 revenue villages randomly assigned to the program group and 207 to the control (no treatment) group, following a baseline survey in

2015. Cross randomization helps to compare different program components across the program areas. Hereby, the proportion between the genders mobilized in a village was experimentally varied. In 141 of the 446 villages more female participants were mobilized, while more male participants were mobilized in 129 of the villages. In the remaining 176 villages there was no gender-specific mobilization for the Child Protection activities. The establishment of alternative, non-formal basic education centers was cross-randomized to the Child Protection activities in a subset of 100 villages. The endline quantitative and qualitative data collection was conducted in late 2020 and ended in August 2021. 7,415 eligible adolescents, parents of adolescents, community leaders and key service providers were interviewed in program and control areas to share their knowledge and experiences on the topics of interest to the program.

In the analysis of this rich set of information, the key evaluation questions are:

1. Did the IALP program influence attitudes and social norms and reduce child marriage rates, decrease the number of early pregnancies, and increase school enrollment for adolescents as aimed?

2. How effective were the Child Protection and NFBE components of the IALP program, as standalone interventions and when combined?

3. How effective was the approach of targeting men and boys to achieve the program's goals?

Additionally, 16 qualitative key informant interviews (KIIs) were conducted with stakeholders of the IALP program as part of the qualitative data collection. These included UNICEF and government officials, as well as staff of the implementing partners from both provinces. KIIs provided additional insight into the experience of key informants with designing and implementing IALP program activities. Hereby the focus was on the challenges faced during implementation, perceived consequences of the IALP program, and recommendations of stakeholders for future programming.

Findings and Implications

The evaluation shows that the IALP program has a statistically significant positive impact on child marriage reduction. Compared to the control group, the overall IALP program reduces child marriage for the relevant age group (10-17) by 33% for the full sample and 50% for girls (with the average child marriage prevalence in the control areas being 0.9% for the full and 1.4% for the girls sample, respectively). For adolescents age 15-17, the overall IALP effect is a reduction by 42% for the full sample and 51% for girls. These effects are sizable given the low average prevalence in the control areas of 1.9% for the full sample and of 3.3% for girls only, i.e., the outcomes for the group of people who had not been assigned to the program. When restricting the sample to the NFBE-eligible sample of villages, the reported child marriage prevalence is higher, and, as expected, program impacts are positive and even slightly higher. Overall, the IALP program has a positive impact on child marriage reduction relative to the low level of child marriage captured in the surveys, with the program effects being higher for girls. Given sample size restrictions, however, no statistically significant and meaningful result can be captured for changes in child pregnancy rates.

Moreover, the IALP program increases self-reported engagement rates for boys age 10-17, while their marriage rates do not change. Accordingly, it increases the (indirectly) reported likelihood that boys (age 11-16) will be married next year. The evidence on changes in engagement and delay of marriage indicates a postponement in line with the original program expectations. Finally, adolescents in the NFBE areas of the IALP program report that the dowry amount paid by their families reduces.

The evaluation randomly sampled children who were eligible for non-formal basic education in program and comparison areas. Based on the comparison of the two groups, it finds an increased attendance of non-formal educational institutions, which further implies an improvement in selected learning outcomes for the population eligible for NFBE relative to the control group. When comparing learning outcomes of NFBE eligible children to comparable children in NFBE non-eligible areas, the evaluation finds positive impacts on reading skills, with the effects being mainly driven by improved outcomes amongst male respondents. Moreover, children residing in areas where the two program components (Child Protection and NFBE) were jointly implemented are not only more likely to attend the non-formal education centers, but also less likely to miss school, and more likely to want to continue with secondary education. The effects here are particularly strong for girls. A higher uptake of non-formal education may have a direct effect on attendance of formal educational institutions. The data reveals potential evidence of crowding-out effects from formal education, as female household members (age 11-23) are reportedly less likely to attend primary and male household members (age 11-23) secondary education due to the program. This effect is particularly strong in NFBE-eligible villages, where the Child Protection program targeted primarily girls. Moreover, potentially because these informal school forms

had not mixed genders in class while formal schools did, the results show selective evidence for a program-induced change in the preference that girls and boys should be educated separately.

To understand the potential underlying drivers of change, the evaluation looks at changes in perceptions, attitudes, actions and interactions of adolescents themselves as well as of the key local stakeholders and service providers surrounding them. While patterns in the intergenerational dialogue do not change, the results indicate an increase of interaction between adolescents. Girls' incidence of intragenerational dialogues among adolescents increases. Selective evidence reflects an increase in communication amongst adolescents regarding their fears and emotions, as well as the likelihood that girls talk to their brother about their future marriage. Moreover, especially where girls were targeted by the program, the results indicate an increase of girls' average score on the interpersonal communication competence scale, reflecting positive impact on girls' communication skills.

In terms of empowerment, especially male respondents in areas where the two program components were jointly implemented (Child Protection (CP) and NFBE) have a higher likelihood to report that they were reached on topics related to "child rights and entitlements". Due to the program, male adolescents do reportedly feel more confident in their interactions with key stakeholders. Eventually, the program also reduces the acceptance of using corporal punishment on children and leads to a decrease in the experience of corporal punishment by adolescents. Male respondents are, for instance, less likely to accept the usage of corporal punishment on daughters when they discuss their marriage timing. However, scope for improvement remains: Adolescents and parents of adolescents in program areas are no more likely to report that the adolescents' confidence in day-to-day situations changes when being compared to control areas. And adolescents, especially girls, still feel uncomfortable performing a set of daily tasks by themselves both in program and control areas. Accordingly, there is no reported evidence on changes in self-efficacy or on changes in actions taken to ensure adherence to adolescent rights in the last year. Eventually, the evaluation results also show a small but negative impact on the reported wellbeing of female adolescents over the last two weeks (before the endline survey), which may be due to the program or the end of it. The effect may be linked to girls feeling left behind after the program ended, as indicated in qualitative interviews, suggesting the need for further programming and engagement. Future implementation would benefit from stronger monitoring and a midline evaluation to identify potential negative effects early on during implementation. Given that the evaluation came about a year after the program ended, future research can help to better unpack the causes of these negative effects.

Targeting men and boys pays off. Findings indicate that the IALP program reduces agreement with the statement that it is OK for parents to marry sons before the age of 18 and reduces agreement to marry girls before the age of 16 in program areas, especially among males. This change may translate to different marriage outcomes in future. However, with every second person accepting marriages of girls before the age of 18, and every third person before the age of 16, there is much room for improvement and additional programming. The importance of engaging men, who may facilitate participation of women and girls, is also stressed in qualitative interviews.

The evaluation captures no general improvement in the overall situation with respect to the provision and access of adolescent-friendly services when comparing program and control areas. However, an improvement in access to non-formal educational services and a moderately better provision of marriage services as reported by the marriage registrars in the villages can be constituted.

Notably, despite being implemented over a relatively short period given the challenges at hand, the program manages to change attitudes and behaviors towards the expected and desired direction, including a reduction in child marriage, an increase in peer dialogues, and changes in attitudes towards marriage age and adolescents in general. To change deeply rooted social norms and traditions and to improve the magnitude of effects, a longer, more intense implementation period with greater outreach is recommended. Given a large number of beneficiaries who were reportedly indirectly engaged through social outreach activities such as street theater, a higher number of directly targeted individuals would have been important. In fact, the overall findings suggests only limited scope of direct program uptake, reducing the ability to capture greater impacts. This is less of a concern in Sindh, where greater uptake of the program is correlated with higher impacts. The involvement of men and boys when addressing topics related to child marriage is recommended. No recommendation is made on whether the educational or the Child Protection activities were more successful in reaching better outcomes, while in some instances the joint implementation of both approaches was reportedly helpful. Moreover, central institutions such as marriage registrars would need to be more intensively and directly targeted to ensure adherence to laws and a change in attitudes towards child marriages and its negative consequences. In this regard, the triangulation of quantitative and qualitative results points to converging insights. Moreover, the importance of community engagement and a careful consideration of sustainability of the intervention was felt to be of great importance by key informants. Last but not least, apart from the impacts the program was able to generate, the evaluation reveals high levels of adolescents' overall vulnerability (in program and control areas) and

thus the need to continue exploring impactful approaches to address these needs. This relates, for instance, to high level of drop-outs, corporal punishments, low levels of confidence and self-efficacy, as well as high levels of inegalitarian views.

A cost benefit analysis reveals that, under the assumptions of analysis, benefits of the IALP program largely cover its costs. The benefit-to-cost ratio is close to 1, i.e., costs exceed benefits only to a small extent. Hereby, the result constitutes only a lower bound since not all benefits were monetized and further positive changes can be expected given changes in attitudes and potential long-term effects.

List of Acronyms

3ie International Initiative for Impact Evaluation

A adolescent

a.c.e. about to complete education at the

AC adolescent-centric

agr. agrees

ALP Accelerated Learning Program

AQ adolescent questionnaire

AR adolescent right

al at least one

B boys

bc. because

BLCC Bunyaad Literacy Community Council

Bunyad Bunyad Literacy Community Council

C4DE Center for Development and Evaluation

C4ED Center for Evaluation and Development

CBA Costs-Benefit Analysis

CC community-centric

CCT Conditional Cash Transfer

CEA Costs-Effectiveness Analysis

CEDAW Convention on the Elimination of all Forms of Discrimination against Women

CL community leader

CLQ community leader questionnaire

CM child marriage

c.d.o. considered dropping out of the

comm. community

c.e. complited education at the

corp. corporal

c.i. correctly identified

c.c. correctly calculated

CP Child Protection

CR child right

CRC Convention on the Rights of the Child

CPF child protection intervention targeting females

CPM child protection intervention targeting males

CPMF child protection intervention targeting males and females

CPU Child Protection Unit

CYAAD College of Youth Activism and Development

dau. daughter

dec. decision

d.n.p. did not participate due to

d.t.c. discussed topics covered in the

d.e.w. discussed event with

d.n.w. discussed about NFBE with

e.org. event organized by

e.p.a. event participants were

e.p.g. event participants were

EQ evaluation question

FGD focus group discussions

FDR False Discovery Rate

f.e. frequency of event

G girls

GDP Gross Domestic Product

GNI Gross National Income

GO local government officials

gov. government

HDI Human Development Index

HH household

HW health worker

h.e. heard about event from

i.l.y. in the last year

ICC interpersonal communication competence

IALP Improving Adolescent Lives in Pakistan

IGD Intergenerational Dialogue

IKEA-I Preventing Child Labour (an IKEA-funded program)

IKEA-II Improving Adolescent Lives in Pakistan (an IKEA-funded program)

IKEAF IKEA Foundation

imp. important

i.e.c. implementation of event by

ind. independently

ind. individual

intergen. intergenerational

intragen. intragenerational

IP implementing partner

IRB institutional review board

IRC Indus Resource Center

ITT Intention-to-Treat Effect

KCI Key Community Influencer

KI key informant

KII key informant interview

KPI Key Performance Indicators

LATE Local Average Treatment Effect

mar. marriage

main reasons attending: main reasons attending:

m.c. multiple choice question

MDES Minimum Detectable Effect Size

MDE minimum detectable effect

MICS Multiple Indicator Cluster Surveys

MIS Monitoring Information System

NRSP National Rural Support Programme

N Nikah Khwan

NFBE Non-Formal Basic Education

NFBEC NFBE Center

NFE Non-Formal Education

NGO Non-Governmental Organisation

NKQ Nikah Khwan questionnaire

NPV net present value

oUN only UNICEF[n1]

o.l.2w. over the last 2 weeks

o.l.4y. over the last 4 years

OLS Ordinary Least Squares

OOS out-of-school

org. organisation

o.p.e. other participants of event were

P parent

part. participated

pp percentage points

PAP Pre-analysis Plan

PDHS Pakistan Demographic and Health Survey

PKR Pakistani Rupee

PQ parent of adolescent questionnaire

PV present value

pn. punishment

RCT Randomized Control Trial

rea. reason

resp. respondent

RSPN Rural Support Programmes Network

RV revenue villages

SDG Sustainable Development Goals

SDM school drop-out or never-enrolled questionnaire/module

SED secondary education

s.r. self-reported

SOC Social Opportunity Cost of Capital

SRSO Sindh Rural Support Organization

ToC Theory of Change

t.e.r. thought education received at

TPFM Third party

t.c.e. topic covered at event was

UC union council

UN United Nations

UNICEF United Nations Children's Fund

USD United States Dollars

VEC Village Education Committee

vill. village

WHO World Health Organization

WVS World Values Survey

I. General Information

Funding Research presented in this evaluation has been commissioned by United Nations Children's Fund (UNICEF) Regional Office for South Asia through the International Initiative for Impact Evaluation, Inc. (3ie), with funding of the IKEA Foundation. Program implementation was managed by UNICEF Pakistan. The views expressed in this report are not necessarily those of UNICEF, 3ie (Thematic Window 12) or its members. This impact evaluation has been submitted to 3ie in fulfillment of requirements under grant TW12.1009 issued under TW 12 Improving Adolescents' Lives program.

Researchers and their Role The authors of the evaluation are Alexandra Avdeenko (corresponding author) and Markus Frölich.[1] The views and opinions expressed are those of the authors' and do not necessarily reflect the official policy or position of UNICEF or 3ie. All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Alexandra Avdeenko (avdeenko@c4ed.org). Finally, the researchers had no direct decision-making power over whether and how to implement the program.

Quality Assurance The 3ie technical quality assurance team comprises Sayak Khatua, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Sebastian Martinez.

Researcher Independence and Financial Conflicts of Interest No contractual limitations existed on the ability of the researchers to report the results of the evaluation. The researchers have no financial conflicts of interest with regard to the results of the research. The researchers' contract for this evaluation was with 3ie.

Scope The impact evaluation of the Improving Adolescent Lives in Pakistan (IALP) programme is part of a larger package of impact evaluations covering the IAL program in three countries (Afghanistan, India and Pakistan). The Terms of Reference (ToR) for the assignment covered all these evaluations as well as other M&E services. The Pakistan evaluation included a baseline survey conducted in 2015 and an endline impact evaluation drawing on endline survey data collected in 2020-2021. The endline impact evaluation scope of work was reviewed and adjusted in the first quarter of 2020 and adjusted again in October 2020 to adapt the workplan and design to the COVID-19 pandemic outbreak. The pre-analysis plan agreed in October 2020 defines the scope of work and design of this impact evaluation. The pre-analysis plan is available for consultation.

Evaluation Process The original Terms of Reference (TOR) was between 3ie and UNICEF, and between 3ie and University of Mannheim/ C4ED, with several rounds of extension mostly due to COVID-19 related delays in the endline survey timing. A Pre-analysis Plans (PAPs) was developed over time and documented the changes. The final version is publicly available under AEARCTR-0006438 at the American Economic Association website. The original TOR is available upon request.[2] Despite contractual extensions, the content of the TOR did not change over time.

Timeline 3ie's call for proposal was issued on February 26th, 2015. The baseline data collection spanned from June 2015 (piloting) to January 2016 (full survey). The assignment into the program vs. control group was performed in winter 2016 for the CP activities and summer 2017 for the NFBE activities (cross-randomized). The Program implementation took place between summer 2017 and winter 2019, both program components started after the respective program assignment plan. The implementation was conducted by three local Non-Governmental Organisations (NGOs), including a change in implementing partners in Punjab. The prioritization of non-formal basic education (over an original focus on secondary school education) was introduced over the course of the program. The endline data collection officially began on December 15th 2020 (first registered interviews) and ended on August 15th, 2021. Originally, the endline data collection was planned for January 2020, the difference in timing being caused by the outbreak of the COVID-19 pandemic. The qualitative interviews were conducted between February 2021 and June 2021. The first draft of the final

---

[1]Both researchers from the Center for Evaluation and Development (C4ED). Dr. Avdeenko is also affiliated with University of Heidelberg and CEPR, email avdeenko@c4ed.org. Prof. Frölich is Professor at the University of Mannheim and affiliated with J-PAL and IZA, email froelich@c4ed.org.

[2]Consistent with how the evaluation was commissioned, there is no separate TOR for the IALP evaluation.

evaluation report was shared on September 15th, 2021.

Evaluation Pre-registration A pre-analysis plan was registered prior to the endline data collection, on November 29th 2020, under AEARCTR-0006438.

Institutional Review Board (Ethics Approval) The necessary ethics approvals have been requested from the Institutional Review Board of the University of Mannheim and are also, additionally, under review at Research and Development Solutions in Pakistan.

Declaration of Interest None.

Scarcity Budget and other non-research considerations were the critical factors limiting scale and as such the random assignment of the program merely influenced who got which program activity.

II. Acknowledgements

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

I.   Existing Evidence and Contribution

Adolescence, age 10-19, is characterized by a range of biological, psychological, and social changes.Transitioning into new roles in their families and communities, children quickly outgrow their protective environment. The age at which they collect their first independent, intimate experiences is decisive for their long-term wellbeing and development. Good knowledge and application of legal rights can offer a new layer of protection, acquiring more education another.

The Relevance of the Problem   Above all, marrying at a young age can lead to emotional scars for underage girls and boys. As childbearing often starts shortly after marriage, the consequences for girls are especially damaging, including lower levels of education (Field and Ambrus 2008), lower decision-making power, long-term economic disadvantages, and higher levels of gender discrimination (Amin, Ainul, Akter, Masudul Alam, Ahmed, and Rob 2014), as well as more health complications (for early mothers (UN Human Rights Council 2017) and for their children (Garcia-Hombrados 2021)). Moreover, also boys' childhood ends with early marriage as they have to take over adult responsibilities and roles as early bread-winners for their families.

In South Asia, shortcomings in the legal enforcement of existing child rights exacerbate adolescents' vulnerabilities. Here, children have the highest probability world-wide to be married according to UNICEF South Asia 2016. Instead of following the rules, officials in Pakistan often carry out marriage registrations by own judgment and therewith respond to societal pressures or are simply unaware about the details of the law (Field and Vyborny 2018). Domestic violence is a widespread problem (Solotaroff and Pande 2014), especially among married girls under 18 (UNICEF 2021), and non-egalitarian norms are accepted even amongst youth themselves (Asadullah, De Cao, Khatoon, and Siddique 2020).[1] 13.5% of Pakistani girls and 2.6% of Pakistani boys age 15-19 were reportedly married in 2017-18 (UNICEF 2021a).[2] While child marriage rates have stagnated over the past decades, before the pandemic, school attendance and literacy rates have remained low and drop-out rates high along with child labor, especially for girls (Government of Pakistan 2017).

> **The Context: Adolescents in South Asia and Pakistan**
>
> Pakistan's adolescents face a number of challenges, with some being amongst the most demanding ones in South Asia. Pakistan's population was 220,892,331 in 2020, thereof about every fifth person is in the age group of 10-19. According to the Pakistan Demographic and Health Survey (PDHS) 2017-18, 3.6% of the 20- to 24-year-old women were married by the age of 15, and 18.3% by the age of 18 (National Institute of Population Studies 2018).

[1]According to UNICEF, 25.4% (32.1%) of women and 23.6% (31.1%) of men age 15-49 years in Punjab (Sindh) state that a husband is justified in hitting or beating his wife under certain circumstances (Bureau of Statistics 2020; Bureau of Statistics, Bureau and Board, Planning & Development and of the Punjab, Government 2018).

[2]In Sindh, UNICEF reports that 15% of girls and 4.8% of boys age 15-19 are currently married (Bureau of Statistics 2020). For Punjab these figures change to 10.5% of girls and 2.9% of boys (Bureau of Statistics, Bureau and Board, Planning & Development and of the Punjab, Government 2018).

The Program　　This evaluation assesses whether empowerment and non-formal education for out-of-school children reduces child marriage and pregnancies and increases schooling in rural Pakistan.[3] The "Improving Adolescent Lives in Pakistan" program, or shortly the IALP, was implemented with support of UNICEF through different implementation partners in rural villages in four districts: Bahawalpur and Rahim Yar Khan in Punjab, as well as Ghotki and Khairpur in Sindh.[4]

The implementation focused on three groups: First, on adolescents as the intended rights-holders to be supported by the program. The goal was to give them the abilities and knowledge to be "agents of change" for themselves and their communities, who are critical in enabling adolescents to exercise their rights. The idea was to increase their knowledge about child rights, and about opportunities to seek protection and higher education, ultimately also improving both. Second, it targeted families, communities and local decision-makers, providing them with knowledge about adolescents' rights and how to protect and advocate in their favor. Third, for the same purpose, it also involved public authorities and service providers, as primary duty-bearers of children's rights.

To achieve the overarching goals, two broad types of interventions were implemented: A Child Protection (CP) and an educational intervention with a focus on Non-Formal Basic Education. The Child Protection intervention aimed at changing the norms and skills of beneficiaries, primarily through life-skills training, workshops, intergenerational dialogues, and community mobilization. These activities were targeted foremost at adolescents, yet also involved other community members. The activities aimed at achieving the program's overall objectives (1) by enhancing socio-emotional skills such as self-esteem, empathy and respect, as well as skills such as communication and expression, coping skills with stress and managing emotions, and (2) improving adults' attitudes toward adolescents.[5] In selected areas, boys and men were particularly encouraged to participate in the activities. In addition to the Child Protection intervention, NFBE Centers were established in 2017 in order to directly improve enrollment rates for drop-out children and therewith to also contribute to the achievement of the higher goals of the program.

Background on the Evaluation Framework: Objectives

The **main objective** is to identify and measure the causal impact (effectiveness) of the IALP program on the lives and environment of adolescents in selected areas of Sindh and Punjab, Pakistan. In the process of causal analysis, a goal in itself is also to describe and assess the program implementation process. Finally, causal impacts of introducing variations in program implementation are assessed, such as varying the targeting (gender of primary beneficiaries) and content and combination of the program activities (education vs. child protection). Once the impacts of IALP program are identified, a **secondary objective** is to relate them to the costs for the program and therewith assess program efficiency. This evaluation is supplemented by qualitative evidence from interviews. For all research questions, please refer to Box I..

Identification of Causal Program Effects　　Prior to program roll-out, a large-scale baseline survey with adolescents, parents, community leaders, and marriage registrars was implemented. Following, a cluster Randomized Control Trial (RCT) with several treatment arms was implemented in order to identify the impacts of the program. The Child Protection randomization and the baseline survey were conducted in autumn 2015 and the NFBE randomization in July 2017, all prior to program roll-out. Out of 653 villages, 446 randomly selected rural communities received the relevant workshops of the program. Thereby, out of the 446 Child Protection-treatment villages, in 141 villages more female participants and in 129 villages more male participants were mobilized, and in 176 villages no gender-specific mobilization took place for the Child Protection activities (i.e., both genders were equally strongly mobilized). Additionally, in 212 rural communities with poor educational outcomes, 100 villages were cross-randomized to establish

---

[3]According to UNICEF, "adolescent empowerment" refers to "A personal journey during which an adolescent (age 10-19), through increased assets and critical awareness develops a clear and evolving understanding of themselves, their rights and opportunities in the world around them, and through increased agency, and voice and participation, has the power to make personal and public choices for the improvement of their lives and their world" (UNICEF 2021).

[4]The program is part of the "Improving Adolescent Lives in South Asia" initiative which aims at providing the necessary means to reduce some of the challenges faced by adolescents. It is a program in the three countries Afghanistan, Pakistan, and India, and implemented under the supervision of UNICEF and funded by the IKEA Foundation (IKEAF). For Pakistan, the implementation partners were Rural Support Programmes Network (RSPN), Indus Resource Center (IRC), College of Youth Activism and Development (CYAAD), Bunyaad Literacy Community Council (BLCC), Sindh Rural Support Organization (SRSO), National Rural Support Programme (NRSP), Bunyad, and the Directorate Government of Sindh. In all three countries, the program has three main goals: (1) To sustainably reduce child marriages and (2) early pregnancies, as well as (3) to increase schooling for both young female and male adults.

[5]Both selected adolescents (so-called "adolescent champions") and key community influencers were encouraged to pass on this new knowledge, enabling a "trickle down" effect of the initial workshops. Other activities under the Child Protection intervention included interactive theaters, action plan implementation, the establishment of youth information centers, and the provision of vocational skills trainings.

new NFBE Centers, with 60 villages which received both types of interventions. The implementation of the Child Protection and NFBE activities started at full scale in 2017, with regional variation in intensity. In December 2019, the program implementation was stopped and the endline survey was conducted between winter 2020 and summer 2021.

Background on the Evaluation Framework: Scope

The **thematic scope** of the evaluation focuses on whether empowerment and non-formal education reduces child marriage and pregnancies and increases schooling. The **evaluatory scope** of work did **not** cover an assessment of the relevance, coherence or sustainability of the program. The evaluation was guided by a theory of change developed in detail jointly with UNICEF. In particular, it defined and guided the outcomes and assumptions as well as the hypotheses tested in the impact assessment (Section O.I.A.1 for reference). An ex-ante assessment of the coherence, relevance, and validity of the ToC was not part of the scope of the evaluation. Based on this, a pre-analysis plan was developed and publicly logged before the endline data collection started. The **time scope** of the impact evaluation covers a period from baseline (2015) to endline survey (2020/21). The evaluation measures the impacts one year after the official end of the IALP program in 2019. The end of specific program activities may have varied. Finally, the **geographic scope** covers four program districts of rural Sindh and Punjab, Pakistan. The education component was only assessed in one province (Sindh) where it was implemented in a subset of education component-eligible villages. **Changes in scope over time** included the drop of urban areas which were no longer part of the eligible program pool. Moreover, the inclusion of the evaluation of the educational component was added in summer 2017 as a cross-randomization to the original design covering only child protection interventions.

The **N**etwork on **D**evelopment **E**valuation by the **O**rganisation for **E**conomic **C**o-operation and **D**evelopment **D**evelopment **A**ssistance **C**ommittee defines six evaluation criteria, relevance, coherence, effectiveness, efficiency, impact and sustainability (OECD 2021). The focus of the evaluation is on "*effectiveness*" (Is the intervention achieving its objective?, see Section II.A), "*efficiency*" (How well are the resources being used?, see Section II.D), and "*impact*" (Has the intervention generated or is expected to generate significant positive or negative, intended or unintended, higher-level effects? Effects of the intervention that are longer term or broader in scope than those already captured under the effectiveness criterion, see Section II.B). The evaluation does not explicitly assess (a) trade-offs between different priorities or needs (i.e., "*relevance*"), or (b) an assessment on the complementarity, harmonization and co-ordination with other interventions in the country, sector, or institution, the extent to which the intervention of interest is adding value while avoiding duplication of effort (i.e., "*coherence*"), or (c) the "*sustainability*" of the intervention (i.e., the extent to which the net benefits of the intervention continue, or are likely to continue).

Contribution to the Global Evidence   The evaluation makes a number of contributions to the literature on how to reduce child marriage. To achieve this goal, a wide range of approaches have been implemented and assessed. Programs focusing on increased education, empowerment, and financial incentives have been shown to have the potential to delay the age of marriage and reduce fertility.[6] One key lesson from the literature is that a combination of approaches seems to be most effective: Empowerment without changing opportunities may be insufficient, if not even harmful. Informing that child marriage is legally forbidden, without enabling a supporting environment can be shortsighted. Similarly, teaching girls about their educational rights, without providing access to schooling, can exacerbate vulnerabilities.

Given the higher prevalence of child rights violations in poor households, financial incentives have been identified to be particularly effective to reduce such violations.[7] In a recent study, Buchmann et al. (2021) show that providing direct small financial transfers for not marrying reduced child marriage (unconditional on school attendance). They find that girls eligible for the financial incentive treatment arm - conditional on marriage alone - were 24% less likely to be married under the age of 18. In order to identify the most cost-effective solutions, rigorous studies simultaneously test combinations of program activities. Buchmann et al. (2021), for example, directly compare the impacts from financial incentives to a version of the program which implemented only girl's empowerment and norm-changing activities.[8]

---

[6]See Ashraf, Bau, Nunn, and Voena (2014); Baird, McIntosh, and Özler (2011); Breierova and Duflo (2004); Buchmann, Field, Glennerster, Nazneen, and Wang (2021); Currie and Moretti (2003); Duflo, Dupas, and Kremer (2015); Hahn, Islam, Nuzhat, Smyth, and Yang (2018a); Hahn et al. (2018a); Muralidharan and Prakash (2017); Osili and Long (2008). Moreover, Lee-Rife, Malhotra, Warner, and Glinski (2012) conducts a meta-study of 23 programs targeting child marriage. The authors focus on five approaches through which child marriage is commonly addressed: Empowering girls, educating and mobilizing parents, improving access and quality to formal schooling for girls, offering economic incentives for girls and their families, and enhancing and improving the overall legal and policy framework. Out of those categories, Lee-Rife et al. (2012) find that interventions attempting to empower girls as well as those providing incentives seem to be the most promising approaches.

[7]Amongst the first, for instance, Baird et al. (2011) rigorously examined the effect of cash transfers in Malawi, designed to provide incentives to girls to remain in school. They find that unconditional transfers led to a reduction in the likelihood of ever being pregnant and ever marrying by 27% and 44%, respectively, compared to the control group. Similarly promising results of cash transfers were identified for a program in Punjab, Pakistan. Here, the independent evaluation group of the World Bank evaluated a Conditional Cash Transfer (CCT) program in Pakistan and found that it successfully increased the likelihood of adolescent girls finishing middle school, working less, and marrying later (World Bank 2011).

[8]The goal of empowerment programs is to improve knowledge on rights and skills (see for more evidence Adoho, Chakravarty, Korkoyah Jr., Lundberg, and Tasneem (2014); Bandiera, Buehren, Goldstein, Rasul, and Smurra (2019); Buchmann et al. (2021); Calero, Diez, Soares, Kluve, and

They can show that the empowerment program alone was successful in promoting more progressive gender norms, but unintentionally encouraged early marriage, while also increasing dowry payments in empowerment program areas. In other words, their work demonstrates that policies aimed at changing preferences may be misdirected.

Given the importance of being able to translate preferences into behavior, the program of interest here aimed at changing the intragenerational and intergenerational dialogue, explicitly also involving elderly and parents. A relatively new body of evidence shows promising results for innovative approaches teaching children persuasion skills in order to enable intergenerational dialogue.[9] For instance, in an RCT, Ashraf et al. (2020) find that training girls in negotiation skills not only improves their educational outcomes but may also lead to more optimal parental investment in their daughters' education. While the underlying evaluation finds little evidence of changes in intergenerational dialogue, it constitutes that the IALP program increases the prevalence of intragenerational dialogue among peers, especially amongst girls.

Program effect sizes depend also on the targeting group. Typically, the programs have been primarily targeting females, training girls in life skills, negotiation, legal rights of women, and nutritional and reproductive health knowledge.[10] The literature so far, however, has largely missed out on causally identifying the added value of involving boys and men in adolescent empowerment programs. Given that one of the prevalent problems of high levels of child marriage are societal preferences (social conservatism, harmful traditional gender norms, and strong beliefs about the benefits of marrying young), directly targeting men's preferences seems reasonable and doing so already at a young age could therefore be promising. In fact, there are a number of positive benefits for men from empowered women and reasons for why men could want to agree with granting more rights to women (Doepke and Tertilt 2009; Geddes and Lueck 2002). According to the work by Doepke and Tertilt (2009), for instance, men would prefer wives to have rights, at least wives of other men, because men would care about their own daughters (especially to increase their marital bargaining power, because they would not want their daughters to be treated badly by their sons-in-law) and because an expansion of women's rights increases educational investments in children. In India, Dhar et al. (forth.) evaluate a program centered around classroom discussions about gender equality in schools. Students reflected on their own and society's views and discussed gender stereotypes. The researchers find that gender attitudes are more progressive for students that participated in the program and that in the long-term, the effect was larger for boys' attitudes on gender equality and boys were even more likely to translate it into behavioral change than girls. Boys did more chores and encouraged their sisters to pursue college. Dhar et al. (forth.) draw an important conclusion about empowerment per se and the feasibility of translating it into change. They argue that behavioral change would require not just a desire but also an ability to change. Thus, the very fact of boys' and men's greater power in society renders it important to include them in interventions aimed at increasing girls' and women's power. This hypotheses is further supported by Edmonds, Feigenberg, and Leight (2021) who evaluate an empowerment program in Indian schools and find no impacts on changes in aspiration and child marriage, arguing that girls were not able to translate their enhanced agency to negotiate with parents. This evaluation contributes to this literature by identifying the additional impacts of making especially adolescent boys and men more perceptive to sensitive gender-related topics. Instead of class-room trainings, the setting is rural villages. Here, it is experimentally varied which village is assigned to which gender-targeting approach. The results confirm the pattern identified in India: Villages where more boys/ men were targeted with Child Protection interventions, display the biggest reduction in child marriages rates for girls age 15-17. The findings also suggest that especially men changed their acceptance levels of female marriages under the age of 16, being more likely to reject them.

Background on the Evaluation Framework: Purpose

The ultimate goal is to add the learning generated to a greater **global evidence base on what works to empower adolescents and end child marriage**. Apart from the **research community**, the report speaks to the community of **implementers, policy makers, and donors**. In particular, the evaluation findings are expected to be used by policy makers across South Asia seeking to identify approaches that will yield better results for children. Policy makers as well as UNICEF and the IKEA Foundation are expected to gain insights that can be used to strengthen future programs and inform policy advocacy.

Corseuil (2017); Edmonds, Feigenberg, and Leight (2019); Ibarraran, Ripani, Taboada, Villa, and Garcia (2014)).

[9]See Amin, Ahmed, Saha, Hossain, and Haque (2016); Ashraf, Bau, Low, and McGinn (2020); Bandiera, Buehren, Burgess, Goldstein, Gulesci, Rasul, and Sulaiman (2020); Dhar, Jain, and Jayachandran (forth.).

[10]In their theoretical model, for instance, Buchmann et al. (2021) hold social conservatism of men and older adults constant. Apart from girls, the pervasive approach in these interventions has been to engage communities as a homogeneous construct and/ or to engage girls and women exclusively (Rasmussen, Maharaj, Sheehan, and Friedman 2019).

The IALP program aimed at directly improving schooling by changing existing norms and beliefs as well as increasing girls' desire to pursue their educational aspirations. Yet, empowerment without additionally changing the opportunities available to the youth may be less effective. This evaluation investigates the added value of combining the empowerment program with the establishment of new centers for out-of-school children, which aimed at helping adolescents to catch-up with schooling in a shorter period of time. With this approach, particularly vulnerable girls, those who are not in school, could be targeted. By measuring the additional and pure impact of establishing alternative learning centers (NFBEs) on the outcomes of interest, this evaluation adds to the body of evidence on evaluations of multifaceted policy intervention.[11] The effects differ for boys and girls. The empowerment program alone does not have direct impacts on educational outcomes of either girls or boys. Establishing alternative schools without changing norms, has no impact on girls' outcomes either. However, learning outcomes for randomly sampled eligible boys improves if compared to children in control villages. This is an important finding as studying the impacts of approaches targeting out-of-school children is at the top of policy attention due to the COVID-19 lockdowns and subsequent economic crises.[12] For out-of-school girls' learning outcomes, it is the combination of the two program elements which is most beneficial. Learning outcomes for girls improve especially in areas where the empowerment activities targeted more boys. This speaks to the literature and findings described above on the added value of involving boys in empowerment programs and of carefully designing targeted programs to achieve the desired goals.

Finally, the long-term effects of empowerment programs, after support and opportunities end, are unclear. On the one hand, adolescents who know their rights may fare better in times of hardship. On the other hand, a sudden end of a program after having being equipped with new knowledge about rights and opportunities, may be particularly discouraging. Programs usually end due to the end of funding resources or external factors. In this case, the formal end of the program coincided with the onset of the COVID-19 pandemic, which has further complicated the transition to adulthood for many adolescents. It has cut social contacts and access to safety networks, and moreover has restricted mobility and exchange of information. For millions of girls world-wide it results in early child marriage, followed by negative long-term consequences. The evaluation analyzes whether the empowerment program has made adolescents more resilient and more prepared to face hardship. Prior changes in preferences and harmful practices, if effective, could matter when shocks occur. Bandiera et al. (2019), for instance, analyze an intervention to empower young women in Sierra Leone, leveraging the quasi-random between-village variation in the severity of disruptions due to the Ebola outbreak. They find that in highly-affected control villages, women spent more time with men compared to the baseline, and the likelihood of being pregnant increased. In contrast, in treatment villages (both highly affected or little affected), girls decreased the time spent with men and were less likely to be pregnant outside marriage.[13] Apart from the direct program effects, the underlying evaluation is thus interested in whether children who were prior empowered were indeed more resilient once the COVID-19 pandemic hit.

The rest of this evaluation is structured as follows. Section II. describes the program and its context. In Section I. the evaluation design is described. Section II. offers a discussion of the data, with clear definitions for the variables used in the program. Section 3 presents all findings of the report, including a section testing the robustness of the main results. The report ends with Section 4, a conclusion, followed by lessons learned and recommendations.

II.   The Program

The IALP program is part of a greater initiative implemented by UNICEF in South Asia (Afghanistan, India, and Pakistan) that aims at (1) reducing child marriage, (2) increasing adolescent participation in education, and (3) reducing teenage pregnancies over a time span of three and a half years.[14]

---

[11]Note, Non-Formal Basic Education is also sometimes referred to as the Accelerated Learning Program (ALP).

[12]Per se, while an innovative approach to address high rates of dropouts is necessary, the effects of establishing alternative learning centers are not clear ex-ante. Eventually, they may lead to crowding out effects from formal education if students receive the same education in a shorter period of time. And parents who send their children to work may use the system to achieve both goals: Income-generation from child labor for their households and education through accelerated learning.

[13]Relatedly, Corno, Hildebrandt, and Voena (2020) show the relevance of marriage institutions when unanticipated droughts occur, which reportedly increased rates of child marriage in brideprice societies, but decreased them in those that traditionally pay dowries.

[14]Notably, prior to the baseline of the IALP program in 2015, UNICEF had implemented earlier adolescents interventions in Pakistan as part of the Preventing Child Labour (an IKEA-funded program) (IKEA-I) program. Starting in 2011, the IKEA-I program came to a close in December 2017, and aimed to provide backstopping support to multiple levels of government, as well as communities, families and other stakeholders in order to strengthen the protective environment for children for the enhanced realization of child rights in the targeted areas.

> **Sustainable Development Goals (SDG)**
>
> UNICEF has been mandated by the United Nations General Assembly to advocate for the rights of children aged between 0-18 years. UNICEF is guided by the Convention on the Rights of the Child (CRC), which was adopted in 1989. Adolescents' rights are enshrined in the CRC and in the Convention on the Elimination of all Forms of Discrimination against Women (CEDAW), which applies to all adolescent girls. Globally, adolescent programs have been integrated into UNICEF's Strategic Plan 2014-2017. This encompasses all sectors and will build on current strengths and requires development of an investment framework for 10-19 year olds. In South Asia, UNICEF's regional office has prioritized adolescents and participation in its Regional Office Management Plan 2014-2017, and a full-time position of Regional Advisor, Adolescents Development and Participation has been established (UNICEF South Asia 2019a).
>
> The IALP program targets the areas of quality education and gender equality, promoting the United Nations (UN) SDG 4 and 5. In accordance with target 5.3 of the SDGs, it aims to eliminate harmful practices against women, such as child, early, and forced marriage. Additionally, it focuses on SDG target 5.c with the goal of "strengthening enforceable legislation for the promotion of gender equality and the empowerment of all women and girls at all levels". Regarding quality education it strengthens SDG 4.7 to "ensure that all learners acquire the knowledge and skills needed to promote sustainable development, including, human rights, gender equality, promotion of a culture of peace and non-violence".

In Pakistan, the program directly reached 269,143 adolescents and 321,040 other direct beneficiaries (UNICEF 2020).[15] Note, a detailed theory of change framework was developed before the endline survey and guided the development of the survey instruments and hypotheses. It is outlined in Section O.I.A of the Appendix. The cost of the IALP program adds up to around 1.5 million United States Dollars (USD), where around 1.1 million USD can be attributed to the Child Protection component.[16]

II.A   Context

Pakistan's adolescents face a number of challenges, with some being amongst the most demanding ones in South Asia.[17] Adolescents are a big share of Pakistan's population, which was 220,892,331 in 2020, whereby about every fifth person was in the age group of 10-19.

Child Marriage   Legally, the *1929 Child Marriage Restraint Act* has for many years set the minimum age for marriage to 18 years for boys and 16 years for girls in Pakistan. The Pakistan Senate later aimed at amending it by raising the minimum age for girls to 18 as well (Rehman 2019). The amendment was ultimately rejected by the National Assembly. Previously, the province of Sindh had already raised the minimum age for marriage for girls to 18 and the province of Punjab has been working on a similar bill (of Sindh 2013). In other words, other than the province of Sindh, which has increased the marriageable age of a girl to 18 years, the remaining provinces and the federal laws stipulate that girls may marry at 16 years and boys at 18 years.[18] Moreover, accurate record of age is important to prevent child marriage. Almost 60 million children under the age of 18 years remain without the ability to prove their birth date, place, parentage and citizenship legally (UNICEF Pakistan 2019). And thus while the legal framework in Pakistan is improving, de-facto unofficially girls are still often urged to marry before the age of 16.

One in five women in Pakistan marry early, even though Malé, Chata and Wodon, Quentin (2016) indicate that child marriage before the age of 18 has significantly decreased over the past decades.[19]

---

[15]Direct adolescent beneficiaries: In 2015: 1,222; in 2016: 17,332 ; in 2017: 21,728; in 2018: 53,292; in 2019: 175,569. Other direct beneficiaries: In 2015: 45; in 2016: 1,932; in 2017: 67,260; in 2018: 108,380; in 2019: 143,423.

[16]Please find more detailed information on costs as well as benefits in Section I.D.

[17]Pakistan is a lower-middle income country (World Bank Indicator Website) and belongs to the group of so-called developing countries with its Human Development Index (Human Development Index (HDI)) score of 0.557 in 2019. This positions the country at 154 out of 189 countries and territories for which the HDI scores are available. With a Gross National Income (Gross National Income (GNI)) per capita of 1,410USD in 2019 - a decrease from the previous year - it ranges below the average of 1,533USD amongst the group of fragile and conflict affected countries. Another indicator for slow economic growth is the development of the Gross Domestic Product (Gross Domestic Product (GDP)). Since the millennium, Pakistan experienced an average GDP per capita growth of about 1.6% per year ranging from -1% up to 5%. Recently, the country's GDP per capita decreased by more than 1% in 2019. Furthermore, the Pakistan Bureau of Statistics (2020) from 2018-2019 indicates that 16% of the population lives with moderate food insecurity and 2.3% even with severe insecurity. The worst situation in Pakistan is seen in Sindh with 19% moderate or severe food insecurity (in Punjab with 14%).

[18]A move to raise the minimum age for marriage to 18 years for boys and girls followed a recommendation by the *Chief Justice of the Supreme Court*. Since crime statistics are not separately disclosed for child marriage, only the numbers of kidnapping/ abduction can be reported with 15,188 in Punjab and 3,111 in Sindh for 2019 according to the Pakistan Bureau of Statistics.

[19]While in the group of women aged 41-49 years, 43% have been married before being 18 years old, according to Malé, Chata and Wodon, Quentin (2016), this number dropped to 27% for 23-30 year-olds and 18.7% for 18-22 year-olds. Their data also displays that women from wealthier families were less likely to marry below the age of 18. In the wealthiest quintile, only 7.3% were married before 18 compared to 36% for the poorest quintile. Additionally, the education level would be correlated with child marriage as only 10% of women with completed secondary education were

Child Pregnancies   With only every third woman having access to contraceptives in 2019, early child birth follows early marriages. In 2019, reportedly 38 out of 1,000 girls age 15-19 have given birth. This is close to the average in South Asian regions (not considering high-income countries), with Afghanistan, Bangladesh, and Nepal displaying higher rates. Early childbirth follows the path of child marriages since most early childbirths take place after marriage according to PDHS data 2017-18 (National Institute of Population Studies 2018). 6.8% of women aged 18-22 have had an early childbirth compared to 18.6% for women aged 41-49. Factors like wealth and education are strongly correlated with early childbirth.

Education   Early school drop-out follows from early marriages and pregnancies. Compulsory education is 10 years in Pakistan, yet when children were about to start lower secondary education at age 10, only 67.3% of girls had completed primary education in 2019, while the rate is 11.52 percentage points (PP) higher for boys.[20] This is the lowest rate amongst girls and also boys in South Asia, with the gender gap being bigger only in Afghanistan. Moreover, only 52.4% of boys and 46.3% of girls completed lower secondary education in the relevant age group in 2019.[21] Only Afghanistan displays lower levels in South Asia, and even here the rates for boys have been higher. Governmental survey data draws a similar picture for children out of school aged 5 to 16 years: 36% of girls do not attend school (compared to 25% for boys; 49% in Sindh) (Pakistan Bureau of Statistics 2020). One of the associated reasons for this discrepancy would be that girls are not allowed to attend school. The survey found that parents not allowing the daughter to go to school was the main reason (25%) for girls to leave school before completing primary (age 10-18). In contrast to that, only 1% of boys were reportedly not allowed to complete school for the same reason.[22]

Eventually, and maybe little surprisingly, only every fifth woman (20.1%) participated in the labor force in 2019, which was the lowest rate in South Asia. Only 49% of women can read compared to 71% of men and even though both numbers have increased since 2014, the gender gap in literacy rates is not closing (Pakistan Bureau of Statistics 2020). The Sindh region performs below average with female literacy rates of 44%, while Punjab ranges well above the country average. Against this background of high illiteracy, gender inequalities, and low access to basic services, addressing the needs of adolescents has been a core priority of programming in Pakistan.

II.B   Program Activities and Expected Changes

Changes in the three main outcomes - marriage, pregnancy, and education - are core to the program and therewith crucial for the impact evaluation (Figure O.2). The changes are expected to be caused by two interventions, a Child Protection intervention and an educational intervention, which address *two different mechanisms*, broadly defined as (1) a change in knowledge about rights (Section II.B.1) and gender roles, as well as (2) education (Section II.B.2).[23]

II.B.1   Child Protection Activities: Life-Skills Workshops

The program is built upon the premises (1) that addressing gender roles and children's rights matters for the above-mentioned main outcomes, and (2) that it is important to encourage boys and men to question institutional practices and social norms.

   The Child Protection intervention included a four-day workshop for adolescents age 10-19. The participation in the workshops was restricted to a small group of adolescents, denoted as "adolescent champions".[24] The selected adolescents were expected to become group leaders and to conduct a number of peer-learning workshops and meetings by themselves with assistance from a social worker. These activities are referred to as "trickle-down" events. To do so, the adolescents received material from UNICEF and a manual for the modules to be implemented. The training contained two sub-components. The first focused on sets of skills and behavior. The adolescent champions learned

---

married early compared to 32% for women with no education.

   [20]Latest indicators according to the World Bank databases (wbopendata), September 2021.

   [21]In 2018, the country had 22.8 million children out of school of which 74% are between 10 and 16 years old (UNICEF 2019). This indicator shows that many children are not getting any education opportunities since the official age for primary education is from 5 to 9 and schools therefore cannot admit them. The report also shows that 43% of female adolescents aged 10-19 are neither in education, employment or training compared to 7% of males.

   [22]See on the causes of school (non-)attendance in Pakistan also the work by Jacoby and Mansuri (2011).

   [23]For more information on the program, the detailed theory of change, intended effects, and implementation data please refer to Section O.I.A.

   [24]Both boys and girls, and sometimes both at the same time, participated in the workshops. The program intended to select adolescent champions with the support of the community and village organizations.

how to communicate effectively, make decisions, solve disputes, and facilitate group work. In addition, their training dealt with specific topics such as health and hygiene, child rights, gender norms, and the importance of education and life-skills such as sewing, cooking or designing. It was expected that the adolescents include these components in their own workshops. The overarching goal of these workshops was to facilitate personal development.[25] Moreover, community-centric (CC)-related events such as street theaters were implemented as activities under the Child Protection intervention with the goal of mobilizing different generations within communities.

At the same time, Key Community Influencers, public authorities, and service providers were targeted in order for them to better understand and address the development, needs, and challenges of adolescents. This group was meant to explicitly discuss gender-roles in their societies, and their impact on the upbringing of sons and daughters in their communities. Central, reputable, and motivated individuals were selected to engage in voluntary work.[26] One specific activity includes community intergenerational dialogues. These were planned to take place on a yearly basis at provincial level, and once through the project cycle on a district level. Here, all stakeholder of the program, adolescent champions, KCIs, religious leaders, Nikah Khwans (a marriage registrar who is often also a religious leader) but also government officials or police officers were meant to be invited to initiate relevant dialogues. In addition to such mobilization events, community-based media campaigns created awareness for the topics mentioned above.

Child Protection Target Groups (Experimental Treatment Arm Variation)   Moreover, the program divides the activities into three groups, targeting male and female adolescents and champions to a different extent. In other words, the intensity with which the different genders were mobilized was (experimentally) varied. In villages labeled as "male" more boys were mobilized (child protection intervention targeting males (CPM)), in "female" villages more girls (child protection intervention targeting females (CPF)), and in "mixed" villages both genders were equally strongly mobilized (child protection intervention targeting males and females (CPMF)). In more detail, it meant that while the communities and their key community influencers were mobilized similarly, some variation at the adolescent-level was introduced in the treatment arm villages. Thus, in girls-centered villages, girls were the adolescent champions, and in the boys-centered villages, the boys. It meant that further on, the peer-to-peer sessions were then also primarily performed amongst the same genders. Eventually, in CPM villages, for instance, the idea was that only male adolescent groups would be formed. Apart from the different target group, there was no specific toolkit that was otherwise different in the villages. Boys groups were explicitly exposed to learn about gender rights, as were girls themselves.

> **Equity-Based Approach in Program Implementation and Evaluation**
> The program puts a particular emphasis on equity through its focus on topics such as adolescent marriages and harmful, discriminating norms, which disproportionally affect girls negatively. However, girls and boys are similarly targeted by the program, which is also reflected in the sampling for this evaluation. Thus, despite the choice of topics which prioritize especially girls' unique needs, the target group of beneficiaries is male and female. Therewith, the roots of harmful practices are addressed for all adolescent children, considering also boys who suffer from harmful norms in their communities as well. Apart from the content on reproductive health, all content of the trainings remained the same for both genders.

This enables testing whether, for instance, the program unfolds its impact more when particularly altering the attitudes of males. By directly targeting boys and men, the architecture of the program is different from common approaches which tackle child marriage by either engaging communities as a homogeneous group or by raising awareness on gender imbalances in women and girls only. Instead, the program also addresses men and boys as "agents of change" in preventing child marriage, promoting education, and discouraging early pregnancies in an attempt to nurture a feeling of positive masculinity. Apart from the content on reproductive health, all content of the trainings remained the same for the three different groups.

---

[25]In particular, UNICEF designed modules to raise and strengthen: (1) Identity and self-esteem, (2) empathy and respect, (3) communication skills, and (4) stress management and emotions. These workshops were mostly implemented through small games in which adolescents share experiences or build-up awareness on their thoughts, ideas, or accomplishments. To a lesser extent other modules and activities were also implemented, for example by focusing on practical life skills such as sewing, designing, drawing, and rescue training, as well as by offering sport events, the exchange of small gifts, speech or poetry competitions, quizzes or singing lessons. A final group of activities attempted to create awareness for topics such as health and hygiene, gambling, selling of cigarettes to minors, the consequences of school dropout or the establishment of complaint boxes in schools, to name just a few.

[26]Key Community Influencers (KCIs) could participate in workshops lasting three days, covering topics such as "Understanding the Child Rights", "Who Are Adolescents & Stages of Adolescent Development?", and "Gender & Equality in Adolescents Programming". They were then expected to return to their communities and reach out to at least 50 adults and 6 adolescents through community mobilization events and activities.

Progress: Beneficiaries, Changes over Time, and Partners   According to UNICEF's reports, 2017 was the first full year of implementation of activities related to the IALP program. This included the launch of U-Report PakAvaz as a communication platform for adolescent participation and a mass media campaign as part of the regional initiative #WeTheFuture. The latter aimed at promoting empowerment and participation of adolescents through the inclusion of videos by selected adolescent champions who were supposed to act as role models for their peers (UNICEF 2017). Furthermore, 2017 saw the implementation of life-skills training for adolescent champions who "trickled down" their achieved competencies and knowledge to reportedly more than 91,000 adolescents through peer-to-peer sessions. Adolescents also jointly developed action plans that were to be implemented in their respective communities. Implementation in 2017 focused on Sindh. Given that the implementing partner (IP) changed in Punjab throughout 2018, the newly contracted IP Bunyad took up work only in September 2018. There, training of 651 KCIs on adolescent rights took place. The trained KCIs, in turn, conducted sessions with parents and community members. Moreover, the service provision and abuse report mechanism of the Child Protection Unit (CPU) were strengthened. In 2018, the provision of life-skills training and peer-to-peer sessions continued. With the inclusion of the new IP in Punjab, project activities were rolled out in Rahim Yar Khan and Bahawalpur (UNICEF 2018). 4,024 adolescents (2,430 girls) actively participated in life-skills training sessions which addressed child marriage, gender, and social norms. These adolescents in turn reached an additional 36,255 adolescents (15,095 girls) with key information on child protection through the implementation of so-called action plans. In the following year, service providers from all involved districts discussed and further developed their referral system for child and youth protection. adolescent-centric (AC) videos were disseminated as part of the #WeTheFuture campaign (UNICEF 2019). In 2019, UNICEF reportedly supported 116,416 adolescents (42,602 girls, 73,814 boys) in capacity-building activities, using a tested life-skills toolkit for adolescents. The adolescents were also supported in developing and implementing action plans reaching 209,753 of their peers (101,855 girls, 107,928 boys) with messages on adolescents' right to protection, right to education, right to participation, and freedom of expression, and choice to delay marriage. With training and peer-to-peer interactions, these adolescents reportedly articulated new visions of their future, reducing risk of child marriage.[27] Moreover, UNICEF and its implementing partners directly reached 26,656 community influencers (11,972 women, 14,684 men) with adolescent protection messages addressing child marriage and the importance of birth registration among others. Social mobilization efforts, including the use of street-theatre, reached 214,180 people (96,297 females, 117,883 males) with messages, with the goal to help shift social norms on child marriage, gender-related norms, and adolescent protection (UNICEF 2019).

II.B.2   Education Activities: Non-formal Basic Education Centers

One solution to high drop-out rates between primary and secondary education is non-formal basic education, which is delivered to children and youth.[28] The idea is that children and adolescents come to special centers, where independent of their age and social background they are attending the same grades. In these centers, children experience accelerated learning, which is important to catch up with their age-group if they dropped out of school earlier. Children undergo special assessment which is designed to support children and teachers to identify their actual learning level. To do so, entrance and exit tests as well as needs-based curricula are implemented.[29] Depending on their needs, students then undergo basic literacy classes (Urdu writing and reading, numeracy, English, resulting in a literacy certificate), life skills training (Islamic education, values, rights, and responsibilities), learn about income generating skills, and also receive individual additional coaching. The training under the NFBE intervention is believed to improve educational outcomes in particular for girls, who have disproportionally more limited access to education.

The Education Section of UNICEF tailored the NFBE interventions to the following categories of out-of-school adolescents: (1) Those who have never been enrolled in school, (2) those who have dropped out, and (3) those who are at the risk of dropping out. In other words, the NFBE intervention was adapted for out-of-school adolescents that have never enrolled in school and need a more comprehensive condensed curriculum. It was also adapted for those that have had some form of schooling and need a shorter bridge course to get back on formal education at key transition points, which are from primary (classes: 1-5) to middle (classes: 6-8), to lower secondary (classes: 9-10), and to upper

---

[27]Due to community norms, however, reach amongst girls had reportedly remained lower than boys. To address this, concerted efforts were made to organize separate events for girls and women.

[28]In 2017, the UN Human Rights Council (2017) expressed concerns about persistent gender disparity in a range of social indicators, including school enrollment rates. The government in Pakistan has made commendable efforts to address this concern and to increase school attendance rates, for example by providing stipends for girls. However, attendance rates remain low and literacy rates are relatively low as well, especially for women in rural areas. Moreover, drop-out rates are high in Pakistan, especially for girls. Nationwide, 1.022 million (60%) boys are enrolled in higher secondary schools/ inter colleges stage, whereas this figure is 0.674 million (40%) for girls (Government of Pakistan 2017).

[29]The materials for these NFBE initiatives were prepared with technical support, based on a comprehensive assessment of needs.

secondary (classes: 11-12). The program followed a planned, yet flexible approach in responding to students' needs so that they could be mainstreamed to formal school according to their age and performance. The centers worked with vulnerable adolescents for two to three years. The goal was to help them integrating back into formal schools or prepare them for the labor market.

The NFBE efforts in Sindh aimed at contributing to the overall effect of IALP by increasing service providers' capacities to promote and provide access to education.

Progress: Beneficiaries, Changes over Time, and Partners   A feasibility study for the implementation of the NFBE component was conducted in 2017. Under the responsibility of IRC, training for teachers for the newly established NFBE Centers in Sindh was realized and the first students were enrolled. More than half of all teachers and students were female. Village Education Committees (VECs) were established in close proximity to NFBE Centers (UNICEF 2017). By December 2017, 60 NFBE Centers were successfully established and 1,899 adolescents (1,174 girls) were enrolled. In 2018 the number of established NFBE Centers and VECs in Sindh totaled 100 each. In 2018, 1,959 out-of-school adolescents were enrolled in the certified accelerated learning programs (UNICEF 2018). In 2019, 3,058 adolescents attending the NFBE Centers were promoted to the third and final learning Package C, which is equivalent to grade 5 of the formal education (UNICEF South Asia 2019b).

Activities at the Non-formal Basic Education Centers

▶ *Source*: UNICEF Education Unit.

II.B.3   Overlap of the Interventions

The two types of programs, Child Protection and NFBE, were partly implemented in the same villages. This, however, did not mean a joint implementation. NFBE eligible children were not specifically targeted for Child Protection interventions, and vice versa. In essence, this meant the interventions were implemented in parallel. Partly, NFBE Centers were established more centrally in the revenue villages (RV), while the Child Protection interventions could have been more scattered. The gender mobilization of the NFBE Centers was not aligned with the CP treatment arms, i.e., villages that were assigned to the CPF treatment arms did not necessarily receive an NFBE Center which targeted only girls as well. Specifically, the gender mobilization for the NFBE Centers depended on the demand of the communities and parents.[30]

---

[30]Comparing data on the implementation of NFBE Centers with the treatment assignment reveals only minimal overlap in the gender mobilization of NFBE Centers and CP treatment arms, as displayed in Table A.5 in the Appendix.

CHAPTER 2

Methodology

I. Evaluation Design

The goal of this evaluation is to identify the causal impact of the IALP program as a whole and to learn about the mechanisms (elements) of change described above. An RCT was implemented at two levels and for two major types of interventions.

The core question of this evaluation is whether the IALP program had the desired positive impact on the communities, particularly the adolescents it targeted. The goal is to capture the size of this effect and the statistical confidence with which it can be claimed, considering thereby different variations and combinations of program components. While desired outcomes are important, given that this program was implemented for the first time in this context, this work aims to also capture potential unintended effects. To achieve this goal, it compares differences in outcomes between the program and control areas over a year after the end of program implementation, it reports effects on indicators of adolescent empowerment (awareness on child rights, confidence, corporal punishment, self-efficacy, general wellbeing), child marriage rates and practices, engagement rates and practices, child pregnancy rates, as well as education attendance, attitudes towards education, and learning outcomes. Moreover, it captures differences with respect to communication (intergenerational and intragenerational dialogue as well as interpersonal communication competence), service provision (access to services over time, adolescent-friendly services, marriage service provision), and social norms and practices (on child marriage, child rights, equal opportunities).

Evaluation Questions

In the following chapter, the evaluation approach is described. The core of the evaluation design is a quantitative, experimental approach using an RCT at two levels (Union Councils and villages) to address the main evaluation questions (described in Sections I. - I.B). The quantitative approach is complemented with a smaller qualitative research component (Section I.C). The following pages will address the strategies to systematically and rigorously answer the following evaluation questions (EQs):

EQ.1 *First and foremost, what evidence on IALP program implementation is captured and to which extent were similar activities implemented in non-program (control) areas?*

EQ.2 *What is the impact of the IALP program on the outcomes of interest (child marriage, child pregnancy, education, adolescent empowerment, communication, social norms, and service provision)?*

For each of the outcomes, further sub-hypotheses are defined by laying out the exact indicators that are looked at. In addition, the indicators are mapped against the original strategic goals set out by UNICEF. This supportive information is provided in Table O.3. The more refined information was logged in a PAP, which defined all outcomes indicators, before collecting the endline data, along with the overarching hypotheses in line with the evaluation questions displayed here (see Section O.I.A.2 for the more detailed PAP-logged original hypotheses guiding the evaluation design and survey tools development).

EQ.3 *Considering program-eligible villages of similar characteristics (so-called Non-Formal Basic Education (NFBE) eligible areas), which program element has a higher impact in reaching the goals, the Child Protection (CP) activities or the establishment of NFBE Centers (NFBECs)?*

EQ.4 *Considering program-eligible villages of similar characteristics, what is the joint effect of implementing Child Protection (CP) and Non-Formal Basic Educations (NFBEs) interventions? And is this effect larger than implementing the program activities independently?*

EQ.5 *Which targeting approach leads to the highest impacts for the outcomes of interest, the one focusing on boys/men, girls/women, or both genders?*

EQ.6 *How do the impacts differ by core characteristics, including gender, poverty level, implementation area (Sindh vs. Punjab), and self-reported COVID-19 exposure?*

> *Finally, the impacts are related to costs and the approach is described in Section I.D, where the methodological strategy to answer the following question is laid out:*
>
> *EQ.7 Do the benefits of the IALP program surpass its costs?*
>
> The hypotheses were registered in a PAP as estimation strategies and as presented in Section O.I.B.

I.A   Levels of Randomization and Treatment Arms

Child Protection-Level Randomization   The impact evaluation was initially set up for the activities implemented by the Child Protection (CP) Unit of UNICEF. In a first step, 139 treatment union councils (UCs) and 65 control UCs were selected. Thereby, of all the prioritized UCs in the district, 30% were randomly assigned to control, and 70% to treatment. In treatment and control UCs, 446 RVs in the treatment UCs and 207 RVs in the control UCs were sampled for this purpose.[1]

Additionally, UNICEF's Child Protection Unit implemented three modalities (treatment arms) of the Child Protection intervention: (a) A mixed-gender life-skills intervention targeting males and females (CPMF), (b) an intervention targeting primarily females (CPF), and (c) an intervention targeting primarily males (CPM). In other words, of the 446 Child Protection-treatment villages, 176 were assigned to treatment arm 1 (the mixed approach, CPMF), 141 were assigned to treatment arm 2 (female-only approach, CPF), and 129 were assigned to treatment arm 3 (male-only approach, CPM).[2]

> The three Child Protection treatment arms are based on their primary gender mobilization: Mixed, female-only, male-only.

NFBE-Level Randomization   Moreover, extending the Child Protection design, NFBE Centers were built. Their inclusion was integrated into the experimental design via a cross-randomization, on top of the Child Protection design described above. The NFBE Centers cross-randomization was performed in a way that it can measure the joint and also independent effects of Child Protection's and education section's activities. This, however, happened only in a non-random subset of Child Protection-eligible villages, considering particularly areas with high educational needs in Sindh. In other words, following pre-defined eligibility criteria, villages were excluded out of the original sample. Moreover, UNICEF implemented the NFBE Centers only in the rural districts of Sindh, Ghotki and Khairpur, which further reduces the NFBEC-eligible pool to 212 RVs.[3]

> The NFBEC-eligible pool refers to the 212 villages eligible for NFBEs Centers based on pre-defined criteria (e.g., worse educational outcomes).

Consequently, 212 villages are referred to as the NFBEC-eligible pool, whereas the other 441 villages are the NFBEC-non-eligible pool. These two different sets of villages differ systematically on various characteristics, because the 212 villages were purposefully chosen as those in need for additional education support. The randomization procedure, however, randomizes within these two strata and the impact evaluation of the treatment effects is therefore unbiased.[4] Hence, out of the 212 villages, 100 villages were randomly selected for the NFBE Center implementation, including an almost equal split for the Child Protection-treatment arms in NFBEC-eligible areas.[5] Thereby, in 40 RVs,

---

[1]All UCs were considered as per the census. From there, the eligible pool was then shortlisted in close consultation with UNICEF based on the criteria described in the text. Note, no random sampling was applied and the selected UCs are thus not representative for the entire districts. Instead, they were purposively selected based on the criteria described in Section II.B.1, reflecting UNICEF's expansion plans and programming needs. The final eligible sample for the study is also the eligible sample for the program (i.e., no further sampling was taken). The impact evaluation design has thus a very high internal validity. However, against original plans, UNICEF did not implement the program in urban areas and thus the urban areas were dropped from the endline sample in Summer 2020. The original sample was 761 RVs in rural and urban areas. The decision to not collect data was made after the official end of the project and before any data had been collected. It was based on the insight that program implementation was difficult in urban areas and thus stopped early. The reasons: First, overall project implementation was delayed and started only in 2018, not at the end of 2017 as originally planned. Generally, finding local partners for implementation proved difficult in urban areas for UNICEF. Finally, access to urban areas for program implementation was limited. Although reducing the sample size decreases power, leaving these areas in the sample while knowing that no implementation took place there would reduce the power of the analysis as well. It is emphasized that, following the decision to drop urban areas, any results of the analysis only apply to rural areas. In the remaining parts of this report, all numbers refer to the rural sample only. Original baseline tables and figures including the urban areas are presented in Appendix A.I.A.

[2]Originally, 40% of the villages were assigned to treatment arm 1, i.e., 208 RVs, 30% of the villages to treatment arm 2, i.e., 157 RVs, and 30% of the villages to treatment arm 3 i.e., 154 RVs.

[3]For a full list of eligibility criteria for the 212 villages, please refer to Table A.3.

[4]In other words, at baseline, the evaluation was based on a total 653 villages, which were stratified in two classes: A subset of 212 villages in which the NFBE treatment was allocated. And the remaining 653-212 =441 villages which were considered as not eligible for the NFBE treatment. Within these two strata, treatment arms were randomized. The two strata of 212 and 441 villages, however, differ in various characteristics.

[5]The allocation within those 212 villages is summarized in more detail in Table A.1 and Figure A.1 of the Appendix.

NFBE Centers were established in areas which were originally Child Protection-control areas in order to measure the pure impact of NFBE interventions.

This translates to 60 villages out of the 100 which received both types of interventions, and 40 RVs (in 17 UCs) from the original Child Protection control group which received only the education section's activities.

The final numbers of treated villages is 486 RVs with at least one IALP intervention and 167 RVs with no interventions, which corresponds to 108+17=125 treatment UCs and 50-17=33 control UCs.

Figure 2.1 below illustrates the randomization to the different treatment arms, where randomization was within strata, that is, within the 212 villages eligible for NFBE and the remaining 441 villages.

> The final number of treated villages is 486 villages with at least one IALP intervention and 167 villages with no interventions.

Figure 2.1: Impact Evaluation Design
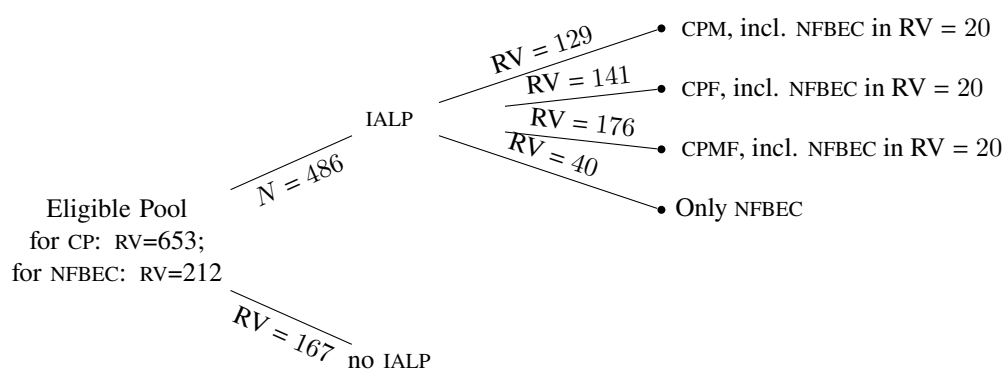
Note: Figure 2.1 displays the allocation of villages to treatment arms. For alternative, graphical presentations of the evaluation design content, see Appendix A.I.A.1.

Timeline   Figure 2.2 illustrates the timing of the assignment to the different treatment arms. The program implementation started at the end of 2017. In Sindh, implementation was most intense in 2018 and 2019. It ended by the end of 2019. In Punjab, the implementing partner was exchanged in September 2018.[6] Here, the implementation intensity was also most intense in 2018 and 2019. The most intense program implementation phase was in the two years 2018 and 2019. While the first wave of randomization compared the assignment of Child Protection and control areas, the later randomization added the evaluation of the NFBE activities in a subset of villages. Randomization for the Child Protection intervention happened in autumn 2015, following a baseline. Cross-randomization of NFBE interventions occurred in summer 2017. Both assignments were implemented prior to program implementation.

Figure 2.2: Timeline
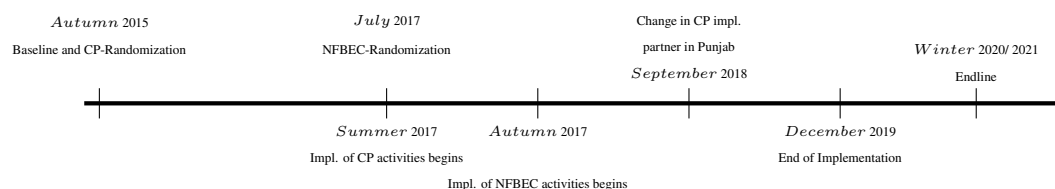
Note: Figure 2.2 displays the timing of the randomization, and of the endline survey. Due to the outbreak of the COVID-19 pandemic, the data collection could not take place as originally planned. The data collection started via phone for a subsample of all targeted respondents. See also Section II.B.3.

---

[6]The NGO CYAAD ended work in December 2017 and Bunyad started in September 2018.

Table 2.1: Balance Table

Variable	(1) Control		(2) CP, not NFBE pool		(3) CP, NFBE pool		(4) CP and NFBE		(5) only NFBE		(6) Normalized Differences (1)-(2)	(7) Normalized Differences (1)-(3)	(8) Normalized Differences (1)-(4)	(9) Normalized Differences (1)-(5)
	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE				
RV: Share adolescent (A) married	167	0.062	303	0.066	83	0.063	60	0.054	40	0.062	-0.046	-0.009	0.127	0.012
	[49]	[0.006]	[91]	[0.005]	[35]	[0.008]	[34]	[0.008]	[17]	[0.012]				
RV: Share A engaged	167	0.091	303	0.085	83	0.130	60	0.115	40	0.146	0.060	-0.412	-0.257	-0.567
	[49]	[0.008]	[91]	[0.006]	[35]	[0.013]	[34]	[0.016]	[17]	[0.019]				
RV: Share A pregnant under 18	167	0.122	303	0.131	83	0.127	60	0.114	40	0.123	-0.088	-0.042	0.097	-0.001
	[49]	[0.010]	[91]	[0.007]	[35]	[0.012]	[34]	[0.011]	[17]	[0.014]				
RV: Share A with any education	167	0.262	303	0.242	83	0.261	60	0.250	40	0.259	0.098	0.006	0.060	0.017
	[49]	[0.026]	[91]	[0.019]	[35]	[0.018]	[34]	[0.025]	[17]	[0.024]				
RV: ∅ approval child protection for girls	167	0.600	303	0.590	83	0.595	60	0.629	40	0.570	0.075	0.039	-0.221	0.225
	[49]	[0.014]	[91]	[0.011]	[35]	[0.023]	[34]	[0.021]	[17]	[0.022]				
RV: ∅ approval child protection for boys	167	0.490	303	0.478	83	0.461	60	0.492	40	0.469	0.087	0.222	-0.023	0.166
	[49]	[0.014]	[91]	[0.012]	[35]	[0.017]	[34]	[0.017]	[17]	[0.021]				
RV: Share in favor of child marriage (CM)	167	0.440	303	0.436	83	0.442	60	0.476	40	0.428	0.023	-0.014	-0.232	0.078
	[49]	[0.016]	[91]	[0.010]	[35]	[0.025]	[34]	[0.027]	[17]	[0.031]				
RV: Share see no risk in CM	167	0.252	303	0.261	83	0.335	60	0.290	40	0.304	-0.046	-0.407	-0.190	-0.278
	[49]	[0.019]	[91]	[0.014]	[35]	[0.034]	[34]	[0.038]	[17]	[0.036]				
RV: Share informed about child right (CR)	167	0.305	303	0.304	83	0.248	60	0.312	40	0.268	0.007	0.261	-0.030	0.170
	[49]	[0.028]	[91]	[0.020]	[35]	[0.024]	[34]	[0.033]	[17]	[0.043]				
RV: Share advocating CR	167	0.189	303	0.190	83	0.223	60	0.163	40	0.240	-0.006	-0.167	0.126	-0.254
	[49]	[0.017]	[91]	[0.015]	[35]	[0.024]	[34]	[0.030]	[17]	[0.040]				
RV: Share parent (P) allow dau. go to doctor	167	0.232	303	0.244	83	0.194	60	0.206	40	0.195	-0.050	0.161	0.113	0.156
	[49]	[0.024]	[91]	[0.021]	[35]	[0.029]	[34]	[0.032]	[17]	[0.041]				
RV: Share household (HH) with P preventing children's education	167	0.080	303	0.093	83	0.127	60	0.156	40	0.138	-0.102	-0.324	-0.516	-0.445
	[49]	[0.012]	[91]	[0.009]	[35]	[0.021]	[34]	[0.028]	[17]	[0.029]				
RV: Share mobile phone access	167	0.322	303	0.301	83	0.264	60	0.298	40	0.314	0.084	0.238	0.097	0.033
	[49]	[0.025]	[91]	[0.020]	[35]	[0.024]	[34]	[0.031]	[17]	[0.034]				
RV: Share watch TV	167	0.464	303	0.454	83	0.509	60	0.563	40	0.570	0.036	-0.169	-0.376	-0.402
	[49]	[0.033]	[91]	[0.020]	[35]	[0.029]	[34]	[0.031]	[17]	[0.049]				
RV: Share know legal marriage age	167	0.199	303	0.164	83	0.154	60	0.210	40	0.227	0.126	0.154	-0.037	-0.095
	[49]	[0.032]	[91]	[0.020]	[35]	[0.037]	[34]	[0.050]	[17]	[0.071]				
RV: Share HH head with no education	167	0.617	303	0.623	83	0.685	60	0.661	40	0.746	-0.020	-0.231	-0.149	-0.426
	[49]	[0.038]	[91]	[0.028]	[35]	[0.031]	[34]	[0.034]	[17]	[0.049]				
RV: ∅ poverty score	167	26.643	303	27.762	83	21.881	60	21.833	40	22.940	-0.124	0.577	0.566	0.428
	[49]	[1.254]	[91]	[0.905]	[35]	[0.652]	[34]	[0.798]	[17]	[1.234]				
RV: Share part. any event	167	0.124	303	0.100	83	0.128	60	0.226	40	0.164	0.139	-0.024	-0.550	-0.216
	[49]	[0.021]	[91]	[0.012]	[35]	[0.024]	[34]	[0.027]	[17]	[0.036]				
RV: ∅ # CRs known	167	0.250	303	0.262	83	0.182	60	0.198	40	0.169	-0.089	0.492	0.365	0.555
	[49]	[0.020]	[91]	[0.012]	[35]	[0.010]	[34]	[0.017]	[17]	[0.017]				
RV: Share Nikah Khwan (N) refuse marriages below legal age	167	0.353	303	0.386	83	0.331	60	0.267	40	0.250	-0.068	0.046	0.184	0.219
	[49]	[0.044]	[91]	[0.036]	[35]	[0.060]	[34]	[0.067]	[17]	[0.090]				
RV: Share N name benefit of CM	167	0.156	303	0.139	83	0.199	60	0.167	40	0.075	0.049	-0.115	-0.030	0.232
	[49]	[0.038]	[91]	[0.024]	[35]	[0.047]	[34]	[0.045]	[17]	[0.040]				
RV: Share N identify as Imam	167	0.503	303	0.455	83	0.705	60	0.692	40	0.775	0.096	-0.408	-0.379	-0.546
	[49]	[0.048]	[91]	[0.035]	[35]	[0.052]	[34]	[0.067]	[17]	[0.067]				

▶ *Note*: Table 2.1 compares treatment and control areas. Each indicator is a revenue villages average measured at baseline.
▶ Sample: All villages.
▶ Squared brackets indicate number of UCs, N refers to the number of villages; The mean and standard errors are reported for a number of outcomes at baseline. Standard errors are clustered at the UC level and unreported in the table is the control variable whether the village is eligible for the NFBE program activities. Columns (6) to (9) report the normalized differences between mean values.

Comparing IALP and Control Villages Prior to Implementation   In how far were treatment villages comparable to the control villages before the onset of any program activities? Table 2.1 compares baseline survey characteristics of each of the two different program activities, Child Protection and NFBE activities, to the averages in the control group (column (1)). Notably, given that NFBE activities were implemented only in a subset of villages with higher educational needs, while Child Protection was randomly assigned across all villages, there are two versions of Child Protection villages: (a) *Outside* of the NFBE-village pool (with the characteristics being reported in column (2)), and (b) *inside* this pool (column (3)). Column (4) reports the averages in villages where later both types of interventions were implemented, and finally in column (5) the averages where only NFBE interventions were rolled-out are reported. For instance, it is observed that the average share of adolescents married at baseline is about 6.2% in the control group. The share is not statistically different from the prevalence in Child Protection villages (with 6.6 and 6.3% in columns (2) and (3), respectively) or NFBE villages (with and without Child Protection interventions; with 5.4% and 6.2% on average). To compare the means to the t-statistic, it is controlled for eligibility to the NFBE program (unreported in the table) and the standard errors are clustered at the UC-level (the first level of randomization). Across almost all combinations in which the treatment arms vs. control group are compared, little statistical differences in pre-treatment characteristics are observed.[7] Differences between the groups of villages are considered in the final estimations of the program impacts through conditioning on most disbalanced baseline characteristics. While baseline differences are important, note that there is no reason to believe that they occurred systematically.[8]

I.B   Estimation Strategy

In order to estimate the treatment effects, the Intention-to-Treat Effect (ITT) is estimated. The ITT compares the outcomes of individuals from the experimentally assigned treatment group with the outcomes from the control group. As mentioned before, two main treatments were randomly allocated: The Child Protection and the NFBE Centers, the latter being restricted to two districts and a set of pre-specified criteria.[9]

Strata   The empirical analysis proceeds within *strata*, i.e., via interacting with the stratum indicator (NFBE-eligible-pool).[10] Thereby, the evaluation relies on six different groups in total. In those villages which were classified as not eligible for NFBE, there are two groups: Those villages assigned to Child Protection and those not. In those villages classified as eligible for NFBE, there are four groups: Villages participating in the Child Protection and in the NFBE program, those villages that participate only in Child Protection, those villages that participate only in NFBE and those villages that receive neither. There are thus six non-overlapping groups in total (the six numbered cells in Table 2.2). The broad intuition is that the impact of the Child Protection intervention is estimated by comparing outcomes after the program ended in program assigned areas (CP=1) with control areas (CP=0). Thereby the impact of the Child Protection interventions within the NFBE-eligible-pool (NFBE_pool=1) and outside of the NFBE-eligible-pool pool (NFBE_pool=0, Table 2.2) are considered. Accordingly, the impacts of the areas assigned to the NFBE intervention (NFBE=1) are compared to areas that were not (NFBE=0), within the pool of NFBE-eligible areas (NFBE_pool=1). In 60 villages, the Child Protection and NFBE intervention were jointly implemented (cell (6), , Table 2.2).

Following the allocation of treatment arms displayed in Table 2.2, the regression model also contains 6 coefficients (including the constant) and all regressors are binary (as summarized in the estimation approach displayed in equation 2.1). Econometrically speaking this means that the regression model estimates are identical to differences between the outcome averages of these six groups. This means, even though a linear regression model is employed, it obtains exactly the same estimates as it would have obtained had the village averages been calculated. The regression model,

---

[7]Table A.2 presents the balance statistics by Child Protection treatment arms, with no ex-ante evidence of systematic differences between the different types of villages.

[8]As shown in the Appendix, the first round of randomization of Child Protection activities was well-balanced. This is when the original sample is considered, which included urban areas, but also when these were later dropped. Please refer to Section O.I.D for more information.

[9]Note, the evaluation does not assess the ToC (e.g., its relevance, coherence, validity).

[10]This pool of villages defines the communities eligible for educational interventions due to their particularly low educational outcomes in the baseline survey. Moreover, the program targeted only areas in Gothki and Khairpur.

Table 2.2: Program Effects

CP	non eligible for the NFBE intervention (441 villages) NFBE_pool=0 (non_NFBE_pool)	eligible for the NFBE intervention (212 villages) NFBE_pool=1 (NFBE_pool)	
		NFBE=0	NFBE=1
	1: control	2: control	3: receive only NFBE
CP =0	CP=0, NFBE_pool=0 (ass_control_not_NFBE_pool)	CP=0, NFBE=0, NFBE_pool=1 (ass_control_NFBE_pool)	CP=0, NFBE=1, NFBE_pool=1 (ass_only_NFBE)
	estimation outcome: α β_5, , villages: 138	estimation outcome: α, villages: 29	estimation outcome: α β_1, villages: 40
	4: receive only CP	5: receive only CP	6: receive CP and NFBE
CP =1	CP=1, NFBE_pool=0 (ass_CP_not_NFBE_pool)	CP=1, NFBE=0, NFBE_pool=1 (ass_CP_NFBE_pool)	CP=1, NFBE=1, NFBE_pool=1 (ass_treat_CP_NFBE)
	estimation outcome: α β_5 β_4, villages: 303	estimation outcome: α β_3, villages: 83	estimation outcome: α β_2, villages: 60

▸ *Note*: Table 2.2 describes the sample size by program activities and summarizes the corresponding coefficient names as defined in equation 2.1.

however, allows for also calculating other statistical properties and adjusting for covariance patterns.

$$
\begin{aligned}
Y = \alpha\ &\beta_1 \cdot NFBE\_only \\
&+\beta_2 \cdot NFBE\_and\_CP \\
&+\beta_3 \cdot CP\_only \cdot NFBE\_pool\_1 \\
&+\beta_4 \cdot CP\_only \cdot NFBE\_pool\_0 \\
&+\beta_5 \cdot NFBE\_pool\_0 \\
&+controls+\epsilon
\end{aligned}
\tag{2.1}
$$

In more detail, Y in equation 2.1 refers to the main outcome variables described in the subsequent sections, e.g., a set of different indicators measuring for example marriage prevalence, school attendance, pregnancies, empowerment and social practices. The respective outcome indicator is measured via endline survey, i.e., at $t = 1$. The measures are either derived from responses in a household roster information or from targeted, program-eligible respondents.

Intention-to-Treat Effects   The coefficient β_1 measures the effect of NFBE for the 212 NFBE-eligible villages. Since the NFBE intervention was randomly allocated only within the stratum of 212 villages, the impact of NFBE can thus only be estimated for this stratum. For the other stratum of 441 villages, the treatment effects cannot be estimated. Similarly, the coefficient β_2 measures the effect of being allocated to the NFBE and the Child Protection intervention together. Again, this effect can only be estimated for the 212-villages stratum. In contrast, the effects of Child Protection intervention alone can be estimated for both strata. For ease of comparison, the effects of Child Protection are estimated and reported separately for each stratum. The coefficient β_3 thus measures the effect of Child Protection in the subset of 212 villages, whereas β_4 measures the effect of Child Protection in the subset of 441 villages. Thus, the estimated effect is likely to be more precise as it is based on larger sample size. And since the villages to be considered for the NFBE intervention were not randomly selected, there is no reason to expect these two coefficients to be identical since the effects of Child Protection might depend, e.g., on village characteristics (effect modifier).

Weighting   Now the effect sizes estimated in equation 2.1 are adjusted considering that the sample sizes within the six strata differed, i.e., for the fact that the Child Protection effect size estimated in 212 Non-Formal Basic Education-eligible villages may differ from the effect size in 441 Non-Formal Basic Education-non-eligible villages. Hereby, the number of villages in each stratum that actually was allocated to Child Protection is used as weights, that is, 83 and 303 villages. Consequently, as a summary measure of the effects of Child Protection for both strata (the average effect of Child Protection), the weighted effect is calculated as follows:

β_3 as the effect of Child Protection for the 212 villages;

β_4 as the effect of Child Protection for the 441 villages.

One needs to note that for statistical reasons, the uncertainty around β_3 will be larger due to the smaller sample size. In order to report more precise estimates, the weighted average of these effects is shown as

$$
average \text{ effect of CP} = \frac{83}{386}\beta_3 + \frac{303}{386}\beta_4.
\tag{2.2}
$$

   The evaluation applies an analogous approach for the overall effect of IALP below. The overall effect of the IALP in Pakistan is:

$$average \text{ effect of IALP} = \frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4. \tag{2.3}$$

This is the estimated effect for a randomly selected observation from those villages which were randomly allocated to UNICEF interventions.

Covariates   The coefficient β_5 measures the observed differences in outcomes between the villages classified as NFBE-eligible and the NFBE-non-eligible. β_5 does not measure any treatment effect (i.e., has no causal meaning), but simply controls for the fact that the 212 villages were selected based on their higher levels of educational needs. Moreover, the regression model employed for the impact evaluation includes further, additional control variables.[11] For instance, given that different respondents provide the responses to the questions, the respondent type (Res^T) is controlled for, i.e., whereby different indicator variables differentiating between the four main types of respondents (a community leader and/ or a Nikah Khwan and/ or a parent of an adolescents and/ or an adolescent) are added.[12] A set of further control variables in the robustness section includes a vector of the baseline covariates at the RV level, i.e., variables that were used for the randomization (Table 2.1).[13] Additionally, the evaluation considers basic exogenous individual-level variables (gender and age) and geographical information variables (district fixed effects). The evaluation also takes into account enumerator fixed effects and allows the *error term to be clustered at the level of union councils*.[14]

Further Sample Restrictions   For the estimation of program impacts on the educational outcomes, the analysis is restricted to children i who were eligible for the NFBE Center program. This captures whether adolescents have either been never enrolled in school or have dropped out of school.

Additional Treatment Arms Capturing Experimental Program Variation in Targeting Approaches   Finally, it can be distinguished between the various Child Protection treatment arms, i.e., CPM, CPF, and CPMF. Accordingly the Child Protection treatment group indicators are CP^F for areas randomly assigned to mobilize more females, CP^M for more males, and CP^{MF} for the equal mobilization of both genders. To estimate the treatment effects, CP in the equation is replaced with the three treatment arm indicators, CPM, CPF, and CPMF, with the control group as the reference category. This enables to better understand the importance of focusing on one gender, e.g., mobilizing boys/ men (by looking at the coefficient of CPM). The analysis also tests whether the composition of the groups matter, i.e., whether the effect of the program differs depending on whether more female, more male, or an equal numbers of both were mobilized for the program. For instance, the intervention's impact may unfold itself by altering the attitudes of male in particular. Then, one would expect a higher coefficient for CPM compared to the other treatment arms.

> **Interpretation of the Results**
>
> This evaluation focuses on the so-called Intention-to-Treat (ITT) effects when discussing the results. Hereby, de facto, the effects of an adolescent's community being randomly assigned to the program, or not are measured. This approach ignores the influence of factors that happened after the random assignment to the program, such as non-delivery or partial delivery of the program, non-compliance, withdrawal, or a general lack of program uptake, as well as deviation from the original assignment (such as program implementation in control communities). Therewith more conservative, yet also more credible (unbiased) estimates for the impacts are received, given that a lack of program implementation or deviations from the program are non-random and thus a relationship may exist between the selected program delivery/ uptake and the outcomes of interest (e.g., a higher level of empowerment would be captured not only because of the program, but also because it was potentially selectively delivered to more empowered individuals to start with). Due to non-compliance, the ITT effects remain an underestimate of the true effect (Angrist, Imbens, and Rubin 1996). In other words, the effects on actual direct program participants/ beneficiaries (such as adolescent champions or NFBE participants) are certainly larger in size. Therefore, it is recommended to read and interpret the results of this evaluation as being more conservative, yet easier generalizable for policy purposes. What needs to be kept in mind is the community's eligibility criteria (original characteristics) for the original program assignment (i.e., the

---

[11]To simplify the intuition for the estimation process, consider first the regression model without any additional control variables. While the linear regression model implemented also adds linear controls for other covariates, this does not change the interpretation of the coefficients.

[12]In a set of heterogeneity analyses the regressions are run for each type separately.

[13]In situations where collinearity or near-collinearity does not permit to linearly control for baseline covariates, e.g., in case of subgroup analysis in small subsets and if substantial non-response occurred during endline, randomization inference without the introduction of controls is applied.

[14]Please see Section A.IV.A for robustness results with the error term clustered at the village level, a multi-level clustered error term (combination of village and union council) and the error term clustered at the household level.

> type of communities that were in the original pool to be either assigned to the treatment or control group).

Local Average Treatment Effect (LATE)   The ITT provides a measure of the overall effect of the IALP intervention on all the villages that were assigned to the program. One important implication is that the ITT is low if, for some reason, the program did not reach the beneficiaries. Thus, with the help of a two-stage least squares (2SLS) treatment-on-the-treated estimate using the Wald estimator the LATE is identified, i.e., the effect on those individuals and villages for whom evidence of actual participation or take-up following the random assignment is found. In some interventions the discrepancy between the intended treatment status and actual one can be large for various reasons.[15] Substantial differences between ITT and LATE could reveal that either the treatment is correlated with unobserved characteristics, and some of the potential beneficiaries did not find it attractive to participate in the program (compliance depends on unobservable traits), and/ or suggest that the intervention could not reach a significant share of the beneficiaries. For the estimation, MIS information on implementation is considered. Moreover, the endline survey contains information on whether the individuals were reached by the program but did not like to uptake it or whether there is no evidence that they have ever heard about the program. The LATE estimate is discussed in more detail in Section II.C.

I.C   Mixed Methods

A smaller qualitative part complements the findings from the quantitative impact evaluation. To gather qualitative information, 16 key informant interviews (KIIs) using open questions were conducted with stakeholders of the IALP program. These include UNICEF and government officials, as well as staff of IPs from both provinces, Sindh and Punjab. KIIs provide insight into the experience of key informants (KIs) with designing and implementing, as well as following up on IALP activities. Using the MAXQDA software, a qualitative content analysis is conducted. For this, codes are drawn mainly deductively from the Theory of Change (ToC) and the qualitative evaluation matrix, as well as inductively based on insight from the transcribed KIIs. The focus in the analysis of the KIIs lies on challenges faced during implementation, perceived consequences of the IALP program, and recommendations by stakeholders for future programs that might plan to employ a comparable strategy or pursue a similar aim. The goal is to triangulate findings from the quantitative and qualitative part of the evaluation in an effort to provide a more comprehensive picture of the IALP program. Information from KIIs is hereby used to gain insight into results of the quantitative analysis that seem puzzling or warrant an explanation. The information collected through KIIs is further triangulated with secondary data drawn from reports examined during document review.

I.D   Cost-Benefit and Cost-Efficiency

For the cost-benefit analysis and cost-efficiency analysis of the IALP program, the monetary impact of the implemented program is calculated. The results of this analysis have important implications for policymaking. In particular, finding that the benefits surpass the costs would suggest that extending this program would have an overall positive impact. However, finding the opposite would not lead to the contrary conclusion since only short-term monetary benefits are looked at. Guiding this exercise is the work by Field, Glennerster, Buchmann, and Murphy (2016) who conduct a study titled "Cost-Benefit Analysis of Strategies to Reduce Child Marriage in Bangladesh".

> The primary goal of conducting **cost-benefit analysis** is to make program activities comparable. Once the problems to be tackled and goals of the program are defined and the impact evaluation is conducted, the impacts are related to the costs of the program to make comparisons with alternatives possible. Thereby, it is explicitly acknowledged that programs often produce outcomes that are not sold on markets and do not generate direct revenues, such as is the case in empowerment and educational programs. In this case, the impacts of the program need to be translated into "benefits" if they were directly or indirectly positive, or, in case they were harmful, into "cost". This evaluation focuses on marriage benefits, whereby the delays in marriage are translated into additional years of education and therewith potential future income effects - all these steps are made under a set of assumptions and considering discount factors. Additionally, the costs of implementing the program (administrative costs, inputs, etc.) are considered, which requires an extensive costs analysis. In this evaluation, it involved a review of the costs at UNICEF and its IPs. Eventually, the costs are related to the benefits of a particular program. For the cost-effectiveness analysis, the costs are related to the effects as a ration: $\frac{costs}{effects}$ (Asian Development Bank, 2013; Field et al., 2016).

---

[15]For example, not all of the targeted program participants actually participated in the program, administrative reasons might have impeded them from participating or there was selective migration due to nonconformity with the treatment status. For a detailed explanation see Gertler, Martinez, Premand, Rawling, and Vermeersch (2016).

Estimating Benefits   The analysis captures a **lower bound of the actual benefits**. In fact, if the program is successful, an increase in knowledge about rights and education are expected to have a beneficial impact for many months and years to come. Finally, only monetary outcomes are measured. Thus, positive but non-monetary outcomes such changes in life satisfaction are excluded from the analysis. For the analysis, program costs and benefits are brought to a common unit (USD).

In more detail, program benefits are calculated based on the rate of return to additional education induced by the IALP program and being projected over a lifetime.[16] The idea is the following: The program increases the adolescent age of marriage (reduces the child marriage rate). These additional years that adolescents are not married, adolescents are assumed to spend in school which would translate to additional years of education. To calculate the education premium in terms of additional years of schooling, every percentage point reduction in child marriage is multiplied with the conversion factor 0.04 as suggested in Field et al. (2016).[17] The conversion factor is applied because no statistically significant program impacts on years of schooling are found in this analysis. Column (2) in Table 2.3 reports the result: 0.8 percentage point reduction in child marriage[18] multiplied by conversion factor 0.04. To put a monetary value to these benefits, Field et al. (2016) suggest using a modified Mincer equation to estimate adolescents' expected wages in each year of their working life as a function of education and experience (see Section A.IV.D). The underlying data source for this estimation is the Labor Force Survey Pakistan from 2010-11 (Federal Bureau of Statistics 2011). The estimation reports an average rate of return for education of 6.3%.[19]

To calculate annual benefits, the results from a Mincer regression (Table A.61) are used: The income for adolescents receiving the program (adolescents from program-assigned areas) and not receiving the program (adolescents from control area) is estimated. Table 2.3 reports the results. The estimated median numbers for non-treated areas of years of education is 5 (column (3)) and of potential experience is at 14 years (column (4)).[20] Column (5) shows both annual incomes with regards to the median levels of education and experience. Non-program participants have an annual estimated income of 851.594 USD, while program participants have 852.693 USD. The difference in income between the individuals in program-assigned and those from non-assigned areas is caused by the education premium, which is added to the years of education in column (3), and subtracted from the years of experience for adolescents in the program-assigned areas (in column (4)). The present value lifetime income in USD is reported in column (6), with 8031.771 for non-program participants and 8038.95 for program participants. The aggregation of annual income to lifetime income requires the following assumptions. First, constant rate of returns to education over working ages are assumed. Similar to Field et al. (2016), working ages are defined as beginning at the age of 17 and ending at the age of 60, which is the age of pension entry according to the civil servants pension scheme in Pakistan (WB 2020). Second, equal returns of education for adolescents who enter the workforce upon graduation and for adolescents who stay absent from the workforce are assumed.[21] Third, it is assumed that an extension in education delays entry into the workforce and that people start working right after finishing their education.[22] Finally, increasing wages are accounted for by multiplying with the GDP per capita growth rate.[23] Moreover, the evaluation applies a Social Opportunity Cost of Capital (SOC) of 12% for Pakistan to discount the two cost streams to the present value at the base year of implementation (2017), which allows to account for beneficiaries' time preference (Zhuang, Liang, Lin, and De Guzman (2007)). To recover the annual benefits, the difference of both present values (PVs) are taken and inflated to the year of analysis 2021 USD using the average annual USD inflation rate. Finally, the benefits are multiplied by the number of beneficiaries of the IALP program. Using the sum of directly reached adolescents in Pakistan reported by UNICEF results in 269,143 adolescents (UNICEF 2020). Columns (7) and (8) in Table 2.3 show the individual and aggregated benefits of the program, which are the differences between the two income streams.

---

[16]The difference between the median lifetime income and the lifetime income with additional years of education are calculated.

[17]In other words, a reduction in child marriage is assumed to go in hand with an increase in additional years of schooling.

[18]The information will be taken from the **ITT ! (ITT !)** results in Table 3.7, Panel A, column (7), coefficient for the IALP overall effect.

[19]Table A.61 in Section A.IV.D in the Appendix shows the coefficients corresponding to the Mincer equation 1. This estimate is in line with the estimates by Montenegro and Patrinos (2014) who find an average rate of return for an additional year of schooling in Pakistan for the years 1992 to 2010 ranging from 6.3% to 10.8%.

[20]Estimates from the Labor Force Survey (Federal Bureau of Statistics 2011).

[21]This enables applying the rate of return to the entire population and disregarding differences between people in the workforce and the unemployed.

[22]Year 2018 is defined as the year where beneficiaries enter the workforce. The UNICEF MIS reports that beneficiaries were 16 years old on average at the beginning of implementation in 2017. The average marriage age is defined as the age of entry to the workforce, which was 17 in the endline data collection. Hence, 2018 is chosen as the year where the average beneficiary turns 17 and enters the workforce by assumption. 2018 is therefore the base year of the income streams used for the calculation over the working lives of beneficiaries.

[23]Annual GDP per capita growth rates for Pakistan are taken from the World Development Indicators (WB 2021). The actual rates for the years from 2018 to 2020 and the ten-year average growth rate for future growth rates is used. Equations 2 and 3 in the Appendix illustrate both income streams corresponding to the specifications in Field et al. (2016).

Table 2.3: Benefit Calculation

Program received	Education Premium	Years education	Years of experience	Estimated annual income in USD	PV lifetime income in USD	Individual benefit in USD	Aggregated benefit in USD (269,143 beneficia-ries)
No	0.032	5.000	14.000	851.594	8,031.771	9.184	2,471,962.808
Yes		5.032	13.968	852.639	8,040.956		

► *Notes*: The table outlines the benefit estimation for the IALP program. Column (2) reports the education premium in form of additional years of schooling (calculation: 0.8 (pp reduction in CM) * 0.04). Columns (3) and (4) show the median years of education and experience that predict annual earnings in column (5) using the Mincer estimation. Column (6) shows the discounted lifetime income in terms of 2021 USD. Columns (7) and (8) show the individual and aggregated benefits of the program, which are the differences between the two income streams.

Estimating Costs   Social costs are calculated, i.e., costs to the implementer and costs to the beneficiaries. Moreover, costs incurred to the beneficiaries are considered as the opportunity costs (counterfactual income, i.e., outcome for the control group).

To estimate the **costs to the implementer**, the gross-estimations of costs are calculated using total program expenditure of each treatment, i.e., the NFBE and Child Protection activities, which include expenses such as offices space, program management, etc. To obtain a comparable cost specification, all recorded expenditures are converted with regards to inflation, exchange rates, and the year of program implementation. Hereby, the recommendations made in Dhaliwal, Duflo, Glennerster, and Tulloch (2013) are followed.[24]

When looking at costs of a program to society as a whole, costs to beneficiaries need to be considered. Beneficiaries, who were assigned to the program, benefit from extended years of education. However, it is assumed that they simultaneously forgo potential work experience which is associated with income (Table A.61).[25] To estimate the **costs to the beneficiaries**, the idea is thus to estimate the costs of remaining in school for longer instead of earning money and gaining working experience in the meantime. Similar to the estimation of benefits, two income streams are calculated using again the Mincer equation. The education premium induced by the program is considered as the loss of such alternative experiences. Table 2.5 outlines these calculation with the assumptions made for estimating the benefits. The results of cost-benefit analysis will be presented in Section II.D.

---

[24]Table A.62 in the Appendix outlines the present value of total costs in terms of 2021 USD (Section A.IV.D).

[25]In other words, the difference in income caused by an adolescent having less experience over their lifetime are the opportunity costs for beneficiaries by participating in the program (Field et al. 2016).

Table 2.5: Costs to Beneficiaries Calculation

Program received	Education Premium	Years education	Years of experience	Estimated annual income in USD	PV lifetime income in USD	Individual costs in USD	Aggregated costs in USD (269,143 beneficiaries)
No	0.032	5.000	14.000	851.594	8,031.771	7.092	1,908,970.321
Yes		5.032	13.968	850.913	8,024.678		

▶ *Notes*: The table outlines the costs to beneficiaries estimation for the IALP program. Column (2) reports the education premium in form of additional years of schooling (calculation: 0.8 (pp reduction in CM) * 0.04). Columns (3) and (4) show the median years of education and experience that predict annual earnings in column (5) using the Mincer estimation. The difference in annual income is predicted by the experience beneficiaries lack by participating in the program. The underlying assumption is that the extended length of education induced by the program causes experience to decrease. Column (6) shows the discounted lifetime income in terms of 2021 USD. Column (7) and (8) show the individual and aggregated costs to beneficiaries of the program, which are the differences between the two income streams.

II. Data

II.A Outcomes

In this section, the indicators that are used in the analysis are shortly introduced.

Main Outcomes   The outcome indicators for this evaluation were developed along the key performance indicators and reflect the program's theory of change. The evaluation reports outcomes for the following topics, wherever meaningful reverting back to using summary scales: First, the evaluation reports the impacts on the three core topics - child marriage, child pregnancy, and education; second, it reports on adolescent empowerment, communication, and social norms; third, it captures information on program delivery and service delivery. The next section reports the values in the overall sample as captured in the endline survey, without distinguishing between program and control areas.

Measuring the Strategic Outcomes

The main three strategic outcomes - marriage, pregnancy, and education - are measured by referring to information captured in the <u>household roster</u>: The interviewers asked the household head or, in case of absence, a knowledgeable adult to provide answers on the whole household by answering a household roster. In the analysis, the information gathered is then restricted to the relevant age group.

- **Marriage**: To capture the child marriage incidence, the household heads were asked about the age of all household members, and, few questions later, whether this household member was married. See Table 3.7 for the main marriage results.

- **Pregnancy**: In order to capture the prevalence of child pregnancy, the household heads were asked about the age of all household members, and, few questions later, whether this household member has any children to whom she has given birth, conditional on them being a female and being married, widowed, or divorced. The interviewers additionally asked about the age of the oldest child they have given birth to in order to capture age at first child. See Table 3.9 for the main pregnancy results.

- **Education**: To capture educational information, the interviewers asked the household heads about the type of educational institutions all household members had ever attended. See Table 3.10 for the main education results.

Intermediate Outcomes   Intermediate outcomes are potential drivers of change. Following the program's theory of change, they capture factors that were envisioned to become mechanisms of transformative processes. They are more closely and directly related to the program activities, capturing change in well-being, levels of protection, attitudes, gender norms, actions taken to support adolescents, and general interactions on program-related outcomes.

Qualitative Outcomes   Using MAXQDA, a qualitative content analysis of the transcribed and translated KIIs is conducted. Qualitative content analysis here refers to the "subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns" (Hsieh and Shannon, 2005, p. 1278). In line with direct content analysis, the codes are developed mainly deductively by drawing from the ToC as well as from the qualitative evaluation matrix. Additionally, some categories are generated inductively based on insight from transcribed interviews.[26]

II.B Large-Scale Data Collections

In this section, the geographical coverage of the impact evaluation as well as the baseline and endline data collection are first briefly described, highlighting COVID-19-related changes. These changes include the split of the endline data collection into two waves. In the first wave, community representatives were interviewed via phone. In the second wave, adolescents, parents, and marriage registrars were surveyed in person. All changes are outlined in greater detail in Section II.B.3 below.

---

[26]All interviews were coded by a coder of the Center for Evaluation and Development (C4ED) office in Mannheim in close cooperation with the Senior Research Manager in Pakistan who conducted the interviews.

II.B.1   Geographical Coverage

The sampling universe covered all program areas: Two districts in Punjab province, Bahawalpur and Rahim Yar Khan, and two districts in Sindh province, Ghotki and Khairpur.[27] *Following, within these districts, UCs were first purposively selected.* UCs were selected following several criteria that had been jointly agreed upon by UNICEF and relevant provincial government stakeholders. The selection criteria for programming were: (1) High prevalence of child marriage in Punjab (according to Multiple Indicator Cluster Surveys (MICS) 2011 statistics), (2) achieving a balance between geographic contiguity for easier operation and the effort to minimize spillovers by introducing buffer UCs, (3) considering security constraints incorporating local knowledge provided by UNICEF field staff and local NGOs, and (4) covering 30% of the rural population (i.e., considering population size).[28] To sum up, UCs were selected in such way that they would cover a large part of the population, with a focus on high child marriage rates, and accessibility for effective implementation. At this stage, only UCs were included which UNICEF would consider for their interventions. Prior to the baseline, UNICEF had implemented earlier adolescents interventions in these districts and the program studied here was part of an "expansion program" of UNICEF. Since UNICEF funds were insufficient to cover the entire population, the evaluation team coordinated with UNICEF program team planning the expansion of the adolescents program.[29] Finally, it was aimed to choose non-adjacent UCs in order to have larger coverage within a Tehsil and reduce program spill-overs. *In the second stage, revenue villages were sampled.* This second stage was pure random sampling of villages, where four villages were randomly selected in smaller UCs and five villages were randomly selected in the larger UCs.[30] For these sample revenue villages, baseline data was collected. The random allocation to treatment and control was done at a later stage by a computer program of the authors of this evaluation. This random allocation process was completely independent of UNICEF.[31] On average, about 4.2 villages per UC are considered in the final estimation sample of this evaluation. Figure 2.3 shows the geographic location of treatment and control villages in the sample.

II.B.2   Baseline Data

Timing of the Survey   The baseline data collection (including two pilots) took place between June 2015 and December 2015.[32]

Target Group   The selection of respondents reflects the three-pillar structure of the program and is consistent with the multi-stakeholder approach of the program, including adolescents as the rightholders, parents of adolescents, community representatives and marriage registrars as duty-bearers.

Respondents Sampling Strategy   In each village four adolescents, four parents, one community leader, and at least one Nikah Khwan (marriage registrar) were interviewed during the baseline data-collection. Within each village, further selection criteria were set to prevent under-reporting and under-sampling of some group(s) of interest. The evaluation sets additional targets for the sample: (1) Interviewing 50% female and 50% male adolescents since the program focuses on both female and male adolescents; (2) applying poverty score card and considering 12.5 acres of land-ownership as an upper limit, the sample had to cover 50% of the most vulnerable (poor) and 50% not so poor households (excluding the richest households), (3) an additional, yet not binding, instruction was to cover all age groups (12-19) in the adolescents sample. Parents interviewed with the parental questionnaire and adolescents interviewed with the adolescent questionnaire are not related. Instead, they were sampled from different households. The parent

---

[27]The two originally selected cities, i.e., Lahore in Punjab and Karachi in Sindh, later became ineligible for the project and were thus excluded from the sample.

[28]Since it was planned that the RCT would be designed at a later stage to allocate about 70% to treatment and 30% to controls, the overall plan for the UC selection was to aim for a coverage of about 40% of the population, so that about 30% (i.e., 40%*70%) would be covered. Thus, for instance, areas in the desert were excluded due to small population coverage.

[29]In more detail, the evaluation included both IKEA-I and expansion areas (IALP areas) in the sample. Included were all IKEA-I operational areas so that after the baseline, 70% of them could be selected to also receive the new IALP program. IKEA-I Tehsils with low prevalence of child marriage in Punjab were not included. For example, in Bahawalpur and Rahim Yar Khan, the predecessor program had operated in Tehsils with very high and also with low child marriage rates. For the expansion plan it was agreed to focus particularly on UCs with high child marriages rates (and not use Tehsils, such as Yazman in Bahawalpur, with low rates). In Bahawalpur, this meant going to a new Tehsil and dropping one. In Rahim Yar Khan, the expansion was planned to focus on the two Tehsils with highest child marriage incidence.

[30]The final number of villages was defined based on power calculations.

[31]The procedure is described in more detail in Section I.A.

[32]The evaluation team conducted two pilots, in June 2015 (incl. 2 days of training) and September 2015 (incl. 3 days of training). Final baseline data was collected between mid-September and mid-December (128 person-days). Enumerator teams were mostly gender-balanced. Enumerators were requested to interview first the parent of their own gender, implying some sample balance with regard to parents' gender.

Figure 2.3: Program and Control Villages: Geographical Distribution
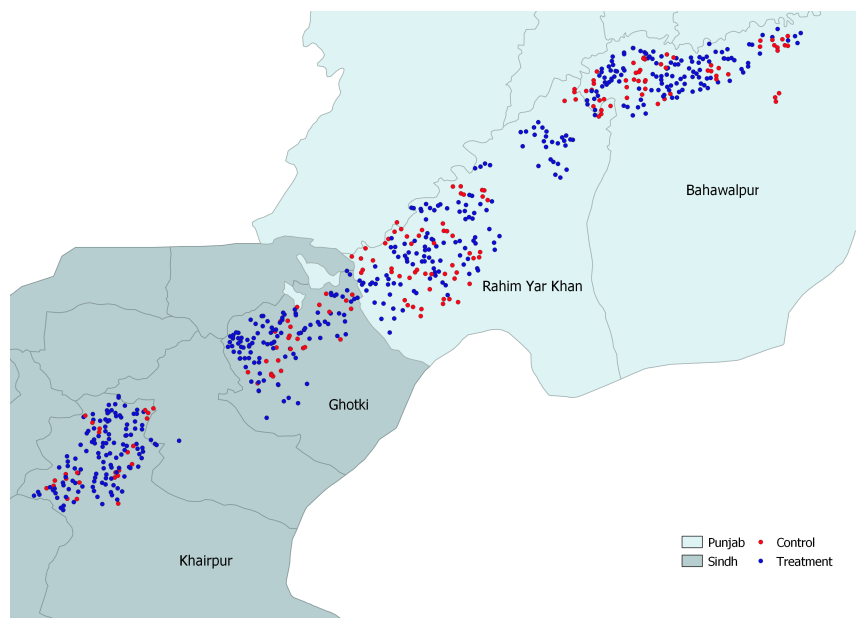
▶ *Notes*: This figure displays program and control areas in the four districts. Figure O.7 in the Appendix displays the vegetation, displaying the areas that are not covered by the intervention because they are deserted areas with little population.

sample is comprised of parents who have at least one adolescent daughter or son. Parents' and adolescents' gender is balanced in each district.

Sample Size   In total, the baseline rural interviews took place in 653 villages. Thereof in Bahawalpur in 181 villages, in Rahim Yar Khan in 189 villages, in Ghotki 131 in villages, and in Khairpur in 152 villages. The final baseline sample in rural areas includes a total of 2,744 adolescents (1,382 girls and 1,362 boys). The sample contains a total of 2,775 parents, out of which 1,399 are mothers and 1,376 are fathers. Additionally, the sample includes 657 community leaders and 663 Nikah Khwans. Therewith, each village has at least one community leader and at least one Nikah Khwan.[33]

A HH head or a knowledgeable adult provided answers on the whole household by answering a household roster.

Questionnaires   A separate questionnaire was developed for each respondent type: An adolescent questionnaire (AQ) for adolescents, a parent of adolescent questionnaire (PQ) for parents, a Nikah Khwan questionnaire (NKQ) for marriage registrars, and a community leader questionnaire (CLQ) for community representatives. In a household roster, respondents were additionally asked about each household member and about siblings/ offspring who no longer live in the household.

[33]In few cases a village had more than one leader or Nikah Khwan, and in those cases both were interviewed, respectively, in order to prevent internal conflicts. For this reason, although 653 villages are in the sample, there are 657 community leaders and 663 Nikah Khwans, respectively.

II.B.3   Endline Data

Timing of the Survey   Given the need to respond to the changing COVID-19 situation, the endline data collection took place in different waves between September 2020 and August 2021. In more detail, the endline data collection started in September 2020 with a subset of interviews (targeting community leaders only) via phone in order to assess the pandemic situation. Based on the insights from these assessments, the in-person data collection was decided on October 22, 2020. The interviews with parents, adolescents and marriage registrars started in winter 2020/ 2021 (Figure 2.2).[34] Data from community leaders (CLs) was collected via phone from September 2020 onwards. All respondents that could be reached were interviewed by mid October 2020.[35] In-person endline data in Sindh was collected between December 2020 and March 2021, followed by another break due to the global pandemic. Data collection then resumed in Punjab in June 2021 and was completed in August 2021.

Target Group   As in the baseline survey, four different respondent types were targeted, for each of who a different questionnaire was administered. For adolescents, the age eligibility criteria was chosen so as to include the group of young people that were adolescents at some point during program implementation and, therefore, eligible to participate in program activities, hence 11- to 23-year old respondents were sampled due to their eligibility for the program during its implementation. Specifically, the following groups were targeted besides Nikah Khwans and community leaders as the duty-bearers: Adolescents and young adults (aged 11 to 23) as the rightholders, and parents of adolescents and young adults that have at least one child aged 11 to 23. Depending on their age, adolescents and young adults (age 11-23) were targeted by and, therefore, exposed to the program for a different duration of time. For this reason, and also to capture the impact of the program on its beneficiaries with more power, the random draw of the adolescent among all HH members aged 11 to 23 was weighted to account for this difference in program exposure.

> Four different respondent types were targeted: Adolescents and young adults (age 11 to 23), parents of adolescents that have at least one child aged 11 to 23, Nikah Khwans, and community leaders.

In NFBE villages, NFBE-eligible children were additionally targeted. Eligibility for this program activity was defined by the eligibility to access NFBE centers. These were designed for students that have dropped out of school, that are at risk of dropping out, or have never enrolled in school. Moreover, with the aim to conduct a panel analysis on the individual level, a subset of *baseline households* was specifically targeted again at endline. For this purpose, prior to the data collection, a list of baseline participants to be interviewed was shared with the enumerators.[36] Thus, in each village, the enumerator teams tried to track 3 households that also participated in the baseline (2 baseline adolescents' households and 1 baseline parent's household).

Respondents Sampling Strategy   New households were sampled through a random walk, whereby it was also accounted for the fact that RVs often consist of several different settlements. In each village, respondents were sampled from the main and one additional *randomly* selected nearby settlement.[37]

To identify eligible adolescents and parents for the survey, a listing was implemented. Following a short household roster to identify people aged 11 - 23 currently living in that household, the main respondent was drawn *randomly* from

---

[34]Originally, the endline data collection was planned to take place in spring 2020. Due to the COVID-19 pandemic, the data collection was split into two waves, starting with interviews with key community representatives or community leaders via phone. This questionnaire was shortened.

[35]Later, field teams collected additional numbers in the field to complete the CL survey. As these new numbers became available only with the teams being in the villages, remaining interviews with CLs were conducted until August 2021.

[36]This list contained the respondent's gender, age, and the name of all household members. The order of the names on this list was random. Per village, enumerators were asked to track 2 adolescents and 1 parent from the baseline. They started by looking for the first adolescent or parent on the list. If they could not find them, they looked for the second one, and so on. If an adolescent was found, the next adolescent from the other gender was searched next.

[37]For the selection of random respondents, enumerators conducted a random walk to all four sides of the village, starting from the right hand side of the center point. The center point could be geographically central or located at any major site such as school, mosque, community center or any such prominent place recognized by the community as central point. During the random walk, every 5th household was selected for the interview. If the household was eligible, i.e., had an adolescent or young adult of 11 to 23 years and gave consent, it was selected, otherwise the team would move to next household. This was done until the required number of respondents was reached. Moreover, about 70% were selected in the main settlement of the revenue village, and 30% in a randomly selected settlement of the revenue village at around 30 minutes walking distance from main settlement. Note, if every 5th household was poor or very poor it was selected (judging by the appearance of the dwellings). During the selection of the last two to three households, enumerators tried to balance poor (houses of backed bricks but not cemented) and very poor (unbaked bricks and not cemented and straw house or a mere shade) along with respondents' gender balance.

all household members in that age group.[38] Consent for the full interview was acquired[39], and the enumerator teams returned for the full survey with the chosen adolescent or the chosen parent.

In the 212 NFBE-villages, five households with at least one NFBE-eligible member aged 11 to 23 were added to the sample. In those households, the youngest NFBE-eligible adolescent above 11 was sampled. Consequently, the evaluation has a higher share of young respondents in NFBE villages. This accounts for both the fact that the NFBE intervention was targeted at younger children and that higher impacts of this component for the younger age range were expected.[40] Baseline households were targeted using household information from the baseline survey.

{ A subset of baseline households were specifically targeted again at endline. }

The interviews with the community leaders were based on phone numbers provided by UNICEF and its implementing partners, whereby the baseline names of the community leaders were shared. Eventually, 67% were community leaders and the remaining ones were knowledgeable community representatives, such as community members, teachers, and social workers. For simplicity, all are referred to as community leaders.

Finally, the sampling aimed at a balance between male and female respondents. Furthermore, two types of poverty levels were approximated, very poor and "not so poor" households in order to identify the group of adolescents who would have benefited most from the program.[41] The full sampling approach (incl. numbers of observations and targeting) is described in detail in Appendix A.II.A.

Sample Sizes   The estimated, PAP-registered, intended sample size for the face-to-face endline survey for 653 RVs was a maximum of 8,243 observations.[42] In reality, the endline interviews took place in 648 villages and therein for the NFBE evaluation in 212 villages.

{ Interviews took place in 648 villages; for the NFBE evaluation in 212 villages. }

The following Table 2.7 summarizes the final number of observations per type of respondent. A total of 7,415 respondents were interviewed.[43] Panel A provides information on the main respondents to the full survey. Among adolescents, 9.5% were traced from the baseline and 42.6% were randomly sampled at endline. Among parents, 6.1% were traced from the baseline and 27.8% were randomly sampled. In terms of geographical coverage, 49.2% of observations are in Sindh and 50.8% in Punjab. The final endline sample includes a total of 4,902 main respondent adolescents age 11-23 (2,419 boys and 2,483 girls). Panel B displays the age distribution of NFBE-eligible respondents only. Here, the endline sample comprises a total of 2,421 NFBE-eligible adolescents and young adults age 11-23 (1,058 boys and 1,363 girls). Overall, the original face-to-face targets for the sampling were achieved as planned. There were, as mentioned, limitations in being able to reach all individuals targeted who were targeted via phone.

Questionnaires   The endline questionnaires (four questionnaires: One for adolescent, one for parent, one for Nikah Khwan, one for community leader) captures key characteristics and program goals. Questionnaires, aligned with the baseline questionnaires, begin with screening questions concerning the economic status and presence of female/male adolescents in the household - key information to meet the sampling criteria. Apart from questions capturing the Key Performance Indicators (KPI), the adolescent and parental questionnaire includes a household roster. In the questionnaire, adolescents are defined as the group aged 10 to 19.[44]

---

[38]For parent interviews, one adolescent or young adult between 11 and 23 was randomly selected. One of their parents was then selected as main respondent for the PQ.

[39]If the main respondent was below 18, additional consent was obtained from their parent or legal guardian.

[40]The different types of survey respondents are depicted in Figures A.II.A and A.2 with original material used in the enumerator training. For more graphical information on the selection of the NFBE respondents see Figure A.II.A.

[41]This is done by targeting households living in bricks but not cemented houses versus unbacked bricks/ mud houses (very poor households) in each respondent group (adolescents and parents). Therewith, in the final sample, the average poverty level is 0.8. Note, the same sampling strategy as in the baseline survey was followed whereby the non-poor group only included households that own up to 12.5 acres of land. This threshold was chosen in order to keep the focus on the adolescents that should have benefited most given an access to the interventions.

[42]In more detail, the sampling plan was as following: In all non-NFBE villages (441 villages in total), the target was to list 10 HHs for the full survey, resulting in around 12 interviews per village (6 AQs, 4 PQs, 1 NKQ, and 1 CLQ). In the NFBE-eligible villages, around five eligible additional adolescents were sampled per village, resulting in about 17 targeted interviews in these villages. The final planned sample size was: Around 1,959 adolescents and parents in a panel, as a follow-up to the baseline. Around 5,224 individuals newly sampled (3 parents, 4 adolescents, 1 marriage registrar). Thereof the evaluation team estimated to sample 5 additional NFBE-eligible children in 212 villages (around N=1,060).

[43]Note, if the main adolescent respondent fulfilled the NFBE-eligibility criteria, no further NFBE respondent needed to be sampled; resulting in a smaller sample size.

[44]At baseline, adolescent respondents aged 12 to 19 were targeted. While defining the targeted age group at endline, the evaluation followed UNICEF's request to apply a different definition of adolescents. The endline questionnaire contains some changes to the baseline questionnaire: New survey modules of questions, in particular on program awareness and take-up, were included.

Table 2.7: Number of Observations

	(I)		(II)		(III)	
Panel A: Main Respondents						
			Gender			
	Male		Female		Total	
	No.	%	No.	%	No.	%
Respondent type:						
adolescent (11-23, from baseline)	352	9.8%	349	9.1%	**701**	**9.5%**
adolescent (11-23, randomly sampled)	1,590	44.2%	1,566	41.0%	**3,156**	**42.6%**
adolescent (additional for NFBE)	477	13.3%	568	14.9%	**1,045**	**14.1%**
parent (of 11-23, from baseline)	211	5.9%	242	6.3%	**453**	**6.1%**
parent (of 11-23, randomly sampled)	966	26.9%	1,094	28.6%	**2,060**	**27.8%**
Total	3,596	100.0%	3,819	100.0%	**7,415**	**100.0%**
Province:						
Punjab	1,732	48.2%	1,918	50.2%	**3,650**	**49.2%**
Sindh	1,864	51.8%	1,901	49.8%	**3,765**	**50.8%**
Total	3,596	100.0%	3,819	100.0%	**7,415**	**100.0%**
Adolescent's age:						
11	122	5.0%	143	5.8%	**265**	**5.4%**
12	272	11.2%	302	12.2%	**574**	**11.7%**
13	247	10.2%	322	13.0%	**569**	**11.6%**
14	237	9.8%	310	12.5%	**547**	**11.2%**
15	278	11.5%	292	11.8%	**570**	**11.6%**
16	184	7.6%	237	9.5%	**421**	**8.6%**
17	193	8.0%	166	6.7%	**359**	**7.3%**
18	239	9.9%	248	10.0%	**487**	**9.9%**
19	142	5.9%	123	5.0%	**265**	**5.4%**
20	208	8.6%	139	5.6%	**347**	**7.1%**
21	88	3.6%	66	2.7%	**154**	**3.1%**
22	132	5.5%	84	3.4%	**216**	**4.4%**
23	77	3.2%	51	2.1%	**128**	**2.6%**
Total	2,419	100.0%	2,483	100.0%	**4,902**	**100.0%**

Panel B: NFBE-eligible Adolescents (in NFBE-RVs)						
			Gender			
Adolescent's age:	Male		Female		Total	
	No.	%	No.	%	No.	%
11	43	4.1%	89	6.5%	**132**	**5.5%**
12	115	10.9%	168	12.3%	**283**	**11.7%**
13	89	8.4%	166	12.2%	**255**	**10.5%**
14	90	8.5%	167	12.3%	**257**	**10.6%**
15	134	12.7%	161	11.8%	**295**	**12.2%**
16	83	7.8%	163	12.0%	**246**	**10.2%**
17	84	7.9%	103	7.6%	**187**	**7.7%**
18	119	11.2%	130	9.5%	**249**	**10.3%**
19	61	5.8%	63	4.6%	**124**	**5.1%**
20	98	9.3%	60	4.4%	**158**	**6.5%**
21	39	3.7%	27	2.0%	**66**	**2.7%**
22	69	6.5%	38	2.8%	**107**	**4.4%**
23	34	3.2%	28	2.1%	**62**	**2.6%**
Total	1,058	100.0%	1,363	100.0%	**2,421**	**100.0%**

▸ *Note*: Table 2.7 describes the sample size, overall and by key characteristics. Panel A provides information on the main respondents to the full survey. Panel B takes displays the age distribution of NFBE-eligible respondents only.

The adolescent and parent questionnaires are built following a parallel structure, i.e., the same questions are asked to adolescents and to the parents (with reference to the behavior of their randomly selected adolescent child). The community leader questionnaire collects key leaders' and village characteristics.[45] The questionnaire designed for Nikah Khwan captures basic characteristics concerning Nikah Khwan themselves (their age, the UCs that Nikah Khwan work in, whether they are religious leaders) and their professional experience. All questionnaires were subject to several rounds of revisions with the IPs and UNICEF. The questionnaires were translated into Urdu and Sindhi.[46]

The additional NFBE-respondents answered the full AQ as well as the module on NFBE awareness and the math and literacy test. In NFBE villages, the two latter were also completed by adolescents in the main survey households in case of eligibility.[47]

Data Collection (I)

► Photographs depict data collection, from left-to-right, in RV Fareed Malik, and RV Abad Jagir.
► *Source*: Center for Development and Evaluation (C4DE) Pakistan.

In the households that were surveyed using the standard questionnaire, NFBE-related questions were asked and the skills test was administered with an NFBE-eligible adolescent that was identified during the listing.[48]

Following a pre-household roster used for the listing, a HH head or a knowledgeable adult provided answers on the whole household by answering a household roster. Applying an out-of-household roster, basic information was obtained from the household head about adolescents who moved out from the respective household in the past four years. Here, questions about the marriage status of the children, educational levels, and children, amongst others were asked.[49] The household roster respondent also answered questions about basic characteristics of the household. This covered the number of rooms, type of toilet, ownership of refrigerator, air conditioner, microwave oven, engine driven vehicles, TV, and livestock.

The following survey modules were asked to the main respondent (the parent of an adolescent or an eligible young person). First, the main respondent was asked how long she or he has been living in the household, followed by a *module on education*. In the parent questionnaire, the same questions were asked about the child. In a *participation module*, the aim was to better understand if and how the program was implemented. Adolescents and parents were asked whether they were aware of the activities, whether they participated in them, and additional information on the activities. The survey continued with *modules on actions taken and self-efficacy, on protection and safety, and on personality and wellbeing*.

[45]For CL, some indicators are included that originally were not asked to allow for greater learning, especially since in-person interviews did not take place. The questionnaire was shortened in order to limit the duration of the interview to 40 minutes.

[46]Complete endline questionnaires: Section O.IV..

[47]Hence, up to 10 additional adolescents answered the respective modules (up to 2,120 adolescents).

[48]If the randomly selected adolescent for the main interview was also eligible for the NFBE program, they answered the additional education-related questions during the full interview. Otherwise, another eligible adolescent or young adult was randomly drawn. If no eligible person aged 11 to 23 resided in the household, the additional modules were not administered in that household.

[49]Note that as long as the interviewing approach is the same in treatment and control areas, and as long as household-attrition is non-differential between treatment and control, all impact estimates are valid.

Following, the adolescents were asked questions regarding discussions of the adolescents with parents, religious leaders and others, in which situations the adolescent's opinion is taken into consideration, whether the adolescent experienced corporal punishment and similar questions. A *module on marriage* included questions regarding the marriage status, arranged marriages, plans of getting married or engaged and for having children, main reasons for marriage and similar questions. The main respondents were asked a set of questions testing their knowledge concerning the law in Pakistan. Several of these questions first asked the adolescent whether a legal age exists for certain activities. Final modules of the main questionnaire included *information on role models*, where the respondent's opinion regarding certain statements about adolescents and expected behavior of men and women were asked about. This was followed by questions on social norms and practices, such as *attitudes about marriage and education*, asking about the opinion with respect to certain statements regarding adolescents and expected behavior of men and women. Moreover, respondents were asked about their knowledge of child rights and whether they were reached on child rights topics through the IALP program. The very last questions asked about *experience with COVID-19*, in particular asking the respondent about how strong she or he was affected by the pandemic and how well he or she coped with it.

To receive information on the implementation and capture the impacts of this intervention, questions on awareness and attendance of the centers, among others, as well as a math and literacy test were developed to be administered to NFBE-eligible adolescents. Thus, all NFBE-eligible children were asked a special *NFBE module* measuring experiences with NFBE Centers and learning outcomes (the so-called school drop-out or never-enrolled questionnaire/ module (SDM)). Mostly, questions such as whether the eligible children heard of the centers and how, whether they were invited, whether they attended them and, if so, for how long were of interest. They were also tested in math and literacy skills.[50] The questions were increasing in difficulty, with three levels being assessed and three questions per level. The content of the questions follows the curriculum from UNICEF and each question tries to test some particular competence/ skill.[51]

> All NFBE-eligible children were asked a special NFBE module measuring experiences with NFBE Centers and learning outcomes.

Community representatives (mostly leaders) were interviewed via phone with a shortened questionnaire.

Data Collection (II)

▶ Photographs depict data collection, from left-to-right in RV Fareed Abad, and RV Lutif Gadani.
▶ *Source* C4DE Pakistan.

II.C   Qualitative Interviews

Timing of the Survey   key informant interviews (KIIs) with different stakeholders involved in the program were conducted between February and June 2021.

Target Group   The quantitative impact evaluation of the IALP program is supported by qualitative evidence from KIIs with UNICEF Child Protection and Education Officers, government officials from the Social Welfare and District

---

[50]Based on the UNICEF curriculum and the tests used at the centers.
[51]UNICEF provided a document with some tests and the curriculum. See Section O.IV.A for the detailed questions asked.

Education Departments, as well as managerial staff from the IPs in both provinces. The qualitative interviews were conducted with UNICEF, IPs, and government staff. The interviews with government officials in Sindh and implementation partners in Sindh were conducted face-to-face. However, due to COVID-19, interviews with government officials, IPs, and UNICEF were conducted virtually, via phone.

Respondents Sampling Strategy   The evaluation used purposive sampling, aided by a list of 35 prospective interviewees provided by UNICEF, to identify study participants, which ultimately consisted of UNICEF staff, representatives of IPs (for Sindh: RSPN and IRC; for Punjab: BLCC) as well as government officials from Sindh and Punjab.[52] In total, 17 interviews were conducted locally by a Senior Research Manager of C4DE in Pakistan, 18 prospective interview partners could not be reached or did not respond to requests. The analysis is based on a sample of 16 KIIs.[53] A detailed, anonymized breakdown of the final sample of respondents can be found in the Appendix (Tables A.66 and A.67).[54]

Questionnaires   For each stakeholder group, a separate questionnaire with open questions was designed. Based on the interviewee's preferences, KIIs were conducted in English, Urdu or Sindhi. All analyzed interviews were audio recorded, first transcribed in the original languages, and later translated to English. The qualitative part provides insight into the evaluation of the implementation process, challenges, and recommendations from the perspective of stakeholders. The latter can inform the design and implementation of future programs. To ensure the anonymity of the KIs, identifying information (such as names, workplace or job title) was omitted or enciphered.

II.D   Ethical Considerations, Integration of Gender, and Data Quality Assurance

This section discusses ethics and data quality assurance. Further details are provided in Appendix O.I.C.

II.D.1   Ethical Considerations

The evaluation follows the UNICEF's Procedure on Ethical Standards in Research, Evaluation, Data Collection, and Analysis (UNICEF 2021b). This evaluation adheres to the five key principles respect, beneficence, justice, integrity and accountability, during all work stages.

Conflict of Interest   There is no conflict of interest; the evaluation is conducted independently by C4ED and its local branch C4DE. The impartiality standard is also followed, and the results are presented in an unbiased manner, with a goal to transparently describe the program's strengths and weaknesses. The credibility of the evaluation is granted by the employed methodology as well as the data quality assurance process throughout the data collection and multiple rounds of the review process. For more information, please refer to Section I..

Data Collection   C4ED assured that the data collection firmly adheres to the principles of the conduct of evaluations laid out in the UNICEF procedures for Ethical Research Involving Children (Graham, Powell, Anderson, Fitzgerald, and Taylor 2013) and UNICEF's Procedure on Ethical Standards in Research, Evaluation, Data Collection, and Analysis (UNICEF 2021b). To address the ethics of the entire research process, ethical clearance was acquired from the institutional review board (IRB) at the University of Mannheim. Approval for the evaluation was granted in May 2015. In September 2020, the IRB further approved the start of the endline data collection. Moreover, the evaluation team additionally requested approval from the IRB at Research and Development Solutions in Islamabad which was approved in November

---

[52]In more detail, UNICEF provided the list of 35 potential KIIs and also offered support in reaching out to them. C4ED contacted all potential respondents via email, SMS, and telephone call. Some government officials had retired or were posted in other districts or had gone abroad by the time the evaluation was conducted. One of the respondents was very ill and therefore refused to participate. Most of the senior government officials and senior UNICEF staff did not respond to requests for interviews. To address this limitation, interview findings were triangulated with secondary data and quantitative results where possible.

[53]As one interview partner did not consent to be audio recorded, this KII is not included in the analysis.

[54]The methodology for qualitative data collection was slightly changed from the one originally envisaged. Initially, focus group discussions (FGD) were planned to be conducted with parents, adolescents (in school and out of school), and the community. The objectives of FGDs were to gather information on different perspectives on interventions, documenting best practices, recommendations, lessons learned on the ground, and examine social and community dynamics around participation and empowerment. However, due to restrictions for the prevention of COVID-19 placed by Government of Pakistan for any group gatherings, FGDs could not be carried out. Normally, FGDs are very useful for uncovering community and social dynamics, however, the inability to carry out FGDs means that the only information regarding community and social dynamics comes through KIIs. This limits the ability to gain an insider or community perspective as KIIs may lack insight into the the complexity of social changes experienced by the community. It further limits the scope of qualitative data collection to KIIs only.

2020.[55] Prior to engaging respondents for data collection, C4ED followed a thorough informed consent procedure for participation in the evaluation. This included the provision of prospective respondents with the opportunity to ask questions and to terminate data collection without repercussions at any time. For respondents under 18 years of age, parental informed consent was acquired together with an adolescent's assent indicating their willingness to take part in the interview. C4ED keeps all participant data confidential by encrypting and/ or removing from the data any identifying information prior to sharing it beyond the core research team. The research team made sure that settings, questions, and the entire process are free of stigmatization and that all respondents are comfortable with their engagement in the evaluation. The questionnaires and interview grids were adapted to the local context with the help of UNICEF and the IPs. Moreover, all the enumerators and qualitative researchers received trainings on basic ethics and, in particular, ethical behavior during the interview process. Female adolescents were only interviewed by female enumerators due to the sensitivity of the questionnaire and research goals.

COVID-19   The highest priority of C4DE is the health and safety of staff, field teams, and the communities in which C4DE works. Together with UNICEF Pakistan, the country situation was critically assessed and necessary measures implemented to protect staff and respondents (in line with all national guidelines and restrictions).[56]

II.D.2   Data Quality Assurance

During data collection, real-time checks were conducted to assure high-quality data. Those checks included, i.e., average interview duration, response distribution of key outcome variables and outliers. The progress of completed interviews by day and enumerator was closely monitored during the full period of data collection. Additionally, back checks as well as spot checks of a random sample were realized. In case that any suspicious patterns were detected, the most convenient actions were taken to improve data quality. Finally, interviewer assignment to villages were independent of the treatment status of the villages.

During analysis and report writing, International Initiative for Impact Evaluation (3ie), peer reviewers, and UNICEF reviewed the working documents to assure quality. In this course, draft findings were also presented to UNICEF ROSA, UNICEF Pakistan, and 3ie as reference group to provide feedback. Quality of results are also assured applying different robustness checks (e.g., running different regression specifications, multiple hypothesis testing, Section II.C). All flagged interviews - due to suspicious data quality - were excluded from the main analysis. Moreover, all estimations include fixed effects for each enumerator.

II.D.3   Limitations

Implementation Intensity and Timing of Measuring Outcomes   The overall duration of the main implementation phase has been about 2,5 years (summer 2017 to end of 2019). Moreover, the program has been implemented by several local partners. In Sindh, the partner was present over the years, while in Punjab the partner changed in the midst. The implementation duration was shorter in Punjab due to a change in implementing partners. Program intensity varied between the provinces and over time. To change deeply rooted social norms and traditions a longer period of implementation may need to be required, at least this was the original concern. Moreover, against original plans the pool of community leader respondents is incomplete and the questionnaire shortened in order to be implemented via phone (Section II.B.3).

NFBE Evaluation   The evaluation encompasses interventions from two different programming teams at UNICEF, one situated in Child Protection, and one in Education. The evaluation of the educational component was added at a later point in time (Section I.A). Hereby, no further resources could be dedicated to study the impacts of this component comprehensively. In the endline survey, the pool of respondents was enlarged. Thereby out-of-school (OOS) children were targeted within the randomly selected households with adolescents. Schools were not visited and school records on re-enrollment were not tracked.

Qualitative Evaluation   Due to limited resources, the qualitative evaluation was limited in scope at endline. UNICEF provided the list of potential key informants, yet not all of them could be reached (Section II.C). Therefore, to address the limitation of low response, quantitative findings were used to triangulate qualitative data.

---

[55]See Section O.V. for the approval letters.
[56]For more details, please see Section O.I.C in Appendix.

CHAPTER 3

Findings

In this chapter the results of the evaluation are presented, which are based on the endline survey collected after the end of the program and the MIS data of the program collected during program implementation. Section I.A starts with a description of key characteristics in the analyzed sample of respondents, with comparisons being made to official statistics and related studies. The idea is to describe the specific population of interest - individuals eligible for the program - for whom the results presented thereafter hold. Section I.B continues to lay out evidence on program delivery, awareness, and uptake; essentially capturing the main inputs of the program and their limitations. Given this background information, Section II. presents the main results. The section starts with an analysis of evidence on program implementation (Section I.B). It continues with the identification of program impacts on the main outcomes of interest of this program, i.e., the identification of causal changes in child marriage, child pregnancies, and education (Section II.A). The core outcomes are presented along with related measures. To better understand the drivers of change or the lack thereof, Section II.B therefore focuses on the program's impacts on changes in adolescent empowerment, communication, and social norms.

I. Descriptive Statistics

In this section, the characteristics of the program-eligible survey respondents are described (Section I.A). Thereupon the first evaluation question is answered: "What evidence on IALP implementation is captured and how are similar activities implemented in non-program (control) areas?" (Section I.B).

I.A Local Context at Endline (2020/ 2021)

The following paragraphs display the descriptive statistics of the endline sample collected from December 2020 to April 2021 (in Sindh) and June to July 2021 (in Punjab). The focus is on parents and children only. Overall, referring to all households in treatment and control RVs, the sample includes a total of 4,057 adolescents (thereof 4,902 with full interviews), and a total of 17,727 parents (thereof 2,513 with full interviews). Of the adolescents (2,143 girls and 1,914 boys), 2,102 are in the 15-19 age group (thereof 2,102 with full adolescent survey). Apart from presenting the statistics for both groups together, they are also disaggregated by the respondents' gender. In more detail, Table 3.1 displays in Part I the full sample, then Part II restricts the sample to respondents of the full adolescent survey only and Part III to respondents of the full parent survey only. Presented are the number of observations and the mean and standard deviation of each variable (outcome of interest), followed by t-test differences by gender of respondent.

   In the households sampled, 5.6% of adolescents aged 15-19 reported being currently married. When disaggregating by gender, it becomes clear that girls are more likely to be married than boys. While 8.1% of girls age 15-19 are married, not even 3.3% of boys in the same age category are. The numbers for girls are relatively low compared to earlier official figures for Pakistan, which report 15% [10.5%] of girls and 4.8% [2.9%] of boys in Sindh [Punjab] age 15-19 currently married (Bureau of Statistics 2020; Bureau of Statistics, Bureau and Board, Planning & Development and of the Punjab, Government 2018). Although engagement is not a strategic result per se, it can be a good predictor for future marriages. Gender differences are also visible here, as 12.2% of girls age 15-19 are engaged whereas only 7.6% of boys in the same age group are. With respect to out-of-household adolescents, 10.7% of the in total 431 adolescents

aged 11-23 not living in the household cited marriage as a reason for moving out.[1] With regard to underlying reasons, 83.6% of marriages occurred for any other reason than love and appreciation. 28.4% of girls and 30.1% of boys age 11-23 reported having friends their same age that are married.

In terms of educational attainment in our sample, 26.1% of adolescents aged 10-14 attend secondary school. This number is lower than pre-pandemic national statistics, according to which the net attendance ratio for secondary school children (age 10-14) was 37.7% in 2017-18 (Pakistan and ICF 2019).[2] Across most of the education indicators, gender differences are visible. For example, only 39.6% of girls age 11-14 have completed primary education compared to 47% of boys in the same age group. This is consistent with the overall situation in Pakistan, where girls and young women are disproportionately constrained in their access to education (UNICEF 2020).

Drop-out rates in the 12-14 age group are high and girls are more likely to be affected. As much as 32% of all adolescents in our sample have ever dropped out of school - 38% of girls, and 27.1% of boys.[3] At the same time, girls are significantly less likely than boys to have aspirations to continue education: Only 39.2% of girls reported actually planning to continue their education, compared to 54.2% of boys. Education-related attitudes seem to be strongly biased towards boys. For example, almost half of the respondents (43.2%) believe that educating boys is more important than girls. And boys are more likely to agree with this statement (48.8%) than girls (40.8%). This gender imbalance is also reflected in the World Values Survey, according to which 51.7% of females and 66.7% of males in Pakistan agreed with a similar that statement that university is more important for a boy than a girl (Haerpfer, Inglehart, Moreno, Welzel, Kizilova, Diez-Medrano, Lagos, Norris, Ponarin, and Puranen 2020).

> In the sample, more than every third girl (age 12-14) has ever dropped out of school.

> More than every second adolescent (age 11-14) experiences corporal punishment.

In terms of adolescent empowerment, the confidence of adolescents to express their opinion in the public sphere was captured, which seems low overall and also displays large gender differences. For example, while 30.6% of boys feel confident expressing community needs to local government officials, only 14.7% of girls feel the same way. In addition, adolescent's confidence in day-to-day situations is low, especially among girls. For example, only 11.5% of girls and 49.7% of boys reported feeling confident in going to the local market alone. As a short self-reported measure of current mental wellbeing, the World Health Organization-Five Wellbeing Index (WHO-5) was employed. On average, adolescents reached 71.0% on the scale, with 100% representing the best imaginable wellbeing. In order to see whether adolescents feel they can influence their fate, a short locus of control scale was applied.[4] On average, adolescents reached 64.0% on the scale, with 100% indicating a very strong feeling of being in control. One item of the scale asks respondents whether the statement "If I work hard, I will succeed" applies to them. In our sample, 55.3% of females and 58.8% males agreed with this statement. A similar level can be observed in the World Values Survey, where 41.2% of females and 41.6% of males in Pakistan agreed with the statement "In the long run, hard work usually brings a better life" in 2018 (Haerpfer et al. 2020).

Use of violence against children was captured by asking about the acceptance of using corporal punishment on children, as well as by asking adolescents about their experience. When adolescents and parents were asked about the usage of corporal punishment in a situation where daughters do not obey, 65.7% agree on average with the parental usage of corporal punishment in this situation. There are no significant gender differences observable. Regarding actual prevalence of corporal punishment, 59.3% of adolescents age 11-14 in our sample reported experiencing corporal or physical punishment in any place. This is a slightly lower prevalence compared to national statistics, according to which 62.8% of adolescents in Punjab age 10-14 reported experiencing any form of physical punishment in the last month (Bureau of Statistics, Bureau and Board, Planning & Development and of the Punjab, Government 2018).

> "If I work hard, I will succeed" - only slightly more than every second adolescent agrees to this statement.

To capture some prevailing social norms, adolescents and parents were asked about their view on gender roles. To

---

[1]The results indicate stark differences between in- and out-of-household adolescents. While 97.5% of women aged 20-23 and 43.4% of men aged 20-23 who are not living in the household are married, on average only 25.5% of adolescents aged 20-23 that live in the household are married (implying that the main reason for not living in the household is marriage).

[2]Official numbers are: In Sindh, 20.9% of children age 10-12 attend secondary school or higher in 2018-19 (Bureau of Statistics 2020). In Punjab, 36.7% of children age 12-14 attend secondary school or higher in 2017-18 (Bureau of Statistics, Bureau and Board, Planning & Development and of the Punjab, Government 2018).

[3]According to national estimates, approximately 36.7% of adolescents aged 12-14 in Punjab have been out of school in 2017-18 (Bureau of Statistics, Bureau and Board, Planning & Development and of the Punjab, Government 2018).

[4]Originally developed by Rotter (1966) and shortened by Lumpkin (1985) and Kovaleva (2012).

test this, a scale from Waszak, Severy, Kafafi, and Badawi (2001) was used, which captures the role of women/ girls vs. men/ boys in the public and private sphere. On average, respondents answered 22.0% of the questions on the gender role models scale (0-14 items) positively, i.e., presented egalitarian beliefs about gender norms. If one looks within the scale into its components, when asked whether daughters should have the same chance to work outside the homes as sons, 65.2% of adolescents and 65.9% of parents agree with the statement. Another way of measuring social norms refers to perceptions and attitudes surrounding child marriage. For example, respondents were asked whether they agreed with the statement that girls should finish education before marriage - which a slight majority did (57.6% of all respondents). Adolescents and parents were also asked what rights existed for children. On average, respondents were able to correctly mention about two out of four rights.

With respect to communication, respondents were asked if parents take their children's opinion into account for a series of decisions. Both girls' and boys' opinion is frequently not taken into account, and girls' opinions are considered relatively less. On average, adolescents' opinion was only taken into account in 25.0% of situations (out of seven). In more detail, only 26.9% of girls compared to 56.5% of boys felt that their opinion on when they will get married is considered by their parents. In terms of education, while 81.4% of boys feel that their opinion on whether to go to school is taken into account, only 60.6% of the girls have the same feeling. In addition, the established psychological scale on communication by Rubin and Martin (1994) was applied. On average, adolescents reached 58.0% on the interpersonal communication competence scale, with 100% representing full competence across 10 dimensions. In the empathy sub-domain, respondents were asked whether they feel they can put themselves in other shoes: On a scale of 1 to 5, adolescents reached a score of 2.73 on average, with 5 representing full empathy competence. Similar to the overall ICC scale, girls reached a lower competence score than boys: On average, girls reached an empathy score of 2.69 while boys obtained a score of 2.76.

Lastly, respondents were also asked about their self-reported program uptake. On average, across both treatment and control villages, the share of overall Child Protection-program awareness among adolescents is 12.1%, whereas it is slightly higher among parents at 15.5%. Adolescents and parents were most likely to be aware of group sessions (6.8% and 8.8%, respectively). Of those aware of the activities, 55.5% of adolescents and 55.8% of parents reported participating in any child protection program activities. Regarding non-formal basic education centers, 29.9% of adolescents indicated being aware of these centers in villages within NFBE-eligible villages (19% in the full sample). Of those aware, 6.8% of adolescents reported ever attending or currently attending a non-formal basic education center in their NFBE-eligible village (4.9% in all villages).

This section concludes with selected background characteristics. The average respondent is 26.03 years old. Within revenue villages, average household size was 7.47, more than half of the households (59.0%) had a household head with no education, and the average poverty score was 17.35.

Table 3.1: Characteristics of the Estimation Sample

	Obs./Clusters (1)	Mean/SD (2)	(I) All			(II) Adolescents (11-23)			(III) Parents		
			Females Mean/[SD] (3)	Males Mean/[SD] (4)	T-test Diff. (5)	Girls Mean/[SD] (6)	Boys Mean/[SD] (7)	T-test Diff. (8)	Mothers Mean/[SD] (9)	Fathers Mean/[SD] (10)	T-test Diff. (11)
Panel A: Child Marriage (all household members)											
adolescent currently married (15-19)	9440	0.056	0.081	0.033	0.048***						
	[158]	[0.416]	[0.411]	[0.241]							
adolescent currently engaged (15-19)	9440	0.098	0.122	0.076	0.046***						
	[158]	[0.474]	[0.420]	[0.368]							
ever married (20-23)	4186	0.255	0.343	0.195	0.148***						
	[158]	[0.792]	[0.665]	[0.553]							
ever married (20-23; out-of-household)	213	0.840	0.975	0.434	0.541***						
	[86]	[0.432]	[0.157]	[0.434]							
adolescent (11-23) moved out of the household because of marriage	431	0.107	0.146	0.015	0.131						
	[119]	[0.487]	[0.546]	[0.128]							
marriage due to any other reason than love	1583	0.836	0.838	0.833	0.005**						
	[152]	[0.657]	[0.527]	[0.458]							
Child Marriage (only main respondents)											
adolescent has friends his/her age & gender that are married*	7353	0.275	0.265	0.285	-0.020	0.284	0.301	-0.017	0.230	0.252	-0.022
	[158]	[0.751]	[0.670]	[0.556]		[0.667]	[0.581]		[0.481]	[0.455]	
Panel B: Education (all household members)											
adolescent (10-14) attends secondary school	4449	0.261	0.261	0.261	-0.000						
	[125]	[0.678]	[0.634]	[0.550]							
adolescent (12-14) ever dropped out of school	5098	0.320	0.380	0.271	0.109***						
	[158]	[0.779]	[0.701]	[0.650]							
Education (only main respondents)											
adolescent (11-14) completed primary education	2445					0.387	0.491	-0.104***			
	[158]					[0.524]	[0.537]				
adolescent plans to continue education*	7270	0.470	0.417	0.526	-0.109***	0.392	0.542	-0.150***	0.463	0.494	-0.031
	[158]	[0.852]	[0.911]	[0.671]		[0.823]	[0.608]		[0.611]	[0.552]	
index for negative education-related attitudes	7398	0.422	0.417	0.426	-0.009***	0.430	0.438	-0.009***	0.394	0.402	-0.007
	[158]	[0.723]	[0.626]	[0.515]		[0.614]	[0.470]		[0.378]	[0.354]	
● agrees: educating boys is more important than girls	7269	0.432	0.402	0.463	-0.061***	0.408	0.488	-0.080***	0.390	0.412	-0.022
	[158]	[1.000]	[0.816]	[0.820]		[0.795]	[0.716]		[0.569]	[0.625]	
Panel C: Adolescent Empowerment											
index for confidence of adolescent in community	7335					0.112	0.225	-0.114***			
	[158]					[0.411]	[0.509]				
● can express community needs to local government officials	7227	0.251	0.186	0.320	-0.134***	0.147	0.306	-0.158***	0.255	0.349	-0.095***
	[158]	[1.002]	[0.748]	[0.842]		[0.579]	[0.705]		[0.652]	[0.677]	
index for adolescents confidence in doing activities themself	2754					0.162	0.391	-0.229***			
	[158]					[0.328]	[0.528]				
● confident going to local market alone	2711					0.115	0.497	-0.382***			
	[158]					[0.392]	[0.766]				
index for World Health Organization 5 items wellbeing*	3703	0.712	0.701	0.723	-0.022**	0.704	0.729	-0.025**	0.696	0.709	-0.014
	[158]	[0.339]	[0.285]	[0.263]		[0.263]	[0.248]		[0.223]	[0.208]	
● has felt active & vigorous over the last 2 weeks*	3733	0.773	0.758	0.789	-0.031	0.750	0.793	-0.043	0.773	0.780	-0.007
	[158]	[0.669]	[0.583]	[0.546]		[0.566]	[0.508]		[0.464]	[0.450]	
index for short locus of control*	3229	0.641	0.630	0.652	-0.023***	0.624	0.657	-0.033***	0.639	0.645	-0.006
	[158]	[0.168]	[0.149]	[0.145]		[0.141]	[0.136]		[0.133]	[0.125]	
● agrees: if I work hard, I will succeed*	3302	0.570	0.553	0.588	-0.034**	0.548	0.601	-0.053**	0.562	0.564	-0.003
	[158]	[0.812]	[0.761]	[0.614]		[0.646]	[0.577]		[0.615]	[0.560]	
index for agrees with corporal punishment for any reason	4936	0.226	0.231	0.221	0.010	0.232	0.225	0.007	0.230	0.212	0.017
	[158]	[0.355]	[0.291]	[0.265]		[0.257]	[0.245]		[0.196]	[0.196]	
● agrees to punish daughters who don't obey	2457	0.657	0.667	0.647	0.020	0.654	0.640	0.014	0.690	0.661	0.030
	[158]	[0.860]	[0.723]	[0.690]		[0.645]	[0.637]		[0.536]	[0.559]	
adolescent (11-14) experiences corporal punishment at any place	1938					0.552	0.643	-0.091***			
	[158]					[0.566]	[0.577]				
● experiences corporal punishment at community center	4117					0.114	0.100	0.014			
	[158]					[0.466]	[0.373]				
Social Norms											
index for positive items on gender role models mentioned	7362	0.223	0.236	0.209	0.027***	0.235	0.207	0.028***	0.236	0.212	0.024
	[158]	[0.283]	[0.265]	[0.226]		[0.241]	[0.203]		[0.211]	[0.195]	

Continued on next page

Table 3.1 – *Continued from previous page*

		(I) All			(II) Adolescents (11-23)			(III) Parents		
		Females	Males	T-test	Girls	Boys	T-test	Mothers	Fathers	T-test
	Obs./ Clusters	Mean/ SD	Mean/[SD]	Mean/[SD]	Diff.	Mean/[SD]	Mean/[SD]	Diff.	Mean/[SD]	Mean/[SD]	Diff.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
● agrees: daughters should have same chance to work as sons	3661	0.654	0.687	0.621	0.066**	0.687	0.617	0.070***	0.686	0.628	0.058
	[158]	[0.774]	[0.664]	[0.637]		[0.594]	[0.562]		[0.554]	[0.577]	
index for positive attitudes about marriage	7347	0.457	0.453	0.461	-0.007**	0.457	0.457	0.000**	0.447	0.469	-0.022
	[158]	[0.476]	[0.394]	[0.365]		[0.355]	[0.329]		[0.272]	[0.264]	
● agrees: girls should finish secondary education before marriage	7155	0.576	0.579	0.572	0.007	0.579	0.575	0.004	0.578	0.566	0.013
	[158]	[0.992]	[0.834]	[0.802]		[0.737]	[0.699]		[0.607]	[0.638]	
index for recognized child rights	6574	0.468	0.443	0.494	-0.052***	0.445	0.497	-0.053***	0.439	0.488	-0.049***
	[158]	[0.525]	[0.430]	[0.395]		[0.410]	[0.363]		[0.285]	[0.274]	
● knows child right: right to education	6574	0.839	0.821	0.858	-0.038***	0.823	0.868	-0.046***	0.817	0.838	-0.021
	[158]	[0.696]	[0.633]	[0.486]		[0.540]	[0.470]		[0.491]	[0.400]	
Communication											
index for situations in which parents consider child's opinion	7353	0.248	0.209	0.289	-0.080***	0.192	0.301	-0.109***	0.242	0.265	-0.023***
	[158]	[0.254]	[0.215]	[0.210]		[0.196]	[0.173]		[0.164]	[0.185]	
● parents consider child's decision on the timing of marriage	6435	0.416	0.304	0.536	-0.232***	0.269	0.565	-0.296***	0.362	0.477	-0.114***
	[158]	[0.992]	[0.745]	[0.831]		[0.683]	[0.743]		[0.567]	[0.650]	
● parents consider child's decision on going to school	7120	0.706	0.631	0.785	-0.154***	0.606	0.814	-0.207***	0.674	0.727	-0.053
	[158]	[0.772]	[0.676]	[0.638]		[0.648]	[0.546]		[0.515]	[0.543]	
index for interpersonal communication competence*	3130					0.577	0.593	-0.015***			
	[158]					[0.150]	[0.144]				
● interpersonal communication competence: empathy (1/5)*	3623					2.694	2.761	-0.067**			
	[158]					[1.208]	[1.229]				
Panel D: Self-Reported Program Uptake											
awareness of events (only UNICEF[n1])	7384	0.133	0.129	0.136	-0.007	0.121	0.122	-0.001	0.144	0.166	-0.022
	[158]	[0.864]	[0.680]	[0.627]		[0.539]	[0.473]		[0.589]	[0.609]	
aware of event type: group session	7384	0.075	0.074	0.075	-0.001	0.070	0.066	0.004	0.083	0.095	-0.012
	[158]	[0.674]	[0.538]	[0.482]		[0.450]	[0.367]		[0.454]	[0.477]	
aware of event type: community meeting	7384	0.062	0.064	0.060	0.004	0.059	0.051	0.008	0.075	0.079	-0.004
	[158]	[0.592]	[0.492]	[0.413]		[0.402]	[0.332]		[0.418]	[0.390]	
aware of event type: theater session	7384	0.035	0.032	0.040	-0.008	0.028	0.036	-0.008	0.038	0.048	-0.009
	[158]	[0.352]	[0.258]	[0.312]		[0.208]	[0.265]		[0.233]	[0.260]	
heard about NFBECs in revenue villages	7401	0.196	0.187	0.206	-0.019**	0.184	0.195	-0.011	0.193	0.229	-0.036***
	[158]	[1.251]	[0.968]	[0.894]		[0.826]	[0.734]		[0.697]	[0.680]	
participated in event	980	0.556	0.566	0.546	0.020	0.545	0.565	-0.019	0.599	0.518	0.081
	[126]	[0.911]	[0.808]	[0.769]		[0.638]	[0.619]		[0.703]	[0.745]	
ever attended/ currently attending NFBE Center	3754					0.046	0.053	-0.006			
	[158]					[0.300]	[0.368]				
● answered question: NFBE Center attendance	1130	1.000	1.000	1.000		1.000	1.000		1.000	1.000	
	[64]	[0.000]	[0.000]	[0.000]		[0.000]	[0.000]		[0.000]	[0.000]	
Panel E: Background											
age at endline	7415	26.028	25.404	26.690	-1.286	15.588	16.210	-0.622***	43.647	48.228	-4.580***
	[158]	[16.063]	[15.386]	[13.139]		[3.854]	[3.982]		[9.233]	[9.536]	
RV: household size	7415	7.466	7.475	7.457	0.019	7.457	7.436	0.021	7.509	7.499	0.010
	[158]	[6.685]	[4.628]	[4.886]		[3.924]	[4.178]		[2.597]	[2.636]	
RV: share household head with no education	7415	0.593	0.593	0.592	0.001	0.587	0.590	-0.003	0.604	0.596	0.008
	[158]	[0.915]	[0.670]	[0.632]		[0.544]	[0.519]		[0.407]	[0.380]	
RV: ⊘ poverty score	7415	17.349	17.343	17.355	-0.012	17.382	17.375	0.007	17.270	17.314	-0.043
	[158]	[12.859]	[9.199]	[9.087]		[7.954]	[7.534]		[4.977]	[5.253]	

▸ *Notes*: Table 3.1 displays the summary statistics. The averages displayed are not weighted.
▸ Source: Endline survey.
▸ Part I considers the full sample, Part II restricts the sample to adolescent respondents only, Part III to parents only. Columns (1) and (2) present the number of observations and the mean and standard deviation of each variable (outcome of interest) for the whole sample. Columns (3) and (4) report the mean and standard deviation by gender separately, followed by t-test differences (5). Columns (6) and (7) present the mean and standard deviation for girls and boys (aged 11-23) separately, followed by corresponding t-test differences (8). Columns (9) and (10) present the mean and standard deviation for mothers and fathers separately, followed by corresponding t-test differences (11).
▸ * indicates that parents were asked the question about their child (and not themselves). The child marriage and education variables were generated from the household roster information, whereas all remaining variables were only generated for respondents of the full survey.
▸ Bullet points indicate variables that are part of the preceding, composite indicators. For example, "share if negative education-related attitudes" is a composite indicator, of which "agrees: educating boys is more important than girls" is part.
▸ Differences in sample size (Obs.) can be due to the fact that some respondents refused to answer certain questions, some questions were not asked to all respondents, and some indicators restrict the sample to certain age groups. For example, the sample size for the indicator "marriage due to any other reason than love" is smaller because this was only asked for married individuals.

I.B   Program Implementation and Exposure

In this section, UNICEF MIS (until November 2019) and endline survey data is used to explore program implementation and exposure in program-assigned (treatment) areas.[5]   Secondly, program implementation and uptake is compared between program-assigned and non-assigned areas, i.e., between the treatment and control group.

I.B.1   Program-Assigned Areas

Implementing Partners   UNICEF worked with several local IP and government departments. For the Sindh region, Rural Support Programmes Network (RSPN) was the main IP, who further cooperated with Sindh Rural Support Organization (SRSO). Indus Resource Center (IRC), an NGO with expertise in the education sector, implemented the NFBE program in Sindh in cooperation with the Directorate of Literacy with Non-Formal Education, School Education and Literacy Department, Government of Sindh. In Punjab, College of Youth Activism and Development (CYAAD) was initially the partner but, after roughly a year, the contract was terminated. After a time gap of another year, Bunyaad Literacy Community Council (BLCC) was selected as the new IP for the remaining time period and completed the project.

Evidence from Monitoring and Information System Records: Child Protection (CP) Activities   The available MIS data is reported in Table 3.2, divided into adolescent-centric (AC) (Panel A) and community-centric (CC) activities (Panel B), for all RVs, RVs in Punjab and RVs in Sindh.[6] Table 3.2 shows that AC activities were implemented in Sindh, e.g., there were on average 7.9 adolescent champions per RV.[7] With on average zero adolescent champions in a village, the MIS evidence for Punjab might be either indicating a gap in implementation or incomplete, which is likely given the changes in implementing partners described above. In fact, official documents suggest that adolescent champions were, indeed, also trained in Punjab (on average 11.1 per RV, UNICEF (2019)). Turning to CC events, a greater variety of topics were covered in Sindh and more beneficiaries reached. In both provinces, the most important topic was education (followed by social issues in Punjab, health, and sanitary facilities in Sindh).

Evidence from Self-Reported Endline Data: Child Protection (CP) Activities   Table 3.3 presents implementation and compliance indicators for the treatment group by main respondent (adolescents, parents, community leaders, and Nikah Khwans). For this evidence self-reported information from the endline survey is considered.[8] To capture whether respondents were aware of any Child Protection event in their RV, they were asked to indicate which of the following events took place in their RV in the last four years: (1) Training, workshop, or group session, (2) community meeting, (3) and theater session (e.g., street theater). Further answers were captured in an open "other" category. For the coding of the outcome indicator, the evaluation considers only those responses Child Protection events (and thus part of the IALP) where the respondents additionally mentioned UNICEF or its partners as the implementing partners.[9] Notably, respondents may have had problems to recall the events. Thus, it was attempted to address this problem by encouraging the enumerators to use additional techniques so that respondents are encouraged to remember the events.[10]

Panel A of Table 3.3 captures the awareness of any such Child Protection event in the treatment group: 13.3% of all respondents were aware of such events. The awareness variables on the village level show that in more than 80% of all RVs at least one main respondent was aware of a Child Protection event. Panel B captures characteristics of the indicated events. The number of observations corresponds to the respondents being aware of at least one Child Protection event, regardless whether they participated in the event or not. 42.5% were invited to participate in the event with the highest share for Nikah Khwan. Figure 3.1a show that the main topics covered in events were health (51%), education (50.6%),

---

[5]Since the available MIS data captures the information only until a certain point in time and was only recorded via an online platform, further information about the IALP program implementation might be missing. For further information, see Section A.III. in Appendix.

[6]Further tables covering MIS are presented in Section A.III.B.

[7]The most prominent topics discussed at AC-related events in Sindh were identity and self-esteem, empathy and respect, communication and expression, sports as well as awareness and rights.

[8]Table A.11 in the Appendix covers all indicators gathered from the endline survey related to program awareness and implementation.

[9]For this, the questions about the official name, the organizer and the topic of the event was used.

[10]Indeed recall problems may have occurred, but (1) given that the majority of interventions happened towards 2019 (Please see number of direct beneficiaries in UNICEF 2020) and (2) given that there is no reason why the recall bias should have occurred more likely in treatment areas, it is believed the problem is not biasing the impacts of interest. Enumerators were trained to probe especially when asking questions in the participation module, e.g., enumerators were asking about any joint big gatherings conducted at community centre or a central place at community during last four years where community men, women and adolescent gathered or adolescent issues were discussed.

Table 3.2: Child Protection Interventions: MIS Evidence by Topics

	(I) All			(II) Punjab			(III) Sindh		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
Panel A: AC Activities									
RV-level ⊘ # of adolescent champions	3.368	4.273	446	0.000	0.000	255	7.864	2.686	191
Indicator for any adolescent champion in the RV	0.424	0.495	446	0.000	0.000	255	0.990	0.102	191
RV-level ⊘ # of adolescent members	6.796	8.532	446	0.000	0.000	255	15.869	5.074	191
Indicator for Trickle Down	0.460	0.499	446	0.059	0.236	255	0.995	0.072	191
Indicator for Action Plan	0.648	0.478	446	0.388	0.488	255	0.995	0.072	191
Workshop Topics									
Indicator for Topic B1 - Identity & Self-Esteem	0.453	0.498	446	0.059	0.236	255	0.979	0.144	191
Indicator for Topic B2 - Empathy & Respect	0.460	0.499	446	0.059	0.236	255	0.995	0.072	191
Indicator for Topic B3 - Communication & Expression	0.460	0.499	446	0.059	0.236	255	0.995	0.072	191
Indicator for Topic B4 - Coping with Stress and Managing Emotions	0.460	0.499	446	0.059	0.236	255	0.995	0.072	191
Indicator for Topic C - Gender Issues and Rights	0.318	0.466	446	0.004	0.063	255	0.738	0.441	191
Indicator for Topic D - Health and Safety	0.000	0.000	446	0.000	0.000	255	0.000	0.000	191
Indicator for Topic E - Education and Skills	0.031	0.175	446	0.055	0.228	255	0.000	0.000	191
Indicator for Topic F - Others/Unknown	0.413	0.493	446	0.004	0.063	255	0.958	0.201	191
Indicator for Topic G - Sports	0.430	0.496	446	0.055	0.228	255	0.932	0.253	191
Indicator for Topic H - Arts & Drama	0.244	0.430	446	0.000	0.000	255	0.571	0.496	191
Indicator for Topic I - Sessions, Awareness & Rights	0.448	0.498	446	0.043	0.204	255	0.990	0.102	191
Indicator for Topic J - Health & Others	0.222	0.416	446	0.282	0.451	255	0.141	0.349	191
Panel B: CC Activities									
Indicator for KCI Rollout Mobilization	0.424	0.495	446	0.004	0.063	255	0.984	0.125	191
Indicator for Community Mobilization	0.291	0.455	446	0.271	0.445	255	0.319	0.467	191
Indicator for Inter-Generation Dialogue	0.372	0.484	446	0.329	0.471	255	0.429	0.496	191
Indicator for Review Meetings	0.000	0.000	446	0.000	0.000	255	0.000	0.000	191
Indicator for Street Theaters	0.487	0.500	446	0.314	0.465	255	0.717	0.452	191
Indicator for Text Messages	0.204	0.403	446	0.004	0.063	255	0.471	0.500	191
Indicator for KCI Training	0.040	0.197	446	0.067	0.250	255	0.005	0.072	191
Workshop Topics									
Indicator for Topic L - Education	0.265	0.442	446	0.204	0.404	255	0.346	0.477	191
Indicator for Topic M - Transport & Energy	0.045	0.207	446	0.000	0.000	255	0.105	0.307	191
Indicator for Topic N - Health & Sanitary Facilities	0.108	0.310	446	0.020	0.139	255	0.225	0.419	191
Indicator for Topic O - Sports & Play	0.031	0.175	446	0.004	0.063	255	0.068	0.253	191
Indicator for Topic P - Social Issues and Others	0.137	0.344	446	0.169	0.375	255	0.094	0.293	191

▸ *Note*: Table 3.2 displays MIS data on the village level for selected variables.
▸ Source: MIS data until November 2019.
▸ Note, MIS data might be incomplete as it captures the information only until a certain point in time (until November 2019). This is why qualitative documents are also considered (UNICEF 2019).

safety (44.2%), consequences of child marriage (39.4%) and child rights (34.2%).[11] Most invitations came from NGOs (38.6%), a befriended adult (29.3%) and from an acquainted or befriended adolescents (29.1% and 25.5%, respectively, Figure 3.1b). The participants' gender across all age groups was mostly mixed; most events were attended by a mix of age groups (Figure 3.2).

Figure 3.1: Reported Program Characteristics

a) Topics of Events

health (incl. reproductive health)	51.
education	50.6
safety	44.2
consequences of early marriage	39.4
child rights and entitlements	34.2
vocational skills	19.8
consequences of teenage pregnancy	19.5
sports/ play	19.2

Percent of Respondents
N=1196

b) Invitation

invited by NGO/ comm. org. representative	38.6
invited by family/friend (adult)	29.3
invited by acquaintance (A)	29.1
invited by family/friend (A)	25.5
invited by CL	23.0
invited by acquaintance (adult)	21.8
invited by teacher/ other school officials	16.8

Percent of Respondents
N=508

▶ *Notes*: Figure 3.1a shows the percentage distribution of different topics covered in project activities.
Figure 3.1b shows the percentage distribution of answers related to the question by whom participants were invited.
▶ Sample: Endline survey. Topic of events - Respondents that were aware of at least one CP event.
Invitation - Respondents that were aware of at least one CP event and were invited to CP event.
▶ Question: Topic of events - "In general, what topics or skills have been covered?" Invitation - "Were you personally invited to participate?"
▶ Answer options: Topic of events - identity/ self-esteem/ self-confidence, empathy/ respect, communication/ expression,
coping with stress/ managing emotion, gender issues/ rights, health, safety/ sanitary facilities, education, sports/ play,
journalism/ arts/ drama, transport/ energy, child rights/ entitlements, consequences of teenage pregnancies, consequences of early marriage,
life skills, vocational skills, preparation of an action plan. Invitation - family member or friend (adolescent) family member/ friend (adult),
acquaintance (adolescent), acquaintance (adult), community leader, non-governmental organization/ community-based organization representative,
teacher/ other school-officials, health worker, religious leader/ nikah khwan, police, ombudsman, media, other local/ governmental authority.
Refuses to answer, does not know and other (specify) were answer options for all questions.
▶ Question type: Topic of events - multiple choice question (m.c.). Invitation - m.c..

---

[11]Answered by respondents that were aware of at least one Child Protection event.

Table 3.3: Child Protection Activities: Awareness, Uptake, and Self-reported Consequences in Program-Assigned Areas

	(I) All		(II) Adolescents		(III) Parents		(IV) Community Leader		(V) Nikah Khwan	
	(1) Share	(2) N	(3) Share	(4) N	(5) Share	(6) N	(7) Share	(8) N	(9) Share	(10) N
Program Delivery (MIS)	0.85	486								
Program Delivery (Self-reported)	0.83	486	0.44	482	0.35	482	0.74	429	0.17	409
Program Delivery (MIS and Self-reported)	0.96	486								
Panel A: Program Awareness										
awareness of events (only UNICEF[n1])	0.18	6466	0.13	3753	0.17	1864	0.74	429	0.16	420
aware of event type: group session	0.13	6466	0.08	3753	0.09	1864	0.70	429	0.11	420
aware of event type: community meeting	0.11	6466	0.06	3753	0.09	1864	0.64	429	0.05	420
aware of event type: theater session	0.06	6466	0.03	3753	0.05	1864	0.38	429	0.05	420
aware of event type: other	0.03	6466	0.04	3753	0.03	1864	0.00	429	0.03	420
Panel B: Program Characteristics										
invited to event	0.42	1196	0.56	502	0.57	309	0.00	316	0.78	69
who else participated[m.c.]										
other participants of event were family/friend (adolescent)	0.24	1196	0.35	502	0.32	309	0.00	316	0.25	69
other participants of event were family/friend (adult)	0.27	1196	0.37	502	0.36	309	0.00	316	0.41	69
other participants of event were acquaintance (adolescent)	0.29	1196	0.40	502	0.45	309	0.00	316	0.19	69
other participants of event were acquaintance (adult)	0.25	1196	0.31	502	0.41	309	0.00	316	0.22	69
other participants of event were NGO/ community (comm.) organisation (org.) representative	0.19	1196	0.25	502	0.29	309	0.00	316	0.19	69
Panel C: Participation and Self-reported Consequences										
participated in event	0.69	1196	0.56	502	0.57	309	0.97	316	0.86	69
would have liked to participate	0.50	374	0.54	223	0.48	132	0.00	9	0.40	10
reasons for not participating[m.c.]										
did not participate due to disinterest in topic	0.22	374	0.23	223	0.21	132	0.00	9	0.20	10
did not participate due to lack of guardian's permission	0.19	374	0.21	223	0.17	132	0.00	9	0.10	10
did not participate due to transport	0.09	374	0.09	223	0.11	132	0.00	9	0.00	10
did not participate due to domestic work	0.39	374	0.38	223	0.42	132	0.00	9	0.40	10
did not participate due to study for school	0.13	374	0.13	223	0.15	132	0.00	9	0.00	10
did not participate due to caregiving for family member	0.07	374	0.04	223	0.12	132	0.00	9	0.00	10
where[m.c.]										
participation in village (vill.) or city of residence	0.57	822	0.92	279	0.95	177	0.00	307	0.69	59
participation in vill. or city in this union council	0.05	822	0.07	279	0.06	177	0.00	307	0.17	59
last participation (months ago)	17.07	494	18.05	267	15.96	168	.	0	15.78	59
total times joined event in last 4 years	3.13	822	5.08	279	4.48	177	0.00	307	6.14	59
implementation (apart from facilitators)[m.c.]										
implementation of event by an adolescent	0.47	822	0.52	279	0.51	177	0.43	307	0.31	59
implementation of event by an adult	0.27	822	0.42	279	0.39	177	0.08	307	0.20	59
implementation of event by a respondent	0.39	822	0.07	279	0.10	177	0.90	307	0.15	59
discussed event with at least 1 other person	0.44	822	0.68	279	0.67	177	0.00	307	0.90	59
satisfaction with event[m.c.]										
satisfied with event	0.98	822	0.97	279	0.97	177	0.99	307	1.00	59
changes in[m.c.]										
self-reported personality changed due to event	0.36	822	0.61	279	0.56	177	0.00	307	0.49	59
self-reported behavior changed due to event	0.35	822	0.53	279	0.58	177	0.00	307	0.68	59
self-reported no changes due to event	0.07	822	0.12	279	0.10	177	0.00	307	0.14	59

▸ *Note*: Table 3.3 displays shares and totals for the awareness, uptake, and self-reported consequences questions in the endline survey at the individual level for the treated population.

▸ Sample: Endline survey. Respondents are As, Ps, CPs and Ns.

▸ The table is restricted to the treated population and sections II through IV are split by respondent type: adolescents, parents, community leaders, and Nikah Khwans respectively.

▸ Section I displays for the all individuals their shares (column (1)) and totals (column (2)) for the different variables. The succeeding sections restrict the sample to the treated adolescents (II), to the treated parents (III), and to the treated community leaders (IV). For each of these sections shares are described in the first column within the section and totals in the second one. All of the variables are at the individual level, except for the awareness variables in rows two and four.

▸ Only respondents that were aware of at least one event were asked the questions displayed in the table. In the survey, not all of the questions were asked to all of the study participants. Therefore, the variables presented in the table are restricted to subgroups of respondents. All of the variables on the table are restricted to study participants that were aware of at least one CP event, i.e., variables are restricted to the first row (14.4). Panel C was only asked to event participants, i.e., restricted to the individuals that participated in any CP event (811). Other variables are also restricted in an intuitive way. For instance, *invited to event by* was only asked to individuals who were invited to an event, *reasons for not participating* to individuals who did not participate, etc.

▸ [n1]: as only UNICEF should be understood UNICEF and its implementation partners. Those implementation partners were RSPN, IRC, CYAAD, BLCC, SRSO, NRSP, Bunyad Literacy Community Council (Bunyad), and the Directorate Government of Sindh. A question about the organizer of the event is used in the survey for this. Additionally, the question asked to the people that were aware about the official name of the event is employed.

▸ [m.c.]: these questions were multiple choice. Some of them allowed the respondent to give multiple answers (*invited to event by whom*, *topics of the event*, *who else participated*, *reasons for not participating*, *organizer of the event*, *implementation of the event*, *discussed with whom*, *satisfaction with event*, and *changes in*), while others only allowed the selection of one option (*age group of participants*, *gender of participants*, and *where did the event take place*).

▸ Other similar tables are Table 3.3, Table A.11, Table A.12, Table A.13, Table A.14, Table A.17, and Table A.18.

Figure 3.2: Reported Participant Characteristics
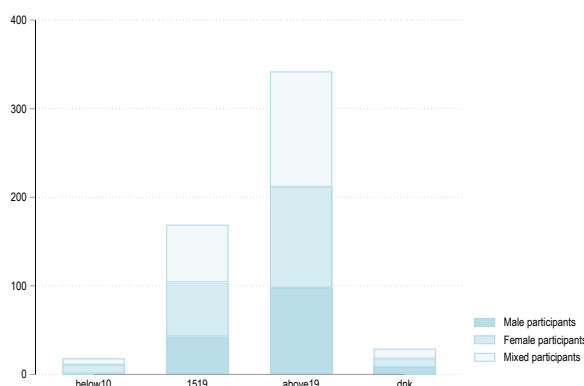
► *Notes*: Figure 3.2 depicts on the x-axis the participants' age grouped by categories and on the y-axis number of respondents.
The different blue shaded areas present the participants' gender.
► Sample: Endline survey.
► Question: Age - "What was (usually) the age group of most of the participants?" Gender - "What was the gender of the participants?"
► Answer options: Age - younger than 10 years, age 10 to 14, age 15 to 19, older than 19, all age groups (mixed).
Gender - male, female, mixed (male and female).
Refuses to answer and does not know were answer options for both questions.
► Question type: Age - m.c.. Gender - single choice.

Panel C of Table 3.3 reports details about the events and self-reported consequences. Only respondents that participated in the event were asked the following questions. Around 60% across all respondents participated in an Child Protection event, whereby the share for adolescents is 55.6% and the share of community leaders is .%. Around 50% of adolescents would have liked to participate (50.3% for all respondents). The main reason that adolescents did not participate was domestic work with 37.7%. The reported events took place in their own villages. Apart from facilitators, the events were mostly implemented by an adolescent (as reported by 46.8% across all respondent types). The survey identified 6.8% as adolescent champions (who implemented the event).[12]

After taking part in the event, most attendees discussed the event with family and friends. 67.7% and 66.7% of adolescents and parents report having discussed the event with another person. The percentage for Nikah Khwans is even higher at 89.8%.[13] Finally, satisfaction across all respondents who eventually participated in at least one event was high, 98.1% (considering individuals who were fully or somewhat satisfied with the event). Only few respondents (7.2%) reported that the participation at the event did not change their behavior.

Evidence from Monitoring and Information System Records: NFBE Centers (NFBECs)   Relying on information from the MIS, it can be seen that 62.5% of NFBE Centers were offering special space for girls education. Most NFBEs were established in community places (71.3%), followed by their physical integration in schools (Table O.2). The number of male participants was on average 7.7, with a standard deviation of 14.21. A session was on average about 3.3 hours long, with only minor differences in the two implementing districts. In every third NFBE village, a Community Participation Event took place.

Evidence from Self-Reported Endline Data: NFBE Centers (NFBECs)   Table 3.4 presents self-reported implementation and compliance indicators for the NFBE-assigned program areas.[14] Respondents (adolescents in AQ and PQ) were asked if they ever heard about the NFBE program in their village.[15] 34.1% report having heard about the program in program-assigned areas (N=1,465). The remaining indicators in the table are restricted to individuals who have heard

---

[12]The low percentage might be due to the random sampling approach which would in this care reflect that on average 6.8% of all adolescents in a program villages were champions. In a village with 500 adolescents, for instance, this would amount to 34 adolescents champions responsible to spread the messages and mobilize further adolescents.

[13]Of the individuals that discussed the event with someone else, 58.6% discussed it with an adolescent family members or an adolescent friends, and 53.1% with an adult family members or an adult friends.

[14]Table A.15 in the Appendix displays the full range of questions asked in the survey related to the awareness and uptake of the NFBE treatment.

[15]In the survey the study participants were asked "Have you heard about the existence of non-formal basic education centers in your revenue village?".

Figure 3.3: Discussion
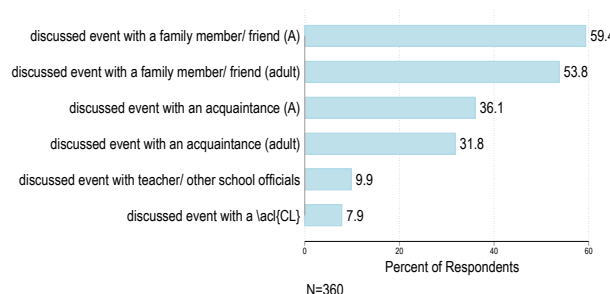
► *Notes*: Figure 3.3 shows the percentage distribution
of answers to the question with whom the project activity/ event was discussed.
► Sample: Endline survey. Respondents that participated in at least one CP event.
► Question: "After your participation, did you usually discuss the topics covered in the event with others? Who did you talk to?"
► Answer options: Family member or friend (adolescent), family member/ friend (adult), acquaintance (adolescent), acquaintance (adult), community leader, non-governmental organization/ community-based organization representative, teacher/ other school-officials, health worker, religious leader/ nikah khwan, police, ombudsman, media, other local/ governmental authority.
Refuses to answer, does not know and other (specify) were answer options for this question.
► Question type: m.c..

about the program.[16] Of the adolescents that heard about NFBE, 57.7% were reportedly invited to take part.[17] 45% of the adolescents aware considered joining a NFBE Center (N=140), while 23.8% in our sample reportedly attended a center (N=122).[18]

> **"Heterogeneous Effects"**
>
> The program may have had different effects for different individuals. "Heterogeneity" refers to differences across the units being studied. In the evaluation, more refined statements about the impact on particular groups of people can therewith be made. Throughout the analysis, the results are not only presented for all respondents jointly, but also distinguish the results by male and female respondents (restricting the sample to the groups when estimating the impacts of the program). The evaluation further focuses on three main types of respondents: (1) respondents residing in Punjab vs. Sindh, which captures potential variation in the different provincial characteristics, the differences in implementing partners, and type of activities conducted (CP only vs. CP and NFBE), (2) respondents residing in poor vs. very poor households, and (3) respondents reporting different levels of exposure to COVID-19. In further analysis, the evaluation refers to results being split by whether a parent or an adolescent report the information.[a]
>
> [a]Please note that the heterogeneity analysis focuses on the most relevant heterogeneous effects. Further analyses suggested in the PAP will be conducted over the course of the coming months, i.e, were beyond the scope of the current time frame. Also, since community leader interviews were entirely conducted over phone, some key variables are missing and thus, sample sizes are too small.

I.B.2   Comparison between Treatment and Control Areas

In the following, the self-reported occurrence of selected events are compared to similar events as reported by respondents in non-program areas. That such events are reported in control areas as well is likely to happen since the questions were asked on purpose very generally.[19] The question hereby is whether the likelihood to participate in a Child Protection-related event is higher in program areas than in control areas. High frequencies of reported similar events in control areas may eventually indicate to us that the evaluation may not be capturing the full effects of the program when comparing the prevalence in the outcomes of interest between program and control areas. It can also limit our ability to estimate local treatment average effects, as will be discussed in Section A.IV.A. Whether the evidence on the prevalence of similar events in control areas effects are due to spill-over (e.g., due to mass-media campaigns or

---

[16]Most individuals heard about the program from a neighbor or a relative (39.7%), followed by an adolescent (30.9%), and by a teacher (22.9%).

[17]Individuals were reportedly mostly invited by an NGO or a community organization representative (24.4%), a family member or a friend who was an adolescent (30.2%), and by an acquaintance who was an adolescent (16.3%).

[18]The main reason for not attending were the long distances (10.8%).

[19]Survey questions on program activities and uptake were on purpose asked in a very general way in order to capture the potential prevalence of similar events in control areas.

Table 3.4: NFBE Program Awareness and Attendance in NFBE-Program Assigned Areas

	(1) Share	(2) N
heard about NFBECs in revenue villages	0.34	1880
heard about NFBE through[m.c.]		
heard about NFBECs from an adolescent	0.31	647
heard about NFBECs from a teacher	0.23	647
heard about NFBECs from neighbour or relatives	0.40	647
heard about NFBECs from a NGO	0.16	647
invited to take part in a NFBEC	0.58	149
invited to event by[m.c.]		
invited to event by family/ friend (adolescent)	0.30	86
invited to event by family/ friend (adult)	0.16	86
invited to event by acquaintance (adolescent)	0.29	86
invited to event by acquaintance (adult)	0.24	86
invited to event by community leader	0.22	86
invited to event by NGO/ comm. org. representative	0.36	86
invited to event by teacher/ other school officials	0.08	86
considered attending a NFBEC	0.45	140
attended a NFBEC	0.24	122
did not attend NFBEC because of long distances	0.11	93

▸ *Note*: Table 3.4 displays shares and totals for the NFBE program awareness, atten-
dance, and characteristics questions in the endline survey at the individual level.
▸ Sample: Endline survey. Respondents are As from AQ and PQ.

▸ Column (1) displays shares for the different variables, and column (2) the number of
respondents to the question.
▸ [m.c.]: these questions were multiple choice. Some of them allowed the respondent to
give multiple answers (*heard about NFBEC through*, *invited to event by*, *did not attend
NFBEC because*, *main reasons for attending NFBEC*, *discussed with whom*, and *satis-
faction with NFBEC*), while others only allowed the selection of one option (*participa-
tion in NFBEC took place in*, *gender of participants*, and *age group of participants*).
▸ Related: Table A.15 and Table A.16.

exchange of information between villagers) or whether the events captured in control areas are fundamentally different in character is an investigation beyond the scope of this report.

What can be constituted at this stage is that - in line with the expectations that the program was implemented as planned and that the evaluation team targeted the appropriate program-eligible population in endline survey - a higher level of awareness of UNICEF program activities in program-assigned areas is captured than in control areas. However, the effects are limited in size due to, among others, potential limitations to greater outreach in program areas and the reported prevalence of similar events in control areas.

Program Awareness   Tables 3.5 and 3.6 display the effects from a set of Ordinary Least Squares (OLS) estimation results using endline survey and MIS data on a set of indicators directly or indirectly related to program delivery. Hereby, equation 2.1 presented in Section I.B is estimated. The first indicator combines awareness about the Child Protection and NFBE component of the IALP program (columns (1)-(3)). In all program-assigned villages, except in villages where the treatment modality is Child Protection and not NFBE pool, the awareness about the program is significantly higher (significant at the 1% level) compared to the non-assigned areas (column (1)). The effect is largest for female individuals (column (3)). In total, the IALP program awareness is 4.8pp larger in program than in non-program areas. For the share of Child Protection and NFBE program awareness, further heterogeneity analyses reveal that the observed increases are driven by observations in Sindh, which makes sense given that this program component was implemented only in Sindh (Table A.45). Moreover, the positive effect is larger for individuals living in poor households compared to individuals living in very poor households. Splitting the first awareness indicator into the Child Protection and NFBE component, the results indicate that female respondents have a larger percentage point change in Child Protection awareness than male respondents (columns (6) vs. (5)). This could mean that either female respondents better recalled the implemented events or were more exposed to the program. The larger shares of NFBE awareness in the program-assigned areas stem from male respondents (columns (8) vs. (9)). For both indicators, there are significantly more people aware in program-assigned areas (significant at the 1% level). The observed increases in both variables are driven by observations in Sindh (Table A.45). The last indicator reports the share of program delivery according to the MIS data (column (10)). The results clearly show that only in villages where the treatment modality Child Protection and not NFBE pool was implemented the program was not delivered as assigned (only 76%). This information plays a crucial role in the estimation of the LATE in Section A.IV.A. The heterogeneity results reveal that the non-delivery sources from Punjab while Sindh the delivery was at 100%.

Table 3.5: Program Awareness

	share of program awareness (CP and NFBE)			share of program awareness (CP)			share of program awareness (NFBE)			share of program delivery (MIS)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	All	Male	Female	All	Male	Female	All	Male	Female	All
only NFBE	0.116***	0.110**	0.123**	0.077***	0.052	0.099***	0.123***	0.154***	0.092*	1.000***
	(0.043)	(0.046)	(0.050)	(0.026)	(0.035)	(0.029)	(0.045)	(0.049)	(0.052)	(0.000)
CP and NFBE	0.143***	0.142***	0.139**	0.086***	0.070**	0.100***	0.133***	0.153***	0.108**	1.000***
	(0.042)	(0.041)	(0.054)	(0.027)	(0.034)	(0.029)	(0.043)	(0.042)	(0.054)	(0.000)
CP, NFBE pool	0.101***	0.100***	0.101**	0.095***	0.085**	0.105***	0.064*	0.067*	0.061	1.000***
	(0.036)	(0.037)	(0.045)	(0.030)	(0.036)	(0.034)	(0.034)	(0.037)	(0.042)	(0.000)
CP, not NFBE pool	-0.015	-0.015	-0.017	0.012	0.006	0.017	-0.023	-0.017	-0.032	0.766***
	(0.017)	(0.019)	(0.021)	(0.014)	(0.016)	(0.015)	(0.018)	(0.020)	(0.020)	(0.030)
Control mean	0.223	0.251	0.189	0.126	0.158	0.087	0.036	0.036	0.036	0.000
SD	0.416	0.433	0.391	0.332	0.365	0.283	0.070	0.070	0.070	0.000
CP overall*	0.010	0.010	0.009	0.030	0.023	0.036	-0.005	0.001	-0.012	0.816
CP p-value	0.030	0.041	0.091	0.002	0.025	0.001	0.297	0.240	0.533	0.000
IALP overall**	0.035	0.034	0.034	0.041	0.031	0.049	0.023	0.032	0.011	0.854
IALP p-value	0.001	0.001	0.009	0.000	0.020	0.000	0.007	0.001	0.083	0.000
N	9,232	4,940	4,292	8,519	4,630	3,889	8,642	4,356	4,286	653
RVs	653	653	648	653	653	648	653	653	648	653
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

▸ *Notes*: Table 3.5 displays the program effects on indicators of program delivery. Effects from separate OLS estimations are presented.
▸ Sample: Endline survey. Full sample.
▸ Program effects on the following indicators of program awareness are considered: Share of program awareness, share of Child Protection program awareness, share of NFBE-program take-up, and awareness (UNICEF). Program effects on the following indicators of program implementation are considered under Part II: Invited to event, participated (part.) in intergenerational (intergen.) dialogue, participated in CPF activities, and part. in CPM activities. "Share of program awareness" is captured by asking respondents (adolescents, parents, community leaders, and Nikah Khwans) to indicate which of the following events, that they aware of, took place in their revenue village in the last four years: (1) Training, workshop, or group session, (2) community meeting, (3) and theater session (e.g., street theater). Moreover, adolescents, parents, and community leaders were asked whether they have heard about the existence of non-formal basic education centers in their revenue village. First, a variable indicating the total number of Child Protection-events respondents were aware of (max. three) and the NFBE component are generated. The summary variable then combines awareness of child protection and educational interventions, and indicates, out of the four program components, the share of overall program awareness on a scale from 0 to 1. "Share of CP program awareness" is captured by first generating indicator variables for all three types of Child Protection-events which indicate whether respondents were aware of them individually. Thereafter, a summary variable indicating the share of overall Child Protection-program awareness on a scale from 0 to 1 is generated. "Share of NFBE-program take-up" by first generating an indicator for attending NFBE. This indicator captures having ever attended or currently attending a NFBEC, and attendance was in own revenue village, for adolescents in the relevant age group. The second self-reported index is an indicator for satisfaction with activity. Conditional on having ever attended a NFBEC, adolescents and parents were asked how satisfied, on a scale from 1 (fully satisfied) to 4 (fully dissatisfied), they (or their child) were after their participation in the NFBE-program. The indicator for satisfaction with activity thus indicates full satisfaction (=1) with the program. Taking the two self-reported indices of NFBE-program take-up together, a summary variable indicating the share of overall NFBE-program take-up on a scale from 0 to 1 is constructed. "Awareness (UNICEF)" indicates whether adolescents, parents, community leaders, and Nikah Khwans were able to name the program name or one of the organizers (UNICEF and partners). Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
▸ The equation 2.1 is estimated with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4 \ \frac{83}{386}\beta_3 \ \frac{303}{386}\beta$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.
▸ Related table(s): Table A.45.

Program Uptake   Turning to Table 3.6, self-reported data indicates that 8.3pp (p-value 0.073) more female individuals were invited to events compared to the non-program areas (column (3)). This pp-change is statistically significant and higher than for male respondents. This is in line with the evidence presented above, namely that the change in program-awareness is higher for female than male respondents. The significant change for females is also driving the overall results of a 2.9pp-change. Note, however, that the estimations are made on a subset of villages; those with at least some reported evidence on program awareness. In particular the sample for females is estimated using evidence from only 285 villages. The share of individuals participating in an event related to Child Protection is 4.4pp larger (with p-value of 0.011) in program-assigned than non-program areas (column (4)). The pp-change is higher for female than male respondents which is equally line with the before-mentioned results about invitation and awareness. The variable "share participated with other female activities" and "share participated with other male activities" indicates whether respondent participated in any of the Child Protection-program events where the gender of the other participants was female or male, respectively (columns (7)-(12). Overall, the indicator "share participated with other female activities" is significantly larger in program-assigned than non-program assigned areas. Yet especially female respondents indicate that they attended events where other participants were also female (column (9)). And this share is higher in program than in control areas.

Table 3.6: Child Protection Activities Uptake

	share invited to event			share part. in event			share part. with other female participants			share part. with other male participants		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
only NFBE	0.086*	0.019	0.167**	0.128***	0.090*	0.176**	0.109**	0.008	0.214**	-0.019	-0.030	0.028
	(0.047)	(0.046)	(0.077)	(0.048)	(0.050)	(0.082)	(0.047)	(0.043)	(0.094)	(0.055)	(0.085)	(0.020)
CP and NFBE	0.045	0.050	0.039	0.050	0.067	0.032	0.028	-0.028	0.114	-0.004	-0.023	0.024
	(0.042)	(0.043)	(0.082)	(0.043)	(0.046)	(0.085)	(0.033)	(0.026)	(0.074)	(0.053)	(0.079)	(0.016)
CP, NFBE pool	0.060	0.029	0.096	0.102**	0.076	0.143*	0.030	-0.024	0.089	-0.010	-0.011	0.014
	(0.037)	(0.043)	(0.069)	(0.044)	(0.047)	(0.076)	(0.030)	(0.026)	(0.071)	(0.056)	(0.084)	(0.017)
CP, not NFBE pool	0.009	-0.025	0.076	0.015	-0.016	0.099	0.019	-0.009	0.122*	-0.026	-0.041	0.006
	(0.033)	(0.037)	(0.076)	(0.034)	(0.033)	(0.071)	(0.023)	(0.023)	(0.065)	(0.042)	(0.059)	(0.026)
Control mean	0.391	0.337	0.513	0.672	0.745	0.513	0.100	0.046	0.220	0.263	0.375	0.014
SD	0.488	0.474	0.502	0.469	0.437	0.502	0.297	0.209	0.411	0.439	0.483	0.106
CP overall*	0.020	-0.013	0.081	0.034	0.004	0.108	0.021	-0.012	0.115	-0.023	-0.035	0.008
CP p-value	0.156	0.948	0.090	0.036	0.311	0.020	0.188	0.342	0.029	0.601	0.617	0.489
IALP overall**	0.029	-0.003	0.083	0.044	0.019	0.104	0.029	-0.013	0.123	-0.020	-0.033	0.012
IALP p-value	0.049	0.486	0.073	0.011	0.075	0.039	0.043	0.534	0.012	0.710	0.672	0.127
N	1,447	883	564	1,447	883	564	1,447	883	564	1,447	883	564
RVs	514	454	285	514	454	285	514	454	285	514	454	285
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

‣ *Notes*: Table 3.6 displays the program effects on indicators of program delivery. Effects from separate OLS estimations are presented.
‣ Sample: Endline survey. Full sample.
‣ Program effects on the following indicators of program awareness are considered under Part II: "Invited to event" is generated by asking respondents (adolescents, parents, community leaders, and Nikah Khwans) whether they were personally invited to participate, conditional on them being aware of the event. "Share part. in event" is captured by asking respondents (adolescents, parents, community leaders, and Nikah Khwans) whether they participated in a Child Protection event. "Share part. with other female participants" is generated by asking respondents (adolescents, parents, community leaders, and Nikah Khwans) to indicate what the gender of the other participants was. The variable thus indicates whether respondent participated in any of the Child Protection-program events where the gender of the other participants was female. "Share part. with other male participants" indicates whether respondents participated in any of the Child Protection-program events where the gender of the other participants was male. Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
‣ We estimate the equation 2.1 with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
‣ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$ $\frac{83}{386}\beta_3$ $\frac{303}{386}\beta$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
‣ Standard errors are clustered at the union council level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.
‣ Related table(s): Table A.45.

II. Program Effectiveness

In this section, the main evaluation questions are addressed: The main outcomes (child marriage, child pregnancy, education, Section II.A) are presented, followed by a closer investigation of potential drivers of change (adolescent empowerment, communication, social norms, and service provision, Section II.B). All results are further disaggregated by comparing the program's impacts in specific geographical areas (Non-Formal Basic Education (NFBE) eligible areas, Sindh vs. Punjab), by implementation characteristics (variation in target group and composition of core activities), and by household characteristics (poverty level, COVID-19 exposure). The results reported here are - with minor deviations - robust to alternative estimation specifications, such as removing all baseline control variables, including all variables that are randomized on, and clustering the standard errors at the household instead of the community level (Section II.C). The section ends by addressing potential limitations (Section II.C) and by relating the program's impacts to program costs (Section II.D).

Assistance in Reading and Interpreting the Estimation Tables

A. Statistical Significance

In this evaluation, the program effects are assessed based on a sample of individuals, not the entire population. Is the effect calculated true for the whole population? Or is it simply captured by chance? In other words, what is the likelihood the results are caused by chance (and not the program)?

To answer these questions, many factors are considered such as the sample size (which might be too small to generalize effects to the whole population with much confidence), the size of the observed effect, etc.

Moreover, it is common practice to display different levels of confidence, marking the results according to the **p-values** (a p-value is the probability that random chance could explain the results; thus, the lower the probability, the better; in this evaluation: The more stars one observes, the more confident one can be about the results). While all results are displayed (significant and not significant ones), to remain careful, only those outcomes that are significant are interpreted and discussed (marked with additional stars). In this case, the focus is on the interpretation of outcomes that are **significant at the 5% or 1% significance level (two or three stars)**, since these effects are more robust. They indicate that, with a chance of only 5% or 1%, what is discussed is caused by random chance and not the program itself. If an effect is below this significance level, one has to be more careful with its interpretation given that the observed effect could be due to chance and not actually caused by the program. (Because many outcome variables are considered and at some point some of them would turn out significant by chance, a conservative approach is chosen.)

B. Percentage Change vs. Percentage Points

Both terms, percentage change vs. percentage points, are used to describe the relationship between two data points. However, **percentage change refers to the rate of change**, whereas **percentage point measures the actual amount of change**.

To make this point clearer, please consider the following figure:

30% married, or every third person

decrease in marriage by 20 percentage points (e.g., due to the program) in other words, it falls from 30% to 10% (or 2 fewer people out of 10)

in the areas without the program we may still see an average of 30%, thus the relative change w.r.t. the control group is 20/30 or 67%.
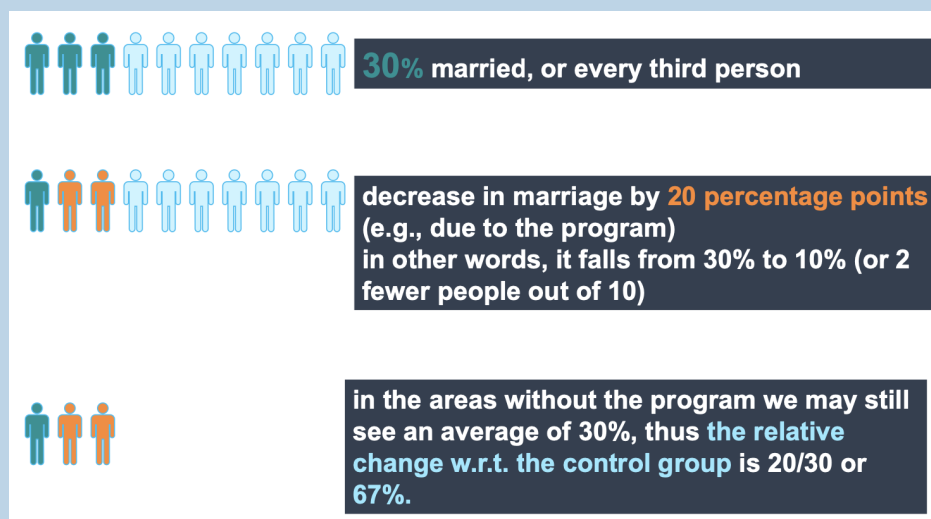
The figure illustrates the difference between **percent, percentage points and percentage change**.

Percent is a number or ratio expressed as a fraction of 100 and means "out of 100". In the example above, 3(0) people out of 10(0) are married. This means that 30%, or every third person, is married, i.e., the marriage rate is 30%.

Now assume that a program reduces the marriage rate by 20pp in treatment areas from its counterfactual value of 30% (i.e., its control mean). In other words, the marriage rate in treatment areas falls from 30% to 10%. One can also say that the program induced 2 fewer people out of 10 to be married.

Since the average marriage rate might still be at 30% in areas without the program, the relative change observed with respect to the control group is $\frac{20}{30} \cdot 100$ or 67%. This means that the program decreased the marriage rate in treatment areas by 67%.

C. Control Mean

The **control mean reports mean values of each outcome variable for the control group**. Since the treatment was randomized, this mean can be interpreted as counterfactual for the treatment group, i.e., without the treatment, the treatment group would have experienced the same mean outcomes. The control mean helps to get a sense of the relative size of the effect.

For example, the IALP program has an overall negative effect of 0.007 on the marriage rate of girls age 10-17, with a control mean of 0.014. This means that treatment decreased, on average, the marriage rate by 0.7pp from its counterfactual value of 0.014%.

II.A   Main Outcomes: Child Marriage, Child Pregnancy, Education

The core, final outcomes of the IALP program focus on reducing child marriages and pregnancies, and increasing education. These outcomes are the focus of the following paragraphs.

II.A.1   Marriage

Key Findings on Marriage Outcomes

One of the strategic goals set in 2015 was to achieve a reduction in child marriage by 10% by the year 2019 (UNICEF South Asia 2020). Consequently, a core evaluation question is whether child marriage rates can be changed in the desired direction and to which extent this is possible.

- Overall, the IALP program has a positive impact on child marriage reduction relative to the low level of child marriage captured in the surveys, especially for girls;
- CP effects are particularly strong in NFBE areas: Here, the IALP program decreases child marriage by 2.2pp for adolescents age 15-17 and by 3.9pp for girls;
- The results indicate a reduction in the prevalence for boys having friends of their age (11-17) and gender who are married, as well as for adolescents with high self-reported COVID-19 exposure;
- Instead, the IALP program increases engagement rates for boys age 10-17, while their marriage rates do not see a change. Accordingly, it increases the (indirectly) reported likelihood that boys (age 11-16) will be married next year. The evidence on changes in engagement and delay of marriage for boys indicates a postponement in line with the original program expectations. The lack of such indirect evidence on the effects amongst girls is not clear and may be related the source of information. For direct information on child marriage rates the household head was asked, while the indirect information on engagement and friends was captured by asking the adolescent girls/ parents of adolescent girls directly;
- The NFBE component of the IALP program reduces the dowry amount paid by adolescents' and parents' own families at marriage by 17.4%.

Table 3.7 displays marginal effects from a set of OLS regression results using endline survey information on child marriage. To capture the child marriage incidence, information from the household roster was consulted: Responses about the age, marital status, and length of marriage for all household members were provided by a household head or, in case absence, a knowledgeable adult (Section II.A). The main results are presented, first for adolescents (following the formal age definition; see Part I), then considering the age group 10-17 (Part II), narrowing down the age group at which most child marriages occur and considering the age of 15-17 in Part III, and finally investigating changes to people who were program-exposed yet are no longer adolescents. Moreover, the impacts of the different treatment arms or program variations are more closely investigated, in the full sample of villages (Panel A and C) or a subset thereof

(Panel B, focusing only on 212 NFBE eligible villages), distinguishing between the different combinations of program components that were implemented on the ground. The aggregated program effects of the IALP program (based on equation 2.3) are presented in row *IALP overall*). Accordingly, the results based on equation 2.2) are presented in row *CP overall*, with the corresponding p-value below.

The IALP program has no impact on the marriage prevalence among adolescents age 10-19. Considering the age group of adolescents (age 10-19, Part I), the results indicate no impacts of the program or its separate components for this age range for the full sample (Panel A).

    The IALP program has a positive impact on child marriage reduction for adolescents age 10-17, especially for girls. The results indicate a reduction in child marriage for girls by 1.4pp in areas where the Child Protection and educational interventions were jointly implemented (column (6), significant at the 5% level). Hereby, this program component is compared with an unconditional control group, i.e., villages where no program was implemented at all and where girls are likely to be married at a rate of 1.4%.[20] When Child Protection interventions are implemented in the NFBE-eligible areas, the results show an overall negative impact of 0.7pp (significant at the 5% level). Moreover, in column (6) the results indicate that the effect is driven by changes for the main target group, namely girls. Here, the negative impact is 1.3pp compared to the control group. Overall, across all treatment effects, the coefficients are negative. The combined IALP program effect captures a negative impact of 0.3pp, with a p-value of 0.049. While the effect may seem small at first, it is large relative to the average prevalence in the control areas, where the child marriage rate is reportedly 0.9% in the full sample. Moreover, in column (6) the results again indicate that the effect is driven by changes for girls. Here, the negative impact is 0.6pp compared to the control group.

    The positive impact of the program on child marriage reduction is particularly strong for adolescents age 15-17. Child marriage is more likely to occur in the age range of 15-17, as discussed before. In Part III of the table the focus shifts to this age group and their reported child marriage rate. As expected, all impacts increase in size. In column (9) the results again indicate that the Child Protection effects are particularly strong in NFBE areas. This may be due to a different way of implementing the program there, but also due to an initially higher vulnerability in those specifically targeted areas.[21]

    The largest overall effect can be observed in areas where Child Protection interventions were implemented in the NFBE-eligible areas. Here, the findings show an overall negative impact of 2.2pp (significant at the 5% level). The effect is large considering a relatively low level of child marriage captured in the control group (1.9%) and then relative to this share a large program impact (of 0.8pp). Moreover, in column (9) the results reveal that the effect is driven by changes for girls. Here, the negative impact is 3.9pp compared to the control group. Girls of age 15-17 have a 3.3% reported probability to marry in control areas. According to UNICEF's strategic goal to reduce child marriage below 18 years of age by 10%, the IALP program exceeds these expectations as child marriage rates overall decrease by 42.1% for the age group of 15-17 year olds relative to the control group and its mean.

    The evaluation captures no impact on the marriage prevalence for the age group 18-23, the program-exposed adolescents and young adults. Herewith, also legal marriages are considered, from age 18 onwards (16 in Punjab for women). The results in Part IV of Table 3.7 show no effects on female household members in this age group. Moreover, the overall IALP program effect is insignificant.

{ Considering 100 adolescents, the IALP program causes two fewer child marriages. }

While Panel A of Table 3.7 considers the effects for the full sample, Panel B takes a closer look at the NFBE-eligible sample (212 villages). One can immediately constitute that for age groups 10-17 and 15-17, the reported child marriage prevalence is higher in this subsample of villages. This is not surprising given that the sample was selected due to its worse educational outcomes, an indicator which correlates with child marriage prevalence. The results reveal that all program impacts hold when restricting to this sample and are in many cases slightly higher in size. For instance, in areas where the Child Protection and NFBE activities where jointly implemented, the child marriage rate drops by 4.5pp for girls in the age group of 15-17 (column (9), control group mean is 5.4%). Relative to this, a 4.5pp negative impact in child marriage corresponds to a reduction of $\frac{4.5}{5.4} \cdot 100$ or 83%.

In Panel C, the full sample of villages is considered again, now with all possible combinations of the three different Child Protection treatment arms. The impacts are estimated relative to the control group in the non-NFBE pool. The

---

[20]Note that the control mean reports the overall child marriage rate in control areas (i.e., across all program variations), whereas the coefficients refer to program effects in specific treatment arms or program variations. Therefore, the child marriage control mean might be higher for specific coefficients. This explains why the percentage point reduction might be equal to or larger than the overall control mean reported in the tables.

[21]See Section I.A for a description on how the NFBE pool was selected.

negative effects for girls age 10-17 and 15-17, in areas where Child Protection and educational interventions were jointly implemented, seem to be driven by the CPF intervention. Similarly, in the age group 15-17 a reduction in female marriages (column (9)) can be constituted in child protection intervention targeting males villages, which are outside of the NFBE-eligible village pool. Moreover, there is some evidence in column (12) that the CPMF intervention, i.e., the one targeting both genders, leads to an increase in child marriage for girls of age 18-23, which may indicate the occurrence of delayed marriages.

In order to capture how effects differ for certain groups, *heterogeneous analyses* are conducted for the following three groups: The estimations are split by province (Sindh vs. Punjab), poverty levels (very poor vs. poor), and self-reported COVID-19 exposure (high vs. low).[22] The heterogeneous analysis reveals that the observed reductions in child marriage among adolescents age 10-17 and 15-17 are driven by observations in Sindh. Moreover, the results indicate that the decrease is larger for individuals living in very poor households compared to individuals living in poor households. The average prevalence of marriage among adolescents age 15-17 in the control areas is reportedly 2.5% for very poor individuals and therewith lower than in poor areas where it is 0.9% in this sample. This difference between the two type of groups is, however, not significant (Table A.47).[23] When comparing households with high vs. low levels of self-reported COVID-19 exposure, the evaluation finds that the reduction in child marriage for the age group 10-19 caused by CP interventions in the NFBE pool is now significant and higher in villages with high COVID-19 exposure. And this difference in effects between high and low COVID-19 exposure now also statistically significant, indicating that the program might have had a higher impact in the group with a high level of COVID-19 exposure. Similar effects are prevalent for other age groups. However, this program effect is hard to pin down causally since endogeneity may play a role here (i.e., it is hard to disentangle the sequence of events of the program impacts involved).

Finally, a number of *robustness checks* for these results are presented in Section II.C analyzing the stability of the observed results to alternative model specifications. The results indicate that the significance level for the child marriage results drops after clustering the standard errors at the village level; though not considerably.

Friends and Engagements   In what follows, the evaluation takes a closer look at indirect measures of child marriage in Table 3.8. This includes a closer investigation of alternative child marriage measures, reports on marriages of same-age friends, as well as changes in engagement rates and practices.

Using an alternative measure, the IALP program has no significant impact on age at marriage (10-17) for program-exposed adolescents and young adults. Now, the marital status and current age as reported in the household roster are again employed, but now also considering length of marriage.[24] With the information on how long household members were married and their age, a variable indicating age at marriage is constructed. Therewith, the first dependent variable is age at marriage (restricted to ages 10-17), reported in columns (1) to (3).[25] Overall, the results indicate no impacts of the program or its separate components on this indicator. Note, that the information was provided by the household head and potential effects may be obscured by measurement errors in the duration of marriage variable. Note also, that the sample size for this estimation is very small (with information from RVs 343 only). These results, however, are not confirmed if we use an alternative outcome measure, e.g., self-reported marriage status.

The program reduces the prevalence for boys having friends of their age (11-17) and gender who are married, as well as for adolescents with high self-reported COVID-19 exposure. Turn to reports by the adolescents themselves: An indicator for whether adolescents aged 11-17 have friends of their age and gender that are married is generated, which adolescents and parents were asked (about their child) directly. In columns (4) to (5), the results indicate a negative impact of 7.0pp in the prevalence for boys having a married friend of their age (11-17) in areas where only the NFBE-program was implemented (significant at the 5% level). This effect is large relative to the average prevalence for males in the control areas, which is reportedly 9.9%. Moreover, while the results indicate no impacts of the program in areas where Child Protection and educational interventions were jointly implemented for the full sample, the heterogeneous analysis

---

[22]While the NFBE interventions happened only in Sindh (and thus the corresponding coefficients drop in the Punjab estimations), the effects of the Child Protection interventions in both provinces can be estimated and analyzed whether they varied by district.

[23]Similarly, Table O.6 reports the same heterogeneous analysis for female individuals only. The results show that, for female individuals only, the decrease in child marriage in the age group 15-17 is driven by very poor households and in the group reporting low exposure to COVID-19.

[24]Responses were provided by a household head or, in case absence, a knowledgeable adult. The sequence of questions - marital status and duration of marriage - was embedded in a set of other general questions about education, school attendance, gender, etc., which made them less evident or suspicious to respondents.

[25]The sample size drops to only few observations in these estimations, considering that only people up to age 23 were considered (i.e., only adolescents and young adults who were potentially exposed to the program).

Table 3.7: Child Marriage (I)

	(I) Adolescents Age 10-19			(II) Adolescents Age 10-17			(III) Adolescents Age 15-17			(IV) Young adults Age 18-23		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
Panel A: All Villages												
only NFBE	0.001	0.002	-0.003	-0.003	0.001	-0.008	-0.017**	-0.008	-0.027*	-0.023	-0.035	0.001
	(0.005)	(0.006)	(0.007)	(0.003)	(0.003)	(0.006)	(0.008)	(0.006)	(0.017)	(0.025)	(0.030)	(0.030)
CP and NFBE	-0.005	-0.002	-0.009	-0.007*	-0.000	-0.014**	-0.019*	0.000	-0.042**	-0.010	-0.031	0.019
	(0.005)	(0.006)	(0.008)	(0.004)	(0.003)	(0.006)	(0.010)	(0.008)	(0.018)	(0.024)	(0.026)	(0.032)
CP, NFBE pool	-0.006	-0.003	-0.011	-0.007**	-0.001	-0.013**	-0.022**	-0.005	-0.039**	-0.028	-0.048*	0.000
	(0.005)	(0.006)	(0.007)	(0.004)	(0.003)	(0.006)	(0.009)	(0.008)	(0.017)	(0.023)	(0.026)	(0.030)
CP, not NFBE pool	-0.001	-0.002	0.000	-0.002	-0.000	-0.004	-0.001	0.002	-0.004	0.003	-0.002	0.009
	(0.004)	(0.004)	(0.006)	(0.002)	(0.002)	(0.003)	(0.005)	(0.004)	(0.009)	(0.016)	(0.015)	(0.022)
Control mean	0.029	0.018	0.041	0.009	0.005	0.014	0.019	0.008	0.033	0.197	0.154	0.246
CP overall*	-0.002	-0.002	-0.002	-0.003	-0.000	-0.006	-0.006	0.000	-0.012	-0.004	-0.012	0.007
CP p-value	0.256	0.514	0.229	0.035	0.702	0.010	0.035	0.680	0.026	0.365	0.095	0.804
IALP overall**	-0.002	-0.002	-0.003	-0.003	-0.000	-0.007	-0.008	-0.000	-0.017	-0.006	-0.016	0.008
IALP p-value	0.399	0.798	0.187	0.049	0.901	0.018	0.023	0.560	0.024	0.366	0.125	0.721
N	19,371	10,065	9,306	15,894	8,178	7,716	5,759	3,033	2,726	7,529	4,292	3,237
RVs	648	648	648	648	648	648	648	635	633	647	639	638
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols
Panel B: NFBE Villages Only												
only NFBE	0.001	0.002	-0.000	-0.003	0.000	-0.007	-0.017**	-0.009	-0.027*	-0.017	-0.032	0.008
	(0.005)	(0.007)	(0.006)	(0.003)	(0.003)	(0.006)	(0.008)	(0.006)	(0.016)	(0.023)	(0.027)	(0.030)
CP and NFBE	-0.005	-0.002	-0.008	-0.007*	-0.000	-0.014**	-0.021**	-0.000	-0.045**	-0.010	-0.033	0.020
	(0.005)	(0.006)	(0.007)	(0.004)	(0.003)	(0.006)	(0.010)	(0.008)	(0.018)	(0.024)	(0.025)	(0.033)
only CP	-0.006	-0.003	-0.010	-0.008**	-0.001	-0.014**	-0.024**	-0.006	-0.042**	-0.024	-0.044*	0.004
	(0.005)	(0.006)	(0.006)	(0.003)	(0.003)	(0.005)	(0.009)	(0.008)	(0.016)	(0.023)	(0.025)	(0.030)
Control mean	0.019	0.013	0.025	0.011	0.004	0.018	0.029	0.010	0.054	0.129	0.113	0.151
$\beta_1 = \beta_3$	0.122	0.392	0.172	0.137	0.634	0.097	0.253	0.611	0.162	0.736	0.568	0.883
$\beta_2 = \beta_1 + \beta_3$	0.974	0.923	0.783	0.381	0.832	0.322	0.041	0.097	0.192	0.251	0.166	0.843
N	8,459	4,469	3,990	6,913	3,579	3,334	2,628	1,399	1,229	3,207	1,931	1,276
RVs	210	210	210	210	210	210	210	208	208	210	208	207
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols
Panel C: All Treatment Arms												
only NFBE	0.003	0.002	0.003	0.001	0.001	-0.001	-0.001	-0.001	-0.002	-0.027	-0.037	-0.007
	(0.004)	(0.005)	(0.007)	(0.003)	(0.003)	(0.005)	(0.007)	(0.004)	(0.013)	(0.020)	(0.023)	(0.026)
only CPMF, not NFBE pool	0.003	0.001	0.005	-0.002	0.000	-0.005	-0.003	0.001	-0.006	0.033*	0.018	0.053**
	(0.005)	(0.005)	(0.007)	(0.003)	(0.002)	(0.004)	(0.007)	(0.006)	(0.011)	(0.020)	(0.019)	(0.026)
only CPMF, NFBE pool	-0.006	-0.005	-0.008	-0.001	0.002	-0.006	0.001	0.011	-0.008	-0.033	-0.045*	-0.018
	(0.006)	(0.006)	(0.009)	(0.004)	(0.003)	(0.005)	(0.010)	(0.008)	(0.014)	(0.026)	(0.024)	(0.041)
CPMF and NFBE	0.015	0.008	0.022	-0.002	0.001	-0.004	-0.002	0.008	-0.015	0.050*	0.013	0.100**
	(0.010)	(0.009)	(0.013)	(0.005)	(0.003)	(0.008)	(0.012)	(0.008)	(0.018)	(0.028)	(0.025)	(0.046)
only CPF, not NFBE pool	-0.007	-0.006	-0.007	-0.003	-0.001	-0.006	-0.005	0.001	-0.010	-0.007	0.000	-0.011
	(0.006)	(0.006)	(0.008)	(0.003)	(0.003)	(0.004)	(0.007)	(0.006)	(0.012)	(0.023)	(0.023)	(0.030)
only CPF, NFBE pool	-0.001	-0.001	0.001	-0.004	-0.002	-0.006	-0.011	-0.006	-0.017	-0.032	-0.053*	-0.001
	(0.009)	(0.009)	(0.010)	(0.003)	(0.003)	(0.005)	(0.008)	(0.007)	(0.011)	(0.031)	(0.028)	(0.047)
CPF and NFBE	-0.012**	-0.006	-0.017**	-0.005	0.001	-0.012***	-0.008	0.007	-0.027**	-0.027	-0.031	-0.023
	(0.006)	(0.007)	(0.007)	(0.003)	(0.005)	(0.004)	(0.008)	(0.011)	(0.013)	(0.028)	(0.025)	(0.042)
only CPM, not NFBE pool	-0.004	-0.000	-0.007	-0.004	-0.001	-0.009**	-0.014*	-0.004	-0.026**	-0.023	-0.027	-0.017
	(0.004)	(0.005)	(0.008)	(0.003)	(0.003)	(0.004)	(0.007)	(0.006)	(0.012)	(0.024)	(0.021)	(0.034)
only CPM, NFBE pool	-0.006	-0.006	-0.008	-0.006	-0.003	-0.009	-0.010	-0.002	-0.016	-0.029	-0.051**	0.000
	(0.006)	(0.007)	(0.009)	(0.004)	(0.003)	(0.007)	(0.011)	(0.009)	(0.017)	(0.025)	(0.025)	(0.039)
CPM and NFBE	-0.011	-0.009*	-0.012	-0.002	-0.001	-0.004	-0.000	0.004	-0.007	-0.069**	-0.083***	-0.051
	(0.007)	(0.005)	(0.011)	(0.003)	(0.004)	(0.004)	(0.009)	(0.008)	(0.018)	(0.028)	(0.022)	(0.046)
Control mean	0.028	0.018	0.040	0.009	0.005	0.014	0.020	0.008	0.035	0.191	0.150	0.241
N	19,371	10,065	9,306	15,894	8,178	7,716	5,759	3,033	2,726	7,529	4,292	3,237
RVs	648	648	648	648	648	648	648	635	633	647	639	638
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

▸ *Notes*: Table 3.7 displays the program effects on the child marriage rate for different adolescent and young adult age groups. The dependent variable, marriage, considers the marital status of household members reported in the household roster. Responses were provided by a household head or, in case absence, a knowledgeable adult. Marginal effects from separate OLS estimations are reported.
▸ Sample: Endline survey. Panel A and C full sample. Panel B only NFBE-eligible villages.
▸ Part I considers adolescents age 10-19 (standard definition of adolescents), Part II restricts the sample to the age group 10-17, Part III considers child marriage prevalence amongst adolescents age 15-17, and Part IV focuses on the more directly program-exposed adolescents and young adults (age 18-23). Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
▸ Panel A estimates equation 2.1 with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.
▸ Note that in Panel C the impacts are estimated relative to the control group in the non-NFBE pool. ▸ Related table(s): Table A.47 and Table O.6.

reveals a negative impact in the prevalence by 8.1pp for adolescents with high self-reported COVID-19 exposure.[26] In other words, the program effects as measured by this outcome are stronger if children were reportedly more affected by COVID-19. The average prevalence of child marriage using the friend's marital status indicator in the control areas is reportedly 13.5% for individuals with high COVID-19 exposure; for less affected individuals the incidence is lower (10.0%).[27]

The IALP program increases engagement rates for boys age 10-17, while their marriage rates do not see a change.

Engagement rates for boys age 10-17 increase due to the program. In Part II of Table 3.8, the focus shifts to engagement prevalence and the dependent variable is an indicator of engagement for adolescents of age 10-17. Changes in engagement practices may present an alternative strategy to marriage at young age, without the more immediate negative impacts of being married due to early child birth. Engagement is measured similar to marital status by asking about engagement status and current age of household members in the household roster. The average engagement prevalence in the control areas is reportedly 1.9% for boys in the full sample; for girls the prevalence is considerably higher (7.1%). The results on the program effects indicate an increase in engagement incidence of boys age 10-17 by 1.7pp in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas (significant at the 5% level). Moreover, in areas where only the education intervention was implemented, the engagement prevalence of boys age 10-17 decreases by 1.3pp (significant at the 5% level).[28] To sum up, while the previous findings indicated no changes in marriage rates for boys for the age group of 10-17, an increase in their engagement rates can be observed instead.

The results on increased engagement rates for boys are confirmed in *list experiments* which were conducted to capture the prevalence of child marriage with indirect questions.[29] One list experiment (indirectly) asked 11-16 year old boys about the likelihood that they would get married "next year". The results reveal that the program increase this reported likelihood. While the evidence on changes in a delay of marriage for boys may capture a postponement in line with the original program expectations, a lack of such evidence for girls is not clear.

The evidence on changes in engagement and delay of marriage for boys indicates a postponement in line with the original program expectations.

The IALP program reduced the dowry amount paid by adolescents' and parents' own families at marriage by 17.4%. Finally, in Part III of the table, potential changes in engagement practices are considered. Here, the dependent variable is the logarithm of the dowry amount in Pakistani Rupees (divided by 1000) paid by adolescents' and parents' own families at marriage. In the estimation the evaluation imposes no restriction on the age of the bride or groom. The average dowry price paid by adolescents' and parents' own families at marriage in the control areas is reportedly 11,118 Pakistani Rupees. The results indicate a reduction in the dowry amount by 17.4% in areas where only the educational intervention was implemented (significant at the 5% level). In contrast to the work by Buchmann et al. (2021) for Bangladesh, the results indicate no increase in dowry prices from being assigned to the Child Protection components of the empowerment program. When analyzing the effect by groups, the findings reveal that the observed reduction of the dowry amount is driven by observations in Sindh (Table A.48), which is in line with the results reported above which indicated that the reduction in child marriages was more likely to occur there as well.[30]

Evidence from Qualitative Interviews *Assessment of Changes in Patterns of Child Marriage by Key Informants* Statements from KIs offer a mixed picture regarding the perceived effect the program might have had on the prevalence of child marriage and teenage pregnancy. Some reported that KCIs were enabled to identify child marriages in their

---

[26]This effect is significant at the 5% level, as seen in Table A.48. Moreover, the difference of -8.1pp between individuals with high and individuals with low COVID-19 exposure is significant at the 5% level.

[27]Moreover, when analyzing the indirect indicators of child marriage for adolescent and parent respondents separately, Table O.7 reveals no change in the prevalence of adolescents having friends their age and gender that are married for females when only the responses of the adolescent girls themselves are considered, while a weakly significant reduction can be observed for females when only the parents' responses are considered (Table O.8). This suggests that the lack of indirect evidence on the effects amongst girls may be related the source of information.

[28]The overall Child Protection effect for boys is an increase of 1.6pp with a p-value of 0.006. Similarly, the overall IALP effect for boys is an increase of 1.6pp with a p-value of 0.016.

[29]See Section A.IV.A (in detail) and II.C (short).

[30]Moreover, the reduction is stronger for individuals living in poor compared to individuals living in very poor households, as well as for individuals with low self-reported COVID-19 exposure compared to individuals with high self-reported COVID-19 exposure. These differences are not significant, however.

Table 3.8: Child Marriage (II)

	(I)						(II)			(III)		
	Child Marriage Rate						Engagement			Engagement Practices		
	age at marriage (10-17)			A (11-17) has friends his/her age & gender that are married			engaged (10-17)			amount of dowry paid by own family		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
only NFBE	-0.145	-0.328	-0.116	-0.041*	-0.070**	-0.018	-0.000	0.013**	-0.018	-0.174**	-0.240*	-0.120
	(0.373)	(0.736)	(0.435)	(0.025)	(0.029)	(0.037)	(0.008)	(0.007)	(0.014)	(0.077)	(0.136)	(0.103)
CP and NFBE	0.177	0.485	0.133	-0.039	-0.037	-0.038	0.015	0.020*	0.009	-0.040	-0.166	0.055
	(0.441)	(0.750)	(0.529)	(0.028)	(0.037)	(0.037)	(0.011)	(0.011)	(0.014)	(0.079)	(0.123)	(0.141)
CP, NFBE pool	-0.256	-0.706	-0.108	0.008	-0.017	0.026	0.007	0.013	0.001	-0.060	-0.044	-0.087
	(0.394)	(0.704)	(0.440)	(0.028)	(0.037)	(0.035)	(0.009)	(0.008)	(0.013)	(0.071)	(0.119)	(0.118)
CP, not NFBE pool	0.092	-0.071	0.234	-0.005	0.016	-0.012	0.003	0.017**	-0.011	0.185	0.350	0.040
	(0.196)	(0.235)	(0.240)	(0.013)	(0.019)	(0.020)	(0.007)	(0.007)	(0.011)	(0.137)	(0.296)	(0.124)
Control mean	15.346	15.578	15.224	0.120	0.099	0.138	0.044	0.019	0.071	11.118	11.000	11.242
SD	1.529	1.465	1.556	0.325	0.298	0.346	0.205	0.136	0.257	1.438	1.783	0.880
CP overall*	0.017	-0.207	0.161	-0.002	0.009	-0.004	0.004	0.016	-0.008	0.132	0.265	0.013
CP p-value	0.715	0.308	0.801	0.935	0.981	0.728	0.375	0.006	0.572	0.421	0.366	0.777
IALP overall**	0.023	-0.132	0.134	-0.010	-0.003	-0.009	0.005	0.016	-0.007	0.086	0.170	0.007
IALP p-value	0.906	0.762	0.914	0.269	0.250	0.652	0.250	0.001	0.574	0.696	0.829	0.722
N	795	270	525	4,824	2,219	2,605	15,894	8,178	7,716	1,539	733	806
RVs	343	179	306	648	636	645	648	648	648	320	244	246
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

▸ *Notes*: Table 3.8 displays the program effects on child marriage-related indicators. Effects from separate OLS estimations are presented.
▸ Sample: Endline survey. Full sample.
▸ The program effects on the following indicators are presented: age at marriage (10-17), adolescent (11-17) has friends his/ her age and gender that are married, engaged (10-17), and amount of dowry paid by own family. "Age at marriage (10-17)" is constructed by using information on marital status, length of marriage, and current age of household members reported in the household roster. "Adolescent (11-17) has friends his/ her age and gender that are married" is generated by asking this as a yes/ no question to adolescents and parents (about their child) directly. "Engaged (10-17)" is constructed by using information on marital status and current age of household members reported in the household roster. "Amount of dowry paid by own family" is generated by asking adolescents and parents to indicate in Pakistani Rupees how much their own family paid as dowry. Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
▸ The equation 2.1 is estimated with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$. N refers to the number of respondents.
▸ Related table(s): Tables A.48, O.7 and O.8.

communities and were actively seeking out parents who intended to marry their children, in order to discourage them from doing so. Few parents and adolescents were also reportedly advocating in their communities for a stop or delay in the marriage of underage boys and girls. Another interview partner disclosed that child marriages had been prevented due to attendees calling the IP who then informed the police. Others were more cautious about the impact the program had on child marriage, stating that a change in knowledge and, possibly, attitude regarding this topic had taken place, but that this has not necessarily translated into a change in practices. Further quantitative analysis of impacts on social norms will confirm this change in attitude below (Section II.B). Another interview partner perceived that child marriages are still common in rural villages and suggested further action is required. One KI mentioned that a potential reduction in teenage pregnancies would be difficult to measure.

> "In Punjab, many parents reported that they stopped child marriages and they were advocating to the communities. Some of the key community influencers going house to house, having heard that any parents are planning to do child marriage (...) So, I think there is a change, at least the parents and communities are aware of the consequences of child marriage now. Though we cannot say the child marriages are completely stopped now (...), but there is improvement." (Key Informant No.1)
>
> "In my opinion, people in the community got a strong message regarding early marriages and education." (Key Informant No.12)
>
> "Early marriages are still in trend in villages, hence there is still a need for further working, process so that the numbers may go down." (Key Informant No.9)

II.A.2   Child Pregnancy

> **Key Findings on Pregnancy Outcomes**
>
> Another strategic impact to be achieved by the program was postulated as a reduction in adolescent pregnancies by 3%. The evaluation was set up to test the hypothesis in an attempt to capture to which extent this goal could be achieved. The estimation results for child pregnancy rates, however, have to be interpreted carefully since the estimates are based on a very small sample.
>
> • The results indicate no significant impacts of the program or its separate components on child pregnancy rates amongst female adolescents.

Table 3.9 displays effects from a set of OLS regression results using endline survey information on child pregnancy. The dependent variable is an indicator for child pregnancy.[31] With the information from the household roster on the adolescent's age and the age of the oldest child, a variable indicating whether a female household member had her first child at certain age is constructed and only members who are now adolescents or young adults (age 11-23), and therefore could have been exposed to the program, are considered. The presentation of the rules is aligned to the one on child marriage, naturally focusing only on females (Table 3.7). Note, however, the estimation results for pregnancy have to be interpreted very carefully since the estimates are based on a small sample, as only married girls in the respective age range are included, in total 467 girls. In those villages where both interventions were implemented, that is, the treatment arm with Child Protection and NFBE, the sample includes only 42 girls in the age range 11 to 23. The average prevalence of having had the first child at age 10-19 (among girls who are now age 11-23) is reportedly 6.9% in the control areas. The results indicate no impacts of the program or its separate components on child pregnancy prevalence for any of the age groups. While Panel A considers the effects for the full sample, Panel B takes a closer look at the

---

[31]To capture the child pregnancy incidence, information from the household roster was considered: As mentioned under Section II.A, responses about the age, children, and age of the oldest child for all household members (conditional on them being a female and being married, widowed, or divorced) were provided by a household head or, in case absence, a knowledgeable adult. Note that child pregnancy is defined in the PAP as an "indicator for having had first child while being an adolescent (19 or younger), considered are only individuals who were adolescents during the time of the program". Although adolescents now age 11 will most likely not have been exposed to the program, the earliest observation for child pregnancy in our sample is for girls age 14, i.e., girls that were targeted by the program. An alternative definition of child pregnancy is initially tested which considers adolescents and young adults in certain age groups that have had a child, without the restriction of having had the first child at certain ages (e.g., "has child (10-19)" indicates adolescents currently age 10-19 that have at least one child). However, employing this definition of child pregnancy results in an extremely small sample size of only 80 girls age 15-17 and 10-17, and therefore the definition is adapted as described in the main text above.

NFBE-eligible sample (212 villages). Here the sample size drops even further and the results again indicate no impacts of the program or its separate components on child pregnancy prevalence for any of the age groups.[32]

II.A.3   Education

Key Findings on Educational Outcomes

One core strategic impact to be achieved by the IALP program is an improvement in access to education. Hereby the focus of the program shifted from the original plan to primarily focus on secondary education enrollment (with an expected increase by 4%) to a focus to improve access to alternative basic education as a primary goal (with additional access for 3,000 adolescents; see Figure O.4). The primary question of interest is thus whether an increase in NFBE participation is captured and, given that adolescents are better informed about their rights, incl. the right to access education, also whether an increase in the access to formal education is observed.

- The IALP program has a sizable positive impact on having ever attended NFBE;

- The program has a negative impact on girls' (past) primary school attendance, in particular for those with low self-reported COVID-19 exposure. Similarly, the program has a negative impact on boys' (past) secondary school attendance. The negative effects observed for primary and secondary school attendance seem to be mostly prevalent in the CPM intervention areas. They may also indicate a crowding out effect from formal education to non-formal education;

- At the same time, children residing in areas where the two program components (NFBE and Child Protection) were jointly implemented are less likely to miss school and more likely to want to continue with secondary education, with effects being particularly strong for girls;

- The program has no impact on the overall index capturing negative education-related attitudes, while agreement with the statement that educating boys is more important than girls sees an increase in Sindh. Moreover, the program increases the preference that girls and boys should be educated separately;

- The evaluation records an improvement in the ability to write and read in Sindhi due to the IALP program activities, especially among males. No improvement in Mathematics outcomes is recorded;

- Positive impacts in learning outcomes are especially visible in areas where both the Child Protection and NFBE treatments were implemented as compared to the control group.

This section now turns to the last strategic goal of the program, namely the promotion of education, whereby the focus is on secondary education and an increase in access to alternative (non-formal) basic education. The NFBE component of the IALP program was expected to directly affect enrollment rates in non-formal educational institutions and, eventually, in formal education. The focus of NFBE program is to target over-age out-of-school children for completing primary certification. Therefore, the main focus of the NFBE activities was the primary education, while the secondary education was an indirect logical outcome of the program following the theory of change.[33] Moreover, community members and stakeholders were expected change their attitudes towards the need for more education, in particular for girls, through the Child Protection activities.

Household Roster Information on Education   To capture educational attendance, information from the household roster is considered.[34] Table 3.10 displays the effects from a set of OLS regressions, whereby the dependent variables of interest are indicators for adolescents and young adults (age 11-23) having ever attended certain types of educational institutions.

The IALP program has a sizable positive impact on (past) NFBE attendance. Part I of the table reports program effects on adolescents and young adults (11-23) having ever attended NFBE. The sample is hereby restricted to 210 NFBE-eligible villages only. The results indicate an increase in this prevalence in the full sample by 1.7pp in areas where the Child Protection and educational interventions were jointly implemented, significant at the 5% level. Moreover, column (3) reveals that the effect is driven by changes for females. Here, the increase is 2.3pp compared to areas without program interventions. An effect of a similar size for males is observed in areas where only the NFBE component was implemented (2.0pp increase in NFBE attendance, significant at the 5% level). This is a sizable impact given a mean

---

[32]The heterogeneous analysis also indicates no impacts of the program on child pregnancy.

[33]Note, the original focus of the program shifted from the original goal of increasing secondary education by 4% (until 2018) to providing the services that 3,000 additional adolescents can access alternative basic education (in 2019) (compare UNICEF South Asia (2019a) to UNICEF South Asia (2020).

[34]As mentioned under Section II.A Outcomes, the household head or, in case absence, a knowledgeable adult was asked about the type of educational institutions all household members had ever attended.

Table 3.9: Child Pregnancy

	(I) had first child 10-19 (now 11-23) (1)	(II) had first child 10-17 (now 11-23) (2)	(III) had first child 15-17 (now 15-23) (3)	(IV) had first child 18-23 (now 18-23) (4)
Panel A: All Villages				
only NFBE	0.027	0.057	0.090	-0.107
	(0.127)	(0.134)	(0.114)	(0.139)
CP and NFBE	0.069	0.163	0.160	-0.181
	(0.136)	(0.136)	(0.155)	(0.152)
CP, NFBE pool	0.136	0.222	0.209	-0.243
	(0.123)	(0.141)	(0.153)	(0.155)
CP, not NFBE pool	-0.017	-0.001	-0.015	0.005
	(0.061)	(0.053)	(0.055)	(0.055)
Control mean	0.691	0.327	0.280	0.706
CP overall*	0.016	0.047	0.033	-0.048
CP p-value	0.397	0.137	0.221	0.140
IALP overall**	0.023	0.062	0.054	-0.069
IALP p-value	0.552	0.247	0.250	0.203
N	467	467	466	450
RVs	305	305	304	298
Model	ols	ols	ols	ols
Panel B: NFBE Villages Only				
only NFBE	0.089	0.047	0.087	-0.076
	(0.130)	(0.148)	(0.118)	(0.158)
CP and NFBE	0.130	0.170	0.186	-0.189
	(0.128)	(0.146)	(0.154)	(0.159)
only CP	0.207	0.194	0.182	-0.213
	(0.133)	(0.155)	(0.151)	(0.162)
Control mean	0.700	0.350	0.300	0.765
$\beta_1 = \beta_3$	0.118	0.096	0.339	0.171
$\beta_2 = \beta_1 + \beta_3$	0.313	0.700	0.616	0.609
N	142	142	141	130
RVs	86	86	85	82
Model	ols	ols	ols	ols

▸ *Notes*: Table 3.9 displays the program effects on the child pregnancy rate for different adolescents' age groups. The dependent variable, pregnancy, considers female household members that are married, divorced or widowed and have a child, and the age of their oldest child as reported in the household roster. Responses were provided by a household head or, in case absence, a knowledgeable adult. Effects from separate OLS estimations are reported.
▸ Sample: Endline survey. Panel A full sample. Panel B only NFBE-eligible villages.
▸ Part I considers adolescents and young adults of age 11-23 having had their first child 10-19. Part II considers adolescents and young adults of age 11-23 having had their first child 10-17. Part III restricts the sample to adolescents and young adults age 15-23 having had their first child 15-17. Part IV considers the more directly program-exposed adolescents and young adults age 18-23 having had their first child age 18-23. Columns (1), (2), (3), and (4) display the sample for girls only.
▸ Panel A estimates equation 2.1 with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. N refers to the number of respondents.

of 0.2% in control areas for the full sample (column (1)).[35] The analysis by groups reveals that the positive effect is driven by observations in Sindh, which makes sense given that this program component was implemented only in Sindh (Table A.49). Moreover, the positive program effect is larger for individuals in poor than for individuals in very poor households (significant at the 5% level).[36]

The program has a negative impact on girls' (past) primary attendance, in particular for those with low self-reported COVID-19 exposure. Part II of the table reports program effects on adolescents and young adults (11-23) having ever attended primary school. Both in areas where only the educational intervention was implemented and where the Child Protection and educational interventions were jointly implemented, the results indicate a reduction in (past) primary school attendance for girls (by 8.2pp and 9.2pp respectively, significant at the 5% level). The average primary school attendance is 59.9% in control areas for girls (column (6)).[37] Splitting the sample into groups, the results indicate a negative impact to be particularly strong for adolescents and young adults with low self-reported COVID-19 exposure.[38] In other words, the negative program effects on (past) primary school attendance of girls (age 11-23) are stronger if they were reportedly less affected by COVID-19.

Similarly, the program has a negative impact on boys' (past) secondary school attendance. Columns (7) to (9) take a closer look at program effects on adolescents and young adults (11-23) having ever attended secondary school, as reported by the household head. The results indicate a decrease in the prevalence in areas where Child Protection interventions were implemented in the NFBE-eligible areas but only for boys (7.4pp, significant at the 5% level, in column (8)). The overall Child Protection effect for boys is a reduction of 2.7%, with a p-value of 0.036. The average prevalence of secondary attendance of boys is reportedly 40.6% in control areas.[39]

Taken together, the results indicate no impact of the program on the prevalence of having ever attended any - formal or informal - type of education. Finally, columns (10) to (12) take a closer look at program effects on adolescents and young adults (11-23) having ever attended any type of education (including primary or secondary school, NFBE Center, college, university, or madrassah). Overall, the results indicate no impacts of the program (or its separate activities) on this indicator. An average overall effect, however, seems to reflect a zero-sum as the evaluation captures negative effects for formal educational attendance and positive impacts for non-formal educational attendance. In NFBE-eligible areas this may be a sign of crowding-out effects.

While Panel A considers the effects for the full sample, Panel B again takes a closer look at the NFBE-eligible sample (212 villages), the area in which the educational component was implemented. The evaluation reveals that all program impacts hold when restricting to this sample, and are in many cases slightly higher in size. Moreover, one can constitute that formal educational attendance is lower in this subsample of villages, which is not surprising given that the sample was selected due to its worse educational outcomes. The results also indicate that in this subsample of villages, a negative and slightly stronger impact on girls' (past) primary attendance can be observed as compared to the full sample. Both in areas where only the educational intervention was implemented, and where the Child Protection and educational interventions were jointly implemented, the results indicate a reduction in (past) primary school attendance for girls (by 8.7pp and 9.3pp respectively, significant at the 5% and 1% level).

The negative effect observed for secondary school attendance seems to be driven most strongly by the CPM intervention. In Panel C, all possible combinations of the three different Child Protection treatment arms (program activities) are considered for the full sample of villages. The impacts are estimated relative to the control group in the non-NFBE-eligible pool. The negative effect for girls' (11-23) past primary school attendance, in areas where Child Protection interventions were implemented in the NFBE-eligible areas, seems to be driven by both the CPM and CPF interventions. For instance, the CPM intervention lead to a negative impact of 16.1pp compared to the control group (significant at the 5% level). The negative impact for boys' secondary attendance, again in areas where Child Protection interventions were

---

[35]The overall IALP effect is an increase of 0.5pp, with a p-value of 0.025.

[36]Both of these differences are not significant, however. Note that since the NFBE component was only implemented in Sindh, the coefficients in Punjab for the first three treatment modalities (only NFBE, CP and NFBE, and CP, NFBE pool) are always equal to zero. When estimating the difference between the two provinces, these null estimates are considered as missing, therefore rendering the difference insignificant.

[37]The overall IALP effect is a reduction of 2.9pp for girls, with a p-value of 0.036.

[38]The negative impact for those with low COVID-19 exposure is strongest in areas where the Child Protection and educational interventions were jointly implemented (-12.0pp and significant at the 1% level, as seen in Table A.49). Moreover, the difference of 10.5pp between individuals with high and individuals with low COVID-19 exposure is significant at the 1% level.

[39]When looking at the effect by groups, the analysis reveals in Table A.49 that the negative impact is higher for adolescents with high self-reported COVID-19 exposure (significant at the 5% level). The difference between individuals with high and low COVID-19 exposure is, however, not significant.

implemented in the NFBE-eligible areas, seems to be driven by both the CPM and CPMF interventions.[40] For attendance of any educational institution, Panel C reveals negative impacts for the full sample and both genders, again in areas where the CPM treatment arm was implemented in NFBE-eligible areas.[41]

Education: Attendance, Attitudes, and Gender Equity   The information presented in Table 3.10 was captured via a household roster. In the next paragraphs the evaluation will additionally consider self-reported information as shared by the adolescent respondents in the main interviews. Although the sample size is smaller, the information suffers from less measurement error (since coming directly from the adolescent itself). Table 3.11 displays the effects from a set of OLS regression results analyzing endline survey information on further education-related variables.

> The negative effects observed for primary and secondary school attendance may indicate a crowding out effect from formal education to non-formal education.

The IALP program decreases the number of missed school days for girls by 0.9 days.   Part I of the table reports program effects on the number of times adolescents missed school last week. Adolescents [and parents] were asked to indicate the number of days they [or their child] missed at the main educational institution that they were currently attending. The results indicate a decrease in the number of missed school days for girls by 0.9 days in areas where Child Protection and educational interventions were jointly implemented (significant at the 1% level). The average number of missed school days in the control areas is reportedly 2.7 for girls.

Construction of Summary or Index Variables

To capture broad concepts and more efficiently summarize information, the report presents summarized information of several responses on a similar topic, so-called "indices". In what follows, the general procedure of generating these indices is outlined shortly. Summary variables (or indices) are constructed by first generating score variables which summarize the responses to several individual questions, and, in a second step, by rescaling this score to range between 0 and 1. Rescaling is achieved by dividing the achieved score by the number of total subcomponents included in the summary variable.

To make this procedure clearer, please consider the following example. Imagine respondents were asked to answer a set of ten questions with yes or no. In order to aggregate the responses to the ten questions into an index, the following steps are followed:

1. Individual indicator variables (responses) for each question are generated, coded as 1 if the respondent indicated an improvement, and 0 otherwise. Thereby all responses first need then to point to the same direction (for instance that a 1 always indicates an improvement);

2. A score variable is generated which indicates the total number of questions the respondent answered with yes (max. ten);

3. The score variable is rescaled to range between 0 and 1, attained by dividing the score by the number of total number of questions included in the summary variable (i.e, by ten);

The resulting summary variable (or index) indices, out of the ten questions, the share of questions that were answered with yes on a scale of 0 to 1. If, for example, a respondent answered four questions with yes, then their corresponding summary variable would lie at 0.4, in other words the respondent answered 40% of the questions with yes.

This procedure is similar to how Bandiera et al. (2019) construct their indices. For example, for their gender empowerment index, they first construct a score indicating the number of questions (out of eight) to which respondents answered that both genders should be equally responsible. Therefore, higher values represent more egalitarian gender norms. They then rescale the score to range between 0 and 100. Their gender empowerment index, thus, indicates the share of questions (out of eight) to which the answer was both genders should be responsible. Note that other approaches to index construction are possible, such as adding weights as done by Dhar et al. (forth.).

In areas where both program components were implemented, the results indicate a positive program effect for females on the prevalence of adolescents wanting to continue secondary education.   Part II and III of the table considers attitudes towards education, including gender norms. To capture education-related attitudes, adolescents and parents were asked whether they or their adolescent child plans to continue their education. The results indicate an increase in the prevalence for females by 11.4pp in areas where the Child Protection and educational interventions were jointly implemented, significant at the 1% level (column (6)). The average prevalence of adolescents wanting to continue secondary education is reportedly 24.8% for girls in control areas. The heterogeneous analysis reveals a positive effect in areas where only the NFBE-program was implemented for adolescents with low self-reported COVID-19 exposure

---

[40]The CPM intervention seems to have the strongest negative impact with -14.5pp, significant at the 1% level.

[41]Here, the CPM intervention leads to a negative impact of 10.8pp compared to the control group (significant at the 1% level).

Table 3.10: Education (I)

	(I) A/ young adult (11-23) ever attended NFBE			(II) A/ young adult (11-23) ever attended primary school			(III) A/ young adult (11-23) ever attended secondary school			(IV) A/ young adult (11-23) ever attended any type of education		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
Panel A: All Villages												
only NFBE	0.017*	0.020**	0.012	-0.027	0.011	-0.082**	-0.044	-0.049	-0.041	-0.014	0.001	-0.040
	(0.009)	(0.010)	(0.010)	(0.034)	(0.038)	(0.040)	(0.037)	(0.039)	(0.050)	(0.030)	(0.029)	(0.043)
CP and NFBE	0.017**	0.012	0.023**	-0.030	-0.022	-0.059	-0.047	-0.039	-0.055	-0.017	-0.013	-0.032
	(0.008)	(0.008)	(0.011)	(0.030)	(0.029)	(0.044)	(0.034)	(0.042)	(0.039)	(0.031)	(0.029)	(0.046)
CP, NFBE pool	0.006	0.010	-0.000	-0.049	-0.026	-0.092**	-0.060*	-0.074**	-0.039	-0.038	-0.030	-0.059
	(0.008)	(0.008)	(0.009)	(0.033)	(0.033)	(0.043)	(0.031)	(0.037)	(0.038)	(0.029)	(0.030)	(0.042)
CP, not NFBE pool	0.000	0.000	0.000	-0.009	-0.021	0.002	-0.010	-0.014	-0.004	-0.004	-0.018	0.009
	(.)	(.)	(.)	(0.020)	(0.019)	(0.026)	(0.018)	(0.019)	(0.024)	(0.019)	(0.017)	(0.027)
Control mean	0.002	0.001	0.002	0.651	0.697	0.599	0.345	0.406	0.276	0.732	0.801	0.654
CP overall*	0.001	0.002	-0.000	-0.017	-0.022	-0.018	-0.021	-0.027	-0.012	-0.012	-0.021	-0.006
CP p-value	0.441	0.217	0.995	0.134	0.201	0.072	0.049	0.036	0.324	0.223	0.156	0.309
IALP overall**	0.005	0.005	0.004	-0.020	-0.019	-0.029	-0.026	-0.030	-0.019	-0.012	-0.018	-0.012
IALP p-value	0.025	0.017	0.105	0.186	0.509	0.036	0.084	0.107	0.224	0.373	0.460	0.290
N	9,395	5,152	4,243	21,602	11,528	10,074	21,602	11,528	10,074	21,602	11,528	10,074
RVs	210	210	210	648	648	648	648	648	648	648	648	648
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols
Panel B: NFBE Villages Only												
only NFBE	0.017*	0.020**	0.012	-0.033	0.006	-0.087**	-0.042	-0.047	-0.036	-0.014	-0.000	-0.034
	(0.009)	(0.010)	(0.010)	(0.032)	(0.036)	(0.038)	(0.036)	(0.038)	(0.048)	(0.027)	(0.027)	(0.040)
CP and NFBE	0.017**	0.012	0.023**	-0.033	-0.031	-0.055	-0.040	-0.034	-0.045	-0.013	-0.015	-0.018
	(0.008)	(0.008)	(0.011)	(0.028)	(0.026)	(0.045)	(0.032)	(0.041)	(0.038)	(0.029)	(0.028)	(0.045)
only CP	0.006	0.010	-0.000	-0.054*	-0.033	-0.093**	-0.055*	-0.069*	-0.033	-0.035	-0.030	-0.047
	(0.008)	(0.008)	(0.009)	(0.032)	(0.030)	(0.042)	(0.030)	(0.037)	(0.035)	(0.026)	(0.028)	(0.039)
Control mean	0.009	0.007	0.011	0.554	0.590	0.512	0.342	0.444	0.220	0.667	0.766	0.548
$\beta_1 = \beta_3$	0.228	0.413	0.187	0.472	0.208	0.891	0.634	0.440	0.935	0.343	0.139	0.737
$\beta_2 = \beta_1 + \beta_3$	0.657	0.157	0.560	0.188	0.940	0.015	0.189	0.081	0.667	0.315	0.666	0.236
N	9,395	5,152	4,243	9,395	5,152	4,243	9,395	5,152	4,243	9,395	5,152	4,243
RVs	210	210	210	210	210	210	210	210	210	210	210	210
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols
Panel C: All Treatment Arms												
only NFBE	0.018*	0.021**	0.013	-0.036	-0.015	-0.068**	-0.025	-0.043	-0.005	-0.025	-0.022	-0.035
	(0.009)	(0.010)	(0.010)	(0.023)	(0.028)	(0.029)	(0.027)	(0.028)	(0.040)	(0.021)	(0.020)	(0.033)
only CPMF, not NFBE pool	0.000	0.000	0.000	-0.030	-0.039*	-0.026	-0.041**	-0.041*	-0.041	-0.036*	-0.043**	-0.031
	(.)	(.)	(.)	(0.022)	(0.022)	(0.028)	(0.019)	(0.021)	(0.025)	(0.020)	(0.020)	(0.026)
only CPMF, NFBE pool	0.007	0.010	0.001	-0.001	-0.002	-0.010	-0.038	-0.063**	-0.006	-0.012	-0.033	0.004
	(0.007)	(0.011)	(0.009)	(0.026)	(0.029)	(0.039)	(0.027)	(0.031)	(0.034)	(0.024)	(0.025)	(0.040)
CPMF and NFBE	0.012	0.017	0.004	-0.029	-0.031	-0.045	-0.014	-0.020	-0.012	-0.031	-0.028	-0.047
	(0.013)	(0.014)	(0.014)	(0.038)	(0.027)	(0.062)	(0.035)	(0.043)	(0.038)	(0.035)	(0.028)	(0.058)
only CPF, not NFBE pool	0.000	0.000	0.000	0.020	0.011	0.027	-0.013	-0.010	-0.013	0.025	0.007	0.040
	(.)	(.)	(.)	(0.024)	(0.024)	(0.031)	(0.025)	(0.027)	(0.030)	(0.023)	(0.022)	(0.031)
only CPF, NFBE pool	-0.001	0.002	-0.006	-0.074**	-0.063*	-0.101**	-0.013	-0.030	0.008	-0.051*	-0.051	-0.062
	(0.007)	(0.006)	(0.011)	(0.035)	(0.034)	(0.046)	(0.025)	(0.033)	(0.033)	(0.028)	(0.031)	(0.040)
CPF and NFBE	0.021	-0.001	0.043*	-0.040	-0.033	-0.063	-0.014	-0.006	-0.017	-0.019	-0.021	-0.025
	(0.014)	(0.006)	(0.025)	(0.029)	(0.032)	(0.042)	(0.038)	(0.045)	(0.041)	(0.037)	(0.032)	(0.053)
only CPM, not NFBE pool	0.000	0.000	0.000	0.000	-0.001	-0.003	0.011	0.010	0.013	0.019	0.016	0.021
	(.)	(.)	(.)	(0.022)	(0.023)	(0.030)	(0.025)	(0.029)	(0.032)	(0.023)	(0.022)	(0.032)
only CPM, NFBE pool	0.018	0.025	0.009	-0.130***	-0.117**	-0.161**	-0.089**	-0.145***	-0.015	-0.108***	-0.090***	-0.140**
	(0.021)	(0.022)	(0.019)	(0.048)	(0.050)	(0.068)	(0.044)	(0.046)	(0.056)	(0.033)	(0.031)	(0.055)
CPM and NFBE	0.019	0.019	0.020	-0.044	-0.073**	-0.024	-0.053*	-0.068**	-0.029	-0.028	-0.055*	-0.002
	(0.013)	(0.015)	(0.016)	(0.029)	(0.033)	(0.037)	(0.032)	(0.034)	(0.045)	(0.027)	(0.030)	(0.043)
Control mean	0.009	0.007	0.011	0.642	0.686	0.592	0.345	0.410	0.271	0.726	0.798	0.646
N	9,395	5,152	4,243	21,602	11,528	10,074	21,602	11,528	10,074	21,602	11,528	10,074
RVs	210	210	210	648	648	648	648	648	648	648	648	648
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

▸ *Notes*: Table 3.10 displays the program effects on educational attainment of adolescents and young adults (11-23) at different educational institutions. The dependent variables are indicators for adolescents and young adults (age 11-23) having ever attended certain types of educational institutions as reported in the household roster. Responses were provided by a household head or, in case absence, a knowledgeable adult. OLS estimates are reported.
▸ Sample: Endline survey. Panel A and C full sample (exception: Part I reporting program effects on adolescents and young adults (11-23) having ever attended NFBE restricted to NFBE-eligible villages). Panel B only NFBE-eligible villages.
▸ Part I considers adolescents and young adults (11-23) having ever attended NFBE, Part II considers adolescents and young adults (11-23) having ever attended primary school, Part III considers past educational attendance of secondary school, and Part IV focuses on adolescents and young adults (11-23) having ever attended any type of education. Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
▸ Panel A estimates equation 2.1 with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.
▸ Related table(s): Table A.49.

(significant at the 1% level, Table A.50).[42]   In other words, the positive program effects on adolescents wanting to continue secondary seduction are stronger if they were reportedly less affected by COVID-19.

The program had no impact on the index capturing negative education-related attitudes.   To capture gender and education-related attitudes (Part III), respondents were asked to answer nine specific questions with yes or no based on their personal opinion. These included, for example, "Should boys should be punished more severely than girls?", "Should girls and boys always be educated separately?", "Do girls and boys have different learning abilities?", etc.[43] The summary variable indices, out of the nine questions, the share of questions that were answered with yes, i.e., to which the respondent presents negative education-related attitudes. Overall, the results indicate no impacts of the program or its separate components on this index. On average, respondents in control areas answer 39.7% of the questions with yes, i.e., present negative education-related attitudes at almost 4 out of 9 questions.

The program increased agreement with the statement that educating boys is more important than girls. An example of one of the items, which is part of the index for education-related attitudes, can be found in columns (10) to (12) (Part III). Here, program effects on agreement with the statement that educating a boy is more important than educating a girl is considered. The results indicate an increase in the prevalence of agreement by 6.0pp for the full sample in areas where Child Protection interventions were implemented in the NFBE-eligible areas (significant at the 5% level, column (10)). The average prevalence in the control areas is reportedly 39.1% for the full sample. Further results in indicate that the positive effect is driven by observations in Sindh (Table A.50).[44]

The program increases agreement with the statement that girls and boys should be educated separately, which could be due to the gender-separating nature of the NFBE Centers. Additionally, the full "index for negative education-related attitudes" is graphically presented, displaying responses to the nine statements. The results indicate in Figure 3.4a that the effect of the program on most indicators of agreement with the statements about gender and education are statistically insignificant. However, the IALP program seems to have a positive and significant effect on agreement with the statement that girls and boys should be educated separately. While the average in the control group is about 75%, the effect caused by the program is an increase by 3.1pp in responses to this question (significant at the 1% level).[45]

---

[42]Moreover, the difference between individuals with high and individuals with low COVID-19 exposure is significant at the 5% level. The average prevalence in the control areas is reportedly 27.3% for individuals with low COVID-19 exposure; for less affected individuals the incidence is slightly lower (26.7%).

[43]The individual variables are first converted to binary variables for each negatively phrased question, which are coded as 1 if the respondent answered yes, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions the respondent answered yes to, i.e., the number of questions to which the respondent presents negative education-related attitudes. In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of questions included in the summary variable (i.e, by nine).

[44]Moreover, the analysis reveals that the positive effect is larger for individuals in poor than for individuals in very poor households. The difference between very poor and poor households is not significant, however. The average prevalence of agreement in the control areas is reportedly 34.7% for individuals in poor households and therewith lower than in very poor areas where it is 42.0%.

[45]While NFBE Centers are gender-segregated, so are most formal government schools in Sindh according to national statistics (National Education Management Information System (NEMIS) and Academy of Educational Planning & Management (AEPAM) 2018). Note that NFBE Centers were not necessarily gender segregated. However, in order to be culturally sensitive, this depended on the demand of the communities and parents.

Table 3.11: Education (II)

	(I) Attendance			(II) Attitudes towards Education			(III) Gender & Education					
	# of times A missed school last week			A wants to continue secondary education			index for negative education-related attitudes			agrees (agr.): educating boys is more important (imp.) than girls		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
only NFBE	-0.200	-0.210	-0.303	0.023	-0.040	0.055	0.016	0.005	0.026	0.039	0.029	0.045
	(0.278)	(0.350)	(0.323)	(0.038)	(0.073)	(0.037)	(0.022)	(0.022)	(0.036)	(0.035)	(0.034)	(0.048)
CP and NFBE	-0.392	-0.126	-0.823***	0.068*	0.003	0.114***	0.014	0.002	0.026	0.026	0.046	0.007
	(0.243)	(0.272)	(0.285)	(0.039)	(0.073)	(0.041)	(0.020)	(0.023)	(0.030)	(0.029)	(0.033)	(0.043)
CP, NFBE pool	-0.178	-0.073	-0.418	0.023	-0.002	0.038	0.034*	0.024	0.048	0.060**	0.048*	0.068*
	(0.243)	(0.293)	(0.297)	(0.036)	(0.061)	(0.039)	(0.019)	(0.021)	(0.029)	(0.028)	(0.027)	(0.038)
CP, not NFBE pool	0.199	0.220	0.177	-0.008	0.027	-0.034	-0.011	-0.022	-0.003	-0.012	-0.015	-0.009
	(0.191)	(0.231)	(0.191)	(0.022)	(0.034)	(0.028)	(0.012)	(0.016)	(0.013)	(0.020)	(0.025)	(0.023)
Control mean	2.607	2.503	2.706	0.270	0.299	0.248	0.397	0.422	0.374	0.391	0.420	0.365
SD	2.759	2.707	2.815	0.443	0.459	0.431	0.273	0.276	0.269	0.486	0.491	0.481
CP overall*	0.118	0.157	0.049	-0.001	0.021	-0.019	-0.001	-0.012	0.008	0.004	-0.001	0.008
CP p-value	0.947	0.692	0.498	0.726	0.724	0.928	0.330	0.957	0.177	0.170	0.367	0.197
IALP overall**	0.029	0.092	-0.088	0.009	0.014	0.004	0.002	-0.009	0.012	0.009	0.007	0.011
IALP p-value	0.401	0.817	0.085	0.318	0.947	0.109	0.344	0.887	0.281	0.165	0.182	0.327
N	3,500	1,802	1,698	3,231	1,366	1,865	7,264	3,523	3,741	7,138	3,466	3,672
RVs	641	617	584	640	579	619	648	648	648	648	648	648
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

▸ *Notes*: Table 3.11 displays the program effects on several education-related indicators. Effects from separate OLS estimations are reported.
▸ Sample: Endline survey. Full sample.
▸ The program effects on the following indicators are presented: number of times adolescent missed school last week, adolescent wants to continue secondary education, index for negative education-related attitudes, and agrees: educating boys is more important than girls. "Number of times adolescent missed school last week" is generated by asking adolescents and parents to indicate the number of days they (or their child) missed at the main educational institution that they were currently attending. "Adolescent wants to continue secondary education" is captured by asking adolescents and parents whether they or their adolescent child plans to continue their education. "Index for negative education-related attitudes" is captured by asking respondents to answer nine specific questions with yes or no based on their personal opinion. The individual variables are first converted to binary variables for each negatively phrased question, which are coded as 1 if the respondent answered yes, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions the respondent answered yes to, i.e., the number of questions to which the respondent presented negative education-related attitudes. In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of questions included in the summary variable (i.e, by nine). The summary variable indices, out of the nine questions, the share of questions that are answered with yes, i.e., to which the respondent presents negative education-related attitudes. "Agrees: educating boys is more important than girls" is an example of one of the items part of the education-related attitudes scale. Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
▸ The equation 2.1 is estimated with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. N refers to the number of respondents.
▸ Related table(s): Table A.50.

Figure 3.4: Educational Attitudes
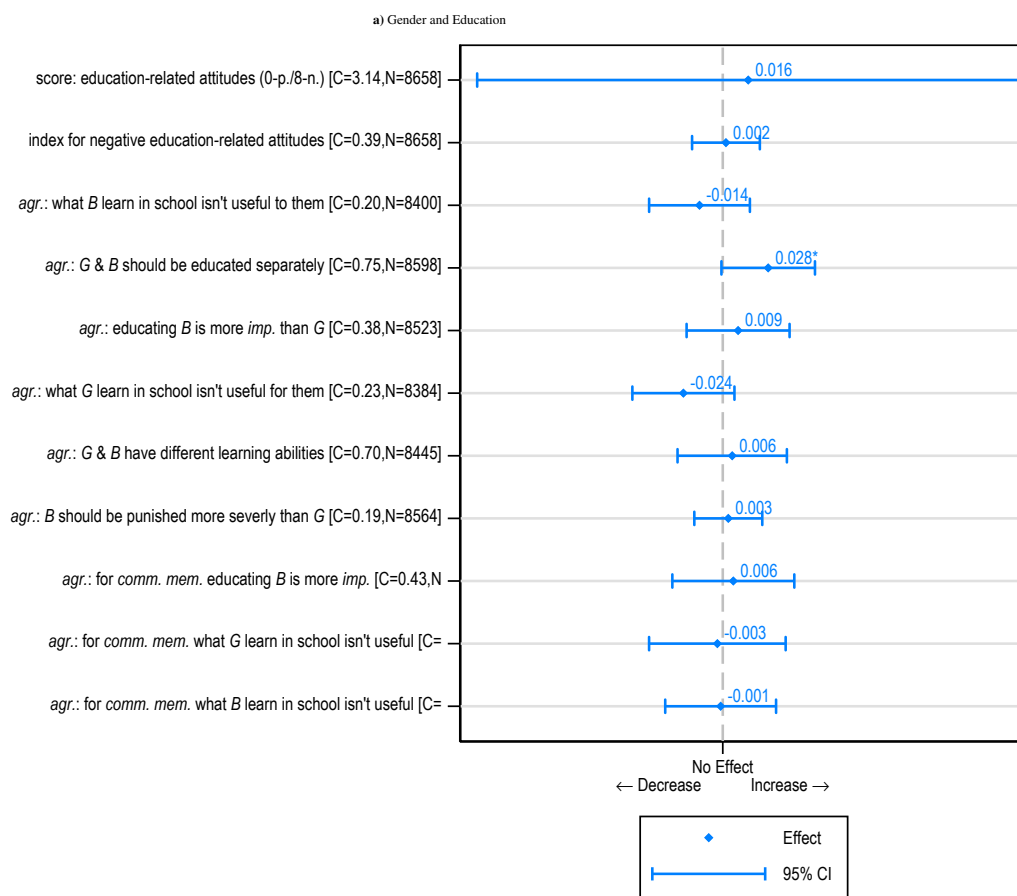
▶ *Notes*: This figure shows treatment effects capturing changes in educational attitudes. To capture gender and education-related attitudes, respondents are asked to answer nine specific questions with yes or no based on their personal opinion. The individual variables are first converted to binary variables for each negatively phrased question, which are coded as 1 if the respondent answered yes, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions the respondent answered yes to, i.e., the number of questions to which the respondent presented negative education-related attitudes. In a final step, a summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of questions included in the summary variable (i.e, by nine). The summary variable "index for negative education-related attitudes" thus indices, out of the nine questions, the share of questions that are answered with yes, i.e., to which the respondent presents negative education-related attitudes.

▶ The treatment variable is assignment to any type of program activities.

▶ Sample: Endline survey.

▶ Significance levels are indicated by $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$. N refers to the number of respondents.

Learning Outcomes   The NFBE is an innovative approach to education. The primary purpose of an NFBE curriculum is to help the learner complete the requisite level of education, in line with the objectives of the formal curriculum (Government of Sindh 2017). The idea is that children come to the centers at different ages and with different social backgrounds in order to learn together (e.g., 9 and 15 years old children learn together at grade 1). They go through an accelerated learning through need-based curricula, materials and teaching-learning methods. A number of assessments were designed to support children and teachers to identify their actual learning level (at entrance and exit). All centers were supposed to implement customized standards, curricula, assessment methods/tools, and learning materials. In what follows, a closer look is taken at potential changes in learning outcomes, which may reflect the quality of education provided at these institutions.

> The **NFBE-curriculum** covered languages (Sindhi, Urdu, English), Mathematics, Social Studies, Islamyat, and Sciences.
>
> In **Package A** children amongst others learned letters, shapes, the composition and decomposition of words, made role plays, and read paragraphs. In Mathematics, they learned how to count from one to hundred, learned addition of one and two digit numbers, the identification of the shapes circles, triangles, rectangles, and squares, watching time on clock, days of a week, and the concepts of upward to downward and right to left, front or behind, and away and near.
>
> In **Package B** children read paragraphs and stories, learned the usage of words in sentences (incl. masculine, feminine, singular/ plural), and learned grammar. In Mathematics, they learned measurements (distance, length), symbols (lesser or greater), mathematical operations (addition, subtraction, multiplication, division).
>
> In **Package C** the knowledge on Sciences was transferred. In Sindhi, they learned poems, read stories, proses and pictorial stories. In Mathematics, students expanded their understanding of numbers (billions, division, uniqueness, compound numbers, as well as weights, volume, time), learned fraction and multiplication, learned how to calculate percentages and basic statistics.

The NFBE curriculum was set up to ensure that learning objectives of the regular curriculum would be covered. Yet the expectation is for the learning process to be completed in the minimum possible time (Government of Sindh 2017). A Government-approved "Accelerated Non-Formal Education Curriculum" was used in the NFBE Centers. It consists of three separate packages, each one having direct equivalence and accreditation with the formal system. The duration of Package A, which is equivalent to grade 1 and 2 of formal education is 8 months; Package B, which is equivalent to grade 3 and 4 of formal education, is of 10 months and Package C, which is equivalent to grade 5 of the formal education, is of 14 months. The total duration of the non-formal education is 32 months (UNICEF 2017). Therewith, the three packages cover five grades over a three-year period. Children were admitted depending on their level of prior education. For instance, out-of-school children were admitted to Package A (grade 1) for a duration of eight months. They acquired knowledge up to grade 1. Children could continue with package B. Additionally, out-of-school dropouts (with completed grade 1) could join Package B (reflecting grades 2-3) to acquire core competency in about eight months (Package B).[46] The subjects covered varied by package (Government of Sindh 2016). Completion of Packages A and B was internally assessed and Package C was assessed through external examination.[47]Thus, the transition between packages was internal and after Package C a child could transition to any formal or non-formal post-primary institution as he/she will have acquired the required certification. In case a child wanted to enroll in formal school after completing Package A or B only, he/she could sit for placement tests (internal to the school).

> **Who is defined as NFBE-eligible in the evaluation report?**
>
> The NFBE eligibility is defined as adolescents that have actually dropped out of school or have been at risk of doing so. To identify them, the following questions were added into a so-called pre-roster (a short listing of all program eligible adolescents): "Are there any household members of the age group 10-23 who never went to any kind of formal school?" and "Did any household member of the age group 10-23 ever drop out of school?" Additionally, individuals who have been drop-out or never-enrolled-to-school children are considered (Section A.II.A). Moreover, these individuals are from 212 villages in Sindh (the so-called NFBE-eligible areas). Only NFBE-eligible children in NFBE-eligible areas (in program and control areas) were assessed in their learning abilities. If several individuals fulfill this eligibility criteria in a household, the youngest child is invited to the test.

---

[46]Accordingly, children who had dropped out after grade 2 could acquire grade 3 level in eight months by also joining Package B. After completing Package B or being a drop-out after grade 3, children could join package C for 14 months and complete grades 4 and 5. If a child was a drop-out after grade 4, he could join Package C for 7 months to complete grade 4. If a child dropped out at grade 5, he could also join package C for 7 months to complete grade 5.

[47]Certificates of Package A and B were issued by school principals (and co-signed by an area education officer) and certificates of Package C (primary school certificate) are issued by the examination section of the Federal Directorate of Education.

For the purpose of this evaluation, out-of-school children in NFBE-eligible areas were targeted. They were asked about their knowledge of the centers, their participation therein, and tested on their learning outcomes. For this, a special module was developed following UNICEF's shared material on the curriculum and the usual tools used to internally test the knowledge of the children (Government of Sindh 2016). All questions used are displayed in Section O.IV.A. Note, the following analysis focuses on learning outcomes for the age group 11-22.

Table 3.12 compares the conditional means in NFBE-treatment and NFBE-control areas.[48] Column (1) describes the means in the NFBE program areas, column (2) in control areas. Column (3) displays the difference-in-means between the groups (the p-values for the corresponding t-tests are indicated with stars). Comparing core background characteristics, the results indicate that there are as many female eligible respondents as male and that also the poverty level is comparable. There is a small significant difference in the age variable, with the adolescents in control villages being on average 0.38 years younger.[49]

The evaluation records an improvement in the ability to write and read in Sindhi due to the IALP program activities, while no improvement in Mathematics outcomes is recorded. Children in treatment areas have a significant improvement in the ability to write and read in Sindhi, skills taught under Package A.[50]

> In program areas more than every second child can write a sentence, while this is the case for only 45.8% in control areas.

The results indicate that in the treatment group more than every second child can write a sentence, this is the case for only 45.9% in control areas. The difference of 5.5pp is statistically significant at the 5% level. A similar results can be observed for reading skills which were tested with two questions. In both, eligible adolescents from NFBE areas perform better, with a significant difference of 5.3pp ($p < 0.05$) and 6.4pp ($p < 0.01$), respectively. Overall, the share of correct answers in Package A is significantly higher in treatment areas as compared to control areas. This is, however, not the case when analyzing more difficult questions asked, i.e., questions asked under Package B or C.[51]

Note, the number of respondents for package A, B and C questions are different: Only individuals who passed level A [B], were asked level B [C] questions. The number of respondents thus falls with every level and the sample composition changes accordingly, with higher level respondents having a higher level of knowledge. Within each level of similar skills, the program and control areas are compared. Finally, without further conditioning, the raw summary statistics indicate that eligible adolescents from NFBE areas do not perform better in Mathematics in program as compared to control areas. The only question they are more likely to answer correctly is how to read the clock (5.8pp difference, $p < 0.05$).

---

[48]The means are conditional on enumerator fixed effects and clustered at the RV-level.

[49]Note, however, there are no statistically significant age difference between treatment and control areas when comparing people who reached level B questions and level C questions. The age for people who reached level B is 15.79 in control and 15.98 in treatment (*p*-value of no difference 0.2137). The age for people who reached level C is 15.94 in control and 16.15 in treatment (*p*-value of no difference 0.2578).

[50]For a full list of original questions asked please refer to Section O.IV.A.

[51]Differences by gender are presented in Table 3.14 below.

Table 3.12: NFBE Learning Outcomes: Sindhi Reading and Literature

Variable	N/[Clusters]	(1) Treatment Mean/SE	N/[Clusters]	(2) Control Mean/SE	(3) t-test Difference (1)-(2)
respondent is female	1151	0.555	1270	0.570	-0.015
	[100]	[0.011]	[110]	[0.010]	
age at endline	1151	16.089	1270	15.706	0.384**
	[100]	[0.125]	[110]	[0.113]	
poverty indicator	1151	0.578	1270	0.580	-0.003
	[100]	[0.022]	[110]	[0.020]	
Package A: share of correct answers	992	0.437	1109	0.391	0.046**
	[100]	[0.015]	[110]	[0.014]	
correctly identified the first 3 letters of the alphabet	1017	0.412	1140	0.382	0.030
	[100]	[0.026]	[110]	[0.024]	
correctly wrote the sentence	1017	0.515	1140	0.458	0.057**
	[100]	[0.021]	[110]	[0.022]	
correctly read the sentence (1)	1017	0.518	1140	0.464	0.054**
	[100]	[0.021]	[110]	[0.022]	
correctly read the sentence (2)	1017	0.517	1140	0.453	0.065***
	[100]	[0.021]	[110]	[0.021]	
correctly identified the pictures that have plants or animals	1017	0.228	1140	0.211	0.018
	[100]	[0.024]	[110]	[0.022]	
Package B: share of correct answers	501	0.269	474	0.276	-0.007
	[96]	[0.015]	[106]	[0.015]	
correctly identified the verb	492	0.313	477	0.333	-0.020
	[96]	[0.026]	[107]	[0.023]	
correctly identified the sentence that should contain an interrogation sign	492	0.236	477	0.237	-0.001
	[96]	[0.021]	[107]	[0.020]	
Package C: share of correct answers	480	0.683	459	0.664	0.019
	[96]	[0.026]	[106]	[0.026]	
correctly said the sentence in past tense	492	0.768	477	0.719	0.049*
	[96]	[0.026]	[107]	[0.025]	
correctly identified any feature that helps Saharan ants survive in hot environments	492	0.587	477	0.610	-0.023
	[96]	[0.034]	[107]	[0.030]	

▸ *Notes*: Table 3.12 displays the summary statistics for performance in reading and literature.
▸ Sample: Endline survey. Only NFBE-eligible individuals (until age 22) in NFBE-eligible villages.
▸ Column (1) describes the means in the NFBE-assigned program areas, column (2) in control areas. Column (3) displays the difference-in-means between the groups (the p-values for the corresponding t-tests are indicated with stars).
▸ Enumerator fixed effects are controlled for.
▸ Standard errors are clustered at the village level. Significance levels are indicated by $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$. N refers to the number of respondents.

Table 3.13: NFBE Learning Outcomes: Mathematics

Variable	N/[Clusters]	(1) Treatment Mean/SE	N/[Clusters]	(2) Control Mean/SE	(3) t-test Difference (1)-(2)
respondent is female	1151	0.555	1270	0.570	-0.015
	[100]	[0.011]	[110]	[0.010]	
age at endline	1151	16.089	1270	15.706	0.384**
	[100]	[0.125]	[110]	[0.113]	
poverty indicator	1151	0.578	1270	0.580	-0.003
	[100]	[0.022]	[110]	[0.020]	
Package A: share of correct answers	992	0.656	1109	0.629	0.027
	[100]	[0.010]	[110]	[0.012]	
correctly identified the number eight	1017	0.805	1140	0.780	0.025
	[100]	[0.019]	[110]	[0.020]	
correctly identified the number missing	1017	0.689	1140	0.668	0.021
	[100]	[0.019]	[110]	[0.019]	
correctly identified the circle	1017	0.888	1140	0.865	0.023
	[100]	[0.013]	[110]	[0.013]	
correctly calculated the number of horses you have after receiving the present	1017	0.728	1140	0.683	0.044
	[100]	[0.021]	[110]	[0.023]	
correctly identified the figure with a smaller width	1017	0.499	1140	0.457	0.042
	[100]	[0.029]	[110]	[0.025]	
correctly identified the placement of the circle	1017	0.862	1140	0.843	0.019
	[100]	[0.017]	[110]	[0.016]	
correctly identified the largest number	1017	0.506	1140	0.510	-0.003
	[100]	[0.017]	[110]	[0.017]	
correctly identified the even numbers	1017	0.272	1140	0.277	-0.005
	[100]	[0.019]	[110]	[0.021]	
correctly identified the time	1017	0.656	1140	0.598	0.058**
	[100]	[0.018]	[110]	[0.021]	
Package B: share of correct answers	747	0.496	797	0.483	0.013
	[100]	[0.013]	[110]	[0.013]	
correctly calculated the answer	767	0.292	823	0.278	0.014
	[100]	[0.022]	[110]	[0.019]	
correctly calculated the number of hours passed since waking up and until having lunch	767	0.580	823	0.537	0.043
	[100]	[0.027]	[110]	[0.024]	
correctly calculated the perimeter of the triangle	767	0.296	823	0.313	-0.018
	[100]	[0.027]	[110]	[0.026]	
correctly identified which team has won	767	0.799	823	0.813	-0.014
	[100]	[0.032]	[110]	[0.028]	
correctly calculated the difference in height between the two shirts	767	0.399	823	0.382	0.017
	[100]	[0.025]	[110]	[0.023]	
correctly calculated the quantity of chillies that Rabia's mother has	767	0.355	823	0.341	0.013
	[100]	[0.020]	[110]	[0.020]	
correctly identified how many seconds there are in 2 minutes	767	0.733	823	0.730	0.002
	[100]	[0.020]	[110]	[0.021]	
Package C: share of correct answers	347	0.204	343	0.190	0.014
	[93]	[0.008]	[99]	[0.007]	
correctly calculated the answer to the expression on the screen	352	0.304	355	0.296	0.008
	[93]	[0.027]	[103]	[0.031]	
correctly identified the lowest common multiple of 12, 36, and 3	352	0.017	355	0.017	0.000
	[93]	[0.007]	[103]	[0.007]	
correctly calculated the answer to the expression on the screen	352	0.080	355	0.045	0.034
	[93]	[0.025]	[103]	[0.013]	
correctly rounded up to the first decimal	352	0.003	355	0.000	0.003
	[93]	[0.003]	[103]	[0.000]	
correctly identified how many yellow balls there are	352	0.861	355	0.817	0.044
	[93]	[0.018]	[103]	[0.026]	
correctly calculated how many milliliters of mixture (water and juice) you will have	352	0.000	355	0.000	N/A
	[93]	[0.000]	[103]	[0.000]	
correctly identified a right triangle & provided correct reasoning for choice	352	0.057	355	0.048	0.009
	[93]	[0.014]	[103]	[0.012]	
correctly marked two parallel lines	352	0.315	355	0.293	0.022
	[93]	[0.037]	[103]	[0.031]	

▸ *Notes*: Table 3.13 displays the summary statistics for performance in Mathematics.
▸ Sample: Endline survey. Only NFBE-eligible individuals (until age 22) in NFBE-eligible villages.
▸ Column (1) describes the means in the NFBE-assigned program areas, column (2) in control areas. Column (3) displays the difference-in-means between the groups (the p-values for the corresponding t-tests are indicated with stars).
▸ Enumerator fixed effects are controlled for.
▸ Standard errors are clustered at the village level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.

In a set of linear regressions, differences in age can be controlled for and whether the impacts differ by age can also by analyzed. Here, it can also be controlled for the fact Child Protection interventions took place in part of the villages. Table 3.14, Panel A, describes the effects of the program for all eligible individuals randomly sampled in the village. Unreported in the table are gender, age, and district control variables. The standard errors are clustered at the community level and it is additionally controlled for enumerator fixed effects.

Positive impacts in learning outcomes are especially visible in areas where both the Child Protection and NFBE *treatments were implemented as compared to the control group.* In column (1), the results indicate that the share of children with correctly responded questions are higher in program areas. Especially the coefficient in areas where both types of treatments were implemented indicates an increase in 5.8pp, from a control mean of about 39%, an effect significant at the 5% level. If children responded to more than 50% of the questions correctly, they were displayed level B questions. In the treatment group, 10.1pp more eligible individuals progressed to the next level (i.e., responded to more A-level questions correctly). For those individuals who reached level B, the share of level B questions correctly responded in the treatment group is not higher than in the control group. Still, out of all respondents, on average the individuals are 8.1pp more likely to reach level C. The results indicate no impacts on aggregated Mathematics outcomes, neither on reaching a higher level nor on the share of correct responses within a level, when comparing treatment and control areas.

The overall positive effects on learning outcomes are largely driven by the improvement in outcomes amongst males. The evaluation now compares the impacts of the program for male and female adolescents separately. To this extent, the following two panels split the sample by gender: Panel B of Table 3.14 reports results only for male adolescents (N=913) and Panel C only for female adolescents (N=1,188). What becomes clear is that the overall effects are largely driven by an improvement in outcomes amongst males. For instance, the NFBE effect on reaching Package B in literature (i.e., answering sufficient questions correctly in Panel A) is an increase by 19.5pp in NFBE villages only and 12.0pp in NFBE and Child Protection villages (column B). The overall NFBE effect is statistically significant with a p-value of 0.014. Moreover, the results show no statistical difference in the impact of the two different approaches to programming (Child Protection with NFBE and NFBE alone) - in other words, there is no additional effect that can be claimed with sufficient certainty of adding Child Protection interventions to improve learning outcomes at level B.[52] Additionally, the results indicate that more male adolescents correctly respond to Package A Mathematics questions. A significant improvement by 4.9pp, from a control mean of about 65.2%, an effect significant at the 10% level. Combining the results for Mathematics and Sindhi, it can be observed that boys in program areas where only the NFBE program was implemented had on average a 12.1pp higher likelihood to reach level C questions (column (12), Panel B).

For female adolescents, the results indicate an impact of 6.4pp in Package A share of correctly responded questions, given a control mean of about 35.3% (an effect significant at the 10% level, reported in Panel C, column (1)). Notably, the impact for NFBE and Child Protection villages here is statistically different from the impact achieved in NFBE villages only, suggesting that the combination of the different program activities yields better outcomes for girls. In other words, adding Child Protection interventions pays off for girls. At the same time, across all outcomes measured, the coefficients in NFBE villages only are instead negative and in one case even weakly statistically significant (column (8)).

Especially the reading and writing outcomes of children and adolescents age 15-18 were positively impacted by the program. The effects by age are reported in Table 3.15, age group 11-14 with 837 respondents, 15-18 with 859 respondents, 19-22 with 405 respondents in program and control areas. The results indicate positive impacts for the age group of 11-14 year old respondents in reading and writing as captured by the higher share of correctly responded

> Boys' and girls' learning outcomes differ, as does their improvement through the program. In Sindhi and Mathematics, girls' learning outcomes are worse than boys'. And while boys benefit from NFBE interventions alone and there seems to be no added value of CP activities for their learning outcomes, girls benefit only from the combination of both types of interventions. This is in line with the results that only in areas where both interventions took place, they were more likely to attend NFBE Centers. However, overall, there is only little certainty about the improvement in girls' learning outcomes.

---

[52]In the tables it is indicated with which certainty (p-value) an overall zero effect of NFBE can be rejected, e.g. NFBE effect=0. The overall effect considers the effect from Child Protection with NFBE and NFBE villages. Similarly, it is analyzed with what certainty the null hypothesis that the effect size in NFBE villages only (ass_treat_NFBE_only) is equal to the effect of the villages where both types of interventions were implemented (ass_treat_CP_NFBE=1) can be rejected.

questions in Package A (column (1)). However this effect is only weakly statistically significant. The age group of 15-18 year old respondents is more likely to respond to more than 50% of the questions correctly (13.1pp higher likelihood to reach level B given a share of 40% in the control group and 13.2pp to reach level C given a control mean of 41%). However, in Mathematics, there is weak evidence for this group of having a lower likelihood to correctly respond to questions in Package B-Mathematics. Effects on eligible individuals of higher age (19-22) who have ever dropped out from school in the past are also negative.

Table 3.14: Impacts on Learning Outcomes, By Packages and Gender

	(I) Sindhi					(II) Mathematics					(III) Any	
	A	B		C		A	B		C		B	C
	(1) share correct	(2) level B reached	(3) share correct	(4) level C reached	(5) share correct	(6) share correct	(7) level B reached	(8) share correct	(9) level C reached	(10) share correct	(11) share reached	(12) share reached
Panel A: All												
only NFBE	0.013	0.063	-0.000	0.044	0.034	0.007	-0.016	-0.016	0.018	-0.010	-0.018	0.005
	(0.028)	(0.043)	(0.036)	(0.042)	(0.038)	(0.018)	(0.035)	(0.024)	(0.036)	(0.016)	(0.035)	(0.045)
CP and NFBE	0.058**	0.101**	0.027	0.081*	-0.014	0.010	-0.014	-0.036	-0.042	0.018	0.001	0.001
	(0.028)	(0.044)	(0.031)	(0.044)	(0.037)	(0.018)	(0.033)	(0.024)	(0.034)	(0.017)	(0.035)	(0.044)
only CP	0.007	0.015	0.016	0.002	-0.019	0.002	-0.021	-0.028	-0.024	-0.001	-0.028	-0.047
	(0.027)	(0.042)	(0.028)	(0.043)	(0.037)	(0.017)	(0.032)	(0.023)	(0.032)	(0.016)	(0.032)	(0.045)
Control Mean	0.390	0.442	0.256	0.431	0.698	0.611	0.714	0.501	0.312	0.198	0.810	0.550
NFBE effect=0 (p-value)	0.149	0.038	0.637	0.112	0.755	0.592	0.627	0.220	0.688	0.779	0.779	0.948
Equality of CP with NFBE and NFBE alone treatment arms (p-value)	0.089	0.305	0.422	0.294	0.149	0.854	0.946	0.360	0.109	0.021	0.555	0.917
Ind.	2,101	2,101	975	2,101	939	2,101	2,101	1,544	2,101	690	2,101	2,101
R^2	0.182	0.146	0.090	0.156	0.408	0.293	0.244	0.286	0.235	0.262	0.157	0.168
Panel B: Male												
only NFBE	0.090**	0.195***	0.027	0.177***	0.035	0.049*	0.064	0.021	0.054	-0.020	0.080	0.121**
	(0.041)	(0.067)	(0.052)	(0.066)	(0.055)	(0.026)	(0.054)	(0.032)	(0.052)	(0.020)	(0.050)	(0.059)
CP and NFBE	0.048	0.120*	0.034	0.099	-0.027	0.022	0.036	-0.028	-0.061	0.005	0.066	0.008
	(0.040)	(0.068)	(0.048)	(0.066)	(0.051)	(0.024)	(0.053)	(0.031)	(0.052)	(0.019)	(0.049)	(0.057)
only CP	0.021	0.065	0.022	0.049	-0.022	0.009	-0.001	-0.022	-0.059	-0.027	0.021	-0.019
	(0.040)	(0.067)	(0.047)	(0.065)	(0.051)	(0.024)	(0.053)	(0.030)	(0.047)	(0.018)	(0.050)	(0.058)
Control Mean	0.441	0.500	0.232	0.491	0.727	0.652	0.786	0.547	0.429	0.214	0.830	0.634
NFBE effect=0 (p-value)	0.071	0.014	0.513	0.027	0.932	0.125	0.328	0.898	0.939	0.668	0.125	0.232
Equality of CP with NFBE and NFBE alone treatment arms (p-value)	0.172	0.100	0.850	0.085	0.147	0.170	0.391	0.073	0.026	0.110	0.651	0.011
Ind.	913	913	507	913	490	913	913	749	913	401	913	913
R^2	0.189	0.165	0.152	0.175	0.440	0.262	0.197	0.302	0.247	0.283	0.127	0.185
Panel C: Female												
only NFBE	-0.045	-0.027	-0.041	-0.046	0.006	-0.029	-0.077	-0.063*	-0.028	-0.005	-0.093*	-0.083
	(0.040)	(0.063)	(0.059)	(0.062)	(0.062)	(0.026)	(0.055)	(0.034)	(0.052)	(0.022)	(0.052)	(0.064)
CP and NFBE	0.064*	0.088	0.002	0.070	-0.033	-0.002	-0.051	-0.053	-0.042	0.038	-0.049	-0.007
	(0.036)	(0.056)	(0.050)	(0.057)	(0.060)	(0.024)	(0.045)	(0.032)	(0.047)	(0.023)	(0.045)	(0.061)
only CP	-0.011	-0.028	-0.006	-0.039	-0.052	-0.007	-0.034	-0.048	-0.017	0.028	-0.064*	-0.076
	(0.032)	(0.051)	(0.049)	(0.053)	(0.060)	(0.022)	(0.042)	(0.030)	(0.044)	(0.022)	(0.038)	(0.057)
Control Mean	0.353	0.401	0.278	0.389	0.672	0.582	0.662	0.462	0.229	0.177	0.796	0.490
NFBE effect=0 (p-value)	0.775	0.565	0.689	0.823	0.811	0.469	0.137	0.051	0.422	0.418	0.081	0.427
Equality of CP with NFBE and NFBE alone treatment arms (p-value)	0.007	0.046	0.382	0.033	0.400	0.288	0.634	0.766	0.764	0.022	0.422	0.160
Ind.	1,188	1,188	468	1,188	449	1,188	1,188	795	1,188	289	1,188	1,188
R^2	0.198	0.159	0.144	0.167	0.414	0.314	0.269	0.263	0.210	0.342	0.179	0.159

▸ *Notes*: Table 3.14 displays the program effects on adolescent's learning outcomes. The dependent variables as follows: Share of correct responses in Package A reported in column (1) and (6). Whether level B was achieved recorded in columns (2), (7), and (11). Share of correct responses in Package B reported in column (3) and (8). Whether level C was achieved recorded in columns (4), (9), and (12). Share of correct responses in Package C reported in column (5) and (10). Effects from separate OLS estimations.
▸ Sample: Endline survey. Only NFBE-eligible individuals (until age 22) in NFBE-eligible villages.
▸ Part I reading and literature outcomes (in Sindhi). Part II Mathematics outcomes. Part III either of the two subjects. Panel A all respondents. Panel B only male respondents. Panel C only female respondents.
▸ Apart from the indicated variables, it is controlled for enumerator fixed effects, gender, age and district fixed effects.
▸ Standard errors are clustered at the village level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.

Table 3.15: Impacts on Learning Outcomes, By Packages and Age Groups

	(I) Sindhi					(II) Mathematics					(III) Any	
	A	B		C		A	B		C		B	C
	(1) share correct	(2) level B reached	(3) share correct	(4) level C reached	(5) share correct	(6) share correct	(7) level B reached	(8) share correct	(9) level C reached	(10) share correct	(11) share reached	(12) share reached
Panel A: Ages 11 -14												
only NFBE	0.034	0.074	-0.024	0.030	0.125*	0.017	0.028	0.006	0.033	-0.014	0.047	0.019
	(0.032)	(0.055)	(0.062)	(0.050)	(0.066)	(0.021)	(0.054)	(0.035)	(0.049)	(0.023)	(0.049)	(0.051)
CP and NFBE	0.058*	0.087	0.028	0.058	0.046	0.001	-0.059	-0.038	-0.121***	0.025	-0.017	-0.048
	(0.034)	(0.055)	(0.058)	(0.055)	(0.061)	(0.021)	(0.048)	(0.031)	(0.043)	(0.020)	(0.051)	(0.053)
only CP	0.002	-0.002	-0.006	-0.030	0.014	-0.014	-0.067	-0.014	-0.077*	0.009	-0.060	-0.096**
	(0.029)	(0.048)	(0.051)	(0.048)	(0.058)	(0.020)	(0.044)	(0.030)	(0.041)	(0.018)	(0.044)	(0.047)
Control Mean	0.355	0.411	0.255	0.411	0.698	0.571	0.659	0.484	0.295	0.174	0.767	0.535
NFBE effect=0 (p-value)	0.084	0.091	0.967	0.336	0.131	0.627	0.731	0.548	0.242	0.775	0.722	0.747
Equality of CP with NFBE and NFBE alone treatment arms (p-value)	0.530	0.821	0.422	0.598	0.169	0.471	0.099	0.248	0.003	0.061	0.216	0.210
Ind.	837	837	333	837	319	837	837	553	837	227	837	837
R²	0.215	0.233	0.170	0.240	0.500	0.351	0.279	0.352	0.285	0.402	0.230	0.236
Panel B: Ages 15 -18												
only NFBE	0.032	0.096	-0.003	0.108	-0.012	0.026	0.003	-0.049	0.014	-0.010	-0.017	0.026
	(0.045)	(0.068)	(0.061)	(0.070)	(0.064)	(0.031)	(0.059)	(0.039)	(0.055)	(0.029)	(0.060)	(0.071)
CP and NFBE	0.071	0.131**	0.047	0.132*	-0.035	0.023	0.035	-0.061*	-0.028	0.023	0.018	0.032
	(0.043)	(0.065)	(0.054)	(0.067)	(0.061)	(0.028)	(0.053)	(0.035)	(0.053)	(0.030)	(0.053)	(0.068)
only CP	-0.003	0.023	0.040	0.031	-0.072	0.021	0.036	-0.066*	-0.022	0.000	-0.003	-0.038
	(0.040)	(0.064)	(0.051)	(0.067)	(0.060)	(0.027)	(0.051)	(0.034)	(0.050)	(0.028)	(0.051)	(0.068)
Control Mean	0.400	0.440	0.239	0.410	0.707	0.622	0.720	0.538	0.330	0.205	0.830	0.560
NFBE effect=0 (p-value)	0.188	0.061	0.682	0.057	0.692	0.360	0.705	0.099	0.883	0.814	0.992	0.647
Equality of CP with NFBE and NFBE alone treatment arms (p-value)	0.336	0.531	0.298	0.658	0.618	0.933	0.503	0.710	0.369	0.079	0.452	0.915
Ind.	859	859	419	859	403	859	859	654	859	286	859	859
R²	0.207	0.129	0.138	0.148	0.432	0.300	0.275	0.290	0.247	0.320	0.161	0.161
Panel C: Ages 19 -22												
only NFBE	-0.115*	-0.066	0.016	-0.095	-0.100	-0.055	-0.147**	0.031	0.049	-0.068**	-0.147**	-0.013
	(0.065)	(0.096)	(0.071)	(0.096)	(0.086)	(0.036)	(0.064)	(0.063)	(0.094)	(0.027)	(0.061)	(0.089)
CP and NFBE	0.000	0.044	-0.024	0.002	-0.122	0.007	-0.031	0.035	0.110	-0.043*	-0.006	0.075
	(0.065)	(0.098)	(0.069)	(0.096)	(0.080)	(0.037)	(0.064)	(0.066)	(0.100)	(0.024)	(0.060)	(0.087)
only CP	0.006	-0.003	-0.018	-0.022	-0.044	-0.005	-0.040	0.019	0.092	-0.068***	-0.015	0.055
	(0.064)	(0.096)	(0.065)	(0.097)	(0.081)	(0.036)	(0.065)	(0.068)	(0.099)	(0.025)	(0.060)	(0.090)
Control Mean	0.475	0.550	0.295	0.550	0.682	0.711	0.875	0.465	0.325	0.250	0.900	0.575
NFBE effect=0 (p-value)	0.345	0.900	0.952	0.605	0.147	0.477	0.134	0.594	0.387	0.015	0.166	0.704
Equality of CP with NFBE and NFBE alone treatment arms (p-value)	0.014	0.104	0.461	0.143	0.744	0.029	0.018	0.911	0.331	0.292	0.004	0.207
Ind.	405	405	223	405	217	405	405	337	405	177	405	405
R²	0.242	0.208	0.251	0.227	0.463	0.285	0.232	0.330	0.235	0.420	0.159	0.193

‣ *Notes*: Table 3.15 displays the program effects on adolescent's learning outcomes. The dependent variable as follows: Share of correct responses in Package A reported in column (1) and (6). Whether level B was achieved recorded in columns (2), (7), and (11). Share of correct responses in Package B reported in column (3) and (8). Whether level C was achieved recorded in columns (4), (9), and (12). Share of correct responses in Package C reported in column (5) and (10). Effects from separate OLS estimations.
‣ Sample: Endline survey. Only NFBE-eligible individuals (until age 22) in NFBE-eligible villages.
‣ Part I reading and literature outcomes (in Sindhi). Part II Mathematics outcomes. Part III either of the two subjects. Panel A respondents of age 11-14. Panel B respondents of age 15-18. Panel C respondents of age 19-22.
‣ Apart from the indicated variables, it is controlled enumerator fixed effects, gender, age and district fixed effects.
‣ Standard errors are clustered at the village level.

The introduction of NFBE Centers leads to an increased participation and better learning outcomes, especially for boys. However, the evaluation also finds evidence of potential crowding-out from formal education.

Evidence from Qualitative Interviews   *Assessment of Effect on Education by Key Informants* According to some interview partners, attendance of NFBECs empowered children and adolescents to learn how to read and write, and to ultimately obtain a primary school diploma. A KI stated that attendance of an NFBEC also enabled adolescents to express to their parents their desire for further education. After completion of all three NFBE packages, adolescents reportedly have the opportunity to enroll in grade 6 of a nearby school. With regard to gender inequality, according to an interviewee, male family members became aware of the importance of education for girls. As with the results concerning child marriage, there was also a more critical voice concerning the effect that participation in the NFBE had on adolescents, while still acknowledging that the activities had generated awareness for the importance of education and the possibility for adolescents to return to school.

> "The project is very limited in scope and limited in time. So, we cannot say that X activity resulted in doing that change. However, what we hope is that through the education component we were able to raise a lot of awareness, allowing adolescents who were out of school for a long time to reconnect to their learning opportunity. And giving ideas to the community members that yes, if a child has missed out on their education (...) they can reconnect again. (...) But if it directly resulted in behavior change, you know, that is something that takes time (...)." (Key Informant No.2)

Recommendations of Key Informants Regarding Inclusiveness Regarding the program design, several interview partners suggested to always use an integrated approach in which adolescents are simultaneously part of the CP and the NFBE intervention, as they often require both. Some interviewees proposed not separating groups by gender slightly contradicting observations of other KIs who recounted that separation by gender was important in NFBECs to overcome obstacles to the inclusion of girls. A concrete suggestion to ease the participation of girls in future programs was to provide money for transportation or safe travel opportunities "because walking to these service centers for adolescent girls becomes difficult" (Key Informant No.2).[53] This advice is in line with findings from the literature that observed an increase for girls' enrollment rates in India when funds for transportation were provided (Muralidharan and Prakash 2017).

Interviewees also offered suggestions for changes to the program content, such as including vocational training from the start, focusing skills development on those most relevant nowadays, such as IT skills, and including visits to other schools and factories in the NFBE curriculum. KIs identified the underlying problem of lacking opportunities for adolescents, which the interviewees saw as important for translating the awareness raised among adolescents into further impacts. Other interview partners demanded that adolescents be offered jobs or given cash transfers for economic activities.

> "There should be opportunities so that they [adolescents] can contribute. We have taught them, gave awareness, but when there will be no opportunities, no platform, then what is the use of telling them. All will be wasted." (Key Informant No.12)

The Alternative Learning Program as a Way Back to Formal Education? out-of-school adolescents were taught NFBE, and NFBECs as well as VECs were established.[54] According to a stakeholder, NFBECs were established in Sindh in 2016 with the curriculum being approved by the School Education and Literacy Department of Sindh later that year.[55] Interview partners stated that overall, the implementation of the educational intervention in Sindh was facilitated by the previous implementation of the CP component as well as by the existing infrastructure of the local IP, IRC, which would have helped with hiring teachers and networking. According to interview partners, most NFBECs would have been separated by gender and, in most cases, female teachers were assigned to girls.[56] One interviewee attributed the

---

[53]As part of the IALP project, transport allowances were provided to 250 adolescents attending vocational skills training courses (UNICEF 2019). This was suggested to be extended to all adolescents who require means of transportation.

[54]In Punjab, the Government is responsible for the provision of non-formal education. Nevertheless, in both districts, Rahim Yar Khan and Bahawalpur, Bunyad established 84 non-formal education facilities in cooperation with the American Refugee Committee starting from the end of 2019 (UNICEF 2019). As these, however, present an intervention outside the IALP program, the following section will address the NFBE intervention only with regard to Sindh.

[55]By December 2017, 60 NFBECs had been established in Sindh (UNICEF 2017). As stated above, the curriculum consisted of three packages that cover different grades of primary school with the total duration of the NFBE mounting up to 32 months (UNICEF 2017).

[56]In Sindh, 58 of 100 teachers trained in the NFBE curriculum were female (UNICEF 2019).

perceived success of the NFBE to the teachers selected. Attendance was reportedly flexible to accommodate for working adolescents who could only join in the evening. Harvest season was also taken into consideration for planning the time schedule. Other KIs highlighted that NFBECs were also providing for the participation of teenage mothers and married adolescents.

> "Timing was flexible if children had to go to work in the morning, they were able to join center in the evening." (Key Informant No.7)
> "I think ALP is a very successful program (...). For example, in some instances, married adolescent girls also participated in the ALPs and they used to bring their kids with them and be supported at the same time. While others would leave their children at their mother-in-law's place. In case they were breastfeeding or wanted to see their kids, space was provided to them." (Key Informant No.3)

To foster an enabling environment for adolescents to return to education, the creation of and support for the VECs was perceived as being crucial.[57] According to an interview partner, IRC regularly sensitized communities and parents regarding the importance of adolescents' education and facilitated the attendance of government officials in the NFBE activities, which was perceived as being a motivational factor for the community. Moreover, it was reported that UNICEF conducted capacity building sessions with government staff involved in non-formal education.

> "Community engagement is key for any work to be done at the field level, especially for the non-formal education. So, some Village Education Committees were formed, and this was the thing that played a critical role in ensuring that we can reach out to out-of-school adolescent children." (Key Informant No.2)

Institutional and Infrastructural Challenges Similar to the institutional problems mentioned with regard to the CP intervention, an interview partner described how non-formal education was only recently introduced at district level. Several issues regarding the selection of the location for the NFBECs came up in the interviews. First, it was described that the selected villages were quite remote and disagreements about where centers should be located arose with communities. Difficulties in locating the centers were also reported in connection with tribes living in the area who did not want their children to mix with those of other tribes. This also entailed questions from the community when people from other castes entered their neighborhood to attend NFBECs. Another problem involved the absence of proper sanitation in some NFBECs, especially in poor areas.

> "(...) in some places, there were the tribal and baradari issues. In some places the baradari makes the center, they are not allowing their children to the other baradari or tribes." (Key Informant No.5)
> "Problems like location of the centers, as you know about the typical thoughts. 'Why would people from any other caste enter our area?', 'Why center is in that area, why isn't here?'. And conflicts are raised by these statements." (Key Informant No.7)

A further challenge in establishing NFBECs involved the *lack of qualified teachers* who could teach the NFBE and who could cope with having students of very different ages in one class. According to one KI, teachers were initially incapable of teaching the NFBE content. An effort was made to find female teachers for girls to ease parents' concerns related to girls' attendance of NFBECs in the presence of boys or male teachers, though this proved challenging.

> "It's not easy for a teacher to teach in that setting where you have children coming from various age groups, various backgrounds in one center and going through an accelerated curriculum." (Key Informant No.2)
> "There were lots of challenges. First, to mobilize the parents of the children, especially female children. There should be female teachers for female students. Mostly people are hesitant to send their females where there are male children or teachers." (Key Informant No.11)

A few KIs reported problems with, first, identifying OOS adolescents, and second, enrolling them in NFBE facilities and ensuring attendance. According to one interview partner, the identification of OOS adolescents was complicated by the structure of the revenue villages, each of which consists of several smaller villages. On the other hand, it was argued that the large number of OOS adolescents in rural Sindh made it difficult to select just a few. Furthermore, reaching OOS adolescents was described as challenging due to their engagement in child labor or other employment as well as their married status. Enrolling younger children in NFBECs appeared to be easier than enrolling older children who are exposed to economic or social pressure to provide for their families. Moreover, one interview partner remarked that adolescents with previous exposure to formal education might feel that they do not profit from attending school, making

---

[57]Similarly, the UNICEF ROSA Annual Progress Report 2018 also identified VECs as playing an essential role (UNICEF 2018).

it difficult to motivate such OOS adolescents to return to school. In some centers low attendance was noticed, which was attributed to seasonal work in agriculture, economic migration, and the impact of COVID-19.

> "That's a very difficult age group to reach out to, because then, most of the children are either in working or, you know, negative coping mechanisms are taking place, girls' mobility is restricted, girls get married off at this age, then there is child labor, including household labor. So, a lot of community mobilization is needed to be done and to engage along with those structures." (Key Informant No.2)
> "And then the mobilization to get these children back into formal or structured learning because they have been out of the system for long. So, it is more difficult to get them back and make it more relevant, so they find it something that they want to invest their times in." (Key Informant No.2)

Obstacles for Adolescent Girls were not only limited to concerns surrounding interaction with male NFBEC staff or attendees. As other interview partners pointed out in relation to training female adolescent champions, travelling to NFBE facilities was not always possible for girls due to security concerns, financial issues or cultural reasons. Another interviewee suggested that at the heart of the problem was the lack of schools close to home, and that, contrary to common perception, parents were eager to allow their daughters to attend school.

> "(...) [the] challenge was to find a conflict free, neutral and safe place, especially for the girls, so that parents do not feel hesitant to send their female children." (Key Informant No.11)
> "Everywhere you would hear that 'We want to teach our girls but there is no school for them'. To send the girls 5 to 6 km in the city was a challenge. They used to feel insecure and had protection issues as well, transportation issues, affordability issues. The poor couldn't afford private transport, so these were the issues." (Key Informant No.3)

II.B   Intermediate Outcomes: Empowerment, Communication, Social Norms, and Service Provision

In the following paragraphs, a closer look is taken at intermediate outcomes that are more directly linkable to the activities, especially as implemented under the Child Protection component. For instance, according to the program's theory of change, the greater strategic goals in changing child marriage and education were to be achieved by encouraging exchange and interaction between adolescents as well as adolescents and adults. Moreover, key local stakeholders were to be equipped with the knowledge and skills to protect adolescents against rights violations and to take active action against it. Eventually, the need to underline the importance of protection of adolescents requires a change in norms in an environment in which corporal punishments are still likely to occur. Social mobilization efforts and messages around child marriage, gender related norms and adolescent protection were envisaged to shift attitudes, and norms which may eventually translate to (further) changes in behavior. Finally, under the so-called Pillar III of the program, a core output was to target service providers to promote and provide access to services for adolescents.

II.B.1   Adolescent Empowerment

Key Findings on Adolescent Empowerment Outcomes

One of the core program activities was to train adolescent to become champions and to trickle this knowledge down to the communities, a fundamental idea to increase adolescent empowerment (UNICEF South Asia 2019b). Adolescent champions were educated based on the *Life Skilled Based Education* toolkit, with topics ranging from building confidence and courage to advocating for changes in their communities that will promote their wellbeing. The following paragraphs test, in how far this driving factor of further change was achieved by the program. Moreover, according to the theory of change, one of the program's core outputs under program Pillar II was to equip key local stakeholders with the knowledge and skills to protect adolescents against rights violations and to take active action against it.

- The IALP program has a positive impact on the share of people reached on child rights topics. This effect is dominantly recorded by male respondents in program areas;

- In terms of the program's impact on adolescents' confidence, the results indicate no impact on adolescents' confidence in day-to-day situations, whereas their confidence in expressing their needs increases. Again, this program effect is captured for male adolescents only;

- Both adolescent's confidence in day-to-day situations and expressing their opinion in their community/ with public representatives remain at low levels and are significantly lower for females, which leaves room for future improvements;

- In terms of self-efficacy, the findings indicate no program impacts on adolescents' feelings of being able to influence their fate, nor on actions taken to ensure adherence to adolescent rights in the last year;
- The results show a small but negative impact on the reported wellbeing of female adolescents over the last two weeks. The results may be due to the program or the end of it;
- The program seems to reduce the acceptance of using corporal punishment on daughters when discussing their marriage timing, especially among male respondents (i.e., adolescent boys or fathers of an adolescent);
- The program leads to a decrease in the experience of corporal punishment by adolescents, especially in NFBE areas.

Table 3.16 displays the effects from a set of OLS regression results studying several empowerment-related variables.

Selected activities under the IALP program had a positive impact on the share of boys reached on child rights topics. Part I of the table considers awareness of child rights.[58] The "reached on child rights" indicator captures that respondents mentioned "topics on child rights and entitlements" as topics covered by the event they attended. In areas where both the Child Protection and educational intervention were implemented, the results indicate an increase in the share of people reached on child rights topics by 14.0pp for males only (significant at the 5% level, relative to a control mean of 12.2%). Note, however, the RV sample size in these estimations is small due to non-response.

Randomization of Questions in Survey Modules In order to reduce the length of the questionnaire, certain scales were randomized across respondents so that not all respondents were asked all questions.

For example, for the topic on attitudes towards corporal punishment, one third of respondents was asked a set of questions on their agreement with the use of corporal punishment at specific locations, one third was asked a set of questions regarding their agreement with the use of corporal punishment against daughters in specific situations, while the last third of respondents was asked a set of questions regarding their agreement with the use of corporal punishment against sons n specific situations.

The scales that were randomized (and thus exhibit smaller sample sizes) are the following:

- Attitudes towards corporal punishment scales (see example above);
- Personality and wellbeing scales: These include adolescents' confidence in day-to-day situations scale, interpersonal communication competence (ICC) scale, WHO wellbeing scale, and locus of control scale. Here, all respondents were randomly allocated two (out of the four) scales;

The program had no impact on adolescents' confidence in day-to-day situations, which remains at low levels and displays significant gender differences. Part II of Table 3.16 considers program effects on adolescents' confidence. In order to capture adolescents' confidence in day-to-day situations, adolescents were asked to indicate on a scale from 1 (not at all confident) to 7 (very confident) how confident they are in performing a number of activities by themselves, such as going to the local market alone or going to school or NFBE Center.[59] The summary variable captures, out of the six activities, the share of activities the adolescent feels (very) confident about performing on a scale from 0 to 1.[60] The results indicate no impacts of the program or its separate components on this indicator compared to the control group. In program and control areas, children feel similarly confident in performing the activities. Note the low levels and the gender difference, which leave space for improvement in future: While only 35.0% of boys in control group areas are confident, the share is even lower for girls at 14.8%.

Adolescents' confidence expressing their opinion in their community/ with public representatives increases due to the program amongst male respondents only. However, overall levels remain low and again disclose significant gender differences. In order to capture adolescents' confidence to express their opinion in the public sphere, adolescents and parents were asked whether they (or their child) express community needs to local government officials, whether they feel confident expressing their needs to the local government officials, and whether, over the last four years, adolescents in their community have gotten together to request local government officials or political leaders for anything benefiting

---

[58]In order to capture whether respondents (adolescents, parents, community leaders, and Nikah Khwans) were reached on child rights-related topics, respondents were asked to indicate the topics of the event (or activity) they were aware of (i.e., conditional on being aware of any event (or activity) taking place in the village in the last four years).

[59]The individual variables are first converted to binary variables, which are coded as 1 if the adolescent answered being confident (=6) or very confident (=7), and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of activities the adolescent feels (very) confident about performing (max. six). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of questions included in the summary variable (i.e, by six).

[60]Note that the adolescents' confidence in day-to-day situations scale was randomized, i.e., not all adolescents were asked these questions. For a brief explanation of the randomization of scales see Box II.B.1 above.

adolescents.[61] The summary variable indices, out of the three situations, the share of questions to which the answer was yes, i.e., the share of situations in which adolescents feel confident in their interactions with key stakeholders, in particular government officials or political leaders for anything benefiting adolescents. In control areas, the average share of adolescents' confidence expressing their opinion in their community and with public representatives is reportedly 18.1%. Again, note the low levels and the gender difference, which leave space for improvement in future: While only 23.6% of boys in control group areas are confident in their interactions with key stakeholders, the share is even lower for girls at 13.0%. The results indicate an increase by 4.4pp for males only, in areas where both program components were implemented jointly (significant at the 5% level).[62]

The evaluation results show a small but negative impact on the reported wellbeing of female adolescents over the last two weeks, which may be due to the program or the end of it. Part III of the table focuses on adolescents' wellbeing. As a short self-reported measure of current mental wellbeing, the World Health Organization-Five Wellbeing Index (WHO-5) is employed, which asks about personal wellbeing over the last two weeks. Adolescents and their parents were asked to indicate how frequently on a scale from 1 (all the time) to 6 (at no time) they (or their child) felt certain positive emotions, such as cheerful and in good spirits or calm and relaxed.[63] The summary variable indices the share reached on the WHO-5 scale, with 1 representing the best imaginable wellbeing.[64] The results indicate a negative impact in wellbeing by 4.3pp in areas where the Child Protection and educational interventions were jointly implemented (significant at the 1% level). Moreover, column (12) reveals that the effect is driven by changes for girls. Here, the negative impact is 6.6pp compared to the control group. In addition, a negative impact is observed in areas where Child Protection interventions were implemented in the NFBE-eligible areas (2.8pp reduction, significant at the 1% level). Column (12) again reveals that the effect is driven by changes for girls. Here, the reduction is 3.9pp compared to the control group.[65] Additional analysis of the treatment arms reveals that negative effects for girls seem to be strongest in the CPMF intervention areas (Table A.58). While the effects are negative, they are small in size relative to the average share reached on the WHO-5 scale in the control areas, which is reportedly 70.6% in the full sample.[66]

The results show a worsening of self-reported girls' wellbeing in the aftermath of the program, as compared to the control group. Note again, the variable refers to the state in the last two weeks, i.e., long after the empowerment program ended. Given that the evaluation came about a year after the program ended, it does not make any statements about the timing of these changes in wellbeing. It could be a direct effect of the program. It could also be that wellbeing levels have improved during program implementation, yet may have fallen thereafter due to girls' feeling of being left behind after the program ended. Overall, it is an important finding in itself needed further research on changes in wellbeing during program implementation (e.g., through the implementation of midline surveys). This would allow to chose the appropriate strategy of either to adjust the program during the implementation phase (if the effects are indeed negative) or to prolong the program if the effects on wellbeing are negative due to its ending. Further investigations reveal that the observed reductions in wellbeing are driven by observations in Sindh (Table A.51).[67]

> One major criticism voiced by KIs concerned the short time span of the intervention; achieving long-term effectiveness would require a longer intervention and it remains unclear what will happen after the project has ended.

---

[61]The individual variables are first converted to binary variables, which are coded as 1 if the adolescent answered yes, i.e., if they feel confident and expressed their opinion in their community, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of situations in which adolescents feel confident in their community (max. three). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by three).

[62]The overall IALP effect for the full sample is an increase by 1.4pp, with a p-value of 0.032. The heterogeneous analysis reveals that the increase was larger for individuals with high self-reported COVID-19 exposure compared to individuals with low exposure (Table A.51). The difference is not significant, however.

[63]In a first step, the variables are recoded such that higher values indicate a better state (i.e., so that 5 corresponds to "all the time" and 0 to "at no time"). Thereafter, the answers across the five questions are summarized into a score variable (25 being the maximum). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by 25.

[64]Note that the WHO-5 scale was randomized, i.e., not all adolescents and parents were asked these questions. For a brief explanation of the randomization of scales see Box II.B.1 above.

[65]The overall Child Protection effect for the full sample is a negative impact of 0.7pp, with a p-value of 0.012. The overall IALP effect is a negative impact by 1.3pp, with a p-value of 0.001.

[66]Alternative definitions of the scale are tested, simply using the standard definition of 25×4 being the maximum value, which generates comparable results.

[67]Moreover, the analysis indicates that the reduction is larger for individuals with high COVID-19 exposure compared to individuals with low exposure, as well as for individuals living in poor compared to individuals living in very poor households. Both differences between the two groups are not significant, however.

The program seems to have reduced the acceptance of using corporal punishment on daughters when they discuss their marriage timing, especially among male respondents. Part IV of Table 3.16 reports program effects on the prevalence of violence against children, which is captured by asking about the acceptance of using corporal punishment on children in specific situations as well as by asking adolescents about their actual experiences with corporal punishment. To do so, parents and adolescents were asked to indicate to what extent they agree on a scale from 1 (fully agree) to 4 (fully disagree) with the parental usage of corporal punishment of daughters and sons in certain situations.[68] The indicator in columns (1) to (3) captures that respondents fully or somewhat agree with the parental usage of corporal punishment of daughters when they talk to their parents about the timing of their marriage. On average, 40.2% of respondents report agreement with this statement in control villages. In program areas, the results indicate a decrease in the prevalence for the full sample by 6.4pp in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas (significant at the 5% level).

{ The group of "male respondents" refers to adolescent boys or fathers of an adolescent. }

Moreover, column (2) reveals that the effect is driven by changes for boys. Here, the reduction is 14.3pp compared to the control group. The average prevalence of agreement with the statement in the control areas exhibits significant gender differences: For boys it is reportedly 46.0%, while for girls it much lower at only 34.0%.[69] An additional analysis of the treatment arms which investigate different targeting strategies reveals that negative effect for boys in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas seems to be strongest in the CPMF intervention areas, i.e., where both genders were jointly targeted (Table A.58).

Self-reported experiences with corporal punishment of adolescents decreased in selected program areas. To capture actual prevalence of corporal punishment, adolescents were asked whether they experience corporal or physical punishment as a disciplinary measure at specific places, such as at home, at school, or at a learning center.[70] The summary variable captures, out of the seven places, the share of places at which adolescents experience corporal punishment.[71] The results indicate a decrease in the prevalence for the full sample by 5.1pp in areas where only the NFBE-program was implemented (significant at the 5% level). This effect is relatively large compared to the average share in the control areas, which is reportedly 18.9% for the full sample. The heterogeneous analysis reveals that the negative effect is driven by observations in Sindh (Table A.52).[72] Altogether, it can be constituted that the acceptance of and the self-reported experiences with corporal punishment changes in selected program areas, while the overall impact of the IALP remains statistically insignificant.

The IALP program has no impact on adolescents' overall feeling of control about their lives. Part V of Table 3.16 considers program effects on whether adolescents are taking action and their degree of self-efficacy as they become "agents" of their lives and fates. In order to see whether adolescents feel they can influence their fate, a short locus of control scale is applied.[73] Adolescents and parents were asked to what extent, on a scale from 1 (not at all) to 5 (completely), the following four statements apply to them (or their child): "I'm my own boss", "Whether at work or in my private life: What I do is mainly determined by others", "Bad luck often gets in the way of my plans", and "If I work hard, I will succeed".[74] By summing up their responses, a summary score is constructed which indicates adolescents' overall feeling of control, with 20 indicating a very strong feeling of being in control. Overall, the results indicate no impacts of the program or its separate components on this indicator.

The program has no impact on the number of actions taken to promote and ensure the adherence to adolescent rights.

---

[68]Note that the acceptance of using corporal punishment on daughters in specific situations was randomized, i.e., not all parents and adolescents were asked these questions. For a brief explanation of the randomization of scales see Box II.B.1.

[69]The heterogeneous analysis reveals that the negative effect is larger for individuals with low self-reported COVID-19 exposure compared to individuals with high exposure (Table A.52).The difference is not significant, however.

[70]The individual variables are first converted to binary variables, which are coded as 1 if the adolescent answered yes, i.e., if they experience corporal punishment at that place, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of places at which adolescents experience corporal punishment (max. seven). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by seven).

[71]Note that the experience of corporal punishment scale was randomized, i.e., not all adolescents were asked these questions. For a brief explanation of the randomization of scales see Box II.B.1.

[72]Moreover, the analysis indicates that the negative effect is larger for individuals in very poor households compared to individuals in poor households, as well as for individuals with high self-reported COVID-19 exposure compared to individuals with low exposure. These differences are not significant, however.

[73]Originally developed by Rotter (1966) and shortened by Lumpkin (1985) and Kovaleva (2012).

[74]Note that the short locus of control scale was randomized, i.e., not all parents and adolescents were asked these questions. For a brief explanation of the randomization of scales see Box II.B.1.

Finally, an index for actions taken to promote and ensure the adherence to adolescent rights is constructed. To this extent, respondents (adolescents, parents, community leaders, and Nikah Khwans) were asked whether, in the last year, they reported a case of sexual and/ or physical violence to someone under the age of 18 or a case of someone performing a child marriage to a religious or legal authority with the intent to initiate legal action or the filing of a case. In addition, they were asked whether, over the last four years, adolescents in their community have gotten together to request local government officials or political leaders for anything benefiting adolescents (age 10 to 19).[75] The summary variable indices, out of the five actions, the share of actions taken by respondents to ensure adherence to adolescent rights in the last year on a scale of 0 to 1. Overall, the results indicate no impacts of the program or its separate components on this index.

Evidence from Qualitative Interviews   *Role of Community in Empowering Adolescents and Creating an Enabling Environment* To implement activities related to reducing child marriage and teenage pregnancies as well as increasing school attendance, **engaging** with and earning the trust of the communities was a prerequisite. One KI suggested that community support was particularly important during the implementation of the IALP program, as it targeted a vulnerable group (adolescents) and addressed sensitive topics (child marriage and teenage pregnancy). Interviewees stated that being able to use the community-based organizational structure of the IP RSPN in Sindh greatly facilitated activities.[76] Generally, it was reported that the IPs provided information on the project to groups of parents and community members before engaging adolescents, especially to ensure girls' attendance in the activities as they were more likely to face cultural or religious obstacles when trying to attend. An interview partner mentioned the difficulty of convincing parents of the benefit that the trainings can bring, due to the program activities consisting mostly of intangible interventions the impact of which might be less visible to the community in comparison to e.g., infrastructure projects. When conveying the aim and content of the activities to the community, one KI recounted that it was helpful to stress that the activities would not contravene the beneficiaries' religion and, instead, are valuable for their children's health. Interview partners recounted gender differences in the willingness of parents to allow adolescents to participate: On the one hand, it was reported that mothers were supportive of participation of their children from the beginning. Another interviewee perceived that it was crucial to also engage men, and thus facilitate participation of women and girls. A third KI mentioned that fathers were easily motivated to cooperate whereas brothers were reluctant to let their adolescent sisters participate.

> "Having a community structure in place due to the Rural Support Programme was very helpful because they have women organizations who work with communities (...)." (Key Informant No.3)
>
> "It is very interesting because when we told the community about child marriages, the parents were reluctant, but when they were told that it is for the sake of health of their children and we are not going against the religion, it impacted positively (...)." (Key Informant No. 13)
>
> "We utilized services of men so they would provide better opportunities to their women to participate in such platforms and raise voice for themselves. Usually, in our society men are dominant and overshadow every role that must be played by a woman, so we engaged men and convinced them to include women as well, so they could get better chances of participation." (Key Informant No.4)

Adolescent Champions and the Adolescent Empowerment Kit Interviewees stated that adolescent champions in Punjab were chosen based on their age, communicative skills, and willingness to participate. According to KIs, specific efforts were taken to include minorities. For instance, it was reported that the IP strived to include Hindus as adolescent champions in Ghotki. On the other hand, a KI brought up that some areas have few people belonging to any minority. An official of the health department, Punjab, informed us that adolescent champions were selected in Punjab from among children who had volunteered. There was an attempt to create a diverse group where different ethnic groups, children with disabilities and from the Christian community were all included (Key Informant no. 12). Adolescent champions were tasked with the role of *change agents* and were meant to trickle-down their newly achieved knowledge on life skills to peers through training sessions. Interviewees' assessments of the Toolkit itself were mixed: Some described it

---

[75]The individual variables are first converted to binary variables, which are coded as 1 if the respondent answered yes, i.e., if the respondent took action, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of actions taken by respondents to ensure adherence to adolescent rights in the last year (max. five). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by five).

[76]RSPN maintains a mobilization structure at neighborhood, village, and union council level for rural communities. Strengthening community-based organizations, especially when taking on sensitive topics, was also included as a lesson learnt from Pakistan in the 2019 Annual Progress Report of UNICEF ROSA (UNICEF 2019).

Table 3.16: Adolescent Empowerment

	(I)			(II)						(III)		
	Awareness Child Rights			Confidence						Wellbeing		
	reached on CR topics			index for A confidence in doing activities themself			index for confidence of A in comm.			index for WHO 5 items wellbeing		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
only NFBE	-0.042	0.007	-0.058	-0.005	0.027	-0.039	0.036*	0.055	0.014	-0.018	0.004	-0.037*
	(0.074)	(0.069)	(0.097)	(0.032)	(0.033)	(0.049)	(0.020)	(0.034)	(0.015)	(0.013)	(0.022)	(0.019)
CP and NFBE	0.053	0.140**	-0.005	-0.030	-0.047	-0.027	0.027	0.044**	0.007	-0.043***	-0.024	-0.066***
	(0.064)	(0.063)	(0.081)	(0.029)	(0.037)	(0.039)	(0.017)	(0.021)	(0.019)	(0.012)	(0.020)	(0.018)
CP, NFBE pool	-0.022	0.063	-0.080	-0.034	-0.043	-0.033	0.025*	0.037*	0.011	-0.028***	-0.017	-0.039***
	(0.062)	(0.063)	(0.075)	(0.023)	(0.033)	(0.038)	(0.015)	(0.022)	(0.017)	(0.010)	(0.019)	(0.014)
CP, not NFBE pool	-0.022	-0.011	-0.028	0.020	0.012	0.010	0.005	0.002	0.008	-0.002	0.006	-0.009
	(0.032)	(0.040)	(0.064)	(0.017)	(0.023)	(0.023)	(0.011)	(0.015)	(0.014)	(0.007)	(0.011)	(0.008)
Control mean	0.143	0.122	0.166	0.250	0.350	0.148	0.181	0.236	0.130	0.706	0.709	0.703
SD	0.349	0.322	0.374	0.299	0.326	0.228	0.298	0.324	0.259	0.182	0.188	0.177
CP overall*	-0.022	0.005	-0.039	0.008	-0.000	0.001	0.009	0.009	0.009	-0.007	0.001	-0.016
CP p-value	0.522	0.462	0.279	0.594	0.404	0.605	0.105	0.153	0.380	0.012	0.591	0.002
IALP overall**	-0.015	0.022	-0.036	0.002	-0.004	-0.006	0.014	0.017	0.009	-0.013	-0.002	-0.024
IALP p-value	0.864	0.281	0.492	0.486	0.553	0.453	0.032	0.030	0.386	0.001	0.569	0.000
N	1,355	682	673	2,724	1,354	1,370	7,201	3,503	3,698	3,641	1,768	1,873
RVs	355	298	283	640	579	583	648	648	648	647	627	628
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

	(IV)						(V)					
	Corporal Punishment						Self-Efficacy					
	agr. to punish daughter (dau.) who discuss marriage (mar.) timing with P			index for experience of corporal (corp.) punishment (pn.) of children			score: short locus of control scale (4-l./20-h.)			index for actions taken to ensure adherence to adolescent right (AR) in the last year (i.l.y.)		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
only NFBE	0.030	-0.003	0.084	-0.051**	-0.047*	-0.054*	0.200	0.261	0.173	-0.007	0.001	-0.015*
	(0.045)	(0.066)	(0.058)	(0.021)	(0.027)	(0.029)	(0.272)	(0.310)	(0.317)	(0.007)	(0.011)	(0.008)
CP and NFBE	-0.004	0.029	-0.018	-0.015	-0.019	-0.009	0.104	0.131	0.220	-0.006	-0.005	-0.009
	(0.047)	(0.073)	(0.059)	(0.021)	(0.026)	(0.027)	(0.227)	(0.232)	(0.318)	(0.007)	(0.008)	(0.007)
CP, NFBE pool	0.065	0.061	0.077	-0.012	-0.005	-0.013	0.194	0.099	0.299	-0.006	-0.002	-0.012
	(0.048)	(0.072)	(0.061)	(0.020)	(0.028)	(0.024)	(0.218)	(0.208)	(0.334)	(0.007)	(0.009)	(0.008)
CP, not NFBE pool	-0.064**	-0.143***	0.010	0.002	-0.005	0.011	0.063	0.229	-0.110	-0.003	-0.004	-0.003
	(0.028)	(0.039)	(0.036)	(0.010)	(0.011)	(0.014)	(0.107)	(0.164)	(0.130)	(0.003)	(0.004)	(0.003)
Control mean	0.402	0.460	0.340	0.189	0.214	0.165	12.947	13.112	12.787	0.024	0.028	0.020
SD	0.486	0.490	0.474	0.236	0.243	0.225	2.294	2.255	2.325	0.075	0.080	0.070
CP overall*	-0.036	-0.099	0.024	-0.001	-0.005	0.006	0.091	0.201	-0.022	-0.004	-0.003	-0.005
CP p-value	0.984	0.275	0.235	0.667	0.747	0.962	0.309	0.234	0.606	0.209	0.561	0.081
IALP overall**	-0.027	-0.075	0.024	-0.007	-0.010	-0.001	0.102	0.197	0.024	-0.004	-0.003	-0.006
IALP p-value	0.826	0.760	0.339	0.167	0.283	0.367	0.372	0.256	0.515	0.261	0.712	0.075
N	2,350	1,147	1,203	4,793	2,363	2,430	3,161	1,519	1,642	6,821	3,339	3,482
RVs	632	567	565	648	647	647	642	605	611	648	648	648
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

▸ *Notes*: Table 3.16 displays the program effects on empowerment-related indicators. Effects from separate OLS estimations are presented.
▸ Sample: Endline survey. Full sample. Note that missing observations are due to non-response (refused to answer or does not know).
▸ Program effects on the following indicators are presented: Reached on CRs topics, index for As confidence in doing activities themself, index for confidence of adolescent in community, index for World Health Organization (WHO) 5 items wellbeing, agr. to punish dau.s who discuss marriage timing with Ps, index for experience of corp. pn. of children, score: short locus of control scale (4-l./20-h.), and index for actions taken to ensure adherence to ARs i.l.y.. "Reached on CRs topics" is generated by asking respondents (adolescents, parents, community leaders, and Nikah Khwans) to indicate what the topics covered by the Child Protection-program activity they attended were (i.e., conditional on participating in Child Protection-program events). The indicator captures that respondents mentioned "topics on child rights and entitlements" as topics covered by the Child Protection-event they attended. "Index for As confidence in doing activities themself" is generated by asking adolescents to indicate on a scale from 1 (not at all confident) to 7 (very confident) how confident they are in performing a number of activities by themselves. The individual variables are first converted to binary variables, which are coded as 1 if the adolescent answered being confident (=6) or very confident (=7), and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of activities (very) confident about performing (max. six). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of questions included in the summary variable (i.e, by six). The summary variable indices, out of the six activities, the share of activities the adolescent feels (very) confident about performing on a scale from 0 to 1. "Index for confidence of adolescent in community" is constructed by asking adolescents whether they express and feel confident expressing community needs to local government officials, and whether, over the last four years, adolescents in their community have gotten together to request local government officials or political leaders for anything benefiting adolescents. The individual variables are first converted to binary variables, which are coded as 1 if the adolescent answered yes, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of situations in which adolescents felt confident in their community (max. three). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by three). The summary variable indices, out of the three situations, the share of questions to which the answer was yes, i.e., the share of situations in which adolescents feel confident in their interactions with key stakeholders. "Index for WHO 5 items wellbeing" is constructed by asking adolescents and their parents to indicate how frequently on a scale from 1 (all the time) to 6 (at no time) they (or their child) felt certain positive emotions over the last two weeks. In a first step, the answers are recoded such that higher values indicate a better state (i.e., so that 5 corresponds to "all the time" and 0 to "at no time"). Thereafter, the answers across the five questions are summarized into a score variable (25 being the maximum). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by 25. The summary variable indices the share reached on the WHO-5 scale, with 1 representing the best imaginable wellbeing. "Agrees to punish dau.s who discuss marriage timing with Ps" is generated by asking parents and adolescents to indicate to what extent they agree on a scale from 1 (fully agree) to 4 (fully disagree) with the parental usage of corporal punishment of daughters when they talk to their parents about the timing of their marriage. The indicator captures that respondents fully or somewhat agree with the parental usage of corporal punishment in this situation. "Index for experience of corp. pn. of children" is captured by asking adolescents whether they experience corporal or physical punishment as a disciplinary measure at specific places, such as at home, at school, or at a learning center. The individual variables are first converted to binary variables, which are coded as 1 if the adolescent answered yes, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of places at which adolescents experience corporal punishment (max. seven). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by seven). The summary variable indices, out of the seven places, the share of places at which adolescents experience corporal punishment. "Score: short locus of control scale (4-l./20-h.)" is based on the scale originally developed by Rotter (1966) and shortened by Lumpkin (1985) and Kovaleva (2012). Adolescents and parents were asked to what extent, on a scale from 1 (not at all) to 5 (completely), the following four statements apply to them (or their child): "I'm my own boss", "Whether at work or in my private life: What I do is mainly determined by others", "Bad luck often gets in the way of my plans", and "If I work hard, I will succeed". By summing up their responses, a summary score is constructed which indicates adolescents' overall feeling of control, with 20 indicating a very strong feeling of being in control. "Index for actions taken to ensure adherence to ARs i.l.y." is constructed by asking respondents (adolescents, parents, community leaders, and Nikah Khwans) whether, in the last year, they reported a case of sexual and/ or physical violence to someone under the age of 18 or a case of someone performing a child marriage to a religious or legal authority with the intent to initiate legal action or the filing of a case. In addition, they were asked whether, over the last four years, adolescents in their community have gotten together to request local government officials or political leaders for anything benefiting adolescents (age 10 to 19). The individual variables are first converted to binary variables, which are coded as 1 if the respondent answered yes, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of actions taken by respondents to ensure adherence to adolescent rights in the last year (max. five). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by five). The summary variable indices, out of the five actions, the share of actions taken by respondents to ensure adherence to adolescent rights in the last year on a scale of 0 to 1. Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
▸ The equation 2.1 is estimated with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.
▸ Related table(s): Tables A.51, A.52, and A.58.

as being interactive and inclusive of illiterate adolescents thanks to pictorial material. Another KI mentioned that the Toolkit enabled disadvantaged adolescents to voice their opinion. Others criticized it as being too difficult to teach to adolescent champions and requiring a certain level of education which came at the expense of inclusion. Several KIs reported that refresher trainings with adolescent champions were conducted in Sindh, as some adolescent champions had reportedly not grasped the full range of the Toolkit that they should then transmit to peers. RSPN assisted with the selection of adolescent participants for the peer-to-peer sessions provided by adolescent champions using data on household financial support and social protection programs. During the selection process, the gender of adolescents was reportedly taken into consideration in accordance with the treatment arm to which the village had been assigned to.[77] During the peer-to-peer sessions, topics such as creating confidence, discussing one's own identity, empathy, strengths and weaknesses were covered. One interviewee perceived that informal exchange with peers facilitated learning among adolescents.

> "The most interesting thing for me was providing chance to the boys and girls, who have never got a chance to come forward. They got the opportunity to come forward and share their views that this is happening in society and this we want. So, that was the power through that Toolkit." (Key Informant No.10)
>
> "They [the adolescents] learned a lot of new things because when a person talks with their age-fellows they do informal talks and learn more." (Key Informant No.12)

Assessments of Improvement in Adolescent Empowerment by Key Informants A variety of stakeholders perceived an increase in confidence as well as in participation among the adolescents and noted that the workshops empowered them to speak up for their rights.[78] Some of the adolescent champions reportedly joined the staff of IPs in Sindh and one former adolescent champion even became an Adolescents Program Officer. One interviewee had the impression that adolescent champions even displayed "too much confidence" (Key Informant No.10) after having taken part in the Intergenerational Dialogues (IGDs). A few KIs felt that parents had grown more supportive of their adolescent children and had started talking about issues concerning them, thus indicating positive steps towards creating an enabling environment for adolescents. Some of the parents reportedly committed to sending their children to school. At the same time, there were questions by parents on the perspectives after the end of the NFBE and demands for further support having perceived a change in their children.

> "Their confidence was amazing. The same girls and boys at the start of the training were reluctant to speak out. But after the training, their level of participation and training their peers was a different scenario. They could speak much more confidently (...)." (Key Informant No.3)
>
> "I also observed through many capacity building trainings the community members, the parents, they were also talking about internally feeling that the adolescents' issues were the main issues (...) And that they need to resolve these issues." (Key Informant No.1)

Differences in implementation between Punjab and Sindh In Punjab, the organization which was originally implementing the project (CYAAD), was dismissed for poor performance after about an year. The implementation was then carried out by Bunyaad Community Literacy Council, who, however, only had 16 months to implement the programme which was originally planned for 3 years.(Key Informant no. 11). The shortage of time was felt to be a major challenge, especially as changes in mindset would be expected to take much longer (Key Informant no. 12). The non-performance and dismissal of the initial implementing partner, with additional time taken to hire the partner who eventually did the work were major challenges (Key Informant no. 1). This may have contributed to the poorer performance of the adolescent champions component in Punjab.

---

[77]The three CP treatment arms concern the extent to which adolescents of a specific gender were mobilized as part of the CP intervention. Mobilization focused either on male or on female adolescents, or both genders were targeted equally strongly. A village's assignment to a CP treatment arm seems to have been adhered to during implementation (UNICEF 2019).

[78]This is in line with findings from the literature that programs in which adolescents can play a very active role lead to empowerment and a high sense of ownership (Larson, Walker, and Pearce 2005).

II.B.2   Communication

> **Key Findings on Communication Outcomes**
>
> One of the key outputs according to the program's theory of change is an encouragement to exchange and interact. Adolescents and key local stakeholders were meant to participate in groups, share information amongst peers, and connect with each other. To this extent, participatory communication platforms were created and intergenerational dialogues were organized. In the following paragraphs, in how far communication between peers and across genders changed due to the program is tested.
>
>   • Overall, the IALP program has no impact on changes in the incidence of intergenerational dialogue. However, the program increases the share of situations in which parents consider their children's opinions as reported by adolescent boys and fathers;
>   • The program has a mostly positive impact on adolescents' communication skills: The results indicate an increase of the average score on the ICC scale, especially in CPF intervention areas;
>   • Also the incidence of intragenerational dialogues among adolescents increases, especially for females;
>   • Single indices reveal that the positive effect is mainly driven by significant increases in the prevalence of adolescents talking to friends regarding their fears and emotions, and talking to their brother about their future marriage.

The evaluation continues to study whether the IALP program facilitated the dialogue among adolescents as well as between parents and their adolescent offspring as originally envisioned. In other words, whether it made cross-generational discussions more likely, especially on young adult- and adolescent-related topics. Table 3.17 displays the effects from a set of OLS regression results analyzing information on several communication-related variables. Information on communication is captured in the main survey by asking adolescents and parents about their (or their adolescent child's) activities and competences.

The IALP program has no impact on the overall incidence of intergenerational dialogue. Part I of the table considers two indicators of intergenerational dialogue. To capture the overall share of intergenerational dialogue adolescents engaged in, adolescents and parents were asked a series of questions. Conditional on having ever participated in any of the Child Protection-events (training, workshop, or group session, community meeting, or theater session) or having ever attending a non-formal basic education center, adolescents and parents were asked whether they (or their child) usually discussed the topics covered in the classes with others.[79] Moreover, they were asked whether they (or their child) did any of the following activities in the last three months: Contacted their sibling's teachers regarding their education, convinced or tried to convince their parents to send their brothers or their sisters to school (separate questions), negotiated with their parents on behalf of their sisters so that they can get skills training, talk to their parents regarding their fears and emotions, and talked to their parents about their future marriage.[80] The summary variable indices, out of the seven activities, the share of activities at which adolescents engage in intergenerational dialogue on a scale from 0 to 1. Overall, the results indicate no impacts of the program on this indicator, neither for male nor female respondents, as compared to the control group. In order to capture how effects differ for certain groups, an extended heterogeneous analysis for the following four groups is conducted: Province (Sindh vs. Punjab), poverty (very poor vs. poor), COVID-19 exposure (high vs. low), and adolescent boys vs. adolescent girls. This analysis reveals no detectable differences between the groups (Table A.53).

The program increased the share of situations in which parents consider their child's opinion as reported by fathers and boys. In columns (4) to (6) of Table 3.17, program effects on the index for situations in which parents consider their child's opinion are considered. Adolescent and parents were asked in which situations their parents (or they) take their (or their child's) opinion into consideration, on a scale from 1 (fully considered or accepted) to 3 (not taken into consideration or accepted). These included, for example, the decision to go to the local market alone, the decision whether to go to the doctor or medical facility, the decision on whether to speak out in public, etc.[81] The summary

---

[79]Adolescents were asked "After the participation, did you usually discuss the topics covered in the classes with others?", while parents were asked "After the participation, did your child usually discuss the topics covered in the classes with others?"

[80]The individual variables are first converted to binary variables, which are coded as 1 if the adolescent answered yes, i.e., if they engaged in the dialogue, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of situations in which respondents engaged in intergenerational dialogue (max. seven). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by seven).

[81]In a first step, the seven individual variables are converted to binary variables, which are coded as 1 if the respondent answered full (=1) or somewhat (=2) consideration, and 0 otherwise. In a second step, a score variable is generated which indicates the total number of situations in which parents (fully or somewhat) consider their child's opinion (max. seven). Thereafter, a summary variable is constructed by rescaling the score to range

variable indices, out of the seven situations, the share of situations in which parents (fully or somewhat) consider their child's opinion on a scale from 0 to 1. The results indicate an increase by 1.5pp amongst boys and fathers in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas as compared to the control group (significant at the 5% level). In control areas, the average share is reportedly 26.7% for males. The analysis also reveals a small negative impact of 1.4pp for adolescent girls only in areas where both program components were implemented, i.e., adolescent girls in these areas report that their parents consider their opinion in less situations (Table A.53). Additional analysis of the treatment arms reveals that positive effect for boys in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas seems to be mostly prevalent in the CPM intervention areas (Table A.59).

The program had positive impacts on adolescents' communication skills, especially in CPF intervention areas. Part II of Table 3.17 focuses on program effects on adolescents' communication skills. An important scale thereby is an established psychological scale on communication by Rubin and Martin (1994): The ICC scale. Adolescents and parents were asked to indicate whether (ten) statements such as "I allow friends to see who I really am" and "I am comfortable in social situations" fit them (or their child) on a scale from 1 (never) to 5 (always).[82] Using this scale, different domains can be generated: Identity and self-esteem (indicator for social relaxation), empathy and respect (using the indicators for empathy, assertiveness, and altercentrism), communication and expression (with indicators for self-disclosure for interaction management, expressiveness, immediacy, supportiveness) and finally coping with stress and managing emotions (using the indicator for environmental control). Taking their answers for the ten domains together, the interpersonal communication competence score indicates adolescents' overall communication competence on a scale from 10 to 50, with 50 representing full competence across 10 dimensions. The results indicate no impacts of the program or its separate components on this indicator for the full sample. However, the analysis reveals an increase of the score by 1.6 for females in areas where only the NFBE component was implemented, which is 29.4% of a standard deviation in the control group (significant at the 1% level). Moreover, an effect for girls of a similar size is observed in areas where Child Protection interventions were implemented in the NFBE-eligible areas (1.7 increase, significant at the 1% level). For males, the findings are mixed. In areas where only the NFBE component was implemented, the results indicate a reduction of the score by 1.5 for males (significant at the 5% level). Whereas, in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas, the analysis reveals an increase of the score by 1.1 for males (significant at the 1% level). The overall Child Protection and IALP effects for males remain insignificant, however, whereas the overall IALP effect for girls is an increase of the score by 0.7 (with a p-value of 0.018).[83] In control areas, the average score on the ICC-scale is reportedly 28.7 for girls. A heterogeneous analysis confirms the positive findings for females: The program significantly increases the ICC score for adolescent girls, while adolescent boys are not impacted (Table A.53). An additional analysis of the different program activities reveals that both - the positive effect for girls in areas where Child Protection interventions were implemented in the NFBE-eligible areas and the positive effect for males in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas - seem to be mostly prevalent in the CPF intervention areas (Table A.59).

"We utilized services of men so they would provide better opportunities to their women to participate in such platforms and raise voice for themselves." (Key Informant No.4)

The incidence of intragenerational dialogues among adolescents increased, especially for girls. Finally, in Part III of Table 3.17 changes in the intragenerational dialogue among adolescents (peers) are considered. To this extent, adolescents and parents were asked whether, in the last three months, they (or their child) talked to friends regarding their fears and emotions, or to their brothers and sisters regarding their future marriage. In addition, adolescents and parents were asked whether they (or their child), after the participation in Child Protection or NFBE activities, usually discussed the topics covered in the classes with peers, such as adolescent family members, fiends, or acquaintances.[84][85] The summary variable indices, out of the four situations, the share of situations in which

---

between 0 and 1, attained by dividing the score by the number of total number of situations included in the summary variable (i.e, by seven).

[82]Note that the ICC scale was randomized, i.e., not all parents and adolescents were asked these questions. For a brief explanation of the randomization of scales see Box II.B.1.

[83]The overall Child Protection effect for girls is an increase of the score by 0.5, with a p-value of 0.034.

[84]Adolescents were first asked "After the participation, did you usually discuss the topics covered in the classes with others?", and if yes, "Who did you talk to?". Parents were asked "After the participation, did your child usually discuss the topics covered in the classes with others?", and if yes, "Who did your child talk to?".

[85]Taking respondents' answers to the four questions together, a score variable is generated which indicates the total number of situations in which respondents engaged in intragenerational dialogue (max. four). Thereafter, a summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of situations included in the summary variable (i.e, by four).

adolescents engage in dialogue among peers. The results indicate an increase in the share by 4.1pp for the full sample in areas where Child Protection and educational interventions were jointly implemented (significant at the 5% level). Moreover, column (12) reveals that the effect is driven by changes for females. Here, the increase is 7.0pp compared to the control group (significant at the 1% level).[86] In control areas, the average share is reportedly 10.5% for girls, and therewith slightly lower than for boys (13.4%). The heterogeneous analysis reveals that the positive effect is driven by observations in Sindh (Table A.53). Moreover, the positive effect is larger for individuals with high self-reported COVID-19 exposure compared to individuals with low exposure.[87] In other words, the positive program effects on dialogue amongst peers for females are stronger if they are reportedly more affected by COVID-19. Moreover, when looking at the effect among adolescents only, the results indicate a positive impact of 6.5pp for adolescent boys in areas where only the educational intervention was implemented (significant at the 5% level).

Separating by indicators, the results reveal that the positive effect is mainly driven by significant increases in the prevalence of adolescents talking to friends regarding their fears and emotions, and talking to their brother about their future marriage. The individual indicators are additionally presented separately in Figure 3.5. The interventions seem to have a positive impact on all indicators of intragenerational dialogue, with statistically significant increases (at the 10% level) in the indicators "talked to friends regarding one's fears and emotions" (share of those who do it in control group is 27%, the impact caused by the program is 2.8pp) and "talked to brother regarding one's future marriage" (average in control group is 14%, the impact caused by the program is 2.0pp). The figure additionally indicates program effects on two summary variables of intragenerational dialogue. The score variable indicates the total number of situations in which respondents engaged in intragenerational dialogue (max. four). The second summary variable, also referred to as index for intragenerational dialogue, indices, out of the four situations, the share of situations in which adolescents engage in dialogue among peers. The graph displays a positive impact on both summary variables.

---

[86]Effects of similar sizes can be observed in areas where only the NFBE-program was implemented, and in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas. The overall Child Protection effect for girls is an increase of 1.6pp, with a p-value of 0.016. The overall IALP effect for girls is an increase of 2.5pp, with a p-value of 0.003.

[87]The difference of 5.7pp between individuals with high and individuals with low COVID-19 exposure is significant at the 1% level.

Table 3.17: Communication

	(I) Intergenerational Dialogue						(II) Interpersonal Communication Competence			(III) Intragenerational Dialogue		
	index for intergen. dialogue			index for situations in which P consider child's opinion			score: ICC scale (10-l./50-h.)			index for intragenerational (intragen.) dialogue amongst A		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
only NFBE	0.020	0.035	0.013	0.009	0.013	0.000	0.066	-1.493**	1.638***	0.039**	0.029	0.052**
	(0.021)	(0.027)	(0.020)	(0.009)	(0.010)	(0.018)	(0.334)	(0.741)	(0.560)	(0.017)	(0.028)	(0.022)
CP and NFBE	0.006	0.017	0.001	0.001	0.004	-0.002	-0.192	-1.151	0.868	0.040**	0.014	0.070***
	(0.017)	(0.023)	(0.016)	(0.007)	(0.010)	(0.012)	(0.444)	(0.795)	(0.644)	(0.017)	(0.026)	(0.022)
CP, NFBE pool	-0.006	0.005	-0.012	-0.002	0.001	-0.005	0.377	-0.876	1.689***	0.030**	0.013	0.047**
	(0.014)	(0.020)	(0.014)	(0.007)	(0.010)	(0.012)	(0.479)	(0.835)	(0.607)	(0.014)	(0.024)	(0.020)
CP, not NFBE pool	-0.005	-0.008	-0.000	0.007	0.015**	0.002	0.475	1.118***	-0.092	0.002	-0.003	0.007
	(0.007)	(0.010)	(0.009)	(0.005)	(0.007)	(0.006)	(0.311)	(0.408)	(0.410)	(0.007)	(0.009)	(0.009)
Control mean	0.176	0.190	0.162	0.236	0.267	0.207	28.935	29.136	28.746	0.119	0.134	0.105
SD	0.210	0.216	0.203	0.143	0.139	0.141	5.520	5.354	5.570	0.205	0.220	0.187
CP overall*	-0.005	-0.005	-0.003	0.005	0.012	0.000	0.454	0.689	0.291	0.008	0.001	0.016
CP p-value	0.521	0.885	0.476	0.526	0.188	0.808	0.140	0.797	0.032	0.043	0.692	0.016
IALP overall**	-0.002	0.001	-0.001	0.005	0.011	0.000	0.342	0.282	0.473	0.015	0.005	0.025
IALP p-value	0.747	0.452	0.969	0.404	0.225	0.898	0.525	0.263	0.011	0.010	0.468	0.003
N	7,267	3,526	3,741	7,219	3,512	3,707	3,130	1,512	1,618	7,175	3,487	3,688
RVs	648	648	648	648	648	648	643	602	601	648	648	648
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

▸ *Notes*: Table 3.17 displays the program effects on communication-related indicators. Effects from separate OLS estimations are reported.
▸ Sample: Endline survey. Full sample.
▸ Program effects on the following indicators of intergenerational dialogue are considered under Part I: Index for intergen. dialogue and index for situations in which Ps consider child's opinion. Part II of the table reports program effects on the score on the ICC scale (10-l./50-h.). Finally, Part III considers program effects on the share of intragen. dialogue amongst As. "Index for intergen. dialogue" is generated by asking adolescent and parents a series of questions. Conditional on participating in any Child Protection or educational interventions, adolescents and parents were asked whether they (or their child) usually discussed the topics covered in the classes with others. Moreover, they were asked whether they (or their child) did any of the following activities in the last three months: Contacted their sibling's teachers regarding their education, convinced or tried to convince their parents to send their brothers or their sisters to school (separate questions), negotiated with their parents on behalf of their sisters so that they can get skills training, talk to their parents regarding their fears and emotions, and talked to their parents about their future marriage. The individual variables are first converted to binary variables, which are coded as 1 if the adolescent answered yes, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of situations in which respondents engaged in intergenerational dialogue (max. seven). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by seven). The summary variable indices, out of the seven activities, the share of activities at which adolescents engage in intergenerational dialogue on a scale from 0 to 1. "Index for situations in which Ps consider child's opinion" is generated by asking adolescents and parents in which situations their parents (or they) take their (or their child's) opinion into consideration, on a scale from 1 (fully considered or accepted) to 3 (not taken into consideration or accepted). In a first step, the seven individual variables are converted to binary variables, which are coded as 1 if the respondent answered full (=1) or somewhat (=2) consideration, and 0 otherwise. In a second step, a score variable is generated which indicates the total number of situations in which parents (fully or somewhat) consider their child's opinion (max. seven). Thereafter, a summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of situations included in the summary variable (i.e, by seven). The summary variable indices, out of the seven situations, the share of situations in which parents (fully or somewhat) consider their child's opinion on a scale from 0 to 1. "Score: ICC scale (10-l./50-h.)" is constructed by asking adolescents to indicate whether (ten) statements such as "I allow friends to see who I really am" and "I am comfortable in social situations" fit them on a scale from 1 (never) to 5 (always). Taking their answers for the ten domains together, the interpersonal communication competence score indicates adolescents' overall communication competence on a scale from 10 to 50, with 50 representing full competence across 10 dimensions. "Index for acintragen. dialogue among As" is captured by asking adolescents and parents whether, in the last three months, they (or their child) talked to friends regarding their fears and emotions, or to their brothers and sisters regarding their future marriage. In addition, adolescents and parents were asked whether they (or their child), after the participation in Child Protection or NFBE activities, usually discussed the topics covered in the classes with peers, such as adolescent family members, fiends, or acquaintances. Taking respondents' answers to the four questions together, a score variable is generated which indicates the total number of situations in which respondents engaged in intragenerational dialogue (max. four). Thereafter, a summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of situations included in the summary variable (i.e, by four). The summary variable indices, out of the four situations, the share of situations in which adolescents engaged in dialogue among peers. Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
▸ Equation 2.1 is estimated with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3+\frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 +\frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.
▸ Related table(s): Tables A.53 and A.59.

Figure 3.5: Intragenerational Dialogue
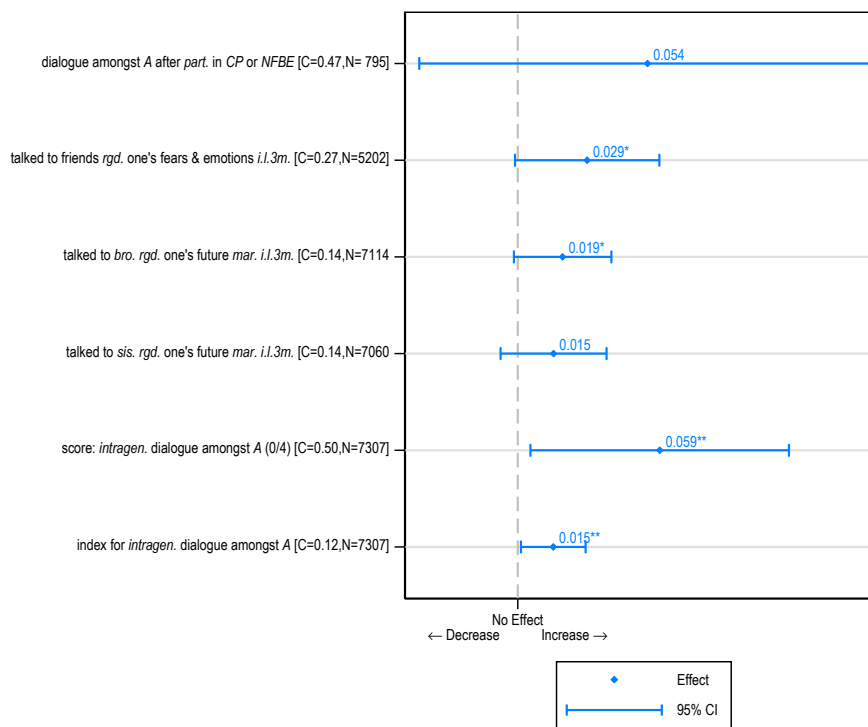
► *Notes*: This figure shows treatment effects capturing changes in intragenerational dialogue. In order to capture intragenerational dialogue amongst adolescents (peers), adolescents and parents were asked whether, in the last three months, they (or their child) talked to friends regarding their fears and emotions, or to their brothers and sisters regarding their future marriage. In addition, adolescents and parents were asked whether they (or their child), after the participation in Child Protection or NFBE activities, usually discussed the topics covered in the classes with peers, such as adolescent family members, fiends, or acquaintances. Taking respondents' answers to the four questions together, a score variable was generated which indicates the total number of situations in which respondents had engaged in intragenerational dialogue (max. four). Thereafter, a summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of situations included in the summary variable (i.e, by four). The summary variable "index for intragenerational dialogue" indices, out of the four situations, the share of situations in which adolescents engaged in dialogue among peers on a scale from 0 to 1.

► The treatment variable is assignment to any type of program activities.

► Sample: Endline survey.

► Significance levels are indicated by $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$. N refers to the number of respondents.

Evidence from Qualitative Interviews   *Communicating Sensitive Topics* Some interviewees had initially been concerned about how to convey content on sensitive topics such as child marriage and teenage pregnancy, fearing "retaliation (...) which could have a negative impact on [the] project" (Key Informant No.4). Yet, their concerns were quickly allayed as these issues tended to come up automatically in sessions with adolescents, who "have seen their friends and relatives in their community being married at an early age and bearing the consequences of teenage pregnancies" (Key Informant No.3). One KI recounted that an effort was made to communicate that child marriage and teenage pregnancy are detrimental to the development not only of girls, but also of boys. For adolescents who are already married, a session on reproductive health was perceived as being effective.

"When addressing the early child marriage issue, we told the same to boys as well that it is not good for them too. They are living a pretty simple life and as soon as they would get married, they will have children and look after their needs, medicines and food etc. and that how difficult it is to find a job. (...) But thankfully we were able to make them understand that early marriage is synonymous to instability economically. Because they suddenly have so many new responsibilities and, as you know, rural areas have joint family culture that is more

> difficult for both the children to manage." (Key Informant No.13)

Communication was also enhanced through a module targeting *KCIs* aimed at enabling them to identify violations of children's and adolescents' rights such as child marriage, abuse, or refusal of formal education. KCIs were to receive knowledge on relevant topics in trainings and subsequently spread their acquired learning and function as a "watchdog" (Key Informant No.10) in their communities. It was mentioned that RSPN contributed to the selection of KCIs through their community structures.

As part of a community-based media campaign, *street theater* performances were launched. While one KI indicated that a professional theater group was hired, another stated that the community, and especially adolescent champions and KCIs worked together on the organization. According to several interviewees, street theaters were mainly used to transmit the topics of children's rights, child marriage, abuse and education to the communities. Due to the plays being staged locally, mobility issues were not a concern. KIs reported that the interactive plays were well received by residents and surmised that they raised awareness as attendees could relate to the content. One interview partner recalled several adolescents joining or forming similar theater groups after attending the performances.

Regarding *Intergenerational Dialogues*, qualitative data showed a complex picture, especially in the real and perceived communication skills of girls. At various administrative levels, *IGDs* were held in an effort to bring adolescents and older generations together to exchange views and concerns on issues pertaining to adolescents.[88] It was planned that adolescent champions and KCIs would also share insight gained from previous training with IGD participants. Several interviewees reported that Government and UNICEF officials participated as well. According to one KI, issues that were discussed in the IGDs included child marriages, obstacles to girls' school attendance and the lack of schools. One interview partner stated that interaction with government officials also had the positive side effect of making people aware of the institutions that could be of help.

> [on IGDs] "I observed many girls were talking about the issue of having no education facilities or their parents not allowing them to go to school and I have also observed some very confident and blunt adolescent girls who were talking about the child marriage issues. Some adolescent girls who already got married they were also talking about the consequences and impact of child marriages on the adolescents. They were urging their parents and communities, that please do not go for early child marriages for their children. This is not good for the development of children and for the development of adolescents. It is good if your girls are getting education and at the appropriate age they could go for marriage." (Key Informant No.1)
>
> "People found out that there are a lot of institutions in the government and private sector that can help. The link which youth developed with these sectors was very good." (Key Informant No.12)

II.B.3   Social Norms

> **Key Findings on Social Norms Outcomes**
>
> With the underlying hypothesis that the need to improve adolescents lives will have to eventually address social norms and societal attitudes (including the need to eliminate stereotypical gender roles and discrimination, the acceptance of corporal punishment and harmful traditional practices (UNICEF South Asia 2019a), the IALP program laid out its social mobilization efforts and messages around the need to help shifting social norms on child marriage and other gender related norms and adolescent protection (UNICEF South Asia 2019b). Whether these goals were achieved by the program is tested in this section. Thus, in this section, the attention is dedicated to the measurement of shifts in attitudes and norms' shifts which may eventually translate to (further) changes in behavior.
>
> - The IALP program reduces agreement with the statement that it is OK for parents to marry sons before the age of 18. Especially female respondents decrease their levels of agreement. No such effect is detectable for girls' marriages before the age of 18 which are still acceptable to about every second respondent;
> - However, the results indicate a decrease in the acceptance to marry girls before the age of 16 in program areas;
> - Overall, with every second person accepting marriages of girls before the age of 18 and every third person before the age of 16,

[88]Interview partners mentioned IGDs being held at union council, tehsil, and district level.

> there is still much room for improvement and programming;
> - The results indicate no additional impacts of the program on the respondents' level of knowledge of child rights;
> - However, the program increases the prevalence of respondents being able to name at least one specific risk of child marriage;
> - The program increases also other respondents' progressive gender norms w.r.t. marriage and its timing. For instance, the agreement among male respondents with the statement that girls should finish secondary education before marriage is positively impacted;
> - While the program reduces the overall extent of egalitarian beliefs about gender norms among males, single indices reveal that the program increases the level of disagreement with two inegalitarian statements, namely that sons should be more educated so to look after their parents and that a good woman never questions her husband's opinions.

Finally, the IALP program was expected to change how stakeholders advocate and promote adolescents rights, and how adolescents are perceived among all community members. On average, it was hypothesized that stakeholders will be more favorable towards adolescents, be more likely to know and promote their rights and capabilities. Table 3.18 displays the effects from a set of OLS regression results analyzing changes in indicators which attempt to capture potential changes in social norms.

Part I of the table considers program effects on social norms related to child marriage. To capture the level of (dis)agreement with the practice of child marriage, parents and adolescents were asked to indicate to what extent they agree, on a scale from 1 (fully agree) to 4 (fully disagree), with the following three statements: "It is ok for parents to marry their sons before their 18th birthday" (asked in Punjab only), "It is ok for parents to marry their daughters before their 18th birthday", and "It is ok for parents to marry their daughters before their 16th birthday". For each of these statements, indicator variables are generated which indicate agreement (combines full (=1) and somewhat (=2) agreement) with the statement.

The IALP program reduced agreement with the statement that it is OK for parents to marry sons before the age of 18, especially among female respondents. In columns (1) to (3),"agrees: ok for parents to marry sons before 18" is considered as the dependent variable (Punjab only). The results indicate a reduction in the prevalence of agreement by 7.4pp in areas where the Child Protection and educational interventions were jointly implemented (significant at the 5% level). Moreover, column (3) reveals that this effect is mainly driven by changes for females. The difference is a reduction by 13.7pp compared to the control group (significant at the 1% level).[89] The average prevalence in the control areas is reportedly 44.3% in the full sample, i.e., roughly every second person in the sample accepts boys' marriages before the age of 18.

The program has no impact on the acceptance to marry daughters before the age of 18 for the full sample but increases acceptance among respondents in the Sindh.

> The IALP program reduces the acceptance of child marriage in general and of girls before 16 in particular, especially among male respondents.

In columns (4) to (6) agreement with the statement that it is OK for parents to marry daughters before the age of 18 is considered as the dependent variable. Roughly every second person in the sample accepts girls marriages before the age of 18 (control mean of 44.3%). The results indicate no impacts of the program or its separate components for this indicator. However, compared to the control group, the results indicate an increase in the prevalence by 10.2pp for individuals in Sindh who resided in CP areas outside of the NFBE-eligible pool (Table A.54).[90]

The analysis shows a decrease in the acceptance to marry girls before the age of 16 in program areas. So far, the results indicate that in program areas, the acceptance to marry boys before the age of 18 drops, while no such effect is detectable for girls' marriages. However, the evaluation shows a decrease in the acceptance to marry girls before the age of 16 in program areas as compared to control areas. In more detail, the results are reported in columns (7) to (10) of Table 3.18 where "agrees: ok for parents to marry daughters before 16" is considered as the dependent variable. The results capture a reduction in the prevalence of agreement for males by 8.0pp in areas where only the NFBE-program was implemented, and by 8.4% in areas where Child Protection interventions were implemented in the NFBE-eligible areas (both effects significant at the 5% level). In areas where Child Protection interventions were implemented in the non-NFBE-eligible areas, the results also indicate a reduction in 4.4% for the full sample. Moreover, column (9) reveals that this effect is mainly driven by changes for females. Here,

---

[89]The overall Child Protection effect for the full sample is a reduction in 4.6pp, with a p-value of 0.026. The overall IALP effects is a reduction in 4.7pp, with a p-value of 0.013.

[90]The difference between Sindh and Punjab is statistically significant.

the reduction is 5.3pp compared to the control group (significant at the 5% level).[91] To put these effects in relation, consider that the average acceptance of girls marriages before the age of 16 in control areas is 31.2% among males, thus the fact that program areas display on average 5.2pp lower acceptance among males is a change of about 16%. Further subgroup analyses reveal that the negative effect in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas is driven by observations in Punjab. It is also higher both among individuals living in poor compared to individuals living in very poor households, and among individuals with low compared to high COVID-19 exposure (Table A.54). When comparing adolescent boys and girls, the analysis captures a reduction of agreement among adolescent boys only. The evaluation constitutes that a change in acceptance of early child marriages for girls is possible, yet the threshold is still lower than the legal requirement and effects differ by provinces. To reach a higher level of rejection of marriages before the legal age of 18 (16 for women in Punjab), further efforts would be necessary. With every second person accepting marriages of girls before the age of 18 and every third person before the age of 16, there is much room for improvement and programming.

The program increased the prevalence of respondents being able to name at least one specific risk of child marriage. Why would respondents change their attitudes towards early marriages? In order to asses whether respondents know about the possible negative consequences of child marriage, adolescents, parents, and community leaders were first asked whether there are any risks or negative consequences of getting married as a child. Conditional on answering yes, they were asked to name the specific risks they know, such as "education is likely to stop" and "child-bearing is likely to start early". Based on their answers, a variable indicating whether respondents were able to correctly name at least one specific risk of child marriage is constructed.

The results indicate an increase in the prevalence of knowledge by 8.2pp for the full sample in areas where only the NFBE-program was implemented (significant at the 1% level). Additionally, columns (11) and (12) reveal that the effect seems to be driven by changes for both males and females. Here, the differences are an increase in 7.0pp and 10.0pp respectively, compared to the control group. The average prevalence in the control areas is reportedly 83.5% for the full sample. Furthermore, heterogeneous analyses reveal that the observed positive effect is again driven by observations in Sindh (Table A.54). Moreover, the results indicate that the positive effect is larger both for individuals in very poor households compared to individuals in poor households, and for individuals with high compared to low COVID-19 exposure.[92] The average prevalence of knowledge about the risks of CM in the control areas is reportedly 79.2% for very poor individuals and therewith lower than in poor areas where it is 90.1%.

> "(...) I think there is a change, at least the parents and communities are aware of the consequences of child marriage now. Though we cannot say the child marriages are completely stopped now (...), but there is improvement." (Key Informant No.1)

The program increased respondents' overall progressive gender norms with respect to marriage, especially for adolescent boys. Part II of Table 3.18 focuses on program effects on attitudes about gender roles and marriage in general. To capture respondents' overall attitudes about marriage, adolescents and parents were asked to indicate to what extent they agree, on a scale from 1 (fully agree) to 4 (fully disagree), with eight statements such as: "It is an adolescent girl's fault to be sexually harassed if she is not appropriately dressed", "Girls should be married as soon as possible to avoid dishonoring (to keep honor of the family)", and "Girls should wait until they have finished their secondary education to marry".[93] The summary variable indices, out of the eight questions, the share of questions at which respondents present progressive gender norms with respect to marriage, here also referred to as positive marriage-related attitudes, on a scale from 0 to 1. The results indicate an increase in the index for positive attitudes about marriage by 3.1pp in areas where only the educational component was implemented (significant at the 5% level). The average share is reportedly 44.8% in control areas, i.e., respondents present positive attitudes about marriage on average at nearly four out of the eight questions asked. Additionally, heterogeneous analyses

---

[91]The overall Child Protection effect for male respondents is a reduction of 4.7pp, with a p-value of 0.005. Similarly, the overall IALP effect for males is a negative impact of 5.2pp, with a p-value of 0.007.

[92]These differences are not significant, however.

[93]In a first step, indicator variables are generated for each statement, which are coded as 1 for either agreement (combines full (=1) and somewhat (=2) agreement) with positively phrased statements, e.g., "Girls should wait until they have finished their secondary education to marry", or disagreement (combines somewhat (=3) and full (=4) disagreement) for negatively phrased statements, e.g., "Girls should be married as soon as possible to avoid dishonoring", and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions to which respondents presented progressive gender norms with respect to marriage (max. eight). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by eight).

reveal that the positive effect is driven by observations in Sindh (Table A.54).[94] Furthermore, the analysis reveals that adolescent boys exhibit a stronger increase in their positive attitudes about marriage compared to adolescent girls.

The program increased agreement among adolescent boys and fathers with the statement that girls should finish secondary education before marriage. An example of one of the eight statements included in the index for positive attitudes about marriage can be found in columns (4) to (6). The variable "agrees: girls should finish secondary education before marriage" is coded as 1 to indicate adolescents and parents fully (=1) or somewhat (=2) agreeing with the statement that girls should wait until they have finished their secondary education to marry, and 0 otherwise. More than every second person in the sample believes that girls should finish secondary education before marriage (control mean of 56.6%). The results indicate an increase in the prevalence of agreement by 7.7% among boys and fathers in areas where only the educational component was implemented (significant at the 5% level). Further investigations reveal an increase in the prevalence by 8.5pp in areas where Child Protection interventions were implemented in the non-NFBE-eligible areas for observations in Sindh (Table A.55).[95] Finally, the information discussed (Part I and II of Table 3.18) is also graphically illustrated in Figure 3.6a below.

The results indicate no impacts of the program on respondents' knowledge of child rights. Part III of Table 3.18) considers program effects on knowledge of child rights. In order to find out whether knowledge on child-rights has improved in the treatment areas as compared to the control areas, adolescents, parents, and community leaders were asked to correctly identify the four rights which exist for children under the age of 18: A right to have birth registration, a right to education, a right to get information that is important to their health and wellbeing, and a right to be protected from violence and abuse.[96] The summary variable indices, out of the four rights, the share of child rights that respondents are able to correctly identify on a scale from 0 to 1. The results indicate no impacts of the program or its separate components for this indicator as compared to the control group. The average share of child rights that respondents are able to correctly identify in control areas is reportedly 48.4% for the full sample, i.e., on average, respondents recognize two out of the four existing child rights.

The program reduced the extent of egalitarian beliefs about gender norms among males. Finally, Part IV of the table focuses on program effects on prevailing norms about equal opportunities, especially for girls. To capture prevailing social norms, adolescents and parents were asked about their view on gender roles. A scale from Waszak et al. (2001) was employed which captures the role of women/ girls vs. men/ boys in the public and private sphere. Adolescents and parents were asked about their agreement, on a scale from 1 (disagree) to 4 (agree), with 14 statements, such as: "Mothers should be able to work outside the home after they have children if they want to", "It is important that sons receive more education than daughters", and "Women should leave politics to men".[97] The summary variable indices, out of the 14 items of the gender role models scale, the share of questions which are answered positively, i.e., at which respondents present egalitarian beliefs about gender norms. The results indicate a reduction in the share of questions at which respondents present egalitarian beliefs about gender norms by 3.0pp for males in areas where Child Protection interventions were implemented in the NFBE-eligible areas (significant at the 5% level). The average share in control areas is reportedly 21.5% for males. In other words, male respondents in control areas present egalitarian beliefs about gender norms, on average, at about three out of the 14 questions on the gender role models scale. The negative effect observed is more prevalent among individuals with low compared to individuals with high self-reported COVID-19 exposure (Table A.55).[98]

---

[94]Moreover, the results indicate that the positive effect is both larger for individuals in poor households compared to individuals in very poor households, and individuals with high self-reported COVID-19 exposure compared to individuals with low exposure. The differences between the two groups are not statistically significant, however.

[95]The difference between Sindh and Punjab of 01.7pp is statistically significant at the 1% level.

[96]In a first step, four indicator variables are generated for each right, which are coded as 1 if the respondent is able to identify it, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of rights the respondent was able to identify (max. four). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by four).

[97]In a first step, indicator variables are generated for each statement, which are coded as 1 for either agreement (combines full (=4) and slight (=3) agreement) for positively phrased statements, e.g., "Mothers should be able to work outside the home after they have children if they want to", or disagreement (combines slight (=2) and full (=1) disagreement) for negatively phrased statements, e.g., "It is important that sons receive more education than daughters", and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions of the gender role models scale which are answered positively, i.e., at which respondents present egalitarian beliefs about gender norms (max. 14). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by 14).

[98]The difference between the two groups is statistically significant at the 5% level.

Nevertheless, the program increases the level of disagreement with the in-egalitarian statements that sons should be more educated so to look after their parents, and that a good woman never questions her husband's opinions. When looking at all single items entering the scale one by one in Figure 3.6b, positive program impacts can be observed for two items. Namely, the graph indicates an increase in the level of disagreement that sons should be more educated so to look after their parents (the control mean is 37% and the impact is an increase in 4.7pp, significant at the 5% level), and an increase in the level of disagreement that a good woman never questions her husband's opinions even if she is not sure she agrees with him (the control mean is 26% and the impact is an increase in 3.4pp, significant at the 5% level).

Table 3.18: Social Norms

(I) Child Marriage

	agr.: ok for P to marry sons before 18			agr.: ok for P to marry dau. before 18			agr.: ok for P to marry dau. before 16			can name risks of CM		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
only NFBE	-0.009	-0.041	0.027	-0.003	0.026	-0.035	-0.017	-0.080**	0.041	0.082***	0.070**	0.100**
	(0.033)	(0.084)	(0.052)	(0.034)	(0.033)	(0.050)	(0.025)	(0.037)	(0.039)	(0.029)	(0.031)	(0.042)
CP and NFBE	-0.074**	-0.008	-0.137***	0.008	0.014	0.002	-0.025	-0.059	0.007	0.023	0.043	0.006
	(0.031)	(0.094)	(0.034)	(0.036)	(0.039)	(0.047)	(0.029)	(0.036)	(0.046)	(0.033)	(0.039)	(0.041)
CP, NFBE pool	-0.054	-0.006	-0.080	-0.012	0.003	-0.024	-0.023	-0.084**	0.035	0.034	0.061*	0.020
	(0.034)	(0.086)	(0.050)	(0.035)	(0.037)	(0.046)	(0.027)	(0.035)	(0.041)	(0.032)	(0.035)	(0.041)
CP, not NFBE pool	-0.044	-0.064*	-0.034	-0.004	0.002	-0.014	-0.044**	-0.037	-0.053**	0.015	0.011	0.018
	(0.027)	(0.034)	(0.028)	(0.019)	(0.026)	(0.024)	(0.020)	(0.026)	(0.022)	(0.015)	(0.018)	(0.022)
Control mean	0.443	0.439	0.446	0.443	0.425	0.460	0.332	0.312	0.350	0.835	0.835	0.834
SD	0.498	0.497	0.499	0.497	0.495	0.499	0.471	0.464	0.477	0.370	0.370	0.370
CP overall*	-0.046	-0.051	-0.044	-0.005	0.002	-0.017	-0.039	-0.047	-0.034	0.019	0.022	0.018
CP p-value	0.026	0.459	0.048	0.700	0.910	0.458	0.050	0.005	0.698	0.168	0.060	0.421
IALP overall**	-0.047	-0.045	-0.049	-0.004	0.006	-0.016	-0.036	-0.052	-0.023	0.025	0.029	0.024
IALP p-value	0.013	0.625	0.046	0.917	0.650	0.580	0.111	0.007	0.783	0.084	0.049	0.232
N	4,087	1,937	2,150	6,982	3,376	3,606	6,969	3,362	3,607	6,483	3,150	3,333
RVs	431	426	425	648	648	648	648	648	648	648	648	648
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

	(II) Child Marriage						**(III) Child Rights**			**(IV) Equal Opportunities**		
	index for positive attitudes about mar.			agr.: girls should finish secondary education before mar.			index for recognized CR			index for positive gender role models mentioned		
	(1) All	(2) Male	(3) Female	(4) All	(5) Male	(6) Female	(7) All	(8) Male	(9) Female	(10) All	(11) Male	(12) Female
only NFBE	0.031**	0.036*	0.025	0.049	0.077**	0.017	0.008	0.037	-0.025	0.001	-0.007	0.007
	(0.014)	(0.019)	(0.017)	(0.034)	(0.036)	(0.046)	(0.016)	(0.023)	(0.018)	(0.012)	(0.015)	(0.017)
CP and NFBE	0.010	0.013	0.006	0.026	0.017	0.030	-0.002	-0.002	-0.007	-0.006	-0.025*	0.010
	(0.015)	(0.018)	(0.017)	(0.034)	(0.041)	(0.046)	(0.017)	(0.021)	(0.019)	(0.011)	(0.015)	(0.017)
CP, NFBE pool	0.014	0.021	0.003	0.049	0.065*	0.024	-0.001	-0.001	-0.004	-0.012	-0.030**	0.005
	(0.015)	(0.017)	(0.017)	(0.035)	(0.038)	(0.047)	(0.017)	(0.022)	(0.020)	(0.011)	(0.014)	(0.017)
CP, not NFBE pool	0.009	0.012	0.009	-0.003	0.000	-0.005	-0.004	-0.013	0.003	0.005	0.005	0.005
	(0.010)	(0.012)	(0.010)	(0.019)	(0.026)	(0.024)	(0.011)	(0.013)	(0.012)	(0.007)	(0.009)	(0.008)
Control mean	0.448	0.453	0.444	0.566	0.561	0.571	0.484	0.509	0.461	0.227	0.215	0.237
SD	0.210	0.203	0.215	0.495	0.496	0.494	0.240	0.252	0.227	0.168	0.165	0.170
CP overall*	0.010	0.014	0.008	0.009	0.014	0.001	-0.003	-0.010	0.002	0.002	-0.002	0.005
CP p-value	0.198	0.113	0.538	0.232	0.158	0.721	0.798	0.579	0.973	0.615	0.136	0.571
IALP overall**	0.012	0.016	0.009	0.014	0.020	0.006	-0.002	-0.005	-0.002	0.001	-0.006	0.006
IALP p-value	0.105	0.095	0.327	0.185	0.107	0.599	0.986	0.718	0.531	0.717	0.118	0.550
N	7,213	3,498	3,715	7,024	3,414	3,610	6,453	3,183	3,270	7,228	3,507	3,721
RVs	648	648	648	648	648	648	648	648	647	648	648	648
Model	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols	ols

▸ *Notes*: Table 3.18 displays the program effects on social norm indicators. Effects from separate OLS estimations are reported.
▸ Sample: Endline survey. Full sample.
▸ Program effects on the following child marriage-related indicators are presented in Part I and II of the table: Agrees: ok for Ps to marry sons before 18 (asked in Punjab only), agr.: ok for Ps to marry dau.s before 18, agr.: ok for Ps to marry dau.s before 16, can name risks of CM, index for positive attitudes about mar., and agr.: girls (G) should finish secondary education (SED) before mar.. In Part III of the table, program effects on the index for recognized CRs are reported. Part IV of the table considers program impact on the index for positive items on gender role models mentioned. "Agrees: ok for Ps to marry sons before 18" (Punjab only), "agr.: ok for Ps to marry dau.s before 18", and "agr.: ok for Ps to marry dau.s before 16" are generated by asking parents and adolescents to indicate to what extent they agree, on a scale from 1 (fully agree) to 4 (fully disagree), with the following three statements: "It is ok for parents to marry their sons before their 18th birthday", "It is ok for parents to marry their daughters before their 18th birthday", and "It is ok for parents to marry their daughters before their 16th birthday". For each of these statements, indicator variables are generated which indicate agreement (combines full (=1) and somewhat (=2) agreement) with the statement. " Can name risks of child marriage" is generated by asking adolescents, parents, and community leaders whether there are any risks or negative consequences of getting married as a child. Conditional on answering yes, they are asked to name the specific risks they know, such as "education is likely to stop" and "child-bearing is likely to start early". Based on their answers, a variable indicating whether respondents are able to correctly name at least one specific risk of child marriage is constructed. "Index for positive attitudes about mar." is captured by asking adolescents and parents to indicate to what extent they agree, on a scale from 1 (fully agree) to 4 (fully disagree), with eight statements such as: "Girls should be married as soon as possible to avoid dishonoring", and "Girls should wait until they have finished their secondary education to marry". In a first step, indicator variables are generated for each statement, which are coded as 1 for either agreement (combines full (=1) and somewhat (=2) agreement) with positively phrased statements, e.g., "Girls should wait until they have finished their secondary education to marry", or disagreement (combines somewhat (=3) and full (=4) disagreement) for negatively phrased statements, e.g., "Girls should be married as soon as possible to avoid dishonoring", and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions to which respondents presented progressive gender norms with respect to marriage (max. eight). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by eight). The summary variable indices, out of the eight questions, the share of questions at which respondents present progressive gender norms with respect to marriage, here also referred to as positive marriage-related attitudes, on a scale from 0 to 1. "Agrees: Gs should finish SED before mar." is an example of one of the eight statements included in the index for positive attitudes about marriage, and indicates adolescents and parents fully (=1) or somewhat (=2) agreeing with the statement that girls should wait until they have finished their secondary education to marry, zero otherwise. "Index for recognized child rights" is generated by asking adolescents, parents, and community leaders to correctly identify the four rights which exist for children under the age of 18: A right to have birth registration, a right to education, a right to get information that is important to their health and wellbeing, and a right to be protected from violence and abuse. In a first step, four indicator variables are generated for each right, which are coded as 1 if the respondent was able to identify it, and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of rights the respondent is able to identify (max. four). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total rights. "Index for positive items on gender role models mentioned" is generated by asking adolescents and parents about their view on gender roles. A scale from Waszak et al. (2001) is employed which captures the role of women/ girls vs. men/ boys in the public and private sphere. Adolescents and parents were asked about their agreement, on a scale from 1 (disagree) to 4 (agree), with 14 statements. In a first step, indicator variables are generated for each statement, which were coded as 1 for either agreement (combines full (=4) and slight (=3) agreement) for positively phrased statements, e.g., "Mothers should be able to work outside the home after they have children if they want to", or disagreement (combines slight (=2) and full (=1) disagreement) for negatively phrased statements, e.g., "It is important that sons receive more education than daughters", and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions of the gender role models scale which were answered positively (max. 14). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by 14). The summary variable indices, out of the 14 items of the gender role models scale, the share of questions which are answered positively, i.e., at which respondents present egalitarian beliefs about gender norms. Columns (1), (4), (7), and (10) display the full sample. Columns (2), (5), (8), and (11) display the sample for boys only. Columns (3), (6), (9), and (12) display the sample for girls only.
▸ The equation 2.1 is estimated with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{480}\beta_1 + \frac{60}{480}\beta_2 + \frac{83}{480}\beta_3 + \frac{303}{480}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.
▸ Related table(s): Tables A.54, A.55, A.60.

Figure 3.6: Rights and Opportunities

a) Child Rights   **b)** Equality of Opportunities
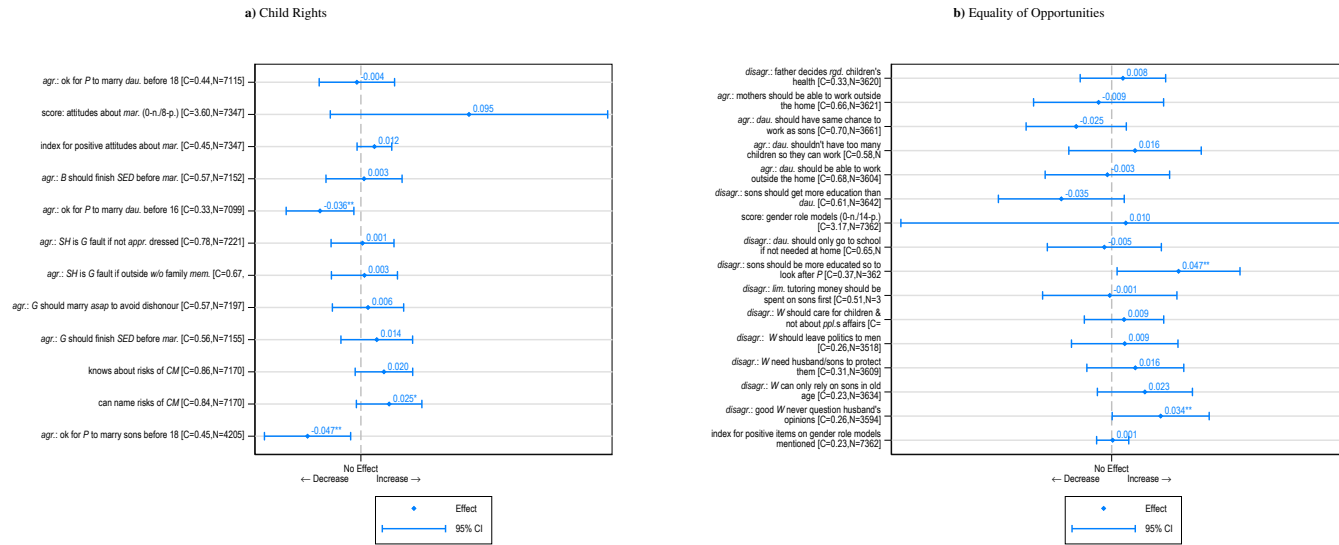
► *Notes*: Figure 3.6a shows treatment effects capturing changes in attitudes towards child marriage. To capture the level of (dis)agreement with the practice of child marriage, parents and adolescents were asked to indicate to what extent they agree, on a scale from 1 (fully agree) to 4 (fully disagree), with the following three statements: "It is ok for parents to marry their sons before their 18th birthday", "It is ok for parents to marry their daughters before their 18th birthday", and "It is ok for parents to marry their daughters before their 16th birthday". For each of these statements, indicator variables are generated which indicate agreement (combines full (=1) and somewhat (=2) agreement) with the statement. To capture respondents' overall attitudes about marriage, adolescents and parents were asked to indicate to what extent they agree, on a scale from 1 (fully agree) to 4 (fully disagree), with eight statements such as: "It is an adolescent girl's fault to be sexually harassed if she is not appropriately dressed", "Girls should be married as soon as possible to avoid dishonoring (to keep honor of the family)", and "Girls should wait until they have finished their secondary education to marry". In a first step, indicator variables are generated for each statement, which are coded as 1 for either agreement (combines full (=1) and somewhat (=2) agreement) with positively phrased statements, e.g., "Girls should wait until they finished their secondary education to marry", or disagreement (combines somewhat (=3) and full (=4) disagreement) for negatively phrased statements, e.g., "Girls should be married as soon as possible to avoid dishonoring", and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions to which respondents present progressive gender norms with respect to marriage (max. eight). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by eight). The summary variable "index for positive attitudes about marriage" indices, out of the eight questions, the share of questions at which respondents present progressive gender norms with respect to marriage on a scale from 0 to 1. "Knows about risks of CM" indicates whether respondents (adolescents, parents, and community leaders) answered yes to the question whether there are any risks or negative consequences of getting married as a child. Conditional on answering yes, respondents were asked to name the specific risks they know, such as "education is likely to stop" and "child-bearing is likely to start early". Based on their answers, a variable indicating whether respondents are able to correctly name at least one specific risk of child marriage is constructed. Figure 3.6b shows treatment effects capturing changes in gender role models. A scale from Waszak et al. (2001) is employed which captures the role of women/ girls vs. men/ boys in the public and private sphere. Adolescents and parents were asked about their agreement, on a scale from 1 (disagree) to 4 (agree), with 14 statements, such as: "Mothers should be able to work outside the home after they have children if they want to", "It is important that sons receive more education than daughters", and "Women should leave politics to men". In a first step, indicator variables are generated for each statement, which are coded as 1 for either agreement (combines full (=4) and slight (=3) agreement) for positively phrased statements, e.g., "Mothers should be able to work outside the home after they have children if they want to", or disagreement (combines slight (=2) and full (=1) disagreement) for negatively phrased statements, e.g., "It is important that sons receive more education than daughters", and 0 otherwise. Thereafter, a score variable is generated which indicates the total number of questions of the gender role models scale which are answered positively (max. 14). In a final step, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of components included in the summary variable (i.e, by 14). The summary variable "index for positive items on gender role models mentioned" indices, out of the 14 items of the gender role models scale, the share of questions which are answered positively, i.e., at which respondents present egalitarian beliefs about gender norms.
► The treatment variable is assignment to any type of program activities.
► Sample: Endline survey.
► Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N refers to the number of respondents.

Evidence from Qualitative Interviews   *Social Norms as Challenges and Obstacles during Implementation* A major obstacle mentioned in most KIIs were the difficulties in achieving involvement of female adolescents in the program in the face of conservative cultural and religious norms.[99] Accomplishing equal participation of female adolescents was especially challenging in cases where adolescent champions were required to travel from the village to the city for trainings, even though this was reported as having been overcome in the end. Interviewees indicated that IPs addressed this cultural barrier by involving the entire family and community members in close communication. Furthermore, some girls were reportedly accompanied by a family member to the first training sessions. Communication with parents and the community was said to have been lacking in Punjab at the beginning of the implementation resulting in male family members questioning the participation of girls and women in the program. Difficulties were also reported in engaging male adolescents, who were often working, and poor adolescents.

> "In the start, it happened that when women attended the meetings, trainings, elder brother, or father found out, so they started questioning. So, we already told the family members about our project (...)." (Key Informant No.12)
>
> "I saw a lot of change because in the beginning, they were reluctant to send their girls even for the training. In the start, someone from the household would accompany the girls as well to the training such as their mother or somebody, just for their own trust. They used to go one day and feel that this is okay." (Key Informant No.3)

II.B.4   Service Provision

> **Key Findings on Service Delivery Outcomes**
>
> Under Pillar III of the program, a core output of the program was to target service providers to promote and provide access to services for adolescents. In particular, according to the theory of change, public authorities were to increase capacity to enact policies and laws to prevent and protect adolescents from abuse. In this section, the hypothesis is tested by assessing changes in service provision, especially with a focus on educational and marriage services. A core public authority to know and enact the laws are the marriage registrars. Given the community-level mobilization efforts and the fact that in this evaluation the marriage registrars working in the respective communities were interviewed, the strategy is to capture changes in behavior due to direct and indirect exposure to the program.
>
> Note, in the ToC service provision is different from output delivery. The latter is under control of the program; the former is meant to be influenced by the program but ultimately under control of the service providers (marriage registers, adolescent service providers). It is their knowledge, attitudes and practices that are meant to change by the program, which is eventually assumed to contribute to the final outcomes and likely reinforces other intermediary outcomes. In this evaluation, this hypothesis is tested by interviewing the community leaders, parents, and adolescents on the general situation of services in the villages as well as by focusing on marriage registrars as core providers of marriage services.
>
> - Compared to control areas, the results indicate no additional impacts of the IALP program on indicators of marriage service provision, access to services over time, and on the extent to which adolescent-friendly services are available in the communities;
> - The program has a sizable positive impact on the prevalence of improved access to non-formal education in program areas;
> - Marriage services in CP areas seem to be positively changed. The findings show a positive effect of the program on the share of Nikah Khwans reporting to have observed a decrease in child marriage in their village over the last four years. Moreover, the program increases the likelihood that a birth certificate is reportedly requested by Nikah Khwans before or during the marriage ceremony.

Table 3.19 displays the effects from a set of OLS regression results for several indicators of service delivery, whereby program villages are compared with non-program areas. The evaluation takes a closer look at the responses provided by community leaders and marriage registrars. While the latter were not necessarily everywhere targeted by the program, they are relevant service providers and capturing their reflection of potential changes in marriage market or even behavioral changes provides a core source of primary, reliable information.

The IALP program had no impact on the extent of correct marriage services provided by marriage registrars. Part I of the table considers program effects on marriage service provision. This is captured by asking marriage registrars about

---

[99]Additionally, the UNICEF ROSA Annual Progress Report 2019 mentions low literacy rates as having impacted negatively on the initial participation of girls (UNICEF 2019).

how they provide marriage services and what their knowledge about child rights is. The summary indicator in this part includes the marriage registrar indicating no benefit of child marriage, being able to name at least one specific harm of child marriage and not accepting the word of parents as proof of age, i.e., only accepting a national identity card or a birth certificate as proof of age.[100] The summary variable thus indices, out of the three services, the proportion of marriage services that are correctly provided by the marriage registrar on a scale from 0 to 1. The results indicate no impacts of the program or its separate components for this indicator. The control mean of 0.829 indicates that marriage registrars in control areas (self-reportedly) provided, on average, 82.9% of marriage services correctly. In order to capture how effects differ for certain groups, a heterogeneous analysis by province (Sindh vs. Punjab) is conducted and also finds no detectable differences (Table A.56).

The program had no impact on the prevalence of respondents reporting non-existence of schools or learning centers as the main community challenge. Part II of Table 3.19 focuses on the program effects on access to services over time. In order to assess challenges faced by adolescents, respondents (adolescents, parents, and community leaders) were asked what, in their opinion, the main challenge is that adolescents face in their community. The dependent variable in column (2) indicates that respondents report the non-existence of schools or learning centers as the main community challenge. The results indicate no impacts of the program or its separate components for this indicator. The control mean of 0.023 indicates that 2.3% of respondents in control areas report non-existence of schools or learning centers as the main community challenge.

The share of marriage registrars knowing the correct legal age of marriage for females was not impacted by the program. In order to capture whether the marriage procedure is in adherence with child rights, marriage registrars were tested to see whether they know the correct legal age for marrying for females, which is 16 in Punjab and 18 in Sindh. The variable in column (3) thus indicates whether the marriage registrar is able to correctly name the legal age of marriage for females. In the control group, almost every second marriage registrar (46.6%) is able to name the correct legal age. The results indicate, however, no impacts of the program or its separate components for this indicator.

The results indicate no impact of the program on the extent of child-friendly service delivery. Part IV of Table 3.19 takes a closer look at program effects on the presence of adolescent-friendly services. To this extent, community leaders were asked to consider the services that are available in their community and to indicate whether service provision for or access to a number services has changed during the last three years. More specifically, they were asked to indicate the change on a scale from 1 (much better) to 5 (much worse). For this summary indicator, only the following four services are considered: Information on rights and entitlements, birth registration, life-skills training, and skills training course (such as apprentices (shaagird)).[101] The summary variable indices, out of the four services, the share of services to which access was indicated to have improved to during the last three years on a scale from 0 to 1. The results indicate that after the program ended, a positive change in the extent of child-friendly service delivery can not be detected in the full sample when compared to the control areas.

The program had a sizable positive impact on the prevalence of improved access to non-formal education in program areas. In Part IV of Table 3.19, further individual indicators of service delivery are considered, as the evaluation is not able to capture positive effects for summary indicators. In column (1) under Part IV, results for an indicator capturing improved access to Non-Formal Education (NFE) are presented, whereby community leaders could report a positive change in access to this service (combines much better (=1) and small positive change (=2)) or not. The results indicate an increase by 27.9pp in areas where Child Protection and educational intervention were jointly implemented (significant at the 5% level). Moreover, a positive effect of smaller magnitude can be observed in areas where the Child Protection-program was implemented in non-NFBE-eligible areas. Here, the increase is by 14.0pp compared to control areas. The average prevalence of improved access to NFE in control areas is reportedly 24.8% and therewith the

---

[100]In a first step, the three individual variables are converted to binary variables, which are coded as 1 if the marriage registrar (i) indicates that there are no benefits of child marriage, (ii) is able to state at least one specific reason why marrying under the age of 18 might be harmful to the married person, and (iii) only accepts a national identity card or birth certificate as proof of age before or during a marriage ceremony, and 0 otherwise. In a second step, a score variable is generated which indicates the total number of marriage services that are correctly provided by the marriage registrar (max. three). Thereafter, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of service components included in the summary variable (i.e, by three).

[101]In a first step, the four individual variables are converted to binary variables, which are coded as 1 if the community leader indicated a positive change (combines much better (=1) and small positive change (=2)) in the respective service provision, and 0 otherwise. In a second step, a score variable is generated which indicates the total number of services to which access has improved to during the last three years (max. four). Thereafter, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of service components included in the summary variable (i.e, by four).

overall program effect is sizable. Thereby, the results indicate that the positive effect in areas where Child Protection and educational intervention were jointly implemented is driven by observations in Sindh (Table A.56). In contrast, the positive effect observed in areas where the Child Protection-program was implemented in non-NFBE-eligible areas is larger in Punjab. The differences in means between the provinces are, however, not statistically significant. The average prevalence of improved access to NFE in the control areas is reportedly 8.4% for Sindh and therewith lower than Punjab where it is 25.7% in this sample.

The findings show a positive effect of the program on the share of Nikah Khwans reporting to have observed a decrease in child marriage in their village over the last four years. In column (2) under Part IV, an indicator variable of whether Nikah Khwans report having observed a decrease ("strong decrease" or "decrease") in child marriages in their village over the last four years is considered. The average prevalence of Nikah Khwans observing a reduction in child marriage in control areas is reportedly 71.4%. The results indicate an increase in such reporting by 12.7pp in areas where the Child Protection-program was implemented in non-NFBE-eligible areas. Investigations by province reveal that the positive effect is larger and significant only for observations in Punjab (Table A.56). The difference between provinces, however, is not significant.

The program increased the likelihood that a birth certificate is requested by Nikah Khwans before or during the marriage ceremony. Column (3) focuses on marriage service provision by Nikah Khwans. The dependent variable is an indicator of whether the Nikah Khwan requests birth proof documents before or during the marriage ceremony. Nikah Khwans were asked to indicate how often, on a scale from 1 (always) to 3 (never), they request birth proof documents before or during the marriage ceremony. Based on their answers, an indicator variable is generated which is coded as 1 if Nikah Khwasn always (=1) or sometimes (=2) requested birth proof documents before or during the marriage ceremony, and 0 otherwise. Even in control areas, almost all (98.2%) of the marriage registrars reportedly request proof of age at marriage. Nevertheless, the results indicate an increase of the likelihood that a birth certificate is requested by 2.5pp in areas where both educational and Child Protection interventions where implemented (significant at the 5% level). Moreover, the positive effect is driven by observations in Sindh (Table A.56).[102]

The findings show no program impacts on the share of boys (10-19) who married last month, as reported by marriage registrars. Finally, column (4) under Part IV of Table 3.19 considers program effects on the proportion of boys (age 10-19) who married last month. In order to obtain this ratio, Nikah Khwans were asked to indicate the total number male children between the age of 10 and 19 who live in their village as well as the total number of male children between the age of 10 and 19 in their village who married in the last month. By dividing the number of those who married by the total number of boys age 10 to 19, the share of boys (age 10-19) who married last month, as indicated by Nikah Khawns, is generated. The results indicate no impacts on this indicator. The results also indicate no effects on potential program-induced changes for girls.

---

[102]The difference between Sindh and Punjab is statistically significant at the 1% level.

Table 3.19: Service Provision

	(I) Marriage Service Provision	(II) Access to Services over Time		(III) Adolescent-Friendly Services
	index for correct mar. service provision of N	comm. challenge: non-existence of schools/ learning centers	N knows correct legal age of mar. for females	index for access to comm. services
	(1) All	(2) All	(3) All	(4) All
only NFBE	0.048	-0.014	0.003	0.145*
	(0.035)	(0.021)	(0.054)	(0.079)
CP and NFBE	-0.007	-0.023	0.006	0.105
	(0.043)	(0.019)	(0.049)	(0.081)
CP, NFBE pool	-0.017	0.005	-0.018	0.033
	(0.042)	(0.019)	(0.052)	(0.084)
CP, not NFBE pool	-0.012	-0.004	0.013	-0.008
	(0.020)	(0.014)	(0.045)	(0.032)
Control mean	0.829	0.156	0.466	0.546
SD	0.223	0.363	0.444	0.301
CP overall*	-0.013	-0.002	0.007	0.001
CP p-value	0.535	0.957	0.950	0.783
IALP overall**	-0.007	-0.005	0.006	0.026
IALP p-value	0.899	0.470	0.977	0.170
N	682	7,934	647	416
RVs	646	653	611	416
Model	ols	ols	ols	ols

	(IV)			
	improved access: NFE	N observed decrease in CM over the last 4 years (o.l.4y.)	N requests birth proof documents before or during mar. ceremony	share of boys (10-19) who married last month
	(1) All	(2) All	(3) All	(4) All
only NFBE	0.200*	0.063	0.004	0.001
	(0.107)	(0.095)	(0.010)	(0.001)
CP and NFBE	0.279**	-0.050	-0.007	-0.000
	(0.111)	(0.098)	(0.008)	(0.001)
CP, NFBE pool	0.010	-0.024	0.002	0.000
	(0.087)	(0.097)	(0.009)	(0.003)
CP, not NFBE pool	0.140***	0.127***	0.025**	-0.009*
	(0.050)	(0.046)	(0.012)	(0.005)
Control mean	0.248	0.714	0.982	0.011
SD	0.429	0.452	0.120	0.055
CP overall*	0.112	0.095	0.020	-0.007
CP p-value	0.141	0.359	0.070	0.109
IALP overall**	0.140	0.074	0.015	-0.006
IALP p-value	0.021	0.684	0.342	0.200
N	400	674	678	479
RVs	400	640	644	452
Model	ols	ols	ols	ols

▸ *Notes*: Table 3.19 displays the program effects on indicators of service-provision. Effects from separate OLS estimations are reported.
▸ Sample: Endline survey. Full sample.
▸ Program effects on the following indicators are considered under Parts I, II, and III: Index for correct mar. service provision of N, comm. challenge: non-existence of schools/ learning centers, N knows correct legal age of mar. for females, and index for access to comm. services. Under Part IV, program effects on the following indicators are considered: Improved access: NFE, N observed decrease in CM o.l.4y., N requests birth proof documents before or during mar. ceremony, and proportion of boys (10-19) who married last month. "Index for correct mar. service provision of N" is captured by asking marriage registrars about how they provide marriage services and what their knowledge about child rights is. In a first step, three binary variables are coded as 1 if the marriage registrar (i) indicates that there are no benefits of child marriage, (ii) is able to state at least one specific reason why marrying under the age of 18 might be harmful to the married person, and (iii) only accepts a national identity card or birth certificate as proof of age before or during a marriage ceremony, and 0 otherwise. In a second step, a score variable is generated which indicates the total number of marriage services that are correctly provided by the marriage registrar (max. three). Thereafter, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of service components included in the summary variable (i.e, by three). The summary variable thus indices, out of the three services, the proportion of marriage services that are correctly provided by the marriage registrar on a scale from 0 to 1. "Community challenge: non-existence of schools/ learning centers" is generated by asking respondents (adolescents, parents, and community leaders) what, in their opinion, the main challenge is that adolescents face in their community. The dependent variable here indicates that respondents report the non-existence of schools or learning centers as the main community challenge. "N knows correct legal age of mar. for females" is captured by asking marriage registrars to state the correct legal age for marrying for females, which is 18 in Sindh and 16 in Punjab. "Index for access to comm. services" is constructed by asking community leaders to consider the services that are available in their community, and to indicate whether service provision for or access to a number services has changed during the last three years. More specifically, they were asked to indicate the change on a scale from 1 (much better) to 5 (much worse). In a first step, the four individual variables are converted to binary variables, which are coded as 1 if the community leader indicated a positive change (combines much better (=1) and small positive change (=2)) in the respective service provision, and 0 otherwise. In a second step, a score variable is generated which indicates the total number of services to which access has improved to during the last three years (max. four). Thereafter, the summary variable is constructed by rescaling the score to range between 0 and 1, attained by dividing the score by the number of total number of service components included in the summary variable (i.e, by four). The summary variable indices, out of the four services, the share of services to which access is indicated to have improved to during the last three years on a scale from 0 to 1. "Improved access: NFE" is captured by asking Nikah Khwans whether they observed a positive change in access to this service (combines much better (=1) and small positive change (=2)) or not. "N observed decrease in CM o.l.4y." is captured by asking Nikah Khwans whether they observed a decrease ("strong decrease" or "decrease") in child marriages in their village over the last four years or not. "N requests birth proof documents before or during mar. ceremony" is generated by asking Nikah Khwans to indicate how often, on a scale from 1 (always) to 3 (never), they request birth proof documents before or during the marriage ceremony. The variable indicates whether Nikah Khwasn always (=1) or sometimes (=2) requested birth proof documents before or during the marriage ceremony, or not. "Share of boys (10-19) who married last month" is generated by asking Nikah Khwans to indicate the total number male children between the age of 10 and 19 who live in their village, as well as the total number of male children between the age of 10 and 19 in their village who married in the last month. By dividing the number of those who married by the total number of boys age 10 to 19, the proportion of boys (age 10-19) who married last month, as indicated by Nikah Khwans, is generated. Columns (1), (2), (5), and (6) display the full sample. Columns (3) and (7) display the sample for males only. Columns (4) and (8) display the sample for females only.
▸ The equation 2.1 is estimated with the following additional control variables: Control variables were selected where needed using a lasso procedure which considers potential imbalance at baseline. The variables are reported in Table A.4. Further covariates are whether the village was NFBE eligible age, gender, indicator for whether the survey was targeting an adolescent or a parent of an adolescent, enumerator fixed effect. For a more detailed description of all outcome variables please refer to Table O.4 in the Appendix.
▸ *Child Protection overall effect: $\frac{83}{386}\beta_3 + \frac{303}{386}\beta_4$; **IALP overall effect: $\frac{40}{486}\beta_1 + \frac{60}{486}\beta_2 + \frac{83}{486}\beta_3 + \frac{303}{486}\beta_4$; Control mean refers to pure control areas in both NFBE eligible and non-eligible areas.
▸ Differences in number of RVs originate from missing data: In the respective RV, the indicator could not be calculated because data is not available.
▸ Standard errors are clustered at the union council level. Significance levels are indicated by $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. N refers to the number of respondents.
▸ Related table(s): Table A.56.

Evidence from Qualitative Interviews   Several interview partners mentioned the *unclear distribution of responsibilities* regarding empowerment of adolescents between government departments and a lack of operational capacity by the administration as a challenge for service provision. One KI specifically attributed the challenges in coordination with the administration to the introduction of the 18th Amendment of the Constitution of Pakistan in 2010, which entailed the devolvement of several ministerial portfolios from federal to provincial level. Others criticized that Youth Affairs are usually part of broad portfolios and that it was unclear who has the mandate for topics concerning adolescents. Furthermore, a KI complained about the lack of operational capacity of ministries, while another criticized that the bureaucracy in rural areas usually does not pursue adolescent empowerment. Additionally, a lack of infrastructure for adolescents at the UC-level was noted in Punjab. This was the perception of KIs who were part of the implementation teams, so there is a potential bias towards claiming success. The data sources of quantitative survey are direct beneficiaries, so that data should be given greater credence.

> "In fact, we had one challenge, that adolescent empowerment is not kind of institutionalized with any of the government departments." (Key Informant No.1)
> "The ministries functioning on the provincial level do not have an outreach to directly undertake an intervention like this. They do not have any direct role; they are just working for protection and they want to perform as best as they can, but they cannot perform well due to lack of operational capacity. On the district level there are at most two officers, and if they want to carry out any activity, they are restricted due to the limitation of human resources and ultimately must rely on civil society or private sector organization." (Key Informant No.4)
> [referring to bureaucracy in Punjab] "(...) in the rural areas, there is no concept of rights of adolescents or empowerment of the youth." (Key Informant No.13)

Key Informants' Concerns About Program Sustainability One major criticism voiced by several KIs concerned the short time span of the intervention, especially regarding the two districts in Punjab in which implementation began only in 2019 due to a change of the IP (UNICEF 2019). KIs criticized that a shift in attitudes and achieving long-term effectiveness would require a longer intervention and that it remains unclear what will happen after the project has ended. Suggestions regarding the timespan of the implementation ranged from three to five years. KIs– also recommended a stronger focus during the planning process on program sustainability. Moreover, interviewees relayed community demands for a continuous offer of NFBECs and the addition of a post-primary NFBE curriculum to the program.[103] Nevertheless, interview partners from various stakeholders remarked that program sustainability and upscaling to cover more districts would require further financial resources. They also identified the importance of forming a strong link between the communities and relevant departments at district level to ensure communication after the project's cessation. Another concern that was raised in connection with the sustainability of the project.

> "As I said earlier, in society nothing can change in one day. Time and resources are needed." (Key Informant No.15)
> "Temporarily, all of us would say that everything is good, but what about when it ends? Obviously, in a seminar or a discussion, everybody would agree to what you say. Let's suppose we are talking about family planning in a seminar. At that time, everybody would agree with you and would seem convinced, but the point to ponder is, are we making a long-term mark or are we leaving a short-term impact on the masses?" (Key Informant No.8)
> "To change attitudes is the biggest problem in our country and when you work with the rural communities, it all goes to waste in the hustle and bustle." (Key Informant No.13)

Necessity for Improved Service Delivery and Holistic Planning KIs across all interviewed stakeholders stressed the necessity to further strengthen the legal framework and its implementation as well as to focus more on capacity building among government officials of relevant departments. There was a perceived need to further engage in activities related to Pillar III as the targeted outcomes are unlikely to be achieved without proper implementation of existing laws against child marriage. KIs therefore recommended that government departments and law enforcement could be targeted as part of further empowerment interventions to increase their capacity. At the same time, other KIs indicated that the Government could bear more responsibility, instead of delegating tasks to NGOs. In addition to strengthening the capacity and the role of government, interview partners recommended also including other actors in the intervention. Examples are offering training on CP to health service workers, including civil defense groups in the implementation, and putting more focus on working with Nikah Khawans. Furthermore, to provide the adolescents with the necessary services to address their needs and claim their rights, interviewees advocated for strengthening the availability of social services.

---

[103]There was also a suggestion to include Life-Skills Based Education in the national curriculum.

> "I think a holistic approach is required for the adolescents and backup services are also important. Because once we tell the children they have right to education, right to health, right to protection from violence, abuse, exploitation, if the services are not available, then it becomes challenging." (Key Informant No.3)
>
> "If this project would be the combination fo advocacy and service delivery, then effectiveness of this project would be so high." (Key Informant No.12)"

II.C   Robustness Specifications

This section outlines further specifications that are run in order to provide more credibility to the results. This section also discusses potential sources of bias and how they have been addressed.

Local Average Treatment Effect   The evaluation so far focused on the so-called Intention-to-Treat (ITT) effects when discussing the results. Hence, as discussed before, the results will likely remain an underestimate of the true effect due to partial compliance (Angrist et al. 1996). In order to compute the LATE, data that reports actual program delivery and program uptake is considered. In this study, the MIS and self-reported endline data is considered.

The MIS data shows that the different four treatment modalities (only NFBE, Child Protection and NFBE, Child Protection and NFBE pool, Child Protection and not NFBE pool) of the IALP program were implemented as assigned except those villages assigned to Child Protection, which were not in the NFBE-eligible pool (Section I.B Table 3.5): Only in around 76% of villages assigned to Child Protection and not eligible for the NFBE pool the program was implemented according to the assignment. In order to compute the LATE, the ITT estimate can be simply divided by the share of villages where the program was implemented as assigned, i.e., by 0.76. Since assignment equals implementation for three out of the four treatment modalities, estimates do not change. The only change can be seen for the coefficient of Child Protection, not NFBE pool: Considering the outcome of married adolescent aged 15 to 17, the coefficient decreased from -0.001 to -0.0013.[104] In other words, the considering actual variation in implementation using MIS data, the evaluation results indicate, as expected, a greater decrease in child marriage rates. Using the MIS data to compute the LATE at the village level enriches the analysis. However, only because the program is delivered/ offered in a village does not mean that individuals take it up, i.e., participate in the program activities. Thus, self-reported data from the endline survey is used to calculate the LATE at the level of participants/ beneficiaries. We apply a Two Stage Least Square regression, where in the first stage, assignment to the program is supposed to predict actual participation. However, in the First Stage regression, actual treatment assignment does not predict well treatment uptake, i.e., self-reported participation in an event. The F-test for First Stage Regression is low. This violates one of the crucial assumptions for conducting a Two Stage Least Square estimation and, thus, for computing the LATE. One reason might be recall bias which is discussed more in Section I.B.

Multiple Hypotheses Testing   Taking into account that multiple outcomes, multiple subgroups, and multiple model specifications are tested at once, multiple hypothesis testing has been considered.[105] While the ITT framework provides conservative point estimates of treatment effects, the applied strategy of multiple hypothesis testing also yields conservative conclusions with respect to statistical inference. Thereby, p-values of ITT estimates from regressions using individual primary outcomes as the dependent variables are adjusted. For this, the method presented in Anderson (2008) is followed allowing to control for the False Discovery Rate (FDR), based on sharpened q-values.[106] Table A.40 in the Appendix reports the p-values and the corresponding sharpened q-values of the main outcomes child marriage and education. After accounting for the multiple hypothesis testing, many of the treatment effects are no longer significant. For instance, for the age group 15-17 of ever married girls the sharpened q-value of 0.462 provided little confidence in the results. [107]

---

[104]This result is identical to running a Two Stage Least Square regression: In the first stage, initial treatment assignment is predicted by the actual treatment implementation. In the second stage, the predicted treatment assignment is regressed on the outcome variable of interest.

[105]For the multiple hypothesis testing, the main outcomes marriage and education are considered. Following (Anderson, 2008) and (List, Shaikh, and Xu, 2019), all subgroups (Punjab/ Sindh, poor/ very poor, covid/ not covid, adolescent girls/ adolescent boys, male/ female) and model specifications (no baseline covariates, all baseline covariates, different clustered standard errors and robust standard errors) per main outcome/ hypothesis (marriage, education) are considered simultaneously, resulting in 68 hypothesis per main outcome.

[106]Q-values can be seen as the proportion of significant results that appear as false leads.

[107]The selected approach is rather conservative in nature. Other methods (e.g., family-wise error rate) and further elaborations on this topic had been beyond the scope of this version and will be paid attention in future revisions.

Clustered Standard Errors   To test for the robustness of the estimated results to alternative inference methods, methods to adjust standard errors are varied. Section A.IV.A in Appendix reports the estimation results with standard errors clustered at different levels, i.e., combination of UC and village level [108], village and household level as well as robust standard errors. Generally, for standard errors clustered at a lower level than the level of UC, significance of the child marriage outcomes drop: For instance, child marriage reduction for females is captured at a 10% instead of a 5% significance level though here the cut-off remains close to the margin of 5% (Table A.30). Robust standard errors report, for female adolescents aged 15 to 17, the same significance levels than standard errors clustered at the level of UC.

Social Desirability Bias   A limitation of using self-reported outcomes is that the results may be biased due to social desirability. This would be particularly a problem if it correlates with program exposure. One could, for instance, imagine that fathers exposed to the program would less likely report marriages of their children when answering the household roster. This would be a concern if it was not the actual outcome (i.e., actually fewer child marriages) but only a response which is provided to please the enumerator, i.e., it would be an experimenter demand effect. To address this concern information on whether other individuals were present during the interview (when specific responses were provided) was collected. While by official instruction, questionnaires have to be responded in an anonymous one-on-one setting, given the tight living conditions it does occur that people walk into the room during an interview. To capture a confounding effect of the presence of others, the enumerators were asked to record this information. With this, an OLS regression is run which we interact the treatment assignment with the presence of others when responses to the household roster (the main source of information for the strategic outcomes) were provided. The estimation display a negative coefficient for the presence of another person interacted with program assignment on the likelihood to report child marriage of age 15-17. The correlation is, however, not statistically significant at the conventional level. In other words, there seems to be little evidence that respondents in program areas were more likely to change their response behavior to sensitive questions due to the presence of others.

Additionally, three list experiments were conducted to capture the prevalence of child marriage with indirect questions (Section A.IV.A). In list experiments, respondents are presented with a "list" of activities or items that they need to count. However, they are never asked directly about each particular item. The sensitive item is hidden within other items. A key element of the experiment is that half of the sample is randomly asked all five activities, while the other half of the sample is only asked the four non-sensitive items. Because the sensitive item is incorporated randomly in questionnaires for some respondents and not to others (some are presented with four non-sensitive other with a list of five items incl. one sensitive), the average number of activities performed by the group that was asked five items can be estimated and the average number of activities reported by the group that was only asked four items can be subtracted from it. The difference between the two is the prevalence of the sensitive item, in this case child marriage. The results indicate no clear impacts on the (indirectly) reported likelihood that individuals attended a child marriage or that a child marriage took place in the household in the past year. A weakly significant increase in the reported incidence of child marriage is captured for Child Protection with NFBE areas in the sample of male respondents, which could be an indicator that male adolescent respondents might have been more responsive to social desirability in their answering.[109]

II.D   Cost-Benefit and Cost-Efficiency

The evaluation conducts the following analyses: A Costs-Benefit Analysis are presented, a Costs-Effectiveness Analysis, and a net present value per 1,000 USD investment.[110] The results indicate a benefit-to-cost-ratio smaller than one (0.742 or 74.2 cents for a dollar invested) and a negative NPV of the program (-856.279 per 1,000 USD) (Table A.64). In other words, under the made assumptions costs of the IALP program exceed benefits. It is important to stress again that the results only constitutes a lower bound of the actual benefits. Positive but non-monetary outcomes such changes in life satisfaction are not considered. Similarly, the IALP program also reached other and indirect beneficiaries which are not accounted for (Section I.D). How does it relate to the cost-effectiveness of other, similar programs? No Costs-Benefit

---

[108]Multi-level clustering of standard errors in order to account for project implementation on UC and village level.

[109]The evaluation design allows for a panel estimation. Since some baseline respondents were reinterviewed during endline, running regressions with matched baseline HHs was part of the PAP. However, this task would require further extensive data work and is therewith not included in this version of the evaluation.

[110]As discussed above (Section I.D), the Costs-Effectiveness Analysis (CEA) estimates the amount of a given outcome that would be achieved for a given investment. For the CEA, impacts are aggregated using the total number of beneficiaries ($N = 269,143$). The net present value (NPV) is the difference between the discounted benefits and the discounted costs of the program.

Analysis of similar interventions in Pakistan could be identified making the comparison of the obtained Costs-Benefit Analysis more challenging. However, impact evaluations of similar interventions in various countries are at hand that conduced a Costs-Benefit Analysis. Although those interventions differ in scopes and objectives, most of them aim at decreasing child marriage and increasing school enrollment. The most similar study, containing a Costs-Benefit Analysis, evaluates an intervention in Bangladesh where an empowerment program is combined with a conditional financial incentive (Buchmann et al. 2021). It reports a benefit-to-cost-ratio of 3.51. This higher ratio originates in a large reduction in child marriage (particularly in the additional years unmarried), and positive significant effects on additional years of schooling. Another intervention, also implemented in Bangladesh, aimed to increase secondary school enrollment of girls. The Female School Stipend Program was evaluated by different researchers (Hahn, Islam, Nuzhat, Smyth, and Yang 2018b; Hong and Sarr 2012), all finding a significant positive effect on additional years unmarried and additional years of schooling. However, since impact sizes differ strongly across different evaluations, benefit-to-cost-ratios computed from the Costs-Benefit Analysis range from 0.39 to 3.16. The benefit-to-cost-ratio of IALP program evaluation lies within this interval.[111]

---

[111]Outside of South East Asia, an Empowerment and Livelihood for Adolescents Program in Uganda was evaluated, using a cluster randomized trial ((Bandiera, Goldstein, Rasul, Burgess, Gulesci, and Sulaiman, 2010)). The evaluation reports a significant reduction in child marriages. The Costs-Benefit Analysis results in a benefit-to-cost-ratio of 3.02.

CHAPTER 4

Conclusion and Outlook

I.   Conclusion

To achieve the Sustainable Development Goals, progress to end child marriage will need to be significantly accelerated. In many parts of the world, adolescents have faced numerous barriers which have been deeply rooted in social norms such as the acceptance of child marriages, violence, and the inferiority of girls. After years of slow global progress, COVID-19 has made the situation of the youth worse. Child marriages and school dropouts, especially amongst girls, are increasing, while access to information about the wellbeing of the adolescents has been limited by lockdowns and mobility restrictions. This evaluation provides a unique opportunity to shed light on the situation of adolescents in areas of Pakistan with historically high levels of child marriages. Core to its interest is the evaluation of an adolescent empowerment program which ended right before COVID-19 broke out. It provides insight to the question whether the comprehensive program designed by UNICEF and implemented by their local partners was able to defend, empower, and protect adolescents. Thereby, the main survey evidence in this evaluation reflects the situation not only before, but also after the COVID-19 pandemic break-out. This provides important information on the current state of several thousand young people and parents of young people in selected areas of rural Sindh and Punjab.

In more detail, this independent evaluation investigates the effects of the UNICEF program "Improving Adolescent Lives in Pakistan" which had the goal to reduce child marriage rates, decrease the number of early pregnancies, and increase school enrollment. The program applied an equity-based approach, aiming that all adolescent children have an opportunity to develop and reach their full potential without discrimination or bias. The IALP program focused on three target groups (or "pillars"): (1) On adolescents, (2) on their families, communities and local decision-makers, and (3) on public authorities and service providers. It encompassed two approaches to achieve those goals: First, the program implemented empowerment and broad awareness-raising activities on child-rights as well as the encouragement for greater inter- and intragenerational dialogues. Hereby, the program specifically also targeted boys and men to question institutional practices and social norms. Second, the program established special non-formal schools for drop-out children which offered accelerated learning courses with the goal to eventually re-integrate the children into formal education. The program was implemented for over more than two years. A baseline survey helped refining the program elements, while a large-scale village-level RCT together with an endline survey allowed to identify its impacts. A final word of caution on the evidence presented is that it may be limited by a lack of statistical power to capture significant effects, though the work had been a priori supported by power calculations that are presented alongside. Additionally, qualitative interviews have shed light on how the program can be improved. They provide rich insights from 16 experts and implementers, reflecting subjective opinions and thus naturally need to be considered with the necessary caution throughout the results and recommendations. Throughout the evaluation, responses from girls and boys were treated equally and results major differences in program impacts between the genders are being highlighted.

The evaluation has captured a number of positive and negative impacts caused by the program. It has started with an outline of evidence on program implementation, capturing causal evidence on program delivery, awareness, and uptake. It was constituted that the IALP program has increased all implementation-related indicators, increasing the share of awareness of and participation in Child Protection- and NFBE-related program activities. The results indicate that the

program was implemented as envisioned (considering evidence on targeting and types of activities reported). However, the program's impacts on its awareness and uptake remain limited in size which may indicate potential limitations to the scope of program-outreach or recall problems (**EQ.1**).

Following, the evaluation has investigated changes in strategic results capturing a wide range of outcomes related to education, marriage, and pregnancies (**EQ.2**). Comparing 167 (control) villages with *apriori* randomly selected 486 program villages, results show a program-induced reduction in child marriage in the relevant age groups, especially for girls. For the age group 10-17, the overall IALP effect is a reduction of 0.3pp for the full sample and 0.7pp for girls (p-value of 0.049 and 0.018, respectively). For adolescents age 15-17, the overall IALP effect is a reduction of 0.8pp for the full sample and 1.7pp for girls (p-value of 0.023 and 0.024, respectively). These are sizable effects compared to a control group mean of 1.9% for the full sample and of 3.3% for girls only, i.e., the outcomes for the group of people who had not been assigned to the program. Overall, the IALP program had a positive impact on child marriage reduction relative to the low level of child marriage captured in the surveys, especially for girls. Moreover, the IALP program increased self-reported engagement rates for boys age 10-17, while their marriage rates did not change. Accordingly, it increased the (indirectly) reported likelihood that boys (age 11-16) will be married next year. The evidence on changes in engagement and delay of marriage, indicates a postponement in line with the original program expectations. Finally, adolescents in the NFBE areas of the IALP program report that the dowry amount paid by their families reduced by 18.4%. In contrast to Buchmann et al. (2021) no increase in dowry prices can be found from being assigned to the Child Protection components of the empowerment program. Moreover, given a small sample size, no statistically significant and meaningful result could be captured for changes in child pregnancy rates.

The evaluation finds mixed impacts of the program on education-related outcomes. The exogenous introduction of Non-Formal Basic Education Centers in Sindh, especially in combination with Child Protection activities (**EQ.3**), led to an increased participation and better learning outcomes (especially for boys). Children residing in areas where the two program components were jointly implemented are not only more likely to attend the non-formal education centers, but also to less likely to miss school, and more likely to want to continue with secondary education. The effects are particularly strong for girls. However, the data also reveals potential evidence of crowding-out effects from formal education, as girls are reportedly less likely to attend primary and boys secondary education due to the program. This effect is particularly strong in NFBE-eligible villages, where the Child Protection program targeted primarily girls. Moreover, the results show selective evidence for a program-induced change in the preference that girls and boys should be educated separately. This impact opposes UNICEF's strive for equity, as does a change in stated opinions that educating boys would be more important than educating girls and the overall constituted reduction in the extent of egalitarian beliefs about gender norms.

In the evaluation, a closer look is taken at intermediate outcomes that are more directly linkable to the activities, especially as implemented under the Child Protection component. For instance, according to the program's theory of change, the greater strategic goals in changing child marriage and education were to be achieved by encouraging exchange and interaction between adolescents as well as adolescents and adults. While patterns in the intergenerational dialogue did not change, an increase of interaction within the generation of adolescents is observed. Especially where girls were targeted by the program, the results indicate an increase of girl's average score on the interpersonal communication competence scale for girls reflecting positive impact on girls' communication skills. Moreover, girl's incidence of intragenerational dialogues among adolescents increased. Selective evidence reflects an increase in communication amongst adolescents regarding their fears and emotions as well as the likelihood that girls talk to their brother about their future marriage.

Empowerment is tautologically the ultimate goal of empowerment programs. There are countless ways to define this broad concept. In this evaluation, apart from the strategic results which are an ultimate final goal in itself, the focus was of closely relating the empowerment measures to the activities (outputs) delivered on the ground. Therewith, the evaluation constitutes that participants of program-related events, especially male participants in areas where the two program components were jointly implemented (Child Protection and NFBE), had a higher likelihood to report that they were reached on topics related to "child rights and entitlements". Compared to control areas, adolescents and parents of adolescents in program areas were no more likely to report that the adolescents' confidence in day-to-day situations changed. Overall, in program and control areas adolescents, especially girls, still feel uncomfortable performing a set of daily tasks by themselves. The measure captured confidence in walking alone in your revenue village during the day/ night, going to the market, visiting friends, going to an educational institution, choosing clothes, choosing a partner for life, and speaking out in public. Accordingly, there is no reported evidence in changes in self-efficacy or on changes in actions taken to ensure adherence to adolescent rights in the last year. Due to the program male adolescents, however, do reportedly feel more confident in their interactions with key stakeholders, in particular government officials or political

leaders for anything benefiting adolescents. Still, overall, confidence levels of adolescent remain and display gender differences that could be addressed in future revised versions of the program. Eventually, the program reduced the acceptance of using corporal punishment on children and led to a decrease in the experience of corporal punishment by adolescents. Male respondents are less likely to accept the usage of corporal punishment on daughters when they discuss their marriage timing. While this result is encouraging, worrisome are the still high levels of reported corporal punishments which hinder children in their development. Finally, the evaluation results show a small but negative impact on the reported wellbeing of female adolescents over the last two weeks, which may be due to the program or the end of it.

Positive program effects are encouraging, zero-effects might be explainable, yet what is core for learning is the protection against adverse program effects, i.e., effects that were not intended yet unexpectedly occurred. Numerous approaches have been tested and implemented over the past decade in many countries. Hope was placed on empowerment programs, yet despite initial optimism, the attempt to change traditional norms can backfire and even have perverse effects. They can alter the marriage market and create unfulfilled hopes. With the end of the programs, empowered children are left alone and many may lose hope for a positive change. The results show a worsening of self-reported girls' wellbeing in the aftermath of the program. Hereby, however, only wellbeing over the last two weeks is considered, a period which is not directly related to impact of the program as the wellbeing may have improved during implementation, and dropped thereafter.

Eventually, the need to underline the importance of protection of adolescents requires a change in norms in an environment in which corporal punishments are still likely to occur. More generally, sustainable change was to be achieved by addressing underlying causes of the problems at hand, which are embedded in harmful social norms and attitudes. For instance, every second respondent reportedly accepts marriages of girls before the age of 18 and every third person before the age of 16. Similar to the findings by Dhar et al. (forth.), however, the evaluation captures a program-related improvement in progressive gender norms, especially among boys. The results, for instance, indicate an important impact of the program, namely that the IALP program reduced agreement with the statement that it is OK for parents to marry sons before the age of 18 and a reduced agreement to marry girls before the age of 16 in program areas. This change may translate to different marriage outcomes in future.

Finally, under the so-called Pillar III of the program, a core output was to target service providers to promote and provide access to services for adolescents. The evaluation captures no general improvement in the overall situation w.r.t. to the provision and access of adolescent-friendly services when comparing program and control areas. However, an improvement in access to non-formal educational services and an in correct provision of marriage services as reported by the marriage registrars in the villages can be constituted.

Apart from gender differences, the evaluation explored how the impacts of the program vary by province, poverty levels, and self-reported COVID-19 exposure (**EQ.6**). The results differed between the provinces, which may related to cultural reasons but also the fact that implementation was more intense in Sindh (due to the implementing partner and the NFBE intervention. In particular, the evaluation finds higher levels of program awareness in this province. Here, the program increased the prevalence of respondents being able to name at least one specific risk of child marriage. The heterogeneous analysis further revealed that the observed reductions in child marriage among adolescents mostly prevalent in Sindh, where also the reduction in dowry was captured, as well as a reduction in corporal punishments amongst boys, an increase in respondents' overall progressive gender norms with respect to marriage, and an increase in the intragenerational dialogues among girls were captured. All impacts on NFBE learning outcomes were estimated for Sindh only, which is also where we find that respondents state that educating boys is more important than girls. Comparing respondents with self-reportedly high and low exposure to COVID-19, the evaluation constitutes that the program effects are on child marriage (rates, friends' marriage, likelihood to name risks of early marriage) as well as intragenerational dialogues are higher for individuals with higher exposure to COVID-19. The incidence of crowding-out effects from formal to informal educational institutions seems to be taking place for individuals with low COVID-19 exposure (though the measure itself may be correlated with other household- and village-level characteristics). The negative effects on egalitarian beliefs were higher in low COVID-19 exposed households. When it comes to poor vs. very poor households, the program effects on child marriage outcomes (rates, risks named) were higher for very poor households where also the incidents of such events were higher on average. Attendance of NFBE centers was, on the other hand, higher for poor households.

Addressing the evaluation question on how to best design the program, the evaluation has found that in many instances a combination of activities led to higher impacts in line with the goals of the program (**EQ.4**). Overall, male respondents are found to be particularly responsive to the program activities and messages, yet this evaluation cannot confidently and rigorously attribute it to the different targeting approaches as set up by the so-called treatment arms

(**EQ.5**). In many instances men and young adolescents in program-assigned areas shared more progressive views.

Finally, a detailed analysis of the program's costs was conducted by the evaluation team and revealed that the costs were almost justified by the quantifiable benefits of the program (**EQ.7**). Under the restrictive assumptions made, the analysis revealed that a dollar invested into the program has a minimum return of 74.2 cents.

This program and its accompanying evaluation have been designed to provide solutions and answers on how to improve the wellbeing of adolescents in Pakistan and beyond. With every fifth person of about 221 million people in Pakistan being an adolescent, this is an important target group. Improving their situation will likely have long-term impacts for Pakistan's growth. Doing so without harm and with continuous learning will ensure that positive impacts are sustainable and welcomed.

II.   Lessons Learned

In this section, the evaluation summarizes lessons learned on *how* the evaluation can contribute to a better adaptation of innovations in future. It focuses on the improvements through the process and outcomes of generating evidence-based, rigorously evaluated solutions. It focuses on what was learned to conduct an evaluation of a complex cross-sectoral program like IALP.

1. *Need to define the learning purposes and expectations as well as to conduct power calculations before setting the evaluation budgets.* Evaluations with extensive primary data collections are expensive. In this evaluation, the goals to be detected were set before the evaluation budget. Yet measuring small changes, requires big sample sizes and therewith bigger costs. The beginning of this evaluation was overshadowed by the need to align the budget to the sample sizes needed. Moreover, given set budgets, learning can be increased from testing combinations of program activities and implementation approaches. The evaluation of the educational component was, for instance, added at a later point in time and was eventually crucial to increase the learning from this evaluation. However, no further resources could be dedicated to study the impacts of this component comprehensively. Consultation prior to setting TORs and evaluation budgets is crucial when designing an evaluation that will be followed for many years ahead. Hereby, realistic expected changes in core outcomes as well as the actual learning needs (instead of external ambitious commitments) could have been considered earlier in the design and budgeting for the evaluation. A detailed theory of change (TOC) on how UNICEF expects to generate change in certain outcomes through its program activities should have been the basis for the evaluation questions to be answered in the TOR. Hereby, the team could have identified gaps and potential challenges that could then have been addressed in a pilot and eventually a rigorous large-scale evaluation. It would also help defining the learning needs more clearly and to link it better to the activities.

2. *Base programs on detailed theory of change (logical chain) frameworks that explain exactly how outputs will translate to impacts, which assumptions are needed and at which point of time what type of program impacts can realistically be achieved.* While the evaluation team has linked the outputs to a theory of change, the early development of a clearer and more detailed theory of change would have been beneficial. UNICEF's team was very supportive in explaining the details of the program, but eventually the program did vary between the provinces and over time. With a well developed ex-ante theory of change and an aligned monitoring system, these changes could have been tracked over time.

3. *Evaluation is primarily for learning and improves with higher levels of trust and willingness to innovate.* A culture of openness, experimentation, and learning is important to set the most relevant evaluation questions and accordingly adjust programming. In this respect, the evaluation initially team met different opinions and parties, with varying levels of interest to commit to this journey. The level of interest in and support of the evaluation has changed over time, with UNICEF Pakistan office being very supportive throughout, yet may have contributed to missed chances to further improve the design and learning. In general, the consequences of failure (if any) should be openly and early communicated to allow for full buy-in and trust. An evaluation's primary purpose - namely greater learning (as opposed to accounting) - was clearly and repeatedly communicated by the evaluation team, encouraging out-of-box thinking and innovation in programming.

4. *Once a high-level commitment for an evaluation has been made, it needs to be made sure that it is supported by the implementation team and its partners.* An evaluation is only as good as the buy-in by the implementing partners.

It relies on their understanding of the purposes and benefits of the evidence generated and the provision of access to critical information at several points of time. The interest to learn from the results was important to keep the evaluation alive. In particular, given that two units at UNICEF were involved in the program implementation, the coordinating role of the Research and Evaluation Unit at UNICEF Pakistan was very important to align the programming and evaluation and to ensure in a timely manner a common understanding of relevant information.

5. *Given the long duration of an evaluation, documentation is crucial. The evaluation benefited from the positive impacts of clear documentation of the project progress (such as in the PAP, an evaluation design document including the updated theory of change, and the progress reports of 3ie).* The underlying evaluation report is based on six years of close and joint work with several units at UNICEF and 3ie, including natural turnover of key staff in all organizations. It is also part of a multi-country evaluation, generating evidence along simultaneously ongoing evaluation studies in Afghanistan and India. Accordingly, the work required a high level of coordination and documentation. A so-called Pre-Analysis Plan was prepared helping to align the expectations towards the end product, this evaluation report. It documented what is technically actually feasible to answer, and over time kept record of changes in programming, in expected outcomes, and in implementation. The involvement of a regional office, such as ROSA, with its technical expertise to support the evaluation team was important. It not only supported the evaluation team with its requests for additional information and clarifications, but also the country team with further explanations on the needs and relevance of the information inquired by the evaluation team. Close coordination and support for the evaluation team has been core, for instance, in times of uncertainty about the future of the data collection due to the pandemic.

6. *In this evaluation, not all data needs were contractually covered or considered from the onset. A better clarification of these needs (for instance in adjusted TORs) could have improved the quality of evidence on program implementation*:

   (a) *The quantitative evidence relied on MIS, which needs to be uniformly collected across the implementing partners in future projects.* Information on program implementation is to understand how the program worked. Collecting this information only after the end of the program, in an endline survey, can be too late and little informative. Survey information naturally suffers from recall bias, i.e., the fact that participants do not recall activities that were offered or took place in their communities over the past four years.

   (b) *Ideally, MIS information needs to be more detailed, coherent, and informative.* It is not sufficient to say how many individuals attended the activities or how many SMS were sent if it remains unclear how often the interactions took place per person. In other words, was the program intense in outreach or in intensity - or at all intensive? It is also crucial to provide information on how the parameters for the monitoring were collected, including very detailed information on the sampling frame. For instance, if percentages are provided, what was the target group (e.g., how representative for the general population) and how big was the pool of respondents? Without further clarifications, MIS information captured by external partners is hard to understand and use for further evidence generation.

   (c) *Since the MIS data is crucial for the evaluation an early involvement of the evaluation team in the development of the log-frames and the alignment with the evaluation purposes would have paid off even more.* More generally, for the purpose of the evaluation, where this information is being used to describe and analyze the program's impacts, a greater involvement of the evaluation team in the development of the monitoring framework and data collection would have been desirable. A web-platform for monitoring allowed tracking the program's outputs over time. While the evaluation team received access to it, a greater and earlier involvement in its development would have helped better aligning the collected information across implementing partners, making sure it is useful and well-understood for the purposes of the evaluation.

   (d) *Relatedly, no direct communication channel of the evaluation team with the implementing partners was established. The communication was facilitated through UNICEF.* Generally, direct communication channels help clarifying the needs of the evaluation (such as the motivation and need to adhere to a certain design) and allow for timely clarifications, adjustments in design and measures. The goal is then to better shed light on changes in and levels of implementation. While the UNICEF team was supportive in making this link, a timely direct exchange would have been beneficial at times.

   (e) *A midline survey would have clarified several open questions, yet was not budgeted in the TOR.* The program was new and ambitious (being simultaneously implemented at large scale in two provinces with different

implementing partners). Instead, a pilot study and/ or a midline evaluation while working closer in one of the provinces with one implementing partner would have beneficial. It would have allowed to focus and make the necessary adjustments in programming. It would have also allowed to analyze changes in outcomes over time.

(f) *Accordingly, given the long-term nature of the changes to be captured, another wave of data collection could have been planned from the onset to assess long-term impacts. This could have helped shaping a longer research agenda on the important topic at hand.* Program impacts can be assessed at different points of time. The evaluation has continued positive impacts in the desired direction. Several outcomes, however, relate to changes in attitudes and may translate to behavioral changes only later in time. This was something that was early on discussed between the program and evaluation team. Moreover, certain positive behavioral changes, such as the reduction in child marriage, may have further positive long-term impacts in future related to better marriage matches, labor market, and educational outcomes. Even intergenerational impacts can be assessed if the control group is not exposed to the program activities. All these learnings would be important for future planning and would now come at "reduced costs" since they would build on earlier investments made in this evaluation.

(g) *Overall, with more frequent data collections (for instance using rotating subsamples of the targeted population, adjusted MIS data, or midline surveys), UNICEF could measure impacts at repeated points of time.* Several of the effects are small in size, yet they are often big relative to the average prevalence in the population. They are big given that the program was only implemented for two years and the impacts may develop over time, i.e., change in the long-term. On the other hand, some impacts may have been stronger right after the end of implementation (e.g., measures on knowledge acquisition). These adjustments are often easily implemented if considered from the onset (e.g. through a better alignment of monitoring and evaluation designs).

(h) *Costing information on the program was important, allowing to conduct a cost-benefit analysis.*

7. *Early data collection investments and the availability of secondary data were not fully considered or shared with the evaluation team.* For instance, it would have helped to align the baseline survey areas with the target group (the program's eligible pool) early on in order to avoid unnecessary delays and costs. The baseline survey and original program design of the impact evaluation considered program implementation in urban areas (Section II.B.2). Eventually, however, the program was not implemented in urban areas yet the baseline data was collected here. A baseline survey in urban areas captured valuable resources, with evidence which eventually was not used in this evaluation. A change in plans (not to implement in urban areas) was only shared with the evaluation team upon several requests for clarification.

8. *A baseline survey was important to identify the challenges, yet thereafter the program implementation should not be delayed too much.* For instance, the idea of implementing mobile phone interventions in rural areas was critically questioned and the survey evidence on low mobile phone coverage and access was decisive in this process. This practice is recommended, potentially considering a smaller scale baseline survey first (see piloting recommendation below). Moreover, while the baseline survey was conducted in 2015, the implementation started no earlier than almost two years after.

9. *UNICEF's close work and contribution on instruments and documents was crucial.* While many of the indicators to be reported to the donor are set very early on and are being reported periodically, this is not necessarily the best approach to an evaluation. An evaluation, instead, has to consider the changing nature of the program over time. Hereby, original indicators which were eventually unlikely addressed, should receive less attention. Depending on the final focus of the program, the survey instruments and indicators need to be adjusted. Here, the close cooperation with the implementing partners and UNICEF was key to the set up of a PAP and the revision of the final indicators, which reflected the actual program, as well as the adjustment of questionnaires to measure the impacts.

10. *An early involvement of the donor not only with the programming teams but also with the evaluation team was helpful to explain the needs of the evaluation and the value of it expected output.* Having had the opportunity to present the evaluation design and its value allows for a high-level of buy-in, interest, and therewith indirect support. This support acknowledging the value of evidence generated in this report, was crucial given the extensions of the involvement due to external factors.

11. *Given that Pakistan started first with this evaluation team, all tools and insights were shared with India and Afghanistan offices, as well as with qualitative knowledge management work in all of South Asia.* This has allowed to safe costs through the generation of spillovers as well as to generate evidence that is comparable through the application of very similar survey instruments and sampling designs.

12. *Policy exchange.* A better engagement with policy makers in Pakistan could have potentially further increased the relevance of the study.

13. *The independence and research focus of the evaluation improves the quality and credibility of the results generated.* 3ie's presence throughout the time ensured that the evaluation team can act independently and that the rigorous and quality of the evaluation is kept. External referees provided valuable feedback at different points of time. An early commitment to share the results with the wider scientific community and to register the PAP publicly allowed to prepare the evidence generated for a global audience and therewith make a contribution beyond the borders of Pakistan.

III.   Recommendations

The following recommendations are addressed to UNICEF. Given a short period of time since the generation of evidence, the process related to developing and making the recommendations into more actionable plan with support from UNICEF and government has not been initiated. As of December 2021, the recommendations stated in this section are based on the evaluation team's own assessment. They shall be further revised in future, with a set of interactive events being planned for the dissemination of the results. The recommendations are split into three main sections: (A) Focusing on recommendations following the main evaluation questions; (B) Focusing on recommendations based on the needs identified in the descriptive evidence; and (C) Focusing on recommendations from the lessons learned on how to conduct a multi-year large scale impact evaluation.

The following recommendations are addressed to implementing agencies and donors investing in impactful approaches. They are in particular addressed to the teams designing future programs in the fields of child marriage reduction, adolescent empowerment, and education.

A. Recommendations following the Evaluation Questions (as being defined in the Box I.):

1. **EQ.1**: Based on the evidence on *program implementation*, the evaluation team recommends greater outreach and mobilization efforts (longer, more intense exposure, more direct beneficiaries). Overall, evidence on program take-up was limited. It was observed that where exposure to the program was higher (as reported by the target group), the program effects were also higher in size. [Section I.B]

   Program uptake should be improved through the avoidance of delays in implementation, longer and more intensive program implementation in the targeted communities, as well as a greater number of directly exposed beneficiaries.

   Early engagement in local buy-in should be intensified.

   For the program of interest here, the implementation time was rather short given the ambitious original goals to change culturally deeply embedded norms and traditions. Adolescent champions received only few days of training. Given the ambitious goals for the setting, a longer program implementation time for empowerment and educational programs is recommended. This concern was also one of the major criticism voiced by several KIs in the qualitative interviews who were concerned the short time span of the intervention, especially in Punjab.

   Once the program activities are defined, local buy-in is necessary. The quantitative evaluation has shown a limited extent of program saturation with many respondents not recalling the program name or child marriage related activities. One reason is that within villages, the number of beneficiaries was limited or only a short duration of program exposure took place. A greater level of engagement in the communities may be needed. Relatedly, in qualitative interviews, respondents recommended an engagement with communities before starting implementation and spreading information on project (components, aim/why is the project helpful for beneficiaries, who can participate). Moreover, they recommend making use of existing community structures when addressing sensitive topics and targeting vulnerable groups. Hereby, cooperation with NGOs which already have a local community structure, know the communities and local conditions well was recommended.

2. **EQ.2**:

(a) The evaluation team recommends the program as a strategy to reduce *child marriage*. Here, the evidence suggests that more risks were known in program areas, that child marriage rates in the relevant age group reduced, and that harmful attitudes changed. [Section II.A.1]

(b) The evaluation makes no recommendation regarding *child pregnancies* given a limited number of incidents and thus limited statistical power to make any evidence-based claims. Changes in child marriage are, however, likely to translate into changes in child pregnancies if implemented on a larger scale. [Section II.A.2]

(c) With regards to changes in *educational outcomes* due to the program, the evaluation recommends a further investigation into whether the finding on the crowding-out effects from formal to informal educational systems can be confirmed. Moreover, to achieve higher levels of equity, girls' learning outcomes should be further improved as more boys benefited from the new system despite the greater needs and original goals of the program to integrate girls more. Certainly, the challenges faced were greater for girls, but so will be the benefits of overcoming them eventually. [Section II.A.3]

(d) *Adolescent empowerment* will need to be paid attention to in future programming. The levels of adolescents' confidence are low, displaying great gender differences. There is no evidence that the program changed these levels, with potentially the need of more saturation (intensity) or the testing of alternative approaches. There is, however, evidence that the current approach may help in reducing corporal punishments, which are still very common. The finding on the negative impacts w.r.t. self-reported wellbeing should be further investigated. [Section II.B.1]

> Relatedly, clearly communicate and define program closure to manage local expectations and hopes. UNICEF should give communities a perspective on whether parts of the project or the whole project will be continued and in what form. The qualitative interviews stress this need and it can potentially also prevent the deterioration of wellbeing captured in the quantitative interviews later on.

(e) The program can be recommended to increase *communication* between adolescents, yet there is little evidence that it is effective in increasing exchange between adolescents and adults. While parents were directly targeted by the program and it was also implemented by integrating adults and adolescents, the evaluation finds no evidence for a greater dialogue between the generations. Here, the evaluation team recommends more targeted approaches (for instance, teaching parents in communication skills). It seems, however, that raising topics of importance to the adolescents resonated with parents as they were more likely to consider their children's opinions. The program, with one of its approaches to peer-to-peer interactions, raised the incidents of exchange between the adolescents and can thus be recommended for this purpose. [Section II.B.2]

(f) *Social norms* are hard to capture in surveys and even harder to change by programs. The evaluation finds high levels of gender-discriminating norms and high levels of acceptance of child marriages. Still, based on the evidence from this evaluation, the program can be recommended to address such deeply rooted norms. Especially the communication of risks stemming from child marriages were memorized and could be integrated in future communication material. [Section II.B.3]

(g) Improvements in *service provision*, apart from the NFBE Centers, were limited and a greater focus on this outcome is recommended. In particular, a closer focus on marriage registrars is recommended. Only every second marriage registrar knew the correct age of marriage and the program had no impact on changes in this outcome. [Section II.B.4]

> Quantitative results and qualitative evidence from KIs across all interviewed stakeholders stressed the necessity for policy makers to further strengthening of the legal framework and its implementation as well as to focus more on capacity building among key service providers and government officials of relevant departments. In fact, program impacts on the marriage registrars have been limited, despite their core role in ensuring the legislation is followed and the opportunities they have in informing parents about the legal rights of marriage for their children. They are also key representatives of Pillar III of the program, partially residing and always working in the respective villages. They eventually not been a direct targeted focus of the program, albeit having been originally discussed to be more primarily targeted, yet could have been exposed to program interventions as key community-members. Overall, there was little to no evidence that marriage registrars in program areas changed their attitudes toward child marriages. Further engagement is important since marriage registrars are one of the key bottlenecks when it comes to registering (formally or informally) child marriages.

3. **EQ.3**: The effects from NFBE Centers (NFBECs) and Child Protection (CP) activities were similarly strong in their impacts to reduce child marriage, especially in the age group of 15-17. A clear recommendation in favor of one of the approaches as opposed to the other can not be made (i.e., no statistical difference in impacts from NFBEC vs. Child Protection (CP) on child marriage).

4. **EQ.4**: Jointly implementing Child Protection activities and NFBE Centers (NFBECs) is recommended to increase non-formal educational attendance for girls. (Also the qualitative evidence suggests to combine educational and Child Protection intervention as most adolescents who need the one also require the other. It has found evidence that frequently the combination of the two was most beneficial, i.e., has shown the most sizable and statistically significant program impacts).

5. **EQ.5**: The evaluation recommends involving boys and fathers to address sensitive topics, such as child marriage. In several instances, the impacts were particularly strong on this group. It does not mean, however, that the focus needs to be exclusively or primarily on this group as compared to targeting females only or both genders together (as it was experimentally tested in this evaluation). (Also the qualitative evidence suggests to combine educational and Child Protection intervention as most adolescents who need the one also require the other. It has found evidence that frequently the combination of the two was most beneficial, i.e., has shown the most sizable and statistically significant program impacts.

> There is potential for change and men/ boys can play crucial role. The activities tested in this evaluation have indicated that despite the often conservative environment, changes in attitudes are possible. Under the program male acceptance of child marriage was reduced. Men and boys could play a more active role as "agents of change" in also promoting progressive norms and defending adolescent girls and boys from rights violations.

6. **EQ.6**: The program was more successful in Sindh. The evaluation team recommends the scale-up of the program in Sindh. Positive lessons learned (from programming and implementation) from Sindh should be considered in Punjab. The program as implemented in Punjab showed only limited positive impacts on the strategic outcomes. Here, the change in the implementing partner may have hindered the realization of positive impacts. Program effects on marriage-related outcomes were higher in households with higher poverty levels and higher exposure to COVID-19, indicating that targeting by these vulnerability criteria may pay off.

7. **EQ.7**: By focusing on more successful implementing and targeting strategies, the cost-benefit ratio of the program may improve. As it stands, the program has generated slightly more costs than benefits (though more benefits may be realized in future). The evaluation team recommends: (1) Measures at a later point of time to consider and incorporate long-term benefits of the program; (2) a focus on impact-optimizing strategies (e.g., careful choice of the combination of activities, implementing partners, and targeted groups considering the evidence from **EQ.1**-**EQ.6**).

B. Recommendations Based on the Descriptive Evidence and Needs:

The following recommendations are addressed to UNICEF, donors, and policy makers. They can be summarized in one sentence: *There is much work ahead.* The basis for these recommendations is evidence presented in Section I.A. The large-scale data collection involving thousands of adolescents and parents of adolescents, as well as hundreds marriage service providers has revealed not only impacts of a specific program but has provided information on the challenges that children still face. Thus, while in this report the focus lies on the estimation of program impacts (captured through a comparison of program and control areas), the survey evidence is no less - if not more - important. Selected examples of topics we recommend to address:

8. High levels of corporal punishment should need to be addressed in future programs targeting adolescents. This involves its acceptance and prevalence. It also directly relates to high levels of acceptance of corporal punishment when children want to discuss the timing of their marriage.

9. Low levels of confidence and beliefs in self-efficacy need to be address by future programs targeting adolescents.

10. Girls' learning outcomes are overall worse than boys' and should be addressed with targeted support.

11. Programs involving marriage registrars and their need follow legal regulations are recommended.

12. The overall share of people reached on topics related to child marriage remains low and needs to improve.

C. Recommendations to Achieve Greater Measurable Impacts:

The following recommendations are addressed to the teams responsible for setting up and committing to rigorous impact evaluations:

13. *Pilot the program if it contains innovative components (e.g., new targeting, new goals, new activities). Scale-up after a midline assessment.* UNICEF should ideally start the implementation of innovative approaches at a smaller scale with a trusted partner, piloting the program, implementing it in one geographical area (e.g., province) and following a midline evaluation, scaling it up to further provinces. A pilot may also serve as an opportunity for adjustment and learning. Program activities, target groups and the reliability and effectiveness of the implementing partner could be tested. The underlying endline may serve as a midline assessment for further continuation of the program, if funds are available.

14. *For an impact evaluation, working with fewer implementing partners is recommended to ensure adherence to protocol and design as well as comparability of implementation.*

   This program has been implemented by several local partners in several provinces at once, which for an innovative approach requires considerable coordination and adjustments. In Sindh, the Child Protection partner was present over the years and the NFBE implementing partner joined later, while in Punjab the Child Protection implementing partner changed in the midst. For an innovative project, with experimental variation to test the impacts in a rigorous, large-scale evaluation these changes and coordination needs might have withdrawn attention from a closer tracking of the implementation.

15. *Relate desired program impacts (goals to be achieved) to implementation duration and saturation.* The original goals of the program were ambitious to start with, generating fears of failure instead of promoting a culture of learning, innovation, and exploration. With higher, more ambitious goals, invest in longer implementation time. Alternatively, if funds are limited and therewith the duration of program implementation is bounded, consider setting lower goals.

   For instance, if the reduction in child marriage is the goal and the actual duration of program implementation is one year on the ground, the focus in the choice of key performance indicators should be on knowledge of rights and access to information. One approach could be to develop a theory of change first, and then plot the goals to be achieved (KPIs) against the different points in time when these shall be measured. This allows tracking outcomes and developments overtime for rightsholders of the program (adolescents), early adjustments if lower level outcomes are not met, and managing expectations in a complex environment. It also allows to track and reflect upon potential adverse effects that should be captured early on.

   Then, the measurement of program key performance indicators should be clearly defined and transparently adjusted, if needed. The development of the survey instruments jointly with the evaluation team, UNICEF country team, and the implementation partners is hereby recommended. While this may take time, it is beneficial for the ability to capture outcomes since then only relevant data is collected, for outcomes were impacts are expected.

Bibliography

ADOHO, F., S. CHAKRAVARTY, D. T. KORKOYAH JR., M. LUNDBERG, AND A. TASNEEM (2014): "The Impact of an Adolescent Girls Employment Program: The EPAG Project in Liberia," *Policy Research Working Paper 6832*.

AMIN, S., J. AHMED, J. SAHA, I. HOSSAIN, AND E. HAQUE (2016): "Delaying Child Marriage Through Community-Based Skills-Development for Girls: Results from a Randomized Controlled Study in Rural Bangladesh." Report, Population Council.

AMIN, S., S. AINUL, F. AKTER, M. MASUDUL ALAM, J. AHMED, AND U. ROB (2014): "From Evidence to Action: Results from the 2013 Baseline Survey for the BALIKA Project," Report, New York: Population Council.

ANDERSON, M. L. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103, 1481–1495.

ANGRIST, J. D., G. W. IMBENS, AND D. B. RUBIN (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91, 444–455.

ASADULLAH, M. N., E. DE CAO, F. KHATOON, AND Z. SIDDIQUE (2020): "Measuring Gender Attitudes Using List Experiments," IZA Discussion Paper.

ASHRAF, N., N. BAU, C. LOW, AND K. MCGINN (2020): "Negotiating a Better Future: How Interpersonal Skills Facilitate Intergenerational Investment," *The Quarterly Journal of Economics*, 135, 1095–1151.

ASHRAF, N., N. BAU, N. NUNN, AND A. VOENA (2014): "Bride Price and the Returns to Education for Women," Working paper.

ASIAN DEVELOPMENT BANK (2013): *Cost-Benefit Analysis for Development: A Practical Guide*, Asian Development Bank.

BAIRD, S., C. MCINTOSH, AND B. ÖZLER (2011): "Cash or Condition? Evidence from a Cash Transfer Experiment," *The Quarterly Journal of Economics*, 126, 1709–1753.

BANDIERA, O., N. BUEHREN, R. BURGESS, M. GOLDSTEIN, S. GULESCI, I. RASUL, AND M. SULAIMAN (2020): "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa," *American Economic Journal: Applied Economics*, 12, 210–59.

BANDIERA, O., N. BUEHREN, M. GOLDSTEIN, I. RASUL, AND A. SMURRA (2019): "The Economic Lives of Young Women in the Time of Ebola: Lessons from an Empowerment Program," *World Bank Policy Research Working Paper*.

BANDIERA, O., M. GOLDSTEIN, I. RASUL, R. BURGESS, S. GULESCI, AND M. SULAIMAN (2010): "Intentions to Participate in Adolescent Training Programs: Evidence from Uganda," *Journal of the European Economic Association*, 8, 548–560.

BEA, U. (2021): "Table 1.1.7. Percent Change From Preceding Period in Prices for Gross Domestic Product," Tech. rep., accessed: 2021-07-11.

BLAIR, G., K. IMAI, AND J. LYALL (2014): "Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan," *American Journal of Political Science*, 58, 1043–1063.

BREIEROVA, L. AND E. DUFLO (2004): "The Impact of Education on Fertility and Child Mortality: Do Fathers Really Matter Less Than Mothers?" Working Paper Series 10513, National Bureau of Economic Research.

BRUECKNER, H., A. MORNING, AND A. NELSON (2005): "The Expression of Biological Concepts of Race," in *Annual Meeting of the American Sociological Association, Philadelphia, PA*.

BUCHMANN, N., E. FIELD, R. GLENNERSTER, S. NAZNEEN, AND X. Y. WANG (2021): "A Signal to End Child Marriage: Theory and Experimental Evidence from Bangladesh," Tech. rep.

BUREAU OF STATISTICS (2020): "Multiple Indicator Cluster Survey Sindh, 2018-2019, Survey Findings Report," Tech. rep., Bureau of Statistics, Planning & Development Board, Government of the Sindh, Karachi, Pakistan.

BUREAU OF STATISTICS, BUREAU AND BOARD, PLANNING & DEVELOPMENT AND OF THE PUNJAB, GOVERNMENT (2018): "Multiple Indicator Cluster Survey Punjab, 2017-2018, Survey Findings Report," Tech. rep., Bureau of Statistics, Planning & Development Board, Government of the Punjab, Lahore, Pakistan.

CALERO, C., V. G. DIEZ, Y. SOARES, J. KLUVE, AND H. C. CORSEUIL (2017): "Can Arts-Based Interventions Enhance Labor Market Outcomes Among Youth? Evidence from a Randomized Trial in Rio de Janeiro," *Labour Economics*, 45, 131 – 142.

ÇARKOĞLU, A. AND S. E. AYTAÇ (2015): "Who Gets Targeted for Vote-Buying? Evidence from an Augmented List Experiment in Turkey," *European Political Science Review*, 7, 547–566.

CHUANG, E., P. DUPAS, E. HUILLERY, AND J. SEBAN (2021): "Sex, Lies, and Measurement: Consistency Tests for Indirect Response Survey Methods," *Journal of Development Economics*, 148, 102582.

CORNO, L., N. HILDEBRANDT, AND A. VOENA (2020): "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments," *Econometrica*, 88, 879–915.

CORSTANGE, D. (2008): "Sensitive Questions, Truthful Answers? Modeling the List Experiment with LISTIT," *Political Analysis*, 17, 45–63.

——— (2010): "Vote Buying Under Competition and Monopsony: Evidence from a List Experiment in Lebanon," in *Presentation at the Annual Meeting of the American Political Science Association, Washington, DC*.

CURRIE, J. AND E. MORETTI (2003): "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings," *The Quarterly Journal of Economics*, 118, 1495–1532.

DE CAO, E., C. LUTZ, ET AL. (2015): "Measuring Attitudes Regarding Female Genital Mutilation Through a List Experiment," Tech. rep., Centre for the Study of African Economies, University of Oxford.

DHALIWAL, I., E. DUFLO, R. GLENNERSTER, AND C. TULLOCH (2013): "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education," *Education Policy in Developing Countries*, 17, 285–338.

DHAR, D., T. JAIN, AND S. JAYACHANDRAN (forth.): "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India," Tech. rep., American Economic Review.

DOEPKE, M. AND M. TERTILT (2009): "Women's Liberation: What's in it for Men?" *The Quarterly Journal of Economics*, 124, 1541–1591.

DUFLO, E., P. DUPAS, AND M. KREMER (2015): "Education, HIV, and Early Fertility: Experimental Evidence from Kenya," *American Economic Review*, 105, 2757–2797.

DUFLO, E., R. GLENNERSTER, AND M. KREMER (2007): "Using Randomization in Development Economics Research: A Toolkit," *Handbook of Development Economics*, 4, 3895–3962.

EDMONDS, E., B. FEIGENBERG, AND J. LEIGHT (2019): "Advancing the Agency of Adolescent Girls," `http://www.dartmouth.edu/~eedmonds/gep/rtrshort.pdf`.

EDMONDS, E. V., B. FEIGENBERG, AND J. LEIGHT (2021): "Advancing the Agency of Adolescent Girls," Tech. rep., Review of Economics and Statistics.

FEDERAL BUREAU OF STATISTICS (2011): "Labour Force Survey," Tech. rep., Government of Pakistan Statistics Division.

FIELD, E. AND A. AMBRUS (2008): "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh," *Journal of Political Economy*, 116, 881–930.

FIELD, E., R. GLENNERSTER, N. BUCHMANN, AND K. MURPHY (2016): "Cost-Benefit Analysis of Strategies to Reduce Child Marriage in Bangladesh," *Copenhagen Consensus Center*.

FIELD, E. AND K. VYBORNY (2018): "Institutional Reform and De Facto Women's Rights in Pakistan," .

GARCIA-HOMBRADOS, J. (2021): "Child Marriage and Infant Mortality: Causal Evidence from Ethiopia," *Journal of Population Economics*, 1–61.

GEDDES, R. AND D. LUECK (2002): "The Gains from Self-Ownership and the Expansion of Women's Rights," *American Economic Review*, 92, 1079–1092.

GERTLER, P. J., S. MARTINEZ, P. PREMAND, L. B. RAWLING, AND C. M. J. VERMEERSCH (2016): "Impact Evaluation in Practice, Second Edition," *Inter-American Development Bank and World Bank*.

GONZALEZ-OCANTOS, E., C. K. DE JONGE, C. MELÉNDEZ, J. OSORIO, AND D. W. NICKERSON (2012): "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua," *American Journal of Political Science*, 56, 202–217.

GOVERNMENT OF PAKISTAN (2017): "Pakistan Education Statistics 2015-2016," `http://library.aepam.edu.pk/Books/Pakistan%20Education%20Statistics%202015-16.pdf`.

GOVERNMENT OF SINDH (2016): "Curriculum of Non Formal Basic Education (NFBE)," Tech. rep.

——— (2017): "Sindh Non-Formal Education Policy," Tech. rep.

GRAHAM, A., M. A. POWELL, D. ANDERSON, R. FITZGERALD, AND N. J. TAYLOR (2013): *Ethical Research Involving Children*, UNICEF Office of Research-Innocenti Florence.

HAERPFER, C., R. INGLEHART, A. MORENO, C. WELZEL, K. KIZILOVA, J. DIEZ-MEDRANO, M. LAGOS, P. NORRIS, E. PONARIN, AND B. PURANEN (2020): "World Values Survey Wave 7 (2017-2020) Pakistan," Tech. rep., JD Systems Institute & WVSA Secretariat, Madrid, Spain & Vienna.

HAHN, Y., A. ISLAM, K. NUZHAT, R. SMYTH, AND H.-S. YANG (2018a): "Education, Marriage and Fertility: Long-Term Evidence from a Female Stipend Program in Bangladesh," *Economic Development and Cultural Change*, 66, 383–415.

——— (2018b): "Education, Marriage, and Fertility: Long-Term Evidence from a Female Stipend Program in Bangladesh," *Economic Development and Cultural Change*, 66, 383–415.

HEERWIG, J. A. AND B. J. MCCABE (2009): "Education and Social Desirability Bias: The Case of a Black Presidential Candidate," *Social Science Quarterly*, 90, 674–686.

HONG, S. Y. AND L. R. SARR (2012): "Long-Term Impacts of the Free Tuition and Female Stipend Programs on Education Attainment, Age of Marriage, and Married Women's Labor Market Participation in Bangladesh," *Retrieved from ImageBank, the World Bank: http://imagebank. worldbank. org/servlet/WDSContentServer/IW3P/IB/2013/09/16/000442464_20130916134309/Rende red/PDF/810610WP0P10620Box0379826B00PUBLIC0. pdf* .

HSIEH, H.-F. AND S. E. SHANNON (2005): "Three Approaches to Qualitative Content Analysis," *Qualitative Health Research*, 15, 1277–1288.

IBARRARAN, P., L. RIPANI, B. TABOADA, J. VILLA, AND B. GARCIA (2014): "Life Skills, Employability and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design," *IZA Journal of Labor & Development*, 3, 1–24.

JACOBY, H. G. AND G. MANSURI (2011): "Crossing Boundaries: Gender, Caste and Schooling in Rural Pakistan," *World Bank Policy Research Working Paper*.

JANUS, A. L. (2010): "The Influence of Social Desirability Pressures on Expressed Immigration Attitudes," *Social Science Quarterly*, 91, 928–946.

KOVALEVA, A. (2012): *The IE-4: Construction and Validation of a Short Scale for the Assessment of Locus of Control*, vol. 9, DEU.

KUKLINSKI, J. H., M. D. COBB, AND M. GILENS (1997): "Racial Attitudes and the "New South"," *The Journal of Politics*, 59, 323–349.

LARSON, R., K. WALKER, AND N. PEARCE (2005): "A Comparison of Youth-Driven and Adult-Driven Youth Programs: Balancing Inputs from Youth and Adults," *Journal of Community Psychology*, 33, 57–74.

LEE-RIFE, S., A. MALHOTRA, A. WARNER, AND A. M. GLINSKI (2012): "What Works to Prevent Child Marriage: A Review of the Evidence," *Studies in Family Planning*, 43, 287–303.

LIST, J. A., A. M. SHAIKH, AND Y. XU (2019): "Multiple Hypothesis Testing in Experimental Economics," *Experimental Economics*, 22, 773–793.

Lumpkin, J. R. (1985): "Validity of a Brief Locus of Control Scale for Survey Research," *Psychological Reports*, 57, 655–659.

Malé, Chata and Wodon, Quentin (2016): "Basic Profile of Child Marriage in Pakistan," Tech. rep., World Bank.

Martinez, M. D. and S. C. Craig (2010): "Race and 2008 Presidential Politics in Florida: A List Experiment," in *The Forum*, De Gruyter, vol. 8.

McKenzie, D. and M. Siegel (2013): "Eliciting Illegal Migration Rates Through List Randomization," *Migration Studies*, 1, 276–291.

Montenegro, C. E. and H. A. Patrinos (2014): "Comparable Estimates of Returns to Schooling Around the World," *World Bank*.

Moseson, H., M. Massaquoi, C. Dehlendorf, L. Bawo, B. Dahn, Y. Zolia, E. Vittinghoff, R. A. Hiatt, and C. Gerdts (2015): "Reducing Under-Reporting of Stigmatized Health Events Using the List Experiment: Results from a Randomized, Population-Based Study of Abortion in Liberia," *International Journal of Epidemiology*, 44, 1951–1958.

Muralidharan, K. and N. Prakash (2017): "Cycling to School: Increasing Secondary School Enrollment for Girls in India," *American Economic Journal: Applied Economics*, 9, 321–50.

National Education Management Information System (NEMIS) and Academy of Educational Planning & Management (AEPAM) (2018): "Pakistan Education Statistics 2017-18," Tech. Rep. AEPAM Publication No. 291, Pakistan.

National Institute of Population Studies (2018): "PDHS 2017–2018 — Key Findings," https://www.nips.org.pk/abstract_files/PDHS%20-%202017-18%20Key%20indicator%20Report%20Aug%202018.pdf.

OECD (2021): *Applying Evaluation Criteria Thoughtfully*.

of Sindh, P. A. (2013): "The Sindh Child Marriages Restaint Bill, 2013," http://sindhlaws.gov.pk/setup/publications/PUB-13-000734.pdf, accessed: 2021-11-08.

Osili, U. O. and B. T. Long (2008): "Does Female Schooling Reduce Fertility? Evidence from Nigeria," *Journal of Development Economics*, 87, 57–75.

Pakistan, N. and ICF (2019): "Pakistan Demographic and Health Survey 2017-18," Tech. rep., NIPS/Pakistan and ICF, Islamabad, Pakistan.

Pakistan Bureau of Statistics (2020): "Pakistan Social and Living Standard Measurement Survey (2018-19)," Tech. rep., Pakistan Bureau of Statistics.

Rasmussen, B., N. Maharaj, P. Sheehan, and H. S. Friedman (2019): "Evaluating the Employment Benefits of Education and Targeted Interventions to Reduce Child Marriage," *Journal of Adolescent Health*, 65, S16–S24.

Redlawsk, D. P., C. J. Tolbert, and W. Franko (2010): "Voters, Emotions, and Race in 2008: Obama as the First Black President," *Political Research Quarterly*, 63, 875–889.

Rehman, S. (2019): "Ammendment to Child Marriage Restraint Act," http://www.senate.gov.pk/uploads/documents/1557377700_875.pdf, accessed: 2019-11-27.

ROTTER, J. B. (1966): "Rotter's Locus of Control Scale," *Psychological Monographs: General and Applied*, 80, 1–28.

RUBIN, R. B. AND M. M. MARTIN (1994): "Development of a Measure of Interpersonal Communication Competence," *Communication Research Reports*, 11, 33–44.

SOLOTAROFF, J. L. AND R. P. PANDE (2014): *Violence against Women and Girls: Lessons from South Asia*, World Bank Group, Washington, DC.

UN HUMAN RIGHTS COUNCIL (2017): "Report of the Office of the United Nations High Commissioner for Human Rights," `https://www.ohchr.org/EN/AboutUs/Pages/ReportHConCivilSociety.aspx`.

UNICEF (2017): "Improving Adolescents' Lives in South Asia: Format E Annual Progress Report and Next Year Plan," Tech. rep., UNICEF.

——— (2018): "Improving Adolescents' Lives in South Asia: Format E Annual Progress Report and Next Year Plan," Tech. rep., UNICEF.

UNICEF (2019): "Annual Progress Report and Next Year Plan," Tech. rep., UNICEF.

UNICEF (2019): "Improving Adolescents' Lives in South Asia: Format E Annual Progress Report and Next Year Plan," Tech. rep., UNICEF.

UNICEF (2019): "The State of the World's Children 2019. Children, Food and Nutrition: Growing Well in a Changing World," Tech. rep., UNICEF.

UNICEF (2020): "Improving Adolescents' Lives in South Asia: Format E Annual Progress Report and Next Year Plan," Tech. rep., UNICEF.

UNICEF (2020): "Investing in Pakistan's Young People," Tech. rep., UNICEF Pakistan Country Office, Islamabad, Pakistan.

——— (2021): "Child Marriage," .

UNICEF (2021): "Technical Note on Adolescent Empowerment," Tech. rep.

UNICEF (2021a): "UNICEF Data Warehouse," .

——— (2021b): "UNICEF Procedure on Ethical Standards in Research, Evaluation, Data Collection and Analysis (2021)," Tech. rep.

UNICEF PAKISTAN (2019): "Every Child is Protected from Violence and Exploitation," Tech. rep.

UNICEF SOUTH ASIA (2016): "Summary of a Baseline Study to Estimate the Number of Child Marriages in South Asia 2014-2017," Tech. rep.

——— (2019a): "Improving Adolescents Lives in South Asia: Format B Section 1 – Regional Project Document," Tech. rep.

——— (2019b): "Improving Adolescents Lives in South Asia: Format E Annual Progress Report and Next Year Plan," Tech. rep.

——— (2020): "Improving Adolescents Lives in South Asia: Format B Section 1 – Regional Project Document," Tech. rep.

Waszak, C., L. J. Severy, L. Kafafi, and I. Badawi (2001): "Fertility Behavior and Psychological Stress: The Mediating Influence of Gender Norm Beliefs Among Egyptian Women," *Psychology of Women Quarterly*, 25, 197–208.

Watson, C. (2014): "Understanding Changing Social Norms and Practices around Girls' Education and Marriage," London: Overseas Development Institute.

WB (2020): "Pakistan Assessment of Civil Service Pensions," .

——— (2021): "World Development Indicators 2021," Tech. rep., World Bank, accessed: 2021-07-11.

World Bank (2011): "Do Conditional Cash Transfers Lead to Medium-Term Impacts? Evidence from a School Stipend Program in Pakistan," Independent evaluation group report, World Bank, Washington, DC.

Zhuang, J., Z. Liang, T. Lin, and F. De Guzman (2007): "Theory and Practice in the Choice of Social Discount Rate for Cost-Benefit Analysis: A Survey," Tech. rep., ERD Working Paper Series.